REC2021



PROCEEDINGS

9TH INTERNATIONAL WORKSHOP ON RELIABLE ENGINEERING COMPUTING

Risk and Uncertainty in Engineering Computations

May 17-20, 2021, Virtual Conference

EDITORS

Alba Sofi, Giuseppe Muscolino, Rafi L. Muhanna

WORKSHOP ORGANIZATION

HONORARY CHAIR

Michael Beer, Leibniz University Hannover, Germany

CHAIRS

Giuseppe Muscolino, University of Messina, Italy
Alba Sofi, University "Mediterranea" of Reggio Calabria, Italy

FOUNDERS AND GENERAL CHAIRS

Rafi L. Muhanna, Georgia Institute of Technology, USA
Robert L. Mullen, University of South Carolina, USA

KEYNOTE SPEAKERS

Michael Beer, Leibniz University Hannover, Germany
Mario Di Paola, University of Palermo, Italy
David Moens, KU Leuven, Belgium

TECHNICAL SUPPORT

Filippo Giunta, University of Messina, Italy
Alessio Altadonna, University of Messina, Italy
Saverio Panarello, University of Messina, Italy
Tiziana Alderucci, University of Messina, Italy
Federica Genovese, University of Messina, Italy

WORKSHOP SUPPORTING INSTITUTIONS AND NON-FINANCIAL SPONSORS

University of Messina, Italy
University "Mediterranea" of Reggio Calabria, Italy
Bernoulli Society for Mathematical Statistics and Probability
ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering; Part B: Mechanical Engineering

RE2021 WEBSITE

http://ww2new.unime.it/REC2021

UNIVERSITY "MEDITERRANEA" OF REGGIO CALABRIA



ASCE American Society of Civil Engineers
ASME SETTING THE STANDARD

Part A: Civil Engineering
Part B: Mechanical Engineering

UNIVERSITY OF MESSINA



Bernoulli Society for Mathematical Statistics and Probability

INTERNATIONAL SCIENTIFIC COMMITEE

Sondipon Adhikari	Swansea University, UK
Umberto Alibrandi	Aarhus University, Denmark
Siu-Kui Au	Nanyang Technological University, Singapore
Ekaterina Auer	University of Wismar, Germany
André T. Beck	University of São Paulo, Brazil
James L. Beck	California Institute of Technology, USA
Michael Beer	Leibniz University Hannover, Germany
Eleni Chatzi	ETH Zurich, Switzerland
Jianye Ching	National Taiwan University, Taiwan
Mario Di Paola	University of Palermo, Italy
Isaac Elishakoff	Florida Atlantic University, USA
Bruce Ellingwood	Colorado State University, USA
Giovanni Falsone	University of Messina, Italy
Scott Ferson	University of Liverpool, UK
Steffen Freitag	Ruhr University Bochum, Germany
Wei Gao	University of New South Wales, Australia
Wolfgang Graf	Technische Universität Dresden, Germany
Achintya Haldar	University of Arizona, USA
Michael Hanss	University of Stuttgart, Germany
Hongwei Huang	Tongji University, China
Héctor A. Jensen	University Santa Maria, Chile
Marcin Kamiński	Lodz University of Technology, Poland
Zhan Kang	Dalian University of Technology, China
Ioannis A. Kougioumtzoglou	Columbia University, USA
Vladik Kreinovich	University of Texas at El Paso, USA
Rudolf Kruse	Otto-von-Guericke University Magdeburg, Germany
Anna Kučerová	Czech Technical University in Prague, Czech Republic
Tom Lahmer	Bauhaus-Universität Weimar, Germany
David Lehký	Brno University of Technology, Czech Republic
Dagang Lu	Harbin Institute of Technology, China
Sankaran Mahadevan	Vanderbilt University, USA
Mehdi Modares	Illinois Institute of Technology, USA
David Moens	Catholic University of Leuven, Belgium
Jamshid Mohammadi	Illinois Institute of Technology, USA
Giuseppe Muscolino	University of Messina, Italy
Michael Oberguggenberger	University of Innsbruck, Austria
Edoardo Patelli	University of Strathclyde, United Kingdom
Kok Kwang Phoon	National University of Singapore, Singapore
Antonina Pirrotta	University of Palermo, Italy
Mallela Venkata Rama Rao	Vasavi College of Engineering, Hyderabad, India
Nathalie Revol	École Normale Supérieure de Lyon, France
Siegfried M. Rump	Hamburg University of Technology, Germany
Adolfo Santini	University “Mediterranea” of Reggio Calabria, Italy
Michael Shields	Johns Hopkins University, USA
Alba Sofi	University "Mediterranea" of Reggio Calabria, Italy
Christian Soize	Marne la Vallée University, France
George Stefanou	Aristotle University of Thessaloniki, Greece
Mark G. Stewart	The University of Newcastle, Australia
Bruno Sudret	ETH Zurich, Switzerland
Yiannis E. Tsompanakis	Technical University of Crete, Greece
Lev Utkin	St. Petersburg State Forest Technical University, Russia
Marcos Valdebenito	University Santa Maria, Chile
Reinhard Viertl	Vienna University of Technology, Austria
Miroslav Vořechovský	Brno University of Technology, Czech Republic
Roman Wan-Wendner	University of Natural Resources and Life Sciences, Austria
Hao Zhang	The University of Sydney, Australia
Enrico Zio	Polytechnic of Milan, Italy

PREFACE

The Center for Reliable Engineering Computing (REC) organizes biennial international workshops which provide an ideal platform for bringing together engineers and scientists from academic, industrial and governmental institutions addressing the reliability of engineering computations.

Previous REC-meetings have been hosted at Georgia Tech Savannah (United States) in 2004, 2006, and 2008, at the National University of Singapore in 2010, at Brno University of Technology in Brno, (Czech Republic) in 2012, at Illinois Institute of Technology in Chicago (United States) in 2014, at Ruhr University Bochum (Germany) in 2016, and at Liverpool University (United Kingdom) in 2018.

The 9th International Workshop on Reliable Engineering Computing was originally planned to be held on May 17-20, 2020, in Taormina (Italy), an international top tourist destination with stunning sea views and unique historical buildings. Due to the Covid-19 outbreak, the event (REC2020) was postponed to May 17-20, 2021, and renamed REC2021. In spite of the efforts of both the organizers and participants to host REC2021 in hybrid format, the meeting had to be converted to a Virtual Conference as a consequence of the dramatic evolution of the pandemic.

The theme of the 9th International Workshop on Reliable Engineering Computing, Virtual Conference (REC2021) is “Risk and Uncertainty in Engineering Computations”.

REC2021 is conceived as a multi-disciplinary forum to bring together various engineering and associated disciplines with the common focus on methodologies and theories in the field of risk and uncertainty.

As in the previous biennial REC-meetings, the central theme of the discussions will be the reliability of engineering computations. Covering a broad range of topics, REC2021 aims to address the urgent need for innovative theories and advanced computational approaches to ensure the safety of engineering systems starting with the uncertainty quantification of input parameters.

The editors are deeply grateful to all authors and reviewers for their valuable contributions during these challenging times.

The patronage of supporting institutions and non-financial sponsors is gratefully acknowledged.

In addition, the support of Prof. Edoardo Patelli and the efforts of Dr. Filippo Giunta who managed the REC2021 website are highly appreciated.

Alba Sofi
Giuseppe Muscolino
Rafi L. Muhanna
Editors

9TH INTERNATIONAL WORKSHOP ON RELIABLE ENGINEERING COMPUTING
Risk and Uncertainty in Engineering Computations
Virtual Conference: May 17-20, 2021

REC2021

TABLE OF CONTENTS

Uncertainty Analysis of Fatigue Failure Using an Interval Approach 1
M. Desch, M. Modares

Deep Interval Neural Network in Computational Mechanics 11
D. Betancourt, R. L. Muhanna

What's Z−X, when Z=X+Y? Dependency Tracking in Interval Arithmetic with Bivariate Sets 27
A. Gray, M. De Angelis, S. Ferson, E. Patelli

Forward Interval Propagation through the Discrete Fourier Transform 39
M. De Angelis, M. Behrendt, L. Comerford, Y. Zhang, M. Beer

Interval Analysis Using Multilevel Quasi-Monte Carlo 53
R. R. P. Callens, M. G. R. Faes, D. Moens

Topology Optimization with Polymorphic Uncertainties using Artificial Neural Networks 68
S. Freitag, S. Peters, P. Edler, G. Meschke

An Integrated Interval Neural Network for Uncertainty Modeling in Inhomogeneous Materials 86
S. Shahi, R. L. Muhanna, F. Fedele

Damage Detection for Structural Health Monitoring under Uncertainty using Deep Interval Neural Networks 102
D. Betancourt, S. Freitag, R. L. Muhanna

Numerical Simulation of a 3D Concrete Printing Process under Polymorphic Uncertainty 118
A. Schmidt, T. Lahmer

Homogenization of 3D Concrete Microstructures Based on CT Image Reconstruction 127
G. Stefanou, D. Savvas, P. Metsis

Elucidating Appealing Features of Differentiable Autocorrelation Functions: a Study on the Modified Exponential Kernel 140
M. G. R. Faes, M. Broggi, P. D. Spanos, M. Beer

Local Reliability Based Sensitivity Analysis with the Moving Particles Method 154
C. Proppe

Mutual Information for Global Sensitivity Analysis and Adaptive-Learning Surrogate Modelling 164
U. Alibrandi, L. V. Andersen, E. Zio

An Efficient Approach for Reliability-Based Optimization of Linear Dynamical Structures subject to Gaussian Excitation 178
D. J. Jerez, H. Jensen, M. A. Valdebenito, M. Misraji, M. Beer

A Fully Decoupled Approach for a Class of Reliability-Based Optimization Problems in Stochastic Linear Dynamics 187
M. A. Valdebenito, M. G. R. Faes

Estimation of Fragilities by the Modified Intensity Measures (IMs) 195
M. Ciano, M. Gioffrè, M. Grigoriu

Analyses of Resistance of Openings to Pyroclastic Flows 209
M. Iacuaniello, A. Montanino, D. De Gregorio, G. Zuccaro

Towards a Multiphase Finite Element Code for the Evaluation of Volcanic Risk 225
A.Montanino, A. Franci, G. Zuccaro

Design Criteria for Temporary Structures in Uncertain Extreme Natural Hazard Load Environments 238
E. Saleh, J. Mohammadi

Reliability Analysis of Randomly Excited Structures with Interval Stiffness Parameters via Sensitivity Analysis 250
F. Giunta, G. Muscolino, A. Sofi

How to Gauge the Quality of a Testing Method when Ground Truth is Known with Uncertainty 265
N. Gray, S. Ferson, V. Kreinovich

Random Set Solutions to Partial Differential Equations 279
J. Karakašević, M. Oberguggenberger

On the Solution of Forward and Inverse Problems in Possibilistic Uncertainty Quantification for Dynamical Systems 295
D. Hose, M. Hanss

A Bimodal Distribution Function with Fuzzy Regression for Predicting Truck Load Population Including Overloads 310
B. Jang, J. Mohammadi

Sensitivity of an Hysteretic Material Model for Random Vibration Analysis of Base-Isolated Rigid-block Historical Artifacts 328
S. Sessa, D. Pellecchia, N. Vaiana, L. Rosati

An Evolutive Probability Transformation Method for the Dynamic Stochastic Analysis of Structures 339
R. Laudani, G. Falsone

AI Based Bridge Health Assessment 350
G. Marasco, B. Chiaia, G. Ventura

On Constrained Distribution-Free P-boxes and their Propagation 365
M. Daub, S. Marelli, B. Sudret

Distribution-Free P-box Processes: Definition and Simulation 380
M. G. R. Faes, M. Broggi, K.-K. Phoon, M. Beer

Distribution-Free Uncertainty Propagation 395
S. Ferson, A. Gray

Fusion of Probabilistic Knowledge as Foundation for Sliced-Normal Approach 408
M. Beer, O. Kosheleva, V. Kreinovich

Interval Matrix Multiplication Using Fast Low-Precision Arithmetic on GPU 419
K. Ozaki, T. Ogita, D. Mukunoki

Guaranteed Minimization of the Bit Error Ratio for Correlated MIMO Systems 435
E. Auer, A. Ahrens

Quantification of Data and Production Uncertainties for Tire Design Parameters 449
M. Böttcher, W. Graf, M. Kaliske

Serviceability Assessment of Footbridges via Interval Analysis 465
R. Santoro, A. Sofi, F. Tubino

Virtual Corrosion Initiation Monitoring for Infrastructures Considering Environmental Uncertainty 476
Y. Yu, W. Gao, Q. Wang, B. Dong

Free Vibration of Functionally Graded Graphene-Platelets Reinforced Composite (FG-GPLRC) Plate with Interval Parameters 486
Y. Q. Huang, D. H. Luo

Hysteretic Tuned Mass Dampers for Seismic Protection 495
A. Boccamazzo, B. Carboni, G. Quaranta, W. Lacarbonara

Behavior of Structures Equipped with Variable Friction Dissipative Systems 510
L. Cavaleri, M. Di Paola, M.F. Ferrotto, V. Sucato

Using a Tuned Fluid Inerter for Improving the Seismic Performance of Structures Isolated with Friction Pendulum System 524
D. De Domenico, G. Ricciardi

Stochastic Response of Beams Equipped with Tuned Mass Dampers, Considering Spring Inertial Effects, Subjected to Poissonian Loads 538
I. P. Dunn, A. Di Matteo, G. Failla, A. Pirrotta

Explicit Sensitivities of the Stochastic Response of Structural Systems under Spectrum Compatible Fully Non-Stationary Seismic Excitations 552
T. Alderucci, F. Genovese, G. Muscolino

Comparison of Fully Non-Stationary Artificial Accelerograms Generation Methods in Non-Linear Dynamic Analyses 566
F. Genovese, G. Muscolino, A. Palmeri

Interval Analysis of the Forced Vibration of Beams with Uncertain Damage 575
F. Cannizzaro, N. Impollonia, G. Cocuzza Avellino, S. Caddemi, I.Caliò

On the Guarantees Derived from a Possibilistic Interpretation of Ensemble Predictions and their Operational Use 588
N. Le Carrer

Why Ellipsoids in Mechanical Analysis of Wood Structures 604
F. N. Schietzold, J. Urenda, V. Kreinovich, W. Graf, M. Kaliske

Imprecise Subset Simulation for Reliability Analysis 615
D. G. Giovanis, M. D. Shields

Enabling the Evidence Theory through Non-intrusive Parametric Model Order Reduction for Crash Simulations 627
J. S. Jehle, V. A. Lange, M. Gerdts

Computation of Temporal Uncertainties in Construction Project Management Using Imprecise Probability 642
M. Modares, M. Desch, R. L. Muhanna, R. L. Mullen

Author Index A

Uncertainty Analysis of Fatigue Failure Using an Interval Approach

Michael Desch[1*] and Mehdi Modares[1]

[1]*Department of Civil, Architectural and Environmental Engineering, Illinois Institute of Technology, USA,*
[*]*Corresponding author: mdesch@hawk.iit.edu*

Abstract. Structural and mechanical systems are susceptible to fracture under cyclic loading, which may lead to fatigue failure. Fatigue failure constitutes a multi-phase process starting with a crack initiation phase and continuing into a crack growth phase, which may result in system failure. In the conventional crack initiation method, based on Miner's rule, each cycle of load is considered to cause an infinitesimal amount of damage in the material leading to a fatigue failure (Miner 1945). This approach uses the results of experiments on the relationship between the applied stress and number of cycles (S-N curves) for the determination of the fatigue life for a specific stress range. These deterministic curves are developed for different structural components and materials through data regression (Basquin 1910).

As such, the uncertainties in material properties, geometry and applied loads are not considered. In some applications, probabilistic methods are utilized by incorporating uncertainties using probability distributions for the parameters governing the formula of the S-N curve. However, the cumulative damage is highly sensitive to these uncertainties, which is an inherent characteristic of using the traditional probabilistic approaches.

In this work, a new method for crack initiation prediction is developed through quantification of uncertainties using an interval approach. Using this method, the values of stress ranges as well as fatigue parameters are quantified as interval variables. Then, the existing interval damage is calculated, leading to the determination of upper and lower bounds of the remaining fatigue life. A numerical example illustrating the developed method is presented; and the results are discussed.

Keywords: Fatigue, Damage, Uncertainty, Interval

1. Introduction

Structures, subjected to cyclic loads, may experience material fatigue failure. Material fatigue failure in components of a structure occurs when the damage caused at each cycle of load is significantly accumulated. The failure can occur even though the stresses induced by the applied cyclic loads are within allowable design criteria. Fatigue failure is a multi-phase process that includes a crack initiation phase and a subsequent crack propagation phase. There exist numerous analytical schemes for prediction of both crack initiation and crack propagation phases.

The conventional crack initiation method, based on Miner's rule, follows the material behavior phenomenon that each cycle of load causes an infinitesimal amount of damage in the material. The damage accumulated over numerous cycles leads to a fatigue failure (Miner, 1945). The crack initiation method can also predict the structure's fatigue life. This method uses experimental results relating the applied stress and number of cycles to failure (S-N relationship) for the determination of the fatigue life (Basquin, 1910).

© 2021 *by M. Desch and M. Modares.*

However, in the crack initiation method, the uncertainties in material properties, geometry and applied loads are not considered. Moreover, this method does not consider the inherent variability of the experiments. Although, in some approaches, the uncertainties are incorporated using probability distributions for the parameters governing the formula of the S-N relationship (Ang and Munse 1975; Sen 2006; Kwon et at. 2012; Diekfuss and Foley 2016), a lack of sufficient data makes the cumulative damage highly sensitive to those uncertainties. One of the approaches to alleviate this issue and enumerate the aforementioned variability and uncertainties with insufficient data is to use interval variables. Using interval variables, the uncertainty is enumerated by lower and upper bounds with no assumption of the possible distributions. Interval variables have been used in several fatigue problems, such as bootstrapping a non-uniform interval S-N curve (Gu and Ma 2018), determining bounds in the crack propagation problem (Long et al. 2018), and performing fatigue analysis on structures with interval axial stiffness (Sofi et al. 2019).

In this work, a new method is developed that is capable of predicting the remaining fatigue life of a structure. This method considers and quantifies variability and uncertainties using an interval approach. Using this method, the uncertainties in the values of the fatigue parameters (obtained from laboratory test data) and stress ranges (obtained from field test data) are quantified as interval variables. Following that, the existing interval damage is calculated, leading to the determination of upper and lower bounds of the structure's remaining fatigue life.

2. Background

This work is the symbiosis of two historically independent fields, structural fatigue life prediction, and interval analysis. To represent the background for this work, reviews of both fields are presented.

2.1. Fatigue Life Prediction

The relationship between the cyclic strain of a structural component and its cycles to failure is obtained through laboratory test data acquisition. This relationship can be rearranged to represent the relationship between cycles to failure and stress range as (Basquin, 1910):

$$N = C/S^m \tag{1}$$

where, C is the fatigue coefficient, and m is the fatigue exponent. The coefficient and exponent are determined experimentally. N is the number of cycles until failure when the system is subject to stress range S. Tests of the S-N relationship usually exhibit considerable scatter. Data regression is generally used to calculate values for the parameters C and m. Eq. 1 is defined for a single stress range, where the stress varies between a maximum stress and minimum stress. However, components in service can be subject to many stress ranges over the course of their service lives. To consider the effects of multiple stress ranges, a cumulative damage approach can be used. In such an approach, each cycle causes a small amount of damage to the system. Over the course of the system's service life, the damage accrued until it reaches a critical value that causes failure. A cumulative damage rule is defined as (Miner 1945):

$$D = \Sigma\, n_i/N_i \tag{2}$$

where D is the total damage, n_i is the number of cycles undergone at stress range i, and N_i is the number of cycles to failure at stress range i. In Miner's model, failure generally occurs when the total damage equals 1. The damage of a structure, subjected to multiple stress ranges, can be calculated by combining Eqs. 1 and 2 as:

$$D = \frac{1}{C} \cdot \sum_i n_i \cdot S_i^m \tag{3}$$

When a structure is subject to a combination of different stress ranges, a cycle counting method may be used to analyze the stress data. The method will produce an equivalent number of cycles corresponding to a finite number of stress ranges. One of the most common cycle-counting methods used is called the rain-flow method (Matsuishi and Endo 1968).

2.2. S-N Data Regression

Fatigue parameters C and m are regressed from S-N data by first transforming both variables S and N to a logarithmic space as $\ln S$ and $\ln N$. In this transformed space, the S-N data forms a linear relationship, and the parameters are determined by least squares regression as:

$$C = \exp\left(\frac{(\Sigma \ln N)(\Sigma (\ln S)^2) - (\Sigma \ln S)(\Sigma \ln S \ln N)}{p(\Sigma (\ln S)^2) - (\Sigma \ln S)^2}\right) \tag{4}$$

$$m = -\frac{p(\Sigma \ln S \ln N) - (\Sigma \ln S)(\Sigma \ln N)}{p(\Sigma (\ln S)^2) - (\Sigma \ln S)^2} \tag{5}$$

where p is the number of data points. In the least squared procedure, the residual value of each data point e_j is calculated as:

$$e_j = \ln N_j - \ln C + m \ln S_j \tag{6}$$

2.3. Uncertainties Present in Fatigue Life Prediction

The equations as presented in section 2.1 determine the structure's fatigue behavior deterministically. However, the uncertainties present in the fatigue life prediction problem must be considered in order to characterize the variations in the fatigue life prediction.

Uncertainties arise from both laboratory experiments and field data acquisition. These uncertainties may be present in 1) the geometric configuration of the detail, 2) the material composition and behavior of the detail, and/or 3) the stress range measurements. Uncertainties may also arise from the numerical procedures used to process and analyze the data. These include: 1) numerical and regression errors from calculating the S-N relationship, and 2) numerical error and the loss of precision when performing cycle counting.

The parameters affected by these uncertainties are the stress ranges S_i and the fatigue parameters (C and m). As the uncertainty in the fatigue parameters are perfectly correlated (Ang and Munse 1975), all uncertainty associated with these parameters may be described as uncertainty in a single parameter. Therefore, the uncertainty in the fatigue parameters may be described with C whilst m remains deterministic.

2.4. Interval Analysis

An interval is defined as a real, closed set bounded by upper and lower values as:

$$\boldsymbol{X} = \left[\underline{x}, \overline{x}\right] = \left\{x \in \mathbb{R} \mid \underline{x} \leq x \leq \overline{x}\right\} \tag{7}$$

in which $\underline{x}$ is the lower bound and $\overline{x}$ is the upper bound of the interval. In this paper, interval variables are depicted in bold. Interval analysis describes the mathematical methods used to define, operate on, and compare interval sets. The set resulting from an interval operation contains all possible values of that operation between all members of the original interval set(s). Interval Analysis and the definitions for its mathematical methods were introduced by Moore (1966). The basic binary operators are amongst those definitions.

Additionally, there has been recent work to fully define the interval power operation; Heimlich (2013). If positive interval $\boldsymbol{X} = \left\{x \in \mathbb{R}^{+} \mid \underline{x} \leq x \leq \overline{x}\right\}$ is raised to a deterministic power y, the result is evaluated as:

$$\boldsymbol{X}^{y} = \left[\underline{x}, \overline{x}\right]^{y} = \left[\min\left(\underline{x}^{y}, \overline{x}^{y}\right), \max\left(\underline{x}^{y}, \overline{x}^{y}\right)\right] \tag{8}$$

3. Methodology

3.1. Formulation of Interval Fatigue Failure Analysis

The general algorithm and major steps for interval fatigue failure analysis and life prediction are given below.

Step 1. Construct an interval S-N relationship from laboratory test data
The uncertainty in the laboratory test data is quantified by defining interval fatigue coefficient $\boldsymbol{C}$ as:

$$\boldsymbol{C} = \left[\underline{C}, \overline{C}\right] = C \cdot \left[\delta_{C_l}, \delta_{C_u}\right] \tag{9}$$

where, C is the fatigue coefficient, and δ_{C_l} and δ_{C_u} are the lower and upper variation parameters, respectively. These variation parameters are obtained using an enveloping procedure as:

$$\delta_{C_l} = \exp\left(\min\left(e_j\right)\right) \tag{10}$$

$$\delta_{C_u} = \exp\left(\max\left(e_j\right)\right) \tag{11}$$

The interval S-N relationship is constructed using interval fatigue coefficient $\boldsymbol{C}$ and fatigue exponent m.

Step 2. Determine interval stress range and cycle counts from field data
Interval stress ranges $\boldsymbol{S_i}$ are constructed as:

$$\boldsymbol{S_i} = \left[\underline{S_i}, \overline{S_i}\right] = \left[S_i - \delta_{S_{i_l}}, S_i + \delta_{S_{i_u}}\right] \tag{12}$$

where, $\delta_{S_{i_l}}$, and $\delta_{S_{i_u}}$ are the lower and upper variation parameters for each stress range, respectively. Values for the variation parameters can be determined by a) the precision of the sensors used to collect the field data, b) the degree of rounding used in the cycle counting method, and/or c) expert opinion.

Step 3. Calculate the interval damage accumulated over the duration of the field data
The interval damage ($\boldsymbol{D_d}$) that is accumulated over the duration of the field data collection is calculated using parameters m and $\boldsymbol{C}$ (Step 1) and the sets $\boldsymbol{S_i}$ and n_i (Step 2) as:

$$\boldsymbol{D_d} = \frac{1}{\boldsymbol{C}} \cdot \sum_i n_i \,.\, \boldsymbol{S_i}^m = \frac{1}{\left[\underline{C}, \overline{C}\right]} \sum_i n_i \,.\, \left[\underline{S_i}, \overline{S_i}\right]^m \tag{13}$$

Step 4. Calculate the interval existing damage and interval remaining life
The interval mean damage per unit time $\boldsymbol{D_m}$ is calculated as:

$$\boldsymbol{D_m} = \frac{\boldsymbol{D_d}}{t_d} = \left[\frac{\underline{D_d}}{t_d}, \frac{\overline{D_d}}{t_d}\right] \tag{14}$$

where, t_d is the duration of the field data collection. In terms of $\boldsymbol{D_m}$, the interval existing damage $\boldsymbol{D_e}$ is:

$$\boldsymbol{D_e} = \boldsymbol{D_m} \cdot t_a = \left[\underline{D_m} \cdot t_a, \overline{D_m} \cdot t_a\right] \tag{15}$$

where, t_a is the current age of the structure. Then, the interval fatigue life of the structure $\boldsymbol{t_l}$ is:

$$\boldsymbol{t_l} = \frac{1}{\boldsymbol{D_m}} = \left[\frac{1}{\overline{D_m}}, \frac{1}{\underline{D_m}}\right] \tag{16}$$

Finally, the interval remaining life $\boldsymbol{t_r}$ is calculated as:

$$\boldsymbol{t_r} = \boldsymbol{t_l} - t_a = \left[\underline{t_l} - t_a, \overline{t_l} - t_a\right] \tag{17}$$

The lower bound of $\boldsymbol{t_r}$, that is $\underline{t_r}$, can then be conservatively taken as the structure's minimum remaining fatigue life.

4. Numerical Example

4.1. Problem Definition

In this example problem, an uncertainty analysis using the developed method is performed to obtain the interval existing damage and interval remaining life for a bridge. To verify the obtained interval bounds, a Monte Carlo simulation (MCS) is also conducted.

The bridge was constructed 20 years ago. The steel grade used was A514. The most fatigue prone details on the bridge are identified as the cover plates on the bottom flanges of the deck girders (Figure 1).

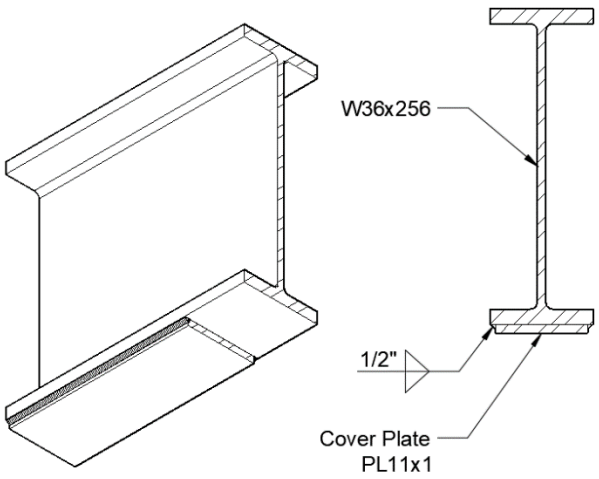


Figure 1. *Cover plate detail, ends un-welded*

The cover plates have un-welded ends. Laboratory test data for this detail/material combination is available from Fisher et al. (1969) (Table I).

Table I. Laboratory S-N Data (adopted from Fisher et al. 1969)	
S_i MPa (ksi)	$N\ (cycles \times 10^3)$ $(Multiple\ tests)$
55.16 (8)	1988.9, 2916.2, 3409.2
82.74 (12)	1031.1, 848.3, 1310.9, 821.7, 1004.7, 1220.0, 755.2
110.32 (16)	514.8, 1227.8, 854.9, 428.5, 542.2, 598.5, 492.9, 412.5, 589.6, 578.0
137.90 (20)	341.3, 429.1, 445.9, 282.3, 192.3, 339.5, 260.0, 238.8, 374.0, 296.0, 207.0
165.47 (24)	156.6, 213.8, 285.2, 192.5

The stress range and cycle count measurements for bridge #0160335 are used for the field test data (adopted from Hahin et al. 1993) (Table II). The cycle counts were obtained over a 24-hour collection period. It is assumed that for each stress range, there is $\pm 1\%$ uncertainty due to sensor precision.

Table II. Stress range and cycle count data (Hahin et al. 1993)	
S_i MPa (ksi)	n_i $(cycles)$
6.89 (1.0)	2843
10.34 (1.5)	991
13.79 (2.0)	386
17.24 (2.5)	111
20.68 (3.0)	88
24.13 (3.5)	64
27.58 (4.0)	33
31.03 (4.5)	32
34.47 (5.0)	15
37.92 (5.5)	7

4.2. PROBLEM SOLUTION

Step 1. The laboratory data is regressed and used to determine interval fatigue coefficient $\boldsymbol{C}$ and fatigue exponent m (Table III).

Table III. Interval fatigue coefficient and fatigue exponent	
Interval Fatigue Coefficient ($\boldsymbol{C}$) $MPa^m \times 10^{10}$ $(ksi^m \times 10^8)$	*Fatigue Exponent (m)*
[1.975,7.478] [2.145,8.122]	2.342

The Interval S-N relationship determined from these parameters is plotted in Figure 2.

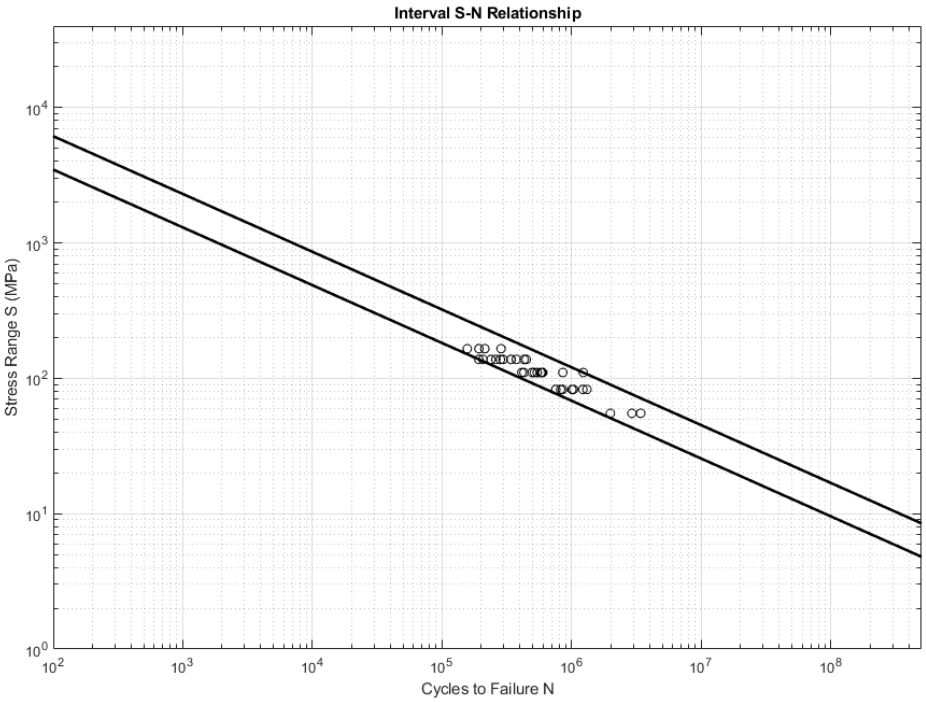


Figure 2. *Interval S-N relationship for an A514 steel girder with cover plates, un-welded ends*

Step 2. The interval stress ranges $\boldsymbol{S_i}$ are determined and enumerated in Table IV.

Table IV. Interval stress ranges

Interval Stress Range ($\boldsymbol{S_i}$)	
MPa	(ksi)
[6.83,6.96]	[0.99,1.01]
[10.24,10.45]	[1.49,1.52]
[13.65,13.93]	[1.98,2.02]
[17.06,17.41]	[2.48,2.53]
[20.48,20.89]	[2.97,3.03]
[23.89,24.37]	[3.47,3.54]
[27.30,27.85]	[3.96,4.04]
[30.72,31.34]	[4.46,4.55]
[34.13,34.82]	[4.95,5.05]
[37.54,38.30]	[5.45,5.56]

Step 3. The interval damage occurring during the one day data collection period ($t_d = 1/365\ years$) is calculated as: $\boldsymbol{D_d} = [1.640, 6.504] \times 10^{-5}$.

Step 4. The values for interval mean damage, interval existing damage ($t_a = 20\ years$), interval fatigue life, and interval remaining life are determined and tabulated (Table 5).

4.3. VERIFICATION

To verify the values and sharpness of the obtained interval bounds, a Monte Carlo simulation is conducted using 10^8 realizations of uniformly distributed random variables. The MCS results are compared to the interval results in Table V.

Table V. Fatigue damage and life prediction results

Parameter	*Interval Results*	*MCS Results*
Interval mean damage ($\boldsymbol{D_m}$)	[0.0060, 0.0237]	[0.0060,0.0236]
Interval existing damage ($\boldsymbol{D_e}$)	[0.1197, 0.4748]	[0.1204,0.4717]
Interval fatigue life ($\boldsymbol{t_l}$) (years)	[42.12, 167.11]	[42.40,166.08]
Interval remaining life ($\boldsymbol{t_r}$) (years)	[22.12, 147.11]	[22.40,146.08]

4.4. OBSERVATIONS

The results depicted in Table 5 shows that the lower bound (conservative) prediction for the remaining life of the bridge is 22.12 years while its upper bound is 147.11 years. This existence of large variation in the remaining fatigue life attests to the hyper sensitivity of fatigue analysis to the presence of uncertainties. This means that large uncertainties in the laboratory or field tests will result in wide uncertainties for the prediction of the remaining fatigue life.

Moreover, and as expected, the results obtained by Monte Carlo simulation are inner bounds of those obtained by the developed method. Furthermore, the interval results are sharp, which can be attributed to the independence of the interval variables in the entire computational scheme.

5. Summary and Conclusions

In this paper, a method to consider uncertainties in fatigue life prediction using interval variables is developed. This method considers uncertainties present in both laboratory and field data as well as uncertainties introduced by the analysis process. Because of its set-based approach, this method computes sharp bounds on the existing damage and remaining fatigue life of the structure with minimal computation effort. The interval results and their sharpness obtained by this method are verified by the Monte Carlo simulations performed. The analysis shows the hyper sensitivity of fatigue analysis to the presence of uncertainties which results in large variations in the fatigue life prediction. Therefore, it is of utmost importance to minimize the errors and uncertainties in the input of the analysis process. The method's simplicity and versatility make it an attractive option to consider uncertainties in the remaining fatigue life problem, especially when limited data is available.

References

Ang, A. H.-S., and W. H. Munse. “Practical reliability basis for structural fatigue.” Meeting preprint 2494, ASCE Annual National Structural Engineering Conference. New Orleans, April 1975: American Society of Civil Engineers. (1975).

Basquin, O.H. “The exponential law of endurance tests”. American Society of Testing Materials, 10: 625-630. (1910).

Diekfuss, J.A. and C.M. Foley. “Modeling Error Uncertainty Characterization for Reliability-Based Fatigue Assessment in Sign Support Structures”. Journal of Structural Engineering, 124(7): 04016042, (2016).

Fisher, J.W., K.H. Frank, M.A. Hirt, and B.M. McNamee. “Effect of weldments on the fatigue strength of steel beams”. Fritz Engineering Laboratory Report No. 334.2. (1969).

Gu, Z. and X. Ma. “A feasible method for the estimation of the interval bounds based on limited strain-life fatigue data”. International Journal of Fatigue, 116: 172-179, (2018).

Hahin, C., J.M. South, J. Mohammadi, and R.K. Polepeddi. “Accurate and Rapid Determination of Fatigue Damage in Steel Bridges”, Journal of Structural Engineering, 119(1), pp 150-168. (1993).

Heimlich, O., Nehmeier M., and Wolff von Gudenberg, J. “Variants of the general interval power function”. Soft Comput, 17:1357-1366, (2013).

Kwon, K., D.M. Frangopol, and M. Soliman. “Probabilistic Fatigue Life Estimation of Steel Bridges by Using a Bilinear S-N Approach”. Journal of Bridge Engineering, 17(1): 58-70, (2012).

Long, X.Y., C. Jiang, K. Liu, X. Han, W. Gao, and B.C. Li. “An interval analysis method for fatigue crack growth life prediction with uncertainty”. Computers and Structures, 210: 1-11, (2018).

Matsuishi, M., and T. Endo. "Fatigue of metals subjected to varying stress". Japan Society of Mechanical Engineering. (1968).

Miner, M.A. “Cumulative Damage in Fatigue”. Journal of Applied Mechanics, 12: A159-A164. (1945).

Moore, R.E. Interval Analysis. Prentice-Hall, Englewood Cliff, NJ, (1966).

Sen, T.K. “Probability of Fatigue Failure in Steel Catenary Risers in Deep Water”. Journal of Engineering Mechanics, 132(9): 1001-1006, (2006).

Sofi, A., G. Muscolino, and F. Giunta. “Fatigue analysis of structures with interval axial stiffness subjected to stationary stochastic excitations”. Meccanica, 54: 1471-1487, (2019).

Deep Interval Neural Network in Computational Mechanics

David Betancourt[1,2)] and Rafi L. Muhanna[1,2)]

[1)] *School of Computational Science and Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA, david.betancourt@gatech.edu* [2)] *School of Civil and Environmental Engineering,*

Georgia Institute of Technology, Atlanta, GA 30332, USA, rafi.muhanna@gatech.edu

Abstract. In computational mechanics, when the input properties to the finite element (FE) model contain uncertainty, it is necessary to model the spatiotemporal variation of the uncertainty and to make predictions with an FE model that can propagate this uncertainty in the system. Often times, we do not have enough information to assume a probability distribution for the input, but only have access to the upper and lower bounds of the data. In such cases, when only the bounds of the input data are known, interval analysis (IA) offers an alternative to model the uncertainty. In addition, we generally do not have direct access to the uncertain property (e.g., elastic modulus) but only to other indirect observations that help characterize it (e.g., strain measurements). As a result, we present an interval deep learning algorithm to model the uncertainty of the input properties for the Interval Finite Element method (IFEM). For this work, we developed a deep interval-valued neural network (DINN) to quantify interval uncertainty of IFEM inputs. The DINN takes as input indirect observations with interval uncertainty (e.g., sensor measurements) and predicts the spatial and/or temporal variation of an unknown IFEM interval input. In this way, interval uncertainty *of the system inputs* is quantified by the DINN and the interval uncertainty *of the mechanical response* is quantified by the IFEM. A numerical experiment is conducted using a dataset of concrete mix measurements with interval uncertainty to predict concrete strength. We conclude by showing that our data-centric method achieves good results compared to ground truth values of the material properties.

Keywords: deep learning, interval uncertainty, interval finite element, interval deep learning

1. Introduction

The analysis and design of complex engineering systems is exposed to a great number of uncertainties in the inputs of the systems' models. Not modeling this uncertainty properly can lead to inaccurate predictions of the system's performance, which in turn can lead to disastrous outcomes. In computational solid mechanics, a system analysis is typically performed with a finite element (FE) model. The FE model receives as input the physical properties of the system under consideration (e.g., material properties, dimensions, support conditions, and applied loading). As output, the FE model computes the resulting responses of the system (e.g., displacements, stresses, and vibration frequencies). The described process is generally referred to as the *forward problem*. When the input properties to the FE model contain uncertainty, it is necessary to model the uncertainty and to make predictions with an FE model that can propagate this uncertainty in the system.

© 2021 *by D. Betancourt and R. L. Muhanna.*

In general, the finite element method (FEM) does not have an integrated mechanism to model the spatial and/or temporal (spatiotemporal) uncertainty variation of the input properties. Given this limitation, the input uncertainty is commonly modeled by either assigning the same FE input property over the entire FE mesh or on an element-basis by the FE analyst, but without an analytical field that considers the spatiotemporal dependencies between mesh elements. Such approaches might vary from the economical to the conservative but they are not always unequivocal. Uncertainty in the input must be duly considered in order to guarantee strength, serviceability, and economics of the structural system being analyzed, as well as the aggregate environmental impact of overly-conservative structures because of excess material. In fact, there are three important considerations for uncertainty quantification in the input properties of the forward problem. First, the spatiotemporal dependence in the domain must be considered—also known as the uncertain field. Second, in many applications in solid mechanics we do not have access to directly measure an unknown property but only have access to indirect measurements. Thus, we need a way to infer the input variable from indirect observations. Third, we need a way to model the aleatory and epistemic uncertainty in the input properties.

Some probabilistic paradigms, such as the Stochastic Finite Element Method (SFEM) (Ghanem and Spanos, 2003), integrate Gaussian random fields with the deterministic FE method to model the uncertain field (i.e., the spatiotemporal-varying uncertainty) in the system inputs and propagate the input uncertainty through the FE model. Nonetheless, random field theory is inadequate for cases where there is epistemic uncertainty and/or not enough data to select a probability distribution (Zhang, 2005). Under such cases, the Interval Finite Element Method (IFEM) presents an alternative paradigm to model system uncertainty because it does not require assumptions on a probability model. Muhanna and Mullen (Muhanna and Mullen, 2001) have shown that the IFEM can be duly beneficial and efficient in computational solid mechanics because there is usually not enough knowledge about the probabilities of the inputs.

However, as with the deterministic FEM in general, the IFEM does not have an integrated procedure to model the uncertainty in the system inputs. As a result, *interval field* models have been recently proposed (Moens and Hanss, 2011; Sofi, 2017; Wu and Gao, 2017) in order to model the uncertainty field for the IFEM input. Previously, we presented a real-valued deep neural network method to model the spatiotemporal variation of an uncertain input IFEM property as a function of indirect observations ((Betancourt et al., 2018)). The resulting output of the real-valued deep neural network could then be given postulated interval bounds based on expert knowledge. For this work, we developed a novel interval-valued deep neural network—Deep Interval Neural Network (DINN)—capable to quantify the aleatory and epistemic uncertainty in the input by using IA. The DINN is able to predict an interval field by using interval-valued indirect measurements of an uncertain IFEM input property.

Our contribution with the DINN is to address the three concerns for uncertainty quantification of the input properties of the FE model: (i) it quantifies the uncertain field by means of an interval field; (ii) it infers the input property (e.g., material properties) by indirect observations (i.e., sensor measurements); and (iii) it quantifies the aleatory and epistemic uncertainty in the input by using IA. Additionally, the DINN can quantify the parameter uncertainty of the neural network itself as part of the algorithm.

This paper is organized as follows. On Sec. 1.1 we explore related work in terms of interval field, interval neural networks, and general data-driven methods in computational mechanics. Sec. 2 explains the process of input uncertainty quantification in the structural forward problem and introduces key concepts of IA. Sec. 3 explains in detail the development of the DINN. Sec. 4 introduces the concept of the Supervised Interval Field (SIF). Finally, Sec. 5 illustrates the overall process through a numerical experiment conducted using a dataset of concrete mix measurements with interval uncertainty to predict concrete strength.

1.1. Interval Field Related Work

The first to incorporate the concept of spatially varying input uncertainty within the IFEM was in (Moens et al., 2011), with subsequent method updates in (Verhaeghe et al., 2011; Imholz et al., 2016). In (Imholz et al., 2016) an interval field concept was developed which allowed to use spatial dependency between elements, which is otherwise not possible using interval parameters. To that end, two methods were introduced: the Local Interval Field Decomposition (LIFD), and the Inverse Distance Weighting Interpolation (IDW). For the LIFD, the dependency parameters are determined from a gradient method developed by Imholz et al., which defines the dependency between parameters as the maximum difference that can occur between them. In this definition, perfect dependency corresponds to a zero gradient and perfect independency corresponds to the maximum gradient between parameters. The dependency parameters are discretely defined at each element of the mesh—thus the dimensionality of the interval field equals the numbers of elements in the mesh. The discrete values for the dependency parameters are determined by a smooth second-order polynomial function presented in (Imholz et al., 2016). For the IDW method, the base functions needed for the interval field are determined and spatially selected based on prior engineering knowledge.

The LIFD is computationally expensive as the dimensionality of the uncertainty equals the number of finite elements in the mesh. The IDW is computationally cheaper, and could produce good results where enough a-priori domain engineering knowledge is available. As a result, expert hand-engineering would be required to produce good results. Both methods were only presented for 1D cases.

(Wu and Gao, 2017) developed a hybrid method for static FE analysis (X-UISS method) that combines random fields and intervals fields. It has an algorithm that easily extends to 2D and 3D domains. For the interval field implementation, the X-UISS method defines the upper-bound UB and lower-bound LB functions from the extrema of a set of measurements at any spatial coordinate. Then, the UB and LB of measured points are linearly interpolated between the $\mathbf{x}$ points where data is unavailable. Then, the conventional IFEM (Muhanna and Mullen, 2001) procedure is implemented using the linearly interpolated values for the domain. The pitfall of this method is that it is computationally expensive by having to solve both SFEM and IFEM methods but more importantly because it uses linear interpolation between data points, which is an oversimplification of the spatial uncertainty variability in a domain.

(Sofi, 2017), (Sofi, 2015) developed a 1D interval field approach for IFEM which uses the extra unitary interval (EUI) in order to reduce the overestimation of the IFEM due to mutual dependency. In contrast to the Hybrid method by Wu et al. and the LIFD method by Moens et al., the dimension

of uncertainty does not depend of the FE mesh size of the model. In that work, the underlying interval field method is based on a closed-form solution of a determinate beam or an approximate solution for statically indeterminate beams. A general EUI method for 2D and 3D domains under general boundary conditions has not been developed to the authors' knowledge.

1.2. Interval Neural Networks Related Work

For deep neural networks, predictive model uncertainty quantification has customarily taken place by using Bayesian Neural Networks (BNNs) or Bayesian approximations. In (Hernández-Lobato and Adams, 2015), the authors present Probabilistic Backpropagation, a Bayesian neural network method to train deep neural networks and obtain parameter uncertainty. Shortly after, (Gal and Ghahramani, 2016) developed a method called Monte Carlo Dropout, which averages Monte Carlo simulations of predictions using Dropout (Srivastava et al., 2014) at test time in order to obtain a predictive uncertainty estimate. In deep reinforcement learning, state uncertainty in Partially Observable Markov Decision Processes (POMDPs) was studied by (Hausknecht and Stone, 2015) using Long Short-Term Memory (LSTM) units in order to *remember* past states.

On the other hand, training deep models end-to-end with uncertain input data has been less studied. Most of the work adds noise to the data with known probability distributions (Czarnecki and Podolak, 2013) or uses sampling methods (McDermott and Wikle, 2019). In many safety-critical applications, we do not have enough information to select a probability distribution for the input, as in conditions of epistemic uncertainty, but can gain access to bounds on the data by expert knowledge or by the simple nature of the data acquisition (Ferson et al., 2007). In such case, when only the bounds of the input data are known, IA (Moore et al., 2009) offers an alternative to model the uncertainty.

More generally, IA algorithms are typically used for rigorous error bounds, result verification, sensitivity analysis, and uncertainty quantification (Moore et al., 2009), (Neumaier, 1990). For uncertainty quantification in particular, IA has been used for inference with imprecise probabilities (Walley, 1991), matrix inequalities (Ben-Tal and Nemirovski, 2002), partial differential equations with finite element methods (Muhanna and Mullen, 2001), and data processing (Kreinovich et al., 2013). Recently, IA has been used as a surrogate to verify DNNs but not in a way that can be directly used for inference (Liu et al., 2019; Adam et al., 2016). Symbolic IA has also been used for security analysis of DNNs (Wang et al., 2018), but not with mathematically rigorous IA algorithms as it's used for other computational applications previously mentioned.

For neural networks (NNs) using IA for inference, prior work that began in the 1990s used interval-valued inputs for shallow NNs for smaller datasets, and with few exceptions not trained using gradient-based optimization. Nonetheless, the developments were valuable to demonstrate the concept of learning with interval-valued inputs. For example, (Ishibuchi and Tanaka, 1991; Hernandez et al., 1993) developed Backpropagation for interval arithmetic using feed-forward two-layer neural networks, (Freitag et al., 2011) developed an algorithms to train recurrent neural networks, (Freitag et al., 2012) trained a recurrent neural network using particle swarm optimization. Other work, has attempted to train neural networks with interval values but without using IA, which makes it is less rigorous (Yang et al., 2019)

In fact, some of this work carried on until recently, using older neural network architectures and optimization methods (Yang and Wu, 2012),(Yang and Liu, 2018),(Yang et al., 2019),(Huang et al., 2020). Consequently, we introduce a deep interval neural network (DINN), which is able to predict with interval-valued data and produce interval estimates using interval gradient-based optimization for deep neural networks.

1.3. Data-Driven Methods in Computational Mechanics

Recently, data-driven methods using machine learning (ML) in computational mechanics have emerged for applications from constitute modeling to damage detection. However, a deeper look at most surveyed papers revealed that the data come from simulations instead of real measured data (Korzeniowski et al., 2019), (He and Chen, 2020), (Nuttall, 2018), (Hauseux et al., 2018), with few exceptions (Reynders and De Roeck, 2014; Yeh, 2006). This is more than likely due to the high costs of obtaining data and the limited availability of public datasets in these applications. Using solely simulation data as training data for ML models is very limiting because it impairs generalization of the learning algorithm—the main goal in ML—and biases the predictions to the simulation assumptions. Nonetheless, simulation data can be in fact used to complement real measured data. Overall, it is imperative to address the challenge of gathering data for computational mechanics applications.

2. Uncertainty Modeling in Computational Mechanics

In this section, uncertainty modeling in computational mechanics is described in mathematical form as a prelude to the DINN development. In the forward problem, a system analysis is typically performed with a finite element (FE) model to determine the state of the system. The FE model receives as input the physical properties of the system. As output, the FE model computes the resulting responses of the system (e.g., displacements, stresses, and vibration frequencies).

Let's now assume that the FE inputs contain interval uncertainty and the system analysis is performed using the IFEM. The stages in the structural forward problem under interval uncertainty is shown on Fig. 1. The input uncertainty is in the spatiotemporal variation of the IFEM input variable as well as the uncertainty in the measurements. Let's also assume that we cannot directly measure the unknown input property but only have access to indirect measurements. We will quantify the input uncertainty in the described forward problem in the following fashion. Let Y be an uncertain system input to the IFEM, such as the material elastic modulus, that cannot be directly measured. Let X be a set of observations that characterize Y indirectly, such as physical properties captured by sensors. Often times, the relationship between X and Y is unknown. Other times, it is obtained through empirical engineering formulas that generally make simplifying assumptions. On the other hand, we can use a predictive model f to find the relationship between X and Y, such that $f : X \rightarrow Y$, to obtain an estimate of the sought-after property $\widehat{Y}$. The measured data for X has inherent aleatory and epistemic uncertainty due to error in the measurement devices, randomness, lack of knowledge about the data generating process, imprecision, ignorance, and poor understanding of physics phenomena. Additionally, there has to be spatial and time coherence in

neighboring points that we seek to also capture with the predictive model f. In this work, the DINN is the choice of predictive model f to quantify the uncertain input properties for the IFEM.

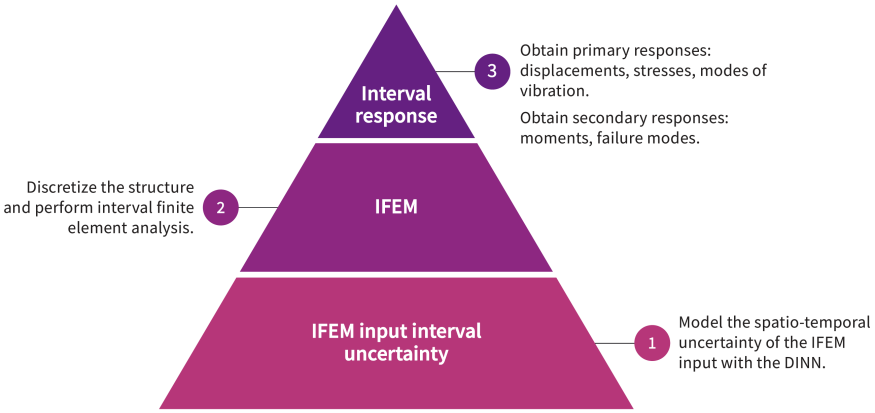


Figure 1. Stages of Forward Problem. The DINN quantifies the IFEM input uncertainty.

In solid mechanics, we often do not have enough knowledge about the probability distributions of X and Y. Indeed, an exclusively probabilistic approach would require to make assumptions on the distributions without enough knowledge, which does not always result in reliable predictions. Therefore, we instead use interval uncertainty on the values of X to obtain interval predictions Y using the DINN—which does not require probabilistic assumptions but only upper and lower bounds of the input.

3. DINN for Computational Mechanics

In this section, the details of the DINN development are explained. The goal of the DINN for this work is to predict an interval field of an IFEM input property using indirect observations $\mathbf{X}$ (i.e., indirect sensor measurements). The DINN is the interval extension of a real-valued DNN to process interval-valued matrices of any dimensionality (i.e., interval tensors). The DINN is a supervised learning algorithm in a regression setting which seeks to learn an interval predictive model $F : [\underline{X}, \overline{X}] \to [\underline{Y}, \overline{Y}] \in \mathbb{IR}$. In order to achieve this goal, the DINN is trained using the $\mathbf{X}$ interval-valued features in a d-dimensional space with n samples, along with its known $\mathbf{Y}$ interval-valued targets. This composes the training set $\mathcal{T} = \{(\mathbf{X}_1, \mathbf{Y}_1), \ldots, (\mathbf{X}_n, \mathbf{Y}_n)\}$ where $\mathbf{X} \in \mathbb{IR}^{n \times d}$ is the feature space and $\mathbf{Y} \in \mathbb{IR}^n$ is the target space. As its output, the DINN computes for each sample i an estimate of the interval target $\widehat{F}(\mathbf{X}_i) = \widehat{\mathbf{Y}}$. In order to do so, the training algorithm iteratively reduces the difference between the true known target $\mathbf{Y}_i$ and the prediction target $\widehat{F}(\mathbf{X}_i)$ at each sample i, by minimizing a loss function $\mathcal{L}(\widehat{F}(\mathbf{X}_i), \mathbf{Y}_i; \mathbf{W})$, such as mean squared error, parameterized by $\mathbf{W}$. Hence, we train the DINN to find the set of interval parameters $\mathbf{W}$

that minimize the loss function. After training, each new future sample i is predicted as

$$\widehat{F}(\mathbf{X}_i) = \mathbf{X}_i^{new}\mathbf{W}, \tag{1}$$

where $\mathbf{W}$ are the trained interval weights of the DINN and $\mathbf{X}_i^{new}$ is a future query of unseen values. The ultimate goal of the DINN is to generalize to predictions of unseen uncertain data where target values $\mathbf{Y}$ are not available. The following sections delve into the details of the DINN.

3.1. Fully-Connected Interval Network

For deep neural networks, the predictive interval function $\widehat{F}(\mathbf{X})$ at the output layer L_{out} is composed of the functions of the hidden layers, such that

$$\widehat{F}(\mathbf{X}) = F^{(L)}\Big(F^{(L-1)}\big(F^{(L-2)}\ldots\big(F^{(\ell)}\ldots\big(\widehat{F}(\mathbf{X})\big)\big)\big)\Big), \tag{2}$$

for $\ell = 1\ldots L$, where ℓ is the layer number.

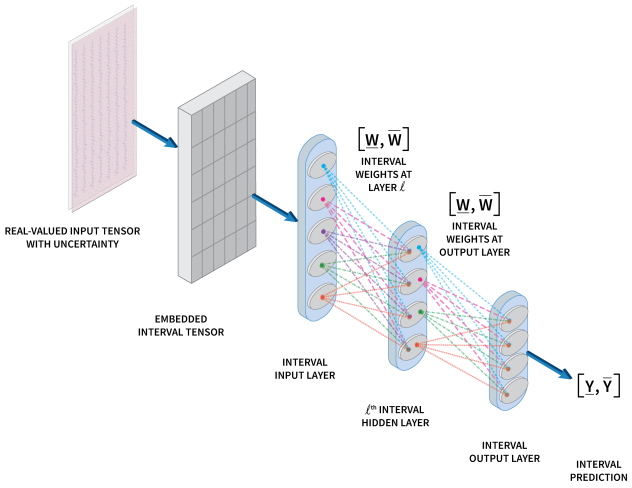


Figure 2. Schematic Representation of DINN.
Nodes represent the activations from previous layers after passing through a `ReLU`
The edges represent the interval weights of the layers.

Hence, the DINN is a composition of interval-valued functions where each of these functions is the output of each of the l layers of the network after passing through a nonlinear activation function. A schematic representation of the DINN is shown on Fig. 2. Deep neural networks have more power to express the functions that they attempt to approximate because of the combinatorial advantage of the composition of functions (Goodfellow et al., 2016). A crucial aspect of the development of the DINN is requiring the nonlinear activation function at each layer $F^{(\ell)}$ be monotonic and Lipschitz so that the final output can also be Lipschitz and as sharp as possible.

3.2. Embedding of Interval Input

The raw real-valued features X and targets Y in the training set are converted as needed to intervals $\mathbf{X}, \mathbf{Y}$ in an embedding layer using the specialized INTLAB toolbox in MATLAB (Rump, 1999), according to their upper and lower bounds. The interval embedding corresponds to layer 0 of the network. Eq. 2 is used in order to define the interval input.

3.3. Training

During training, we find a set of optimal parameters $\mathbf{W} \in \mathbb{IR}$ for each layer of the network by minimizing the loss function. In order to minimize the interval loss function $\mathcal{L}(\widehat{F}(\mathbf{X}_i), \mathbf{Y})$, we use first-order gradient based optimization. In particular, we use mini-batch stochastic gradient descent (SGD) and some of its variants, SGD with momentum (Polyak, 1964) and Adam (Kingma and Ba, 2014). The gradient descent update rule at step $k+1$ for sample i at layer l is defined as

$$\mathbf{W}_{k+1}^{(\ell)} = \mathbf{W}_{k}^{(\ell)} - \alpha_k \nabla_{\mathbf{W}_k}^{(\ell)} \mathcal{L}(\widehat{F}(\mathbf{X}_i), \mathbf{Y}_i), \tag{3}$$

$$\mathbf{b}_{k+1}^{(\ell)} = \mathbf{b}_{k}^{(\ell)} - \alpha_k \nabla_{\mathbf{b}_k}^{(\ell)} \mathcal{L}(\widehat{F}(\mathbf{X}_i), \mathbf{Y}_i), \tag{4}$$

where at each step k, $\mathbf{W}_k$ is the interval weight matrix , $\mathbf{b}_k$ is the interval bias vector, α_k is the step size, $\nabla_{\mathbf{W}_k}\mathcal{L}(\cdot)$ and $\nabla_{\mathbf{b}_k}\mathcal{L}(\cdot)$ are the gradients of the loss function with respect to parameters $\mathbf{W}_k$ and $\nabla_{\mathbf{b}}$. With mini-batch gradient descent, the algorithm chooses uniformly at random a batch of samples of size B from the full training set $\mathcal{T}$ and updates the gradient with Eqs. 3 and 4, which trains the network. There are three major steps to training a deep neural network:

1. Compute the loss at each training iteration via forward propagation of the input activations through the network's layers.

2. Compute the gradient of the loss function with respect to the model's parameters via the Backpropagation algorithm.

3. Minimize the loss function via stochastic gradient descent.

3.3.0.1. *Interval Forward Propagation* Algorithm 1 computes the predictions $\widehat{F}(\mathbf{X})$ and the loss $\mathcal{L}(\widehat{F}(\mathbf{X}_i), y)$ of the DINN through forward-propagation of each mini-batch of data. The layer output $\mathbf{z}^{(\ell)}$ (scores) is computed through an affine transformation as $\mathbf{z}^{(\ell)} = \mathbf{h}^{(\ell-1)}\mathbf{W}^{(\ell)} + \mathbf{b}^{(\ell)}$, at each layer ℓ; $\mathbf{W}^{(\ell)} \in \mathbb{IR}^{H^{(\ell-1)} \times H^{(\ell)}}$, $\mathbf{h}^{(\ell-1)} \in \mathbb{IR}^{B \times H^{(\ell-1)}}$, $\mathbf{b}^{(\ell)} \in \mathbb{IR}^{H^{(\ell)}}$, and $\mathbf{z}^{(\ell)} \in \mathbb{IR}^{B \times H^{(\ell)}}$, $H^{(\ell-1)}$ is the dimension of layer $\ell - 1$, $H^{(\ell)}$ is the dimension of layer ℓ, and B is the number of samples in each mini-batch of data. Then, the layer output $\mathbf{z}^{(\ell)}$ passes through an element-wise activation function $F(\mathbf{z}^{(\ell)})$. The default activation function for the DINN is a `ReLU`, defined as

$$\texttt{ReLU}(\mathbf{z}^{(\ell)}) = \max(\mathbf{z}^{(\ell)}, 0). \tag{5}$$

In the regression setting, for the final layer L, the prediction is simply the affine transformation of final output $\mathbf{z}^{(L)}$, without a `ReLU`. The loss (i.e., error) at the end of each epoch is computed using Mean Squared Error (MSE) for each mini-batch of size B, as

$$\mathcal{L}(\widehat{F}(\mathbf{X}), \mathbf{Y}) = \frac{1}{2B} \sum_{i=1}^{B} (\widehat{F}(\mathbf{X}_i) - \mathbf{Y}_i)^2. \tag{6}$$

The loss is an interval scalar. Notice that `ReLU` is a monotonically increasing function, which allows for calculation of sharp interval enclosures. The `ReLU` is the standard activation function for the DINN but it can be replaced by a different activation function with ease.

Algorithm 1: Forward Propagation Algorithm for Each Mini-batch of Data

Obtain the input raw real-valued features X and targets Y or labels y
Embed X into interval input features $\mathbf{X}$ considering lower $\underline{X}$ and upper bounds $\overline{X}$ for each sample
Embed Y into interval input features $\mathbf{Y}$ considering lower $\underline{Y}$ and upper bounds $\overline{Y}$ for each sample
Choose number of layers L
Require model parameters $\mathbf{W}^{(\ell)}$, for each layer ℓ
$\mathbf{h}^{(0)} = \mathbf{X}$
for $\ell = 1 : L$ **do**
  $\mathbf{z}^{(\ell)} = \mathbf{h}^{(\ell-1)}\mathbf{W}^{(\ell)} + \mathbf{b}^{(\ell)}$
  $\mathbf{h}^{(\ell)} = F(\mathbf{z}^{(\ell)})$
Prediction $\widehat{F}(\mathbf{X}) = \mathbf{h}^{(L)}$
Compute loss $\mathcal{L}(\widehat{F}(\mathbf{X}_i), \mathbf{Y}_i)$

3.3.0.2. *Interval Backpropagation* The first step for SGD is to find the gradients $\nabla_{\mathbf{W}}\mathcal{L}$. These gradients are computed via automatic differentiation in the computational graph of the neural network via Backpropagation. For the DINN, we use the Backpropagation algorithm to take derivatives of interval matrices. In particular, the chain rule is used to compute the gradient as follows

$$\nabla_{\mathbf{W}^{(\ell)}}\mathcal{L} = \frac{\partial \mathcal{L}}{\partial \mathbf{W}^{(\ell)}} = \frac{\partial \mathcal{L}}{\partial \mathbf{z}^{(\ell)}} \frac{\partial \mathbf{z}^{(\ell)}}{\partial \mathbf{W}^{(\ell)}} \tag{7}$$

where at layer ℓ, $\frac{\partial \mathbf{z}^{(\ell)}}{\partial \mathbf{W}^{(\ell)}}$ is the partial derivative of the score with respect to the weights, and $\frac{\partial \mathcal{L}}{\partial \mathbf{z}^{(\ell)}}$ is the partial derivative of the loss with respect to the score at layer ℓ.

This procedure is summarized in Algorithm 2 for general loss and activation functions. In the algorithm, the $\boldsymbol{\delta}$'s denote error terms that recursively update the error in the network at each layer. These interval gradients are accumulated in the $\mathbf{G}$ data structure in the code. Regularization $\alpha\Omega(\mathbf{W})$ is also added to the gradient calculation and explained in the next section. The algorithm requires only an interval subtraction, an interval matrix-vector multiplication, and updating the $\mathbf{G}$ data structure for each mini-batch of data.

Algorithm 2: Backpropagation Algorithm for Interval Loss for Each Mini-batch of Data

Run forward computation algorithm to obtain layer activations $\mathbf{h}^{(\ell)}$
Compute gradient of loss at output layer $\boldsymbol{\delta}_{out}^{(L)}$
$\mathbf{G} = \boldsymbol{\delta}_{out}^{(L)}$
$\nabla_{\mathbf{W}}^{(L)}\mathcal{L} = \mathbf{h}^{(L-1)^T}\mathbf{G} + \lambda\nabla_{\mathbf{W}}^{(L)}\Omega(\mathbf{W})$
$\nabla_{\mathbf{b}}^{(L)}\mathcal{L} = \mathbf{G} + \lambda\nabla_{\mathbf{b}}^{(L)}\Omega(\mathbf{W})$
$\mathbf{G} \leftarrow \mathbf{G}\mathbf{W}^{(L)^T}$
for $\ell = L-1, \ldots, 1$ *layers* **do**
  $\mathbf{G} \leftarrow \boldsymbol{\delta}_{\mathbf{z}}^{(\ell)} = \mathbf{G} \odot F'(\mathbf{z}^{(\ell)})$
  $\nabla_{\mathbf{W}}^{(\ell)}\mathcal{L} = \mathbf{h}^{(\ell-1)^T}\mathbf{G} + \lambda\nabla_{\mathbf{W}}^{(\ell)}\Omega(\mathbf{W})$
  $\nabla_{\mathbf{b}}^{(\ell)}\mathcal{L} = \mathbf{G} + \lambda\nabla_{\mathbf{b}^{(\ell)}}\Omega(\mathbf{W})$
  $\mathbf{G} \leftarrow \mathbf{G}\mathbf{W}^{(\ell)^T}$

For the DINN as presented on this paper, with MSE loss and `ReLU` activations, the partial derivatives in Algorithm 2 specialize as follows. At the final layer L we have

$$\frac{\partial \mathcal{L}}{\partial \mathbf{z}^{(L)}} = \boldsymbol{\delta}_{\mathbf{z}}^{(L)} = \boldsymbol{\delta}_{out}^{(L)} = \frac{\partial \mathcal{L}}{\partial \mathbf{h}^{(L)}}\frac{\partial \mathbf{h}^{(L)}}{\partial \mathbf{z}^{(L)}} = (\mathbf{h}^{(L)} - \mathbf{Y}). \tag{8}$$

For $\frac{\partial \mathbf{z}^{(\ell)}}{\partial \mathbf{W}^{(\ell)}}$, we have

$$\frac{\partial \mathbf{z}^{(\ell)}}{\partial \mathbf{W}^{(\ell)}} = \mathbf{h}^{(\ell-1)^T}. \tag{9}$$

At the hidden layers ($\ell = L-1, L-2, \ldots, 1$), going backwards in the network graph, the partial derivative with respect to the score is

$$\begin{aligned}\boldsymbol{\delta}_{\mathbf{z}}^{(\ell)} = \frac{\partial \mathcal{L}}{\partial \mathbf{z}^{(\ell)}} &= \boldsymbol{\delta}_{\mathbf{z}}^{(\ell+1)}\mathbf{W}^{(\ell+1)^T} \odot \underbrace{F'(\mathbf{z}^{(\ell)})}_{\substack{\text{derivative of activation}}} \\ &= \boldsymbol{\delta}_{\mathbf{z}}^{(\ell+1)}\mathbf{W}^{(\ell+1)^T} \odot \underbrace{\mathbb{1}\{\mathbf{h}^{(\ell)} \geq 0\}}_{\substack{\text{derivative of } \texttt{ReLU} \\ = 1 \text{ if } \mathbf{h}^{(\ell)} \geq 0 \\ = 0 \text{ otherwise.}}}.\end{aligned} \tag{10}$$

where $\mathbf{A} \odot \mathbf{B}$ is the Hadamard (element-wise) product of $\mathbf{A}$ and $\mathbf{B}$, and $\mathbb{1}\{\cdot\}$ is the indicator function.

Finally, the gradient with respect to the weights at each layer is (Eq. 7)

$$\nabla_{\mathbf{W}^{(\ell)}}\mathcal{L} = \mathbf{h}^{(\ell-1)^T}\boldsymbol{\delta}_{\mathbf{z}}^{(\ell)}, \tag{11}$$

and the gradient with respect to the biases at each layer is

$$\nabla_{\mathbf{b}^{(\ell)}}\mathcal{L} = \boldsymbol{\delta}_{\mathbf{z}}^{(\ell)}. \tag{12}$$

3.3.0.3. *Regularization* Regularization is used for two purposes in the DINN. First, regularization is used to prevent overfitting to the training set, with the goal of achieving better performance in unseen data. It achieves so by shrinking the weight contribution of less relevant features. Second, it stabilizes the solution which for finding interval gradients is of crucial importance. Two forms of regularization are used for the DINN. $L1$ (Lasso) regularization uses the l_1, as

$$\Omega(\mathbf{W}) = \|\mathbf{W}\|_1 = \max_{1\leq j\leq n} \sum_{i=1}^{m} |\mathbf{A}_{ij}|, \tag{13}$$

and $L2$ (Ridge) regularization which is defined as

$$\Omega(\mathbf{W}) = \sum_{m} \sum_{n} \mathbf{W}_{m,n}^2. \tag{14}$$

The $L2$ regularization is expressed in a way to directly square the elements of the $\mathbf{W}$ in INTLAB, so as to obtain a monotonic expression. The regularization is added to the data loss, as $J(\widehat{F}(\mathbf{X}), \mathbf{Y}; \mathbf{W}) = \mathcal{L}(\widehat{F}(\mathbf{X}), \mathbf{Y}; \mathbf{W}) + \lambda \cdot \Omega(\mathbf{W})$, where $J(\widehat{F}(\mathbf{X}), \mathbf{Y}; \mathbf{W})$ is the regularized (total) loss and λ is the regularization strength.

4. Supervised Interval Field

At inference time, we query the trained DINN model with new interval data and obtain interval predictions using Eq. 1. These predictions are in turn used to develop the interval field, here named Supervised Interval Field (SIF), by mapping the DINN prediction at each spatial coordinate in the IFEM mesh. Fig. 3 shows the overall process of the DINN in a forward problem. We previously presented a SIF (Betancourt et al., 2018) for the spatiotemporal variation of IFEM input variables though a real-valued deep neural network. In the this work, with the DINN we add the capability to process the interval uncertainty in the domain in addition to the spatiotemporal variation.

4.1. SIF discretization for IFEM

The SIF predictions are independent of the IFEM mesh. Thus a post-processing discretization of the SIF to map to the elements in the mesh of the IFEM is required. For each element in the FE mesh, the average of the SIF predictions covered by the corresponding physical area of the element can be taken. After the discretization, a conventional IFEM analysis can be performed.

5. Experiments

5.1. Concrete Strength Dataset

The concrete strength dataset was retrieved from the UCI machine learning repository (Dua and Graff, 2017) and originated in (Yeh, 1998b). It consists of 1030 samples, each with eight concrete mix features and their corresponding cylinder compressive strength. The eight features are

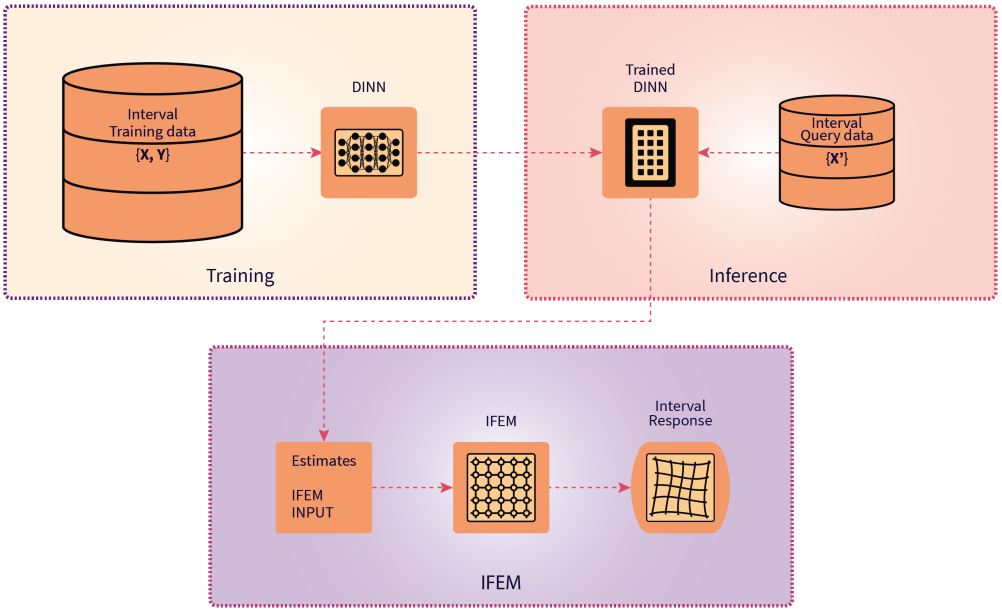


Figure 3. Overall process for forward problem consisting of three stages. **Training**: train the DINN with interval data. **Inference**: Then, make predictions with trained DINN at inference time with new data. **IFEM**:Use the predictions of DINN to form the SIF and perform the IFEM analysis. Finally, get the interval response from the IFEM.

$\mathbf{X} = \{$`Cement, Slag, Fly Ash, Water, Superplasticizer (SP), Coarse Aggregate (CA), Fine Aggregate (FA), Age`$\}$, while the continuous-valued target is $\mathbf{Y} = \{$`Compressive Strength`$\}$. Details about the experiments used to gather the data are included in (Yeh, 1998b). Fig. 4 shows the quartiles of the data and its pair-wise correlations.

Previous works have studied the predictive power of neural networks for this dataset (Yeh, 1998a) (Yeh, 1999), (Yeh, 2003), (Yeh, 2006), (Gal and Ghahramani, 2016). Two important factors area notable in this dataset. First, the data is aggregated from multiple sources, many of which were from experiments done in the 1970s and tested for high strength concrete as defined at the time—which was usually around 40 MPa (5800 psi). Today, high strength concrete is defined by the ACI (363, 2005) as compressive strengths above 55 MPa (8000 psi). Thus, care must be used if predicting concrete strength with *new* high strength data using a trained machine learning model using this dataset. Second, as mentioned in (Yeh, 1998a), the dataset contains uncertainty from experiments being conducted to different standards (i.e., different cylinder sizes), missing features (e.g. Fly Ash content missing for many samples), unspecified type of superplasticizers, unspecified curing temperatures, and lack of knowledge of precision of measurement devices.

The DINN model presented here can have multiple applications from an optimization perspective. However, two very important application can be readily derived from the DINN model. The first, is to use it in a forward problem, to estimate the modulus of elasticity of the structure from

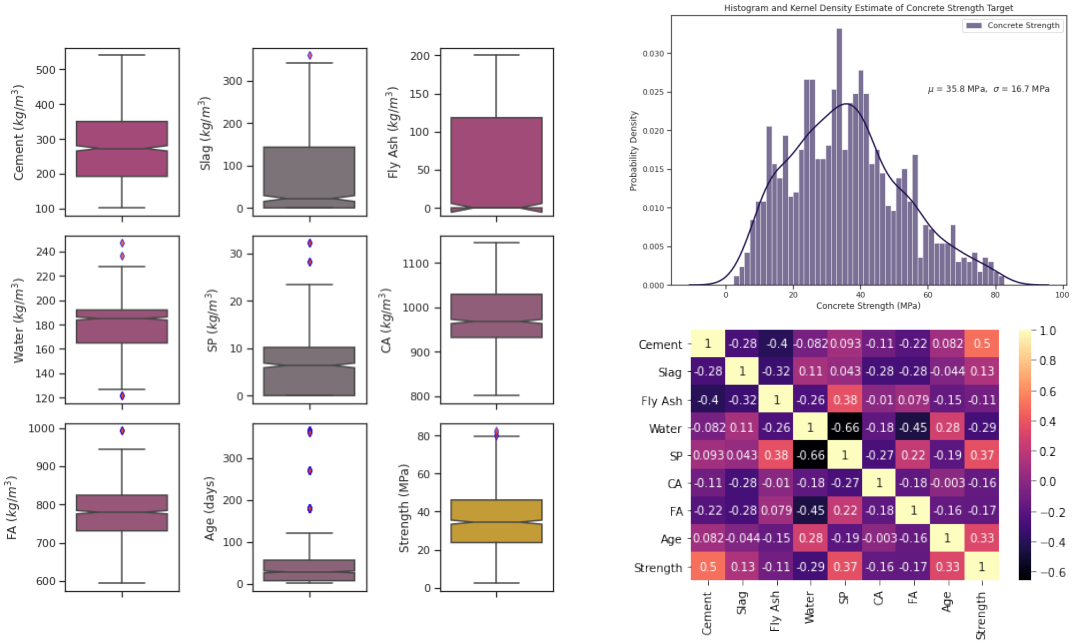


Figure 4. Concrete Strength Dataset Boxplots (left); Concrete Strength (Target) Probability Density (top right); Concrete Strength Dataset feature correlations heatmap (bottom right);

only its mix components. The second, is to design more optimal concrete mixes to achieve structural design concrete strength while decreasing cement content.

5.1.1. *Experimental Details*
The experiments consists of the followings steps (see Fig. 3 for overall process diagram):

1. Perform interval uncertainty quantification on the dataset and convert real-valued data to interval-valued data based on the determined uncertainty levels. This step produces the interval training data (Sec. 3.2).

2. Train the DINN with the interval training data (Sec. 3.3).

3. Perform inference with the trained DINN by querying it with new interval data to obtain the corresponding target predictions (i.e., concrete strength). The results of this query forms the *interval field* which is used as input to the IFEM (Sec. 4).

Once the interval field SIF is computed with the DINN and discretized to map to the IFEM mesh (Sec. 4), we can perform an IFEM analysis to find structural responses.

5.1.1.1. *Interval Uncertainty* Overall, the uncertainty in this dataset is largely epistemic for the reasons given in 5.1. Thus, we use intervals to represent the input uncertainty and propagate it through the DINN by using the procedure in Sec. 3. We use an uncertainty level of $\beta = 15\%$ for

the superplasticizer and $\beta = 5\%$ for all other features in $\mathbf{X}$. We have one set-up with degenerate intervals for the targets $\mathbf{Y}$ (concrete strength), which implies that we have uncertainty in the features but not in the targets.

5.1.1.2. *DINN Model Details* We trained the DINN with SGD and SGD with momentum. We use three hidden layers, 500 units per layer, and 200 epochs. We use randomized search cross-validation (Bergstra and Bengio, 2012) with 10 folds to find optimal hyperparameters for the DINN.

5.1.1.3. *Results* Fig. 5 shows the training loss and the test set predictions. It can be seen that the predictions of the DINN bound the ground truth targets for most test set samples. Additionally, it is observed that the DINN propagated the uncertainty in the inputs without producing overly wide interval predictions.

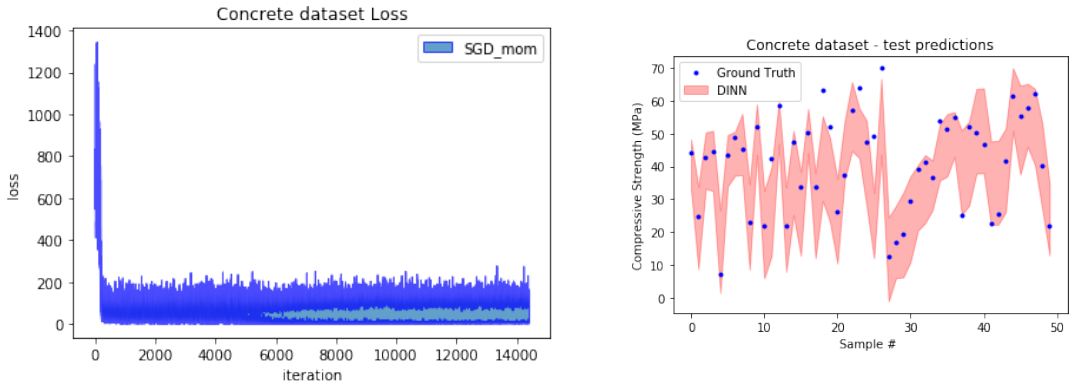


Figure 5. Training loss history (left); Testing Set predictions (right)

6. Conclusion

We have presented a novel interval deep learning algorithm called the Deep Interval Neural Network (DINN), which is capable to make predictions with interval-valued data. Using the DINN, we have presented a method, the SIF, to obtain an interval field from indirect measurements under uncertainty for IFEM input properties.

Our method gains advantage over frequentists or Bayesian methods when we do not have enough information about the uncertainties. It is especially useful in first-of-a-kind safety-critical structural mechanics applications, granted that a data collection program is put into practice.

References

363, A. C. (2005). High-strength concrete. *American Concrete Institute–Special Publication*, 228.

Adam, S. P., Magoulas, G. D., Karras, D. A., and Vrahatis, M. N. (2016). Bounding the search space for global optimization of neural networks learning error: an interval analysis approach. *The Journal of Machine Learning Research*, 17(1):5898–5937.

Ben-Tal, A. and Nemirovski, A. (2002). On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty. *SIAM Journal on Optimization*, 12(3):811–833.

Bergstra, J. and Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13(Feb):281–305.

Betancourt, D., Muhanna, R., and Mullen, R. (2018). Interval field for spatially and temporally dependent uncertainty–machine learning approach. In *Proceedings of the 8th international workshop on reliable engineering computing*, volume 8. University of Liverpool.

Czarnecki, W. M. and Podolak, I. T. (2013). Machine learning with known input data uncertainty measure. In *IFIP International Conference on Computer Information Systems and Industrial Management*, pages 379–388. Springer.

Dua, D. and Graff, C. (2017). UCI machine learning repository.

Ferson, S., Kreinovich, V., Hajagos, J., Oberkampf, W., and Ginzburg, L. (2007). Experimental uncertainty estimation and statistics for data having interval uncertainty. *Sandia National Laboratories, Report SAND2007-0939*, 162.

Freitag, S., Graf, W., and Kaliske, M. (2011). Recurrent neural networks for fuzzy data. *Integrated Computer-Aided Engineering*, 18(3):265–280.

Freitag, S., Muhanna, R. L., and Graf, W. (2012). A particle swarm optimization approach for training artificial neural networks with uncertain data. In *Proceedings of the 5th International Conference on Reliable Engineering Computing, Litera, Brno*, volume 151.

Gal, Y. and Ghahramani, Z. (2016). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, pages 1050–1059.

Ghanem, R. G. and Spanos, P. D. (2003). *Stochastic finite elements: a spectral approach*. Courier Corporation.

Goodfellow, I., Bengio, Y., Courville, A., and Bengio, Y. (2016). *Deep learning*, volume 1. MIT press Cambridge.

Hauseux, P., Hale, J. S., Cotin, S., and Bordas, S. P. (2018). Quantifying the uncertainty in a hyperelastic soft tissue model with stochastic parameters. *Applied Mathematical Modelling*, 62:86–102.

Hausknecht, M. and Stone, P. (2015). Deep recurrent q-learning for partially observable mdps. In *2015 AAAI Fall Symposium Series*.

He, Q. and Chen, J.-S. (2020). A physics-constrained data-driven approach based on locally convex reconstruction for noisy database. *Computer Methods in Applied Mechanics and Engineering*, 363:112791.

Hernandez, C., Espf, J., Nakayama, K., and Fernandez, M. (1993). Interval arithmetic backpropagation. In *Proceedings of 1993 International Conference on Neural Networks (IJCNN-93-Nagoya, Japan)*, volume 1, pages 375–378. IEEE.

Hernández-Lobato, J. M. and Adams, R. (2015). Probabilistic backpropagation for scalable learning of bayesian neural networks. In *International Conference on Machine Learning*, pages 1861–1869.

Huang, D., Fuhg, J. N., Weißenfels, C., and Wriggers, P. (2020). A machine learning based plasticity model using proper orthogonal decomposition. *Computer Methods in Applied Mechanics and Engineering*, 365:113008.

Imholz, M., Faes, M., Cerneels, J., Vandepitte, D., and Moens, D. (2016). On the comparison of two novel interval field formulations for the representation of spatial uncertainty. In *Proceedings of the 7th international workshop on reliable engineering computing*, volume 7, pages 367–378. University of Bochum.

Ishibuchi, H. and Tanaka, H. (1991). An extension of the bp-algorithm to interval input vectors-learning from numerical data and expert's knowledge. In *[Proceedings] 1991 IEEE International Joint Conference on Neural Networks*, pages 1588–1593. IEEE.

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Korzeniowski, T. F., Reppel, T., and Weinberg, K. (2019). A juxtaposition of data driven and stochastic finite element analyses for problems with noisy material data. *PAMM*, 19(1):e201900197.

Kreinovich, V., Lakeyev, A. V., Rohn, J., and Kahl, P. (2013). *Computational complexity and feasibility of data processing and interval computations*, volume 10. Springer Science & Business Media.

Liu, C., Arnon, T., Lazarus, C., Barrett, C., and Kochenderfer, M. J. (2019). Algorithms for verifying deep neural networks. *arXiv preprint arXiv:1903.06758*.

McDermott, P. L. and Wikle, C. K. (2019). Bayesian recurrent neural network models for forecasting and quantifying uncertainty in spatial-temporal data. *Entropy*, 21(2):184.

Moens, D. and Hanss, M. (2011). Non-probabilistic finite element analysis for parametric uncertainty treatment in applied mechanics: Recent advances. *Finite Elements in Analysis and Design*, 47(1):4–16.

Moens, D., Munck, M., Desmet, W., and Vandepitte, D. (2011). Numerical dynamic analysis of uncertain mechanical structures based on interval fields. In *IUTAM symposium on the vibration analysis of structures with uncertainties*, pages 71–83. Springer.

Moore, R. E., Kearfott, R. B., and Cloud, M. J. (2009). *Introduction to interval analysis*, volume 110. Siam.

Muhanna, R. L. and Mullen, R. L. (2001). Uncertainty in mechanics problems—interval–based approach. *Journal of Engineering Mechanics*, 127(6):557–566.

Neumaier, A. (1990). *Interval methods for systems of equations*, volume 37. Cambridge university press.

Nuttall, J. (2018). Estimating spatial correlations from cpt data using neural networks and random fields. In *9th European conference on numerical methods in geotechnical engineering (NUMGE2018), At Porto, Portugal*.

Polyak, B. T. (1964). Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics*, 4(5):1–17.

Reynders, E. and De Roeck, G. (2014). Vibration-based damage identification: the z24 benchmark.

Rump, S. M. (1999). Intlab—interval laboratory. In *Developments in reliable computing*, pages 77–104. Springer.

Sofi, A. (2015). Structural response variability under spatially dependent uncertainty: stochastic versus interval model. *Probabilistic Engineering Mechanics*, 42:78–86.

Sofi, A. (2017). Euler–bernoulli interval finite element with spatially varying uncertain properties. *Acta Mechanica*, pages 1–17.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1):1929–1958.

Verhaeghe, W., Desmet, W., Vandepitte, D., and Moens, D. (2011). Uncertainty assessment in random field representations: an interval approach. In *Fuzzy Information Processing Society (NAFIPS), 2011 Annual Meeting of the North American*, pages 1–6. IEEE.

Walley, P. (1991). *Statistical reasoning with imprecise probabilities*. Chapman & Hall.

Wang, S., Pei, K., Whitehouse, J., Yang, J., and Jana, S. (2018). Formal security analysis of neural networks using symbolic intervals. In *27th {USENIX} Security Symposium ({USENIX} Security 18)*, pages 1599–1614.

Wu, D. and Gao, W. (2017). Hybrid uncertain static analysis with random and interval fields. *Computer Methods in Applied Mechanics and Engineering*, 315:222–246.

Yang, D. and Liu, Y. (2018). L1/2 regularization learning for smoothing interval neural networks: Algorithms and convergence analysis. *Neurocomputing*, 272:122–129.

Yang, D. and Wu, W. (2012). A smoothing interval neural network. *Discrete Dynamics in Nature and Society*, 2012.

Yang, Z., Lin, D. K., and Zhang, A. (2019). Interval-valued data prediction via regularized artificial neural network. *Neurocomputing*, 331:336–345.

Yeh, I.-C. (1998a). Modeling concrete strength with augment-neuron networks. *Journal of Materials in Civil Engineering*, 10(4):263–268.

Yeh, I.-C. (1998b). Modeling of strength of high-performance concrete using artificial neural networks. *Cement and Concrete Research*, 28(12):1797–1808.

Yeh, I.-C. (1999). Design of high-performance concrete mixture using neural networks and nonlinear programming. *Journal of Computing in Civil Engineering*, 13(1):36–42.

Yeh, I.-C. (2003). Prediction of strength of fly ash and slag concrete by the use of artificial neural networks. *J. Chin. Inst. Civil Hydraul. Eng*, 15(4):659–663.

Yeh, I.-C. (2006). Analysis of strength of concrete using design of experiments and neural networks. *Journal of Materials in Civil Engineering*, 18(4):597–604.

Zhang, H. (2005). *Nondeterministic linear static finite element analysis: an interval approach*. PhD thesis, Georgia Institute of Technology.

What's $Z-X$, when $Z = X+Y$?
Dependency tracking in interval arithmetic with bivariate sets

Ander Gray[1,*], Marco De Angelis[2,*], Scott Ferson[3,*] and Edoardo Patelli[4,*,†]
[*] *Institute for Risk and Uncertainty, University of Liverpool, UK*
[†] *University of Strathcylde, UK*

[1] *Ander.Gray@liverpool.ac.uk,* [2] *mda@liverpool.ac.uk*
[3] *Scott.Ferson@liverpool.ac.uk,* [4] *Edoardo.Patelli@strath.ac.uk*

Abstract. In this paper we propose an extension to interval arithmetic to include arithmetic with bivariate sets, which allows for an initial dependency to be propagated, as well as the tracking of complicated dependencies arising from repeated variables. These bivariate sets are represented by Boolean fields which define where two intervals jointly exist or not. We name this extended interval arithmetic *zone arithmetic*. We show how conditional sets may be constructed from bivariate intervals, and how dependent interval arithmetic may be performed with this conditioning. With conditioning we calculate the dependencies between the inputs and outputs of operations, allowing for the extra uncertainty from repeated variables to be greatly reduced. Recent work on using copulas to perform a similar calculations with p-boxes is also reviewed.

Keywords: Uncertainty Propagation, Interval Arithmetic, Imprecise Probabilities, Copulas, Probabilistic Arithmetic, Probability Bounds Analysis, Repeated Variables, Dependency Tracking

1. Introduction

Uncertain computer arithmetics have gained a lot of attention and research over the years (Williamson, 1989; Ferson and Hajagos, 2004; Rump, 1999; Tucker, 2011). They are an intrusive alternative to sampling based methods for uncertainty propagation and quantification. They relay on access to the source code of the computational model. Much like how automatic differentiation calculates the derivative after every line of code, the uncertainty arithmetics calculate the uncertainty. These intrusive methods work by replacing floating point operations with operations defined for uncertain numbers, such as intervals, distributions, fuzzy numbers or p-boxes. This allows for uncertainty to be propagated with the following features:

– Only a single model evaluation needed; albeit more expensive than an evaluations based on floating point.

– A bounded solution to the output uncertainty is achieved. That is, an outer approximation which contracts to the correct result with more computational effort. Interval calculations also allow for automatically verified computation.

© 2021 *by* A. *Gray, M. De Angelis, S. Ferson, E. Patelli.*

– Functions of sets of probability distributions can be computed as cheaply as functions of singular distributions.

One of the main problems for this type of uncertainty propagation is known as the *dependency problem*, or sometimes called *repeated variables* or the *wrapping effect* in interval arithmetic. That when a variable appears multiple times in a sequence of operations, the uncertainty is over-estimated. By the nature of the uncertainty arithmetics, the correct solution is still bounded, however the bounds can be prohibitively large. For example consider the following sequence of binary operations:

$$\begin{array}{l|l} Z = X + Y & [-1,1] + [-1,1] = [-2,2] \\ \tilde{Y} = Z - X & [-2,2] - [-1,1] = [-3,3] \end{array} \tag{1}$$

The left shows a sequence of operations, and the right shows it evaluated with interval arithmetic. X and Y are the intervals $[-1,1]$, and are summed to create Z. X is then subtracted from Z to create $\tilde{Y}$. Clearly since the *same* X was added and then subtracted: $Y = \tilde{Y}$. However the calculation gave us $Y \subset \tilde{Y}$ (i.e. $[-1,1] \subset [-3,3]$). This extra uncertainty came from the fact that X was repeated in the sequence of operations. Operations between uncertain quantities are a function of their dependence. Since Z was created in a binary operation involving X, they are somehow dependent on one another and it was the disregard of this dependence that lead to the over-estimation. The calculation sequence itself introduces a dependence.

There are been several proposed solutions for repeated variables in interval arithmetic, including significance arithmetic (Hyman, 1982), affine arithmetic (Comba and Stolfi, 1993; Rump and Kashiwagi, 2015), zonotopes (Bogomolov et al, 2019) and Taylor models (Makino, 1998). There has also been a method proposed for probability distributions (Li and Hyman, 1998). In this paper we propose an alternative where interval dependencies are propagated and tracked through operations, allowing for an initial interval dependency to be propagated as well as capturing the dependence arising from a sequence of operations with repeated variables. A similar technique for p-boxes but using copulas recently proposed by the authors will also be reviewed.

2. Interval Dependencies

There have been a number of proposed models for dependencies amongst intervals, two which are notable from previous proceedings of the *International Workshop on Reliable Engineering*. Ferson and Kreinovich (2006) describe a dependence between two intervals as regions where the two sets are pair-wise allowed and disallowed (values where the two sets jointly exist). They describe a number of parametric families for interval dependencies, and show how arithmetic operations can be performed with these families. They include *non-interactive* (the standard interval arithmetic case), *perfect* and *opposite* dependencies as special cases. Ceberio et al (2006) similarly describe a joint interval as a set of possible allowed pairs, but allow the set to take any shape. They define a dependence between two intervals x_1 and x_2 to be any proper subset of their Cartesian product $J_{x_1,x_2} \subseteq x_1 \times x_2$. They describe how any arbitrary subset of $x_1 \times x_2$ can be represented on a computer

by a discrete *outer* approximation. They first divide the box $x_1 \times x_2$ into $n \times n$ sub-boxes; and then describe the set J_{x_1,x_2} as the union of sub-boxes which contain a possible pair. This representation of a set is the well known upper approximation used in *rough set* theory (Polkowski, 2002), where dependent interval operations are performed with interval arithmetic on each individual sub-box, much in the style of *sub-intervalisation*. In this section we expand on this, describing in more detail how multivariate intervals may be defined and manipulated. We also describe a method for performing an arithmetic with dependent intervals, which is computationally cheaper than sub-intervalisation and may be used to determine dependencies between inputs and outputs of binary and unary operations, a task which is required for dependency tracking.

2.1. Exclusion Zones

In probability theory, a copula returns a cdf value for a (u, v) pair in the unit square $[0, 1]^2$, $C : [0, 1]^2 \to [0, 1]$, and can be used to define any probabilistic dependency. Following a similar idea, we define any generic interval dependency as an indicator function $\mathbb{Z}(u, v)$ on $[0, 1]^2$ which returns a Boolean value, $\mathbb{Z} : [0, 1]^2 \to \{0, 1\}$, stating whether the pair (u, v) is contained in the bivariate set. Much like how univariate intervals make no statement of how the probability is distributed within them, these bivariate sets only state whether possible pairs are contained or not. Generally the set $\mathbb{Z}$ may be any arbitrarily complicated shape, however we construct an outer approximation of Z by a uniform grid of sub-boxes on $[0, 1]^2$, which may be stored on the computer by a $n \times n$ Boolean field. $\mathbb{Z}$ may also be extended to allow for interval inputs, in which case $\mathbb{Z}$ returns a 1 if *any* real values in the interval are included in $\mathbb{Z}$:

$$Z(u, v) = \begin{cases} 1 & \text{if } (u \times v) \cap \mathbb{Z} \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

Stated another way, Z returns a 0 only if all the values in the interval are not in the original set, and in that sense Z guarantees exclusion. For this reason we label functions like Z as exclusion zones, or *zones* for short. Figure 1 shows a hypothetical interval dependency represented by a zone, with some examples of intervals which would be included and excluded. This set is the union of two rings, and is just an example of a general bivariate set picked to show the flexability of this method. Note that the zone does not need to extend to the entire $[0, 1]^2$ range, which is a modelling choice. Information about the bounds of one variable may inform the bounds of the other, and may not extend to its entire range. Also for practical reasons, any values outside $[0, 1]^2$ will be returned as 0.

Much like how a bivariate distribution may be defined in terms of two marginals and a copula, a joint interval may be defined by two marginal (or univariate) intervals x_1, x_2 and a zone $Z(u, v)$, by the following relation: $J_{x_1,x_2}(x, y) = \operatorname{supp} Z(u, v)$ with $u = (x - \underline{x_1})/(\overline{x_1} - \underline{x_1})$ and $v = (y - \underline{x_2})/(\overline{x_2} - \underline{x_2})$. This defines a grid of sub-boxes following Z by a uniform scaling by the intervals x_1 and x_2, and then taking the support of resulting indicator function. This allows dependencies to be stored and manipulated independently from marginal intervals.

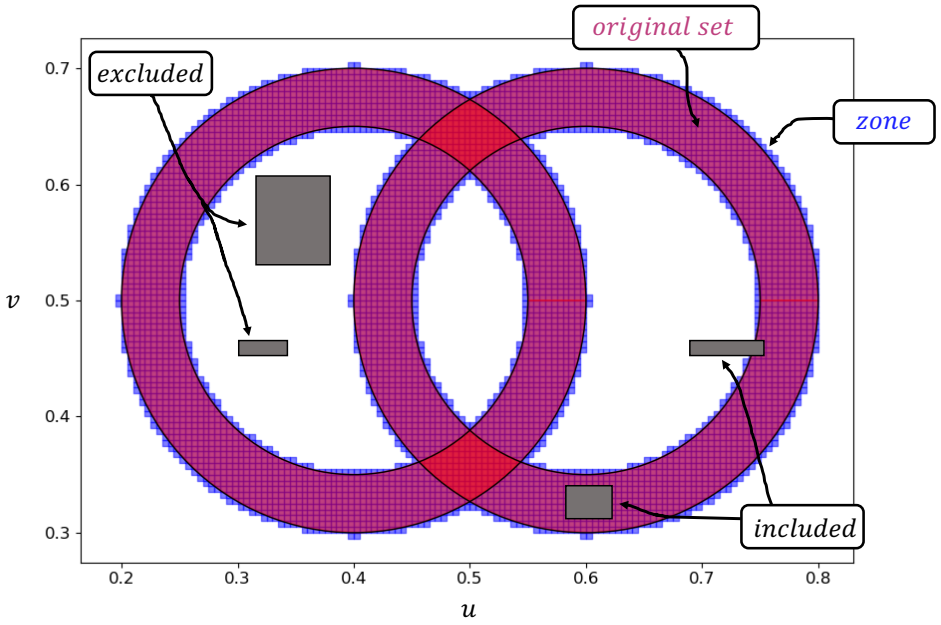


Figure 1. Shows an outer approximation of a set represented by a zone. Also shown are several intervals which would be included and excluded by the zone.

3. Zone Arithmetic

Ceberio et al (2006) suggest that binary operations may be performed with J by performing interval arithmetic on each individual sub-box, which requires of the order of $\sim n^2$ interval operations. We proposed an alternative method which only requires $\sim n$ operations. For this we use another familiar concept for probability theory: *conditioning*. Much like how a univariate distribution can be created from a bivariate distribution by conditioning on a value, it's possible to produce a conditional set from a bivariate set. A conditional set $J(x|\ y = Y)$ may be created from the bivariate interval $J(x, y)$ by intersecting the it with an interval, or real value, Y:

$$J(x|\ y = Y) = J(x, y) \cap ([-\infty, \infty] \times Y) \tag{3}$$

$J(x|\ Y)$ is a univariate set's indicator function, and may be the union of several disjoint intervals. For example, if we condition the zone in figure 1 with $v = [0.48, 52]$, we would get a set made up of 4 disjoint intervals: $[0.195, 0.255] \cup [0.395, 0.455] \cup [0.545, 0.605] \cup [0.745, 0.8]$. Figure 2 shows an illustration of this. Calculations of this sort can be performed quite efficiently in terms of zones, since it's merely a subselection of an array of Z.

We perform dependent interval operations using this conditioning. Instead of n^2 interval operations, we perform n (for each row or column) set operations on the conditionals. Since a conditional may contain more than one interval, multiple interval operations are performed per set operation, however the overall number of interval operations remains lower than n^2. The minimum number of

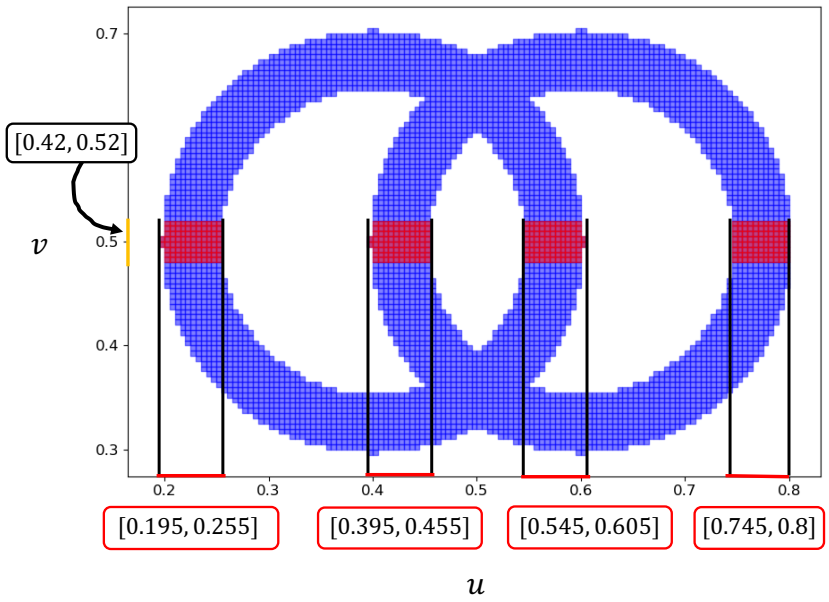


Figure 2. illustration of conditioning a zone with the interval $v = [0.48, 52]$. The resulting set is the union of the disjoint intervals $[0.195, 0.255] \cup [0.395, 0.455] \cup [0.545, 0.605] \cup [0.745, 0.8]$.

interval operations is n, corresponding to the non-interactive case (everywhere true). The maximum number of required interval operations is $\frac{n^2}{2}$, where is a very extreme case where every other element of the zone is empty.

Using conditioning, the dependencies between inputs and outputs of operations may be calculated, which is required for dependency tracking. Take for example the sequence of operations in the title of this paper: $x + y = z$, and then $z - x$, where initially the joint interval J_{XY} is known. For $z - x$ to be evaluated tightly, the joint J_{ZX}, or similarly the zone Z_{ZX}, must be calculated. For this we condition J_{XY} on some subinterval $x = X$, and evaluate the $x + y$ with the conditional set, which gives $J_{ZX}(z|x = X)$. This is then repeated for n subintervals of x, and the joint J_{ZX} can be calculated as the union of the resulting conditionals of z:

$$J_{ZX}(z|x = X) = J_{YX}(y|x = X) + X \tag{4}$$

$$J_{ZX}(z, x) = \bigcup_{i=1}^{n} J_{ZX}(z|x = X_i) \tag{5}$$

where the $+$ operator is an interval sum, and where X_i is a subinterval. Figure 3 shows an example of this, where $x + y = z$ is evaluated for $x, y = [-1, 1]$ and with J_{XY} having the complement of a ring shaped dependence shown on the left. The calculated interval $z = [-2, 2]$ as expected, however the dependence has also been propagated. If $\tilde{y} = z - x$ is evaluated using zone arithmetic, then $\tilde{y} = [-1.08, 1.08]$, which tightly encloses the correct solution of $\tilde{y} = y = [-1, 1]$, and a vast improvement of the standard interval arithmetic answer of $\tilde{y} = [-3, 3]$. The tightness of the result

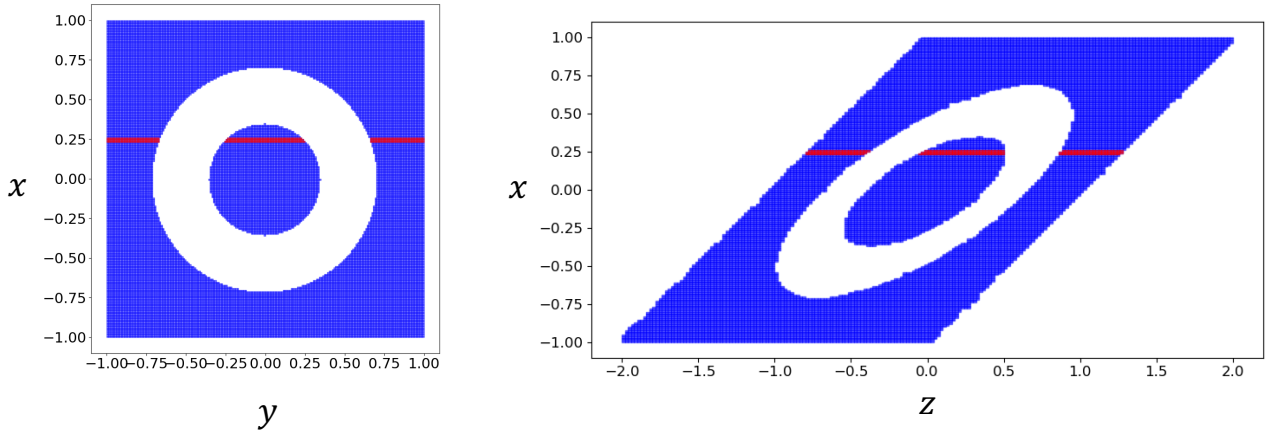


Figure 3. An example of using conditional arithmetic to propagate an interval dependency though a binary operator. Shown is $x + y = z$, with $x, y = [-1, 1]$ having the dependency shown on the left figure. Right shows the calculated bivariate set J_{ZX}. Shown in red is a conditional set, conditioned on a particular subinterval of x, which is used to construct J_{ZX}.

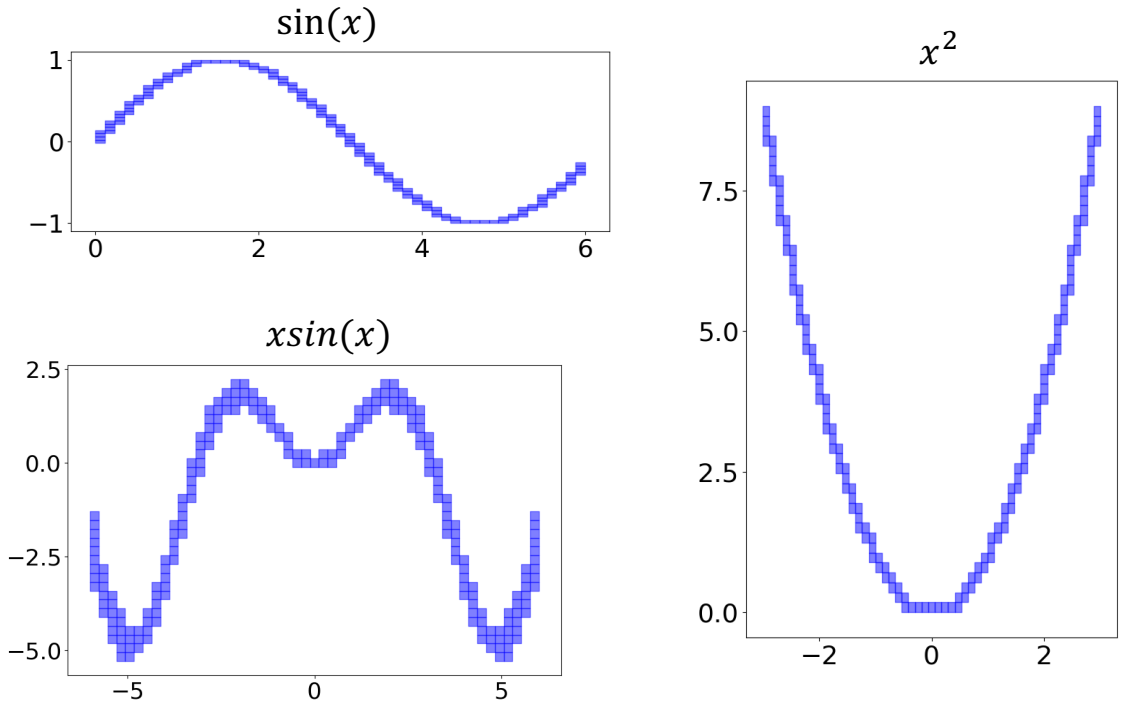


Figure 4. Examples of dependency tracking through unary functions, for discretisations 50×50. Shown is $\sin([0, 6])$ (top left) and $[-3, 3]^2$ (right). Also shown is $x \sin(x)$ for $x = [-6, 6]$ on the bottom left.

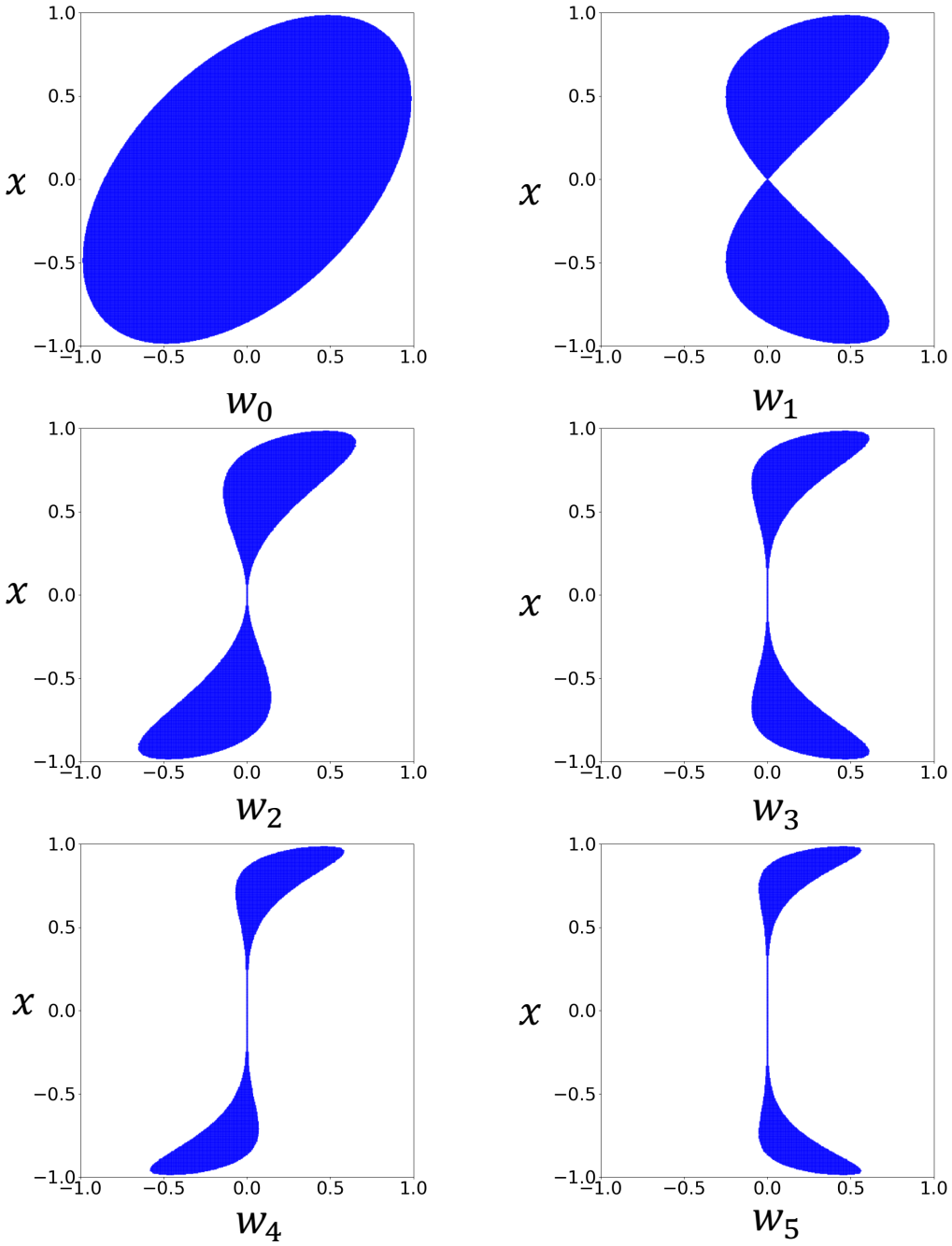


Figure 5. Shown is the evolution of an ellipse through the repeated evaluation of $w_{i+1} = w_i * x$

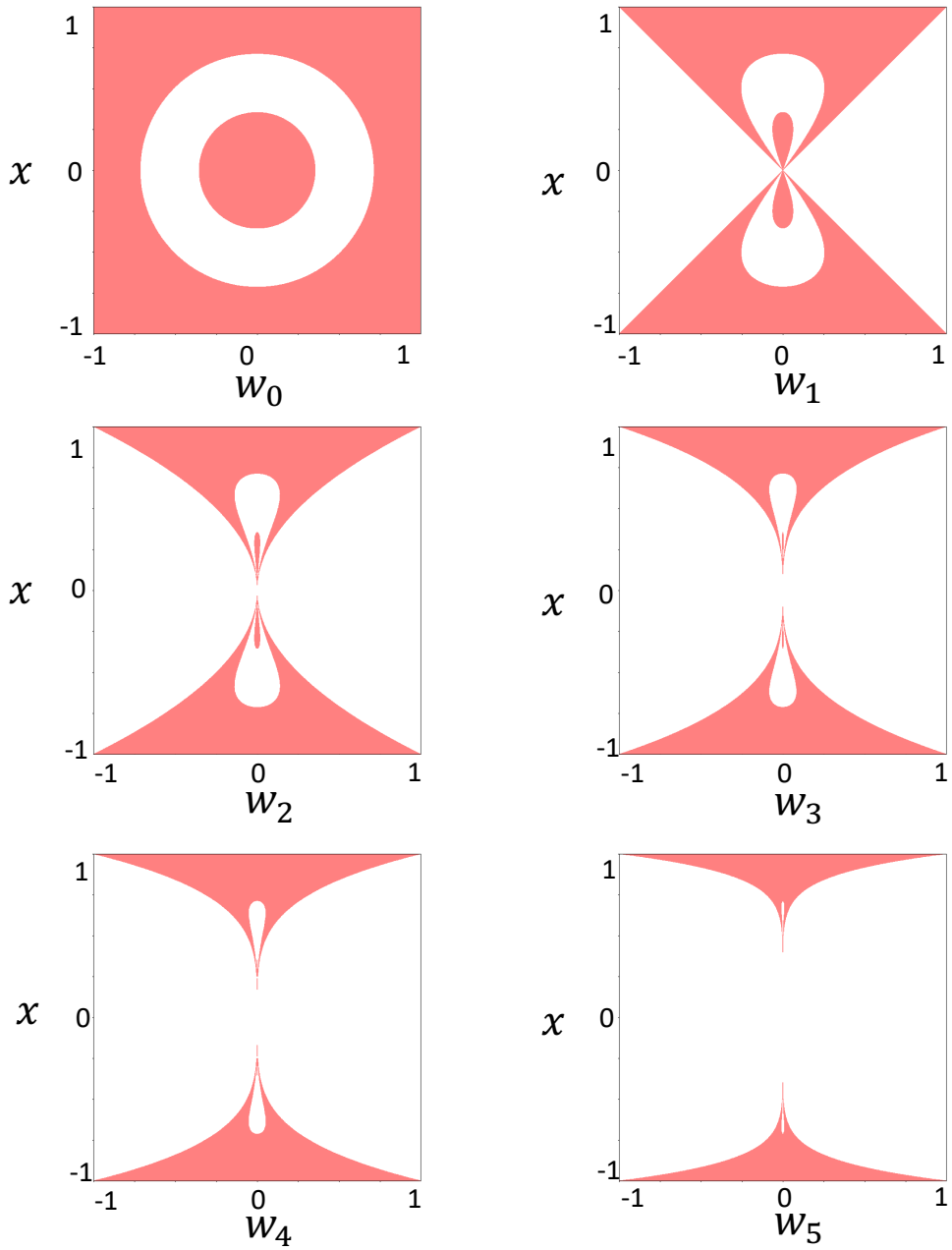


Figure 6. Shown is the evolution of an interval dependence through the repeated evaluation of $w_{i+1} = w_i * x$, starting from the complement of a ring shaped dependence in the top left.

depends on the discretisation used, and for this example a discretisation 200×200 was used. Using fewer sub-boxes will compute the result quicker and lead to a wider result, however it is still guaranteed to bound the correct dependency and marginal. A similar calculation can be performed for unary functions such as $\sin(x)$ or x^2, which is shown in figure 4. It can be seen that the shape of the unary function is enclosed. This method can also be used to propagate dependencies further, such as through $x \sin(x)$ as shown on the bottom left in figure 4. Figure 5 shows the tracking of an elliptical set through repeated evaluation of $w_{i+1} = w_i * x$ for an initial. It can be seen that the dependency does not remain an ellipse. A similar calculation is shown on figure 6 but with an initial ring shaped dependence. It can be seen that this method can tightly track the quite complex dependencies that arise from repeated variables.

4. Copulas and p-boxes

A similar method for dependency tracking with p-boxes has recently been proposed by the authors, and is available through the Julia package *ProbabilityBoundsAnalysis.jl* (Gray, 2020). Here we will review the main elements of the method.

A binary operation of two random variables X and Y with a known joint distribution can be computed by the following Lebesgue-Stieltjes integral or *convolution* (Williamson and Downs, 1990):

$$\sigma_{C,L}(F, G)(x) = \int_{L\{x\}} dC(F(u), G(v)) \tag{6}$$

which is called the σ convolution, written for any two arbitrary inputs cdfs F and G, with a copula C and a binary operation L. The integration domain is $L\{x\} = \{(u, v)|u, v \in \Re, L(u, v) < x\}$, and where $L(u, v)$ is a binary operation, for example $L(u, v) = u+v$. The σ convolution computes the cdf of a binary operation of two distribution functions with a known joint distribution, and is effectively the integral which is approximated when Monte Carlo simulation is performed. A σ convolution may also be perform with p-box inputs (Williamson, 1989). Computing a σ convolution requires a copula C to be specified, which may be available at the beginning of a sequence of operations, but will be lost as the calculation sequence is executed. This leads to the only option being a *Fréchet* convolution, which calculates a binary operation between p-boxes for all possible C. This however, as in the interval case, leads to an overestimation of the uncertainty.

This may be overcome if C is also propagated through binary operations. For some binary operation between random variables $Z = X \,\square\, Y$, where $\square$ is some binary operator (eg $\square = +$), the joint distribution of Z and X is given by:

$$F_{ZX}(z, w) = \iint_{D_{zw}} dF_{XY}(x, y) \tag{7}$$

where the area of integration D_{zw} is the region in the (x, y) plane where $x \,\square\, y < z$ and $x < w$ (Williamson, 1989). Using Sklars theorem:

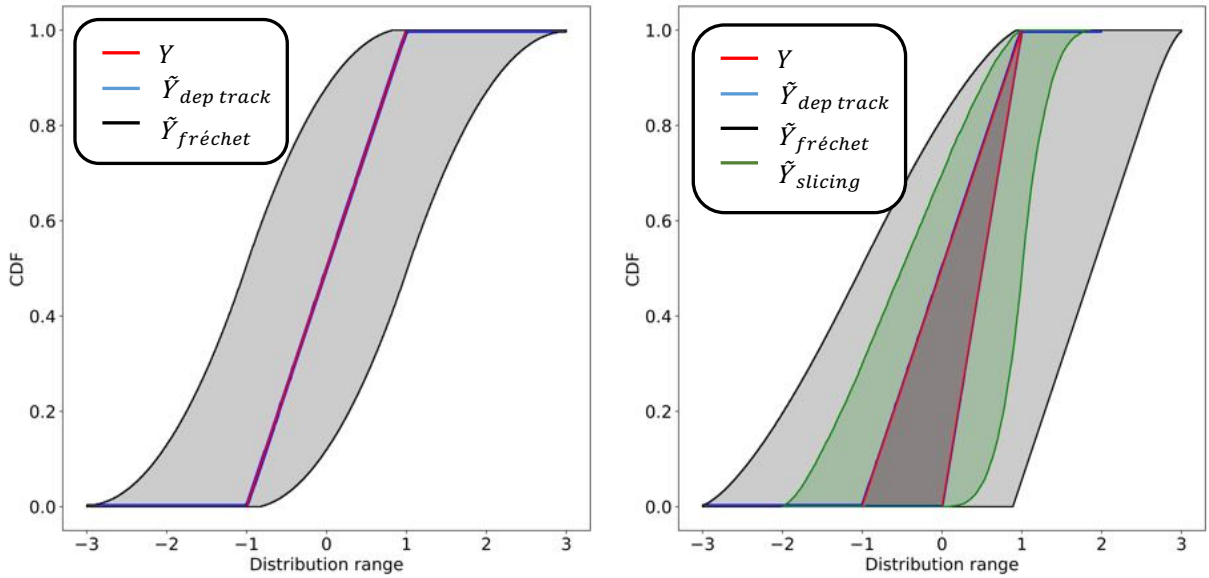


Figure 7. Left shows the calculation sequence of equation 1, with $X, Y \sim U(0,1)$ and independent. The right shows the result the sequence of binary operations but with X and Y starting as the p-boxes $U([-1,0],1)$ and with a gaussian copula with correlation $= 0.5$. The Fréchet result is plotted in both case to show the contraction in uncertainty. The green p-box is the result of random slicing and interval arithmetic.

$$C_{ZX}(u,v) = \iint\limits_{D_{F_Z^\wedge(u),F_X^\wedge(v)}} dC_{XY}(F_X(x),F_Y(y)) \tag{8}$$

where the area of integration is now where $x \,\square\, y < F_Z^\wedge(u)$ and $x < F_X^\wedge(v)$, where $F_Z^\wedge(u)$ is the inverse cdf of Z, which can be found by first computing a σ convolution. Equation 8 may be quite computationally difficult to compute due to the complex integration region. However since the copula is a cdf, it is possible to perform this calulation iteratively, reusing previously calculated values. For some step δ:

$$C_{ZX}(u+\delta,v+\delta) = C_{ZX}(u,v) - C_{ZX}(u+\delta,v) - C_{ZX}(u,v+\delta) + \iint\limits_{A} dC_{XY}(F_X(x),F_Y(y)) \tag{9}$$

with the integration region A being where $F_Z^\wedge(u) \le x \,\square\, y < F_Z^\wedge(u+\delta)$ and $F_X^\wedge(v) \le x < F_X^\wedge(v+\delta)$. Equation 9 uses the 2-d version of the *H-volume* formula (Schweizer and Sklar, 2011), which can be used to find the probability mass in some hyper-rectangle from a n-d cdf. An identical calculation can be performed for Y.

The above methodology was applied to the problem posed in the title of the paper, the sequence of binary operations 1. That is the copula C_{ZX} was calulated for $Z = X + Y$ with X and Y being independently and uniformly distributed $U(-1,1)$. Figure 7 shows the results of using this method.

F_z was first calculated with $\sigma_{\pi,+}(F_X, F_Y)$, the copula C_{ZX} was then calculated with equation 8, defining the calculation dependency between Z and X. $F_{\tilde{Y}}$ was then calulated with $\sigma_{C_{ZX},-}(F_Z, F_X)$. It can be seen that $\tilde{Y}$ and Y are the same, plus some outward directed numerical error from the calculation. The Fréchet calculation is also plotted to show the reduction in uncertainty from dependency tracking. This example was with two independent uniform distributions, however any copula may be initially used. The initial variables may even be p-boxes. The right of figure 7 shows the same calculation but with X and Y being the p-boxes $U([-1, 0], 1)$ and with an initial Gaussian copula with correlation of 0.5. Again it can be seen that the calculation with dependency tracking tightly encloses the correct result. The green p-box is the result of random slicing, where random intervals are generated from X and Y using the gaussian copula. The sequence of binary operations is then evaluated with interval arithmetic and resulting p-box is constructed from the random set. It can be see that random slicing is an improvement over Fréchet (although more expensive to evaluate), but also suffers from repeated variables, since there is no dependency tracking in the interval calculations. Note that when this calculation is performed with p-boxes, equation 8 must be evaluated twice, for the upper and lower copula bound.

5. Conclusions

In this paper an extension to interval arithmetic to include dependence was proposed, where interval dependencies were defined in terms of bivariate sets. These bivariate sets define where two intervals co-exist or not, and may be arbitrarily complex with continuous boundaries. We represent such sets on the computer with finite Boolean fields, like an $n \times n$ Boolean matrix, which defines a grid of sub-intervals. For continuous sets, we create an discrete outer approximation which is guaranteed to enclose all elements of the continuous set. Much in the style of copulas, we describe how interval dependencies may be modelled independently from univariate intervals. We describe how conditional sets are constructed by conditioning bivariate sets on real numbers or intervals, and describe a dependent interval arithmetic in terms of conditioning. Dependent interval arithmetic by conditioning requires of the order of $\sim n$ operations, which is an improvement over the standard *sub-intervalisation* which is around $\sim n^2$ interval operations. We show that when interval dependencies are propagated and calculated this way, the repeated variable problem present in standard interval arithmetic is greatly reduced, and that this method can calculate the complex dependencies that arise in computational models with repeated variables. A similar method for p-boxes using copulas was described, and we show that copulas are capable at modelling the probabilistic dependencies that arise from repeated variables in computations involving distributions and p-boxes. For this task, we calculate the copulas between the inputs and outputs of binary operators involving p-boxes, capturing the dependencies arising from the operation. Uncertainty is greatly reduced when dependency arising from the computation model is tracked when compared to Fréchet convolutions or propagation with random slicing.

Acknowledgements

The authors would like to thank the gracious support from the EPSRC iCase studentship award 15220067. We also gratefully acknowledge funding from UKRI via the EPSRC and ESRC Centre for Doctoral Training in Risk and Uncertainty Quantification and Management in Complex Systems. This research is funded by the Engineering Physical Sciences Research Council (EPSRC) with grant no. EP/R006768/1, which is greatly acknowledged for its funding and support. This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

References

Bezanson, J., Edelman, A., Karpinski, S., Shah, V. B., Julia: A fresh approach to numerical computing, *SIAM review* 59 (1) (2017) 65–98. https://doi.org/10.1137/141000671

Bogomolov, S., Forets, M., Frehse, G., Potomkin, K., Schilling, C., Juliareach: a toolbox for set-based reachability, in: Proceedings of the 22nd ACM International Conference on Hybrid Systems: Computation and Control, 2019, pp. 39–44.

Ceberio, M., Ferson, S., Kreinovich, V., Chopra, S., Xiang, G., Murguia, A., Santillan, J., How to take into account dependence between the inputs: From interval computations to constraint-related set computations, with potential applications to nuclear safety, bio- and geosciences, RELIABLE ENGINEERING COMPUTING (2006) 126.

Comba, J. L. D., and Stolfi, J., A new arithmetic and its applications to computer graphics, In Proceedings of VI SIBGRAPI (Brazilian Symposium on Computer Graphics and Image Processing) (pp. 9-18)

Ferson, S., Kreinovich, V., Modeling correlation and dependence among intervals, RELIABLE ENGINEERING COMPUTING (2006) 115.

Ferson, S. and Hajagos, J. G., Arithmetic with uncertain numbers: rigorous and (often) best possible answers. *Reliability Engineering & System Safety*, 85(1-3):135–152, 2004.

Gray, A., https://github.com/AnderGray/ProbabilityBoundsAnalysis.jl: Imprecise copulas (2020) doi: 10.5281/ZENODO.4281562

Gray, N. and De Angelis, M. and Ferson, S., *COMPUTING WITH UNCERTAINTY: INTRODUCING PUFFIN THE AUTOMATIC UNCERTAINTY COMPILER*. UNCECOMP, 2019.

Hyman, J. M., Forsig: an extension of fortran with significance arithmetic, Tech. rep., Los Alamos National Lab., NM (USA) (1982).

Li, W., Hyman, J. M., Computer arithmetic for probability distribution variables, Reliability Engineering & System Safety 85 (1-3) (2004) 191–209.

Makino, K., Rigorous analysis of nonlinear motion in particle accelerators. PhD thesis, Michigan State University, 1998.

Polkowski, L., Rough sets, Mathematical foundations. Physica Verlag, A Springer-Verlag Co., Heidelberg, Springer, 2002.

Rump, S. M., Intlab—interval laboratory, in: Developments in reliable computing, Springer, 1999, pp. 77–104.

Rump, S. M., Kashiwagi, M., Implementation and improvements of affine arithmetic, Nonlinear Theory and Its Applications, IEICE 6 (3) (2015) 341–359.

Schweizer, B. and Sklar, A. *Probabilistic metric spaces*. Courier Corporation, 2011.

Tucker, W., Validated numerics: a short introduction to rigorous computations, Princeton University Press, 2011.

Williamson, R. C., *Probabilistic arithmetic*. PhD thesis, University of Queensland Australia, 1989.

Williamson, R. C. and Downs, T. Probabilistic arithmetic. i. numerical methods for calculating convolutions and dependency bounds. *International journal of approximate reasoning*, 4(2):89–158, 1990.

Forward interval propagation through the discrete Fourier transform

Marco De Angelis
Institute for Risk and Uncertainty, University of Liverpool, United Kingdom.
marco.de-angelis@liverpool.ac.uk

Marco Behrendt
Institute for Risk and Reliability, Leibniz University Hannover, Germany.
behrendt@irz.uni-hannover.de

Liam Comerford
Institute for Risk and Uncertainty, University of Liverpool, United Kingdom.
l.comerford@liverpool.ac.uk

Yuanjin Zhang
School of Safety Science and Emergency Management, Wuhan University of Technology, China.
ylzhyj@126.com

Michael Beer
Institute for Risk and Reliability, Leibniz University Hannover, Germany.
beer@irz.uni-hannover.de

Abstract. In this paper an algorithm for the forward interval propagation on the amplitude of the discrete Fourier transform (DFT) is presented. The algorithm yields best-possible bounds on the amplitude of the DFT for real and complex valued sequences. We show that computing the exact bounds for the amplitude of the DFT can be achieved with an exhaustive examination of all possible corners of the interval-shaped domain. However, because the number of corners increases exponentially with the number of intervals, such method is infeasible for large interval signals. We provide an algorithm that does not need such an exhaustive search, and show that the best possible bounds for the amplitude can be obtained propagating complex pairs only from the convex hull of endpoints at each term of the Fourier series. Because the convex hull is always tightly inscribed in the respective rigorous bounding box resulting from interval arithmetic, we conclude that the obtained bounds are guaranteed to enclose the true values.

Keywords: Complex intervals; Dependency tracking; Convex hull; Interval algorithm.

1. Introduction

The analysis of engineering structures subject to harmonic stochastic excitation is often conveniently done in the Fourier domain (Newland(1993)). The response of the structure is obtained for each harmonic without the need for step-wise numerical integration. This kind of analysis however, seems to be incompatible with the fact that real data time-series records are often affected by

© 2021 *by M. De Angelis, M. Behrendt, L. Comerford, Y. Zhang, M. Beer.*

uncertainty. Some authors have successfully propagated interval uncertainty in dynamic structural analysis directly in the time domain (Muscolino et al.(2012)). In other instances, the literature on the topic seem to contemplate only purely probabilistic methods to address the problem of characterization of the uncertainty of the power spectrum from real data, as shown in (Comerford et al.(2015)Comerford, Kougioumtzoglou, and Beer) and literature therein. It is currently very difficult to convert an imprecise time signal to its Fourier domain. An algorithm for the interval Fourier transform is much needed to bridge the divide between imprecise signals and the well established harmonic structural analysis, and to enlarge the spectrum of engineering applications that explicitly account for uncertainty.

The discrete Fourier transform (DFT) is used in a variety of different applications in science and engineering, such as in spectral analysis, random vibration, differential equations, data compression, signal processing, image processing, probabilistic programming and many more (Sneddon(1995); James(2011); Newland(1993)). There are various algorithms available for transforming a signal with the DFT, of which the best known is probably the fast Fourier transfom (FFT), presented by Cooley & Tukey (Cooley and Tukey(1965)). Due to increasing computational power, simulations and equivalent calculations can be carried out ever faster, which is particularly important for the numerical analysis. An overview of a variety of algorithms used can be found in abundance in (Cormen et al.(2009)Cormen, Leiserson, Rivest, and Stein; Stoer and Bulirsch(2013)).

Real data records are required for a large number of problems in engineering. A major challenge with data acquisition that should not be neglected is that sensors only work accurately within certain tolerances and are therefore subject to uncertainties. These uncertainties arise, for instance, due to damaged sensors, device failures and measurement errors, or if the sensors are not precisely calibrated. In addition, the data could be captured incorrectly due to threshold limitations. For more accurate simulation results when utilising real data records, uncertainties must be taken into account (Comerford et al.(2015)Comerford, Kougioumtzoglou, and Beer).

In this work, three different methods are investigated to provide bounds on the amplitude of the DFT: (1) a brute-force method, (2) a method based on the convex hull, (3) a method based on complex interval arithmetic and rigorous bounding (Alefeld and Herzberger(2012); Moore(1966); Moore(1979); Moore et al.(2009)Moore, Kearfott, and Cloud). These methods provide a direct relationship between interval signals and the corresponding amplitudes of the DFT. Thus, intervals can replace the exact values when investigating structures while the signal's uncertainty can be captured by these intervals.

2. Problem statement

In this paper the focus is on interval-valued real signals denoted by$\overline{\underline{x}}_n$ (Alefeld and Herzberger(2012)). An interval-valued real number, or real interval, or simply an *interval* is a non-empty, closed, and bounded subset of the real line $\mathbb{R}$,

$$\overline{\underline{x}} := [\underline{x}, \overline{x}] := \{x \in \mathbb{R} | \underline{x} \leq x \leq \overline{x}\}.$$

Similarly a *complex interval* denoted by $\overline{\underline{z}}$, can be defined with a pair of intervals, one for the real component $\overline{\underline{z}}_{\mathrm{re}}$ and the other for the imaginary component $\overline{\underline{z}}_{\mathrm{im}}$,

$$\overline{z} := \overline{z}_{\text{re}} + i\ \overline{z}_{\text{im}} := \{z_{\text{re}} + i\ z_{\text{im}} \mid z_{\text{re}} \in \overline{z}_{\text{re}} \subseteq \mathbb{R}\ \&\ z_{\text{im}} \in \overline{z}_{\text{im}} \subseteq \mathbb{R}\},$$

for more about the space of complex intervals see (Alefeld and Herzberger(2012)). The interval extension of the DFT converts an interval signal (or ordered sequence) $\overline{x}_0, \overline{x}_1, ..., \overline{x}_n$ with $n < N$, into the respective signal of interval complex numbers $\overline{z}_0, \overline{z}_1, ..., \overline{z}_k$, with $k < N$, also known as Fourier sequence

$$\overline{z}_k := \sum_{n=0}^{N-1} \overline{x}_n\ e^{-\frac{i2\pi}{N}kn} = \sum_{n=0}^{N-1} \overline{x}_n\ \left(\cos\frac{2\pi}{N}kn - i\ \sin\frac{2\pi}{N}kn\right). \tag{1}$$

When x is a time signal, each complex number $z_k \in \mathbb{C}$ represents an harmonic of the signal with angular frequency $2\pi kn/N$. The DFT function is sometimes denoted by $\mathcal{F}$, while its interval extension by $\Diamond\mathcal{F}$, where $\overline{z}_k = \Diamond\mathcal{F}\,(\overline{x}_k)$.

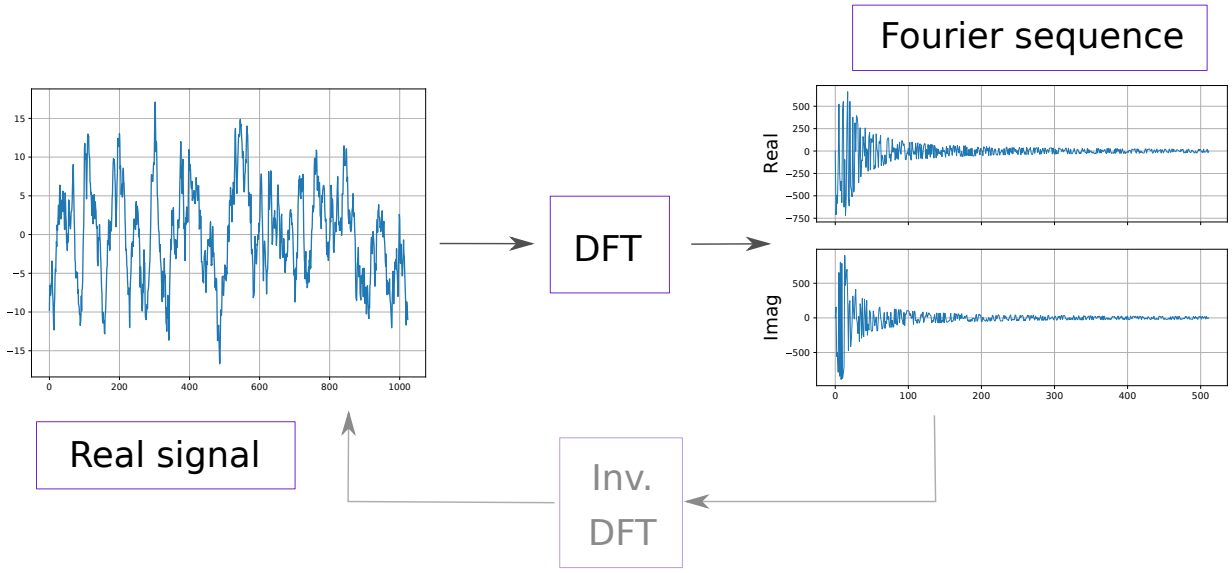


Figure 1. Discrete Fourier transform of a real signal to its Fourier sequence.

The DFT in practice is used to convert a time signal into its Fourier sequence (Figure 1) to study the harmonic properties of the signal where each element corresponds to a frequency. The DFT is often used to reduce the dimensionality of the signal; the inverse Fourier transform is used to reconstruct the signal after compression or simply to convert the signal into the time domain.

The objective of this work is to provide best-possible bounds on the interval extension for the *amplitude* of $\mathcal{F}$,

$$\text{abs}\,(\overline{z}_k) := |\overline{z}_k| := \sqrt{\left(\sum_{n=0}^{N-1} \overline{x}_n\ \cos\frac{2\pi}{N}kn\right)^2 + \left(\sum_{n=0}^{N-1} \overline{x}_n\ \sin\frac{2\pi}{N}kn\right)^2}. \tag{2}$$

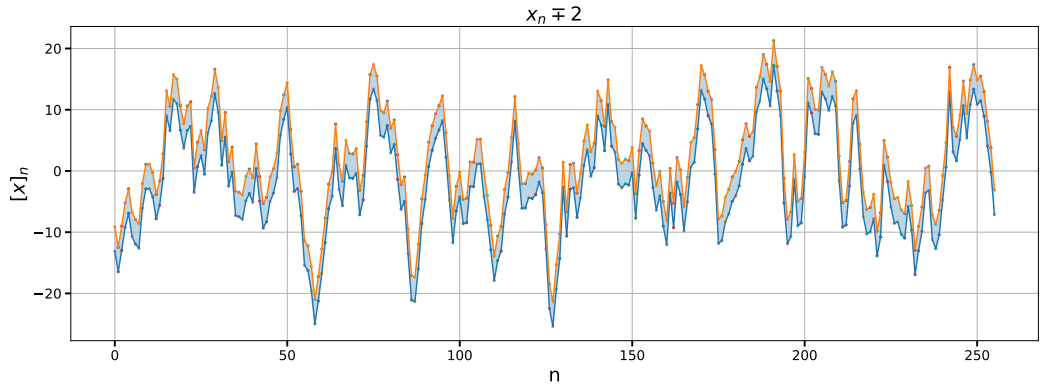


Figure 2. An interval signal $\overline{\underline{x}}_0, \overline{\underline{x}}_1, ..., \overline{\underline{x}}_n$, with N=128 points, and interval uncertainty ∓ 2.

The interval signal of Figure 2 will be considered throughout the manuscript. A convenient way to evaluate the interval extension $\Diamond\mathcal{F}$ is to split precise and uncertain components, as shown in Eq. (3). For example, when the signal is acquired using the same sensor (same precision) the interval uncertainty has the same value throughout. In such case the $\Diamond\mathcal{F}$ becomes

$$\overline{\underline{z}}_k := \sum_{n=0}^{N-1} x_n \; e^{-\frac{i2\pi}{N}kn} + \sum_{n=0}^{N-1} \xi \; \Delta \; e^{-\frac{i2\pi}{N}kn}, \tag{3}$$

where $\Delta = [-1,\ 1]$ is the unit interval, and $\xi \in \mathbb{R}$ is the precision of the interval signal, expressed in the same units as the signal. This form can be used to assess the effect of the cumulation of uncertainty on the amplitude as more and more uncertainty components get added.

3. Interval arithmetic

The interval extensions of Eqs.(1), (2), and (3) can be evaluated with interval arithmetic. The generalized rules of arithmetic for *addition*, *multiplication* and *power elevation* are needed for the purpose. The addition + between any two intervals $\overline{\underline{x}},\ \overline{\underline{y}}$ with $x,\ y \in \mathbb{R}$ is

$$\overline{\underline{x}} + \overline{\underline{y}} := [\underline{x} + \underline{y},\ \overline{x} + \overline{y}]. \tag{4}$$

The multiplication $*$ between an interval and a number $a \in \mathbb{R}$, which symbol is often omitted is

$$a * \overline{\underline{x}} := a\overline{\underline{x}} := \begin{cases} [a\underline{x},\ a\overline{x}],\ a > 0; \\ [a\overline{x},\ a\underline{x}],\ a < 0. \end{cases} \tag{5}$$

The power elevation with exponent two is

$$\overline{\underline{x}}^2 := \begin{cases} [\underline{x}^2,\ \overline{x}^2],\ \overline{\underline{x}} \geq 0 \\ [\overline{x}^2,\ \underline{x}^2],\ \overline{\underline{x}} < 0 \\ [0, \max\left(\underline{x}^2,\ \overline{x}^2\right),\ \overline{\underline{x}} \ni 0. \end{cases} \tag{6}$$

The square root of a positive real interval is

$$\sqrt{\underline{x}} := [\sqrt{\underline{x}},\ \sqrt{\overline{x}}],\ \underline{x} > 0. \tag{7}$$

With these rules the amplitude of the DFT interval extension can be evaluated using Eq. (2) without the need of a particular algorithm. The results provided by the interval arithmetic are guaranteed to enclose the exact bounds. However artefactual uncertainty due to the dependence problem makes the bounds obtained with interval arithmetic wider than they ought to be.

The interval extension of the DFT (Eq.(1)) can be computed using the complex addition operator. Consider two complex intervals $\underline{\overline{z}}_k = [\underline{z}_k, \overline{z}_k],\ k = 1, 2$, their addition is given by

$$\underline{\overline{z}}_1 + \underline{\overline{z}}_2 := [\underline{z}_1 + \underline{z}_2, \overline{z}_1 + \overline{z}_2].$$

With the complex addition operator the program used to compute $\mathcal{F}$ can be recycled to obtain its interval extension. The input *signal* to the program will be an ordered iterable of intervals.

```
1 def DFT(signal):  # Inputs an interval signal
2     F=[]
3     N = len(signal)  # length of the real signal
4     for k in range(N//2):  # for each frequency, with k=0,1,2,...
5         f = 0
6         for n in range(N):
7             f += signal[n]*exp(-1j*2*pi*k*n/N)
8         F.append(f)
9     return F
```

Listing 1: Code sample for a basic Python implementation of the Fourier transform.

Code sample List.1 can be used independently from external interval libraries, replacing an interval with an ordered list of a complex pair. The function in List.1 do need *numpy* for π and exp. The amplitude of $\mathcal{F}(\underline{\overline{\mathbf{x}}})$ for each frequency $\omega_k,\ k = 0, 1, 2, ...$ can be obtained replacing f at line 8 with abs(f), or with the following function.

```
1 def DFT_amplitude(signal):
2     F = DFT(signal)
3     return [abs(f) for f in F]
```

Listing 2: Code sample for the DFT amplitude.

The only caveat in List.1 program, is that the length of the signal must be a power of two in order for the signal to be fully transformed to the Fourier domain. Slightly modified programs can be obtained with little extra effort to map signals that are not divisible by two, see e.g. (Newland(1993)).

3.1. Notes on the dependence problem using interval arithmetic

The bounds on the amplitude of the interval extension $\mathcal{F}$ evaluated with interval arithmetic are *rigorous* in the sense that all the possible signals $\mathbf{x} \in \underline{\overline{\mathbf{x}}}$ are mapped inside these bounds. However when the amplitude is computed, the bounds obtained with Eq. (2) will be suboptimal because of

the dependence problem. The computation of Eq. (2) with interval arithmetic does not track the dependence between the real and the imaginary component, thus the bounds are outer-estimated. The bounds are rigorous because all of the possible dependencies between real and imaginary components are considered by the interval arithmetic bounds.

4. The brute-force method

The *brute-force* method is too inefficient to be used in practical applications, and it is herein used mainly for illustration purposes. The method tracks the dependency between real and imaginary components by constructing a binary tree of all the propagated endpoints at each iteration step. Clearly the number of endpoints increases exponentially in base two and with exponent given by the iteration number. The total number of iterations is the length of the signal, so the longer the signal the more inefficient is the *brute-force* method. Thus this method can only be used for research purposes either on a very short signal or on a signal with very few intervals. On a personal computer the recommended size of the problem is eight, more iterations may slow down the process significantly.

```
def BruteForce(intervalsignal, k, Limit):
    N = len(intervalsignal)  # <- length of the signal
    COMB=[]
    pair = [complex(exp(-2*pi*1j*k*0/N)) * iend for iend in intervalsignal[0]]
    pairs = pair  # initialise set of endpoints
    COMB.append(pairs)
    for n in range(1,Limit): # Limit < N when N is large
        pair  =  [complex(exp(-2*pi*1j*k*n/N)) * iend for iend in intervalsignal[n
    ]]
        pairs = [[pair[0] + ps for ps in pairs],[pair[1] + ps for ps in pairs]]
        COMB.append(pairs) # add to list of endpoints
        pairs = pairs[0] + pairs[1] # update for next iteration
    return COMB
```

Listing 3: Code sample for Python implementation of the *brute-force* method

Because the *brute-force* method is so inefficient, we do not provide the code for computing the bounds out of the set of endpoints output by the program. The code List.3 can be used without external interval libraries; however it does require *numpy* for π and exp. The program 3 outputs the sets of points depicted in Figure 3.

Figure 3 shows that the bounds provided by interval arithmetic in the complex plane are rigorous and best possible. The Figure superimposes the propagated endpoints through the Fourier transform on the interval arithmetic box, for a given *frequency* ($k = 9$) of the Fourier domain, and up to eight iterations. This is because there are no repeated variables in the complex expression of Eq. (1). Because the DFT is a linear function, mapping an *interval* signal to its Fourier counterpart requires the information from the endpoints only. The number of interval components present in the signal does not seem to affect the inflation of the bounds that keep enclosing the endpoints tightly. The final interval box, for a given frequency, seems to always tightly enclose the set of propagated

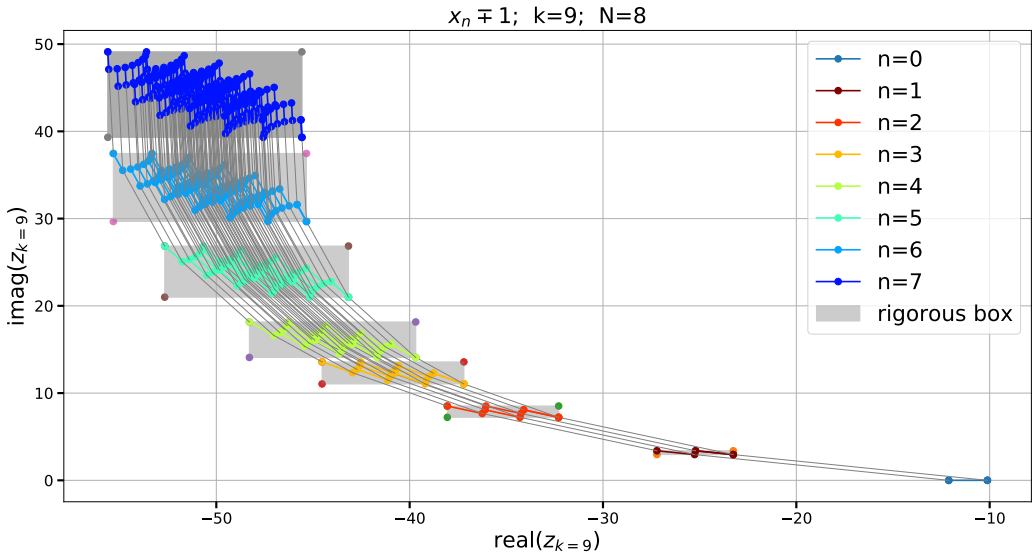


Figure 3. Set of endpoints produced by the *brute-force* method v. rigorous box.

endpoints. For example if the signal had eight components, given $k = 9$, we would be looking at the last gray box ($n = 7$).

5. The selective method

We observe that only the endpoints on the convex hull on the current set of endpoints carry the information about the bounds. Numerical experiments performed using the *brute-force* method have confirmed this intuition, however the investigation is limited by the number of iterations allowed by the brute force. Thankfully, interval arithmetic comes to the rescue. In fact, we can keep propagating the endpoints on the convex hull as well as the interval arithmetic box for as many iterations as needed and for however large signals. Moreover, because interval arithmetic is the most efficient method its box can always be computed alongside the *selective* method.

The numerical experiments done running the *selective* method alongside interval arithmetic have confirmed and *verified* the intuition that the selective method yields rigorous best possible bounds in each tested case. Figure 4, shows on four different frequencies, the three methods altogether. The convex hull is the set of points resting on the set perimeter, which in Figure 4 is depicted by a red closed line. The convex hull is obtained at each iteration by propagating only the endpoints on the perimeter of the previous iteration step.

Interval arithmetic gives us the verified guarantee that the *selective* method is rigorous, i.e. does not underestimate the uncertainty. The fact that the *selective* method makes use of a subset of the endpoints ensures that its bounds are also best possible. In fact, it is always possible to index the

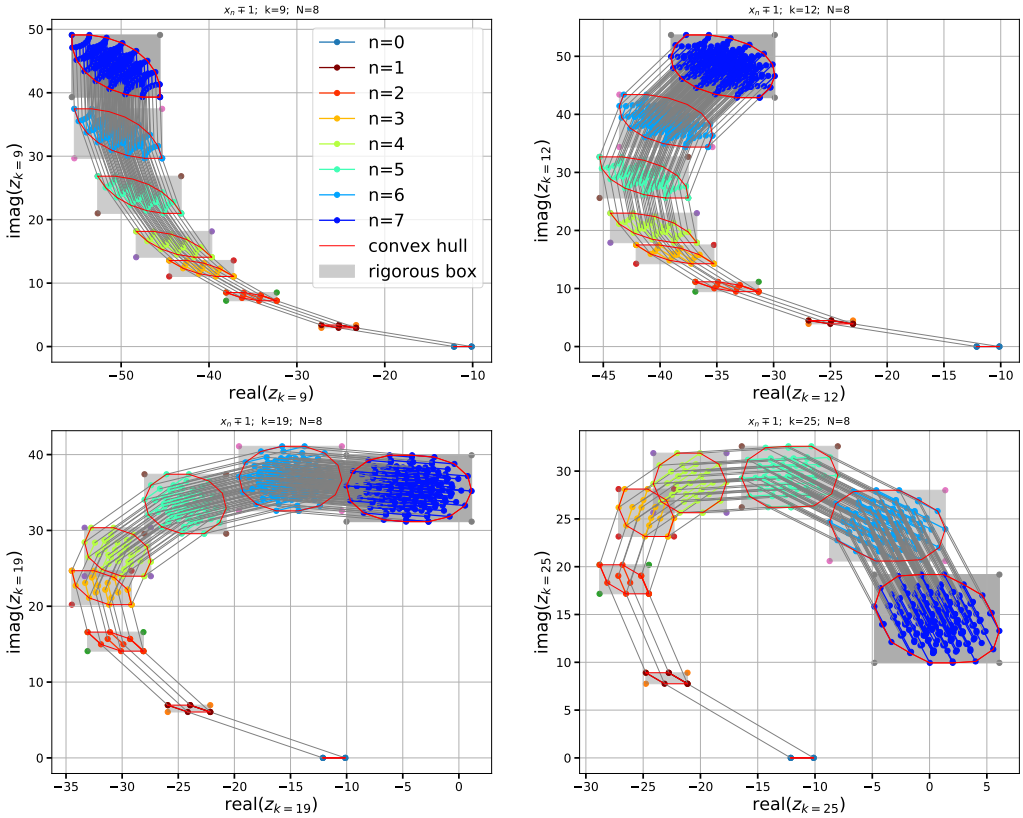


Figure 4. Set of points on the convex hull and interval box for four different frequencies.

convex-hull endpoints resting on the edge of the rigorous box and propagate these points back to identify the combination of endpoints yielding the bounds. However because there is no guarantee that theses endpoints will be unique, the algorithm cannot be used directly to perform the inverse transform.

The code produces the set of points in the convex hull as shown in List.4 and can be obtained as a derivation from the *brute-force* method. The function *Selective* in the code snippet List.4 will be invoked for each frequency k of the Fourier sequence.

```
from scipy.spatial import ConvexHull
def Selective(intervalsignal, k):
    N = len(intervalsignal)
    CHULL = []
    pair = [complex(exp(-2*pi*1j*k*0/N)) * iend for iend in intervalsignal[0]]
```

```
    chullpair = pair
    CHULL.append(chullpair)
    for n in range(1,N):
        pair =  [complex(exp(-2*pi*1j*k*0/N)) * iend for iend in intervalsignal[n
    ]]
        chullpair=[[pair[0]+ch for ch in chullpair],[pair[1]+ch for ch in
    chullpair]]
        chullpair = chullpair[0] + chullpair[1]
        chullpairs_RI = [[p.real, p.imag] for p in chullpair]
        hull = ConvexHull(chullpairs_RI)
        chullpair = [chullpairs_RI[h][0] + 1j*chullpairs_RI[h][1] for h in hull.
    vertices]
        CHULL.append(chullpair)
    return CHULL
```

Listing 4: Code snippet for the Python implementation of the *selective* method.

6. Rigorous bounding method

There are no repeated variables in the interval extension of Eq. (1), thus the final enclosing box is the smallest box possible containing the set of points generated by the *brute force* method. In List.5 we show the procedure to obtain the enclosing interval box at each iteration of the Fourier transform. Note that if the enclosing box was not the smallest or *best possible*, there would have been a gap between the box and the convex hull of points, effectively making it impossible for us to verify the results produced by the selective method. Luckily in our numerical experiments there has been no instance of such a case. The rigorous bounding method can be performed alongside the selective method with very little additional computations, to provide a prompt verification certificate of the bounds outputted by the selective method.

```
def RigorousBox(intervalsignal, k):
    N = len(intervalsignal)
    RBOX = []
    firstinterval = [complex(exp(-2*pi*1j*k*0/N)) * iend for iend in
    intervalsignal[0]]
    RBOX.append(firstinterval)
    ci =firstinterval
    for n in range(1,N):
        ci += complex(exp(-2*pi*1j*k*n/N)) * intervalsignal[k]
        RBOX.append(ci)
    return RBOX
```

Listing 5: Code snippet for the Python implementation of the *interval* method.

The amplitude of the Fourier sequence can be obtained just using the information carried by the corners of the enclosing box. Because the set of extreme endpoints rest on the convex hull, the amplitude obtained with the enclosing box will carry inflated uncertainty. The bounds on the amplitude will be reflective of the fact that the zone outside the convex hull between the box will be propagated too, effectively making the bounds on the amplitude much puffier. The extra puffiness

produced by this method however is compensated by the efficiency of this method, which practically carries no additional computational cost when compared to the *precise* DFT.

7. The final algorithm for the bounds on the amplitude

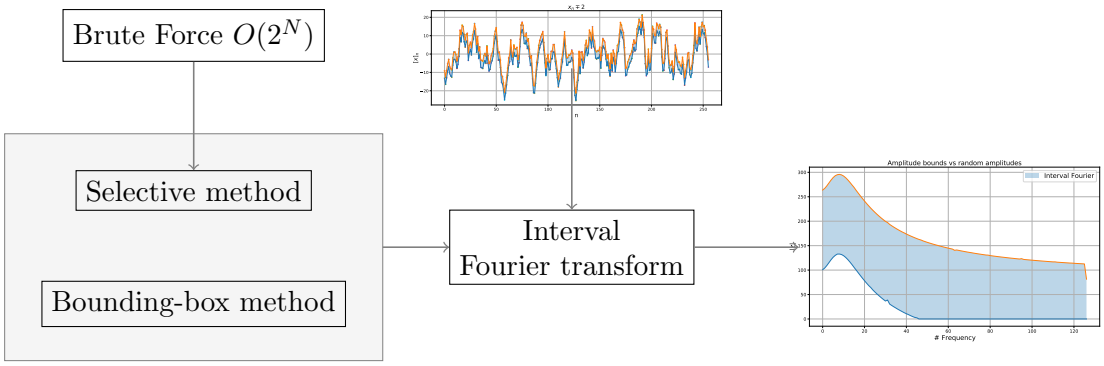


Figure 5. Interval Fourier transform algorithm in summary.

In this section we show how to obtain the bounds on the amplitude for each frequency. A schematic representation of how the three methods interact with each other is provided in Figure5.

The convex hull or the interval box at the last iteration for $n = N - 1$, is used to compute the amplitude at each frequency. We have shown that the endpoints on the convex hull yield the best possible bounds, thus the bounds on the amplitude are obtained by taking the *minimum* and *maximum* amplitude corresponding to this set of endpoints. The points corresponding to the *extrema* are shown in Figure6. Because the amplitude of the Fourier transform is effectively the Euclidean norm on the $\mathbb{R}^2$ isomorphism of the complex plane, *minimum* and *maximum* correspond to the points that are nearest and furthest from the origin of the complex plane. However, this is true as long as the convex hull does not enclose the origin. When the convex hull encloses the origin the *minimum* will be attained on the nearest endpoint to the origin. Because the *selective* approach "forgets" about all the endpoints in the interior of the convex hull, it is not possible to determine exactly the nearest to the origin endpoint.

The algorithm for the amplitude bounds will therefore assume that the *minimum* is attained at the origin of the complex plane in such situation, as shown in Figure7.

A simple implementation of the algorithm for the amplitude bounds with the selective method is shown in List.6. The function *AmplitudeBounds* needs an external function called *origin_in_complex_hull* that returns *True* when the convex hull encloses the origin and a *False* otherwise.

```
def AmplitudeBounds(intervalsignal, k):
    CHULL = Selective(intervalsignal, k)
    convhull = CHULL[-1]
    chull_max = numpy.argmax([abs(h) for h in convhull])
    chull_min = numpy.argmin([abs(h) for h in convhull])
    if origin_in_complex_hull(convhull):
        return 0, abs(convhull[chull_max])
```

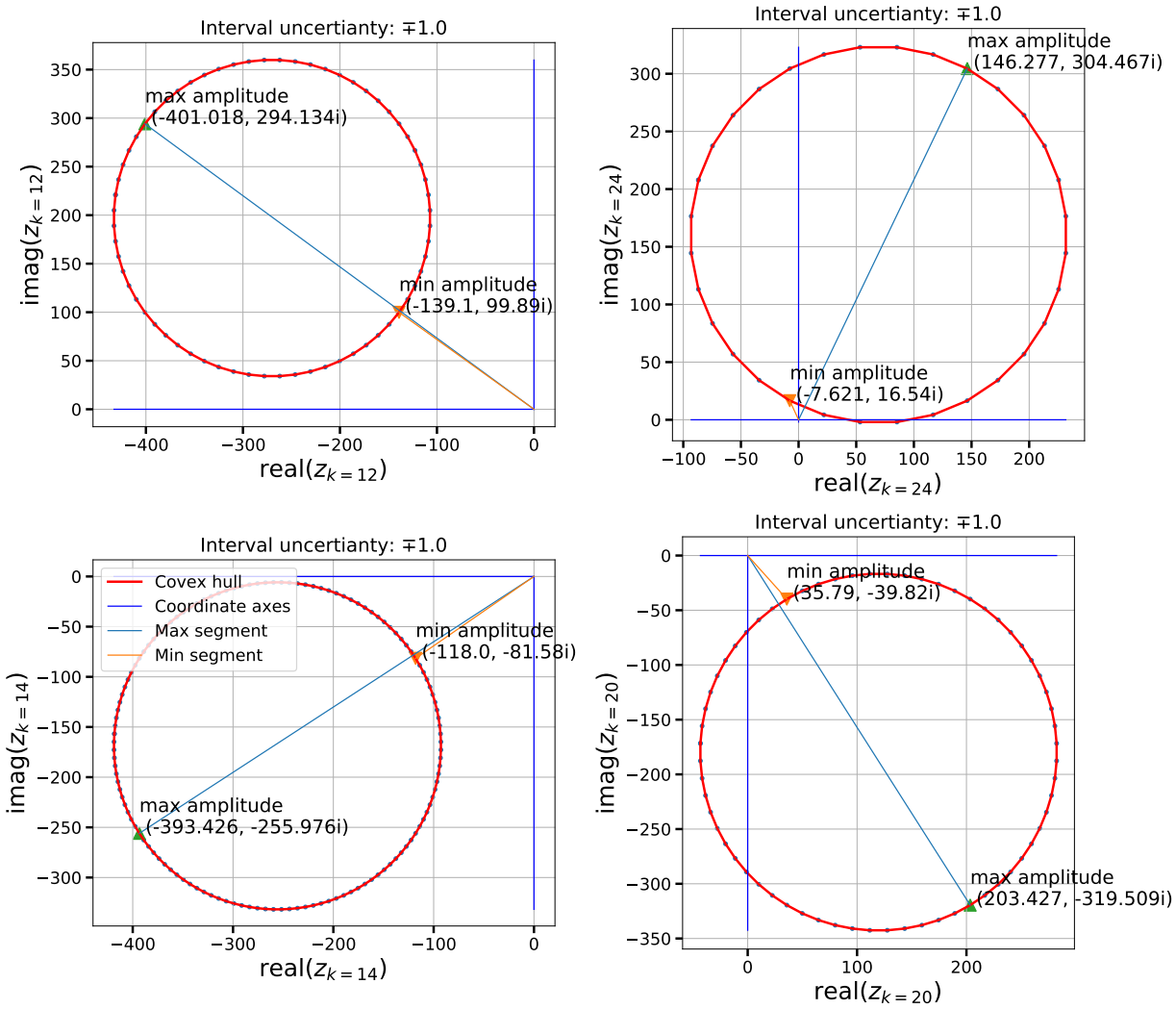


Figure 6. Convex hull of endpoints for an interval signal N=128 long, and four different frequencies.

```
    else:
        return abs(convhull[chull_min]), abs(convhull[chull_max])
```

Listing 6: Python function for the amplitude bounds with the *selective* method.

The procedure for the amplitude bounds with the interval box is equivalent except that is much easier to check if the origin is contained in the box. Iterating over all the frequencies of the Fourier sequence as shown in List.7 leads to the amplitude bounds shown in Figure8.

```
def DFT_bounds(intervalsignal):
    N = len(intervalsignal)
    BOUNDS_I, BOUNDS_C = [],[]
```

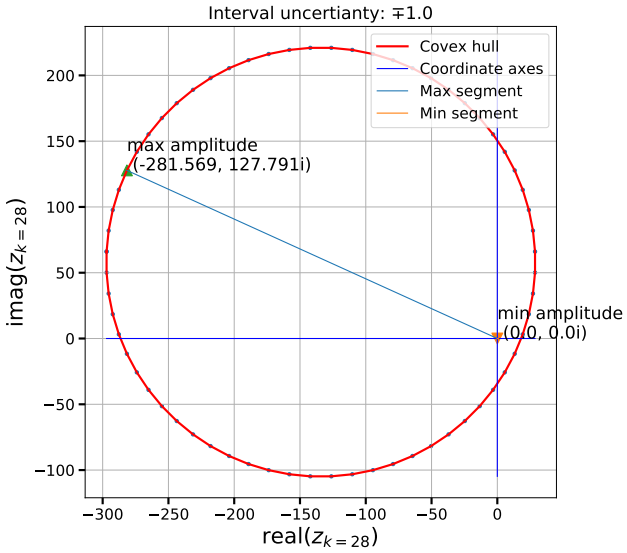


Figure 7. Convex hull of endpoints enclosing the origin of the complex plain.

```
    for k in range(1,N//2):
        bI,bC = AmplitudeBounds(intervalsignal, k)
        BOUNDS_I.append(bI)
        BOUNDS_C.append(bC)
    return BOUNDS_I, BOUNDS_C
```

Listing 7: Python function for the amplitude bounds for every frequency.

8. Conclusions

In this paper we have presented three methods for computing the bounds of the Fourier amplitude when an interval signal is provided. We have shown how the problem can be tackled with a brute-force method. However the brute-force algorithm is NP-complete, thus only working for very small signals (< 20 in length), with no applications in practice. Despite its inefficiency the brute-force method has provided the inspiration for the second more efficient *selective* method, which yields the best possible bounds on the amplitude of the Fourier sequence. In this paper we have questioned the rigour or reliability of the *selective* algorithm, which has led us to our third method, which uses complex interval arithmetic to propagate the bounds through the Fourier transform. The third method called *bounding box*, has provided a viable efficient tool to verify the rigorousness of the *selective* method even for large interval signals. The *bounding box* method has also provided a robust alternative to obtain the amplitude bounds when efficiency is a priority at the expense of losing specificity in the bounds of the amplitude.

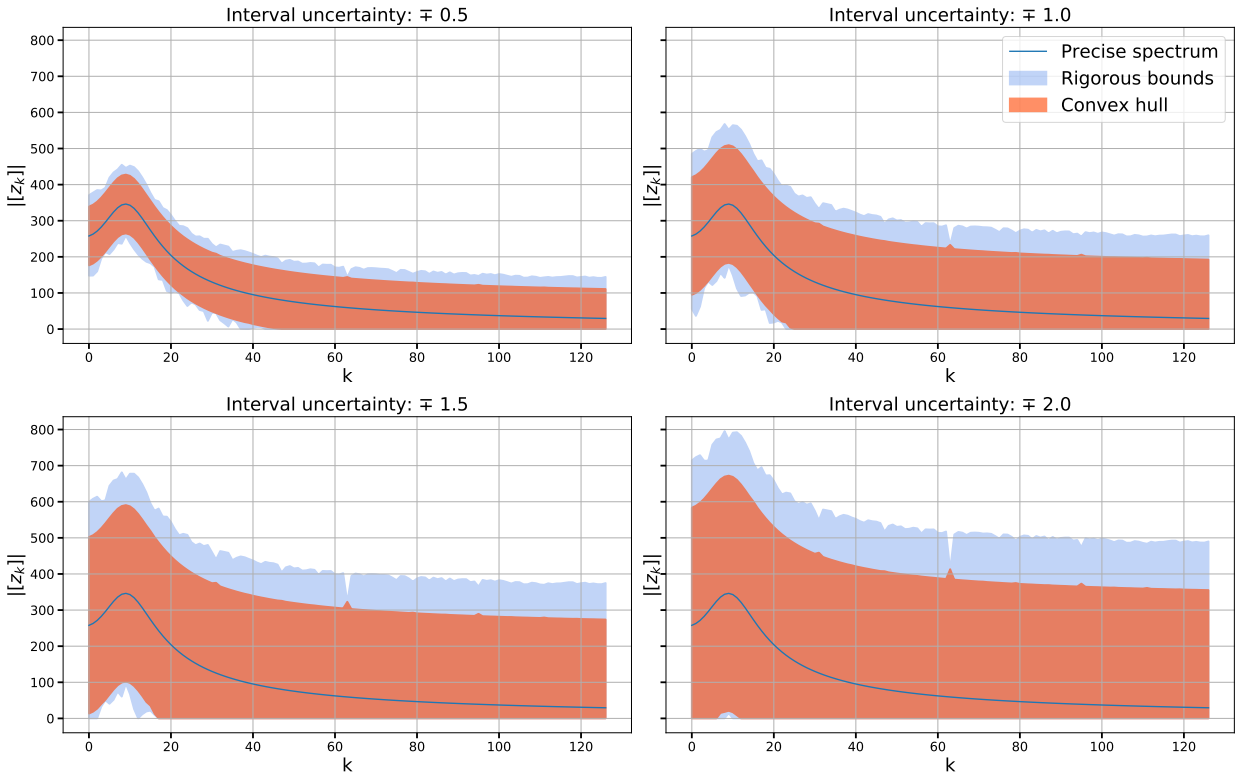


Figure 8. Amplitude bounds for the Fourier transform for different uncertainty levels.

9. Acknowledgements

This research is funded by the Engineering & Physical Sciences Research Council (EPSRC) with grant no. EP/R006768/1. The research is also funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) with grant no. BE 2570/4-1 and CO 1849/1-1. EPSRC and DFG are greatly acknowledged for their funding and support.

References

D. Newland, *An Introduction to Random Vibrations, Spectral and Wavelet Analysis*, Dover books on engineering (Longman Scientific & Technical, 1993), ISBN 9780582215849.

G. Muscolino, R. Santoro, and A. Sofi, in *5th International conference on Reliable Engineering Computing (REC2012). Edited by Brno, Czech Republic: Ing. Vladislav Pokorný-LITERA* (2012), 407–25.

L. Comerford, I. A. Kougioumtzoglou, and M. Beer, On quantifying the uncertainty of stochastic process power spectrum estimates subject to missing data, *International Journal of Sustainable Materials and Structural Systems* **2**, 185 (2015).

I. N. Sneddon, *Fourier transforms* (Courier Corporation, 1995).

J. F. James, *A Student's Guide to Fourier Transforms: With Applications in Physics and Engineering*, Student's Guides (Cambridge University Press, 2011), 3rd ed.

J. W. Cooley and J. W. Tukey, An algorithm for the machine calculation of complex fourier series, *Mathematics of computation* **19**, 297 (1965).

T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to algorithms* (MIT press, 2009).

J. Stoer and R. Bulirsch, *Introduction to numerical analysis*, vol. 12 (Springer Science & Business Media, 2013).

G. Alefeld and J. Herzberger, *Introduction to Interval Computation*, Computer Science and Applied Mathematics (Elsevier Science, 2012), ISBN 9780080916361.

R. E. Moore, *Interval analysis*, vol. 4 (Prentice-Hall Englewood Cliffs, 1966).

R. E. Moore, *Methods and applications of interval analysis* (SIAM, 1979).

R. Moore, R. Kearfott, and M. Cloud, *Introduction to Interval Analysis*, Other titles in applied mathematics (Cambridge University Press, 2009), ISBN 9780898716696.

Interval Analysis Using Multilevel Quasi-Monte Carlo

Robin R.P. Callens, Matthias G.R. Faes, David Moens
KU Leuven, Department of Mechanical Engineering, Division LMSD
St.-Katelijne-Waver 2860, Belgium, robin.callens@kuleuven.be

Abstract. Interval analysis has proven to provide robust bounds on the performance of structures when there is only limited data available on the uncertainty. Calculating the interval bounds on the output side of a computational model typically involves a global optimisation algorithm or vertex analysis for monotonic models, where numerous model evaluations are required. The computational cost, measured as computation time and number of deterministic model evaluations corresponding to this optimisation, increases with the number of interval dimensions and even further for highly detailed models. Recent developments have shown very promising results in the context of reducing the computational cost for probabilistic fields. Multilevel techniques as Multilevel Monte Carlo and Multilevel Quasi-Monte Carlo have been shown to reduce the computational cost significantly for numerical models. This paper presents Interval Multilevel Quasi-Monte Carlo to decrease the computational cost of interval analysis for linear models. In this approach, intervals are represented as Cauchy random variables, which enables the use of probabilistic sampling techniques for interval analysis. A computationally efficient technique for intervals analysis and linear models is then achieved by using the Multilevel Quasi-Monte Carlo framework. The efficiency of this Interval Multilevel Quasi-Monte Carlo technique is illustrated with an academic case study on a linear finite element model to compare the developed technique with a vertex analysis in terms of accuracy and computational cost.

Keywords: Interval analysis, Multilevel Quasi-Monte Carlo, Cauchy distributions

1. Introduction

Interval analysis, is popular when limited or incomplete data are available of the true model parameter values, e.g., in civil engineering (Zou, 2016), in structural engineering (Wang, 2014), in fatigue analysis (Long, 2018). Intervals quantify the uncertainty on the actual model parameters value by an upper and a lower bound, this in contrast to probabilistic techniques, which require distributions of the uncertain parameters (Beer, 2013). Also imprecise probability models have been applied in this context, such as probability-boxes (p-box) (Dannert, 2018), which can be regarded as the combination of non-probabilistic and probabilistic models (Song, 2019). The goal of interval analysis is to quantify the best- and worst-case behaviour of a model, whereas probabilistic methods quantify the moments or distribution of a model. The input of an interval analysis consists by definition of independent intervals, and hence, the joint description of several interval-valued parameters is given by a hyper-rectangle (Faes, 2019). Interval analysis is typically performed with (anti-)optimisation to find the bounds of the output interval. With this approach, a global optimisation algorithm

© 2021 *by R.R.P. Callens, M.G.R. Faes, and D. Moens.*

has to be performed for each bound (Hansen, 1980). This requires numerous model evaluations, depending on the input dimensions. Alternatively, for models that are monotonic inside the interval bounds, the output bounds are also found by vertex analysis. This technique quantifies the output interval bounds by propagating the vertex points of the input interval space trough the model. However, this technique requires numerous model evaluations for high input dimensions, as the required model evaluations scales with $2^{d_{\mathcal{X}}}$ where $d_{\mathcal{X}}$ is the number of input interval dimensions.

For low input interval dimensions, both techniques (vertex analysis and optimisation) can be efficient and accurate to quantify the output interval. However, for medium and high input interval dimensions, the high amount of model evaluations becomes computationally unattractive, especially, when these models are high dimensional. For probabilistic variables, Monte Carlo is typically used for high dimensional input spaces. However, for high dimensional models this also becomes computationally unattractive. In the context of reducing the total computational cost of these model evaluations, recently Multilevel Monte Carlo techniques have shown great potential for probabilistic uncertainties (Robbe, 2019). Multilevel techniques combine models of different types: fast to compute and low in accuracy models and slow to compute and high in accuracy models. These models typically differ in level of accuracy and computational cost (Heinrich, 2001; Charrier, 2013; Giles, 2015). Also, for fuzzy set theory (Mäck, 2019), the use of multi-fidelity methods has shown to reduce the computational cost, while ensuring the accuracy of the high fidelity model. However, these techniques require a probabilistic representation of the uncertainty and thus are not directly applicable for interval analysis.

This paper, proposes the Interval Multilevel Quasi-Monte Carlo technique for linear computational models. This is achieved by extending the approach of representing intervals with Cauchy distributions (Kreinovich, 2004). These distributions have the property that a linear combination of independent Cauchy distribution is again a Cauchy distribution. Furthermore, the parameters of the resulting distribution are found as a linear combination of the input Cauchy distribution parameters, weighted by the linear model. For interval analysis of a linear model, the output interval is also the linear combination of the independent input interval scalars that are weighted by the model. Thus, when representing the intervals as Cauchy distributions, the output interval of an interval analysis is also obtained with sampling technique for Cauchy distributions. The aim of this paper is to combine the approach with the computationally efficient Multilevel Quasi-Monte Carlo for interval analysis.

This paper is organised as follows. Section 2 introduces interval analysis and solution techniques based on optimisation and vertex analysis. The sampling approach based on the Cauchy distribution is discussed in section 3. Next, section 4 proposes the new Interval Multilevel Quasi-Monte Carlo technique to efficiently perform an interval analysis. Finally, section 5 applies the proposed Interval Multilevel Quasi-Monte Carlo technique to a Jet Engine Turbine Blade and compares the results with the Interval Monte Carlo technique and vertex analysis with respect to computational time and accuracy.

2. Interval analysis

Consider an interval scalar $x^I \in \mathbb{IR}$, where $\mathbb{IR}$ is the domain of closed real-valued intervals, which is defined as:

$$x^I = [x_{min}\ x_{max}] = [\underline{x}\ \overline{x}] = \{x \in \mathbb{R} | \underline{x} \leq x \leq \overline{x}\}, \tag{1}$$

where x_{min} and x_{max}, with $x_{min} < x_{max}$ are bounds between which the true values of the uncertain parameter x are deemed to lie. Similarly, the interval midpoint x_m and width x_w are given by

$$x_m = \frac{\underline{x} + \overline{x}}{2}, \tag{2}$$

$$x_w = \frac{\overline{x} - \underline{x}}{2}. \tag{3}$$

An interval vector $\boldsymbol{x}^I \in \mathbb{IR}^{d_{\mathcal{X}}}$ contains $d_{\mathcal{X}}$, by definition, independent interval scalars $x_i^I,\ i = 1, \ldots, d_{\mathcal{X}}$, and as such can be seen as a Cartesian product of the interval scalars: $\boldsymbol{x}^I = x_1^I \times x_2^I \times \cdots \times x_{d_{\mathcal{X}}}^I$. Alternatively, $\boldsymbol{x}^I$ can be represented in set notation as:

$$\boldsymbol{x}^I = \left\{x_1^I\ \ x_2^I\ \ \ldots\ \ x_{d_{\mathcal{X}}}^I\right\}^T = \left\{\boldsymbol{x} \in \mathbb{R}^{d_{\mathcal{X}}} | x_i \in x_i^I\right\}, \tag{4}$$

In the context of interval analysis, the interval vector $\boldsymbol{x}^I$ defines the hyper-rectangular input-space of a function $\mathcal{M} : \mathbb{R}^{d_{\mathcal{X}}} \mapsto \mathbb{R}^{d_d}, \boldsymbol{x} \to \boldsymbol{y}$. In engineering practice $\mathcal{M}$ is represented as a numerical model that consists of d_d deterministic functions $m_1 : \mathbb{R}^{d_{\mathcal{X}}} \mapsto \mathbb{R}, \boldsymbol{x} \to y_i$, with $i = 1, \ldots, d_d$. Mapping $\boldsymbol{x}^I$ via $\mathcal{M}$ yields a solution set $\boldsymbol{y}^S \in \mathbb{R}^{d_d}$ representing the joint bounds on the model responses $\boldsymbol{y}$ of interest. This set is explicitly given as:

$$\boldsymbol{y}^S = \left\{\boldsymbol{y} | \boldsymbol{y} = \mathcal{M}(\boldsymbol{x}), \boldsymbol{x} \in \boldsymbol{x}^I\right\}. \tag{5}$$

Since finding the exact set $\boldsymbol{y}^S$ in the general case constitutes an NP-hard problem, $\boldsymbol{y}^S$ is usually approximated by an interval vector $\boldsymbol{y}^I \in \mathbb{IR}^{d_d}$: $\boldsymbol{y}^I = \left\{y_1^I\ \ y_2^I\ \ \ldots\ \ y_{d_d}^I\right\}^T$. The components $y_i^I = \left[\underline{y}_i\ \overline{y}_i\right]$ of y^I are determined by means of (anti-)optimisation, i.e.,

$$\underline{y}_i = \min_{\boldsymbol{x} \in \boldsymbol{x}^I} \left(m_i(\boldsymbol{x})\right), \tag{6}$$

$$\overline{y}_i = \max_{\boldsymbol{x} \in \boldsymbol{x}^I} \left(m_i(\boldsymbol{x})\right). \tag{7}$$

With this optimisation, the interval of each component y_i^I of $\boldsymbol{y}^I$ is found independently, resulting in an approximation of the solution set $\boldsymbol{y}^S$ as a hyper-rectangle. Note that this approach solves a total of $2d_d$ optimisation problems, each potentially requiring numerous solutions of the model $\mathcal{M}$.

Alternatively for a model $\mathcal{M}$ that is monotonic in $\boldsymbol{x}^I$, the interval of each component y_i^I of $\boldsymbol{y}^I$ is also found by vertex analysis. In this approach, each component is found independently by minimising/maximising the model responses of the set of vertex points $\mathcal{V}$ of $\boldsymbol{x}^I$. This set contains all possible input parameter combinations located at vertex points of the interval input vector $\boldsymbol{x}^I$, yielding $2^{d_{\mathcal{X}}}$ combinations.

3. Interval (Quasi-)Monte Carlo

This section introduces Cauchy sampling tools to efficiently propagate intervals through expensive black-box functions $\mathcal{M}$.

3.1. Cauchy distribution to represent interval uncertainty

The probability density function $f_X(\delta, \gamma)$ of a Cauchy variable $X \sim \mathcal{C}(\gamma, \delta)$ is given as:

$$f_X(\delta, \gamma) = \frac{\gamma}{\gamma\pi(\gamma^2 + (x - \delta)^2)}, \tag{8}$$

and the cumulative density function $F_X(\delta, \gamma)$ is defined as:

$$F_X(\delta, \gamma) = \frac{1}{\pi} \arctan\left(\frac{x - \delta}{\gamma}\right) + \frac{1}{2}. \tag{9}$$

The Cauchy distribution $\mathcal{C}(\delta, \gamma)$ is one of the distributions in the set of α-stable distributions $\mathcal{S}(\alpha, \beta, \delta, \gamma)$, which are defined by four parameters: $\alpha \in (0, 2]$ controls the stability or weight of the tails, $\beta \in [-1, 1]$ represents the skewness of the distribution, $\delta \in \mathbb{R}$ defines the location, and $\gamma \in \mathbb{R}$ defines the half width at half height and therefore acts as scale parameter. Sometimes they are also called *heavy tailed distributions* in case $\alpha < 2$. As a result of these heavy tails, the variance is infinite for $\alpha < 2$ and the mean is undefined for $\alpha \leq 1$. For $\alpha = 1$ and $\beta = 0$ the α-stable distribution becomes the Cauchy distribution and for $\alpha = 2$ and $\beta = 0$ it becomes the normal distribution. In the following paragraphs, the skewness parameter is set as $\beta = 0$, yielding a symmetric distribution. For the more general case of non-symmetric α-stable distributions, i.e., when $\beta \neq 0$, the reader is referred to (Borak, 2005) for more information and properties. In essence, α-stable distributions are distributions that abide by the *stability property*, as defined below.

Stability property: If the variables $X_1 \sim \mathcal{S}(\alpha_1, \beta_1 = 0, \gamma_1, \delta_1)$ and $X_2 \sim \mathcal{S}(\alpha_2, \beta_2 = 0, \gamma_2, \delta_2)$ are independent α-stable distributions with $\alpha = \alpha_1 = \alpha_2$, then $X_1 + X_2 \sim \mathcal{S}(\alpha, \beta = 0, \gamma, \delta)$ with:

$$\gamma = (\gamma_1^\alpha + \gamma_2^\alpha)^{\frac{1}{\alpha}}, \tag{10}$$

$$\delta = \delta_1 + \delta_2. \tag{11}$$

For the proof of this property, the reader is referred to (Borak, 2005). From the stability property the so called *scaling property* can be derived with induction, as defined below.

Scaling property: The sum of n-independent stable random variables $X_j \sim \mathcal{S}(\alpha_j, \beta_j = 0, \gamma_j, \delta_j)$, $j = 1, 2, \ldots, n$, and $\alpha = \alpha_j$ with constants w_j is given as

$$w_1 X_1 + w_2 X_2 + \cdots + w_n X_n \sim \mathcal{S}(\alpha, \beta = 0, \gamma, \delta) \tag{12}$$

where

$$\gamma^\alpha = \sum_{j=1}^{n} |w_j \gamma_j|^\alpha, \tag{13}$$

$$\delta = \sum_{j=1}^{n} w_j \delta_j. \tag{14}$$

For the Cauchy distribution equation (13) becomes $\gamma = \sum_{j=1}^{n} |w_j \gamma_j|$, as $\alpha = 1$.

From the scaling property it follows that the output variable Y of a linear combination of n-independent Cauchy random input variables X_j is again a Cauchy distribution the parameters of which are weighted or scaled linearly (Eq. (13) and (14)). With interval analysis, the resulting output interval of a linear model is also obtained from a linear combination of independent input interval scalars that are weighted and scaled by the model. Based on these observations, Cauchy distributions can be used to propagate interval uncertainty through a linear model. Hereto, the input interval scalars x_j^I are represented by Cauchy random variables $X_j \sim \mathcal{C}(\gamma_j, \delta_j)$, where the location and scale parameter are related to the interval midpoint and width as:

$$\gamma_j = x_{j_w} = \frac{\overline{x_j} - \underline{x_j}}{2}, \tag{15}$$

$$\delta_j = x_{j_m} = \frac{\overline{x_j} + \underline{x_j}}{2}. \tag{16}$$

The output variable Y follows then again a Cauchy distribution $\mathcal{C}(\delta_Y, \gamma_Y)$, with parameters that are weighted (Eq. (13)) and scaled (Eq. (14)) linearly. As a result, these parameters are equal to the output interval midpoint y_m and width y_w (Kreinovich, 2004). The bounds of the interval on the response y of interest are then calculated as:

$$\underline{y} = y_m - y_w, \tag{17}$$

$$\overline{y} = y_m + y_w. \tag{18}$$

In practical application when the model, and hence the linear scaling is unknown, this allows to estimate the output interval with sampling techniques for Cauchy distributions such as Monte Carlo simulation. The next sections will go deeper in how to sample from Cauchy distributions at the input side of the model and estimate a Cauchy distributions on the model responses.

3.2. (Quasi-)Monte Carlo sampling for interval analysis

Sampling from the Cauchy random variable $X \sim \mathcal{C}(\gamma_X, \delta_X)$ is typically performed via the inverse cumulative density function, defined as:

$$F_X^{-1}(\delta_X, \gamma_X) = \delta_X + \gamma_X \tan\left[\pi\left(p - \frac{1}{2}\right)\right], \tag{19}$$

where p is the probability $p \in [0, 1]$. This function also enables to generate Cauchy distributed samples from a uniform distribution of $u \sim \mathcal{U}(0, 1)$. This is particularly relevant for these samples when using Quasi-Monte Carlo, which reduces the variance of the Monte Carlo estimator.

Sampling from the Cauchy distribution generates samples on an infinite range, which follows from the definition of a Cauchy distribution and its ∞ variance. As a result, samples outside the interval bounds are very likely to occur. In case these bounds represent physical bounds on parameter values (i.e., non-negative contact stiffness values) this might be problematic for the underlying numerical solver that propagates these values towards the responses of interest. Also, in case that the model is non-linear outside these bounds (e.g., large displacement issues in finite element calculations), the

stability and scaling property of Cauchy distributions are no longer valid. As they are fundamental to the method, violating these properties leads to incorrect results.

Sampling within the range of the input intervals is achieved by using the *scaling property* (see Eq. (13)) of a Cauchy distribution. Samples c_i with $i = 1, \ldots, N$ are taken from a standard Cauchy distribution $\mathcal{C}(0, 1)$. These samples are then normalised to fit within the interval bounds, this by normalising with the interval width proportional to the largest sample $C = \max(|c_i|)$, resulting in samples of $\mathcal{C}(0, \gamma_X/C)$. The samples are then adjusted to the location of the interval such that they are samples from $X \sim \mathcal{C}(\delta_X, \gamma_X/C)$, samples within the interval bounds. After propagating the samples trough the numerical model the samples are re-normalised from $\mathcal{C}(\delta_Y, w\gamma_X/C)$ to $Y \sim \mathcal{C}(\delta_Y, \gamma_Y = w\gamma_X)$ by multiplying the scale parameter with C where w is a weight of the linear numerical model (Kreinovich, 2004).

Standard Monte Carlo sampling achieves a slow convergence rate for estimating the statistical moments of Y; for instance, the mean value of Y converges at a rate of $1/\sqrt{N}$. Using Quasi-Monte Carlo (QMC) sampling, the samples p_i are chosen more cleverly in $[0, 1]$ to have samples with a low-discrepancy. For integrating a function f with QMC, as is required to determine the statistical moments of a random variable, the error on the moment estimator is bounded by the Koksma-Hlawka inequality:

$$\left| N^{-1} \sum_{i=1}^{N} f(p_i) - \int_0^1 f(p)dp \right| \leq D_N(p_1, \ldots, p_N) Var(f), \tag{20}$$

where $D_N(p_1, \ldots, p_N)$ is the discrepancy of the samples p_i and $Var(f)$ the total variance on the function f. In case the discrepancy $D_N(p_1, \ldots, p_N)$ of the sample sequence can be reduced, a lower upper bound on the error can be obtained, resulting in a faster convergence. QMC sample sequences are designed to minimise the discrepancy between the samples, enhancing the convergence rate over standard MC (Brandolini, 2013; Aistleitner, 2014). These low-discrepancy samples are typically generated from lattice rules (e.g., rank-one lattice rules) and digital nets (e.g., Sobol sets). For the remainder of the paper, and without loss of generality, Sobol sets will be used in the development due to their excellent performance. For more information on lattice rules, the reader is referred to (Hickernell, 1998) and for digital nets to (Dick, 2010).

3.3. Estimation of output interval

Estimating the parameters that describe the Cauchy distribution of the output variable y (i.e., δ_Y and γ_Y) is a challenging task. Sample moment estimators do not exist due to the heavy tails of the distribution, i.e., the sample mean is undefined and sample variance is infinite. Nonetheless, the parameters of a Cauchy distribution can be estimated with several approaches, such as the Maximum Likelihood Estimator (MLE) (as proposed in (Kreinovich, 2004)) and the median adjusted log Hodges-Lehmann Estimator (MHLE) (Kravchuk, 2012).

The MLE maximises a likelihood function to obtain the parameter of the Cauchy distribution. Hereto, the log-likelihood function is defined as:

$$l(s_1, s_2, \ldots, s_n | \delta, \gamma) = -N \log(\gamma\pi) - \sum_{i=1}^{n} \log\left(1 + \left(\frac{s_i - \delta}{\gamma}\right)^2\right), \tag{21}$$

where $s_1, s_2, \ldots, s_N$ are the QMC or MC samples. While the MLE is asymptotically efficient, it is computationally expensive as the likelihood function has $2N-1$ roots and a global optimisation algorithm to locate the global maximum is required. Also for small sample sizes the MLE is reported to be inefficient (Freue, 2007).

In this paper, it is proposed to use the MHLE, which is in comparison to the MLE easy to compute with one explicit formula. The classical Hodges-Lehmann Estimator (HLE) is defined as the half of the median of $N(N+1)/2$ pairwise sums of samples $s_i s_j$ with $1 \leq i, j \leq N, i \leq j$. The MHLE estimates the scale parameter γ of the Cauchy distribution as:

$$\log \mathrm{MHLE}(\gamma) = \frac{1}{2}\mathrm{med}(\ln|(s_i - \delta)(s_j - \delta)|), \tag{22}$$

where δ is the estimated location parameter of the Cauchy distribution and estimated by the sample median. The back-transformed estimator of the $\log \mathrm{MHLE}(\gamma)$ is:

$$\mathrm{MHLE}(\gamma) = \exp(\log \mathrm{MHLE}(\gamma)) \tag{23}$$

While the MHLE estimator is easy to compute and has a unique solution, it requires an estimated median. Therefore in this paper the scale parameter is estimated by first estimating the sample median. Then, the samples s_i with $i = 1, 2, \ldots, N$ (N is sample-size) are re-scaled with the scaling factor C, this is the same C used to generate the set of samples, from $S \sim \mathcal{C}(\delta_S, \gamma_S/C)$ to the output $Y \sim \mathcal{C}(\delta_Y, \gamma_Y)$. With these re-scaled samples the $\mathrm{MHLE}(\gamma_Y) = \exp(\frac{1}{2}\mathrm{med}(\ln|(y_i)(y_j)|)$ estimates the scale parameter. With both Cauchy parameters estimated (δ_Y and γ_Y), the interval bounds are computed with (17) and (18) (Kreinovich, 2004).

The MC and QMC sampling techniques require several deterministic model evaluation before an accurate estimate of the output interval $\boldsymbol{y}^I$ is obtained. For numerical models with a given discretisation level (e.g., element size, time step), MC and QMC perform all these model evaluation on one accurate level, this to achieve an output with sufficient accuracy. Such accurate model is typically high dimensional with a correspondingly high computational cost for each model evaluation, yielding a excessive processing time for both techniques. For instance, in the context of static analysis, this high accuracy level yields large, possibly dense matrices containing up to millions of rows and columns that need to be assembled and inverted to solve a single deterministic problem.

4. Interval Multilevel Quasi-Monte Carlo

4.1. Multilevel (Quasi-) Monte Carlo

Multilevel techniques combine responses y of low accuracy fast to compute models with high accuracy slow to compute models. The levels y^l for $l = 1, \ldots, L$ of a multilevel technique refer to this accuracy and computational cost. MC and QMC typically use the most precise models to achieve the precision of these models. The levels are in general depending on some parameters that determine the numerical accuracy of the computer model, e.g., element sizes and/or basis function orders in finite element models or time steps in dynamical simulations. For instance in a linear elastic finite element analysis, a model on level $l = 0$ is a model with a coarser element size then a

model level $l = L$ that has the finest element size and the largest computational cost. Considering finite element models with discretisation h^l (e.g., element size), which is typically calculated using a geometric sequence with base 2 as $h^l = h^0 2^{l-1}$, note that here the superscript is a power.

The expected value $\mathbb{E}(y^L)$ of the multilevel estimator is defined as:

$$\mathbb{E}[y^L] = \mathbb{E}[y^0] + \sum_{l=1}^{L} \mathbb{E}[y^l - y^{l-1}] = \sum_{l=0}^{L} \mathbb{E}[\Delta y^l], \tag{24}$$

where y^l and y^{l-1} are the responses on level l and level $l-1$ and $\mathbb{E}[y^{-1}] = 0$. Each term $\mathbb{E}[\Delta y^l]$ is then estimated with MC or QMC.

The mean square error on the MLMC estimator is:

$$MSE(y^L) = \sum_{l=0}^{L} Var[\Delta y^l] + (\mathbb{E}[y^L - y])^2 \tag{25}$$

where the first term is the total variance of the multilevel-estimator and the second term is the bias on the MLMC. For a given tolerance e it suffices to ensure both terms are less then $e^2/2$:

$$Var[y^L] \leq \frac{e^2}{2} \quad \text{(variance constraint)}, \tag{26}$$

$$|\mathbb{E}[y^L - L]| \leq \frac{e^2}{\sqrt{2}} \quad \text{(bias constraint)}. \tag{27}$$

In multilevel-techniques, the variance constraint will determine the number of samples on the different levels and the bias constraint will determine the finest level L that is required to ensure that the total error is lower then the given tolerance e. For MLMC there are equations and algorithms available to estimate the optimal number of samples on each level and optimal number of levels (Heinrich, 2001; Charrier, 2013; Giles, 2015; Robbe, 2019). However, for MLQMC no reliable equations for the optimal number of samples and optimal number of levels are available as deriving the convergence rate with respect to N of QMC is particularly hard. The proposed algorithms in literature use therefore an adaptive increase in samples, where additional samples are taken on the level that has the most *profit* (Giles, 2015; Robbe, 2019). The *profit* is typically characterised by a profit indicator, that is the ratio of the variance and the cost of a certain level. Thus, levels with high variance and low cost have a high profit and levels with low variance and high cost have a low profit.

In practical application there are typically multiple outputs, the multilevel technique applies then to each output individually. As such the MLMC is also applicable for estimating the central moments of a distribution. For the variance $Var[y]$ the MLMC estimator is defined as:

$$\mathbb{E}[Var[y^L]] = \mathbb{E}[Var[y^0]] + \sum_{l=1}^{L} \mathbb{E}[Var[y^l] - Var[y^{l-1}]] = \sum_{l=1}^{L} \mathbb{E}[\Delta Var[y^l]], \tag{28}$$

where $\mathbb{E}[Var[y^l]]$ and $\mathbb{E}[Var[y^{l-1}]]$ are positively correlated with the same set of MC or QMC samples. The MSE of this estimator is similar to the mean MLMC estimator, with the difference

that the bias and the variance of the variances have to be calculated, and is defined as:

$$MSE(Var[y^L]) = \sum_{l=0}^{L} Var[\Delta Var[y^l]] + (\mathbb{E}[Var[y^L] - Var[y]])^2, \tag{29}$$

where the variance $Var[\Delta Var[y^l]]$ can be bounded with an upper bound as in (Bierig, 2015). For more details about MLMC for estimating the variance, the reader is referred to this paper.

4.2. Interval Multilevel Quasi-Monte Carlo

In Interval Multilevel Quasi-Monte Carlo (I-MLQMC) the input intervals are represented by Cauchy random variables. The output interval y^I is then estimated with the parameters of the Cauchy output variable (i.e., δ_Y, γ_Y), similar as for the classical MC and QMC.

For each Cauchy parameter the MLQMC sum becomes:

$$\mathbb{E}[\delta_Y^L] = \sum_{l=0}^{L} \mathbb{E}[\Delta \delta_Y^l], \tag{30}$$

$$\mathbb{E}[\gamma_Y^L] = \sum_{l=0}^{L} \mathbb{E}[\Delta \gamma_Y^l], \tag{31}$$

where $\mathbb{E}[\delta_Y^{-1}] = 0$ and $\mathbb{E}[\delta_Y^{-1}] = 0$. The MSE for each parameter is then:

$$MSE(\delta_Y^L) = \sum_{l=0}^{L} Var[\Delta \delta_Y^l] + (\mathbb{E}[\delta_Y^L - \delta_Y])^2, \tag{32}$$

$$MSE(\gamma_Y^L) = \sum_{l=0}^{L} Var[\Delta \gamma_{Y^l}] + (\mathbb{E}[\gamma_Y^L - \gamma_Y])^2. \tag{33}$$

The I-MLQMC technique then estimates the two parameters using the variance and bias constraint of each parameter, thus an expected value and a variance of these quantities have to be computed. However, both variances can only be estimated as no direct solution is available. Therefore, the variance of the MLE and the HLE are typically estimated with the asymptotic variance as both estimators are asymptotic normally distributed. The asymptotic variance yields for small sample sizes, however, a smaller variance then the true variance. Therefore in this paper, an estimate of the true variance is obtained with bootstrapping (Efron, 1994).

Bootstrapping generates B sets of samples from one sample set, to estimate parameters of the true distribution. Assume that $\boldsymbol{s} = (s_1, s_2, \ldots, s_{10})$ is a set of 10 samples, then bootstrapping generates B sets $\boldsymbol{s}_b^*, b \in 1, 2, \ldots, B$ with replacement where, $\boldsymbol{s}_1^* = (s_1^*, \ldots, s_{10}^*)$, $B \in \mathbb{N}$ and * denotes a bootstrapping variable. For each set the quantity of interest y_b^* is calculated, for example the median δ_b^* to estimate the location parameter of a Cauchy distribution. The distribution of all the estimated medians follow the bootstrapping distribution $Y^* \sim \mathcal{B}(m^*, \sigma^*)$, from which the variance is obtained and used to estimated the variance of the sample median. The accuracy of

this bootstrapping estimate is increasing with higher values of B and higher sample sizes N. In the I-MLQMC such a bootstrapping approach is used to estimate the variance and mean of the sample median and the HLE of the scale parameter, respectively $Var[\Delta\delta_Y^l]$, $\mathbb{E}[\Delta\delta_Y^l]$, $Var[\Delta\gamma_Y^l]$ and $\mathbb{E}[\Delta\gamma_Y^l]$. The expected value and variance of the multilevel differences, $\Delta\delta_Y^l$ and $\Delta\gamma_Y^l$, is then estimated with an iterative bootstrapping approach. The iterative approach increases B, to ensure accurate estimates of the expected values and variances. The interactive approach repeats this until the difference between 2 estimates is smaller then 1%.

4.2.1. *Adaptive adding levels*

To estimate the bias it is assumed that the expected value can be calculated at all levels with a function defined as:

$$\mathbb{E}(\Delta\theta_Y^l) \leq c_1 2^{-\alpha l}, \tag{34}$$

where c_1 and α are positive constants independent of l. The original theorem can be found in (Giles, 2008; Robbe, 2019). The constant c_1 can be determined very easily as $\mathbb{E}(\Delta\theta_Y^l) = c_1$ and α by linear regression of $\log_2(|\mathbb{E}^l|), l = 1, 2, \ldots, L$. If $L < 2$ the bias is not calculated, as in this case, not enough levels are available for a reliable estimate of the constants. For $L > 2$ an estimation of the bias is calculated with:

$$|\mathbb{E}[\theta_Y^l - \theta_Y]| = \left| \sum_{l=L+1}^{\infty} \mathbb{E}[\Delta\theta_Y^l] \right| \leq \sum_{l=L+1}^{\infty} c_1 2^{-\alpha l} \leq c_1 \frac{2^{-\alpha L}}{2^{\alpha} - 1} \tag{35}$$

This estimation is then used to check the bias constraint, and if necessary levels are added by I-MLQMC until the bias constraint is obtained.

5. Case study: Thermal Stress Analysis of Jet Engine Turbine Blade

5.1. Problem description

This case study compares the I-MLQMC method with the vertex analysis for a linear elastic thermal stress analysis of a jet engine turbine blade. The turbine has the following boundary conditions: a pressure load on the pressure side p_1, pressure load on the suction side p_2 and a fixed Dirichlet boundary condition on the bottom plane $y = 0.05$ m. Figure 1 visualises the turbine blade with its mechanical boundary conditions.

The discretisation of the model is performed using linear tetrahedral elements with maximal chordal length h. The analysis starts with a linear thermal analysis and continues with a linear stress analysis to obtain the maximal displacement of the blade in the x-direction at the end of the blade (according to the positive y-axis). Other outputs, such as stresses are not considered in this case-study as that this could possibly lead to a non-linear combinations of Cauchy variables. This, would violate the necessary requirements of the method, being a linear combination of Cauchy variables. In total, the model requires 20 input variables to be defined for both the structural and thermal model. For more details about the model, the reader is referred to (MathWorks, 2019).

The input variables for the structural model are Young's modulus E of the blade material, the coëfficient of thermal expansion CTE, the Poisson's ratio ν, the pressure loads p_1 and p_2. The

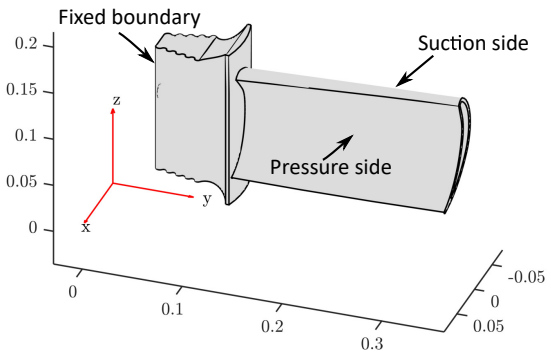


Figure 1. Jet engine turbine blade with the mechanical boundary conditions, a fixed boundary condition and two pressures: pressure side and suction side

thermal parameters consist of the thermal conductivity κ of the material and the temperatures surrounding the blade: T_a ambient, T_p pressure side, T_s suction side, T_t tip side, T_b base side and the T_r root. Also, convection coefficients on these same locations are defined. Table I summarises the intervals that are defined for these parameters.

Table I. Input interval scalars

Parameter	Interval scalar	unit	Parameter	Interval scalar	unit
$1 : E$	$[200 \cdot 10^9; 250 \cdot 10^9]$	[Pa]	$11 : T_b$	$[750; 850]$	[°C]
$2 : CTE$	$[11,5 \cdot 10^{-6}; 13,5 \cdot 10^{-6}]$	[1/K]	$12 : T_r$	$[350; 450]$	[°C]
$3 : \nu$	$[0,2; 0,35]$	[-]	$13 : T_i$	$[130; 170]$	[°C]
$4 : p_1$	$[4,8 \cdot 10^5; 5,2 \cdot 10^5]$	[Pa]	$14 : h_i$	$[28; 32]$	[W/m^{2}K]
$5 : p_2$	$[4,3 \cdot 10^5; 4,7 \cdot 10^5]$	[Pa]	$15 : h_p$	$[45; 55]$	[W/m^{2}K]
$6 : \kappa$	$[18; 22]$	[W/m/K]	$16 : h_s$	$[35; 45]$	[W/m^{2}K]
$7 : T_a$	$[280; 320]$	[°C]	$17 : h_t$	$[18; 22]$	[W/m^{2}K]
$8 : T_p$	$[900; 1100]$	[°C]	$18 : h_b$	$[35; 45]$	[W/m^{2}K]
$9 : T_s$	$[900; 1000]$	[°C]	$19 : h_r$	$[13; 17]$	[W/m^{2}K]
$10 : T_t$	$[900; 1000]$	[°C]	$20 : h_{rb}$	$[900; 1100]$	[W/m^{2}K]

5.2. Interval analysis

For this analysis the I-MLQMC and the I-MC are compared to vertex analysis for the 8 first and all 20 input intervals, according to table I.



Figure 2. Left: coarsest model with maximal element-size $h = 0,03$, Right: Fine model with maximal element-size $h = 0,0075$

In this case the input parameters of the I-MLQMC and I-MC are the input intervals and relative tolerances $e_\delta^r = 0,05$ and $e_\delta^r = 0,15$. The relative tolerance on the interval midpoint is smallest, because the midpoint has typically a higher value relative to the width. On each level, the amount of initial samples is selected as $N_i = 3$, the algorithms will then increase this number adaptively based on the variance. For the I-MLQMC procedure, the coarsest element size is set as $h_0 = 0,03$ m, which is visualised in the left figure of Figure 2. I-MLQMC is then based on the bias constraint decreasing the element-size freely if required, on the right figure of Figure 2 visualises a fine model with element size $h_0 = 0,0075$ m. The I-MC uses that fine model to achieve a low bias on the resulting interval in combination with the same variance constraint as the I-MLQMC method being 50% of the tolerance. The I-MLQMC uses 3 models, to obtain the required tolerance that are summarised in table II.

Table II. Comparison of I-MLQMC and I-MC with vertex analysis with the computational time and output interval

Vertex analysis: element size	Number of input intervals	Computational time	Interval estimation
$h_0 = 0,03$	8	4 min and 29 sec	$[0,19; 1,44]$ mm
$h_1 = 0,015$	8	11 min and 49 sec	$[0,30; 1,03]$ mm
$h_2 = 0,0075$	8	21 min and 20 sec	$[0,334; 1,08]$ mm
I-MC: tolerance	Number of input intervals	Computational time	Interval estimation
$e_\delta^r = 0,05, e_\delta^r = 0,15$	8	17 min and 45 sec	$[0,39; 1,01]$ mm
$e_\delta^r = 0,05, e_\delta^r = 0,15$	20	17 min and 45 sec	$[0,14; 1,26]$ mm
I-MLQMC: tolerance	Number of input intervals	Computational time	Interval estimation
$e_\delta^r = 0,05, e_\delta^r = 0,15$	8	16 min and 15 sec	$[0,30; 1,04]$ mm
$e_\delta^r = 0,05, e_\delta^r = 0,15$	20	17 min and 20 sec	$[0,12; 1,32]$ mm

With I-MLQMC, the resulting output interval y^I when limiting the analysis to the first 8 input intervals from table I is found to be $[0,30; 1,04]$. The total number of samples per level required are: 444 samples on level 0, 177 samples on level 1 and 3 samples on level 2. For the analysis with all 20 input intervals, the result is $[0,12; 1,32]$ with a number of samples listed as: 444 for level 0, 147 for level 1 and 33 for level 2. The total cost of these samples equals 16 min 15 sec[1] for 8 inputs and 17 min 20 sec for 20 inputs. This means there is almost no increase in computational effort for the larger 20-dimensional input space, while yielding similar accuracy. As such, it can be stated that I-MLQMC effectively breaks the so-called curse of dimensionality in interval analysis.

The I-MC achieves similar results to accuracy and with rather small increase in computational cost, 1 min 30 sec and 25 sec for 8 and 20 input intervals respectively. For the reason of comparison, a vertex analysis is performed as reference case and to shown the difference in computational time and accuracy. This techniques uses for the case of 8 input intervals $2^8 = 265$ forward model evaluations on a model with discretisation $h_0 = 0,03; h_1 = 0,015; h_2 = 0,0075$, these are also used in the I-MLQMC technique. The computational cost of performing the vertex analysis with 20 input intervals is prohibitively large as 1 048 576 deterministic model runs are required, yielding for the coarsest model 1 101 004 sec or almost 13 days of computing time. Therefore, these results were not included in this paper. For 8 input intervals the output interval is summarised for the three discretisation levels in table II, with the corresponding computing time.

The computation cost of the I-MC and I-MLQMC is for a given tolerance comparable to the vertex analysis in the case of 8 input intervals. This while reaching a result within the given relative tolerance values of $e^r_\delta = 0,05$ and $e^r_\delta = 0,15$. For the case of all 20 input intervals, the I-MC and I-MLQMC outperforms the vertex analysis with 17 min compared to 13 days. This shows the enormous gain in efficiency. The performance and accuracy difference between I-MC and I-MLQMC has to be studied in future.

6. Conclusion

This paper presents an approach to apply Multilevel Quasi-Monte Carlo to interval analysis. This by representing the intervals with Cauchy distributed variables, as they have the property that a linear combination of these variables is again Cauchy distributed. The input intervals are translated to Cauchy distributions by their interval midpoint and width: the midpoint of the interval equals the location of the Cauchy distribution and the width of the interval equals the scale parameter of the Cauchy distribution. Then on the output side of the linear FE-model again a Cauchy distribution is obtained, which is then translated to an interval by the scale parameter and the median of the Cauchy distribution. With a case-study it was shown that, from a computational perspective the Interval Monte Carlo and Interval Multilevel Qausi-Monte Carlo technique outperforms the vertex analysis for 20 input intervals, and was equal to the vertex analysis for 8 input intervals, while still achieving the accuracy of the vertex analysis for 8 input intervals. As such, it can be stated that Interval Monte Carlo and Interval Multilevel Quasi-Monte Carlo effectively breaks the so-called curse of dimensionality in interval analysis. The application of Interval Monte Carlo and Interval

[1] Computational time on Windows laptop with Intel(R) Core(TM) i7-9850H CPU @ 2.60GHz and 16Gb RAM.

Multilevel Quasi-Monte Carlo to non-linear models and an in depth comparison between Interval Monte Carlo and Interval Multilevel Quasi-Monte Carlo is left for future research.

Acknowledgements

The authors would like to acknowledge the financial support of the Research Foundation - Flanders in the framework of the research grant HIDIF (High dimensional interval fields) under grant number G0C2218N, the PhD fellowship strategic basic research 1SD2421N of Robin R.P. Callens and the postdoctoral grant 12P359N of Matthias G.R. Faes.

References

Matlab, Thermal Stress Analysis of Jet Engine Turbine Blade The MathWorks, Inc., 2019.

Borak, Szymon and Härdle, Wolfgang and Weron, Rafał Stable distributions. In Statistical tools for finance and insurance, 21–44, Springer, 2005

Dick, Josef and Pillichshammer, Friedrich Digital nets and sequences: discrepancy theory and quasi–Monte Carlo integration 2010, Cambridge University Press

Hickernell, Fred J Lattice rules: how well do they measure up? In Random and quasi-random point sets, 109–166, 1998, Springer

Freue, Gabriela V Cohen The Pitman estimator of the Cauchy location parameter In Journal of Statistical Planning and Inference, 137, 6, 1900–1913, 2007, Elsevier

Kravchuk, OY and Pollett, PK Hodges-Lehmann scale estimator for Cauchy distribution In Communications in Statistics-Theory and Methods, 41, 20, 3621–3632, 2012, Taylor & Francis

Heinrich, Stefan Multilevel monte carlo methods In proceedings International Conference LSSC 2001, 58–67, Springer

Giles, Michael B Multilevel monte carlo methods In Acta Numerica, 24, 259, 2015, Cambridge University Press

Charrier, Julia and Scheichl, Robert and Teckentrup, Aretha L Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods In SIAM Journal on Numerical Analysis, 51, 1, 322–352, 2013, SIAM

Robbe, Pieterjan Multilevel Uncertainty Quantification Methods for Robust Design of Industrial Applications Ph. D. Thesis, 2019, KU Leuven, Leuven, Belgium

Giles, Michael B Multilevel monte carlo path simulation In Operations research, 56, 3, 607–617, 2008, INFORMS

Kreinovich, Vladik and Ferson, Scott A A new Cauchy-based black-box technique for uncertainty in risk analysis In Reliability Engineering & System Safety, 85, 1-3, 267–279, 2004, Elsevier

Efron, Bradley and Tibshirani, Robert J An Introduction to the Boostrap 1994, CRC press

Beer, Michael and Kreinovich, Vladik Interval or moments: which carry more information? In Soft Computing, Springer, 2013, 17, 1319-1327

Qiling Zou and Lu Deng and Chao Jiang Predicting the Bounds of Vehicle-Induced Bridge Responses Using the Interval Analysis Method In Journal of Bridge Engineering, 2016, 21

Xiaojun Wang and Lei Wang and Zhiping Qiu A feasible implementation procedure for interval analysis method from measurement data In Applied Mathematical Modelling, 2014, 38, 2377 - 2397

X.Y. Long and C. Jiang and K. Liu and X. Han and W. Gao and B.C. Li An interval analysis method for fatigue crack growth life prediction with uncertainty In Computers & Structures, 2018, 210, 1-11

Dannert, Mona M and Faes, Matthias GR and Fleury, Rodolfo MN and Fau, Amelie and Nackenhorst, Udo and Moens, David Imprecise random field analysis for non-linear concrete damage analysis In Mechanical Systems and Signal Processing, 2020, 150, 107343, Elsevier

Jingwen Song and Pengfei Wei and Marcos Valdebenito and Sifeng Bi and Matteo Broggi and Michael Beer and Zuxiang Lei Generalization of non-intrusive imprecise stochastic simulation for mixed uncertain variables In Mechanical Systems and Signal Processing, 2019, 134, 106316

Faes, Matthias and Moens, David Recent trends in the modeling and quantification of non-probabilistic uncertainty In Archives of Computational Methods in Engineering, 2019, 1-39

Hansen, E. Global optimization using interval analysis the multi-dimensional case. Numer. Math. 34, 247–270, 1980

Markus Mäck and Michael Hanss. Efficient possibilistic uncertainty analysis of a car crash scenario using a multi-fidelity approach In ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg, 5(4), 2019

Brandolini, Luca and Colzani, Leonardo and Gigante, Giacomo and Travaglini, Giancarlo On the koksma–hlawka inequality In Journal of Complexity, 2013, 29, 2, 158–172, Elsevier

Aistleitner, Christoph and Dick, Josef Functions of bounded variation, signed measures, and a general Koksma-Hlawka inequality In arXiv preprint arXiv:1406.0230 2014

C. Bierig, A. Chernov, Convergence analysis of multilevel monte carlo variance estimators and application for random obstacle problems In Numerische Mathematik, 130, 2015, 579-613

M. Pisaroni, S. Krumscheid, F. Nobile, Quantifying uncertain system outputs via the multilevel monte carlo method—part I: Central moment estimation In Journal of Computational Physics, 2020, 109466

Topology Optimization with Polymorphic Uncertainties using Artificial Neural Networks

Steffen Freitag, Simon Peters, Philipp Edler, Günther Meschke
Institute for Structural Mechanics, Ruhr University Bochum, 44801, Bochum, Germany,
steffen.freitag@sd.rub.de, {simon.peters, philipp.edler, guenther.meschke}@rub.de

Abstract. In this paper, a topology optimization approach is presented, where uncertain load and material parameters are considered. The concept of compliance minimization, i.e. stiffness maximization, is applied based on a plane stress finite element formulation. In order to take uncertain structural load parameters and uncertain material behavior into account, the topology optimization is embedded into a reliability-based design optimization approach, where uncertain structural parameters and design variables are quantified as random variables, intervals and probability boxes (p-boxes). This allows to consider aleatory and epistemic uncertainties by means of polymorphic uncertainty models within the topology optimization. Solving optimization problems with random variables, intervals and p-boxes leads to a high computational effort, because the objective functions and constraints have to evaluated millions of times. To speed up the optimization process, the finite element simulation of the topology optimization is replaced by artificial neural networks. This includes not only the topology dependent maximal stresses and displacements of the structure, which are used as constraints, but also the material density distribution inside the design domain. An example is presented, where the material volume of a cantilever structure is minimized, considering interval uncertainty only, combinations of interval and stochastic uncertainty and also polymorphic uncertainty (p-boxes) for the load, the material parameters and the geometry of the structure.

Keywords: Topology Optimization, Intervals, Random Variables, Probability Boxes, Polymorphic Uncertainty, Artificial Neural Network

1. Introduction

Structural optimization is focused on the design of engineering structures using the construction material in an efficient way. Objectives are in general minimizing the self weight, maximizing the stiffness or balancing the stresses of all structural members. Beside these objectives, constraints with respect to the structural load bearing capacity (e.g. strength of materials) and the serviceability (e.g. maximal displacements) have to be considered. The design variables of structural optimization problems are either sizing or shape parameters defining the dimensions of the structural members (shape optimization) or information about the material distribution defining the topology of the structure (topology optimization), see e.g. (Bendsøe and Sigmund, 2004).

Topology optimization has been successfully applied to improve the design of many real world applications, such as in automotive (Cavazzuti et al., 2011) or aerospace (París et al., 2012). Due

© 2021 *by S. Freitag, S. Peters, P. Edler, G. Meschke.*

to innovations in additive manufacturing, its relevance for complex industrial problems has greatly increased. However, in civil engineering the number of applications is limited (Beghini et al., 2014) owing to e.g. the conservative industry in consequence of high-cost combined with high risk projects or the low material cost compared to regular hourly wage rate of civil engineers and construction worker tend to a mentality of over-sized and quickly mounted structures. However, the global cement production is the third-largest source of anthropogenic emissions of carbon dioxide (Andrew, 2018), which could lead to increasing prices for cement-based materials such as concrete due to policy measures such as carbon taxes soon (Lin and Li, 2011). Topology optimization could be one of the promising tools for more ecologically and architecturally appealing buildings (Beghini et al., 2014; Naboni and Paoletti, 2018). In civil engineering, topology optimization of structures is challenging, because of low volume fractions requiring fine design resolutions (Baandrup et al., 2020), which may be hard to be manufactured especially for large scale structures. Often, multi objectives and multi constraints have to be considered. In addition to this, aleatory and epistemic uncertainties of structural parameters should be taken into account in order to get robust designs.

Reliability-based and robust design optimization approaches have been developed to capture stochastic structural parameters and stochastic design variables, see e.g. (Schuëller and Jensen, 2008) and (Valdebenito and Schuëller, 2010). This requires to define appropriate uncertainty measures, e.g. by means of mean values, standard deviations and quantile values, to evaluate the objective functions and uncertain constraints. In case of polymorphic uncertainties, the reliability-based design optimization approaches are extended, which means that the stochastic uncertainty measures have to be combined with non-stochastic uncertainty measures to evaluate the objective functions and constraints, e.g. by lower or upper bounds for intervals (Edler et al., 2019) or by credibility levels for fuzzy numbers, see e.g. (Mäck et al., 2019).

Topology optimization approaches under consideration of stochastic uncertainties can be performed in the framework of reliability-based topology optimization, see e.g. (López et al., 2016) or robust optimization, see e.g. (Lazarov et al., 2012). Approaches for interval uncertainties have been developed e.g. in (da Silva et al., 2019) and (Wang and Gao, 2020). In (da Silva et al., 2020), a comparison on stochastic (robust and reliability-based) and non-stochastic (intervals for worst case) topology optimization approaches is presented. In this paper, a topology optimization approach is introduced, where both, aleatory and epistemic uncertainties are considered by combining stochastic and non-stochastic uncertainty models by means of random variables, intervals and p-boxes within the concept of polymorphic uncertainty quantification.

Solving optimization problems with random variables, intervals and p-boxes leads to a high computational effort, because the objective functions and constraints have to be evaluated millions of times. In order to reduce the computational effort, the finite element simulation model of the topology optimization process is approximated by artificial neural networks. Several feedforward networks are trained sequentially to efficiently evaluate the stress and displacement constraints of the topology optimization problem. This approach is based on the multilevel surrogate modeling strategy for the objective function computation introduced in (Freitag et al., 2020). Moreover, a feedforward network with high dimensional output is trained to learn the dependency between the selected material volume fraction (area fraction for 2D problems) and the corresponding optimal topology by means of the material densities inside the design domain. This allows one to quickly predict the optimal topologies with the artificial neural network.

2. Topology optimization

The first concept in structural optimization known as topology optimization, dating back to 1988 (Bendsøe and Kikuchi, 1988), distributes the densities of the discretized design domain, based on the use of an artificial composite material. Since then many different approaches where investigated, see e.g. (Sigmund and Maute, 2013).

2.1. Compliance minimization

In this paper, the well developed concept of compliance minimization $\min_\rho : c(\rho)$ (stiffness maximization) is used, where the problem for a fixed design domain with multiple deterministic loading conditions ($i = 1, \ldots, N_i$ load cases) and a volume constraint takes the following form:

$$\min_{\rho} : c(\rho) = \sum_{i=1}^{N_i} \mathbf{u_i^T} \cdot \mathbf{K}(\rho) \cdot \mathbf{u_i} \tag{1}$$

$$s.t. \begin{cases} \mathbf{K}(\rho) \cdot \mathbf{u} = \mathbf{f} \\ V(\rho) = V_f \cdot V_0 \\ 0 < \rho_{min} \leq \rho_e \leq 1 \end{cases} \tag{2}$$

In Eqs. (1) and (2), $\mathbf{u}, \mathbf{f}$ and $\mathbf{K}$ are the global displacement vector, the global force vector and the global stiffness matrix, respectively. The material density vector ρ contains the relative material densities ρ_e of all finite elements $e = 1, \ldots, N_e$, where N_e is the number of finite elements. The relative material densities ρ_e are defined in $[0, 1]$, but a minimum relative density ρ_{min} is considered for numerical reasons. V_0 is the material volume of the complete design domain, and

$$V(\rho) = \sum_{e=1}^{N_e} V_e \cdot \rho_e \tag{3}$$

is the material volume of the optimized structure for a prescribed volume fraction V_f, where V_e is the volume of a finite element e.

The problem formulation in Eqs. (1) and (2) follows a density approach (Bendsøe and Sigmund, 2004; Bendsøe, 1989), hence the stiffness matrix $\mathbf{K}$ depends on the material density. This relation could be chosen with anisotropic rank 3 laminates, which leads to an optimal solution of the optimization problem in Eqs. (1) and (2), see e.g. (Allaire et al., 2019; Bendsøe and Sigmund, 2004) for details. But this solution contains many finite elements with intermediate densities between 0 and 1, which are difficult to manufacture. In order to force black-white solutions, the modified Solid Isotropic Material with Penalization (SIMP) approach is used to express this connection. With the SIMP approach, the modulus of elasticity of each finite element is computed by

$$E_e(\rho_e) = E_{min} + {\rho_e}^p (E_0 - E_{min}), \qquad \rho_e \in [0, 1] \tag{4}$$

where E_0 is the modulus of elasticity of the material, E_{min} is a small modulus of elasticity defined for numerical reasons to avoid singularities in the stiffness matrix and p is the penalization parameter

(Andreassen et al., 2011), which is chosen as $p = 3$ in this work. As a result, the compliance for a single load case is calculated element wise as

$$c(\rho_e) = \sum_{e=1}^{N_e} E_e(\rho_e) \cdot \mathbf{u}_e^T \cdot \mathbf{k}_0 \cdot \mathbf{u}_e \ , \tag{5}$$

where $\mathbf{u}_e$ is the element displacement vector and $\mathbf{k}_0$ is the element stiffness matrix for a unit modulus of elasticity ($E = 1$).

2.2. Topology optimization with interval load position

Most of the topology optimization approaches assume deterministic conditions for the input data, obviating the different sources of uncertainties, which may affect significantly the structural performance. In the classical design concepts, the uncertainty of loads and the structural resistance is considered by partial safety factors and worst case analyses. Compared to the safety factor concept, the required volume fraction of a structure can be reduced by the direct implementation of aleatory and epistemic uncertain parameters and design variables into the topology optimization process (López et al., 2016). For the consideration of these uncertainties in topology optimization, two main approaches have been developed, i.e. robust and reliability-based formulations (Lazarov et al., 2012; Thore et al., 2017).

In this work, a reliability-based formulation with polymorphic uncertain parameters is developed by combining the topology optimization with a reliability-based global optimization to consider probabilistic constraints. Within the previously described topology optimization approach, an interval load position $\overline{x}_F = [{}_l x_F, {}_u x_F]$ is taken into account as a set of independent deterministic load positions $x_{F_1}, \ldots, x_{F_i}, \ldots, x_{F_{N_i}}$, where all N_i load positions are defined to be placed between the lower interval bound $x_{F_1} \geq {}_l x_F$ and the upper interval bound $x_{F_{N_i}} \leq {}_u x_F$. In every topology optimization step, the N_i load positions are considered as independent load cases, which are solved and superposed for the sensitivity and compliance calculation (Sigmund, 2001). Therefore, the sensitivities in each finite element are calculated as

$$\frac{dc_e}{d\rho_e} = \sum_{i=1}^{N_i} -p \cdot {\rho_e}^{p-1} \cdot (E_0 - E_{min}) \cdot \mathbf{u}_{e,i}^T \cdot \mathbf{k}_0 \cdot \mathbf{u}_{e,i} \ . \tag{6}$$

3. Reliability-based topology optimization with polymorphic uncertain parameters

3.1. Polymorphic uncertain parameters

In structural optimization, it is distinguished between uncertain design variables and uncertain a priori parameters, see e.g. (Edler et al., 2019), which both may influence the objective function and the constraints of an optimization problem. Uncertain design variables of an objective function to be optimized, allows one to take tolerances or variability of the design variables into account. In general, a reference value of an uncertain design variable, e.g. the mean value, is defined to

be optimized, but each design realization results in an uncertain structural response. Whereas uncertain design variables are varied for solving an optimization problem, uncertain a priori parameters are constant during the optimization, i.e. they cannot be optimized, but they also lead to an uncertain structural response. Within the concept of polymorphic uncertainty modeling, a topology optimization approach is presented, where uncertain parameters are considered as random variables, intervals and p-boxes.

Random variables are quantified by stochastic distribution functions, e.g. lognormal distribution or Gaussian distribution, which are defined by a probability density function (PDF) and the corresponding cumulative distribution function (CDF). In this work, three stochastic a priori parameters are considered as random variables, the structural load F, the modulus of elasticity E and the yield strength f_y.

In addition to the interval load position $\overline{x}_F$ introduced as interval a priori parameter of the topology optimization in Section 2.2, an interval design variable is defined to quantify an imprecise width of the structure $\overline{b}$ within a 2D finite element formulation of the presented topology optimization approach. The interval width $\overline{b}$ is represented by a range with lower and upper bounds

$$\overline{b} = [{}_l b,\ {}_u b] \ . \tag{7}$$

The interval midpoint

$${}_m b = \frac{1}{2} \cdot ({}_l b + {}_u b) \tag{8}$$

is defined to be optimized and the interval radius

$${}_r b = \frac{1}{2} \cdot ({}_u b - {}_l b) \tag{9}$$

is kept constant during the optimization.

Combining random variables and intervals as inputs of a structural simulation results in p-box of the structural response, i.e. the quantities of interest such as the maximal von Mises stress $\overline{\sigma}_V$ and the maximal displacement $\overline{w}$, which are used to evaluate the constraints of the optimization problem. A p-box is defined as an imprecise random variable with a lower bound CDF (e.g. ${}_l F(\sigma_V)$, ${}_l F(w)$) and an upper bound CDF (e.g. ${}_u F(\sigma_V)$, ${}_u F(w)$), which leads to interval probabilities, see e.g. (Ferson et al., 2003). Here, the lower and upper bound CDFs of the resulting p-boxes are represented by empirical distributions obtained from Monte Carlo simulations, which are denoted as free p-boxes according to (Schöbi and Sudret, 2017). In addition to have polymorphic uncertainty only in the results, p-boxes can also be considered directly as uncertain inputs. In this case, parametric p-boxes (Schöbi and Sudret, 2017) are used, which are quantified by a bunch of random variables with interval distribution parameters, e.g. interval mean values.

3.2. Objective function and design variables

The topology optimization problem given by Eqs. (1) and (2) is based on the minimization of the compliance of a structure for a predefined volume fraction V_f. This means that for each realization of V_f, where $0 \leq V_f \leq 1$, an optimal material volume distribution $V(\rho)$ within the design domain volume V_0 is obtained, where the material density ρ_e of each finite element e is a design variable.

In order to minimize the material volume V of the structure, a reliability-based global optimization problem is formulated. Here, a 2D plane stress finite element formulation is used to solve the topology optimization problem, see Eqs. (1) and (2). The volume of the structure, i.e. the objective function to be minimized, is given by

$$\text{min:} \quad {}_mV(A_f, \overline{b}) = A_f \cdot A_0 \cdot {}_mb \tag{10}$$

$$\text{s.t.} \quad \begin{cases} g_1 = {}_uP_{f,lbc} - P_{f,lbc,ac} \leq 0 \\ g_2 = {}_uP_{f,ser} - P_{f,ser,ac} \leq 0 \ , \end{cases} \tag{11}$$

where A_f is the area fraction of the design domain area A_0 and $\overline{b}$ is the interval width of the plane. The first design variable A_f is deterministic defined in $0 \leq A_f \leq 1$ and the second design variable $\overline{b}$ is an interval with midpoint ${}_mb$ to be optimized and a fixed radius ${}_rb$ taking construction impressions into account. Because of the interval design variable $\overline{b}$, the volume of the structure $\overline{V}$ is also an interval.

Here, the midpoint ${}_mV$ of the interval volume $\overline{V}$ is defined as objective function, because an optimization problem with interval design variables cannot directly be solved, see e.g. (Edler et al., 2019). However, it can be seen in Eq. (10), that the optimization problem is linear, which means that also lower or upper bounds of $\overline{V}$ can be used as objective function, and that in this case, the optimum is completely defined by one of the two constraint limit states ($g_1 = 0$ or $g_2 = 0$) in Eq. (11).

3.3. Constraints

The first inequality constraint $g_1 \leq 0$ in Eq. (11) is defined according to the load bearing capacity limit state of the structure. Here, the upper bound of the interval failure probability of the structure with respect to the load bearing capacity ${}_uP_{f,lbc}$ is given as the upper bound (worst case) probability that the p-box of the von Mises stress $\overline{\sigma}_V$ is exceeding the stochastic distributed yield strength of the material f_y in at least one element of the structure. $P_{f,lbc,ac}$ is the accepted failure probability with respect to the load bearing capacity.

The second inequality constraint $g_2 \leq 0$ in Eq. (11) is defined according to the serviceability limit state of the structure. In this work, the upper bound of the interval failure probability of the structure with respect to the serviceability ${}_uP_{f,ser}$ is defined as the upper bound (worst case) probability that the p-box of the displacement $\overline{w}$ is exceeding a tolerated displacement w_{ac} in at least one node of the structural finite element model. $P_{f,ser,ac}$ is the accepted failure probability with respect to the serviceability.

3.4. Solution of the optimization problem

The reliability-based topology optimization problem with polymorphic uncertain parameters is solved by a particle swarm optimization algorithm (Kennedy and Eberhart, 1995). In each optimization step, the objective function and the constraints have to be evaluated for each particle of the swarm. In Figure 1, the computational scheme of the reliability-based topology optimization is shown.

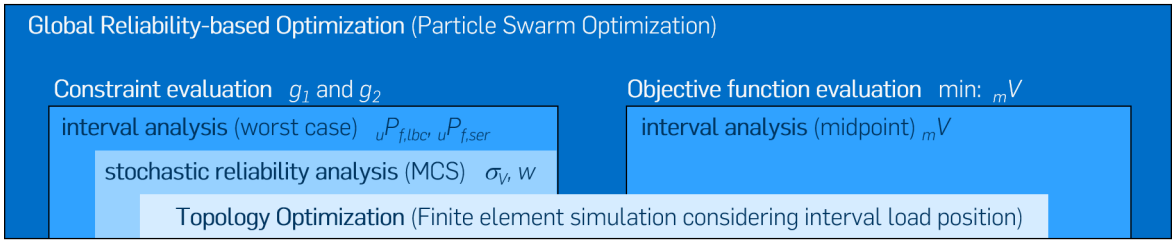


Figure 1. Computational scheme for reliability-based topology optimization.

The basis of the reliability-based topology optimization is the finite element based topology optimization presented in Section 2, which computes the optimal topology (element material densities $\rho_e(A_f)$), the maximal displacement $w^*(A_f)$ and the maximal von Mises stress $\sigma_V^*(A_f)$ for a selected area fraction A_f (deterministic design variable) and under consideration of an interval load position $\overline{x}_F$ (interval a priori parameter). It should be noted, that because of the linearity of the topology optimization problem, a scaled load F^*, a scaled modulus of elasticity E^* and a scaled width of the structure b^* is used for the finite element simulation to compute the scaled displacement w^* and the scaled von Mises stress σ_V^*. This allows one to use the FE results of one realization of the deterministic design variable A_f for the constraints evaluation and for the objective function evaluation by a simple post-processing of the FE results.

For the constraints evaluation, the p-box of the maximal von Mises stress is obtained by

$$\overline{\sigma}_V(A_f, \overline{b}) = \frac{F}{F^*} \cdot \frac{b^*}{\overline{b}} \cdot \sigma_V^*(A_f) \,, \tag{12}$$

and the p-box of the maximal displacement is computed by

$$\overline{w}(A_f, \overline{b}) = \frac{F}{F^*} \cdot \frac{b^*}{\overline{b}} \cdot \frac{E^*}{E} \cdot w^*(A_f) \,. \tag{13}$$

Equations (12) and (13) are solved by Monte Carlo Simulations (MCS) in combination with an interval analysis. Because of the monotonicity, it is sufficient to evaluate the lower bound ${}_l b$ of $\overline{b}$ within the interval analysis. Based on the results of Eqs. (12) and (13), the upper bound of the interval failure probability of the structure with respect to the load bearing capacity ${}_uP_{f,lbc} = P({}_u\sigma_V \leq f_y)$ and the upper bound of the interval failure probability of the structure with respect to the serviceability ${}_uP_{f,ser} = P({}_uw \leq w_{ac})$ and the corresponding values of the limit state functions g_1 and g_2 are computed by Eq. (11).

For the objective function evaluation, the midpoint ${}_mV$ of the interval volume $\overline{V}$ is computed by Eq. (10). In addition to the midpoint, also the corresponding interval volume can be determined by

$$\overline{V}(A_f, \overline{b}) = A_f \cdot A_0 \cdot \overline{b} \,, \tag{14}$$

where the interval bounds of $\overline{V}$ are obtained by interval arithmetic, i.e. by evaluating Eq. (14) for the lower and upper bounds of $\overline{b}$.

4. Artificial neural networks in topology optimization

The solution of the reliability-based topology optimization problem in Section 3 requires multiple runs of the finite element simulation based topology optimization with realizations of the deterministic design variable A_f to compute the element material densities $\rho_e(A_f)$ (deterministic topology design variables), and the corresponding maximal displacement $w^*(A_f)$ (serviceability constraint) as well as the maximal von Mises stress $\sigma_V^*(A_f)$ (load bearing constraint). In order to reduce the computation time, artificial neural networks are trained to approximate the finite element simulation.

4.1. Constraints approximation with artificial neural networks

For the computation of the maximal displacement $w^*(A_f)$ and the maximal von Mises stress $\sigma_V^*(A_f)$ two feedforward networks are generated based on finite element simulation results:

- ANN1_1: $A_f \mapsto \sigma_V^*$
- ANN2_1: $A_f \mapsto w^*$

Both networks ANN1_1 and ANN2_1 have one input neuron and one output neuron. The number of hidden layers and neurons are selected according to the complexity of the finite element simulation model, i.e. the topology optimization design domain. As can be seen in the example presented in Section 5, one hidden layer with a few neurons is sufficient to approximate a 2D plane stress finite element model for a rectangular design domain.

In order to further improve the performance of the optimization algorithm, two additional feedforward networks are generated based on results of ANN1_1 and ANN2_1:

- ANN1_2: $A_f, {}_m b \mapsto D_1\{g_1 = 0\}$
- ANN2_2: $A_f, {}_m b \mapsto D_2\{g_2 = 0\}$

ANN1_2 and ANN2_2 both have two inputs, the design variables A_f and ${}_m b$, and one output, which is the shortest distance $D_1\{g_1 = 0\}$ and $D_2\{g_2 = 0\}$ between a position in the design space and the constraint limit states $g_1 = 0$ and $g_2 = 0$, respectively. In order to compute the distances $D_1\{g_1 = 0\}$ and $D_2\{g_2 = 0\}$, the constraint limit states $g_1 = 0$ and $g_2 = 0$ are approximated by a space subdividing technique, see (Edler et al., 2019). The ANN-based distance function approximation allows one to efficiently move particles of the swarm, which are in the unfeasible region of the design space onto the constraint limit states, where the optimum of the linear optimization problem is located.

For all four artificial neural networks ANN1_1, ANN2_1, ANN1_2 and ANN2_2 the hyperbolic tangent activation function is used in the hidden neurons and the output neurons are activated by a linear activation function.

4.2. Topology prediction with artificial neural networks

In this work, an additional feedforward network is generated based on finite element simulation results to predict the optimal topology, i.e. the optimal material density ρ_e of each finite element e, for a given area fraction A_f:

– ANN3: $A_f \mapsto \rho_e$ with $e = 1, \ldots, N_e$

This ANN3 has one input A_f and a high dimensional (N-dimensional) output ρ_e, where the dimension N corresponds to the number of finite elements of the design domain. For such a high dimensional mapping, a deep neural network with several hidden layers and a lot of hidden neurons is required, see Section 5.

Because the material density is defined as $\rho_e \in [0, 1]$, the sigmoide activation function is used in all hidden neurons and also in the output neurons.

5. Example

In this section, a classical benchmark problem for topology optimization is presented to illustrate the described method for reliability-based topology optimization under consideration of polymorphic uncertainty. In Figure 2, the design domain (80 cm length and 25 cm height) and the boundary conditions of the investigated cantilever structure is shown. At the left boundary of the cantilever, the displacements are completely fixed and a vertically concentrated load $F^* = 10$ kN is applied at the top right side. The imprecise position of this vertical load is quantified as an interval $\overline{x}_F = [72, 80]$ cm with an interval width of 8 cm, see Figure 2.

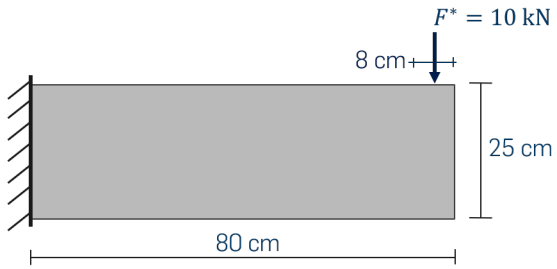


Figure 2. Design domain and boundary conditions for the topology optimization of a cantilever beam.

5.1. Finite element model for the topology optimization

For the finite element analysis of the cantilever, a structured mesh of 80 × 25 four-node bilinear square elements is used over the entire rectangular design domain, i.e. 2000 elements in total with dimensions 1 cm × 1 cm. Plane stress conditions and a width of $b^* = 1$ cm are selected. The whole structure is made of steel S235, therefore the elastic material is assumed to be linearly

isotropic with modulus of elasticity $E^* = 21000$ kN/cm^2, Poisson ratio $\nu = 0.3$ and yield strength $f_y^* = 23.5$ kN/cm^2.

The deterministic topology optimization problem according to Eq. (1) is solved for 65 discrete area fractions A_f in the range of 0.2 to 1.0, i.e. A_f is systematically reduced from 1.0 to 0.2 with a step size of $\Delta A_f = 0.0125$. It should be noted that the area fraction A_f is equivalent to the volume fraction V_f. The interval load position $\overline{x}_F$ is considered by evaluating Eq. (6) for $N_i = 9$ independent load cases. This means that the load is applied at nine different positions of the 8 cm wide loading zone and all 9 load positions are considered for the sensitivity and compliance calculation in each optimization step.

For the finite element simulations, an in house MATLAB code is used, which is based on (Andreassen et al., 2011). In the simulations, a sensitivity filter, where r_{min} is set to 3 cm, is used to prevent checkerboard patterns, see (Bendsøe and Sigmund, 2004) for details about filtering methods in general. For the 65 area fractions, the optimal topologies (final element material density ρ_e for each of the $N_e = 2000$ finite elements) and the corresponding maximal displacements w^* as well as the corresponding maximal von Mises stresses σ_V^* are stored.

5.2. Artificial neural network training

For ANN1$_1$ and ANN2$_1$, feedforward networks with one hidden layer comprising five hidden neurons are trained to approximate the dependencies between the area fraction A_f and the maximal von Mises stress $\sigma_V^*(A_f)$ as well as the dependencies between the area fraction A_f and the maximal displacement $w^*(A_f)$, respectively. Only 21 out of the 65 finite element simulation results have been used to set up the neural networks. Because it has been decided to restrict the design space for the area fraction A_f to the range of 0.2 to 0.7, the neural networks have only been trained for this range. In Figure 3, the neural network predictions are compared with finite element simulation results.

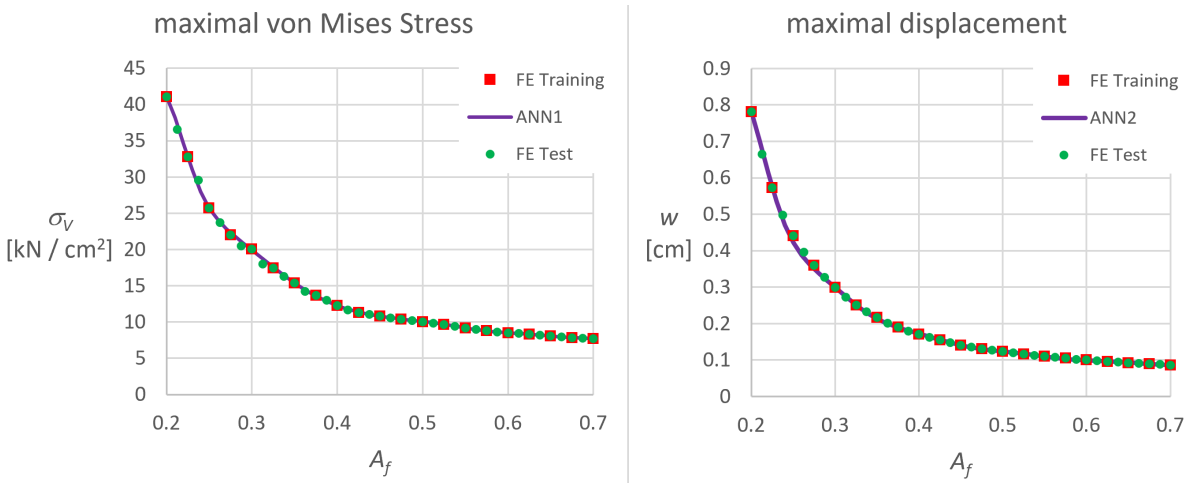


Figure 3. Comparison of ANN1$_1$ (left) and ANN2$_1$ (right) with finite element simulation results (FE) used for training and testing.

Using results of ANN1_1 and ANN2_1 in combination with Monte Carlo simulations to compute the failure probabilities with respect to the load bearing capacity and the serviceability, two additional artificial neural networks (ANN1_2 and ANN2_2) are trained to approximate the shortest distance to the constraint limit states $g_1 = 0$ and $g_2 = 0$, respectively. Also one hidden layer is sufficient for these networks, where the number of hidden neurons is 10 for ANN1_2 and 4 for ANN2_2.

In order to learn the optimal topology (the optimal material density distribution ρ_e of the structure) for specific area fractions A_f, an ANN3 is created, which has 2000 output neurons corresponding to the number of finite elements. For this challenging task, a deep feedforward neural network with three hidden layers and 1000 hidden neurons per layer is used.

It should be noted, that for all ANNs different architectures (with different numbers of hidden layers and hidden neurons) have been investigated and here only the finally used ANN architectures are described.

5.3. Results of the reliability-based topology optimization

The topology optimization is performed with different number of uncertain a priori parameters and finally also with an interval design variable. In order to investigate the influence of the uncertain parameters to the optimal design, three topology optimization problems are formulated according to the following parameters:

- **Problem 1**: optimization only with interval load position
  - deterministic design variables:
    - A_f
    - $\rho_e(A_f)$
    - b
  - uncertain a priori parameters:
    - $\overline{x}_F = [72, 80]$ cm (interval)
  - deterministic constraints:
    - $g_1 = \sigma_V^* - f_y^* \leq 0$
    - $g_2 = w^* - 0.3\ cm \leq 0$
- **Problem 2**: optimization with interval load position and stochastic a priori parameters
  - deterministic design variables:
    - A_f
    - $\rho_e(A_f)$
    - b
  - uncertain a priori parameters:
    - $\overline{x}_F = [72, 80]$ cm (interval)
    - F (normal distributed random variable with mean value $\mu(F) = 10$ kN and standard deviation $\sigma(F) = 2$ kN)

  - E (lognormal distributed random variable with mean value $\mu(E) = 21000$ kN/cm^2 and standard deviation $\sigma(E) = 1050$ kN/cm^2)
  - f_y (lognormal distributed random variable with mean value $\mu(f_y) = 23.5$ kN/cm^2 and standard deviation $\sigma(f_y) = 1.175$ kN/cm^2)
- stochastic constraints:
  - $g_1 = P_{f,lbc} - P_{f,lbc,ac} \leq 0$ (with $P_{f,lbc} = P(\sigma_V \leq f_y)$ and $P_{f,lbc,ac} = 7 \cdot 10^{-5}$)
  - $g_2 = P_{f,ser} - P_{f,ser,ac} \leq 0$ (with $P_{f,ser} = P(w \leq 0.3$ cm) and $P_{f,ser,ac} = 6.7 \cdot 10^{-2}$)

– **Problem 3**: optimization with interval load position, stochastic a priori parameters and interval design variable

- design variables:
  - A_f (deterministic)
  - $\rho_e(A_f)$ (deterministic)
  - $\overline{b}$ (interval, with midpoint ${}_m b$ to be optimized and fixed radius ${}_r b = 1$ mm)
- uncertain a priori parameters:
  - $\overline{x}_F = [72, 80]$ cm (interval)
  - F (normal distributed random variable with mean value $\mu(F) = 10$ kN and standard deviation $\sigma(F) = 2$ kN)
  - E (lognormal distributed random variable with mean value $\mu(E) = 21000$ kN/cm^2 and standard deviation $\sigma(E) = 1050$ kN/cm^2)
  - f_y (lognormal distributed random variable with mean value $\mu(f_y) = 23.5$ kN/cm^2 and standard deviation $\sigma(f_y) = 1.175$ kN/cm^2)
- interval-stochastic constraints:
  - $g_1 = {}_u P_{f,lbc} - P_{f,lbc,ac} \leq 0$ (with ${}_u P_{f,lbc} = P({}_u \sigma_V \leq f_y)$ and $P_{f,lbc,ac} = 7 \cdot 10^{-5}$)
  - $g_2 = {}_u P_{f,ser} - P_{f,ser,ac} \leq 0$ (with ${}_u P_{f,ser} = P({}_u w \leq 0.3$ cm) and $P_{f,ser,ac} = 6.7 \cdot 10^{-2}$)

It should be noted, that problem 3 is the general case, which has been described in Sections 2 and 3. Problem 1 and problem 2 are selected just for the comparison of the results. The uncertainty of the parameters is increased step by step from problem 1 to problem 3. The objective of all three optimization problems is the minimization of the midpoint ${}_m V$ of the required material volume $\overline{V}$ according to Eq. (10), where for problem 1 and problem 2, the required material volume is deterministic, i.e. ${}_m V = V$.

The three optimization problems are solved by particle swarm optimization using five particles. For the evaluation of the constraints and the prediction of the optimal topologies, the artificial neural network surrogate models are applied. In Figure 4, the results of the three optimization problems are presented, see also Table I. The optima are all placed at the constraint limit state functions, which are marked by continuous and dashed lines in Figure 4. The corresponding optimal topologies for the three optimization problems are shown in Figure 5. It can be seen, that the topology predictions based on the high dimensional artificial neural network ANN3 fit very good

with the results of the finite element (FE) simulations. Please note that these results were not used to train the ANN.

As expected, the midpoint ${}_mV$ of the required material volume is increasing with increasing uncertainty in order to guarantee robust designs. However, this is mainly achieved by increasing the midpoint of the width ${}_mb$ of the structure, because surprisingly the area fraction A_f is reduced for increasing uncertainty. One reason could be, that the sensitivities for the varying width of all finite elements are linear, due to the 2D formulation. Moreover lowering the area fraction A_f leads towards other optimized designs of the structure, i.e. only the densities of less important elements (elements with low sensitivities) are reduced to zero. It follows that the trailing optimization process first reduces the densities of all elements with less than linear sensitivities and supports all elements with higher than linear sensitivities.

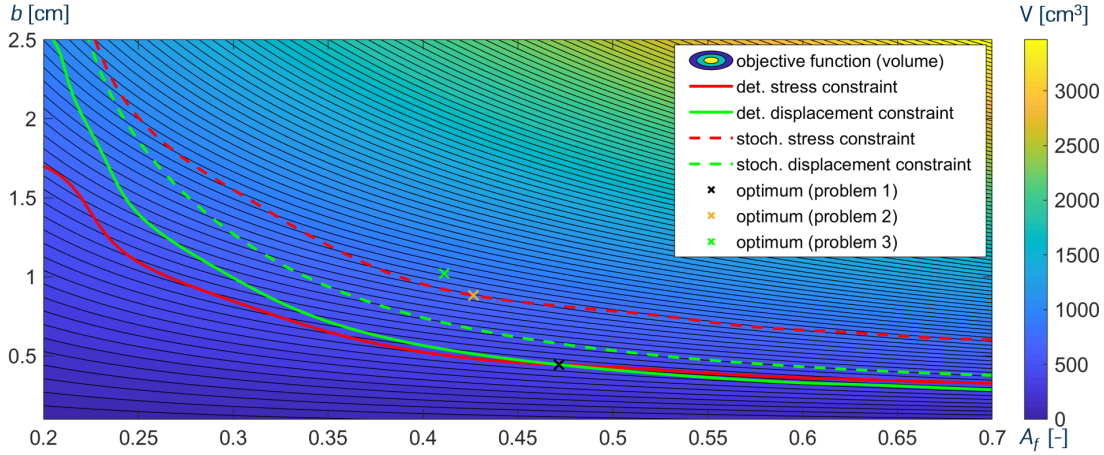


Figure 4. Optimal design variables of the area fraction A_f and the midpoint ${}_mb$ of the interval width of the structure, the three optima are obtained for optimization with interval load position (problem 1), with additional stochastic a priori parameters F, E and f_y (problem 2) and with the additional interval design variable $\overline{b}$ (problem 3).

Table I. Optima of the design variables A_f and ${}_mb$ and corresponding values of the objective function ${}_mV$

	problem 1	problem 2	problem 3
A_f [-]	0.47	0.43	0.41
${}_mb$ [cm]	0.44	0.88	1.02
${}_mV$ [cm^3]	419	753	839

An additional comparison of the results is done by removing the width of the structure (b or $\overline{b}$) from the design variables. This means that in this case, only the area fraction A_f and the element material density $\rho_e(A_f)$ are used as deterministic design variables. The white line in Figure 6 shows

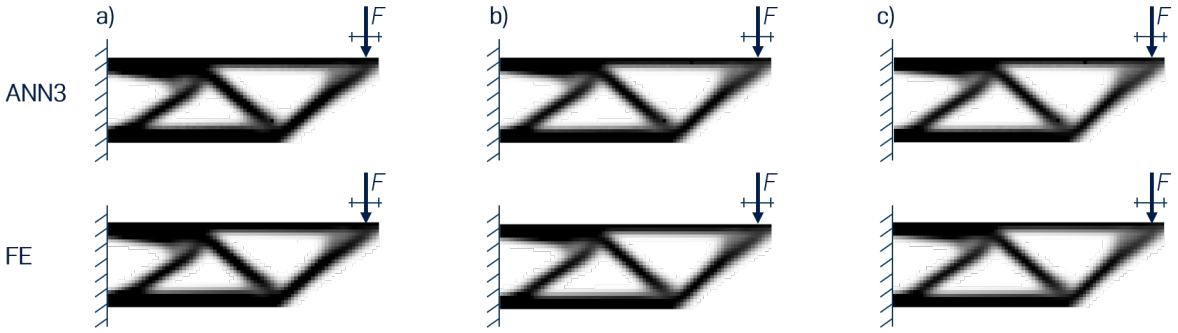


Figure 5. Optimal topologies predicted by an artificial neural network (ANN3) and comparison with finite element simulation (FE); a) interval load position (problem 1); b) with additional stochastic a priori parameters F, E and f_y (problem 2); c) with additional interval design variable $\overline{b}$ (problem 3).

this special case. It can now be seen, that, as expected, the area fraction A_f and consequently also the midpoint ${}_mV$ of the required material volume is increasing with increasing uncertainty, see also Table II and Figure 7.

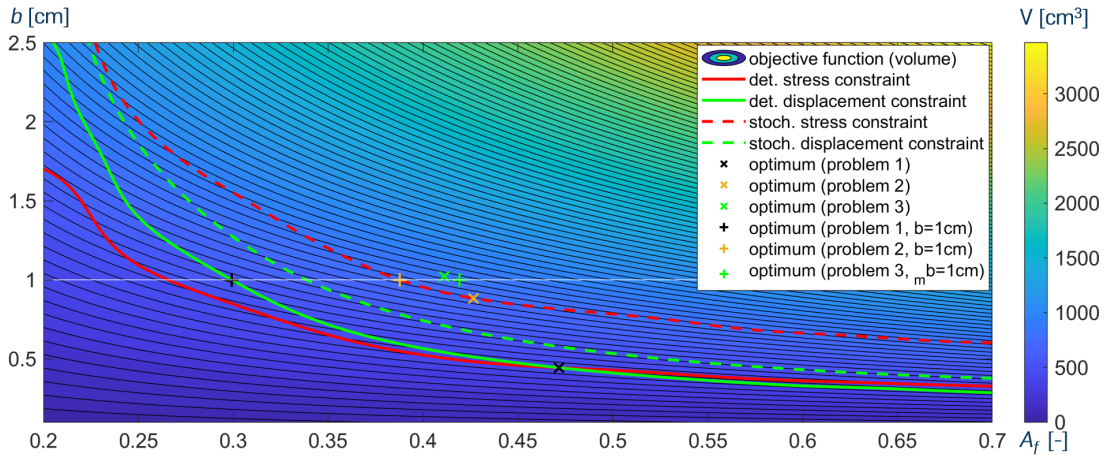


Figure 6. In addition to the results in Figure 4, the optimal area fraction A_f is evaluated for a constant structural width of $b = 1$ cm (for problem 1 and problem 2) and for $\overline{b} = [0.9, 1.1]$ cm (for problem 3), three additional results at the white line for optimization with interval load position only (problem 1), with additional stochastic a priori parameters F, E and f_y (problem 2) and with additional interval a priori parameter $\overline{b}$ (problem 3).

In the presented problems 2 and 3, intervals and random variables are used as uncertain parameters and the resulting maximal von Mises stresses and displacements are obtained as free p-boxes. This means, that the combination of intervals and random variables in the input space leads to polymorphic uncertain results and finally imprecise probabilities, which are used to evaluate the

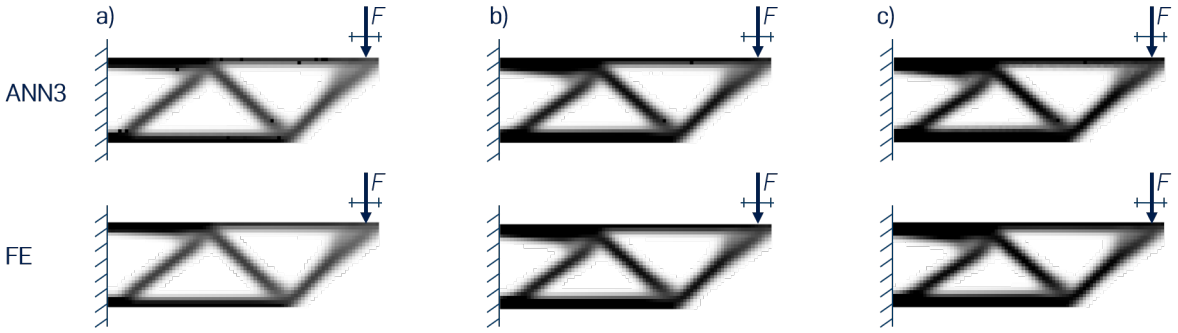


Figure 7. Optimal topologies predicted by an artificial neural network (ANN3) and comparison with finite element simulation (FE); a) interval load position; b) with additional stochastic a priori parameters F, E and f_y; c) with additional interval a priori parameter $\overline{b}$

Table II. Optima of the design variable A_f and the corresponding value of the objective function ${}_mV$ for a constant structural width $b = 1$ cm (for problem 1 and problem 2) and $\overline{b} = [0.9, 1.1]$ cm (for problem 3)

	problem 1	problem 2	problem 3
A_f [-]	0.30	0.39	0.42
${}_mV$ [cm^3]	598	775	838.5

failure probability constraints. However, the presented approach also works, if polymorphic uncertainty is already considered in the input space of the optimization problem. This is demonstrated by an additional problem 4, where the random variables of problems 2 and 3 are extended to parametric p-boxes quantified as normal and lognormal distributed random variables with interval mean values:

– **Problem 4**: optimization with interval load position, polymorphic uncertain a priori parameters defined as parametric p-boxes and interval design variable
  - design variables:
    * A_f (deterministic)
    * $\rho_e(A_f)$ (deterministic)
    * $\overline{b}$ (interval, with midpoint ${}_mb$ to be optimized and fixed radius ${}_rb = 1$ mm)
  - uncertain a priori parameters:
    * $\overline{x}_F = [72, 80]$ cm (interval)

  - $\overline{F}$ (normal distributed parametric p-box with mean value $\overline{\mu}(F) = [9.5, 10.5]$ kN and standard deviation $\sigma(F) = 2$ kN)
  - $\overline{E}$ (lognormal distributed parametric p-box with mean value $\overline{\mu}(E) = [20750, 21250]$ kN/cm^2 and standard deviation $\sigma(E) = 1050$ kN/cm^2)
  - $\overline{f}_y$ (lognormal distributed parametric p-box with mean value $\overline{\mu}(f_y) = [23.2, 23.8]$ kN/cm^2 and standard deviation $\sigma(f_y) = 1.175$ kN/cm^2)
- interval-stochastic constraints:
  - $g_1 = {}_uP_{f,lbc} - P_{f,lbc,ac} \leq 0$ (with ${}_uP_{f,lbc} = P({}_u\sigma_V \leq f_y)$ and $P_{f,lbc,ac} = 7 \cdot 10^{-5}$)
  - $g_2 = {}_uP_{f,ser} - P_{f,ser,ac} \leq 0$ (with ${}_uP_{f,ser} = P({}_uw \leq 0.3$ cm) and $P_{f,ser,ac} = 6.7 \cdot 10^{-2}$)

The results of problem 4 are shown in Table III and Figure 8. Compared to the results of problems 1 to 3, it can be seen, that the midpoint ${}_mV$ of the required material volume is further increasing because of the additional interval uncertainty of the stochastic distribution parameters.

Table III. Optima of the design variables A_f and ${}_mb$ and the corresponding value of the objective function ${}_mV$ of problem 4

	optimization with two design variables	optimization with one design variable and constant width $\overline{b} = [0.9, 1.1]$ cm
A_f [-]	0.42	0.44
${}_mb$ [cm]	1.05	1.0
${}_mV$ [cm^3]	872	874

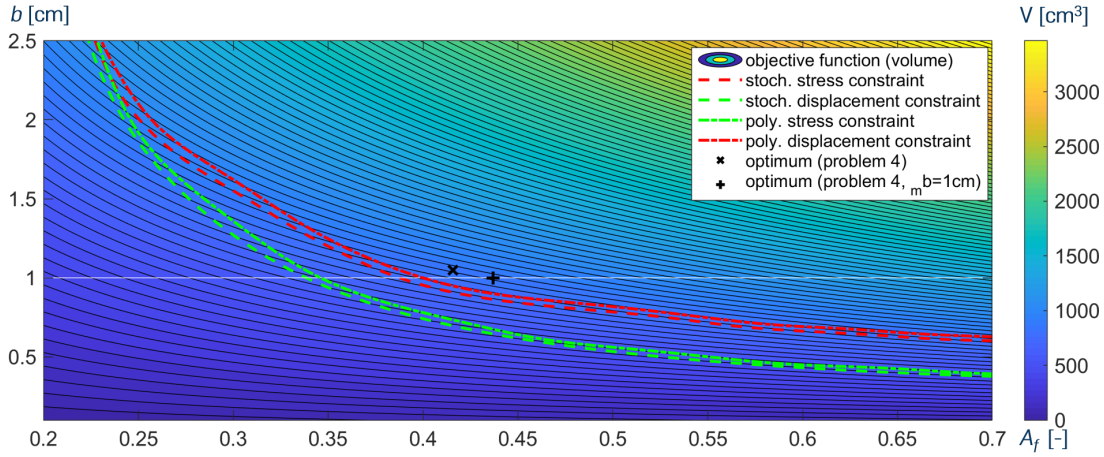


Figure 8. Optimal design variable of the area fraction A_f and the midpoint ${}_mb$ of the interval width of the structure for problem 4, the result at the white line shows the optimal area fraction A_f evaluated for a constant structural width of $\overline{b} = [0.9, 1.1]$ cm.

6. Conclusion

In the paper, a reliability-based topology optimization approach has been presented, where intervals, random variables and p-boxes are considered by means of a polymorphic uncertainty quantification. The finite element based topology optimization has been embedded into a structural optimization approach in order to consider probabilistic constraints (accepted failure probabilities) with respect to the structural load bearing capacity and the serviceability. The results of a cantilever structure optimization have shown, that as expected for increasing uncertainty, more material volume is required. But surprisingly, only the width of the structure is increased and the area fraction is reduced with increasing uncertainty. This can partially be explained by the sensitivities used within the compliance minimization, but needs further investigation in future works.

In order to reduce the numerical effort, several artificial neural networks have been trained to approximate the constraints evaluation as well as the optimal topology computation. Whereas for the constraint approximations small ANN architectures are sufficient, a deep ANN with 2000 output neurons has been created to predict the material density of each finite element in the design domain for a specific area (volume) fraction. Also if the example presented in this paper is an academic one, the potential of deep ANNs for topology prediction has been demonstrated. In future works, the concept of ANN-based topology prediction can be extended to 3D problems and to speed up the finite element simulation also during the topology optimization process, i.e. not only to predict the final optimal topology for a given volume fraction but also the sequential compliance minimization process. It is also planned to perform sensitivity analyses of the uncertain input parameters and to systematically investigate the consequence of FE discretization and ANN training errors, as well as minimization residuals of the optimization approach to the results.

Acknowledgements

This research is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project number: 312921814 within Subproject 6 of the Priority Programme SPP 1886 "Polymorphic uncertainty modelling for the numerical design of structures".

References

Allaire, G., P. Geoffroy-Donders, and O. Pantz. Topology optimization of modulated and oriented periodic microstructures by the homogenization method. *Computers & Mathematics with Applications*, 7:2197–2229, 2019.

Andreassen, E., A. Clausen, M. Schevenels, B. S. Lazarow, and O. Sigmund. Efficient topology optimization in MATLAB using 88 lines of code. *Structural and Multidisciplinary Optimization*, 43:1–16, 2011.

Andrew, R. M. Global CO_2 emissions from cement production. *Earth System Science Data*, 10:195–217, 2018.

Baandrup, M., O. Sigmund, H. Polk, and N. Aage. Closing the gap towards super-long suspension bridges using computational morphogenesis. *Nature Communications*, 11:2735, 2020.

Beghini, L. L., A. Beghini, N. Katz, W. F. Baker, and G. H. Paulino. Connecting architecture and engineering through structural topology optimization. *Engineering Structures*, 59:716–726, 2014.

Bendsøe, M. P. and O. Sigmund. *Topology Optimization*. Springer, Berlin Heidelberg New York, 2004.

Bendsøe, M. P. and N. Kikuchi. Generating optimal topologies in structural design using a homogenization method. *Computer Methods in Applied Mechanics and Engineering*, 71:197–224, 1988.

Bendsøe, M. P.. Optimal shape design as a material distribution problem. *Structural Optimization*, 1:193–202, 1989.

Cavazzuti, M., A. Baldini, E. Bertocchi, D. Costi, E. Torricelli, and P. Moruzzi. High performance automotive chassis design: a topology optimization based approach. *Structural and Multidisciplinary Optimization*, 44:45–56, 2011.

da Silva, G. A., E. L. Cardoso, and A. T. Beck. Non-probabilistic robust continuum topology optimization with stress constraints. *Structural and Multidisciplinary Optimization*, 59:1181–1197, 2019.

da Silva, G. A., E. L. Cardoso, and A. T. Beck. Comparison of robust, reliability-based and non-probabilistic topology optimization under uncertain loads and stress constraints. *Probabilistic Engineering Mechanics*, 59:103039, 2020.

Edler, P., S. Freitag, K. Kremer, and G. Meschke. Optimization Approaches for the Numerical Design of Structures under Consideration of Polymorphic Uncertain Data. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems Part B: Mechanical Engineering*, 5(4):041013 (12 pages), 2019.

Ferson, S., V. Kreinovich, L. Ginzburg, D. S. Myers and K. Sentz. *Constructing probability boxes and Dempster-Shafer structures*. Tech. Rep. SAND2002-4015, Sandia National Laboratories, 2003.

Freitag, S., P. Edler, K. Kremer, and G. Meschke. Multilevel surrogate modeling approach for optimization problems with polymorphic uncertain parameters. *International Journal of Approximate Reasoning*, 119:81–91, 2020.

Kennedy, J. and R. C. Eberhart. Particle Swarm Optimization. In *Proceedings of the IEEE International Conference on Neural Networks IV*, pp. 1942–1948, IEEE Press, Piscataway, 1995.

Lazarov, B. S., M. Schevenels, and O. Sigmund. Topology optimization with geometric uncertainties by perturbation techniques. *International Journal for Numerical Methods in Engineering*, 90:1321–1336, 2012.

Lin, B. and X. Li. The effect of carbon tax on per capita CO_2 emissions. *Energy Policy*, 39:5137–5146, 2011.

López, C., A. Baldomir, and S. Hernández. Deterministic versus reliability-based topology optimization of aeronautical structures. *Structural and Multidisciplinary Optimization*, 53:907–921, 2016.

Mäck, M., I. Caylak, P. Edler, S. Freitag, M. Hanss, R. Mahnken, G. Meschke, and E. Penner. Optimization with constraints considering polymorphic uncertainties. *GAMM-Mitteilungen (Surveys for Applied Mathematics and Mechanics)*, 42(1):e201900005, 1–12, 2019.

Naboni, R. and I. Paoletti. Architectural morphogenesis through topology optimization. In D. D'Uva, editor, *Handbook of Research on Form and Morphogenesis in Modern Architectural Contexts*, Chapter 4, pages 69–92. IGI Global, 2018.

París, J., S. Martínez, F. Navarrina, and M. Casteleiro. Topology optimization of aeronautical structures with stress constraints: general methodology and applications. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 226(5):589–600, 2012.

Schöbi, R. and B. Sudret. Structural reliability analysis for p-boxes using multi-level meta-models. *Probablistic Engineering Mechanics*, 48:27–38, 2017.

Schuëller, G. I. and H. A. Jensen. Computational methods in optimization considering uncertainties – An overview. *Computer Methods in Applied Mechanics and Engineering*, 198(1):2–13, 2008.

Sigmund, O. A 99 line topology optimization code written in Matlab. *Structural and Multidisciplinary Optimization*, 21:120–127, 2001.

Sigmund, O. and K. Maute. Topology optimization approaches. *Structural and Multidisciplinary Optimization*, 48(6):1031–1055, 2013.

Thore, C. J., E. Holmberg, and A. Klarbring. A general framework for robust topology optimization under load-uncertainty including stress constraints. *Computer Methods in Applied Mechanics and Engineering*, 319:1–18, 2017.

Valdebenito, M. A. and G. I. Schuëller. A survey on approaches for reliability-based optimization. *Structural and Multidisciplinary Optimization*, 42(5):645–663, 2010.

Wang, D. and W. Gao. Robust topology optimization under multiple independent uncertainties of loading positions. *International Journal for Numerical Methods in Engineering*, 121:4944–4970, 2020.

An Integrated Interval Neural Network for Uncertainty Modeling in Inhomogeneous Materials

Shahrokh Shahi
School of Computational Science and Engineering Engineering, Georgia Institute of Technology, Atlanta, GA 30332, US, shahi@gatech.edu

Rafi L. Muhanna
School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332, US, rafi.muhanna@gatech.edu

Francesco Fedele
School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332, US, fedele@gatech.edu

Abstract. Engineering fields rely heavily on the Finite Element Method (FEM) as a modeling tool in deterministic systems where no uncertainty is introduced. The effects of uncertainty are of growing concern in the analysis and design of engineering structures and need to be studied to improve the predictability of mathematical models. Recently, in addition to others, Interval Finite Element Method (IFEM) has been introduced to account for uncertainties by incorporating interval arithmetic into the conventional FEM formulation, in which all uncertain parameters are defined as intervals. Nonetheless, in combination with complexity of structures and inhomogeneous materials, the computational and experimental cost remains an inevitable issue in such simulations.

This work aims at integrating Artificial Neural Networks (ANN) and IFEM techniques to establish a flexible and efficient approach for modeling uncertainties in general inhomogeneous structures, in which an Interval Neural Network (INN) is employed as a substitution for the conventional constitutive material model to establish a homogenized representation of structures regardless of material complexity. In this approach, at first, the required dataset is generated by creating and running a set of IFEM simulations. The INN will then be trained to predict the homogenized mechanical behavior of the structure as a function of independent parameters. Afterwards, the trained INN will be integrated in the IFEM procedure to obtain the system's response under uncertainty. The proposed approach is applied to a set of engineering problems to illustrate and verify the capabilities of the methodology.

Keywords: Interval Finite Element Method, Interval Neural Network, Machine Learning, Inhomogeneous Material

1. Introduction

The effects of uncertainty are of growing concern in the analysis and design of engineering structures. In general, uncertainties play an important role in the modeling of various engineering and

© 2021 *by S. Shahi, R.L. Muhanna, and F. Fedele.*

science problems and need to be studied for improving the predictability of mathematical models. Among various proposed approaches for modeling uncertainty, interval-valued data are a natural way to represent uncertainty in engineering systems, e.g. uncertainty in measurements that contain valuable information useful for decision making.

Interval Finite Element Method (IFEM) is an attempt to account for uncertainty in structural analysis by incorporating interval arithmetic into conventional finite element formulation (Muhanna and Mullen, 2001). Within the context of IFEM, all uncertain parameters, such as applied loads, material properties and geometry, are defined as intervals (Rao et al., 2011; Xiao et al., 2015; Muhanna and Shahi, 2020). The main challenge in IFEM is reducing the overestimation due to interval dependency problem which is handled by decomposing matrices into interval and deterministic components, applying Element-By-Element assembly approach and employing iterative enclosure method to solve the system of equations (Muhanna and Mullen, 2001; Rao et al., 2011; Xiao, 2015).

One important component which is an integral part in computational mechanics analysis including FEM and IFEM is constitutive material model. In general, constitutive models describe the material responses to various mechanical conditions and represent the stress-strain relations (Bower, 2009). Such descriptions are mainly presented as mathematical models which are a simplification of complex mechanical behaviors studied through experiments. The level of simplification depends on the purpose and the required precision of the model predictions. For instance, the mechanical behavior of steel can be represented by an elastic-plastic model but to analyze structures under working loads, engineers usually employ linear elasticity model in which Hooke's law represents the material behavior assuming the strain proportional to the applied stress (Bower, 2009; Kenneth et al., 2006).

The conventional process of developing material models mainly includes developing the mathematical formulation of constitutive equations using principles of mathematics and continuum mechanics, and determining the associated parameters to fit experimental measurements (Desai and Siriwardane, 1984). However, establishing an exact closed form constitutive model is not possible for all materials and the requirement for a sufficiently accurate model can lead to obtaining complex and computationally expensive models. Especially, with the introduction and application of composite and inhomogeneous materials in general, providing constitutive models describing complex mechanical behaviors becomes more challenging and there will be an ever-increasing demand for more realistic models (Buyukozturk and Shareef, 1985; Ghaboussi et al., 1999).

Artificial neural networks (ANN) are computational tools inspired by human perception of the biological structure of neurons and the internal operation of brain. A neural network is a nonlinear system usually consisting of a large number of highly interconnected processing units known as artificial neurons and can provide models representing highly nonlinear behaviors. Feedforward supervised neural networks are known as very first and successful learning algorithms (Rumelhart et al., 1985). The universal approximation theorem (Cybenko, 1989; Hornik et al., 1989) guarantees that with even a single hidden layer containing a finite number of neurons, this type of networks can represent arbitrary complex but smooth approximation of any arbitrary input-output mapping (White, 1990; Goodfellow et al., 2016). Such great flexibility in learning nonlinear relationships between input and output parameters makes ANN a robust alternative to mathematical constitutive models, in particular where the complexity of material response to applied mechanical stimuli cannot be easily and accurately described by mathematical models.

Recently, a variety of ANN constitutive models have been proposed to represent complex mechanical behaviors of various materials (Ghaboussi et al., 1999; Sen et al., 2002; Jung and Ghaboussi, 2006) and have also been successfully integrated in FEM procedure (Javadi et al., 2003; Gulikers, 2018). To the best of our knowledge, all the proposed models are deterministic in which the effects of uncertainty have not been taken into account and the input and output of the associated neural networks are crisp (i.e. single-valued) data (Ghaboussi et al., 1999; Sen et al., 2002; Jung and Ghaboussi, 2006).

In general, an ANN is considered INN if at least one of its input, output, or assigned weight sets are intervals (Beheshti et al., 1998). Accordingly, a variety of INNs are proposed. One of the very first version of INNs presented by Ishibuchi et al. in which except the input data, all other numbers are interval-valued and applied for fuzzy regression analysis (Ishibuchi et al., 1993). The universal approximation theorem is also proven for this type of INN (Baker and Patil, 1998). Repeating interval operations in INN with interval-valued inputs, outputs, and weights cause the explosion of interval uncertainty in these networks (Simoff, 1996). One straight-forward method proposed by Rossi and Conan-Guez is the extremal approach in which each interval-valued input is transformed into a pair of real numbers, for instance the corresponding lower bound and upper bound, and the INN is built on top of a standard Multi-layer Perceptron (MLP) (Rossi and Conan-Guez, 2002). Another approach to handle the interval-valued data is interval Multi-layer Perceptrone (iMLP) proposed by San Roque et al. in which inputs and outputs are interval-valued data, but the weights and biases are single-valued data. In this method, the input and output of each neuron are a pair of center-range values and each center and range share the same weight (San Roque et al., 2007). Lately, a regularized artificial neural network (RANN) is proposed by Yang et al. in which a non-crossing regularizer is introduced to control the interval crossing problem (Yang et al., 2019). In this paper, a variant of extremal approach is employed to construct an INN constitutive material model.

This work is an effort to take advantage of interval neural networks (INN) trained by interval-valued data to model the uncertainty, for instance, in experimental measurements of applied stress and strain. Then, the trained INN is employed as a substitution for the conventional constitutive model (mathematical model) to represent a homogenized material model incorporated into IFEM procedure, and establish an integrated approach that accounts for uncertainty in analyzing structures with particularly inhomogeneous materials.

The paper is structured as follows. First, the key features of the extremal approach are briefly presented to handle the interval-valued inputs and outputs data. The INN constitutive model and the integrated framework are explained in the next sections providing the details regarding the substitution of conventional constitutive model with the trained INN material model. Finally, numerical examples are presented and discussed to verify and demonstrate the application of proposed approach.

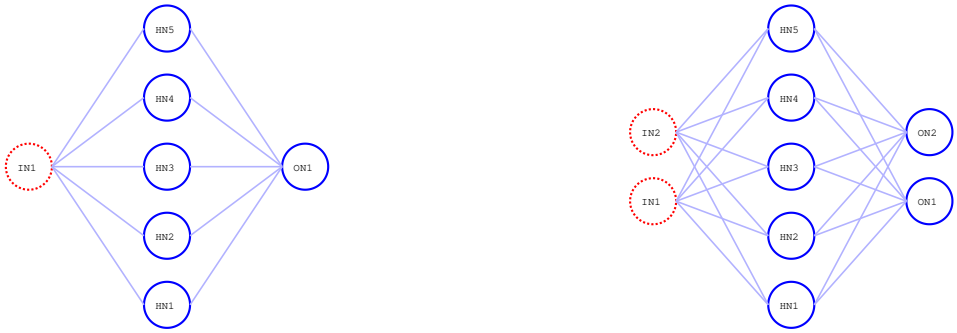


Figure 1. Shallow artificial neural networks. Single input single output network (left). Modified architecture to handle the interval-valued data in extremal approach (right).

2. Formulation

2.1. Extremal Approach

This approach is most probably the simplest approach to deal with interval-valued input and can be easily built on top of regular MLP. Each interval-valued data is defined by a pair of numbers: a lower bound and an upper bound, or a center value and a radius. A natural way for handling interval inputs is treating them as pair of inputs. In this way, instead of training the network with n interval-valued input and output data, $\boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_n}$ and $\boldsymbol{y_1}, \boldsymbol{y_2}, \ldots, \boldsymbol{y_n}$, where $\boldsymbol{x_i} = [\underline{x_i}, \overline{x_i}]$ and $\boldsymbol{y_i} = [\underline{y_i}, \overline{y_i}]$, respectively, $2n$ single-valued numbers $\underline{x_1}, \overline{x_1}, \underline{x_2}, \overline{x_2}, \ldots, \underline{x_n}, \overline{x_n}$ and $\underline{y_1}, \overline{y_1}, \underline{y_2}, \overline{y_2}, \ldots, \underline{y_n}, \overline{y_n}$ are used to train the neural network. Therefore, a single input single output neural network needs to be transformed to a neural network with two input and output neurons to be trained with a set of augmented data (Figure 1). In this work, each interval is considered as a pair of center and radius values denoted by $\boldsymbol{x_i} =< x_i^c, x_i^r >$. Thus, the augmented inputs are as follows:

$$x_1^c, x_2^c, \ \ldots, x_n^c, \tag{1}$$
$$x_1^r, x_2^r, \ \ldots, x_n^r \tag{2}$$

Then, the neural network is trained as a regular MLP with two input and output neurons (Figure 1 (right)).

2.2. INN Constitutive Model

Conventionally, to construct a material model, principles of mathematics and continuum mechanics suggest the form of mathematical equations. The associated parameters are then determined to fit the experimental stress and strain measurements obtained by standard mechanical tests, such as tensile test in which a tensile force is applied to a standard specimen of a material (homogeneous

or inhomogeneous) and the corresponding displacement or strain response is measured (Kenneth et al., 2006). However, these measurements are not deterministic and the uncertainty in these data can naturally be presented as intervals. Such interval-valued data can be used to train an INN where the interval-valued stress is the input data, and the interval-valued module of elasticity is the output of the network.

In this research, the extremal approach explained in Section 2.1 is employed to construct such an INN in which pairs of center and radius values of interval-valued stresses is fed into the input layer and the associated pairs of center and radius values of module of elasticity are used as target values to train the INN. A regular backpropagation algorithm with Bayesian regularization is employed to train the INN and obtain the weights and biases. Such an INN represents the relationship between interval stress and strain and thus can effectively serve as the constitutive material model integrated into an IFEM analysis. Therefore, in this paper, such material descriptions are referred to as INN constitutive models, as opposed to the mathematical constitutive models which are known as conventional material models.

Algorithm 1 Integrated Framework

Input: Problem definition (including geometry, loading , boundary conditions), the trained INN constitutive model
Output: System response (nodal displacements, internal forces, reaction forces, stresses, and strains)

for ($step = 1$ **to** $load_steps$) **do**
  Generating all possible endpoints combinations for loads and material properties
  for ($i \in load_combinations$) **do**
    Forming the global vector of equivalent nodal load for the current load step
    for ($j \in material_properties_combinations$ (`obtained by INN predictions`)) **do**
      for ($k = 1$ **to** $number_of_elements$) **do**
        Forming the element i stiffness matrix
      end for
      Assembling the global stiffness matrix
      Imposing boundary conditions
      Solving the system of equations and obtaining primary unknowns (nodal displacements)
      Calculating secondary variables (stress and strains)
      Calculating principal stresses
      Storing the calculated values for the current combination of loads and material properties
    end for
  end for
  Report the minimum and maximum of the calculated values as the lower bound and the upper bound of the system response in the current load step
  Predicting the interval-valued material properties for the next load step by inputting the vector of interval-valued principal stresses into the trained interval neural network
end for

return The interval-valued results of the system response for the last load step

2.3. Integrated Framework

The proposed INN constitutive model captures the mechanical behavior of material and provides a description of material properties only based on the realistic experimental data without any prior knowledge about the material. Therefore, the integration of such model is straight forward, and it can be directly employed as conventional material models. in this paper, to obtain the exact enclosure, the proposed INN constitutive model is integrated in an endpoint combinations analysis in which the structure will be deterministically analyzed for all possible combinations of the lower bounds and upper bounds of interval parameters, e.g., applied loads and material properties. This method is basically applicable to the structures with a few interval parameters, since the number of combinations increases exponentially as the number of interval parameters increases. However, it is straight-forward to be implemented and a reliable method to validate the proposed approach. Algorithm 1 demonstrates the integrated procedure in which a trained INN-based description of material properties is directly used to form element stiffness matrices in the FEM endpoint combination analysis. The same effort can be made to integrate the INN constitutive model into IFEM procedure.

3. Example Problems

The proposed approach is implemented in a MATLAB program in which the INN is built on top of the MATLAB neural network toolbox. In this section, three numerical examples are presented to illustrate the application of the proposed framework. The first example presents a plane stress analysis of a cantilever beam with a linear elastic homogeneous material under uncertainty. Then, the second example illustrates the same cantilever beam built of an inhomogeneous material. Finally, the application of the proposed framework is further investigated by analyzing a plate built of the same inhomogeneous material in the third example. In this paper, to train the INN constitutive model, a set of synthetic measurements is generated by simulating a uniaxial tensile test of steel material, where the dimensions are chosen due to the standard size of specimens in uniaxial tension tests (Kenneth et al., 2006).

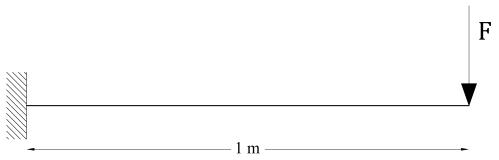


Figure 2. A cantilever beam subjected to a load at the free end.

3.1. Example 1: Plane Stress Analysis of a Cantilever Beam

The cantilever beam shown in Figure 2 is subjected to a vertical surface load $F = 100$ KN/m^2 at its free end. The length and the cross-sectional area are $L = 1$ m and $A = 0.01 \times 0.01$ m^2,

respectively. The synthetic training dataset is generated by simulating a uniaxial tensile test of steel material (Figure 3), where the module of elasticity is set to 200×10^6 KN/m^2. To generate reasonable interval-valued data for stress and strain, an interval tensile load $\boldsymbol{P}$ is applied due to 1% uncertainty. The dataset is created by applying various interval tensile load $\boldsymbol{P}$ and obtaining the strain response of the specimen in an endpoint combination analysis in which the lower and upper bounds of the interval-valued data are determined by minimum and maximum values of the mechanical responses, respectively. Accordingly, a dataset including 400 datapoints is obtained, where 80 percent of them are randomly assigned to the train set and the rest 20 percent form the test dataset which is used to validate the trained INN.

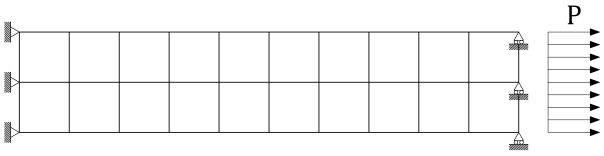


Figure 3. Tensile test simulation for generating synthetic data.

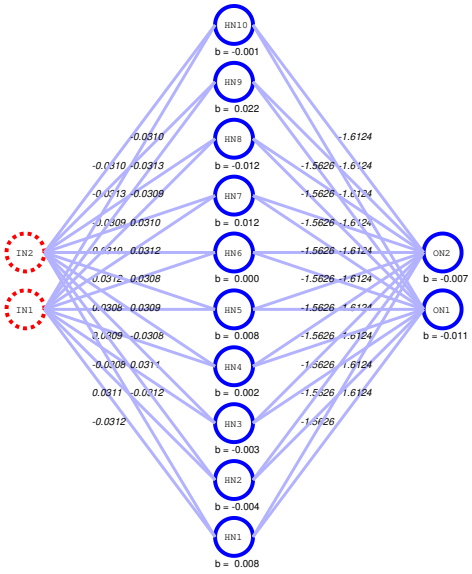


Figure 4. The trained interval neural network with ten neurons on the hidden layer.

The interval-valued stress-strain data is used to train a shallow feedforward INN with ten neurons on the hidden layer by employing the extremal approach explained in section 2.1. The inputs and outputs of this INN are pairs of mid points and radii of the interval-valued stress and strain data, respectively. Figure 4 illustrates the trained network using the synthetic data obtained by simulating a uniaxial tensile test on the homogeneous steel material.

Figure 5 demonstrates the linear relationship between the midpoints of stress and strain interval-valued data. The stress-strain relationship predicted by the trained INN shows that the neural network has successfully captured the material behavior. This INN is employed as a constitutive material model in endpoint combination analysis of the cantilever beam (Figure 2) which is modeled by ten second order quadrilateral elements illustrated in Figure 6.

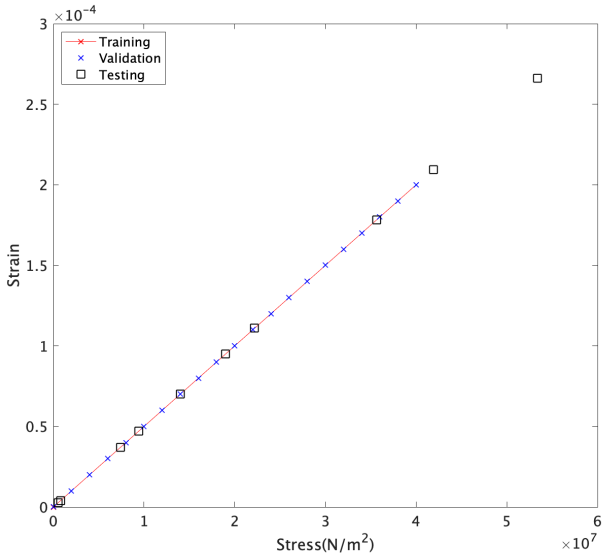


Figure 5. The midpoints of training, validation and test interval-valued data used for building the INN constitutive model.

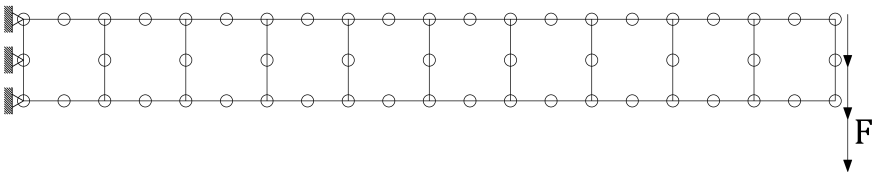


Figure 6. Modeling the cantilever beam with ten second-order quadrilateral elements

In this simulation, 1% uncertainty is considered in the applied surface load, and the material properties is predicted by employing the trained INN at each load step. Figure 7 demonstrates the deformed shape of the beam obtained by the midpoint values of interval-valued nodal displacements

of the endpoint combination analysis with (i) INN constitutive model (red lines) and (ii) the conventional constitutive material model introduced by interval-valued module of elasticity (blue shaded elements). The outputs show a perfect match between the results of two analyses.

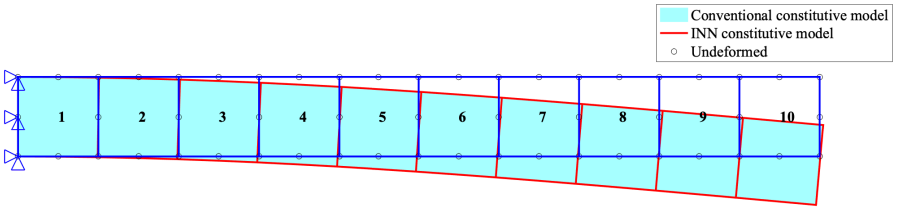


Figure 7. The cantilever beam with homogeneous material - deformed shape.

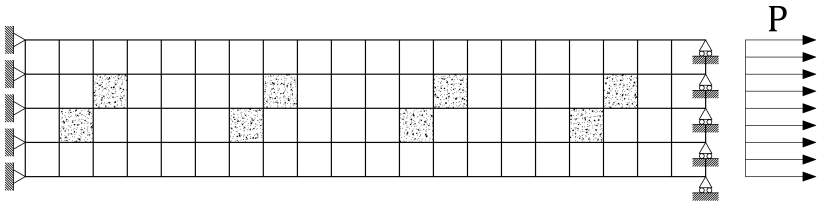


Figure 8. Tensile test simulation of an inhomogeneous material for generating synthetic data.

3.2. Example 2: Plane Stress Analysis of a Cantilever Beam with Inhomogeneous Material

To demonstrate the capability of the proposed approach in analyzing structures built of inhomogeneous materials, the same cantilever beam shown in Figure 2 is analyzed using the INN-based constitutive material model. The same extremal approach is adopted to obtain a homogenized INN constitutive material model of inhomogeneous materials under uncertainty. To train the INN, similar to the first example, a set of 400 datapoints are generated, where 80 percent of them randomly assigned to the train set and the rest 20 percent form the test set. Figure 8 demonstrates the simulation of the standard uniaxial tensile testing of an inhomogeneous material in which different material properties is assigned to some of the elements to model the inhomogenous material structure. Figure 9 illustrates that the stress-strain relationship is successfully captured by the homogenized INN constitutive model. It should be noted that in this example the homogenized material model demonstrates a linear form since the synthetic interval-valued stress and strain data are generated by a linear finite element analysis of a tensile test in which different modulus of elasticity are assigned to the elements. However, in general, the material can have any complicated nonlinear behavior, and due to the universal approximation theorem, the INN can capture such more complex stress-strain relations if enough data is provided for training.

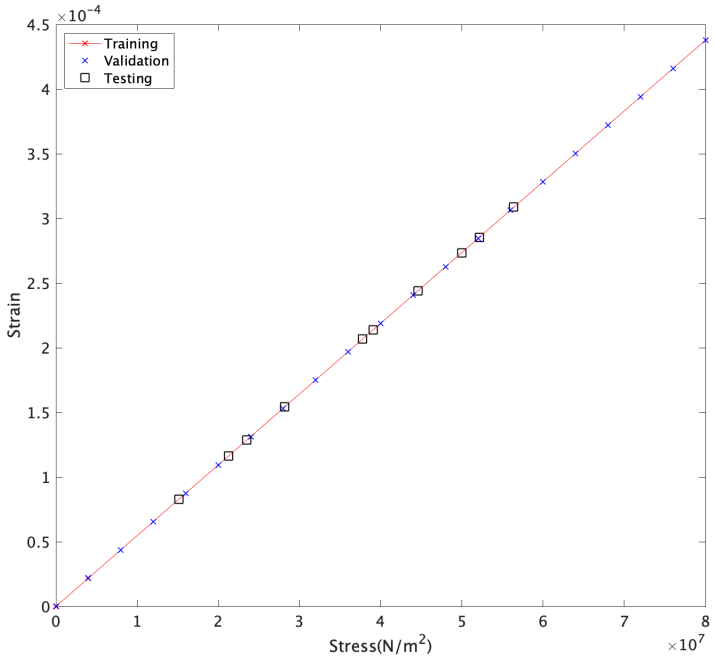


Figure 9. The midpoints of training, validation and test interval-valued data used for building interval neural network constitutive model.

The displacement response of the beam is illustrated in Figure 10. The lower and upper bounds of the horizontal and vertical displacements of eleven selected nodes are demonstrated in Figure 11 which shows the obtained intervals enclose the results of deterministic analysis. For instance, the deterministic numerical value of the vertical displacement at the tip of the cantilever beam (node 11) is -0.0073 mm which is enclosed by the interval $[-0.0075, -0.0072]$ mm obtained by the proposed framework.

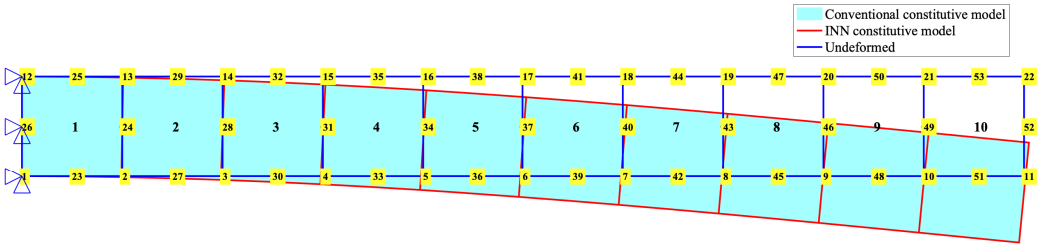


Figure 10. The cantilever beam with inhomogeneous material - deformed shape.

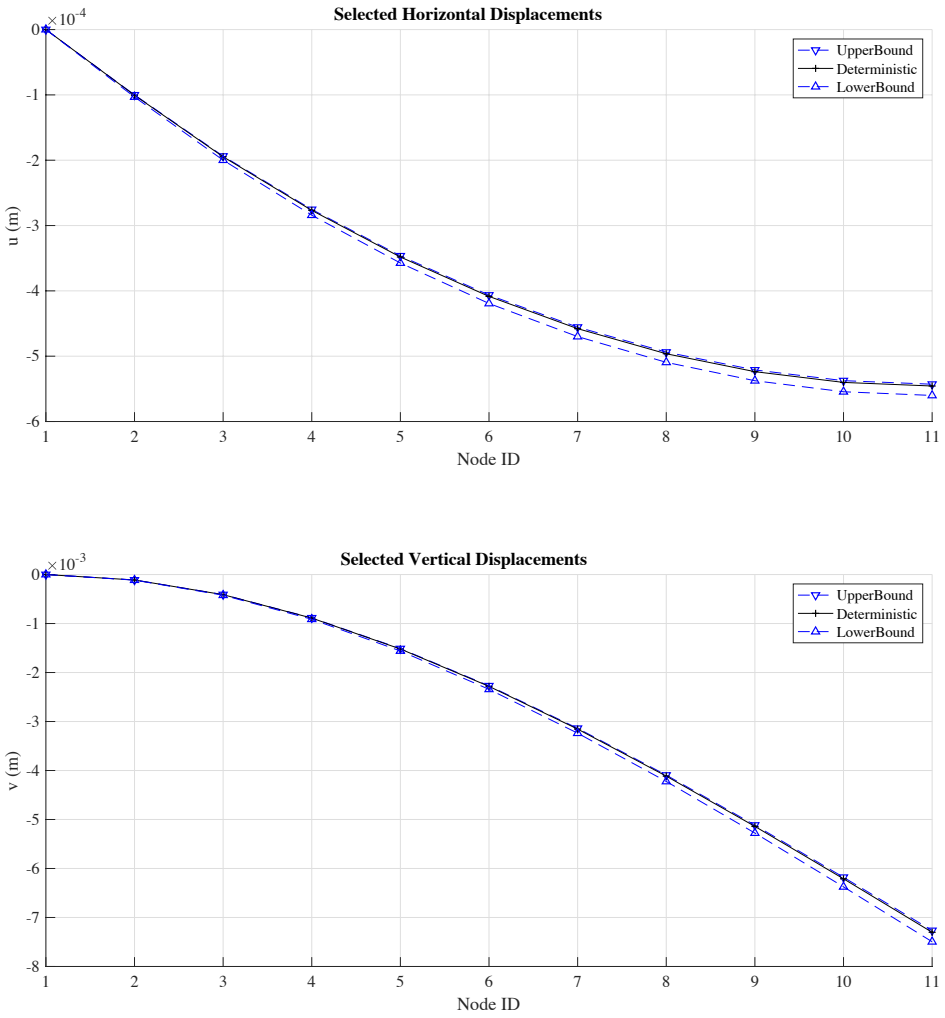


Figure 11. The selected displacement outputs.

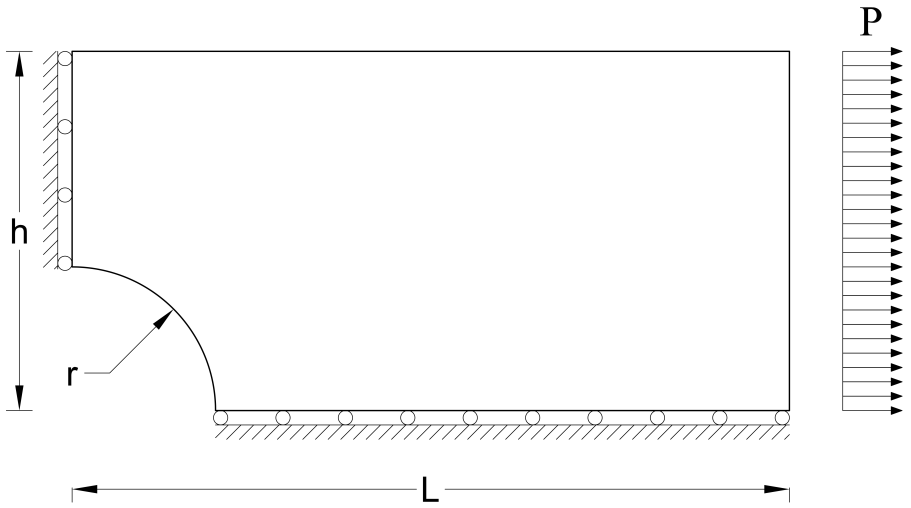


Figure 12. The rectangular plate with circular cutoff.

3.3. Example 3: Plane Stress Analysis of a Rectangular Plate with a Circular Cutoff Subjected to Uniformly Distributed Load

The same approach can be used for analyzing a rectangular plate with a circular cutoff shown in Figure 12. The dimension of this plate is specified by $h = 0.05$ m, $L = 0.1$ m, and $r = 0.02$ m. The thickness of the plate is $t = 0.005$ m and the right edge is subjected to a traction surface load of $P = [19.5, 20.5]$ N/mm^2. It is assumed that this plate is built of the same inhomogeneous material described in Example 2 except the synthetic measurement data is generated for 5% of uncertainty in the applied load. Therefore, it is expected that the interval width increases which is illustrated in Figure 13 where the upper and lower bounds are also added to the graph to demonstrate that the INN can capture the uncertainty in inhomogeneous material behavior. This neural network is trained with the same number and distribution of datapoints in train and test sets.

The deformed shape of the plate using the midpoints of interval displacement values are depicted in Figure 14, where the close match between the two deformed shapes obtained by the conventional and INN-based constitutive models exhibits the accuracy of the proposed approach. For more investigation, the displacements of the selected nodes on the edge of the plate are illustrated in Figure 15 in which the displacement responses obtained by the deterministic analysis when zero uncertainty is introduced in the model is fully enclosed by the bounds of the interval values calculated by the proposed approach.

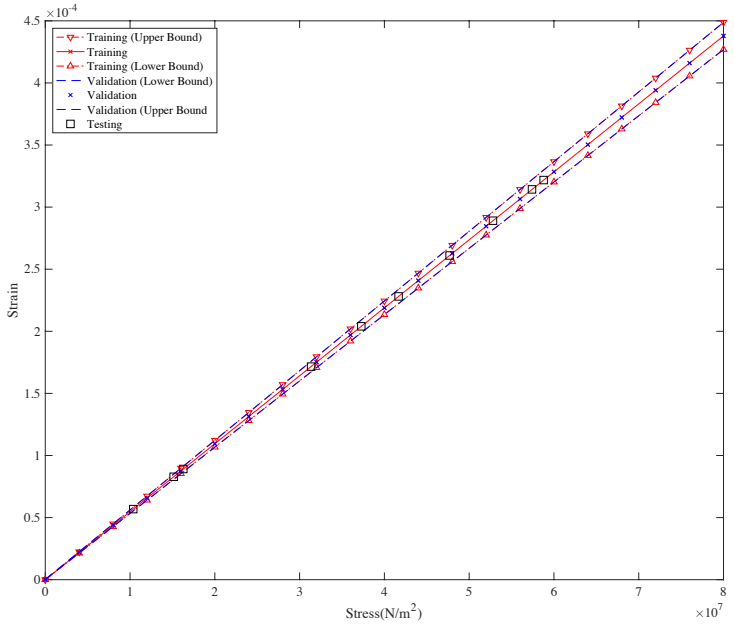


Figure 13. The training, validation, and test interval-valued data used for construction the interval neural network constitutive model. The training data is generated with 5% uncertainty level.

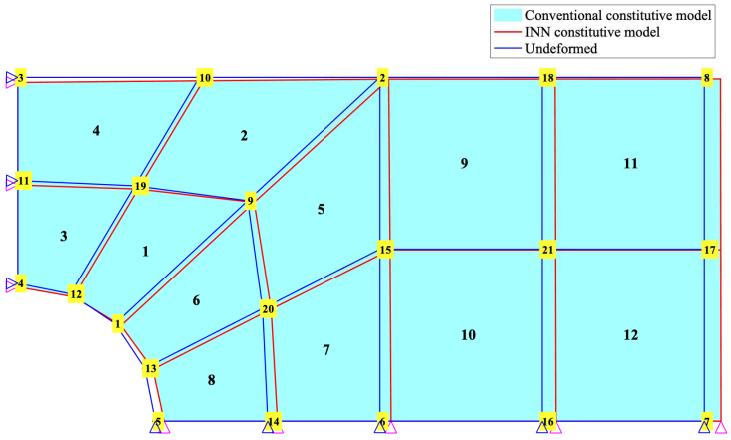


Figure 14. The rectangular plate with inhomogeneous material - deformed shape.

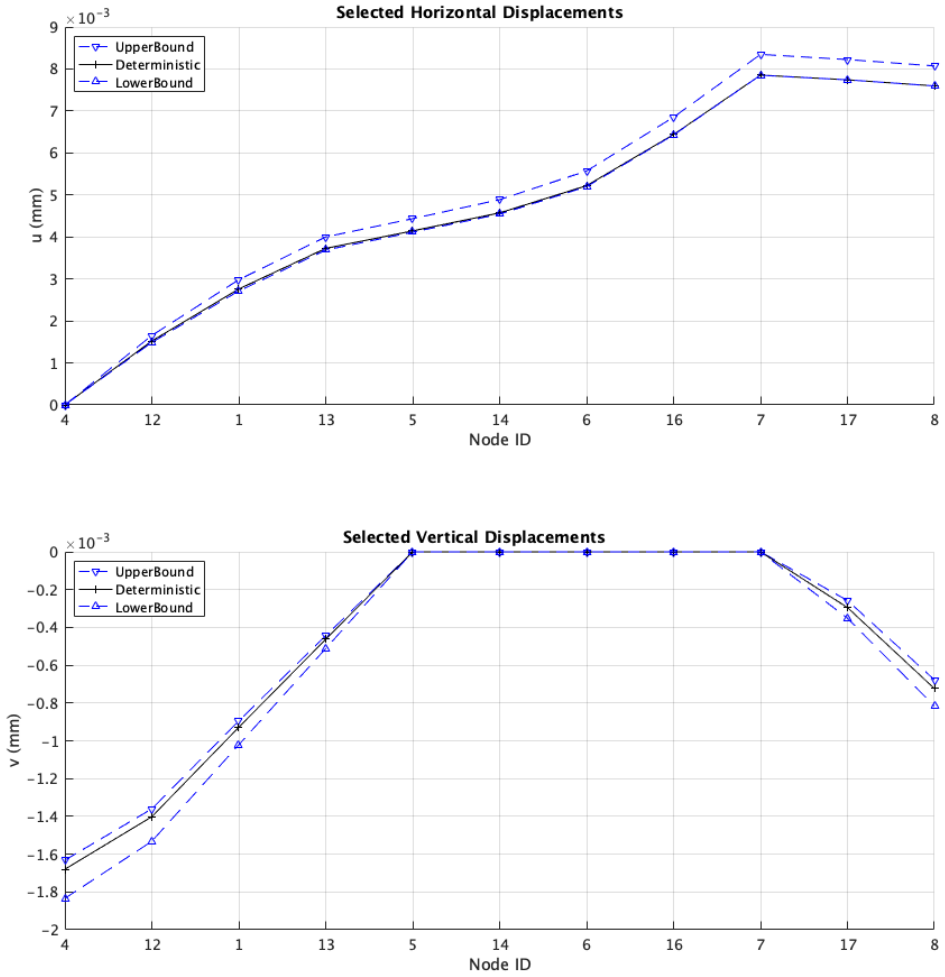


Figure 15. The horizontal and vertical displacements of the selected nodes.

4. Conclusion

This work presents an INN-based constitutive material model and integrates it into mechanical analysis procedure. In general, obtaining a mathematical constitutive material model which can exactly describe the mechanical behavior is not feasible for all materials. Moreover, the uncertainty in the experimental measurements is not taken into account in such conventional material models. The proposed approach provides a flexible substitution for constitutive material models that ac-

counts for uncertainty in measurements which can naturally be presented as interval-valued data. In this method, an extremal approach is used to construct an INN which is trained by interval-valued data obtained from standard mechanical experiments. The trained INN describes the mechanical behavior of the material and is employed as a substitution for conventional material models by predicting the required material properties in IFEM analysis.

The numerical examples demonstrate that the employed INN can perfectly capture the mechanical behavior of a material under uncertainty and this approach can be used as a flexible alternative to constitutive material models in analyzing mechanical structures built of inhomogeneous materials where providing a closed-form constitutive model is not feasible and/or the conventional mathematical models cannot accurately reflect the mechanical behaviors. In this paper, although a set of synthetic interval-valued data is generated by simulating a standard tensile test and used to train the INN, real experimental data can be used with no additional consideration.

Acknowledgements

This work is supported by the National Science Foundation under grant No. 1634483.

References

Baker, M. R. and Patil, R. B. Universal Approximation Theorem for Interval Neural Networks. *Reliable Computing*, 4(3):235–239, 1998.

Beheshti, M., Berrached, A., de Korvin, A., Hu, C., and Sirisaengtaksin, O. On Interval Weighted Three-Layer Neural Networks. *Proceedings of Annual Simulation Symposium*, 188–188, 1998.

Bower, A. F. Applied Mechanics of Solids. *CRC press*, 2009.

Buyukozturk, O. and Shareef, S. S. Constitutive Modeling of Concrete in Finite Element Analysis. *International Journal of Reliability and Safety*, 21(3):581–610, 1985.

Cybenko, G. Approximation by Superpositions of a Sigmoidal Function. *Mathematics of control, signals and systems*, Springer, 2(4):303–314, 1989.

Desai, C. S. and Siriwardane, H. J. Constitutive Laws for Engineering Materials with Emphasis on Geologic Materials. *Prentice-Hall*, 1984.

Ghaboussi J., Wu X., and Kaklauskas G. Neural Network Material Modelling. *Journal of Civil Engineering and Management*, Taylor & Francis, 5(4):250–257, 1999.

Goodfellow, I., Bengio, Y., and Courville, A. Deep Learning. *MIT press*, 2016.

Gulikers, T. *An Integrated Machine Learning and Finite Element Analysis Framework, Applied to Composite Substructures including Damage*. Master thesis, Delft University of Technology, 2018. TU Delft Aerospace Engineering.

Hornik, K., Stinchcombe, M., White, Halbert, and others Multilayer Feedforward Networks are Universal Approximators. *Neural networks*, 2(5):359–366, 1989.

Ishibuchi, H., Tanaka, H., and Okada, H. An Architecture of Neural Networks with Interval Weights and its Application to Fuzzy Regression Analysis. *Fuzzy Sets and Systems*, 57(1):27–39, 1993.

Javadi, A. A., Tan, T.P., and Zhang, M. Neural Network for Constitutive Modelling in Finite Element Analysis. *Central European Association for Computational*, 10(4):523–530, 2003.

Jung, S. and Ghaboussi, J. Neural Network Constitutive Model for Rate-Dependent Materials. *Computers & Structures*, Elsevier, 84(15-16):955–963, 2006.

Muhanna, R. L and Mullen, R. L Uncertainty in Mechanics Problems — Interval–based Approach. *Journal of Engineering Mechanics*, 127(6):557–566, 2001.

Muhanna R. L. and Shahi S. Uncertainty in Boundary Conditions — An Interval Finite Element Approach. *Decision Making under Constraints*, 157–167, 2020.

Rao, MV Rama, Mullen, Robert L and Muhanna, Rafi L. A New Interval Finite Element Formulation with the Same Accuracy in Primary and Derived Variables. *International Journal of Reliability and Safety*, 5(3-4):336–357, 2011.

Rossi, F. and Conan-Guez, B. Multi-layer Perceptron on Interval Data. *Classification, Clustering, and Data Analysis*, 427–434, 2002.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. Learning Internal Representations by Error Propagation. *California Univ San Diego La Jolla Inst for Cognitive Science*, 1985.

Runesson, K., Steinmann, P., Ekh, M. and Menzel, A. Constitutive Modeling of Engineering Materials — Theory and Computation. *Volume I General Concepts and Inelasticity*, 2006.

San Roque, A. M., Maté, C., Arroyo, J., and Sarabia, Á. iMLP: Applying Multi-layer Perceptrons to Interval-valued Data. *Neural Processing Letters*, 25(2):157–169, 2007.

Sen, S., Twomey, J. M., and Ahmad, J. YS. Development of an Artificial Neural Network Constitutive Model for Aluminum 7075 Alloy. *Proceedings of IERC Conference*, 19, 2002.

Simoff, S. J. Handling Uncertainty in Neural Networks: An Interval Approach. *Proceedings of International Conference on Neural Networks (ICNN'96)*, 1:606–610, 1996.

Xiao, N., Muhanna, R. L., Fedele, F., and Mullen, R. L. Uncertainty Analysis of Static Plane Problems by Intervals. *SAE International Journal of Materials and Manufacturing*, 8(2):374–381, 2015.

Xiao, N. *Interval Finite Element Approach for Inverse Problems Under Uncertainty*. PhD thesis, Georgia Institute of Technology, 2015.

White, H. Connectionist Nonparametric Regression: Multilayer Feedforward Networks Can Learn Arbitrary Mappings. *Neural Networks*, 3(5):535–549, 1990.

Yang, Z., Lin, D. KJ, and Zhang, A. Interval-valued Data Prediction via Regularized Artificial Neural Network. *Neurocomputing*, 331:336–345, 2019.

Damage Detection for Structural Health Monitoring Under Uncertainty using Deep Interval Neural Networks

David Betancourt[1,3)], Steffen Freitag[2)] and Rafi L. Muhanna[3)]

[1)] *School of Computational Science and Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA, david.betancourt@gatech.edu*

[2)] *Institute for Structural Mechanics, Ruhr University Bochum, Germany, steffen.freitag@sd.rub.de*

[3)] *School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA, rafi.muhanna@gatech.edu*

Abstract. Most approaches for damage identification tasks in Structural Health Monitoring (SHM) are solid-mechanics based and are generally not designed for real-time damage identification. Such approaches require the comparison of responses from a finite element (FE) model analysis with post-processed sensor data in order to assess damage. On the other hand, more recent data-driven approaches in SHM that use deep neural networks (DNNs) have shown promising results for real-time damage identification. Such DNNs have been used for multiple SHM tasks for real-time damage identification without requiring an FE model. However, DNNs do not have an implicit mechanism to include input and parameter uncertainty. For this work, a novel supervised interval deep neural network (DINN) that can process input sensor uncertainty for damage detection and classification tasks in SHM is proposed. The proposed method can detect and classify damage in an SHM system solely from sensor data. For other SHM tasks, such as damage localization and severity, our method can be combined with FE models.

Keywords: deep learning, structural health monitoring, interval deep learning, epistemic uncertainty, interval finite element, machine learning

1. Introduction

Structural health monitoring (SHM) is typically concerned with the damage identification and system characterization of engineering structures. An SHM program consists of a network of sensors connected to the structural system, which measure an observed physical variable. Hereafter, we refer to *damage* as structural damage and define it as adverse changes to the stiffness of the structural system which compromise its integrity. An important concern in monitoring systems is the uncertainty in sensor measurements. In SHM systems, uncertainty can be due to environmental factors, sensor malfunctions, unknown sensor placements, miscalibrations, and errors in data acquisition which cumulatively contribute to imprecision in the sensor data.

© 2021 *by D. Betancourt, S. Freitag, and R.L. Muhanna.*

There are two general approaches for detecting and identifying damage in SHM. The first is solid mechanics-based[1], while the second one is data-driven. The solid mechanics-based approach models the actual structure, usually via a finite element (FE) model, from which the structural responses are obtained. These responses (e.g., displacements, stresses, modes of vibration) are compared with signal-processed sensor data to predict whether there is damage. On the other hand, the data-driven approach used for this work uses a novel interval deep learning algorithm on the sensor data to detect or classify damage with or without the aid of an FE model, depending on the complexity of SHM task.

Data-driven techniques for SHM have been previously explored, mostly with unsupervised algorithms. In (Figueiredo et al., 2009) the authors present unsupervised methods for damage detection using different small-scale datasets including an 'autoassociative neural network,' which is conceptually equivalent to an autoencoder. Other works, e.g., (Neves et al., 2017), typically use simulation data from an FE model to detect damage. To account for sensor uncertainty, in (Figueiredo et al., 2011) the authors include noise in the observations to account for operational and environmental variability. More recently, (Figueiredo and Santos, 2018) the authors examine some of the same unsupervised algorithms on the Z24 bridge dataset. Around the same time, deep learning techniques have begun to be used in SHM. In (Wang et al., 2018) the authors use autoencoders to detect and identify damage using simulation data for training. In (Sarkar et al., 2016) the authors use autoencoders and image segmentation to detect and identify damage in aircraft structures. They use training data from images of carbon fiber-reinforced coupons subjected to different damage types.

The data-driven SHM techniques surveyed showed some success—albeit largely limited to damage detection at best and a high rate of false positives at worst, since the unsupervised techniques can detect *any* data anomaly. Indeed, any data anomaly does not guarantee correspondence with structural damage. Another setback is that in most cases the training data comes from FE simulations alone—not from real SHM systems—because data in SHM is difficult to obtain. In this work, we consider the problem of SHM real-time damage detection and classification under uncertainty using a supervised deep learning algorithm with interval analysis (IA). In particular, our contribution consists of developing the Deep Interval Neural Network (DINN) in a classification setting for SHM tasks. The Z24 bridge dataset is used to test our proposed methods (Krämer et al., 1999).

2. Methodologies for Damage Identification in SHM

In this section we introduce the relevant background for SHM tasks and the tools to solve them (Sec. 2.1). We also introduce the challenges with uncertainty in SHM sensor data in Sec. 2.2. Finally, we present important discussion on how to take the most advantage of machine learning to solve SHM tasks in Sec. 2.3, and based on that how we use the DINN to solve SHM tasks.

2.1. Damage Identification Tasks in SHM

Damage identification for SHM consists of the following hierarchical tasks (Figueiredo et al., 2011):

[1] solid mechanics-based is also known as physics-based in the SHM literature.

1. Damage Detection: detect whether damage (anomaly) exists.
2. Damage Localization
3. Damage Type
4. Damage Severity
5. Damage Prognosis: remaining useful lifeline.

Task 1 detects whether there is damage in the structure. Tasks 2-4 encompass the diagnosis of the damage, while Task 5 encompasses the prognosis of the damage. Solid mechanics-based or data-driven techniques can be used for each of these hierarchical tasks. Our goal is to demonstrate that the classification DINN developed in this work can be used for these damage identification tasks more efficiently (i.e., in real-time) and accurately than state-of-the-art techniques. For some tasks, such as damage detection, the DINN can be used without a solid mechanics-based model (e.g., an FE model), while for other tasks it can be used in conjunction with the FE model and/or other techniques to enhance the damage identification.

2.2. Uncertainty in SHM sensor data

Environmental and operational changes along with hardware and software errors are significant sources of uncertainty in the data collected by the SHM sensor network. For instance, some SHM damage identification techniques compute inter-story drifts in the structure from double integration of acceleration data. However, because of uncertainty in the measurements, it is difficult to predict accurate story drifts (Kaya and Safak, 2019) that can detect damage.

Overall, the simplifying common approach of adding random noise to the data cannot account for uncertainty in measurement errors when we cannot calibrate the instrument because the measurements performed by this instrument are the most accurate possible (Sun et al., 2017). Many times this is the case in engineering problems—and there is not enough knowledge to build a probabilistic model to account for the uncertainty. For such cases, interval uncertainty can be beneficial for data analysis via interval analysis (IA).

2.3. SHM Anomaly Detection with Machine Learning

SHM damage identification tasks can be framed as anomaly detection and activity recognition tasks for machine learning algorithms. Well-known applications which use anomaly detection with machine learning are transactional fraud, disease detection, and cybersecurity intrusions. Strictly speaking, in anomaly detection we seek to find rare events or observations that vary significantly from the main distribution of the training data. There are two main categories for anomaly detection with machine learning: supervised and unsupervised. Supervised learning constitutes having a fully-labeled set. Unsupervised learning techniques constitute clustering algorithms, which seek to cluster the normal examples and anomalies get classified in a different cluster. Examples of instances of unsupervised models include k-means, Gaussian Mixture models, DBSCAN, and OPTICS. Within

supervised learning, we can use specialized techniques, such as semi-supervised algorithms, which use signals in the training data itself to provide the labels.

In most cases in machine learning, we only have access to training data only with normal examples or with very few examples of anomalies. Thus, because of the class unbalance, unsupervised algorithms are the main option to detect anomalies. As a result, most of the work in this field has been unsupervised due to the nature of the data. However, while unsupervised learning algorithms can prove to be efficient to *detect* anomalies by clustering, they do not provide additional information on what types of anomalies it has found. Even in the simple case of binary anomaly detection, we seek a particular type of anomaly. Thus, detecting any type of data anomaly often results in a high false positive rate, which can be very costly in safety applications. Furthermore, while there is value in binary detection of anomalies, more often than not the main value lies in classifying the types of anomalies—we must know what type of anomaly exits in order to make an informed decision. In order to do this, we need to employ a supervised or semi-supervised algorithm that can train on a balanced enough dataset. The task therein, before we start training, is to gain access to enough instances of anomalies. These instances can be a combination of real measurements—from laboratory experiments or in-situ damage—and from simulations. For instance, the simulation in SHM would consist of using an FE model to simulate different types of damage and produce synthetic training data.

In machine learning, the main goal is to generalize—that is, to make good predictions on unseen data. For example, with disease detection, the algorithm can be trained with instances of a subset of patients $\mathcal{A}$ from the general population, and once trained, to make predictions on a different subset of patients $\mathcal{B}$—which was not used for training the ML model—and detect potential diseases. In the case of SHM, we can use training data from system $\mathcal{A}$ to make predictions on system $\mathcal{B}$. System $\mathcal{B}$ can be fine-tuned with its own SHM data and/or simulations from its own FE model, but the training should not start from scratch. This is indeed a different mindset from, and an advantage on, the solid-mechanics approach to SHM—since in that approach each system requires its own models.

We will use the DINN as a supervised algorithm as explained in the following sections.

3. Interval Deep Learning for SHM under Uncertainty

3.1. Classification DINN

The interval notation for this work is defined as follows. Let $\mathbb{IR}$ be the set of interval real numbers. Interval real numbers are a closed and bounded set of real numbers (Moore et al., 2009), such that if $\mathbf{X} \in \mathbb{IR}$ is an interval, its endpoints are $\underline{X} = \inf(\mathbf{X}) \in \mathbb{R}$ and $\overline{X} = \sup(\mathbf{X}) \in \mathbb{R}$. *Henceforth, boldface is used to represent interval numbers, matrices, and vectors.* The basic arithmetic operations with real intervals are different than with real numbers. For the reader unfamiliar with IA refer to, e.g., (Moore et al., 2009),(Moore et al., 2009), (Neumaier, 1990), (Kreinovich et al., 2013).

The classification Deep Interval Neural Network (DINN) developed in this work is applied to damage identification tasks in SHM. The classification DINN is the interval extension of a real-valued DNN to process interval-valued matrices of any dimensionality (i.e., interval tensors) and output predictive classes with uncertainty bounds around their respective softmax probabilities.

Without loss of generality, the classification DINN is a supervised learning algorithm which seeks to learn an interval predictive model $F : [\underline{X}, \overline{X}] \to y \in \{1, \ldots, C\}$, where C is the number of deterministic classes. To achieve this goal, the DINN is trained using the $\mathbf{X}$ interval matrix inputs (*features*) in a d-dimensional space with n examples, along with its known y categorical labels. This composes the training set of n examples $\mathcal{T} = \{(\mathbf{X}_1, y_1), \ldots, (\mathbf{X}_n, y_n)\}$ where $\mathbf{X} \in \mathbb{IR}^{n \times d}$ is the interval feature space and $y \in \{1, \ldots, C\}^n$ is the deterministic label space. As its output, the DINN computes for each sample i an estimate of the interval label $\widehat{F}(\mathbf{X}_i)$. In order to do so, the training algorithm iteratively reduces the difference between the true known training label y_i and its prediction label $\widehat{F}(\mathbf{X}_i)$ by minimizing a loss function $\mathcal{L}(\widehat{F}(\mathbf{X}_i), y_i; \mathbf{W})$, such as negative log likelihood, parameterized by interval $\mathbf{W} \in \mathbb{IR}$. Hence, we train the DINN to find the set of interval parameters $\mathbf{W}$ that minimize the loss function. After training, at inference time each future sample i is classified as

$$\widehat{F}(\mathbf{X}_i) = \arg\max_{\mathbf{W}} \Big\{ \texttt{mid}\Big(\sigma(\mathbf{X}_i^{new}\mathbf{W}) \Big) \Big\}, \tag{1}$$

where σ is the `softmax` function as defined on Eq. 2, $\mathbf{W}$ are the trained interval weights of the DINN, and $\mathbf{X}_i^{new}$ are the features of a future query of sensor data in the same tensor dimension as the $\mathbf{X}$ used for training. The softmax function $\sigma(\mathbf{X}_i^{new}\mathbf{W})$ outputs interval probabilities for each class, which in turn give an interval bound around the real-valued softmax probability. We take the midpoint `mid()` of each of these interval class probabilities in order to obtain the predicted class through the arg max function. The softmax function for sample i of class $j \in C$, for a vector $\mathbf{z}$ is

$$\sigma(\mathbf{z})_i^j = \frac{e^{\mathbf{z}_i^j}}{\sum_j e^{\texttt{mid}(\mathbf{z}_i^j)}} \ . \tag{2}$$

For the training data of this application, the features $\mathbf{X}$ are usually the acceleration time-series of the sensor network, strain measurements, and environmental data. The labels are obtained either from domain experts that classify each sample i in the training set, from colocated sensors that can help identify a measure of relative damage (e.g., structural displacements)—which amounts to weakly supervised data (because of the weak labeling signals), or a semi-supervised algorithm where one engineers a way to provide more labels for training given a largely unlabeled dataset. The labels y can be binary: {`Damage, No Damage`} or they can identify the damage: {`Damage Type 1,...,Damage Type C`}. Fig. 1 shows a schematic representation of the classification DINN.

Uncertainty in the data is introduced to the DINN by lower bounds and upper bounds of the features $\mathbf{X}$ for every training sample in the training set.

The ultimate goal of the DINN is to generalize to predictions of unseen uncertain data where labels y_i are not available. The following sections delve into the details of the DINN.

3.2. Fully-Connected DINN

For deep neural networks, the predictive interval function $\widehat{F}(\mathbf{X})$ at the output layer L is composed of the functions of the hidden layers, such that

$$\widehat{F}(\mathbf{X}) = F^{(L)}\Big(F^{(L-1)}\big(F^{(L-2)}\ldots\big(F^{(\ell)}\ldots\big(F(\mathbf{X})\big)\big)\big)\Big) \tag{3}$$

for $\ell = 1 \ldots L$, where ℓ is the layer number.

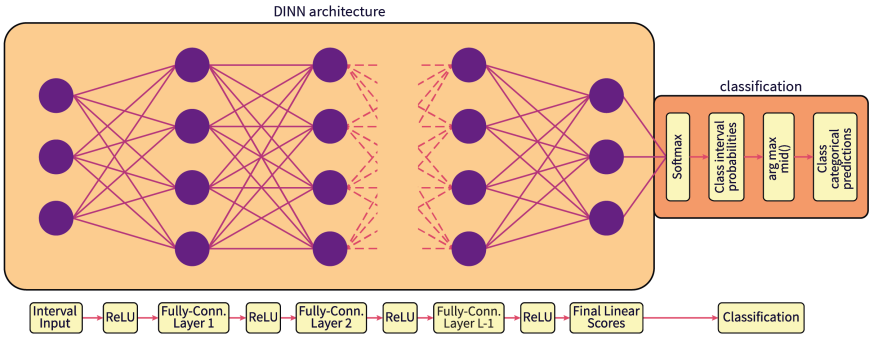


Figure 1. Schematic Representation of Classification DINN.
Nodes represent the activations from previous layers after passing through a `ReLU`
The edges represent the interval weights of the layers.

Hence, the DINN is a composition of interval-valued functions where each of these functions is the output of each of the ℓ layers of the network after passing through a nonlinear activation function. Deep neural networks have more power to express the functions that they attempt to approximate because of the combinatorial advantage of the composition of functions (Goodfellow et al., 2016). A crucial aspect of the development of the DINN is requiring the nonlinear activation function at each layer $F^{(\ell)}$ be monotonic and Lipschitz so that the final output can also be Lipschitz and as sharp as possible. Moreover, the computations of the algorithm to obtain the gradients of the interval functions are ordered in a way to reduce interval dependency by eliminating repeated values of the same variable. With these elements in place, the interval predictions $\widehat{F}(\mathbf{X})$ of the DINN can be reasonably narrow by reducing interval dependency.

3.3. Training

For softmax classification, the corresponding loss at each iteration is computed using the negative log-likelihood of the softmax function, for each mini-batch of size B, as

$$\mathcal{L}(\widehat{F}(\mathbf{X}), y) = \frac{1}{B}\sum_{i}\mathcal{L}_i(\widehat{F}(\mathbf{X}_i), y_i) = -\frac{1}{B}\sum_{i=1}^{B}\log(\mathbf{p}_i^{y_i}) = -\frac{1}{B}\sum_{i=1}^{B}\log\Big(\frac{e^{\mathbf{z}_i^{y_i}}}{\sum_j e^{\mathtt{mid}(\mathbf{z}_i^j)}}\Big), \tag{4}$$

where $\mathbf{p}_i^{y_i}$ is the class softmax interval probability for the target class y_i of sample i, $\mathbf{z}_i$ are the interval linear scores for each sample, $\mathbf{z}_i = \mathbf{X}_i\mathbf{W}$, and the denominator $\sum_j e^{\mathtt{mid}(\mathbf{z}_i^j)}$ is the sum across all classes $j \in C$. Thus, for each sample i we seek to maximize the interval softmax probability of the target class. The loss $\mathcal{L}(\widehat{F}(\mathbf{X}), y)$ it's also called the data loss and it's an interval scalar.

During training, we find a set of optimal parameters $\mathbf{W} \in \mathbb{IR}$ for each layer of the network by minimizing the loss function. In order to minimize the interval loss function $\mathcal{L}(\widehat{F}(\mathbf{X}_i), y_i)$, we use first-order gradient based optimization. In particular, we use mini-batch stochastic gradient descent (SGD) and some of its variants, SGD with momentum (Polyak, 1964) and Adam (Kingma and Ba, 2014) but with an interval modification. The vanilla gradient descent update rule at step $k+1$ for sample i at layer ℓ is defined as

$$\mathbf{W}_{k+1}^{(\ell)} = \mathbf{W}_{k}^{(\ell)} - \alpha_k \nabla_{\mathbf{W}_k^{(\ell)}} \mathcal{L}(\widehat{F}(\mathbf{X}_i), y_i), \tag{5}$$

$$\mathbf{b}_{k+1}^{(\ell)} = \mathbf{b}_{k}^{(\ell)} - \alpha_k \nabla_{\mathbf{b}_k^{(\ell)}} \mathcal{L}(\widehat{F}(\mathbf{X}_i), y_i), \tag{6}$$

where at each step k, $\mathbf{W}_k$ is the interval weight matrix , $\mathbf{b}_k$ is the interval bias vector, α_k is the step size, $\nabla_{\mathbf{W}_k}\mathcal{L}(\cdot)$ and $\nabla_{\mathbf{b}_k}\mathcal{L}(\cdot)$ are the gradients of the loss function with respect to parameters $\mathbf{W}_k$ and $\mathbf{b}_k$ (algorithmic details to obtain the interval gradients are contain in Appendix A). With mini-batch gradient descent, the algorithm chooses uniformly at random a batch of samples of size B from the full training set $\mathcal{T}$ and updates the gradient with Eq. 5 and Eq. 6 or its variants, which trains the network. There are three major steps to training a deep neural network:

1. Compute the loss at each training iteration via forward propagation of the input activations through the network's layers.

2. Compute the gradient of the loss function with respect to the model's parameters via the Backpropagation algorithm.

3. Minimize the loss function via stochastic gradient descent.

Algorithm 1 shows the I-Adam algorithm developed to process interval input and based on the original Adam algorithm (Kingma and Ba, 2014). Both vanilla SGD and SGD with momentum do not require changes to the update rule for interval gradients. However, the Adam optimization algorithm has an adaptive learning rate and implementing it with interval arithmetic causes it to diverge quickly. Thus, we modify the Adam algorithm for interval gradients and call I-Adam. Notice that the algorithm keeps running estimates of the gradients—first moments (mean) and second moments (variance)—in order to perform the weight update. In I-Adam, the second moment biased estimates v_k are modified so that they are real-valued—by squaring only the mid points of the interval gradients in $\texttt{mid}(\mathbf{G}_k)^2$, this modification maintains an adaptive learning rate while reducing divergence.

3.3.1. *Stochastic Gradient Descent with Warm Restarts (SGDR)*
Vanilla SGD and SGD with momentum do not have an adaptive learning rate like Adam. Thus, we use Stochastic Gradient Descent with Warm Restarts (SGDR) developed in (Loshchilov and Hutter, 2016), which follows a cosine decay annealing schedule in order to set the learning rate for every iteration. It is a warm restart because the gradient update restarts from the last iteration, but with a higher learning rate. The learning rate α_k at iteration k, for the i^{th} run of the scheduler is

$$\alpha_k = \alpha^i_{min} + \frac{1}{2}\left(\alpha^i_{max} - \alpha^i_{min}\right)\left(1 + \cos\left(\frac{T_{cur}}{T_i}\right)\right), \tag{7}$$

where $(\alpha^i_{max} - \alpha^i_{min})$ is the range of the learning rate decrease, T_i is the number of iterations performed in a cycle, T_{cur} tracks how many epochs have been performed since the last restart. By using SGDR, we seek to prevent finding suboptimal local minima and to avoid large interval gradients that lead to a blow-up.

Algorithm 1: Adam Algorithm for Interval Input

Input: Step size α
Input: Decay rates $\beta_1, \beta_2 \in [0, 1)$
Input: Initial parameters $\mathbf{W} \in \mathbb{IR}$
Initialize $\mathbf{m}_0 \in \mathbb{IR} = 0$
Initialize $v_0 \in \mathbb{R} = 0$
Initialize $k = 0$
while *stopping criteria not met* **do**
  $k \quad k + 1$
  $\mathbf{G}_k \quad \nabla_{\mathbf{W}} \mathcal{L}_k$ compute interval gradients at step k
  $\mathbf{m}_k \quad \beta_1 \mathbf{m}_{k-1} + (1 - \beta_1)\mathbf{G}_k$ biased interval first moment estimate (interval)
  $v_k \quad \beta_2 v_{k-1} + (1 - \beta_2) \cdot \mathtt{mid}(\mathbf{G}_k)^2$ biased real second moment estimate (real)
  $\widehat{\mathbf{m}}_k \leftarrow \mathbf{m}_k / (1 - \beta_1^k)$
  $\widehat{v}_k \quad v_k / (1 - \beta_2^k)$
  $\mathbf{W}_k \leftarrow \mathbf{W}_{k-1} - \alpha \cdot \widehat{\mathbf{m}}_k / (\sqrt{\widehat{v}_k} + \epsilon)$
return $\mathbf{W}_k$

3.3.2. *Regularization*

Regularization is used for two purposes in the DINN. First, regularization is used to prevent overfitting to the training set, with the goal of achieving better performance in unseen data. It achieves so by shrinking the weight contribution of less relevant features. Second, it stabilizes the solution which for finding interval gradients is of crucial importance. Two forms of regularization are used for the DINN. $L1$ (Lasso) regularization uses the l_1 norm, which is

$$\Omega(\mathbf{W}) = \|\mathbf{W}\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |\mathbf{A}_{ij}|, \tag{8}$$

and $L2$ (Ridge) regularization which is defined as

$$\Omega(\mathbf{W}) = \sum_m \sum_n \mathbf{W}^2_{m,n}. \tag{9}$$

The $L2$ regularization is computed in a way to directly square the elements of the $\mathbf{W}$ in INTLAB, so as to obtain a monotonic expression. The regularization is added to the data loss, as $J(\widehat{F}(\mathbf{X}), y; \mathbf{W}) = \mathcal{L}(\widehat{F}(\mathbf{X}), y; \mathbf{W}) + \lambda\Omega(\mathbf{W})$, where $J(\widehat{F}(\mathbf{X}), y; \mathbf{W})$ is the regularized (total) loss and λ is the regularization strength.

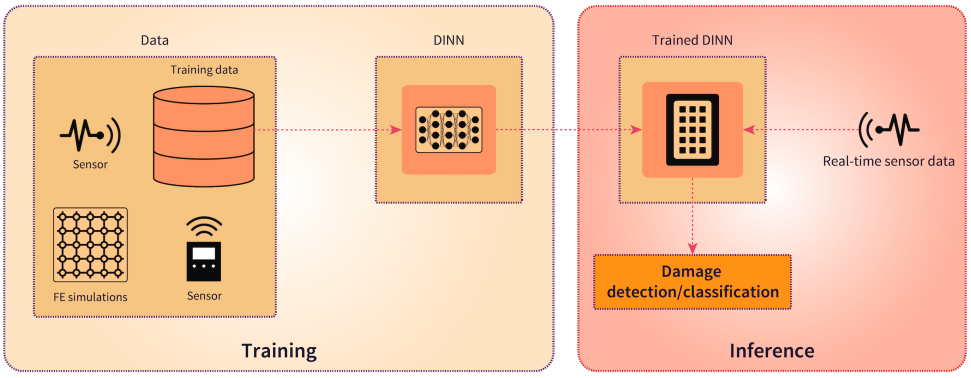


Figure 2. Flowchart for classification DINN for damage detection and classification. The labeled training data consists of sensor measurements and possibly FE simulations. The query data for inference is unlabeled real-time sensor data. The output of the classication DINN for SHM is the detection or classification of the type of structural damage.

3.4. Predictive SHM Tasks Performed with DINN

The DINN is used 'model-free,' i.e., without a structural FE model for the Damage Detection task (see hierarchical tasks in Section 2.1). Nonetheless, the DINN can be used with an interval finite element model (IFEM) for more in-depth damage identification of a particular task or to generate simulation training data. Once the DINN is trained, it can be queried with unlabeled real-time sensor data to detect or classify damage. The overall process is shown on Fig. 2.

4. Experiments

4.1. Iris Dataset

The Iris dataset is a small and simple dataset commonly used to test classifiers, retrieved from the UCI machine learning repository (Dua and Graff, 2017). We thus use this well-benchmarked dataset as a first step to evaluate the skill of the classification DINN to be an efficient discriminator with interval uncertainty in the features. The Iris dataset consists of 150 samples, four features, and three classes. Each class has 50 samples, thus making it a balanced dataset. The features correspond to the sepal and width dimensions, while each class corresponds to a type of iris plant $y = \{\texttt{Setosa, Versicolour, Virginica}\}$.

For the Iris dataset, we conduct separate experiments with increasing uncertainty levels at $\beta = \{2\%, 5\%, 10\%\}$ for all features in $\mathbf{X}$. We use the SGD and the Interval Adam (I-Adam) optimizers. We use a two-layer DINN, with 100 units in the hidden layer, and a maximum of 200 epochs with early stopping. The batch size is 12 samples, while the learning rate is 0.01 and 0.001

for SGD and I-Adam, respectively. A training/validation/test split of 60/10/30 is used for training. The results are summarized on Table I.

Table I. Iris dataset accuracy

Accuracy of two-layer DINN with Iris dataset					
$\beta = 2\%$		$\beta = 5\%$		$\beta = 10\%$	
SGD	I-Adam	SGD	I-Adam	SGD	I-Adam
88.2%	88.9%	85.1%	85.7%	84.5%	84.8%

4.2. Z24 Bridge Dataset

The Z24 bridge dataset has been used as a benchmark to test different SHM damage identification algorithms, both solid mechanics-based and data-driven. It consists of the SHM sensor data from nearly one year of monitoring of a post-tensioned three-span concrete box girder bridge in Switzerland. The bridge was subjected to normal operations for nine months and to artificially induced damage during the last month of operations. In addition to SHM accelerometers, environmental factors were also recorded in parallel. Detailed information about the Z24 bridge SHM program and the different tests is found on (Krämer et al., 1999), (De Roeck et al., 2000).

For the duration of the SHM program on the bridge, different SHM subprograms were conducted on the Z24 bridge with their own goals. The two main SHM subprograms were: (1) a long-term continuous monitoring test for the one-year period before induced damage. This test had the goal of examining the environmental effects on the bridge structural dynamics. (2) short-term progressive damage tests (PDTs) under different damage scenarios for a one-month period prior to the demolition of the bridge. This subprogram had the goal of examining the bridge dynamics under fifteen different damage conditions and two reference undamaged conditions.

For subprogram (2), the bridge contained nine sensor network set-ups, each of which recorded the vibrations associated with the progressive damage tests (PDT). Each set-up has 33 accelerometers. Fig. 3 shows a side view and elevation view from the bridge, based on (Krämer et al., 1999). Refer to sensor locations on (Krämer et al., 1999) and (De Roeck et al., 2000).

4.2.1. *Damage Detection Experiments*

The first task performed with the classification DINN is damage detection (refer to Sec. 2.1). In order to detect damage, we take the sensors readings from `set-up 1` on the bridge deck for the ambient vibrations. This gives a total of 21 accelerometer readings in the vertical, longitudinal, and transverse directions.

The experiments consist of the followings steps:

1. Perform interval uncertainty quantification on the dataset and convert real-valued data to interval-valued data based on the determined uncertainty levels. This step produces the interval training data.

2. Train the classification DINN with the interval training data.

3. Perform inference with the trained classification DINN by querying it with new interval data to obtain the corresponding class predictions.

4. Obtain corresponding interval softmax probability for the predicted class.

4.2.1.1. *Interval Uncertainty* Sensor uncertainty is likely to be present in an SHM system due to multiple sources, as explained in Sec 2.2. By performing these experiments, we seek to assess the classification performance of the DINN under interval uncertainty. We can represent this uncertainty in the sensors via intervals and then propagate it through the DINN in order to predict the class of each sample. The DINN is very flexible as to how the interval input is defined. Intervals can be defined for subsets of samples and features, and a different interval uncertainty can be defined for each sample separately. A scenario could be one in which during some environmental event (e.g., snow, high winds, abnormal traffic), a subset of sensors readings might become corrupted. In order to quantify the uncertainty, the training or query data can be given an interval according to some expert belief. Such interval can be given manually by the SHM operators, by a rules-based system, or even by a separate machine learning model. Furthermore, one can see that the DINN can be beneficial in cases of missing data at some time intervals for some sensors but we decide to exclude the missing data scenario from the experiments and leave it to a future study.

For the Z24 bridge, we use an uncertainty level of $\beta = 10\%$ for sensor `100V` and $\beta = 1\%$ for all other sensors in $\mathbf{X}$ (refer to (Krämer et al., 1999) for sensor locations). Then, we perform another experiment with the same uncertainty level for the same sensor and an uncertainty level of $\beta = 5\%$ for all other sensors in $\mathbf{X}$.

4.2.1.2. *Classification DINN Model Details* We trained the classification DINN with SGD with momentum *SGD_mom* and I-Adam. We use two hidden layers, 500 units per layer, and 300 epochs. The batch size is 2048 samples. For time series data, small batches are sometimes not too informative and this proves to be the case here. We normalize all the features by the l_2 vector norm of each feature. We use a random weight initialization based on a normal distribution with zero mean and unit variance and a weight scale of 0.05. To evaluate the skill of the model, we use a training/validation/test split of 60/10/30.

For SGD with momentum we use a cosine annealing learning rate schedule with warm restarts (SGDR) with an initial learning rate of 0.001 and 5 initial steps, while doubling the number of annealing steps after every cycle. Using the cosine annealing scheduler with warm restarts proved to be highly beneficial when compared with a linearly decaying learning rate without restarts. For I-Adam, we use a learning rate of 0.00005 and linearly decay it by 0.99 after every iteration.

4.2.2. *Results*

Table II shows the comparison of test set results of the trained classification DINN trained with SGDR and Interval-Adam (I-Adam). It is shown that higher accuracy is achieved with SGDR than with I-Adam, and the difference is more significant at higher levels of uncertainty (90.6% vs. 72.2%). Fig. 4 shows the training loss of SGD with momentum with linear learning rate decay (*SGD_mom*),

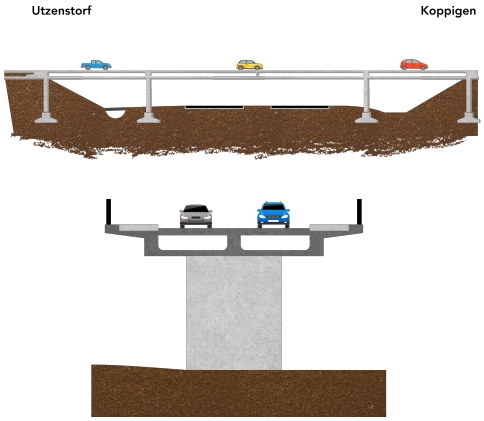


Figure 3. Former Z24 Bridge side and elevation views adapted from (Krämer et al., 1999)
Top: Side view
Bottom: Pier elevation view

SGDR, and I-Adam. On Fig. 4, we can see that SGDR has a flatter loss curve than I-Adam and that *SGD_mom*. Despite the flatter loss curve, SGDR has a narrower loss and outperforms I-Adam in the test set by a wide margin as shown on Table II. One reason for the inferior performance of I-Adam in this dataset could be that the adaptive learning rate of Adam is overly ambitious and goes into interval gradient regions where it gets stuck in local minima. This observation gave us insight into first-order optimization for interval gradients and led us to experiment with SGDR. As its real-valued counterparts, intervals need an efficient learning rate but also benefit greatly from restarts to escape both local minima and avoid exploding interval gradients.

Table II. Z24 test set results

Accuracy			
{ 1%, 10%}		{5%, 10%}	
SGDR	I-Adam	SGDR	I-Adam
95.8%	85.3%	90.6%	72.2%

5. Conclusion and Future Work

We presented a novel interval deep learning algorithm which can be used for real-time damage detection tasks under sensor uncertainty. The algorithm, the Deep Interval Neural Network (DINN) classifier, can be used to detect damage in a structure in real-time, using sensor data alone and subject to uncertainty. The algorithm is in a supervised setting, which implies that both normal and abnormal (structural damage) samples are needed to train the algorithm. We tested the algorithm with the Z24 bridge benchmark and reached an accuracy of 95.8% using an interval uncertainty level of $\beta = 10\%$ for one sensor and $\beta = 1\%$ for all others in the feature set.

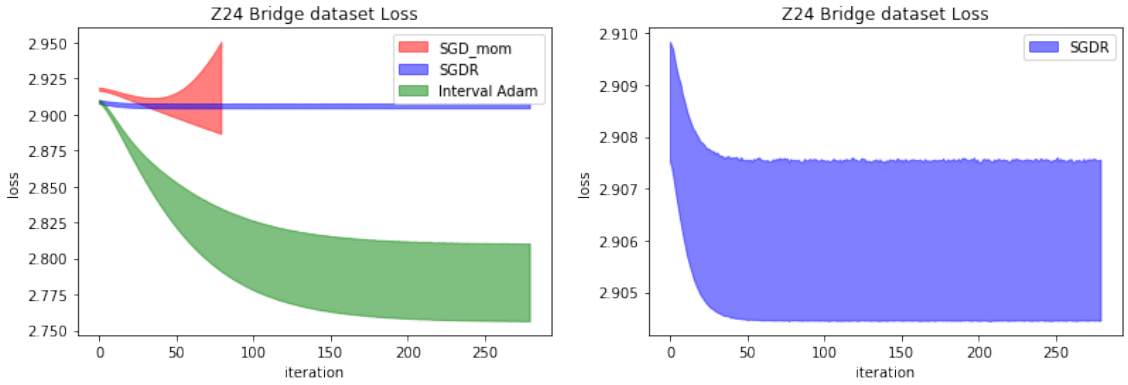


Figure 4. Training loss history comparison between SGD with momentum with only linear learning rate decay, SGD with momentum with cosine decay and restarts (SGDR), and I-Adam (left); Training loss history of SGD with momentum with cosine decay and restarts (SGDR). SGDR is the best performer for the held-out dataset. SGD with momentum with only linear learning rate decay blows up after about 80 iterations. I-Adam gives a smooth loss but the interval is too wide to make correct predictions.

While samples of structural damage were included in this dataset, it is not the case for most SHM system. Nonetheless, one of the advantages of deep learning techniques is that damage samples can be obtained from other SHM systems (either real or from laboratories) to pretrain the DINN and fine-tuned with simulations from FE models. Future promising research directions for this work include weakly or semi-supervised models and more complex SHM tasks, such as damage localization.

Appendix

A. Classification DINN Algorithmic Details

A.1. Interval Forward Propagation

Algorithm 2 of this appendix computes the predictions $\widehat{F}(\mathbf{X})$ and the loss $\mathcal{L}(\widehat{F}(\mathbf{X}_i), y)$ of the DINN through forward-propagation of each mini-batch of data. The layer output $\mathbf{z}^{(\ell)}$ (scores) is computed through an affine transformation as $\mathbf{z}^{(\ell)} = \mathbf{h}^{(\ell-1)}\mathbf{W}^{(\ell)} + \mathbf{b}^{(\ell)}$, at each layer ℓ; $\mathbf{W}^{(\ell)} \in \mathbb{IR}^{H^{(\ell-1)} \times H^{(\ell)}}$, $\mathbf{h}^{(\ell-1)} \in \mathbb{IR}^{B \times H^{(\ell-1)}}$, $\mathbf{b}^{(\ell)} \in \mathbb{IR}^{H^{(\ell)}}$, and $\mathbf{z}^{(\ell)} \in \mathbb{IR}^{B \times H^{(\ell)}}$, $H^{(\ell-1)}$ is the dimension of layer $\ell - 1$, $H^{(\ell)}$ is the dimension of layer ℓ, and B is the number of samples in each mini-batch of data. Then, the layer output $\mathbf{z}^{(\ell)}$ passes through an element-wise activation function $F(\mathbf{z}^{(\ell)})$. The default activation function for the DINN is a `ReLU`.

In the classification setting, for the final layer L, the prediction is the output of the softmax function $F(\mathbf{z}^{(L)}) = \texttt{softmax}(\mathbf{z}^{(L)})$, where $\mathbf{z}^{(L)} \in \mathbb{IR}^{B \times C}$ for B mini-batch samples in C classes. Finally, the loss at the end of each epoch is computed using negative log likelihood for each mini-batch of size B, as

$$\mathcal{L}(\widehat{F}(\mathbf{X}), y) = \frac{1}{B}\sum_{i} \mathcal{L}_i(\widehat{F}(\mathbf{X}_i), y_i) = -\frac{1}{B}\sum_{i=1}^{B} \log(\mathbf{p}_i^{y_i}) = -\frac{1}{B}\sum_{i=1}^{B} \log\Big(\frac{e^{\mathbf{z}_i^{y_i}}}{\sum_j e^{\texttt{mid}(\mathbf{z}_i^j)}}\Big). \tag{10}$$

The loss is an interval scalar. Notice that `ReLU`, `softmax`, and their derivatives are monotonic functions, which allows for calculation of sharp interval enclosures. The `ReLU` is the standard layer activation function for the DINN but it can be replaced by a different activation function with ease.

Algorithm 2: Forward Propagation Algorithm for Each Mini-batch of Data

Obtain the input raw real-valued features X and labels y
Embed X into interval input features $\mathbf{X}$ considering lower $\underline{X}$ and upper bounds $\overline{X}$ for each sample
Choose number of layers L
Require model parameters $\mathbf{W}^{(\ell)}$, for each layer ℓ
$\mathbf{h}^{(0)} = \mathbf{X}$
for $\ell = 1 : L$ **do**
  $\mathbf{z}^{(\ell)} = \mathbf{h}^{(\ell-1)}\mathbf{W}^{(\ell)} + \mathbf{b}^{(\ell)}$
  $\mathbf{h}^{(\ell)} = F(\mathbf{z}^{(\ell)})$
Prediction $\widehat{F}(\mathbf{X}) = \mathbf{h}^{(L)}$
Compute loss $\mathcal{L}(\widehat{F}(\mathbf{X}_i), y_i)$

A.2. Interval Backpropagation

The first step for SGD is to find the gradients of the loss with respect to the weights, i.e., $\nabla_{\mathbf{W}}\mathcal{L}$. These gradients are computed via automatic differentiation in the computational graph of the neural network. For the DINN, we use the Backpropagation algorithm to take derivatives of interval matrices. In particular, the chain rule is used to compute the gradient as follows

$$\nabla_{\mathbf{W}^{(\ell)}}\mathcal{L} = \frac{\partial \mathcal{L}}{\partial \mathbf{W}^{(\ell)}} = \frac{\partial \mathcal{L}}{\partial \mathbf{z}^{(\ell)}}\frac{\partial \mathbf{z}^{(\ell)}}{\partial \mathbf{W}^{(\ell)}} \tag{11}$$

where ℓ is a layer, $\frac{\partial \mathbf{z}^{(\ell)}}{\partial \mathbf{W}^{(\ell)}}$ is the partial derivative of the score with respect to the weights, and $\frac{\partial \mathcal{L}}{\partial \mathbf{z}^{(\ell)}}$ is the partial derivative of the loss with respect to the score at layer ℓ.

This procedure is summarized in Algorithm 3 of this appendix for general loss and activation functions. In the algorithm, the $\boldsymbol{\delta}$'s denote error terms that recursively update the error in the network at each layer. The interval gradients are accumulated in the $\mathbf{G}$ variable, a data structure containing the interval tensors of gradients. Regularization $\lambda \cdot \Omega(\mathbf{W})$ is also added to the gradient computation as shown on Algorithm 3 of this appendix. In the implementation of this algorithms, care must be taken not to repeat interval variables in order to avoid interval dependency.

Algorithm 3: Backpropagation Algorithm for Interval Loss for Each Mini-batch of Data

Run forward computation algorithm to obtain layer activations $\mathbf{h}^{(\ell)}$
Compute gradient of loss at output layer $\boldsymbol{\delta}_{out}^{(L)}$
$\mathbf{G} = \boldsymbol{\delta}_{out}^{(L)}$
$\nabla_{\mathbf{W}^{(L)}}\mathcal{L} = \mathbf{h}^{(L-1)^T}\mathbf{G} + \lambda\nabla_{\mathbf{W}^{(L)}}\Omega(\mathbf{W})$
$\nabla_{\mathbf{b}^{(L)}}\mathcal{L} = \mathbf{G} + \lambda\nabla_{\mathbf{b}^{(L)}}\Omega(\mathbf{W})$
$\mathbf{G} \leftarrow \mathbf{G}\mathbf{W}^{(L)^T}$
for $\ell = L-1, \ldots, 1$ *layers* **do**
  $\mathbf{G} \leftarrow \boldsymbol{\delta}_{\mathbf{z}}^{(\ell)} = \mathbf{G} \odot F'(\mathbf{z}^{(\ell)})$
  $\nabla_{\mathbf{W}^{(\ell)}}\mathcal{L} = \mathbf{h}^{(\ell-1)^T}\mathbf{G} + \lambda\nabla_{\mathbf{W}^{(\ell)}}\Omega(\mathbf{W})$
  $\nabla_{\mathbf{b}^{(\ell)}}\mathcal{L} = \mathbf{G} + \lambda\nabla_{\mathbf{b}^{(\ell)}}\Omega(\mathbf{W})$
  $\mathbf{G} \leftarrow \mathbf{G}\mathbf{W}^{(\ell)^T}$

For the classification DINN as presented on this paper, with cross-entropy loss (negative log likelihood) and `ReLU` activations, the partial derivatives in Algorithm 3 specialize as follows. At the final layer L, for each class $j \in C$, we have

$$\boldsymbol{\delta}_{out}^{(L)} = \boldsymbol{\delta}_{\mathbf{z}}^{(L)} = \frac{\partial \mathcal{L}}{\partial \mathbf{z}^{(L)}} = \frac{\partial \mathcal{L}}{\partial \mathbf{h}^{(L)}}\frac{\partial \mathbf{h}^{(L)}}{\partial \mathbf{z}^{(L)}} = (\mathbf{p}_i^j - \mathbb{1}\{y_i = j\}), \tag{12}$$

where $\mathbb{1}\{\cdot\}$ is the indicator function. For $\frac{\partial \mathbf{z}^{(\ell)}}{\partial \mathbf{W}^{(\ell)}}$, we have

$$\frac{\partial \mathbf{z}^{(\ell)}}{\partial \mathbf{W}^{(\ell)}} = \mathbf{h}^{(\ell-1)^T}. \tag{13}$$

At the hidden layers ($\ell = L-1, L-2, \ldots, 1$), going backwards in the network graph, the partial derivative with respect to the score is

$$\begin{aligned}\boldsymbol{\delta}_{\mathbf{z}}^{(\ell)} = \frac{\partial \mathcal{L}}{\partial \mathbf{z}^{(\ell)}} &= \boldsymbol{\delta}_{\mathbf{z}}^{(\ell+1)}\mathbf{W}^{(\ell+1)^T} \odot \underbrace{F'(\mathbf{z}^{(\ell)})}_{\substack{\text{derivative of activation}}} \\ &= \boldsymbol{\delta}_{\mathbf{z}}^{(\ell+1)}\mathbf{W}^{(\ell+1)^T} \odot \underbrace{\mathbb{1}\{\mathbf{h}^{(\ell)} \geq 0\}}_{\substack{\text{derivative of } \texttt{ReLU} \\ = 1 \text{ if } \mathbf{h}^{(\ell)} \geq 0 \\ = 0 \text{ otherwise.}}}.\end{aligned} \tag{14}$$

where $\mathbf{A} \odot \mathbf{B}$ is the Hadamard (element-wise) product of $\mathbf{A}$ and $\mathbf{B}$.

Finally, the gradient with respect to the weights at each layer is

$$\nabla_{\mathbf{W}^{(\ell)}}\mathcal{L} = \mathbf{h}^{(\ell-1)^T}\boldsymbol{\delta}_{\mathbf{z}}^{(\ell)}, \tag{15}$$

and the gradient with respect to the biases at each layer is

$$\nabla_{\mathbf{b}^{(\ell)}}\mathcal{L} = \boldsymbol{\delta}_{\mathbf{z}}^{(\ell)}. \tag{16}$$

References

De Roeck, G., Peeters, B., and Maeck, J. (2000). Dynamic monitoring of civil engineering structures. *Computational Methods for Shell and Spatial Structures*.

Dua, D. and Graff, C. (2017). UCI machine learning repository.

Figueiredo, E., Park, G., Farrar, C. R., Worden, K., and Figueiras, J. (2011). Machine learning algorithms for damage detection under operational and environmental variability. *Structural Health Monitoring*, 10(6):559–572.

Figueiredo, E., Park, G., Figueiras, J., Farrar, C., and Worden, K. (2009). Structural health monitoring algorithm comparisons using standard data sets. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States).

Figueiredo, E. and Santos, A. (2018). Machine learning algorithms for damage detection. *Vibration-Based Techniques for Damage Detection and Localization in Engineering Structures*, pages 1–39.

Goodfellow, I., Bengio, Y., Courville, A., and Bengio, Y. (2016). *Deep learning*, volume 1. MIT press Cambridge.

Kaya, Y. and Safak, E. (2019). Structural health monitoring: Real-time data analysis and damage detection. In *Seismic Structural Health Monitoring*, pages 171–197. Springer.

Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Krämer, C., De Smet, C., and De Roeck, G. (1999). Z24 bridge damage detection tests. In *IMAC 17, the International Modal Analysis Conference*, volume 3727, pages 1023–1029. Society of Photo-optical Instrumentation Engineers.

Kreinovich, V., Lakeyev, A. V., Rohn, J., and Kahl, P. (2013). *Computational complexity and feasibility of data processing and interval computations*, volume 10. Springer Science & Business Media.

Loshchilov, I. and Hutter, F. (2016). Sgdr: Stochastic gradient descent with warm restarts. *arXiv preprint arXiv:1608.03983*.

Moore, R. E., Kearfott, R. B., and Cloud, M. J. (2009). *Introduction to interval analysis*, volume 110. Siam.

Neumaier, A. (1990). *Interval methods for systems of equations*, volume 37. Cambridge university press.

Neves, A., Gonzalez, I., Leander, J., and Karoumi, R. (2017). Structural health monitoring of bridges: a model-free ann-based approach to damage detection. *Journal of Civil Structural Health Monitoring*, 7(5):689–702.

Polyak, B. T. (1964). Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics*, 4(5):1–17.

Sarkar, S., Reddy, K. K., Giering, M., and Gurvich, M. R. (2016). Deep learning for structural health monitoring: A damage characterization application. In *Annual conference of the prognostics and health management society*, pages 176–182.

Sun, L., Dbouk, H., Neumann, I., Schön, S., and Kreinovich, V. (2017). Taking into account interval (and fuzzy) uncertainty can lead to more adequate statistical estimates. In *North American Fuzzy Information Processing Society Annual Conference*, pages 371–381. Springer.

Wang, R., Li, L., and Li, J. (2018). A novel parallel auto-encoder framework for multi-scale data in civil structural health monitoring. *Algorithms*, 11(8):112.

Numerical simulation of a 3D concrete printing process under polymorphic uncertainty

Albrecht Schmidt[1,*] and Tom Lahmer[1,2]
[1] *Materials Research and Testing Institute at the Bauhaus-Universität*
[2] *Institute of Structural Mechanics, Bauhaus-Universität Weimar, Germany*
[*] Corresponding author: albrecht.schmidt[at]mfpa.de

Abstract. During the past few years, additive manufacturing techniques for concrete have gained extensive attention. In particular, the extrusion-based 3D concrete printing exhibited a rapid development. Previous investigations are mostly based on experimental studies or even trial-and-error tests. A more profound understanding of the relationships between the process and material parameters and the manufactured structure can be advanced by numerical modeling and simulation. It enables to study a wide range of parameters such that dependencies of properties of the printed product on different influencing factors can be identified. Taking into account the uncertain nature of process and material parameters of the extrusion-based concrete printing, the process can be reliably controlled and finally optimized. The presented study uses pseudo-density for an efficient finite element based modeling approach, where predefined elements are activated layer-by-layer. Material parameters are described as temporally and cross-correlated random processes, while an interval description is used to take the vague information on their correlation parameters into account. First results of a reliability estimation are shown for a 2D modeled additively manufactured wall.

Keywords: Random Process, Polymorphic Uncertainty, Concrete Printing, 3DCP

1. Introduction

In the past decade 3D concrete printing (3DCP) technology is getting recognition in the construction industry. The technology has several advantages such as decreased construction time, design flexibility compared to conventional construction methods, cost reduction by avoiding formwork costs, reduced waste, reduced manpower which decreases injuries and fatalities on construction sites and increased sustainability of the construction industry (Wu et al., 2016; Hager et al., 2016). A growing number of projects can be observed in both private enterprises and research institutes worldwide (Buswell et al., 2018). According to Wangler et al. (Wangler et al., 2016) among the different concrete printing processes like particle-bed 3D printing (Le et al., 2012), the extrusion-based process is the most widely used. Therefore, the presented contribution considers the extrusion-based process which has been demonstrated to be applicable to the construction industry (Wu et al., 2016; Hager et al., 2016). However, there are several challenges to be addressed to fully implement the technology. An improved profound understanding of the relationship between design, material behavior, and process parameters is eminent. The materials' rheology and the process parameters,

©2021 *by A. Schmidt, T. Lahmer.*

such as printing speed, time, temperature etc., have an impact on the fresh and hardened state of the printed structure (Panda et al., 2017): Effects like geometric variation of the printed layers (Buswell et al., 2018), stability failure due to local material strength (Wangler et al., 2016) and global buckling (Gosselin et al., 2016) have been reported. Under-filling influenced by printing process parameters (nozzle velocity, pumping rate etc.) and material properties (rheology) are reported (Le et al., 2012). Due to this complex behavior and the variability of material parameters (Buswell et al., 2018) and printing process parameters, adequate process parameters are commonly determined by means of trial-and-error (Suiker, 2018).

By implementing a numerical simulation of the 3DCP process a better understanding of the system parameters on the structure behavior can be achieved. The influence of the parameter uncertainty can be studied. In this contribution a finite element framework is proposed to model the time depending printing process efficiently and is combined with an random process model of fresh concrete material parameters, where auto- and cross-correlation are considered as intervals yielding a polymorphic uncertainty model. A reliability analysis evaluates the dependence on these parameters.

2. Method

2.1. Finite element model of a printing process

According to the layer-wise production process of the concrete structure, the finite element (FE) model "needs to grow". The idea is to work on a previously generated FE mesh, where the elements are activated sequentially in correspondence to the printing process, cf. Fig. 1. This approach enables an efficient simulation avoiding the computational demanding re-meshing procedure. During the

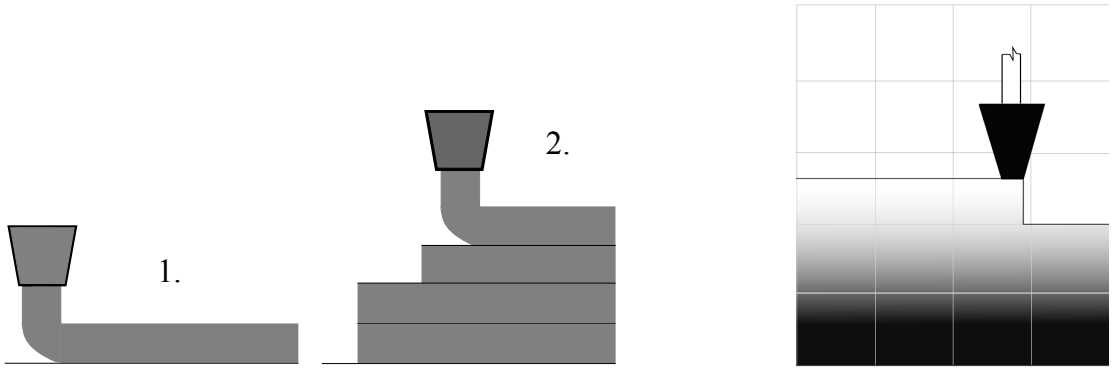


Figure 1. (left) Schematic of a layer-by-layer process; (right) Temporal evolution of a material parameter on a pre-meshed FE model with sequentially activated elements

simulation, i.e. the solution of a finite element system of the form

$$\boldsymbol{K}(\boldsymbol{r},t)\boldsymbol{U}(\boldsymbol{r},t) = \boldsymbol{F}(r,t) \tag{1}$$

is solved, where $\boldsymbol{K}$ is the global stiffness matrix, $\boldsymbol{U}$ is the vector of nodal displacements, the loads due to self-weight are considered in the right-hand-side vector $\boldsymbol{F}$ and $\boldsymbol{r}$ and t denote the space and

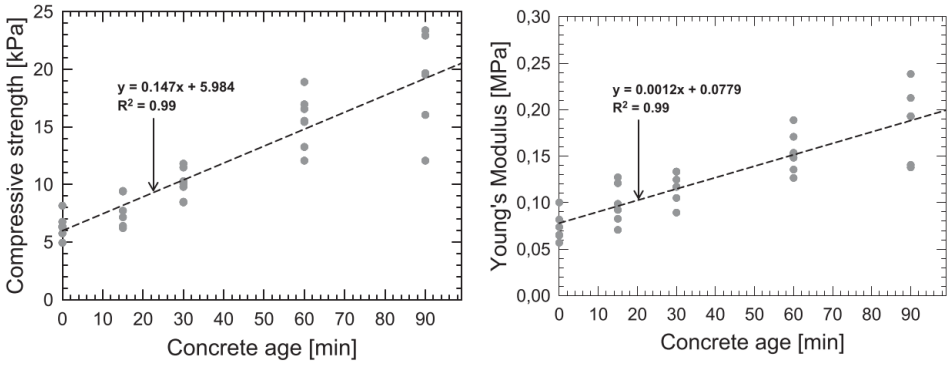


Figure 2. Experimentally determined evolutions of the compressive strength (left) and of the Young's modulus (right), reported in (Wolfs et al., 2018); fit was based on average test results.

time variable, respectively. Additionally, as all parameters do change with time according to the hardening process the material parameters are both spatially and temporarily varying, indicated by the space and time dependency in Eq. 1. An example of the temporal development of the material parameters compressive strength and Young's modulus is shown in Fig. 2 (Wolfs et al., 2018). Obviously, both the compressive strength and the Young's modulus of the printed concrete increase with the passing time. Additionally, the experimental results show a significant scatter of the material properties leading to the necessity of an uncertainty model for the FE model parameters where temporal and spatial correlations are taken into account.

2.2. Polymorphic uncertainty

A principal scheme of the proposed uncertainty description of a 3DCP model is depicted in Fig. 3. The material parameters of the printed structure originate from the time-dependent material parameters $M_1(t)$ and $M_2(t)$ of the pumped concrete, where t denotes the time. While the material properties undergo a natural variation it is reasonable to assume a temporal auto-correlation of the concrete properties since the fresh pumped concrete is taken from the same supply. Hence, a random process description is proposed, described by an auto-correlation function $R(d_t, t_c)$, where d_t denotes the time difference and t_c the auto-correlation time.

Additionally, a cross-correlation parameter c_{12} is introduced to take into account that most of these material parameters are not independent from each other. The same holds for process parameters (e.g. printing velocity, pumping rate), which are fixed here for the sake of clarity.

The stochastic characterization of the material parameters is based on distribution parameters and the auto- and cross-correlation parameters t_c and c_{12}, respectively. While the distribution parameters can usually be determined from experimental data, information on auto- and cross-correlation is typically vague, which constitutes an epistemic uncertainty. Therefore, in this contribution auto- and cross-correlation are considered as interval variables t_c^{i} and c_{12}^{i} to model this kind of uncertainty, where the first one yields a interval valued auto-correlation function $R^{\mathrm{i}}(d_{\mathrm{t}}, t_c^{\mathrm{i}})$.

The uncertainty in the temporal evolution of the material parameters of the printed structure could also be described in an analog manner but is assumed as deterministic for this study. As a

result system responses (for example the maximum displacement u_{mx}) are described in a interval-stochastic space. Hence, a reliability analysis yields an interval failure probability $\hat{p}^{\mathrm{i}}$.

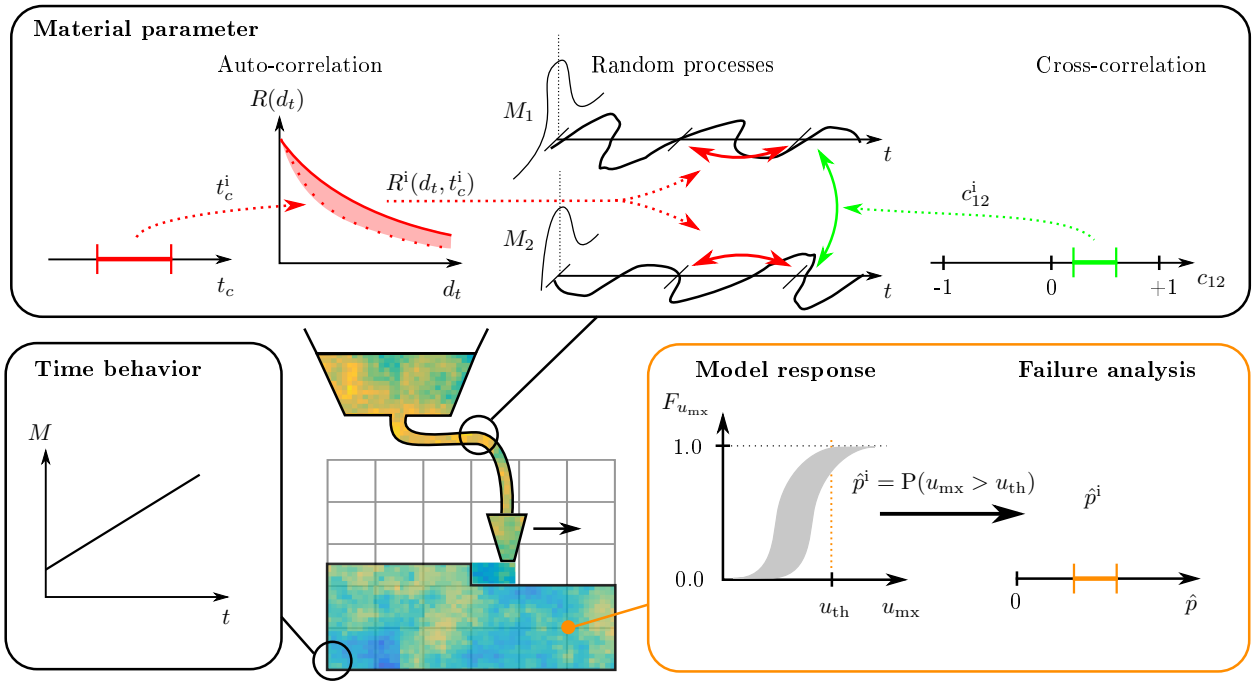


Figure 3. Principle scheme of a 3DCP model with time dependent, uncertain material parameters, which are described as interval probability based random processes

3. Numerical Example

3.1. Description

3.1.1. *Finite Element Model*

We consider a 2D finite element (FE) model of a printed wall with a size of $1\,\mathrm{m} \times 0.2\,\mathrm{m}$. The process parameters are constant: a printing velocity of $v_{\mathrm{p}} = 1\,\mathrm{m/min}$ is chosen while the layer height is set to 10 mm. The total printing time is 1200 s. The wall is modeled with a regular rectangular mesh using 40x20 Q8 elements. A linear-elastic plane-stress material model is assumed. The time-dependent Young's modulus E and compressive strength σ_y are given as

$$E(\boldsymbol{r}, t) = E_0(\boldsymbol{r}) + E_1(t) \tag{2}$$

and

$$\sigma_y(\boldsymbol{r}, t) = \sigma_{y,0}(\boldsymbol{r}) + \sigma_{y,1}(t), \tag{3}$$

where $E_0(\boldsymbol{r})$ and $\sigma_{y,0}(\boldsymbol{r})$ are the starting value as a function of position $\boldsymbol{r}$, E_1 and $\sigma_{y,1}$ are the slopes of the temporal increment, respectively, and t the time. The loading consists of the increasing self-weight, considering a constant mass density of $2020\,\mathrm{kg/m^3}$ (Wolfs et al., 2018). The modeled wall is fixed vertically at the whole bottom edge and horizontally at the middle of the bottom edge.

In order to take the temporal development of the material parameters into account, the total displacement $\boldsymbol{U}$ of the whole system is calculated by activating the elements in layer-wise fashion for all layers $l = 1 \dots n_\mathrm{l}$ (bottom to top), with $n_\mathrm{l} = 20$ in this example, solving

$$\boldsymbol{K}_l \boldsymbol{u}_l = \boldsymbol{F}_l, \tag{4}$$

where the stiffness matrix $\boldsymbol{K}_l$ and the displacement vector $\boldsymbol{u}_l$ take only the elements of the layers 1 to l into account and $\boldsymbol{F}_l$ consists only of the self-weight of layer l. Therefor, the total displacement can be determined as

$$\boldsymbol{U} = \sum_{l=1}^{n_\mathrm{l}} \boldsymbol{u}_l \tag{5}$$

The temporal evolution of the compressive strength can be seen in Fig. 4. An increase of $E_1 = 1.2\,\mathrm{kPa/min}$ and $\sigma_{y,1} = 0.147\,\mathrm{kPa/min}$ are chosen in accordance to (Wolfs et al., 2018), while E_0 and $\sigma_{y,0}$ are modeled via a cross-correlated random process, which is described in the next section.

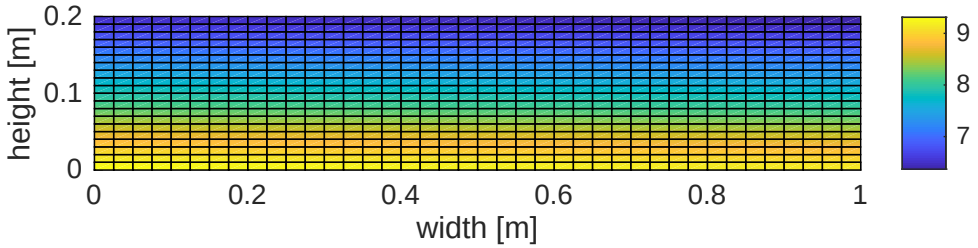


Figure 4. Compressive strength $\sigma_y(\boldsymbol{r}, t)$ in [kPa] for $t = 1200\,\mathrm{s}$ and a constant $\sigma_{y,0} = 6.37\,\mathrm{kPa}$

3.1.2. *Polymorphic Uncertainty Model*

In this contribution the proposed uncertainty scheme of Fig. 3 is applied to the material parameters of the fresh concrete, namely Young's modulus and compressive strength, which are than mapped to E_0 and $\sigma_{y,0}$, respectively.

The variability of the material parameter of such a printing process can be described as lognormally distributed random processes $E^\mathrm{rp}(t)$ and $\sigma_y^\mathrm{rp}(t)$, respectively. Experimental studies suggest a strong positive correlation of the chosen material parameters (Wolfs et al., 2018) which is covered by introducing a cross-correlation c between the processes while auto-correlation is described using a exponentially decaying function with the correlation time t_c. Considering the vague information on cross- and auto-correlation, both parameters are modeled as intervals with $t_\mathrm{c}^\mathrm{i} = [7.5; 75]\mathrm{s}$ and $c^\mathrm{i} = [0.7; 0.9]$, respectively. A mathematically more profound definition of interval probability based random processes can be found in Sec. 3.1 of (Schietzold et al., 2019).

The mean and the standard deviation of the E^{rp} and σ_y^{rp} are taken from (Wolfs et al., 2018) as 74 kPa with 14.5 kPa and 6.37 kPa with 0.95 kPa, respectively. In Fig. 5 realizations of the random processes are plotted for auto-correlation time $t_{\mathrm{c}} = 25\,\mathrm{s}$ and cross-correlation $c = 0.9$.

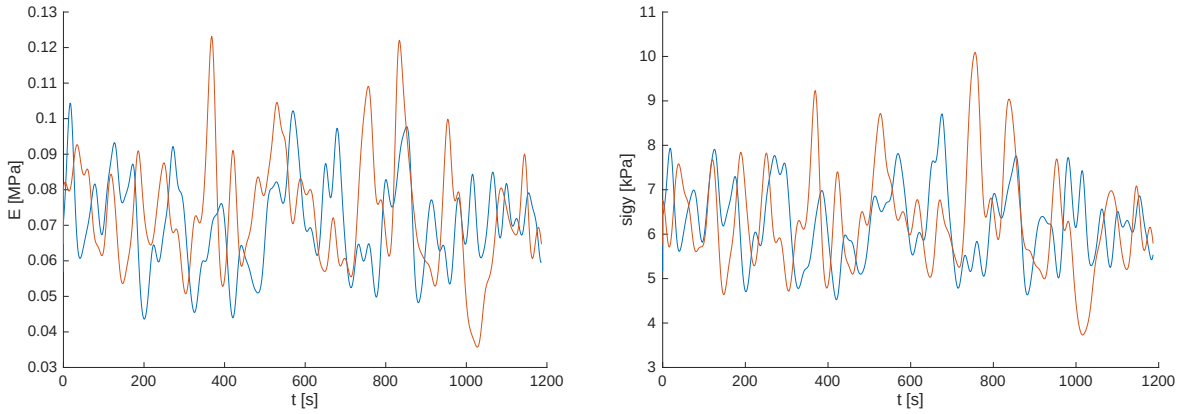


Figure 5. Samples of random processes $E^{\mathrm{rp}}(t)$ (left) and $\sigma_y^{\mathrm{rp}}(t)$ (right) for $t_{\mathrm{c}} = 25\,\mathrm{s}$ and $c = 0.9$

The behavior of the printing process is studied within the interval-stochastic space. A modified Karhunen-Loeve expansion is used to generated the cross-correlated random processes (Vořechovský, 2008). While the interval space $t_{\mathrm{c}}^{\mathrm{i}} \times c^{\mathrm{i}}$ is sampled using a latin hypercube algorithm (50 points), a crude Monte Carlo approach is used in the stochastic space (1000 samples), resulting in 50 000 finite element simulations.

3.2. Results

In Fig. 6 one realization of the Young's modulus at the end of the printing process ($t = 1200\,\mathrm{s}$) is depicted. Both the random nature of the fresh concrete property (E_0) and the increase of stiffness over time can be observed.

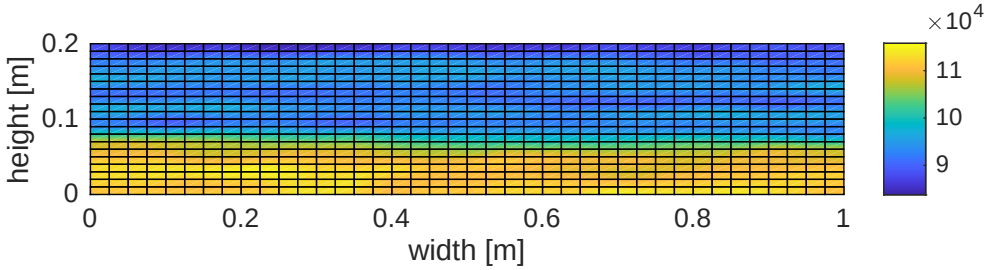


Figure 6. Realization of Young's modulus $E(\boldsymbol{r})$ in [Pa] at $t = 1200\,\mathrm{s}$ for E_0 mapped from the random process $E^{\mathrm{rp}}(t)$

Exemplary results of the structural analysis are shown in Fig. 7. The reasonable deformation shape displays horizontal displacements up to 1.5 cm.

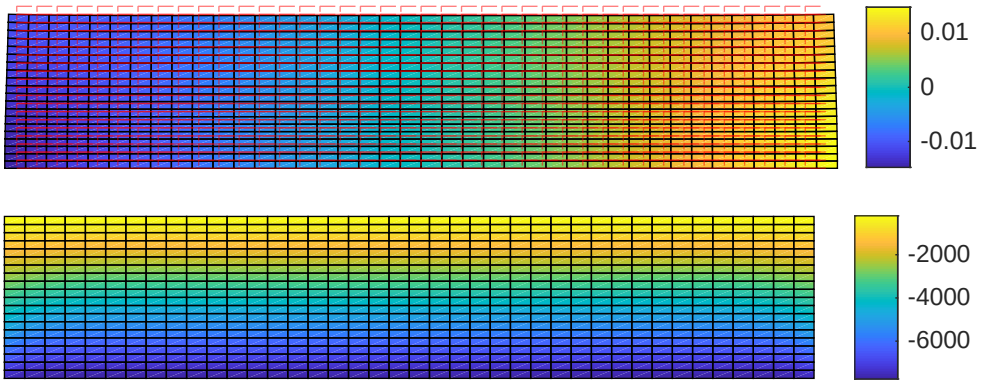


Figure 7. Exemplary results at $t = 1200\,\mathrm{s}$ under self weight: (top) horizontal displacement u_x in [m]; (bottom) normal stress σ_{yy} in [Pa]

For each sample the resulting stresses and strains have been calculated from the simulated nodal displacements. The histogram of the maximum shear strain per sample is displayed in Fig. 8 for two different points in the interval space of t_c and c, displaying a significant difference.

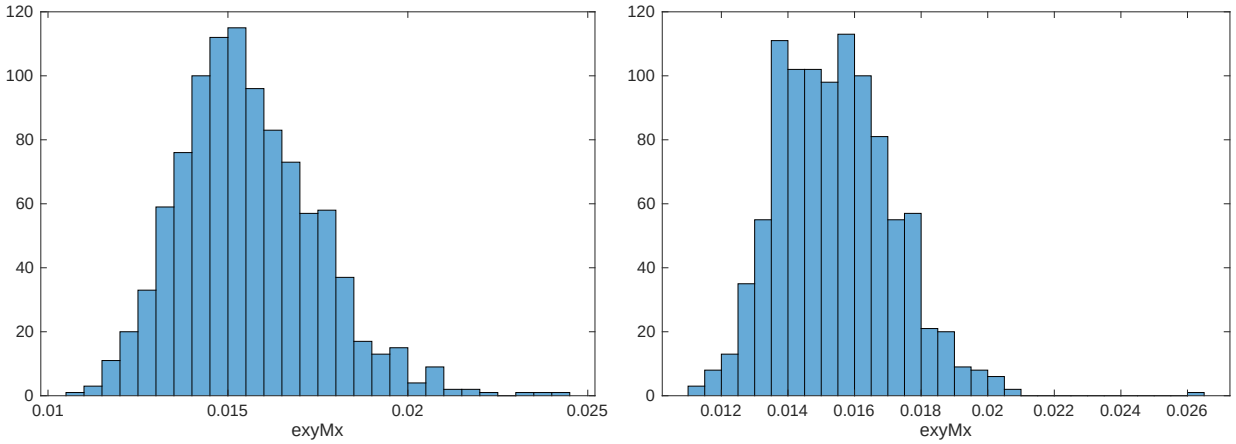


Figure 8. Shear strain histograms for (left) $t_\mathrm{c} = 72.6\,\mathrm{s}$, $c = 0.70$; (right) $t_\mathrm{c} = 40.1\,\mathrm{s}$, $c = 0.89$

Considering a simple limit state of a maximum tolerable deformation of 10% the failure probability for all 50 combinations of auto- and cross-correlation values was estimated. While the failure probability is strongly influenced by the correlation time t_c the cross-correlation c has no effect on the failure probability, cf. Fig. 9

Further, a polynomial-based regression is used to predict the failure probability estimates at the boundaries of the interval, especially for t_c^i. Finally, following the idea of the scheme in Fig. 3 a failure probability interval can be determined as $p_F^\mathrm{i} = [0.23; 0.34]$.

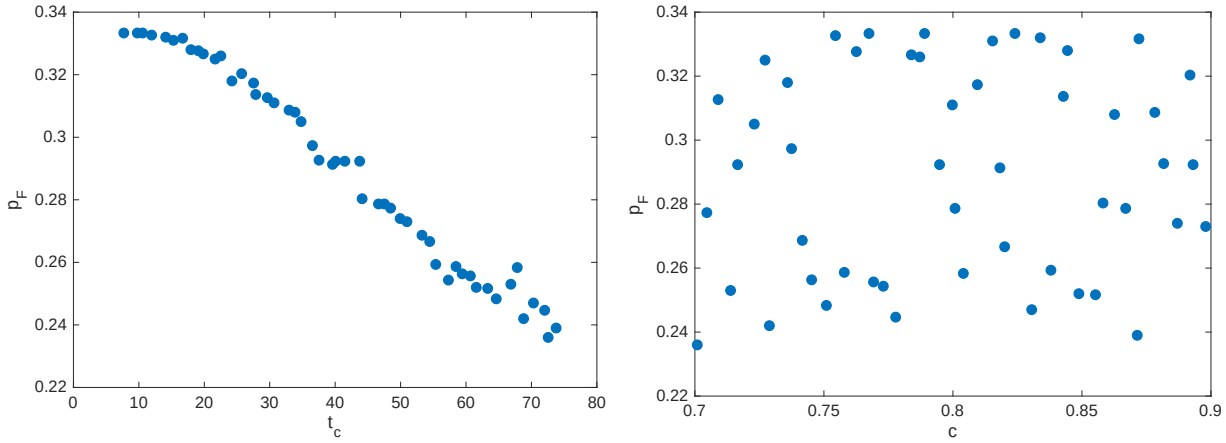


Figure 9. Failure probability p_F as a function of (left) auto-correlation time t_c and (right) cross-correlation c

4. Conclusion

The proposed approach considers an interval probability based random process description to take the uncertain nature of 3D concrete printing material parameters into account. A finite element model utilizing a pre-defined mesh is used to efficiently simulate the time-dependent printing process in a layer-by-layer fashion. For a simple 2D example of a $1\,\mathrm{m} \times 0.2\,\mathrm{m}$ wall, lognormally distributed random processes modelling the Young's modulus and the compressive strength of the fresh, pumped concrete is mapped to a the initial values of the time-dependent stiffness and compressive strength. The statistical properties of the system responses show a strong dependence on the auto-correlation time, which is described by an interval. A reliability analysis limited by the maximal tolerable deformation emphasis the dependence of the failure probability on such vaguely know parameters, which suggests that a realistic, polymorphic uncertainty model is appropriate. Further studies will include additional material and process parameters in the uncertainty model. The effect of the cross-correlation parameter on an stress-based failure mode will be investigated.

Acknowledgements

The demonstrated results and research activities arise from the project "Polymorphic uncertainty modeling of a 3D printing process of concrete" (GO 3270/1-1, KO 1267/24-1, LA 2869/11-1), as part of Priority Program SPP 1886. The funding by the German Research Foundation (DFG) is gratefully acknowledged.

References

Buswell, R. A., Leal de Silva, W. R., Jones, S. Z., and Dirrenberger, J. (2018). 3d printing using concrete extrusion: A roadmap for research. *Cement and Concrete Research*, 112:37–49.

Gosselin, C., Duballet, R., Roux, P., Gaudillière, N., Dirrenberger, J., and Morel, P. (2016). Large-scale 3d printing of ultra-high performance concrete – a new processing route for architects and builders. *Materials & Design*, 100:102–109.

Hager, I., Golonka, A., and Putanowicz, R. (2016). 3d printing of buildings and building components as the future of sustainable construction? *Procedia Engineering*, 151:292–299.

Le, T. T., Austin, S. A., Lim, S., Buswell, R. A., Gibb, A., and Thorpe, T. (2012). Mix design and fresh properties for high-performance printing concrete. *Materials and Structures*, 45:1221–1232.

Panda, B., Chandra Paul, S., and Jen Tan, M. (2017). Anisotropic mechanical performance of 3d printed fiber reinforced sustainable construction material. *Materials Letters*, 209:146–149.

Schietzold, F. N., Schmidt, A., Dannert, M. M., Fau, A., Fleury, R. M. N., Graf, W., Kaliske, M., Könke, C., Lahmer, T., and Nackenhorst, U. (2019). Development of fuzzy probability based random fields for the numerical structural design. *Surveys for Applied Mathematics and Mechanics*, 42(1):e201900004.

Suiker, A. S. J. (2018). Mechanical performance of wall structures in 3d printing processes: Theory, design tools and experiments. *International Journal of Mechanical Sciences*, 137:145–170.

Vořechovský, M. (2008). Simulation of simply cross correlated random fields by series expansion methods. *Structural Safety*, 30(4):337–363.

Wangler, T., Lloret, E., Reiter, L., Hack, N., Gramazio, F., Kohler, M., Bernhard, M., Dillenburger, B., Buchli, J., Roussel, N., and Flatt, R. (2016). Digital concrete: Opportunities and challenges. *RILEM Technical Letters*, 1:67–75.

Wolfs, R. J. M., Bos, F. P., and Salet, T. A. M. (2018). Early age mechanical behaviour of 3d printed concrete: Numerical modelling and experimental testing. *Cement and Concrete Research*, 106:103–116.

Wu, P., Wang, J., and Wang, X. (2016). A critical review of the use of 3-d printing in the construction industry. *Automation in Construction*, 68:21–31.

Homogenization of 3D Concrete Microstructures based on CT image reconstruction

George Stefanou[1,*], Dimitrios Savvas[1] and Panagiotis Metsis[1]

[1] *Institute of Structural Analysis & Dynamics of Structures,Department of Civil Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece, *gstefanou@civil.auth.gr, dimitriosavvas@yahoo.gr, panos@metsis.gr*

Abstract. The purpose of this paper is to determine the random spatially varying elastic properties of concrete at various scales taking into account its highly heterogeneous microstructure. The reconstruction of concrete microstructure is based on computed tomography (CT) images of a cubic concrete specimen. The variability of local volume fraction of the constituents (air, cement paste and aggregates) is quantified and mesoscale random fields of the elasticity tensor are computed from a number of statistical volume elements obtained by implementing the moving window technique on the specimen along with computational homogenization. Based on the statistical characteristics of the mesoscale random fields, useful conclusions are derived regarding the effect of microstructure on the mechanical behavior of concrete.

Keywords: Concrete, CT images, Reconstruction, 3D FEM models, Homogenization, Mesoscale random fields

1. Introduction

The macroscopic mechanical properties of heterogeneous materials such as composites and concrete are significantly influenced by their underlying microstructure and can be efficiently determined using numerical homogenization techniques. The application of homogenization methods in the analysis of multi-phase materials is much more advantageous over direct simulation of the full microstructure in terms of required computational resources. The key issue in homogenization methods is the linking of micromechanical characteristics with the random variation of material properties at higher scales, which is usually established using Hill's macro-homogeneity condition (Hill, 1963). In this framework, it is necessary to identify a representative volume element (RVE) over which a fine-scale boundary value problem is solved. A possible way to achieve this is through computational convergence schemes with respect to specific apparent properties (Ostoja-Starzewski, 2006). In contrast to the RVE, which is characterized by deterministic material properties, the spatial variation of the elasticity tensor at lower scales is quantified by random fields computed on mesoscale models-statistical volume elements (SVEs) (Savvas et al., 2016a).

The scope of this paper is to determine the elastic properties of concrete at various scales taking into account the highly heterogeneous microstructure of this material which is responsible for its complex mechanical behavior. There is a wide variety of methods in the literature for the reconstruction of concrete microstructure mostly based on image processing and/or morphological

© 2021 *by G. Stefanou, D. Savvas, P. Metsis.*

random fields, e.g. (Huang et al., 2016; Huang and Peng, 2018). The reconstruction is based herein on computed tomography (CT) images of a cubic concrete specimen. The variability of local volume fraction of the constituents (air, cement paste and aggregates) is quantified first. Mesoscale random fields of the elasticity tensor are computed subsequently from a number of SVEs obtained by implementing the moving window technique on the concrete specimen along with computational homogenization. Based on the statistical characteristics of the random fields, useful conclusions are derived regarding the effect of microstructure on the mechanical behavior of concrete.

The paper is organized as follows: In Section 2, the process of reconstructing concrete microstructure based on CT images is described. Section 3 presents the framework of computational homogenization used to determine the random spatially varying apparent elasticity tensor of 3D concrete microstructures. The numerical results obtained using the proposed approach are discussed in the next section followed by some concluding remarks in Section 5.

2. Reconstruction of Concrete Microstructure based on CT Images

2.1. CT Images

Computed Tomography or "CT" has its origins in the medical field and refers to a computerized X-ray imaging procedure in which a patient is struck by a narrow X-ray beam that rotates at a certain speed around the patient's body. Although the CT procedure has been applied mainly for medical purposes, the gradual increase in available CT equipment and the improvement of the cost factors have led to other utilizations, such as the exploration of the microstructure of heterogeneous materials.

When the X-ray beam passes through a sample, some of the X-ray radiation is absorbed and scattered (with intensity I_o), while the rest penetrates through the sample (with intensity I), as shown in Figure 1.

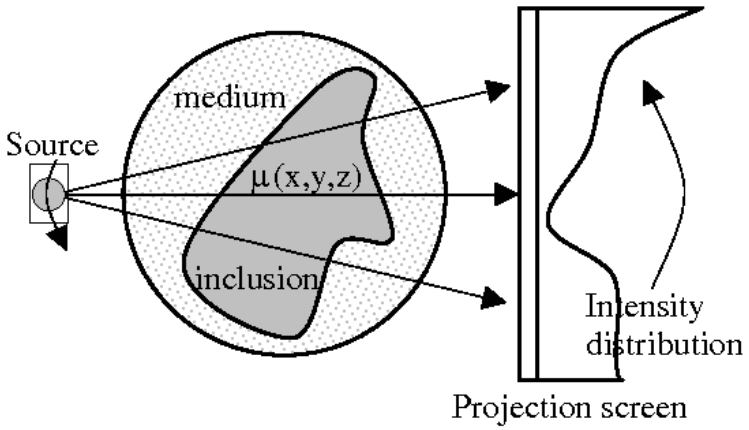


Figure 1. Intensity Distribution of an X-ray array (Daigle et al., 2005)

The radiodensity of the CT scan can be described by the Hounsfield unit (HU) scale. HU scale is a quantitative scale that reflects the linear transformation of the original linear attenuation coefficient measurement (μ) so as the radiodensity of distilled water at standard pressure and temperature is defined as zero HU, while the radiodensity of air, at the same conditions, is defined as -1000 HU. The linear transformation used to convert the attenuation values μ to Hounsfield units HU on each pixel of the CT image is as follows:

$$HU = \mu * slope + intercept \tag{1}$$

where the values of the rescale slope and intercept parameters are specific to the CT scanner system used. Their values can be obtained from the extracted DICOM (Digital Imaging and Communications in Medicine) files.

2.2. Reconstruction of concrete microstructure

The reconstruction of concrete microstructure is based on data obtained from the CT scan. The specimen is cut in 2D slices (images) with their corresponding resolution (Figure 2). Each pixel carries, beside its geometric properties, the attenuation value μ that is converted to the respective HU value using Eq. (1). Then, by utilizing voxels (the generalization of pixels on a regularly spaced, three-dimensional grid), the voxel-based FE method is employed. Using the moving window technique (Baxter et al., 2001), random mesoscale models or SVEs are created, which are discretized by a structured FE mesh based on voxels geometry. More details about the procedure of reconstruction are provided in Section 4.

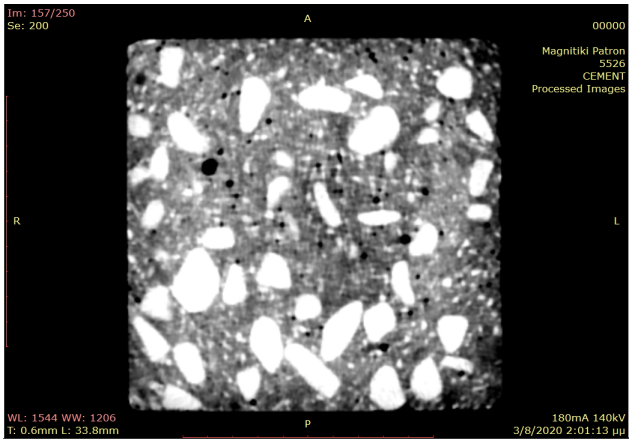


Figure 2. A 2D CT image of concrete.

3. Computational Homogenization

According to 1983, the homogenization process is based on the fundamental assumption of statistical homogeneity of the heterogeneous medium. This implies that all statistical properties of the state variables are the same at any material point and thus an RVE can be identified. While in case of composites with periodic or nearly periodic geometry the RVE is explicitly defined, in case of spatial randomness, the RVE needs to be sought using computational methods (Kanit et al., 2003; Zeman and Sejnoha, 2007; Zohdi and Wriggers, 2008; Wimmer et al., 2016; Savvas et al., 2016a; Savvas et al., 2016b).

According to 1963, the existence of the RVE postulates separation of scales in the form:

$$d \ll L \ll L_{macro} \tag{2}$$

In the above inequality, the microscale parameter d denotes a characteristic size of the fillers, e.g. their diameter in case of circular inclusions, the mesoscale parameter L denotes the size of the volume element and the macroscale parameter L_{macro} denotes the characteristic length over which the macroscopic loading varies in space, or in the case of complete scale separation, the size of the macroscopic homogeneous medium.

The statistical characteristics of the apparent elasticity tensor of concrete are computed by analyzing a set of random mesoscale models or SVEs. These are extracted by implementing the moving window technique on a cubic concrete specimen reconstructed based on a set of 2D CT images (see Section 2). Figure 3 depicts the cubic concrete specimen segmented into a large number of non-overlapping SVE models along with a detail of the voxel based FE mesh of a specific SVE. Each SVE model is characterized by the non-dimensional parameter $\delta = L/d$ with $\delta \in [1, \infty]$, which denotes the scale factor. Note that the length scale of the SVE is smaller than the one of the corresponding RVE. For each placement of the moving window, the apparent material properties are computed through homogenization of the material in the microscale. The resulting mesoscale properties and corresponding moving window centres are used to build the related random field model, as shown in Section 4.

3.1. Computation of the Random Apparent Elasticity Tensor of 3D Concrete Microstructures

The stress and strain fields, $\boldsymbol{\sigma}(\omega, \boldsymbol{x})$ and $\boldsymbol{\varepsilon}(\omega, \boldsymbol{x})$ respectively, developed on the mesoscale realization of concrete $B_\delta(\omega)$ can be expressed as a superposition of means ($\bar{\boldsymbol{\sigma}}$ and $\bar{\boldsymbol{\varepsilon}}$) and of zero-mean fluctuations ($\boldsymbol{\sigma'}$ and $\boldsymbol{\varepsilon'}$) as follows:

$$\boldsymbol{\sigma}(\omega, \boldsymbol{x}) = \bar{\boldsymbol{\sigma}} + \boldsymbol{\sigma'}(\omega, \boldsymbol{x}) \quad , \quad \boldsymbol{\varepsilon}(\omega, \boldsymbol{x}) = \bar{\boldsymbol{\varepsilon}} + \boldsymbol{\varepsilon'}(\omega, \boldsymbol{x}) \tag{3}$$

The means of stress and strain tensors at some point $\boldsymbol{X}$ of the macro-continuum are computed as volume averages over $B_\delta(\omega)$ in the form (Hill, 1963):

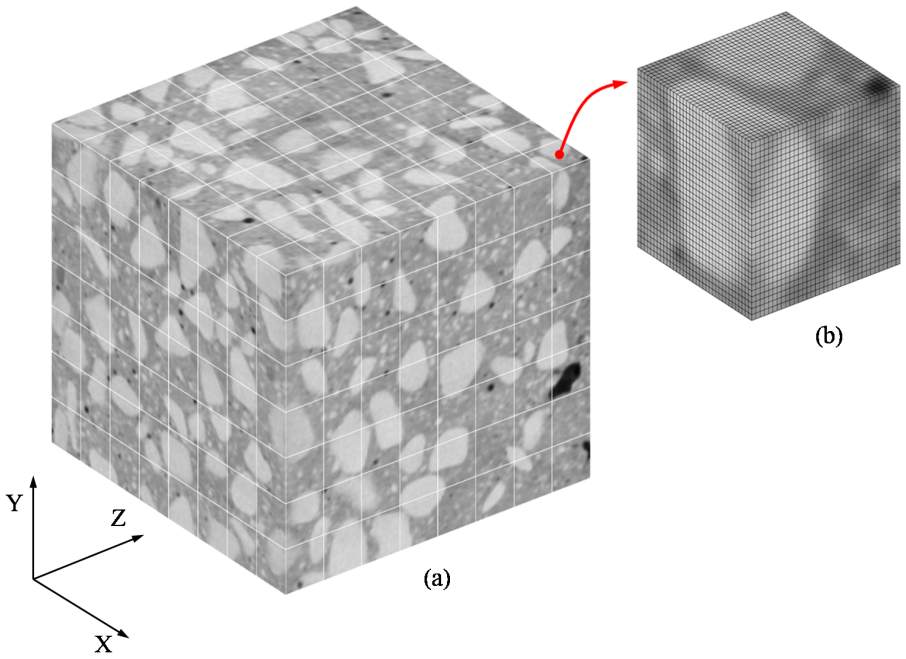


Figure 3. a) Schematic of 3D moving window technique on a cubic concrete specimen reconstructed using 2D CT images, b) Detail of the voxel based FE mesh of a specific SVE model.

$$\bar{\boldsymbol{\sigma}}(\omega, \boldsymbol{X}) = \frac{1}{V_\delta} \int\limits_{B_\delta(\omega)} \boldsymbol{\sigma}\left(\omega, \boldsymbol{x}\right) \mathrm{d}V_\delta, \quad \bar{\boldsymbol{\varepsilon}}(\omega, \boldsymbol{X}) = \frac{1}{V_\delta} \int\limits_{B_\delta(\omega)} \boldsymbol{\varepsilon}\left(\omega, \boldsymbol{x}\right) \mathrm{d}V_\delta \tag{4}$$

with $V_\delta = \int\limits_{B_\delta(\omega)} \mathrm{d}V_\delta$.

Also the volume average of the strain energy can be calculated as:

$$\bar{U} = \frac{1}{2V_\delta} \int\limits_{B_\delta(\omega)} \boldsymbol{\sigma}\left(\omega, \boldsymbol{x}\right) : \boldsymbol{\varepsilon}\left(\omega, \boldsymbol{x}\right) \mathrm{d}V_\delta = \frac{1}{2}\overline{\boldsymbol{\sigma} : \boldsymbol{\varepsilon}} = \frac{1}{2}\bar{\boldsymbol{\sigma}} : \bar{\boldsymbol{\varepsilon}} + \frac{1}{2}\overline{\boldsymbol{\sigma}' : \boldsymbol{\varepsilon}'} \tag{5}$$

Note that for an unbounded space domain ($\delta \to \infty$) the fluctuation terms in Eq. (5) become negligible $\left(\overline{\boldsymbol{\sigma}' : \boldsymbol{\varepsilon}'} = 0\right)$ and thus the following equation holds:

$$\overline{\boldsymbol{\sigma} : \boldsymbol{\varepsilon}} = \bar{\boldsymbol{\sigma}} : \bar{\boldsymbol{\varepsilon}} \tag{6}$$

which is known as Hill's condition. However, at a finite mesoscale size, Hill's condition is valid provided that the following constraint is satisfied (Hazanov and Huet, 1994):

$$\int_{\partial B_\delta} (t - \bar{\boldsymbol{\sigma}} \cdot n) \cdot (u - \bar{\boldsymbol{\varepsilon}} \cdot \boldsymbol{x}) \, \mathrm{d}S = 0 \tag{7}$$

The constraint equation (7) is a priori satisfied by appropriate boundary conditions such as uniform strains (Dirichlet), uniform stresses (Neumann), uniform orthogonal-mixed boundary conditions and periodic boundary conditions. Here, we focus on the case of uniform strains. We now discuss the computational homogenization procedure for this case, following 2002.

In general the 3D elasticity tensor $\bar{\mathbb{C}}$, which is of 4^{th} order, consists of $3^4 = 81$ components. These are reduced to 21 independent components due to symmetries in stress and strain tensors. Thus, the macroscopic linear constitutive relation for an anisotropic material is as follows:

$$\bar{\boldsymbol{\sigma}} = \bar{\mathbb{C}}\bar{\boldsymbol{\varepsilon}} \Rightarrow \begin{bmatrix} \bar{\sigma}_{11} \\ \bar{\sigma}_{22} \\ \bar{\sigma}_{33} \\ \bar{\sigma}_{12} \\ \bar{\sigma}_{23} \\ \bar{\sigma}_{31} \end{bmatrix} = \begin{bmatrix} \bar{C}_{1111} & & & & & \text{sym} \\ \bar{C}_{2211} & \bar{C}_{2222} & & & & \\ \bar{C}_{3311} & \bar{C}_{3322} & \bar{C}_{3333} & & & \\ \bar{C}_{1211} & \bar{C}_{1222} & \bar{C}_{1233} & \bar{C}_{1212} & & \\ \bar{C}_{2311} & \bar{C}_{2322} & \bar{C}_{2333} & \bar{C}_{2312} & \bar{C}_{2323} & \\ \bar{C}_{3111} & \bar{C}_{3122} & \bar{C}_{3133} & \bar{C}_{3112} & \bar{C}_{3123} & \bar{C}_{3131} \end{bmatrix} \begin{bmatrix} \bar{\varepsilon}_{11} \\ \bar{\varepsilon}_{22} \\ \bar{\varepsilon}_{33} \\ \bar{\varepsilon}_{12} \\ \bar{\varepsilon}_{23} \\ \bar{\varepsilon}_{31} \end{bmatrix} \tag{8}$$

The unknown macroscopic elasticity components $\bar{C}_{ijkl}$, with $i, j, k, l = 1, 2, 3$, are equal to the macroscopic stresses $\bar{\sigma}_{ij}$ which are derived by the solution of six independent kinematic uniform boundary value problems (BVPs). In each BVP the displacement boundary conditions are obtained by a prescribed uniform strain tensor as shown below:

$$\bar{\boldsymbol{\varepsilon}} = \left\{ \begin{bmatrix} \bar{\varepsilon}_{11} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \bar{\varepsilon}_{22} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ \bar{\varepsilon}_{33} \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ \bar{\varepsilon}_{12} \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \bar{\varepsilon}_{23} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \bar{\varepsilon}_{31} \end{bmatrix} \right\} \tag{9}$$

The imposed displacements on the boundary nodes of each SVE are calculated through the prescribed uniform strain tensor using the following equation:

$$\boldsymbol{u}_b = D_b^T \bar{\boldsymbol{\varepsilon}} \tag{10}$$

where D_b is a geometric matrix which depends on the coordinates of the boundary node b and is defined as:

$$D_b = \frac{1}{2}\begin{bmatrix} 2x & 0 & 0 \\ 0 & 2y & 0 \\ 0 & 0 & 2z \\ y & x & 0 \\ 0 & z & y \\ z & 0 & x \end{bmatrix} \quad \text{with} \quad (x, y, z) \in \boldsymbol{x} \tag{11}$$

The BVP which has to be solved is as follows:

$$\begin{bmatrix} \boldsymbol{K}_{ii} & \boldsymbol{K}_{ib} \\ \boldsymbol{K}_{bi} & \boldsymbol{K}_{bb} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_i \\ \boldsymbol{u}_b \end{bmatrix} = \begin{bmatrix} \boldsymbol{f}_i \\ \boldsymbol{f}_b \end{bmatrix} \tag{12}$$

where the stiffness matrix $\boldsymbol{K}$ of the FE model is rearranged into four sub-matrices associated with interior nodes i and boundary nodes b. The macroscopic stress tensor is then calculated as a volume average by:

$$\bar{\boldsymbol{\sigma}} = \frac{1}{V}\mathbb{D}\boldsymbol{f}_b \tag{13}$$

where $\boldsymbol{f}_b = \boldsymbol{K}_{bb}\boldsymbol{u}_b + \boldsymbol{K}_{bi}\boldsymbol{u}_i$ is the vector containing the computed reaction forces on the boundary nodes, $\boldsymbol{u}_i = \boldsymbol{K}_{ii}^{-1}(\boldsymbol{f}_i - \boldsymbol{K}_{ib}\boldsymbol{u}_b)$ is the vector containing the calculated displacements of the interior nodes and $\mathbb{D} = [D_1\ D_2\ ...\ D_M]$ with M the total number of boundary nodes b.

4. Numerical Results and Discussion

In this section, numerical results concerning the determination of the random fields of the homogenized elasticity tensor of the 3D concrete specimen are presented. The reconstruction of the microstructure of concrete is based on DICOM data obtained by a medical CT scan conducted at Magnitiki Patron S.A. The equipment used was a LightSpeed VCT of GE Medical Systems with peak potential and current applied to the X-ray tube equal to 140 kV and 180 mA, respectively. The axial resolution of the scan was fixed at 0.625 mm resulting in 211 total slices, while the pixel resolution of the 2D image in each slice plane is $0.4883 \times 0.4883\ mm^2$, resulting in 281×281 pixels. Note that the dimensions of the scanned cubic concrete specimen was approximately $136.724 \times 136.724 \times 131.25\ mm^3$.

As shown in Figure 3, the size of each SVE corresponds to that of the selected 3D window in the moving window technique. In the context of FE simulation, the discretization of the SVEs into hexahedral solid elements is based on the voxels which are used to reconstruct the 3D geometry of the concrete specimen. Information about the pixels (e.g. location, spacing, attenuation values μ) and the slices (e.g. position, thickness) needed in order to construct the voxels and also to identify the constituent materials is included in the DICOM files provided by the CT scan system. In the FE mesh of the SVEs each hexahedral solid element corresponds to a voxel. The nodes of each

element correspond to pixels which are lying in two subsequent slice planes. The material assigned to each integration point of an element depends on the HU value of the integration point which is calculated by interpolating the nodal HU values using the shape functions of the element.

Figure 4 depicts the histogram (empirical distribution) of the Hounsfield units (HU) calculated by Eq. (1) with rescale slope and rescale intercept equal to 1 and -1024, respectively, for the CT equipment used. From this figure, the constituent materials (air, cement and aggregates) can be distinguished by defining specific ranges for their HU values. Based on the observed peaks of the histogram and after careful inspection of the CT images, the HU range can be selected as [-960, 800] for air, [800, 2000] for cement and [2000, 2974] for aggregates. Figure 4 depicts also a gray scale colorbar connected to the HU values of the constituent materials of concrete. The illustration of concrete specimen in Figure 3 is based on this color range.

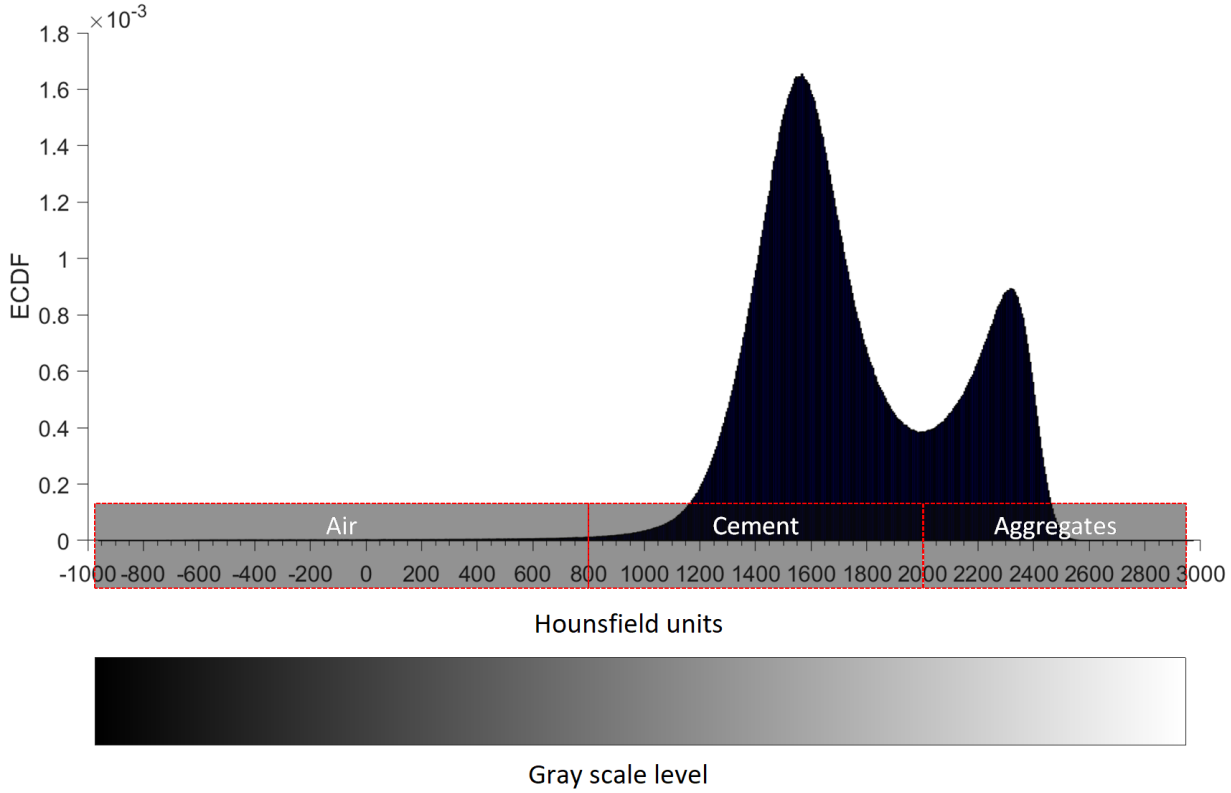


Figure 4. Empirical distribution of HU values derived from all CT images followed by the range of HU for concrete constituent materials.

Based on the aforementioned HU ranges, the local volume fraction (vf) variability of the constituents of concrete can be studied next. Figure 5 illustrates the computed 3D random fields along with the respective empirical distributions of the volume fraction of the constituent materials. For visualization purposes, Figures 5, 6 and 8 provide 2D contour plots of the random fields corresponding to the three orthogonal planes (XY, YZ and ZX) intersected at the center point

of the cubic concrete specimen. Note that, these results have been obtained by selecting the size of the moving window to be $10 \times 10 \times 10$ voxels and the moving step to be equal to that size in each direction so as non-overlapping SVEs to be extracted. For this specific choice of window size and due to the dimensions of the cubic concrete specimen, the total number of the extracted SVEs is 16464. The dispersion of aggregates within cement is not uniform. Also, some air was trapped inside the paste during the mixture process. These facts are obvious in Figure 5, where regions of material rich or poor in aggregates and air can be observed. The mean vf of the constituents in the concrete specimen is 0.5% air, 72.3% cement and 27.2% aggregates.

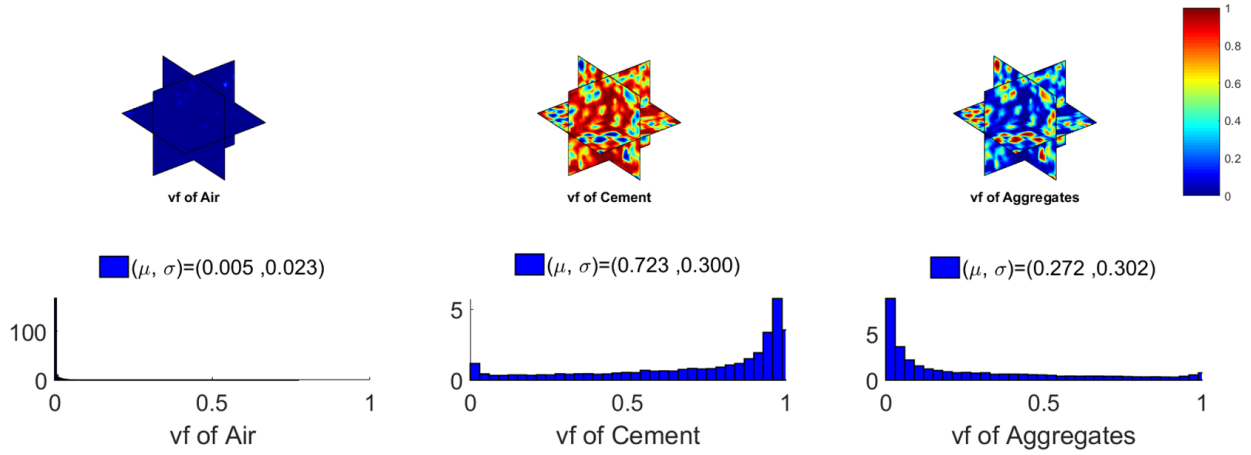


Figure 5. Random fields and histograms of vf of air, cement and aggregates.

Figures 6-9 depict the computed random fields along with the respective empirical distributions of the components of the apparent elasticity tensor (see Eq. (8)). These results have been obtained by implementing the homogenization method presented in Section 3. The constituent materials of concrete are considered linear elastic. Specifically, their Young's modulus and Poisson ratio [E, v] are assumed as [20 GPa, 0.2] for cement and [100 GPa, 0.2] for aggregates. Note that although air is not an elastic material, it can be modeled by setting its Young's modulus equal to 0.2 GPa (1% of the Young's modulus of cement) and its Poisson ratio equal to 0.45 (almost incompressible material). From the calculated mean values of the elasticity components it can be deduced that concrete can be considered as isotropic material in an average sense (Figures 7 and 9). Another observation is that the coefficient of variation (COV) of the axial and shear components of the apparent elasticity tensor is higher than 60%, which means that morphological uncertainty leads to significant random spatial variation of the mechanical properties of concrete.

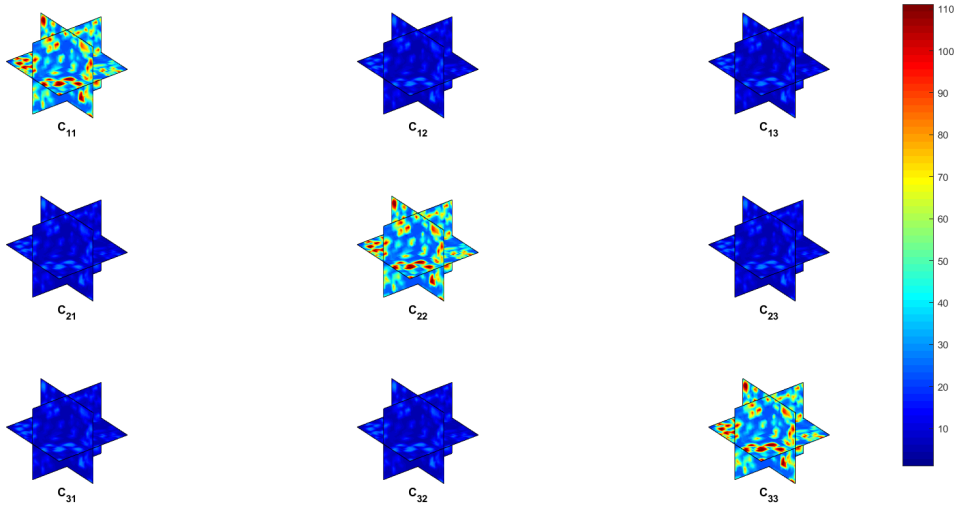


Figure 6. Random fields of $\bar{C}_{iijj}$ (correspondence of subscripts: $1 \to 11$, $2 \to 22$, $3 \to 33$).

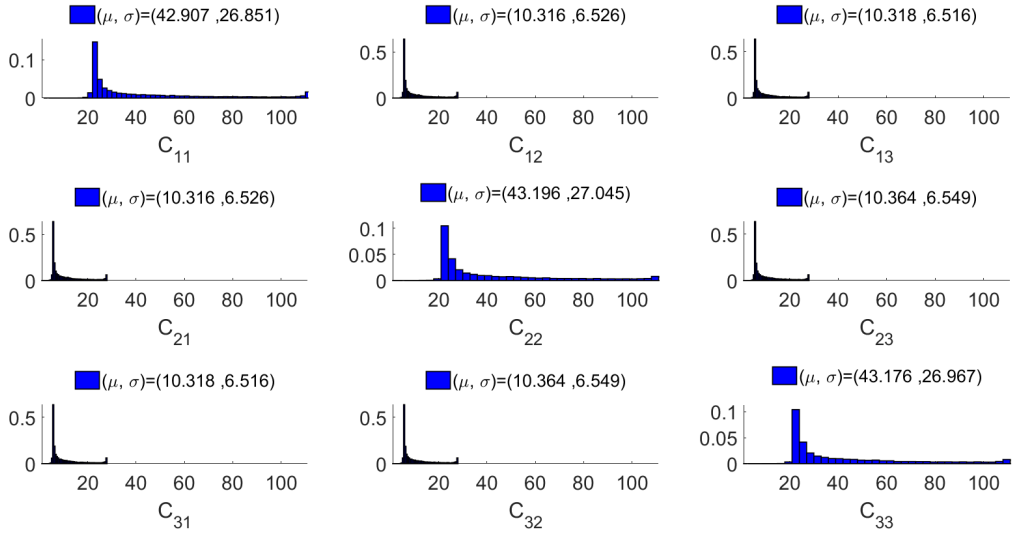


Figure 7. Histograms of $\bar{C}_{iijj}$ (correspondence of subscripts: $1 \to 11$, $2 \to 22$, $3 \to 33$).

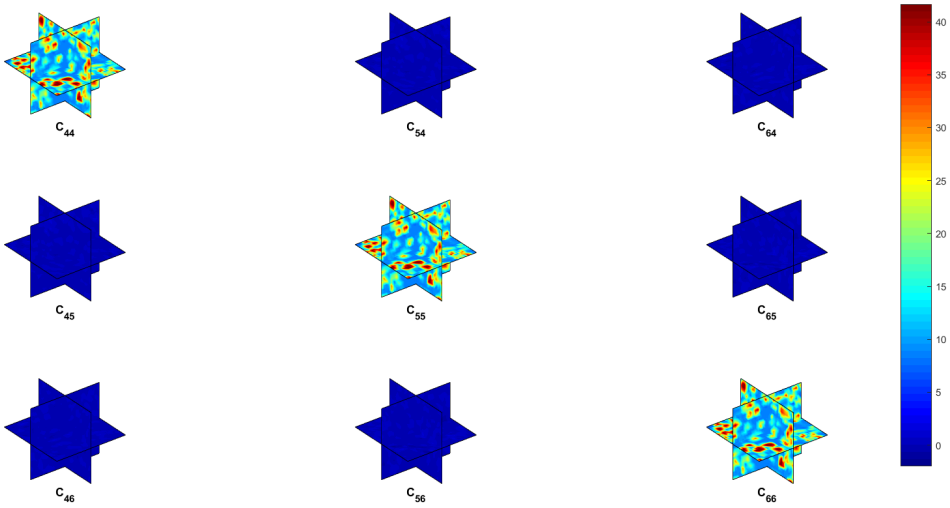


Figure 8. Random fields of $\bar{C}_{ijij}$ (correspondence of subscripts: $4 \to 12$, $5 \to 23$, $6 \to 31$).

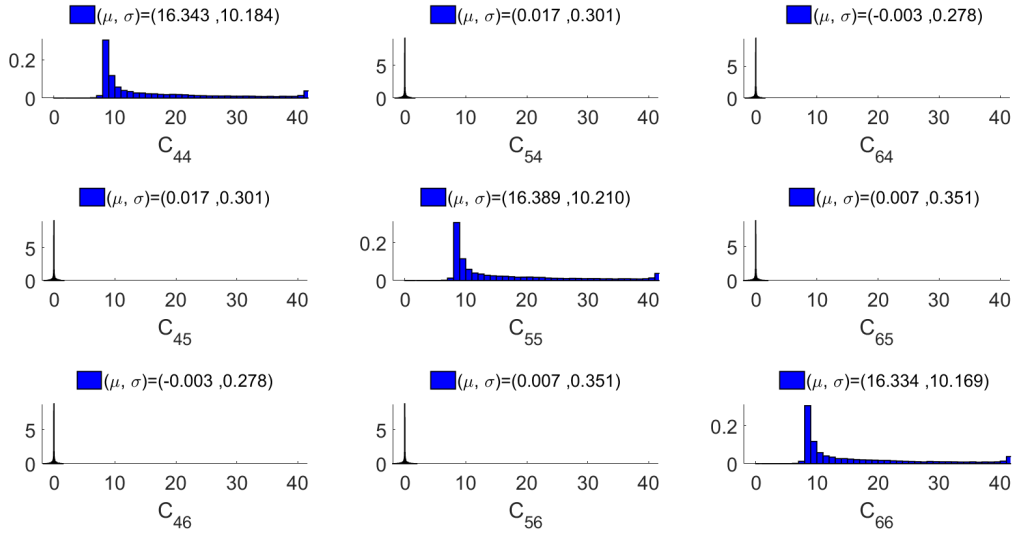


Figure 9. Histograms of $\bar{C}_{ijij}$ (correspondence of subscripts: $4 \to 12$, $5 \to 23$, $6 \to 31$).

5. Conclusions

In this paper, the random spatially varying apparent elasticity tensor of 3D concrete microstructures has been computed using CT images of a cubic specimen for the reconstruction of concrete microstructure and computational homogenization. A large variability of the local volume fraction of the constituents (air, cement paste and aggregates) has been observed. This was reflected on the high COV (>60%) of the axial and shear components of the apparent elasticity tensor, meaning that microstructural uncertainty leads to substantial random spatial variation of the mechanical properties of concrete, which has to be taken into account for a safe design of concrete structures.

Acknowledgements

This research is co-financed by Greece and the European Union (European Social Fund-ESF) through the Operational Programme "Human Resources Development, Education and Lifelong Learning 2014-2020" in the context of the project "Stochastic multiscale modeling of concrete structures reinforced with graphene nanoparticles - Investigation of the effect of nano-inclusions on the mechanical properties of concrete" (MIS 5047867). The authors wish also to acknowledge Magnitiki Patron S.A. (https://www.magnitikipatron.com/) for conducting the CT scans on the concrete specimen and providing all the necessary DICOM files in order to accomplish the present research work.







References

Hill, R. Elastic properties of reinforced solids: some theoretical principles. *Journal of the Mechanics and Physics of Solids*, 11(5):357–372, 1963.

Ostoja-Starzewski, M. Material spatial randomness: From statistical to representative volume element. *Probabilistic engineering mechanics*, 21(2):112–132, 2006.

Savvas, D., G. Stefanou, and M. Papadrakakis. Determination of RVE size for random composites with local volume fraction variation. *Computer Methods in Applied Mechanics and Engineering*, 305:340–358, 2016a.

Huang, Y., D. Yan, Z. Yang, and G. Liu. 2D and 3D homogenization and fracture analysis of concrete based on in-situ X-ray Computed Tomography images and Monte Carlo simulations. *Engineering Fracture Mechanics*, 163:37–54, 2016.

Huang, J. and Q. Peng. A fast algorithm for multifield representation and multiscale simulation of high-quality 3D stochastic aggregate microstructures by concurrent coupling of stationary Gaussian and fractional Brownian random fields. *International Journal for Numerical Methods in Engineering*, 115(3):328–357, 2018.

Daigle, M., D. Fratta, and LB. Wang. Ultrasonic and X-ray tomographic imaging of highly contrasting inclusions in concrete specimens. *Site Characterization and Modeling*, 1–12, 2005.

Baxter, SC., M. Imran Hossain, and L. Graham. Micromechanics based random material property fields for particulate reinforced composites. *International journal of solids and structures*, 38:9209–9220, 2001.

Hashin, Z. Analysis of composite materials: A survey. *Journal of Applied Mechanics*, 50(2):481–505, 1983.

Kanit, T., S. Forest, I. Galliet, V. Mounoury and D. Jeulin. Determination of the size of the representative volume element for random composites: statistical and numerical approach. *International journal of solids and structures*, 40(13):3647–3679, 2003.

Zeman, J., and M. Sejnoha. From random microstructures to representative volume elements. *Modelling and Simulation in Materials Science and Engineering*, 15(4):325–335, 2007.

Zohdi, T. I., and P. Wriggers. An Introduction to Computational Micromechanics, 2nd Edition, Lecture Notes in Applied and Computational Mechanics. *Springer, Heidelberg*, 20, 2008.

Wimmer, J., B. Stier, J.-W. Simon, and S. Reese. Computational homogenisation from a 3D finite element model of asphalt concrete–linear elastic computations. *Finite Elements in Analysis and Design*, 110:43–57, 2016.

Savvas, D., G. Stefanou, V. Papadopoulos, and M. Papadrakakis. Effect of waviness and orientation of carbon nanotubes on random apparent material properties and RVE size of CNT reinforced composites. *Composite Structures*, 152:870–882, 2016b.

Hazanov, S., and C. Huet. Order relationships for boundary conditions effect in heterogeneous bodies smaller than the representative volume. *Journal of the Mechanics and Physics of Solids*, 42(12):1995–2011, 1994.

Miehe, C., and A. Koch. Computational micro-to-macro transitions of discretized microstructures undergoing small strains. *Archive of Applied Mechanics*, 72(4-5):300–317, 2002.

Elucidating appealing features of differentiable autocorrelation functions: a study on the modified exponential kernel

Matthias G.R. Faes[a,b], Matteo Broggi[b], Pol D. Spanos[c], Michael Beer[b,d,e]

[a] *KU Leuven, department of Mechanical Engineering, matthias.faes[at]kuleuven.be*
[b] *Leibniz Universität Hannover, Institute for Risk and Reliability, {broggi,beer}[at]irz.uni-hannover.de*

[c] *Rice University, 232 Mechanical Engineering Building, Houston, TX, United States*
[d] *Institute for Risk and Uncertainty and School of Engineering, University of Liverpool, Peach Street, Liverpool L69 7ZF, UK*

[e] *International Joint Research Center for Engineering Reliability and Stochastic Mechanics, Tongji University, 1239 Siping Road, Shanghai 200092, P.R. China*

Abstract. Research on stochastic processes in recent decades has pointed out that, in the context of modelling spatial or temporal uncertainties, auto-correlation functions that are differentiable at the origin have advantages over functions that are not differentiable. For instance, the non-differentiability of e.g., single exponential autocorrelation functions yields non-smooth sample paths. Such sample paths might not be physically possible or might yield issues when used as random parameters in partial differential equations (such as encountered in e.g., mechanical equilibrium problems). Further, it is known that due to the non-differentiability of certain auto-correlation functions, more terms are required in the stochastic series expansion representations. This makes these representations less efficient from a computational standpoint.

This paper elucidates some additional appealing features of autocorrelation functions which are differentiable at the origin. Further, it focusses on deepening the argumentations in favor of these functions already available in literature. Specifically, focus is placed on single exponential, modified exponential and squared exponential autocorrelation functions. To start, it is shown that the power spectrum of differentiable kernels converges faster to zero wiht increasing frequency as compared to non-differentiable ones. This property allows capturing the same amount of energy of the spectrum with a smaller cut-off frequency, and hence, less stochastic terms. Further, this point is examined with reqards to the Karhunen-Loève series expansion and first and second order Markov processes, generated via auto-regressive representations.

Keywords: stochastic process, exponential kernel, auto-correlation function

1. Introduction

Many of the loads on engineering components, structures and systems, as well as the constitutive properties of these assets exhibit a stochastic nature. This notion is based on the observation that these quantities exhibit apparent variability in time and/or space. In this context, the theoretical framework of stochastic process, and by extension, random fields (Vanmarcke, 1983) has proven

©2021 *by M.G.R. Faes, M. Broggi, P.D. Spanos, and M. Beer.*

to be an excellent means for capturing inherent (aleatory) uncertainty (Shinozuka and Deodatis, 1991). Stochastic processes represent in essence jointly distributed random variables whose correlation function depends on time and/or space. Throughout literature, efficient methods have been introduced to effectively and accurately sample from these potentially high-dimensional joint distributions.

Typically, the auto-correlation of such stochastic process is governed by a pre-defined auto-correlation function (also often referred to as 'kernel'). This function describes the correlation between two random variables in the stochastic process as a function of the distance in time/space between them. Alternatively, in case the process is stationary, the autocorrelation is governed by the relative distance between two points. The generation of samples from these stochastic processes is usually performed using the well-known Karhunen-Loève series expansion (Betz et al., 2014) or the spectral representation method as introduced by (Rice, 1954), (Borgman, 1969) and later by (Shinozuka and Deodatis, 1991). Also extended methods on these techniques have been introduced. Examples of such methods include the Stochastic Harmonic Function representation by (Chen et al., 2013).

The single exponential autocorrelation kernel has been used extensively in engineering applications. For instance, (Desplentere, Verpoest and Lomov, 2009) applied this kernel to represent a stochastic process to model the permeability in resin transfer molding simulation. Further, in (Cao and Wang, 2014), Bayesian updating in a geo-technical context was done based on a single exponential kernel. The single exponential kernel is often selected due to the availability of analytical solutions to the eigenvalue problem corresponding to the Karhunen-Loève series expansion. However, the appropriate selection of the auto-correlation function is of large importance for the correct modelling and simulation of the phenomenon or property under consideration.

More relevantly to the problem discussed herein, (Ching and Phoon, 2018) studied the effect of the auto-correlation function on the probability of failure in several geo-technical examples. In this work, the authors showed that the smoothness of a sample path had a significant effect on the probability of failure. This is particularly true when no spatial averaging is present in the considered problem capable of smoothing out local variations. Further, it was shown that the sample path smoothness depends on the functional form of the auto-correlation function, and, more precisely, on the differentiability of such function at zero-lag. Thus, this work clearly illustrated that the auto-correlation function ends up affecting the probability of failure.

Furthermore in this context, (Wang et al., 2019) noticed that the application of the single exponential auto-correlation kernel tends to underestimate the failure risk in the unsaturated slope risk assessment. In addition, they pointed out how the differentiability of the auto-correlation function affects the truncation order of the Karhunen-Loève series expansion. Also in this regard, (Sudret and Der Kiureghian, 2000) showed that square exponential kernels are preferable over single exponential kernels, allowing for a highly exact series expansion representation with far fewer terms in the expansion. Similar observations were made by (Sclavounos, 2012) in the context of modelling stochastic ocean waves. In essence, they showed that fewer stochastic variables are required to describe stochastic processes with a narrow spectral bandwidth (i.e., spectra that converge relatively fast to zero when the frequency increases).

As a final note to this overview of the wide-spread application of single exponential kernels (versus less frequently used kernels), examination of the literature showed that non-differentiable

auto-correlation functions are used to represent a random 'stiffness' term in partial differential equations. From a mathematical standpoint, this is not correct since the time derivative of the corresponding sample paths is undefined. To address this, in (Spanos, Beer and Red-Horse, 2007) the "modified exponential kernel" was introduced to alleviate the shortcomings of the traditional single exponential kernel caused by non-differentiability. Further, this kernel maintains the enticing properties of the single exponential kernel, such as the availability of analytical solutions and the possibility to characterize quickly varying spatial or temporal phenomena. Also, this kernel is more efficient in terms of the required number of stochastic quantities as compared to the single exponential kernel. These properties are obtained by solving the issue of non-differentiability at zero-lag in the single exponential kernel, while maintaining its functional form to a great extent.

This paper builds upon the work that was presented in (Spanos, Beer and Red-Horse, 2007) in three ways. First, it provides additional explanations of why this particular kernel is highly effective by analytically deriving the energy error rate convergence in the frequency domain. Second, it compares the modified exponential kernel to another widely used autocorrelation kernel, namely the squared exponential kernel. Finally, it provides additional numerical evidence for the effectiveness of the modified exponential kernel in conjunction with the Karhunen-Loève series expansion, and auto-regressive model representations. The paper is structured as follows. Section 2 starts by introducing some basic concepts related to the modelling and simulation of stochastic processes. Section 3 discusses the convergence of the considered autocorrelation kernels both in the frequency and time domain and discusses the implications hereof for numerical analysis purposes. Finally, Section 4 briefly comprises the conclusions that can be drawn from this work.

2. Spectral stochastic process representation

2.1. Definition and stochastic properties

A finite-dimensional stochastic process $f(\mathbf{t},\theta)$ describes a set of correlated random variables $f(\theta)$ that are assigned to a countable number of locations $\mathbf{t} \in \Omega$ in the model domain $\Omega \subset \mathbb{R}^d$ with dimension $d \in \mathbb{N}$. Each such a random variable $f(\theta) : (\Theta, \varsigma, P) \mapsto \mathbb{R}$, with $\theta \in \Theta$ a coordinate in sample space Θ and ς the sigma-algebra, as such maps from a complete probability space to the real domain. This map holds as long as $f(\mathbf{t},\theta) \in \mathcal{L}^2(\Theta, P)$, with $L^2(\Theta, P)$ the Hilbert space of second-order random variables (i.e., finite variance). For a given event $\theta_i \in \Theta$, $f(t,\theta_i)$ is a realization of the stochastic process. A stochastic process is considered Gaussian if the distribution of $(f(\mathbf{t}_1,\theta), f(\mathbf{t}_2,\theta), \ldots, f(\mathbf{t}_n,\theta))$ is jointly Gaussian $\forall \mathbf{t} \in \Omega$.

Consider $f(t,\theta)$ to be a zero-mean one-dimensional univariate stochastic process (i.e, $\Omega \subset \mathbb{R}$) with constant variance over the domain and autocorrelation function $R_{ff}(\tau) : \Omega \times \Omega \mapsto [0,1]$ and $\tau \subset \Omega$ a lag parameter. The autocorrelation $R_{ff}(\tau)$ of such a stochastic process represents the correlation between two random variables $f(t,\theta)$ and $f(t+\tau,\theta)$, separated by a lag τ:

$$R_{ff}(\tau) = \frac{CV\left[f(t_i,\theta) f(t+\tau,\theta)\right]}{\sqrt{V\left[f(t,\theta)\right]}\sqrt{V\left[f(t+\tau,\theta)\right]}}, \tag{1}$$

with $CV[\cdot,\cdot]$ denoting an operator that returns the covariance and $V[\cdot]$ an operator returning the variance of the argument. In this regard, τ may represent a distance in time or space. In practical

applications, often analytical models for the autocorrelation are applied (Ching and Phoon, 2018), including single exponential $R^s_{ff}(\tau)$, squared exponential $R^{sq}_{ff}(\tau)$ and modified exponential $R^m_{ff}(\tau)$ autocorrelation functions. When considering a single exponential auto-correlation function, the correlation $R^s_{ff}(\tau)$ is expressed as

$$R^s_{ff}(\tau) = \sigma^2 \exp(-|\tau|/b). \tag{2}$$

Similarly, the auto-correlation for a modified exponential $R^m_{ff}(\tau)$ (Spanos, Beer and Red-Horse, 2007) and squared exponential $R^{sq}_{ff}(\tau)$ function are given as

$$R^m_{ff}(\tau) = \sigma^2 \exp(-|\tau|/b)(1 + |\tau|/b) \tag{3}$$

and

$$R^{sq}_{ff}(\tau) = \sigma^2 \exp(-\tau^2/b^2). \tag{4}$$

In these equations b is the so-called correlation length.

In the available literature, single and squared exponential functions are normally used in different application domains. As discussed in (Ching and Phoon, 2018), the main difference between these two autocorrelation functions is the smoothness of the resulting sample paths; related applications to soil engineering can be found in references such as (Yue et al., 2018) and (Yue et al., 2020). The main advantage of a single exponential kernel is the availability of analytical solutions in terms of its Karhunen–Loève expansion (Spanos, 1989) and the capability to characterize quickly varying phenomena. However, it exhibits non-differentiability at zero-lag. The modified exponential autocorrelation function aims at combining the strengths of both the aforementioned kernels, as it provides the temporal/geometric characteristics of the sample paths of a single exponential kernel, it solves the problem of the zero-lag discontinuities, and it provides analytical solutions for its expansion (Spanos, Beer and Red-Horse, 2007).

Alternatively, the auto-correlation of $(f(t_1, \theta), f(t_2, \theta), \ldots, f(t_n, \theta))$ can be represented in the frequency domain by means of a two-sided power spectrum $S_{ff}(\omega) : \Gamma \times \Gamma \mapsto \mathbb{R}$, with $\Gamma \subset \mathbb{R}$ the frequency domain. The Wiener-Khintchine theorem allows for the calculation of the autocorrelation function $R_{ff}(\tau)$ of a stochastic process from its two-sided power spectrum $S_{ff}(\omega)$ and vice versa based on following Fourier transforms:

$$S_{ff}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} R_{ff}(\tau) e^{-i\omega\tau} d\tau \tag{5}$$

and

$$R_{ff}(\tau) = \int_{-\infty}^{+\infty} S_{ff}(\omega) e^{i\omega\tau} d\omega \tag{6}$$

The power spectra corresponding to the autocorrelation functions in Eqs. 2–4 can be shown to b:

$$S^s_{ff}(\omega) = \frac{\sigma^2}{\pi} \frac{b}{b^2\omega^2 + 1}, \tag{7}$$

$$S^m_{ff}(\omega) = \frac{\sigma^2}{\pi} \frac{2b}{(b^2\omega^2 + 1)^2}, \tag{8}$$

and

$$S_{ff}^{sq}(\omega) = \frac{\sigma^2 b}{2\sqrt{\pi}} \exp \frac{-b^2\omega^2}{4}. \tag{9}$$

From Eq. 7 it can be deduced that the energy content of the spectrum is spread out over a higher range of frequencies when considering the case of single exponential kernel. Similarly, it can be noted that the energy content of the modified exponential kernel decreases quadratically with respect to the term $b^2\omega^2 + 1$ in Eq. 8. As such, the bandwidth of the single exponential kernel is wider as compared to the modified exponential. Finally, the $\exp(\cdot)$ term in Eq. 9 suggests that most of the energy content is located in the lower frequencies for the squared exponential kernel. Thus, it is shown that the power-spectra of the zero-lag differentiable kernels, i.e., the modified and squared exponential kernel, converge faster to zero as $\omega \to \infty$.

2.2. Simulation of stochastic processes

2.2.1. *Spectral representation*

Following the power spectral representation of the autocorrelation of a the stochastic process, the simulation of paths of this process can be performed according to (Rice, 1954; Shinozuka and Deodatis, 1991) by

$$f(t,\theta) = \sqrt{2} \sum_{n=0}^{N-1} A_n \cos(\omega_n t + \psi_n(\theta)), \tag{10}$$

where

$$A_n = \sqrt{2 S_{ff}(\omega_n)\Delta\omega}, \tag{11}$$

and ω_n is defined as

$$\omega_n = n\Delta\omega, \tag{12}$$

where $n = 1, \ldots, N-1$ and $\Delta\omega = \omega_u/N$. The phase angles $\psi_n(\theta)$ are considered random and distributed as $\mathcal{U}(0, 2\pi)$, with $\mathcal{U}$ the uniform distribution. Thus, samples $f(t, \theta_i)$ of the stochastic process can be generated by sampling from $\mathcal{U}(0, 2\pi)$. The parameter ω_u represents the cut-off frequency, beyond which the power spectral density function $S_{ff}(\omega)$ may be assumed to be zero for either mathematical or physical reasons. In practice, an energy criterion is commonly used. That is,

$$\int_0^{\omega_u} S_{ff}(\omega_n) d\omega = (1 - e_S) \int_0^{\infty} S_{ff}(\omega_n) d\omega, \tag{13}$$

with e_S a measure for the error, which is typically a small value, e.g., $e = 0.01$ or $e = 0.001$. In this regard, the appropriate selection of ω_u is of paramount importance for the accuracy of the analysis. In essence, it represents the degree of approximation of the energy content used in the representation of the stochastic process. As such, if ω_u is selected too small, a significant portion of the energy of the modelled signal is lost. This might lead to an underestimation of the magnitude of the physical quantity under consideration, which can lead to severe consequences, e.g., underestimation of probability of failure. Further, ω_u cannot be selected arbitrarily larger for numerical reasons. Extremely high values of ω_u require a corresponding quite high number of random variables $\phi_n(\theta)$,

leading to computationally untractable high-dimensional problems. Furthermore note that, while herein attention is focussed on the spectral representation, these arguments are also pertinent to other simulation methods based in the frequency domain, e.g., the Stochastic Harmonic Function representation (Chen et al., 2013).

2.2.2. *Karhunen-Loève expansion*

The Karhunen-Loève expansion is a powerful tool to represent stochastic process (Spanos, 1989). Specifically, following the Karhunen-Loève (KL) series expansion, a stochastic process $f(t,\theta)$ is represented as:

$$f(t,\theta) = \mu_x(t) + \sigma_f \sum_{i=1}^{\infty} \sqrt{\lambda_i}\psi_i(t)\xi_i(\theta), \tag{14}$$

where σ_f is the standard deviation of the random field. The quantities $\lambda_i \in (0,\infty)$ and $\psi_i(t) : \Omega \mapsto \mathbb{R}$ are respectively the eigenvalues and eigenfunctions of the continuous, bounded, symmetric and positive (semi-) definite auto-correlation function $R_{ff}(\tau)$. The decomposition of $R_{ff}(\tau)$ is performed in accordance with Mercer's theorem. That is,

$$R_{ff}(\tau) = \sum_{i=1}^{\infty} \lambda_i \psi_i(t)\psi_i(t'). \tag{15}$$

These quantities are in practice obtained by solving the homogeneous Fredholm integral equation of the second kind. That is,

$$\int_{\Omega} R_{ff}(\tau)\psi_i(t')dt' = \lambda_i\psi_i(t), \tag{16}$$

with $t' = t + \tau$. Analytical solutions to this equation exist only for a limited number of auto-correlation functions. In general, discretization schemes are used to solve this equation, as explained in (Betz et al., 2014). Since $R_{ff}(\tau)$ is bounded, symmetric, positive semi-definite, and, in most practical cases, can be even assumed positive definite, the eigenvalues λ_i are non-negative and the eigenfunctions $\psi_i(t)$ satisfy the following orthogonality condition:

$$\langle \psi_i(t), \psi_j(t) \rangle = \int_{\Omega} \psi_i(t)\psi_j(t)dt = \delta_{ij} \tag{17}$$

with δ_{ij} the Kronecker delta. $\langle \cdot, \cdot \rangle : \Omega \times \Omega \mapsto \mathbb{R}$ denotes the inner product in the functional space. Hence, the eigenfunctions form a complete orthogonal basis on an $\mathcal{L}_2$ Hilbert space. In this case, the series expansion in Eq. 15 can be shown to be optimally convergent (Spanos, 1989).

For practical reasons, the infinite series expansion in Eq. 14 has to be truncated after a finite number of terms $n_{KL} \in \mathbb{N}$. That is,

$$f(t,\theta) = \sigma_f \sum_{i=1}^{n_{KL}} \sqrt{\lambda_i}\psi_i(t)\xi_i(\theta), \tag{18}$$

where n_{KL} should be selected such that following inequality holds:

$$1 - \frac{1}{|\Omega|}\frac{1}{\sigma_f^2}\sum_{i=1}^{n_{KL}} \lambda_i \leq e_\sigma, \tag{19}$$

with e_σ the so-called mean error variance (Betz et al., 2014) and $|\Omega|$ denoting the length of the simulation domain.

In contrast to e_S introduced in Eq. (13), the quantity e_σ does not represent some sort of energy loss measure. Instead, e_σ represents the percentage of the variance of the original process that is captured by the truncated Karhunen-Loève expansion. In this respect, if n_{KL} is selected too small, it will lead to a loss of the variance in the representation. Hence, the magnitude of the physical quantity under consideration will be underestimated. Finally, similarly to the spectral representation, setting n_{KL} too large will render the computational cost of the corresponding analysis untractable due to the high dimension of the parameter input space.

2.3. Auto-regressive models

A third commonly used approach to simulate stochastic processes and fields is the auto-regressive representation (AR) method (Spanos & Zeldin, 1996). According to the AR framework, the value of a stochastic process $f(t,\theta)$ at time t_k with order m can be computed as

$$f(t_k,\theta) = -\sum_{i=1}^{m} a_i f(t_{k-I}) + b_0 w(t_k,\theta), \tag{20}$$

where a_i are the AR parameters and b_0 is the gain factor of the AR model. $w(t_k)$ is a band limited $[-\omega_b,\omega_b]$ white-noise process that satisfies

$$E\left[w(t_k)w(t_l)^T\right] = 2\omega_b I_n \delta_{kl}, \tag{21}$$

where $E[\bullet]$ and $\bullet^T$ denotes the operators of mathematical expectation and transpose respectively, ω_b is the cut-off frequency, I_n is the identity matrix and δ_{kl} the Kronecker delta. The representation in in Eq. (20) is the best linear estimator of $f(t_k,\theta)$ by using the m previous values $[f(t_k), f(t_{k-1}),\ldots,f(t_{k-m})]$. The corresponding error ϵ can be expressed as:

$$\epsilon = \frac{\triangle t}{2\pi} E\left[\left(f(t_k,\theta) + \sum_{i=1}^{m} a_i f(t_{k-I})\right)\right] = b_0^2. \tag{22}$$

The parameters a_i in the series expansion defined in Eq. (20) can generally be determined by minimizing ϵ. This leads to the so-called Yule-Walker equations that relate the stationary target autocorrelation function $R_{ff}(\tau) \equiv R_{ff}(t_k - t_i)$ to the AR parameters via a Toeplitz system of equations:

$$\sum_{i=1}^{m} R_{ff}(t_k - t_i)a_i = -R_{ff}(t_k) \qquad k = 1,\ldots,m \tag{23}$$

Alternative approaches, used to fit an AR model to a predefined autocorrelation model, interpret Eq. (20) as the response of a discrete linear system to a white-noise excitation. For a thorough explanation of the fitting of an AR model to a predefined auto-correlation function or power spectral density, the reader is referred to the seminal work of (Spanos & Zeldin, 1996) or (Spanos & Zeldin, 1998).

3. Convergence of the stochastic process representations

In this section analytical expressions are derived for the energy approximation error of the spectral stochastic representation of the of three considered autocorrelation kernels. Then, the paper provides additional illustrations regarding the improved convergence behavior of the modified exponential kernel over the single exponential when they are used in the context of a KL expansion. Finally, a comparison of AR models for these three autocorrelation functions is included for the sake of completeness.

3.1. Spectral stochastic representation

The convergence behavior of the autocorrelation functions in Eqs. 2-4 is studied with respect to Eq. 13. Herein, analytical expressions for the approximation error e_s are derived with respect to the cut-off frequency ω_u. To determine the cut-off frequency ω_u using Eq. (13), the power spectra $S_{ff}(\omega)$ in Eqs. 7-9 needs to be integrated. The right hand side integral of $S_{ff}(\omega)$ w.r.t. ω is obtained

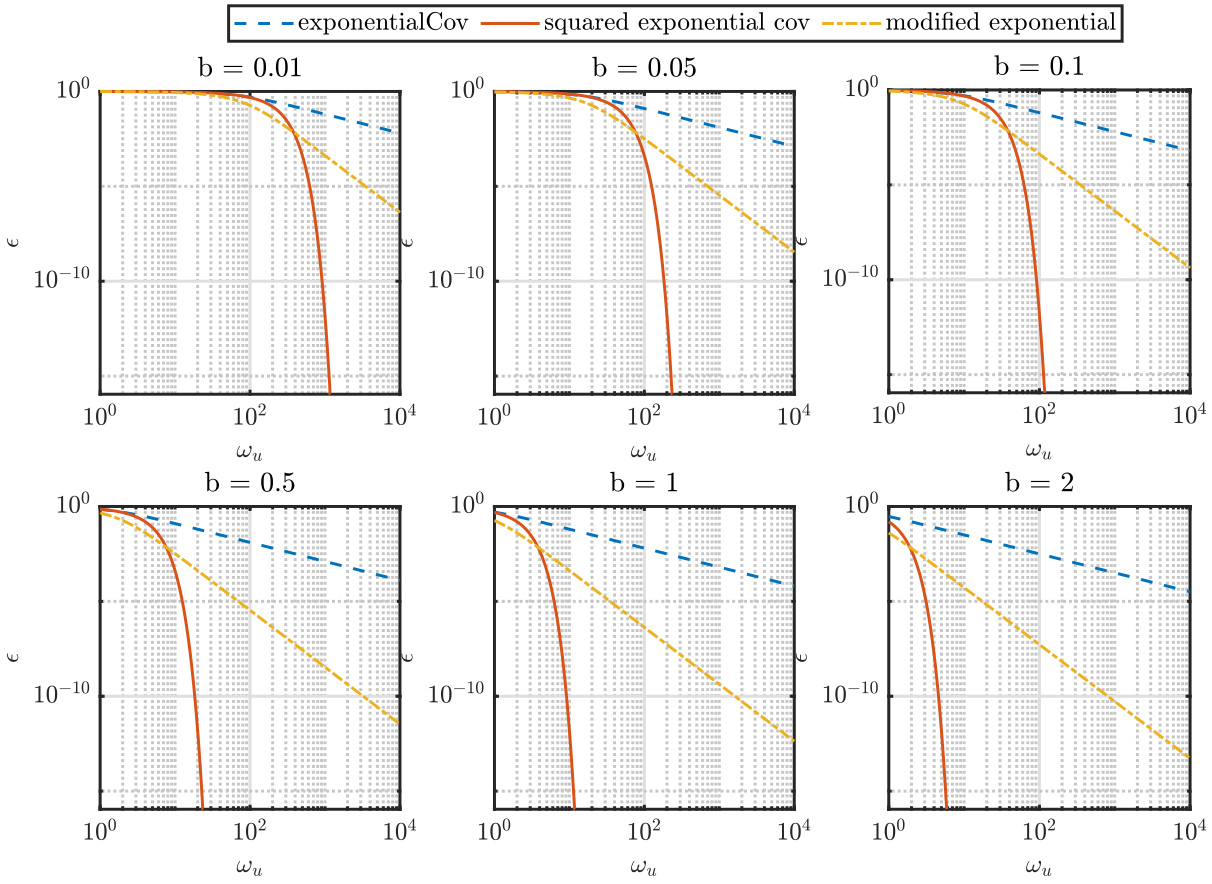


Figure 1. Decay of the energy approximation error e as a function of the truncation frequency ω_u.

by means of substitution and given for these power spectra by the equations

$$\int_0^\infty S_{ff}^s(\omega)d\omega = \left[\frac{\sigma^2}{\pi}\tan^{-1}(b\omega)\right]_0^\infty = \frac{\sigma^2}{2}, \tag{24}$$

$$\int_0^\infty S_{ff}^m(\omega)d\omega = \left[\frac{\sigma^2}{\pi}4b\left(\frac{\omega}{2(\omega^2b^2+1)} + \frac{1}{2b}\tan^{-1}(b\omega)\right)\right]_0^\infty = \sigma^2, \tag{25}$$

and

$$\int_0^\infty S_{ff}^{sq}(\omega)d\omega = \left[\sigma^2\text{erf}\left(\frac{b\omega}{2}\right)\right]_0^\infty = \sigma^2, \tag{26}$$

where $\text{erf}(x) = \frac{1}{\sqrt{\pi}}\int_0^x e^{-t^2}dt$ is the error function.

Similarly, the left-hand side of Eq. (13) for the considered autocorrelation functions can be determined as

$$\begin{aligned}\int_0^{\omega_u} S_{ff}^s(\omega)d\omega &= \left[\tfrac{\sigma^2}{\pi}\tan^{-1}(b\omega)\right]_0^{\omega_u} \\ &= \tfrac{\sigma^2}{\pi}\tan^{-1}(b\omega_u),\end{aligned} \tag{27}$$

$$\begin{aligned}\int_0^{\omega_u} S_{ff}^m(\omega)d\omega &= \left[\tfrac{\sigma^2}{\pi}4b\left(\tfrac{\omega}{2(\omega^2b^2+1)} + \tfrac{1}{2b}\tan^{-1}(b\omega)\right)\right]_0^{\omega_u} \\ &= \tfrac{\sigma^2}{\pi}4b\left(\tfrac{\omega_u}{2(\omega_u^2b^2+1)} + \tfrac{1}{2b}\tan^{-1}(b\omega_u)\right),\end{aligned} \tag{28}$$

and

$$\begin{aligned}\int_0^{\omega_u} S_{ff}^{sq}(\omega)d\omega &= \left[\sigma^2\text{erf}\left(\tfrac{b\omega}{2}\right)\right]_0^{\omega_u} \\ &= \sigma^2\text{erf}\left(\tfrac{b\omega_u}{2}\right).\end{aligned} \tag{29}$$

Using these equations in conjunction with Eq. (13), and isolating the desired truncation error e_S provides relations between the error and the cut-off frequency as. Specifically,

$$e_{S,s} = 1 - \frac{2}{\pi}\tan^{-1}(b\omega_u), \tag{30}$$

$$e_{S,m} = 1 - \frac{2}{\pi}\left(\frac{b\omega_u}{b^2\omega_u^2+1} + \tan^{-1}(b\omega_u)\right), \tag{31}$$

and

$$e_{S,sq} = 1 - \text{erf}\left(\omega_u\frac{b}{2}\right). \tag{32}$$

Figure 1 shows the decay of the truncation error of the energy e_S as a function of the selected ω_u for the considered autocorrelation models, for several different values of the correlation lengths $b = [0.01, 0.05, 0.1, 0.5, 1, 2]$. The decay function is calculated according to Eq. (30) to (32). In particular,

a zero-mean Gaussian stochastic process with unit variance on a domain $t \in [0, 5]$ is considered for this purpose. From this Figure, it is clear that this approximation error decays significantly faster in squared and modified exponential kernels as compared to a single exponential kernel. Specifically, the squared exponential kernel converges super-linearly, the modified exponential kernel converges linearly with a slope of 2.13 and the single exponential kernel converges linearly with a slope of 0.87. These findings are valid for all correlation lengths tested. The difference in the convergence rate between the modified and single exponential kernel originates from the $\frac{b\omega_u}{b^2\omega_u^2+1}$-term in Eq. (31). This term ensures that the error for the modified exponential kernel decreases faster versus the single exponential case when $w_u \to \infty$. However, note that the shape of the decay of the error with respect to ω_u is similar for both single and modified exponential curves.

In Figure 2 sample paths of the corresponding Gaussian stochastic processes are shown for different values of the truncation error e_S. Each sample path is created with all the three correlation structures in the time $t \in [0, 5]$, but only $t \in [0, 0.1]$ is shown to better visualize the local characteristic of the samples. This Figure shows how the sample path from the exponential correlation exhibits non-negligible high frequency oscillations and noise when compared with the other two correlations when lower values of e_S are used. This is a direct result of the slower convergence of the single exponential kernel in energy content, which in it's turn is related to the non-differentiability of R_{ff} at zero lag. As pointed out in the literature, these high-frequency oscillations potentially impact the engineering analysis affected by the properties modeled with this kernel.

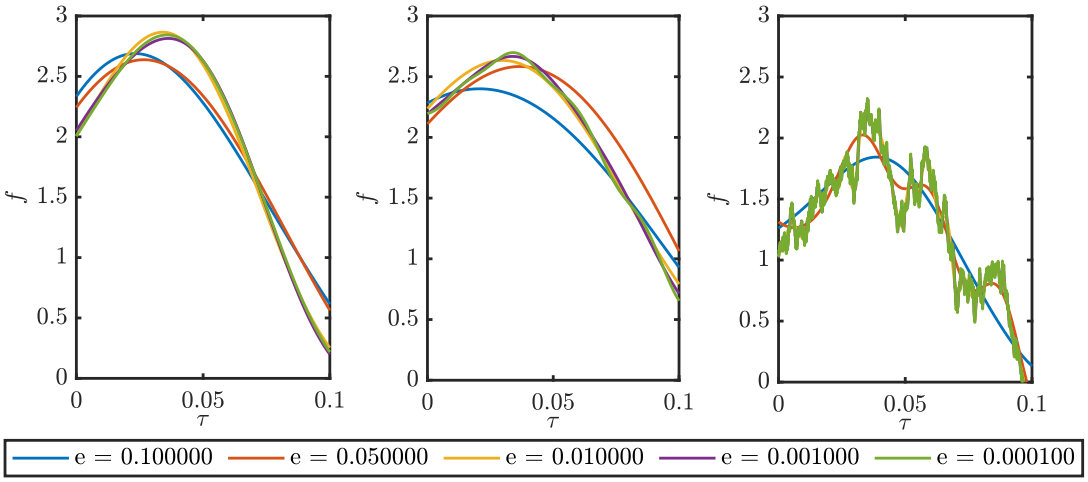


Figure 2. Sample paths of a one-dimensional Gaussian stochastic process, where the autocorrelation is governed by a Squared exponential (left), Modified exponential (middle) and Single exponential (right) kernel function, for different values of the truncation error.

3.2. Karhunen-Loève expansion

The advantage of using the modified exponential kernel over a single exponential one in terms of required terms in the Karhunen-Loève expansion has been illustrated in (Spanos, Beer and Red-Horse, 2007). In this section, additional insights into this topic are given. In particular, it is shown

that the reduction in required number of stochastic terms holds consistently for the entire process series expansion. Further, this property also holds for a wide range of correlation length values. Figure 3 shows the convergence of the mean error variance as a function of the number of terms that are retained in the KL expansion for the same stochastic process considered in Section 3.1. In this paper, the solution of the Fredholm integral equation in Eq. 16 is obtained using a Galerkin procedure with Legendre basis functions. Particularly the solution of the cases where more than 1000 terms are requirede were found to be computationally non-trivial. Clearly, the convergence of the mean error variance corresponding to the squared exponential process is the fastest of the three considered autocorrelation functions. These results further confirm the findings of (Spanos, Beer and Red-Horse, 2007), namely that the modified exponential kernel converges faster than the single exponential kernel when considering the KL expansion for a wide range of correlation length values. This can be also explained by the differentiability of the autocorrelation function at zero-lag.

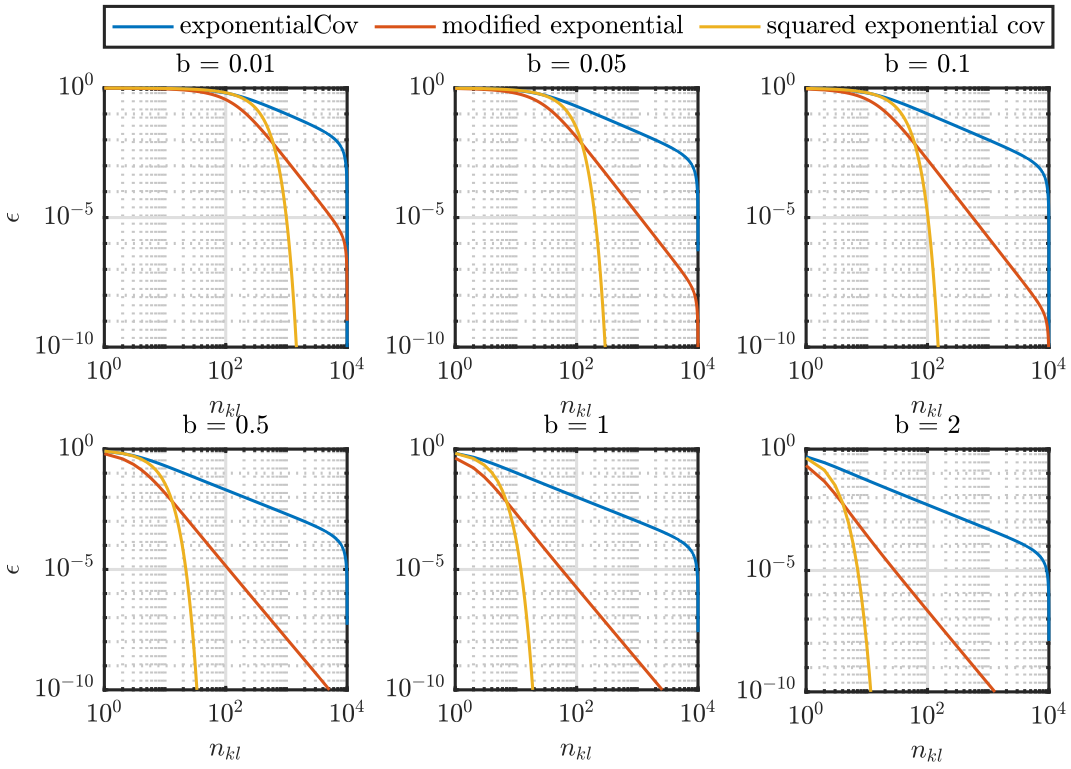


Figure 3. Convergence of the mean error variance with respect to the number of terms in the KL Expansion for the three considered autocorrelation kernels.

3.3. AR model representation

For the study of the applicability of the considered autocorrelation processes with an AR representation, an analytical approach is used. In particular, the required AR models are derived based on the considered auto-correlation functions.

Concerning the single exponential covariance kernel $R^s_{ff}(\tau)$, it can be shown that an AR(1) process, i.e., an AR process with $m = 1$ is capable of directly representing a Gaussian process with $R^s_{ff}(\tau)$ (Vanmarcke, 1983). To capture this, write out Eq. (20) for $m = 1$ and subtract $f(t, \theta)$ from both sides. That is,

$$[f(t + 1, \theta) - f(t, \theta)] + (1 - a)f(t, \theta) = b_0 w(t, \theta). \tag{33}$$

The associated first-order differential equation to this finite-difference equation is

$$\frac{d}{dt} f(t, \theta) + \alpha f(t, \theta) = b_0 w(t, \theta), \tag{34}$$

which can be interpreted as a linear model that links $f(t, \theta)$ to a white-noise excitation. In the limit state for infinitesimally small τ, the transfer function corresponding to Eq.(34) is

$$H(\omega) = \frac{1}{i\omega + \alpha}, \tag{35}$$

that in combination with a constant spectral density S_0 gives rise to the power spectral density $S_{ff}(\omega)$:

$$S_{ff}(\omega) = \frac{S_0}{\omega^2 + \alpha^2}, \tag{36}$$

for $f(, \omega)$. This can be related through Eq. (6) to the autocorrelation function:

$$R_{ff}(\tau) = \sigma^2 \exp(-|\tau|\alpha), \tag{37}$$

which corresponds to a single exponential autocorrelation function.

Following a similar procedure, it can be shown that an AR(1) model of the following form:

$$f(t, \theta) = a\,[f(t - 1, \theta) + f(t + 1, \theta)] + b_o 2(t, \theta) \tag{38}$$

corresponds to a stationary Gaussian process with power spectral density:

$$S_{ff}(\omega) = \frac{4\alpha^3 \sigma^2}{\pi(\omega^2 + \alpha^2)^2}, \tag{39}$$

and hence with auto correlation function

$$R_{ff}(\tau) = \sigma^2 \exp(-\alpha|\tau|)(\alpha|\tau| + 1), \tag{40}$$

that corresponds to the modified exponential kernel.

For the squared exponential autocorrelation function, a one-on-one exact relationship with a corresponding AR model is not trivial to establish and requires the solution of the Yule-Walker equations. This falls outside the scope of this paper.

4. Concluding remarks

In this paper, certain additional aspects relating to the mathematical behavior of common autocorrelation functions have been studied. Specifically, the convergence of the spectral density of the process to zero as the frequency tends to infinity has been examined. In particular, attention has been focussed on the comparison of the spectral convergence of functions that are differentiable at zero-lag versus those that are not-differentiable. It has been shown, both analytically and numerically, that the number of stochastic components required to represent a stochastic process with the single exponential kerne is considerably larger when compared to a modified or squared exponential kernel. This statement holds for both stochastic spectral representation methods, as well as for the well-known Karhunen-Loève series expansion. In this context, it must be noted that the single exponential kernel is not differentiable at zero lag, whereas, the modified and squared exponential kernel are completely differentiable. Further, in the paper AR models have been considered, for which it has been shown that for both a single and modified exponential kernel, a closed form expression for an AR(1) model can be derived. This proves that AR models are highly suited to represent these types of stochastic processes. Further work beyond this initial study will explore the relationship between the differentiability of the autocorrelation function at zero lag, and the convergence of the corresponding spectral density to zero as the frequency tends to infinity.

References

J. Chen, W. Sun, J. Li, and J. Xu. Stochastic harmonic function representation of stochastic processes. Journal of Applied Mechanics, Transactions ASME, 80(1) (2013), 1–11.

J. Ching and K.-K. Phoon, Impact of Autocorrelation Function Model on the Probability of Failure. *Journal of Engineering Mechancis*,145:04018123, 2018.

Z. Cao and Y. Wang, Bayesian model comparison and selection of spatial correlationfunctions for soil parameters. *Structural Safety*,49:10–17, 2014.

F. Desplentere, I. Verpoest and S. Lomov, Stochastic flow modeling for resin transfer moulding. *AIP Conference Proceedings*, 1152:284–292, 2009.

M. Shinozuka and G. Deodatis, Simulation of Stochastic Processes by Spectral Representation. *Applied Mechanics Reviews*, 44:191—204, 1991.

S.O. Rice, "Mathematical Analysis of Random Noise"in Selected Papers of Noise and Stochastic Processes,Edited by N. Wax, Dover Publications, Inc, New York,1954, pp. 180- 18

L.E. Borgman, "Ocean Wave Simulation for Engineering Design," Journal of Waterways and Harbor Div., ASCE,Vol. No. WW4, November, 1969, pp. 557- 583.

Q. Yue, J. Yao, A. H-S Ang, P D. Spanos. Efficient random field modeling of soil deposits properties. Soil Dynamics and Earthquake Engineering, 108: 1-12, 2018.

Q. Yue and J. Yao. Soil deposit stochastic settlement simulation using an improved autocorrelation model. Probabilistic Engineering Mechanics, 59, 103038, 2020.

G. Stefanou & M. Papadrakakis. Assessment of spectral representation and Karhunen-Loève expansion methods for the simulation of Gaussian stochastic fields. Computer Methods in Applied Mechanics and Engineering, 196(21–24) (2007), 2465–2477.

B. Sudret & A. Der Kiureghian (2000). Stochastic finite element methods and reliability. A state-of-the-art-report. In Technical Rep. UCB/SEMM-2000/08, Univ. of California, Berkeley, CA (Issue November).

P.D. Sclavounos (2012). Karhunen–Loève representation of stochastic ocean waves. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 468(2145), 2574–2594.

P. Spanos, M. Beer and J. Red-Horse, Karhunen–Loéve Expansion of Stochastic Processes with a Modified Exponential Covariance Kernel. *Journal of Engineering Mechanics*,133:773-–779, 2007.

L. Wang, C. Wu, Y. Li, H. Liu, W. Zhang and X. Chen, Probabilistic Risk Assessment of unsaturated Slope Failure Considering Spatial Variability of Hydraulic Parameters. *KSCE Journal of Civil Engineering*,23(12):5032–5040, 2019.

E. Vanmarcke, Random Fields: Analysis and Synthesis, MIT Press, Cambridge, 1983.

P. Spanos, R. Ghanem, Stochastic finite element expansion for random media, Journal of engineering mechanics 115 (5) (1989) 1035–1053.

W. Betz, I. Papaioannou, D. Straub, Numerical methods for the discretization of random fields by means of the Karhunen-Loève expansion, Computer Methods in Applied Mechanics and Engineering 271 (2014) 109–129.

S. P. Huang, S. T. Quek, K. K. Phoon, Convergence study of the truncated Karhunen–Loève expansion for simulation of stochastic processes, International Journal for Numerical Methods in Engineering 52 (9) (2001) 1029–1043.

P. D. Spanos, B. A. Zeldin. Efficient Iterative Arma Approximation of Multivariate Random Processes For Structural Dynamics Applications. Earthquake Engineering And Structural Dynamics, 25 (1996), 497–507.

P.D. Spanos, B. A. Zeldin. Monte Carlo treatment of random fields: A broad perspective. Applied Mechanics Reviews, 51 (1998), 219-237

Local Reliability Based Sensitivity Analysis with the Moving Particles Method

Carsten Proppe
Chair of Engineering Mechanics, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany, proppe@kit.edu

Abstract. Local reliability sensitivity methods aim at computing the partial derivatives of the failure probability or the reliability index with respect to model parameters. For efficient local reliability based sensitivity analysis, it is important to avoid repeated evaluations of the performance function. To this end, an extension of the moving particles method to local reliability based sensitivity analysis is presented that is completely based on the already evaluated samples for the reliability estimate and thus avoids repeated evaluations of the performance function. The method is discussed in detail and illustrated by means of examples.

Keywords: local sensitivity analysis, reliability, multilevel splitting, moving particles

1. Introduction

Reliability based sensitivity analysis investigates the dependence of the failure probability on model parameters. In this study, the relevant model parameters are distribution parameters related to the probabilistic characterization of the model input. For sensitivity analysis, local and global approaches can be distinguished. Local reliability sensitivity methods compute the partial derivatives of the failure probability or the reliability index with respect to model parameters. Global reliability sensitivity analysis aims at determining the influence of a model parameter on the failure probability over the entire range of possible values for this parameter. It is often based on Sobol' indices.

(Wu, 1994) considered local reliability based sensitivity analysis and computed partial derivatives of the failure probability by means of importance sampling without any additional evaluations of the performance function. (Song et al., 2009) extended this approach to subset simulation by computing partial derivatives of the conditional probabilities for each of the subsets.

Here, the reliability analysis is carried out by the moving particles method, cf. (Walter, 2015), where a threshold is associated to each sample, samples are moved to new positions in the design space and the number of moves for the initial samples (the particles) to reach the failure region are counted and yield an estimator for the failure probability, which is based on a Poisson process. The estimator is of comparable accuracy and efficiency as the subset simulation estimator. For local reliability based sensitivity analysis, it is important to avoid repeated evaluations of the performance function. To this end, an extension of the moving particles method to local reliability based sensitivity analysis is presented. It is shown that the method is not only completely based

© 2021 *by C. Proppe.*

on the already evaluated samples for the reliability estimate and thus avoids repeated evaluations of the performance function, but that it can be carried out after the last move of a particle and just once after the reliability estimation has been carried out.

The paper is organized in five sections. After the introduction, a brief outline of the moving particles method is given. Following this, local reliability sensitivity analysis is integrated in the moving particles method and it is shown that only a single post processing step is required which is carried out when the last particle has been moved into the failure domain. Finally, the accuracy and efficiency of the proposed method is investigated by examples and conclusions are drawn.

2. Brief Outline of the Moving Particles Method

Structural reliability analysis is concerned with the computation of the failure probability

$$P_F = \int_{g(x)<0} p(x,\theta)\mathrm{d}x, \tag{1}$$

where the real valued function $g(x)$, $x \in \mathbb{R}^n$, denotes the performance function of the structure, $p(x,\theta)$ is the probability density function of those structural parameters that are assumed to be random variables and θ contains the parameters of the probability distributions. The limit state function $g(x) = 0$ separates the safe domain $\{x \in \mathbb{R}^n : g(x) > 0\}$ from the failure domain $F = \{x \in \mathbb{R}^n : g(x) < 0\}$. In engineering applications, the failure probability is small and direct Monte Carlo simulation for its estimation requires a huge number of evaluations of the performance function and thus a large computational effort.

In order to reduce the computational effort, importance sampling and importance splitting have been widely applied to the estimation of rare events (Schuëller et al., 2004). Importance sampling estimates the occurrence of rare events by generating samples from an alternative distribution and correcting for the bias by the introduction of weights. The success of this method relies on the quality of the importance sampling density, which is efficiently constructed in an adaptive way.

Importance splitting allows estimating small failure probabilities efficiently, even for problems that involve a high-dimensional vector of input random variables (Schuëller et al., 2004). The general idea of importance splitting is to discard samples that are too far away from a threshold and to regenerate new samples from the remaining ones. In order to maintain the independence of the samples, Markov Chain Monte Carlo simulation starting from remaining samples is frequently employed.

For time-independent problems, subset simulation (Au and Beck, 2003) is a very efficient method for a wide range of applications. This method considers nested subsets generated by levels of the performance function. These nested subsets yield a multiplicative decomposition of the failure probability in larger conditional probabilities. For the estimation of these conditional probabilities,

conditional samples are generated by means of Markov Chain Monte Carlo simulation.

The moving particles algorithm (Guyader et al., 2011) can be considered as a special case of subset simulation with a maximum number of subsets. It starts with an initial Monte Carlo simulation (MCS) with N samples $x_j^{(1)} \in \mathbb{R}^n$, $j = 1, ..., N$, distributed according to the probability density function $p(x, \theta)$. The values $g(x_j^{(1)})$, $j = 1, ..., N$, of the N samples are ranked. These initial samples (the particles) are then moved to the failure region by the following procedure: For the sample with the maximum value of the performance function (thus furthest away from a limit state), a Markov chain Monte Carlo simulation (MCMC) is carried out starting from one of the other samples and the next state of the Markov chain is accepted, if the value of the performance function can be reduced. The Markov chain can be generated either by application of the classical Metropolis-Hastings algorithm or by direct sampling from a normal transition kernel and requires a burn-in period in order to maintain independence of the samples. This algorithm is repeated and stops if all the particles have been moved into the failure domain. The total number of moves M_T is counted.

Interpreted as a special case of subset simulation with a maximum number of domains, each move of a particle gives rise to a subset $F_i = \{x \in \mathbb{R}^n : g(x) < g_i\}$, $i = 1, ..., M_T$, where $g_i = \max_{1 \leq j \leq N} g(x_j^{(i)})$, $i = 1, ..., M_T$, denotes the maximum value of the performance function of the N particles $x_j^{(i)}$, $j = 1, ..., N$, when carrying out the ith move. For the conditional probability

$$P_i = P(G < g_i | G < g_{i-1}), \tag{2}$$

where for simplicity, the random variable $g(X)$ is abbreviated by G, one finds the estimator

$$\hat{P}_i = \frac{N-1}{N}, \tag{3}$$

because only one particle is moved per each iteration step of the algorithm. $F_0 = \mathbb{R}^n$ together with F_i, $i = 1, ..., M_T$, is a finite sequence of nested subsets and

$$F = \bigcap_{i=0}^{M_T} F_i. \tag{4}$$

Therefore, one finds for the failure probability

$$P_F = \prod_{i=1}^{M_T} P(F_i | F_{i-1}) = \prod_{i=1}^{M_T} P_i \tag{5}$$

and thus

$$\hat{P}_F = \left(\frac{N-1}{N} \right)^{M_T} \tag{6}$$

is an estimator for P_F.

In addition, this algorithm has the advantage that the number of moves of the particles is a Poisson distributed random variable that can be directly related to the failure probability P_F. Thus, properties of the Poisson distribution can be used to analyze the algorithm. For each initial sample j, $j = 1, ..., N$, the number M_j of moves until it reaches the failure region is counted. As has been shown in (Guyader et al., 2011), the number of moves to get an initial sample into the failure region follows a Poisson distribution with parameter $\lambda = -\log P_F$. The estimator for the parameter of the Poisson distribution is obtained from $\lambda = E[M]$ as

$$\hat{\lambda} = \frac{\sum_{j=1}^{N} M_j}{N} = \frac{M_T}{N}. \tag{7}$$

In order to obtain an unbiased estimate, it is mandatory that the trajectories of the Poisson process generated from the initial samples remain independent until the samples finally reach the failure domain. In (Walter, 2015), two means are proposed to maintain the independence:

- Burn-in: The Markov chain simulation is carried out with a burn-in period. The burn-in should ensure the independence of the candidate and the seed of the Markov chain.

- Seed avoidance: Repeated use of the same starting point for the Markov chain should be avoided. Once a sample has been used as starting point, the sample and its offspring should not be used as starting point again.

3. Local Reliability Sensitivity Analysis with the Moving Particles Method

Applying the product rule to equation (5), one finds for the partial derivative of the failure probability with respect to θ:

$$\frac{\partial P_F}{\partial \theta} = P_F \sum_{i=1}^{M_T} \frac{1}{P_i} \frac{\partial P_i}{\partial \theta}. \tag{8}$$

With

$$P(G < g_{i-1}) = \int I_{g<g_{i-1}}(x) p(x) \mathrm{d}x \tag{9}$$

and

$$P_i = P(G < g_i | G < g_{i-1}) = \int I_{g<g_i}(x) \frac{I_{g<g_{i-1}}(x) p(x)}{P(G < g_{i-1})} \mathrm{d}x, \tag{10}$$

where the indicator function is given by

$$I_{g<g_i}(x) = \begin{cases} 1, & \text{if } g(x) < g_i, \\ 0, & \text{otherwise,} \end{cases} \tag{11}$$

the partial derivative of the conditional probability P_i is given by

$$\begin{aligned}\frac{\partial P_i}{\partial \theta} &= \frac{\partial}{\partial \theta}\int I_{g<g_i}(x)\frac{I_{g<g_{i-1}}(x)p(x)}{P(G<g_{i-1})}\mathrm{d}x \\ &= \int I_{g<g_i}(x)\frac{I_{g<g_{i-1}}(x)}{P(G<g_{i-1})}\frac{\partial p(x)}{\partial \theta}\mathrm{d}x - \int I_{g<g_i}(x)\frac{I_{g<g_{i-1}}(x)p(x)}{P(G<g_{i-1})^2}\frac{\partial P(G<g_{i-1})}{\partial \theta}\mathrm{d}x \\ &= \int I_{g<g_i}(x)\frac{1}{P(G<g_{i-1})}\frac{\partial p(x)}{\partial \theta}\mathrm{d}x - \int I_{g<g_i}(x)\frac{p(x)}{P(G<g_{i-1})^2}\frac{\partial P(G<g_{i-1})}{\partial \theta}\mathrm{d}x,\end{aligned} \tag{12}$$

as $g_i < g_{i-1}$ and thus $I_{g<g_i}(x)I_{g<g_{i-1}}(x) = I_{g<g_i}(x)$. The last expression can also be written as

$$\frac{\partial P_i}{\partial \theta} = E_{G<g_{i-1}}[I_{g<g_i}(x)\frac{\partial \ln p(x)}{\partial \theta}] - \frac{P_i}{P(G<g_{i-1})}\frac{\partial P(G<g_{i-1})}{\partial \theta}. \tag{13}$$

The expression in the last term involves the partial derivative of the probability related to a set with larger bound, namely F_{i-1}. Analogously to equation (8), we find the expression

$$\frac{\partial P(G<g_{i-1})}{\partial \theta} = P(G<g_{i-1})\sum_{j=1}^{i-1}\frac{1}{P_j}\frac{\partial P_j}{\partial \theta} \tag{14}$$

for this partial derivative. In summary, the expression for the partial derivative of the conditional probability P_i with respect to θ reads

$$\frac{\partial P_i}{\partial \theta} = E_{G<g_{i-1}}[I_{g<g_i}(x)\frac{\partial \ln p(x)}{\partial \theta}] - \sum_{j=1}^{i-1}\frac{1}{P_j}\frac{\partial P_j}{\partial \theta}. \tag{15}$$

Inserting this expression into equation (8) and observing that $\hat{P}_1 = \hat{P}_2 = ... = \hat{P}_{M_T}$ yields a telescoping sum, such that the estimator for the sensitivity of the failure probability is obtained as

$$\frac{\partial \hat{P}_F}{\partial \theta} = \left(\frac{N-1}{N}\right)^{M_T-1}\left(\frac{1}{N}\sum_{j=1}^{N} I_{g<g_{M_T}}(x_j^{(M_T)})\left.\frac{\partial \ln p(x)}{\partial \theta}\right|_{x=x_j^{(M_T)}}\right). \tag{16}$$

Thus, the score function $\frac{\partial \ln p(x)}{\partial \theta}$, cf. (Rubinstein and Shapiro, 1993), is evaluated and averaged for the remaining $N-1$ particles before the move of the last particle into the failure domain.

For the n-dimensional normal probability density function

$$p(x,\mu,\Sigma) = \frac{1}{\sqrt{(2\pi)^n \det(\Sigma)}}\exp(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)) \tag{17}$$

the partial derivatives with respect to μ and Σ yield the expressions

$$\frac{\partial \ln p(x,\mu,\Sigma)}{\partial \theta} = \Sigma^{-1}(x-\mu) \tag{18}$$

Table I. Distribution parameters for the random variables in example 1.

	q [N/m]	ℓ [m]	A_C [m^2]	E_C [N/m^2]	A_S [m^2]	E_S [N/m^2]
mean value	20000	12	0.04	2×10^{10}	9.82×10^{-4}	1×10^{11}
std. dev.	1400	0.12	0.0048	1.2×10^{9}	5.892×10^{-5}	6×10^{9}

for $\theta = \mu$ and

$$\frac{\partial \ln p(x,\mu,\Sigma)}{\partial \theta} = -\Sigma^{-1} + \frac{1}{2}(\Sigma^{-1} \circ I) + \Sigma^{-1}(x-\mu)(x-\mu)^T\Sigma^{-1} - \frac{1}{2}(\Sigma^{-1}(x-\mu)(x-\mu)^T\Sigma^{-1} \circ I) \quad (19)$$

for $\theta = \Sigma$, where $\circ$ denotes the Hadamard product and I the $n \times n$ identity matrix.

4. Examples

Example 1: Roof truss

Consider the truss structure shown in Figure 1, cf. also (Song et al., 2009). The top boom and the compression bars are made of reinforced concrete (cross section area A_C, Young's modulus E_C), while the bottom boom and the tension bars are made of steel (cross section area A_S, Young's modulus E_S). The concentrated load P represents a constant distributed load q that is applied on the roof truss and thus $P = \frac{q\ell}{4}$. The parameters q, ℓ, A_C, E_C, A_S and E_S are assumed to be independent normally distributed random variables. Table I displays the mean values and standard deviation of the random variables which are the same as in (Song et al., 2009).

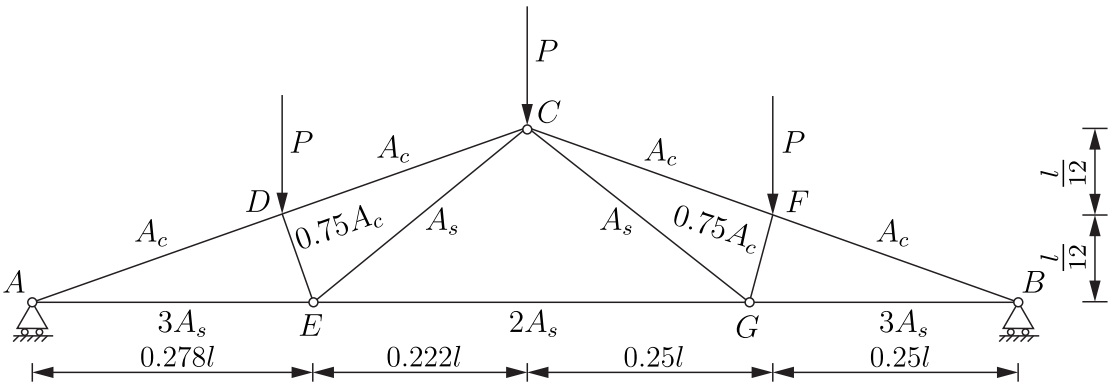


Figure 1. Roof truss structure.

Table II. Example 1: Failure probability and sensitivity of the failure probability with respect to the distribution parameters.

	P_F	$\frac{\partial P_F}{\partial m_q}$	$\frac{\partial P_F}{\partial m_\ell}$	$\frac{\partial P_F}{\partial m_{A_C}}$	$\frac{\partial P_F}{\partial m_{E_C}}$	$\frac{\partial P_F}{\partial m_{A_S}}$	$\frac{\partial P_F}{\partial m_{E_S}}$
estimate	0.00937	1.10e-5	0.0403	-2.110	-3.71e-12	-186	-1.81e-12
elasticity		23.5	51.9	-9.1	-8.0	- 19.6	-19.5
c.o.v.	0.05	0.05	0.08	-0.06	-0.08	-0.07	-0.06
rel. error [%]	0.5	0.7	0.9	1.0	1.2	0.1	0.6
		$\frac{\partial P_F}{\partial \sigma_q}$	$\frac{\partial P_F}{\partial \sigma_\ell}$	$\frac{\partial P_F}{\partial \sigma_{A_C}}$	$\frac{\partial P_F}{\partial \sigma_{E_C}}$	$\frac{\partial P_F}{\partial \sigma_{A_S}}$	$\frac{\partial P_F}{\partial \sigma_{E_S}}$
estimate		1.57e-5	0.0182	2.5047	1.93e-12	204	1.99e-12
elasticity		2.4	0.2	1.3	0.2	1.3	1.3
c.o.v.		0.06	0.22	0.08	0.20	0.09	0.09
rel. error [%]		2.9	1.4	1.3	3.4	0.8	1.0

If the displacement of the top node should not exceed 3 cm, the limit state function reads

$$g(q, \ell, A_C, E_C, A_S, E_S) = 0.03 - \frac{q\ell^2}{2}\left(\frac{3.81}{A_C E_C} + \frac{1.13}{A_S E_S}\right) \tag{20}$$

Table II summarizes the results for the estimated failure probability and the obtained sensitivities with respect to the distribution parameters. In order to be able to compare the influence of the parameters, it also states the elasticity, which is given by

$$\frac{\theta}{P_F}\frac{\partial P_F}{\partial \theta}, \tag{21}$$

cf. (Lemaire, 2010). The initial number of samples, i.e. the particles, has been set to 2000. The relative error has been computed with respect to results obtained by Monte Carlo simulation and published in (Song et al., 2009).

The results indicate that the mean values, notably the mean length ℓ and the mean distributed force q are most influential on the failure probability. The approximation error is rather low. The proposed method based on just 2000 samples is as exact as a Monte Carlo simulation with 10^7 samples. This clearly demonstrates the efficiency of the proposed method. Also, most of the coefficients of variation (c.o.v.) are rather low. However, it was noted that for the sensitivity of the failure probability with respect to the length ℓ and the Young's modulus of concrete E_C, the coefficient of variation is slightly increased. A possible reason could be the fact that the proposed method computes an average on the 2000 samples, which is a rather low number.

Example 2: One storey one bay elastoplastic frame

This example deals with a series system consisting of four failure modes that lead to nearly the same first order reliability index β. Consider the elastoplastic frame structure shown in Figure 2. The four potential failure modes yield the performance functions

$$\begin{aligned} g_1 &= 2M_1 + 2M_3 - 4.5S, \\ g_2 &= 2M_1 + M_2 + M_3 - 4.5S, \\ g_3 &= M_1 + M_2 + 2M_3 - 4.5S, \\ g_4 &= M_1 + 2M_2 + M_3 - 4.5S, \end{aligned} \tag{22}$$

cf. (Song et al., 2009). The performance function is thus given by the minimum of these four functions:

$$g(M_1, M_2, M_3, S) = \min(g_1, g_2, g_3, g_4). \tag{23}$$

M_i, $i = 1, .., 3$ and S are assumed to be independent normally distributed random variables. For M_i, $i = 1, ..., 3$, a mean value of 5.2872 and a standard deviation of 0.1492 is assumed, while for S, the mean value was 3.8378 and the standard deviation 0.3853, so that the reliability index β for each of the failure modes is 3.334 for the first failure mode and 3.364 for the others.

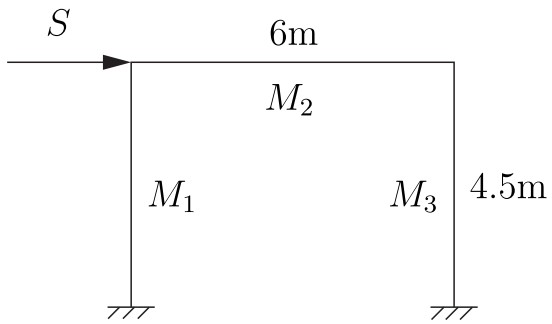


Figure 2. One storey one bay elastoplastic frame.

The results for the estimated failure probability and the obtained sensitivities are collected in Table III. The initial number of samples, i.e. the particles, has been set to 2000. Again, the relative error has been computed with respect to Monte Carlo simulation results with 10^7 samples published in (Song et al., 2009).

As Table III demonstrates, the mean values and the standard deviation related to the applied force S influence the failure probability most. The approximation error is always less than 2%. This confirms the findings of the previous example that the proposed method is very efficient and as exact as a highly accurate Monte Carlo simulation with 10^7 samples. As in the previous example, some of the coefficients of variation are slightly increased (notably for the sensitivities with respect to the standard deviation of the parameters M_i, $i = 1, ..., 3$).

Table III. Example 2: Failure probability and sensitivity of the failure probability with respect to the distribution parameters.

	P_F	$\frac{\partial P_F}{\partial m_{M_1}}$	$\frac{\partial P_F}{\partial m_{M_2}}$	$\frac{\partial P_F}{\partial m_{M_3}}$	$\frac{\partial P_F}{\partial m_S}$
estimate	0.0182	-0.0387	-0.0236	-0.038	0.1132
elasticity		-11.2	-6.9	-11.0	23.9
c.o.v.	0.05	-0.11	-0.18	-0.13	0.04
rel. error [%]	0.4	1.4	1.3	0.5	0.1
		$\frac{\partial P_F}{\partial \sigma_{M_1}}$	$\frac{\partial P_F}{\partial \sigma_{M_2}}$	$\frac{\partial P_F}{\partial \sigma_{M_3}}$	$\frac{\partial P_F}{\partial \sigma_S}$
estimate		0.0177	0.0195	0.0171	0.2318
elasticity		0.1	0.2	0.1	4.9
c.o.v.		0.29	0.3	0.34	0.05
rel. error [%]		1.2	0.6	3.0	0.1

Example 3: High-dimensional problem

This example deals with a component reliability problem of dimension 100. The performance function is given by

$$g(X) = \frac{1}{1000 + \sum_{i=1}^{100} X_i} - \frac{1}{1000 + 3\sqrt{100}}, \tag{24}$$

where the 100 random variables are normal with mean $m = 0$ and standard deviation $\sigma = 1$, cf. (Rahman, 2009). Two cases are considered: (a) the random variables are independent, (b) all random variables are correlated with correlation coefficient $\rho = 0.5$.

As Table IV indicates, the high dimension of the vector of random parameters does not affect the accuracy of the obtained results. Also the correlation of the random variables seems to have no influence on the accuracy. However, it is observed that for the estimator of the partial derivative of the failure probability with respect to the standard deviation, the coefficient of variation is larger than for the uncorrelated case.

5. Conclusions

In this paper, local reliability based sensitivity analysis based on the moving particles method has been introduced. It has been shown that the sensitivity analysis can be carried out in a single post processing step after the last particle has been moved into the failure domain.

Table IV. Example 3: Failure probability and sensitivity of the failure probability with respect to the distribution parameters.

$\rho = 0$	P_F	$\frac{\partial P_F}{\partial m}$	$\frac{\partial P_F}{\sigma}$
estimate	1.35e-3	4.44e-2	1.32e-2
c.o.v.	0.069	0.069	0.083
rel. error [%]	0.3	0.1	1.0
$\rho = 0.5$	P_F	$\frac{\partial P_F}{\partial m}$	$\frac{\partial P_F}{\sigma}$
estimate	0.336	0.514	0.156
c.o.v.	0.024	0.027	0.70
rel. error [%]	0.1	0.1	1.3

The examples demonstrate that the proposed method yields accurate estimates for the partial derivatives of the failure probability with respect to the distribution parameters of the random variables. However, due to the low number of particles on which the estimator is built, the coefficient of variation can be rather high. In this regard, a multilevel approach (Proppe, 2020) would certainly help to reduce the coefficient of variation and thus to further improve the efficiency of the sensitivity analysis.

References

Au, S. K. and Beck, J. (2003). Important sampling in high dimensions. *Structural Safety*, 25:139–163.

Guyader, A., Hengartner, N., and Matzner-Løber, E. (2011). Simulation and estimation of extreme quantiles and extreme probabilities. *Applied Mathematics & Optimization*, 64:171–196.

Lemaire, M. (2010). *Structural Reliability*. Wiley.

Proppe, C. (2020). A multilevel moving particles method for reliability estimation. *Probabilistic Engineering Mechanics*, 59:to appear.

Rahman, S. (2009). Stochastic sensitivity analysis by dimensional decomposition and score functions. *Probabilistic Engineering Mechanics*, 24:278–287.

Rubinstein, R. and Shapiro, A. (1993). *Discrete Event Systems – Sensitivity Analysis and Stochastic Optimization by the Score Function Method*. Wiley.

Schuëller, G., Pradlwarter, H., and Koutsourelakis, P. (2004). A critical appraisal of reliability estimation procedures for high dimensions. *Probabilistic Engineering Mechanics*, 19(4):463–474.

Song, S., Lu, Z., and Qiao, H. (2009). Subset simulation for structural reliability sensitivity analysis. *Reliability Engineering and System Safety*, 94:658–665.

Walter, C. (2015). Moving particles: A parallel optimal multilevel splitting method with application in quantiles estimation and meta-model based algorithms. *Structural Safety*, 55:10–25.

Wu, Y. T. (1994). Computational methods for efficient structural reliability and reliability sensitivity analysis. *AIAA Journal*, 32(8):1717–23.

Mutual information for global sensitivity analysis and adaptive-learning surrogate modelling

Umberto Alibrandi
Department of Civil and Architectural Engineering, AArhus University, Denmark, e-mail: ua@cae.au.dk

Lars V. Andersen
Department of Civil and Architectural Engineering, AArhus University, Denmark, e-mail: lva@cae.au.dk

Enrico Zio
MINES ParisTech / PSL Université Paris, Centre de Recherche sur les Risques et les Crises (CRC), Sophia Antipolis, France
Energy Department, Politecnico di Milano, Via La Masa 34, Milano, 20156, Italy
e-mail: enrico.zio@polimi.it

Abstract. In this paper tools of information theory are applied for probabilistic sensitivity analysis and surrogate modelling with adaptive sampling. One of the authors has recently proposed the adoption of the informational coefficient of correlation as a measure of dependence between the random variables, instead of the largely adopted linear coefficient of correlation. First, it is shown that it can be used for probabilistic sensitivity analysis. Moreover, two novel learning functions for adaptive sampling are proposed. The first, called $\mathcal{H}$- function, gives rise to the method AK-H (Adaptive Kriging - Entropy), and it describes the epistemic uncertainty through the entropy. The second, called $\mathcal{MI}$- function, gives rise to the method AL-MI (Active Learning - Mutual Information), which describes the model error through the Mutual Information. The second has the peculiarity that allows the implementation of adaptive learning in any kind of surrogate modelling, even different from Kriging. A simple numerical example shows the features and the potential of the proposed approach.

Keywords: Information theory, informational coefficient of correlation, global sensitivity Analysis, Active Learning, Kriging

1. Introduction

According to structural reliability theory (Ditlevsen and Madsen, 1996; Madsen et al., 1986; Melchers, 1999), the failure probability with respect to an assigned limit state is defined as

$$P_f = \int_{G(\mathbf{x})\leq 0} f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \tag{1}$$

© 2021 *by U. Alibrandi, L.V. Andersen, E. Zio*

where $\mathbf{x}$ is an n-dimensional vector collecting the basic random variables $x_1, x_2, ..., x_n$, $y = G(\mathbf{x})$ is the limit state function, $G(\mathbf{x}) = 0$ is the limit state surface separating the failure set $G(\mathbf{x}) \leq 0$ from the safe set $G(\mathbf{x}) > 0$ and $f_{\mathbf{X}}(\mathbf{x})$ is the joint probability density function of the random variables $\mathbf{x}$.

The failure probability can be evaluated by using analytical methods, like the First- and Second-Order Reliability Methods (FORM/SORM) FORM has gained wide popularity because of its simplicity and computational cost; however it presents two known main drawbacks: (1) its approximation may not be adequate for limit state surfaces departing significantly from linearity around the region of probabilistic interest, or in the presence of multiple design points (Der Kiureghian and Dakessian, 1998), and (2) it does not provide information about the degree of accuracy achieved.

Monte Carlo Simulation (MCS) is the most robust procedure but requires excessive computational effort for the evaluation of very small failure probabilities, in particular in its crude form. Smart sampling techniques have been proposed, e.g. Importance Sampling (IS) (Melchers, 1999; Rubinstein and Krose, 2008), Subset Simulation (SS) (Au and Beck, 2001), Line Sampling (De Angelis et al., 2015). These approaches are robust, but they typically remain computationally demanding for industrial applications.

To reduce the computational cost, an alternative strategy is given by the Response Surface Methodology (RSM), which builds a surrogate model of the target limit state function, defined in a simple and explicit mathematical form (Faravelli, 1989). FORM/SORM can, then be considered are particular kinds of response surfaces, which approximate the limit state with linear/quadratic response surfaces passing through the design point $\mathbf{u}^*$. Response surface models can be built to find the design point with reduced computational cost (Bucher and Burgound, 1990; Rajashekhar and Ellingwood, 1993; Zheng and Das, 2000). The approximation provided by these formulations however, shares the same drawbacks of FORM/SORM, and response surfaces aiming at modelling the whole limit state surface have been proposed, like the Moving Least-Square Regression (Bucher and Most, 2008), High-Dimensional Model Representation (HDMR) (Li et al., 2001; Chowdury et al., 2005; Alibrandi, 2014) or the Polynomial Chaos Expansion (PCE) (Sudret and Der Kiureghian, 2002; Choi et al., 2004; Berveiller et al., 2006; Blatman and Sudret, 2010).

Recently, increasing popularity has been gained by the Gaussian process models, also known as Kriging interpolation models (Rasmussen and Williams, 2006; Forrester et al., 2008), which are attractive because of their capability to: (i) interpolate data at chosen training samples and, (ii) to provide measures of local uncertainty for the model predictions. In (Kaymaz, 2005), Kriging is applied for structural reliability analysis, and it is shown that the Gaussian function is suitable for limit states nonlinear around the design point. In (Gaspar et al., 2014), it is shown that a polynomial of order zero (constant) can be considered as a basis function for practical applications, and that Kriging is well tailored for adaptive experimental designs. In Adaptive Kriging (AK), a *learning function* allows selecting the suitable new points of the experimental design. AK has been used jointly with different strategies of sampling: Monte Carlo (AK-MCS) (Echard et al., 2011), Importance Sampling (AK-IS) (Echard et al., 2013; Duborg et al., 2013), Subset Simulation (AK-SS) (Huang et al., 2016). A key factor for the efficiency of AK is represented by the choice of the learning function. The most commonly used are the Expected Feasibility Function (EFF) (Bichon et al., 2008), and the $\mathcal{U}$ learning function (Echard et al., 2011). Some alternative learning functions

have been proposed recently (LV et al., 2015; Zhang et al., 2019). This is also the direction pursued in this paper.

In (Alibrandi and Mosalam, 2019; Alibrandi and Mosalam, 2020; Alibrandi and Mosalam, 2021) a framework of data-driven uncertainty quantification and risk analysis based on information theory is proposed; moreover the authors propose the adoption of the *informational coefficient of correlation* as a measure of dependence. This is based on the concept of Mutual Information (MI) and it is capable to describe the full dependencies between random variables. Starting from these findings, in this paper it is proposed an alternative approach for AK based on the Mutual Information (MI), giving rise to AK-MI. First, MI is used for probabilistic sensitivity analysis in order to obtain an initial ranking of the most important variables to be included in the surrogate model. Second, Entropy and MI are adopted to define a learning function for adaptive learning. A simple numerical example shows the potential of the proposed approach.

2. Adaptive Kriging (AK) for structural reliability analysis

At first, as typically done in reliability analysis, a isoprobabilistic transformation toward the normal standard space is pursued, so that the faiure probability is

$$P_f = \int_{G(\mathbf{u})\leq 0} \varphi_n(\mathbf{u}) d\mathbf{u} \tag{2}$$

where $\varphi_n(\mathbf{u})$ is the multivariate normal standard distribution, whereas $G(\mathbf{u})$ is the limit state function defined in the $\mathbf{u}$-space.

2.1. Kriging Metamodel

.

Consider a set of m observations collected in the matrix $\mathcal{D}^{(m)} \equiv [\mathbf{d}^{(1)}, \mathbf{d}^{(2)}, ..., \mathbf{d}^{(m)}]$ of order $(n+1) \times m$, where $\mathbf{d}^{(k)} = \{\mathbf{u}^{(k)}, y_k\}$, $k = 1, 2, ..., m$ with $y_k = G\left(\mathbf{u}^{(k)}\right)$. The idea behind Kriging is to approximate the target limit state function $G(\mathbf{u})$ through a Gaussian stochastic process $\epsilon(\mathbf{u})$

$$\begin{aligned} \hat{G}(\mathbf{u}) &= \mu(\mathbf{u}) + \epsilon(\mathbf{u}) \\ &= \mathbf{h}^T(\mathbf{u})\,\mathbf{w} + \epsilon(\mathbf{u}), \end{aligned} \tag{3}$$

where $\mu(\mathbf{u})$ is the mean function approximating the output, being $h(\mathbf{u})$ a vector collecting a set of p basis functions $h_1(u), h_2(u), ..., h_p(u)$, whereas $\mathbf{w} \equiv \{w_1, w_2, ..., w_p\}$ collects the parameters of the regression. In eq.(3) the contribution $\epsilon(\mathbf{u})$ describes the uncertain departure of $\hat{G}(\mathbf{u})$ from its mean value and it is modelled through a stationary zero-mean Gaussian process, whose autocovariance function is

$$Cov\left(\mathbf{u}, \mathbf{u}'\right) = \sigma_\epsilon^2 R_\epsilon\left(\mathbf{u}, \mathbf{u}'; \theta_\epsilon\right), \tag{4}$$

where $R_\epsilon(\cdot, \cdot)$ is a suitable autocorrelation function providing the dependence structure, σ_ϵ^2 is the variance of the process, and θ_ϵ collects all the parameters of $R_\epsilon(\cdot, \cdot)$. Several correlation functions exist in literature; in practice, square exponential models are often adopted

$$R_\epsilon\left(\mathbf{u},\mathbf{u}';\theta_\epsilon\right)=\exp\left(-\sum_{k=1}^{n}\theta_{\epsilon,k}|\ u-u'_k|^2\right). \tag{5}$$

2.2. Estimation of the Kriging parameters

.

For the given set of m points $\{\mathbf{u}^{(1)},\mathbf{u}^{(2)},...,\mathbf{u}^{(m)}\}$, eq.(3)becomes

$$\mathbf{y}=\mathbf{H}\mathbf{w}+\mathbf{E} \tag{6}$$

where

$$\mathbf{y}=\begin{Bmatrix} G\left(\mathbf{u}^{(1)}\right)\\ G\left(\mathbf{u}^{(2)}\right)\\ ...\\ G\left(\mathbf{u}^{(m)}\right)\end{Bmatrix},\qquad \mathbf{H}=\begin{bmatrix} h_1\left(\mathbf{u}^{(1)}\right) & ... & h_p\left(\mathbf{u}^{(1)}\right)\\ h_1\left(\mathbf{u}^{(2)}\right) & ... & h_p\left(\mathbf{u}^{(2)}\right)\\ ... & & \\ h_1\left(\mathbf{u}^{(m}\right) & ... & h_p\left(\mathbf{u}^{(m}\right)\end{bmatrix},\qquad \mathbf{E}=\begin{Bmatrix} \epsilon\left(\mathbf{u}^{(1)}\right)\\ \epsilon\left(\mathbf{u}^{(2)}\right)\\ ...\\ \epsilon\left(\mathbf{u}^{(m)}\right)\end{Bmatrix}. \tag{7}$$

By assuming that the vector of model error E is Gaussian, the parameters σ_ϵ^2 and w can be determined through the Maximum Likelihood Estimation (MLE):

$$\begin{aligned}\left(\hat{\mathbf{w}},\hat{\sigma}_\epsilon^2\right)&=\underset{\mathbf{w},\sigma_\epsilon^2}{\operatorname{argmax}}\, L\left(\mathbf{w},\sigma_\epsilon^2|\mathbf{y}\right)\\ &=\frac{1}{\left(2\pi\sigma_\epsilon^2\right)^{\frac{n}{2}}\sqrt{|\mathbf{R}_\epsilon|}}\exp\left[-\frac{1}{2\sigma_\epsilon^2}\left(\mathbf{y}-\mathbf{H}\mathbf{W}\right)^T\mathbf{R}_\epsilon^{-1}\left(\mathbf{y}-\mathbf{H}\mathbf{W}\right)\right],\end{aligned} \tag{8}$$

where $\mathbf{R}_\epsilon\left(\theta_\epsilon\right)$ is the matrix of correlation between the training points

$$\mathbf{R}_\epsilon\left(\theta_\epsilon\right)=\begin{bmatrix} R_\epsilon\left(\mathbf{u}^{(1)},\mathbf{u}^{(1)}\right) & ... & R_\epsilon\left(\mathbf{u}^{(1)},\mathbf{u}^{(m)}\right)\\ R_\epsilon\left(\mathbf{u}^{(2)},\mathbf{u}^{(1)}\right) & ... & R_\epsilon\left(\mathbf{u}^{(2)},\mathbf{u}^{(m)}\right)\\ ... & & \\ R_\epsilon\left(\mathbf{u}^{(m)},\mathbf{u}^{(1)}\right) & ... & R_\epsilon\left(\mathbf{u}^{(m)},\mathbf{u}^{(m)}\right).\end{bmatrix} \tag{9}$$

The solutions of eqs.(8) and (9) are

$$\begin{aligned}\hat{\mathbf{w}}&=\left(\mathbf{H}^T\mathbf{R}_\epsilon^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}_\epsilon^{-1}\mathbf{y},\\ \hat{\sigma}_\epsilon^2&=\frac{1}{m}\left(\mathbf{y}-\mathbf{H}\mathbf{W}\right)^T\mathbf{R}_\epsilon^{-1}\left(\mathbf{y}-\mathbf{H}\mathbf{W}\right).\end{aligned} \tag{10}$$

2.3. Kriging prediction

The Kriging estimator at a new point is by definition a Gaussian random variate $\hat{G}(\mathbf{u})$ obtained as the Best Linear Unbiased Estimator (BLUE) of $G(\mathbf{u})$, conditioned to the observed samples. It is a linear combination of the samples and it the unbiased estimator with minimum variance. Given the

set of m training points $\mathbf{d}^{(k)} = \{\mathbf{u}^{(k)}, y_k\}$, $k = 1, 2, ..., m$, the definition of Gaussian process implies that $\{\mathbf{y}, \hat{G}(\mathbf{u})\}$ are jointly Gaussian

$$\begin{Bmatrix} \mathbf{y} \\ \hat{G} \end{Bmatrix} \sim \mathcal{N}\left(\begin{Bmatrix} \mu(\mathbf{u}) \\ \hat{\mu} \end{Bmatrix}, \quad \sigma_\epsilon^2 \begin{bmatrix} \mathbf{R} & \mathbf{r}(\mathbf{u}) \\ \mathbf{r}^T(\mathbf{u}) & 1 \end{bmatrix} \right) \tag{11}$$

where $\mu(\mathbf{u}) = \mathbf{h}^T(\mathbf{u})\,\hat{\mathbf{w}}$ and $\mathbf{r}(\mathbf{u})$ is a vector of order m which collects the correlations between the point $\mathbf{u}$ and the m training points, i.e. $r_i(\mathbf{u}) = R_\epsilon\left(\mathbf{u}, \mathbf{u}^{(i)}\right)$. Using eq.(11), the posterior distribution is obtained

$$f\left(\hat{G}|\mathbf{y}\right) = \mathcal{N}\left(\hat{G}|\mu_{\hat{g}}(\mathbf{u}), \sigma_{\hat{g}}^2(\mathbf{u})\right) \tag{12}$$

where the posterior Kriging mean and variance are

$$\begin{aligned} \mu_{\hat{g}}(\mathbf{u}) &= \mathbf{h}^T(\mathbf{u})\,\hat{\mathbf{w}} + \mathbf{r}^T(\mathbf{u})\,\mathbf{R}_\epsilon^{-1}\left(\mathbf{y} - \mathbf{H}\hat{\mathbf{w}}\right), \\ \sigma_{\hat{g}}(\mathbf{u}) &= \sigma_\epsilon^2\left[1 - \mathbf{r}^T(\mathbf{u})\,\mathbf{R}_\epsilon^{-1}\mathbf{r}(\mathbf{u}) + \mathbf{q}^T(\mathbf{u})\left(\mathbf{H}^T\mathbf{R}_\epsilon^{-1}\mathbf{H}\right)^{-1}\mathbf{q}(\mathbf{u})\right]. \end{aligned} \tag{13}$$

with $\mathbf{q}(\mathbf{u}) = \mathbf{H}\mathbf{R}_\epsilon^{-1}\mathbf{r}(\mathbf{u}) - \mathbf{h}(\mathbf{u})$.

The Kriging surrogate has several interesting features. First, at the training points the prediction is exact and the corresponding variance is zero. Second, when the number of training points increases, the overall variance of the process decreases (i.e. $\sigma_\epsilon \to 0$ when $m \to \infty$). Since the prediction $\hat{G}(\mathbf{u})$ is a Gaussian random variable of mean $\mu_{\hat{g}}(\mathbf{u})$ and variance $\sigma_{\hat{g}}^2(\mathbf{u})$, it is possible to derive confidence bounds over the prediction. The variance process $\sigma_{\hat{g}}^2(\mathbf{u})$ is typically used as an error measure of the epistemic uncertainty. If the prediction of $\hat{G}(\mathbf{u})$ in a point $\mathbf{u}_{new}$ (different from the previously selected m training points) shows high variance $\sigma_{\hat{g}}^2(\mathbf{u}_{new})$, this implies non-negligible epistemic uncertainty of the model prediction $\mu_{\hat{g}}(\mathbf{u}_{new})$. Therefore, to improve the global accuracy of the surrogate, it can be convenient choose a new training point $\mathbf{u}^{(m+1)} \equiv \mathbf{u}_{new}$. This last feature allows to defining adaptive sampling plans for the modelling of the response surface.

2.4. Adaptive Kriging (AK)

To address AK, in literature, is defined a *learning function* capable of detecting a new training point $\mathbf{u}^{(m+1)}$ in a way to improve the overall accuracy of the Kriging metamodel. Typically, in AK the following steps are followed:

1. Initialization
   a) Generate a large set of N_c training candidates $\mathcal{P}_c \equiv \{\mathbf{u}_c^{(1)}, \mathbf{u}_c^{(2)}, ..., \mathbf{u}_c^{(Nc)}\}$
   b) Generate an initial training set of m points $\mathcal{P}^{(m)} \equiv \{\mathbf{u}^{(1)}, \mathbf{u}^{(2)}, ..., \mathbf{u}^{(m)}\}$, with $\mathbf{u}^{(k)} \in \mathcal{P}_c$, $m \ll N_c$
   c) Evaluate the limit state for the m training points, $y_k = G\left(\mathbf{u}^{(k)}\right)$
   d) Collect the dataset in $\mathcal{D}^{(m)} \equiv \{\mathbf{d}^{(1)}, \mathbf{d}^{(2)}, ..., \mathbf{d}^{(m)}\}$, with $\mathbf{d}^{(k)} \equiv \{\mathbf{u}^{(k)}, y_k\}$

2. Build a first Kriging model $\hat{G}_{(0)}\left(\mathbf{u}|\mathbf{w},\sigma_{\epsilon}^{2},\theta_{\epsilon}\right)$ trained over $\mathcal{D}^{(m)}$
3. Evaluate the failure probability $P_{f,(0)}$ using the first Kriging model $\hat{G}_{(0)}$
4. Choose the learning function $\mathcal{L}\left(\mathbf{u}\right)$
5. Adaptive sampling
   a) $r=0$
   b) $r=r+1$
   c) Through the selected learning function $\mathcal{L}\left(\mathbf{u}\right)$, pick from the set $\mathcal{P}_{c}$ the best "new point" of the sampling plan $\mathbf{u}^{(m+r)}$; no evaluation of the limit state is needed in this stage
   d) Evaluate the limit state for the chosen point, $y_{m+r}=G\left(\mathbf{u}^{(m+r)}\right)$
   e) Collect the dataset in $\mathcal{D}^{(m+r)}\equiv\{\mathbf{d}^{(1)},\mathbf{d}^{(2)},...,\mathbf{d}^{(m+r)}\}$
   f) Build a Gaussian model $\hat{G}_{(r)}\left(\mathbf{u}|\mathbf{w},\sigma_{\epsilon}^{2},\theta_{\epsilon}\right)$ trained over $\mathcal{D}^{(m+r)}$
   g) Evaluate the failure probability $P_{f,(r)}$ using the updated Kriging model $\hat{G}_{(r)}$
   h) Check convergence: if satisfied, stop, otherwise go to step 4(b).

A largely adopted learning function is the so-called $\mathcal{U}$-function (Echard et al., 2011)

$$\mathcal{U}\left(\mathbf{u}\right)=\frac{\left|\mu_{\hat{G}}\left(\mathbf{u}\right)\right|}{\sigma_{\hat{G}}\left(\mathbf{u}\right)} \tag{14}$$

which is adopted in the approach AK-MCS (Adaptive Kriging + Monte Carlo Simulation). A small value of $\mathcal{U}\left(\mathbf{u}\right)$ implies that the Kriging model is close to the limit state surface (i.e. $\mu_{\hat{G}}\left(\mathbf{u}\right)=0$) or the uncertainty epistemic $\sigma_{\hat{G}}\left(\mathbf{u}\right)$ is large, or both. The selected new point of the adaptive procedure has the minimum value of $\mathcal{U}\left(\mathbf{u}\right)$

$$\mathbf{u}^{(m+1)}=\underset{\mathbf{u}\in\mathcal{P}_{c}}{\operatorname{argmin}}\left[\mathcal{U}\left(\mathbf{u}_{c}^{(1)}\right),\mathcal{U}\left(\mathbf{u}_{c}^{(2)}\right),...,\mathcal{U}\left(\mathbf{u}_{c}^{(N_{c})}\right)\right]. \tag{15}$$

3. Information Theory for Uncertainty Quantification and Structural Reliability Analysis

Recently, one of the authors has underlined the relationships between Information Theory with Uncertainty Quantification and Risk Analysis (Alibrandi and Mosalam, 2019; Alibrandi and Mosalam, 2020; Alibrandi and Mosalam, 2021). In the following, some findings that are relevant for active-learning surrogate modelling design are discussed.

3.1. Copula

In the basic problem of structural reliability, see eq.(1), the set of basic random variables $x_1, x_2, ..., x_n$ is given, whose marginal Probability Density Functions (PDF) and Cumulative Distribution Function (CDF) are $f_i(x_i)$ and $F_i(x_i)$, respectively. Consider the transformation $v_i = F_i(x_i)$, $i = 1, 2, ..., n$. The vector $\mathbf{v} = \{v_1, v_2, ..., v_n\}$ has uniformly distributed marginals, whereas the copula is the joint CDF $F(v_1, v_2, ..., v_n)$. The copula describes the *stochastic dependence* between the random variables and it is unique. The theorem of Sklar allows to describe any multivariate distribution as the product of its marginal PDFs $f_i(x_i)$ and a copula density $c(v_1, v_2, ..., v_n)$,

$$f\left(x_1, x_2, ..., x_n\right) = c\left(v_1, v_2, ..., v_n\right) f_1\left(x_1\right) f_2\left(x_2\right) ... f_n\left(x_n\right). \tag{16}$$

Many copulas have been proposed to model the dependence between random variables, and each one of them imposes a different dependence structure on the data. One commonly applied parametric copula is the Gaussian, whose copula density is

$$c_G\left(\mathbf{v}\right) = \frac{1}{\sqrt{|\mathbf{R}_c|}} \exp\left[-\frac{1}{2}\mathbf{u}^T\left(\mathbf{R}_c^{-1} - \mathbf{I}\right)\mathbf{u}\right], \tag{17}$$

where $\mathbf{u} = \{u_1, u_2, ..., u_n\}$ collects a set of normal standard random variables mapped from $\mathbf{v}$, i.e. $u_i = \Phi^{-1}\left(v_i\right)$, with $\Phi(\cdot)$ the normal standard CDF, and $\mathbf{R}_c$ the matrix of correlation of the copula (different from the matrix of correlation $\mathbf{R}_\epsilon$ of the model error of the Kriging model), whose elements may be defined through linear correlations $R_{ij} = \rho\left(v_i, v_j\right)$. For further details about copula, see (Embrechts et al., 2000; Dutfoy and Lebrun, 2009; Alibrandi and Mosalam, 2020; Alibrandi and Mosalam, 2021).

3.2. Entropy

In information theory, the entropy of a random variable is a *measure of uncertainty* and it can be interpreted as the degree of information that the observation of the variable gives. The (differential) entropy of a continuous-valued random variable X with probability density function $f(x)$ is

$$H\left(x\right) = -\int f\left(x\right) \log f\left(x\right)\ dx. \tag{18}$$

For a Gaussian random variable of standard deviation σ the entropy is

$$H_G\left(x\right) = \frac{1}{2}\left(1 + \log 2\pi\right) + \log \sigma. \tag{19}$$

In the case of Gaussian distribution, the entropy (i.e. degree of uncertainty) is proportional to the standard deviation; this is expected, since the uncertainty of a Gaussian distribution is described uniquely by σ. Given the limit state function $y = G\left(x_1, x_2, ..., x_n\right)$, the entropy $H\left(G\right)$ provides the uncertainty over G

$$H\left(G\right) = -\int f\left(g\right) \log f\left(g\right)\ dg. \tag{20}$$

If in particular, the limit state function $G(\mathbf{u})$ follows a Gaussian distribution or it is well approximated by a parametric distribution, then its entropy can be determined in closed form as a function of its parameters.

3.3. Mutual Information

The mutual information $I(x_1, x_2)$ of two random variables X_1 and X_2 measures the information that they share. Thus, it is a measure of their dependence and it is defined as

$$MI(x_1, x_2) = \iint f(x_1, x_2) \log \frac{f(x_1, x_2)}{f_1(x_1) f_2(x_2)} \, dx_1 \, dx_2, \tag{21}$$

Different from the coefficient of correlation ρ, the MI is able to describe full dependence between variables, not just linear correlations (i.e. second-order dependencies). While the attractive properties of the mutual information as a measure of dependence are well known since long time, it was not readily adopted in statistics and uncertainty quantification because it was originally applied mostly to discrete random variables, and its evaluation is in such cases quite involved. This gap of knowledge has been analyzed in (Alibrandi and Mosalam, 2020; Alibrandi and Mosalam, 2021), where connections with copulas have been underlined. By replacing eq.(16) into eq.(21) one has

$$MI(x_1, x_2) = \iint f(x_1, x_2) \log c(v_1, v_2) \; dv_1 \, dv_2 = E[c(v_1, v_2)], \tag{22}$$

and the MI is evaluated as the expected value of the logarithm of the copula density. In the case of Gaussian copula, see eqs. (17) and (22), one has

$$MI_G(x_1, x_2) = -\frac{1}{2} \log |\mathbf{R}_c|. \tag{23}$$

3.4. Informational Coefficient of Correlation

The informational coefficient of correlation ρ_{MI} is defined as

$$\rho_{MI}(x_1, x_2) = \sqrt{1 - \exp[-2MI(x_1, x_2)]} \tag{24}$$

and it is a *proper measure of dependence*. In particular $\rho_{MI} = 0$ for independent (not just uncorrelated) random variables, whereas for perfect dependence $\rho_{MI} = 1$. Also, ρ_{MI} does not suffer any limitations of the linear coefficient of correlation, and they coincide in the case of joint Gaussian distribution. For further details see (Linfoot, 1957; Alibrandi and Mosalam, 2020; Alibrandi and Mosalam, 2021)

4. Global Sensitivity Analysis

Given the limit state function $y = G(u_1, u_2, ..., u_n)$, the Mutual Information $MI(G, u_k)$ can be interpreted as the *reduction in uncertainty* about G after observing u_k:

$$MI\left(G,u_k\right)=H\left(G\right)-H\left(G|u_k\right), \tag{25}$$

where $H(G)$ is the entropy (i.e. uncertainty) of G, whereas $H(G|u_k)$ is the conditional entropy of G given u_k. This shows, also, that on average conditioning reduces entropy. In Fig.1 a Diagram of Venn is shown, where the area of the sets represents their uncertainty, whereas their intersection is the mutual information.

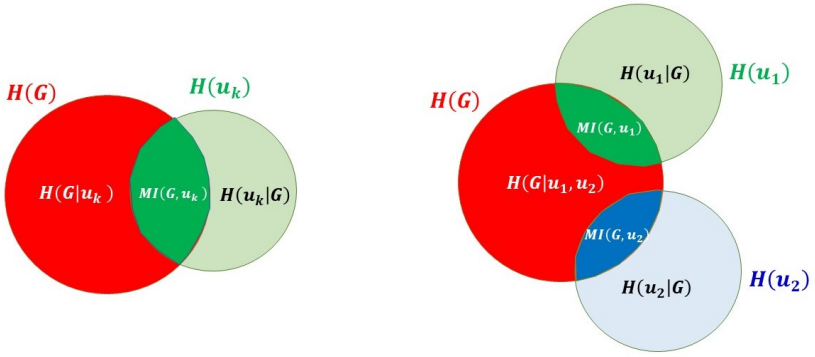


Figure 1. Diagrams of Venn for Mutual Information

Since the normal standard random variables are independent, their mutual informations are zero, i.e. $MI(u_i,u_j)=0$; thus, since there are no interactions between the variables $u_1,u_2,...,u_n$, the informational coefficient of correlation $\rho_{MI}\left(G,u_k\right)$ between G and u_k provides a proper measure of the importance of the variables.

$$\delta_k=\rho_{MI}\left(G,u_k\right) \tag{26}$$

where δ_k are the *informational sensitivity indices*.

5. Adaptive Kriging based on the Mutual Information (AK-MI)

5.1. Adaptive Kriging based on the Entropy: AK-H

Since the entropy can be considered as a generalized measure of uncertainty, which includes the standard deviation in the case of Gaussian random variables, following the idea of the $\mathcal{U}$-function, see eq.(14) (Echard et al., 2011), a learning function based on the entropy of $\hat{G}\left(\mathbf{u}\right)$ is proposed:

$$\mathcal{H}\left(\mathbf{u}\right)=\frac{\left|\mu_{\hat{G}}\left(\mathbf{u}\right)\right|}{H_{\hat{G}}\left(\mathbf{u}\right)}=\frac{\left|\mu_{\hat{G}}\left(\mathbf{u}\right)\right|}{\frac{1}{2}\left(1+\log 2\pi\right)+\log\sigma_{\hat{G}}\left(\mathbf{u}\right)}. \tag{27}$$

A small value of $\mathcal{H}\left(\mathbf{u}\right)$ implies that the Kriging model is close to the limit state surface (i.e. $\mu_{\hat{G}}\left(\mathbf{u}\right)=0$) or the epistemic uncertainty $H_{\hat{G}}\left(\mathbf{u}\right)$ is large, or both. The selected new point of the adaptive procedure has the minimum value of $\mathcal{H}\left(\mathbf{u}\right)$:

$$\mathbf{u}^{(m+1)} = \underset{\mathbf{u}\in\mathcal{P}_c}{\operatorname{argmin}} \left[\mathcal{H}\left(\mathbf{u}_c^{(1)}\right), \mathcal{H}\left(\mathbf{u}_c^{(2)}\right), ..., \mathcal{H}\left(\mathbf{u}_c^{(N_c)}\right)\right]. \tag{28}$$

5.2. Adaptive Learning based on the Mutual Information: AL-MI

Typically, adaptive learning is implemented together with Kriging for its capabilities to describe the error of the epistemic uncertainty. However many well-performing surrogate models do not manage this feature: Thus, they are not used inside a procedure of Adaptive Learning. Alternatively, variants of these methods have been proposed in order to describe the model uncertainty.

Having in mind the adoption of Adaptive Learning with surrogate models different from Kriging, a second learning function is here proposed. Given the limit state function $G = G(u_1, u_2, ..., u_n)$, the total uncertainty is

$$\begin{aligned} H(G, u_1, u_2, ..., u_n) &= H(G|u_1, u_2, ..., u_n) + H(u_1, u_2, ..., u_n) \\ &= H(G|u_1, u_2, ..., u_n) + \sum_{k=1}^{n} H(u_i) = \\ &= H(G|u_1, u_2, ..., u_n) + \frac{n}{2}(1 + \log 2\pi), \end{aligned} \tag{29}$$

and the total uncertainty is reduced to the conditional entropy of G. From Fig.1 it is seen that the maximum uncertainty over G is equivalent to the minimimum of the sum of the mutual information between G and the random variables $u_1, u_2, ..., u_n$, that is

$$\max_{\mathbf{u}} H(G|u_1, u_2, ..., u_n) = \min_{\mathbf{u}} \sum_{k=1}^{n} MI(G, u_k). \tag{30}$$

Since the informational coefficient of correlation is bounded between zero and one, $0 \leq \rho_{MI} \leq 1$, a proxy in eq.(27) for the epistemic uncertainty is $1 - \rho_{MI}(\mathbf{u})$. The proposed learning function based on the mutual information is, then,

$$\mathcal{MI}(\mathbf{u}) = \frac{|\mu_{\hat{G}}(\mathbf{u})|}{1 - \rho_{MI}(\mathbf{u})}. \tag{31}$$

The selected new point of the adaptive procedure has the minimum value of $\mathcal{MI}(\mathbf{u})$

$$\mathbf{u}^{(m+1)} = \underset{\mathbf{u}\in\mathcal{P}_c}{\operatorname{argmin}} \left[\mathcal{MI}\left(\mathbf{u}_c^{(1)}\right), \mathcal{MI}\left(\mathbf{u}_c^{(2)}\right), ..., \mathcal{MI}\left(\mathbf{u}_c^{(N_c)}\right)\right]. \tag{32}$$

6. Numerical Application

Let us consider a limit state function with multiple design points, modified from (Der Kiureghian and Dakessian, 1998)

$$g(u_1, u_2) = 3 - u_2 - 0.3u_1^2, \tag{33}$$

where u_1 and u_2 are normal standard random variables. All the reliability computations have been developed through OpenAIUQ software of data-driven uncertainty quantification and risk-analysis (Alibrandi et al., 2021).

6.1. Probabilistic sensitivity analysis

First, a sensitivity analysis is pursued. The informational sensitivity indices are $\delta_1 \equiv \rho_{MI}(G, u_1) = 0.33$ and $\delta_2 \equiv \rho_{MI}(G, u_2) = 0.92$, see eq.(26).The results show that both variables are important, and that to reduce the uncertainty over G it is more interesting to gain information about u_2. Note that the informational coefficient of correlation ρ_{MI} may detect also nonlinear correlations, which is not so with the linear coefficient of correlation ρ. For sake of illustration, in Fig.2 it is shown the limit state function $y = G(u_1, u_2)$, and the scatter plots of the points (u_1, G) and (u_2, G). The application of the linear coefficient of correlation ρ provides $\rho(G, u_1) = 0.0$ and $\rho(G, u_2) = 0.92$. As expected, only linear correlation may be well estimated by ρ.

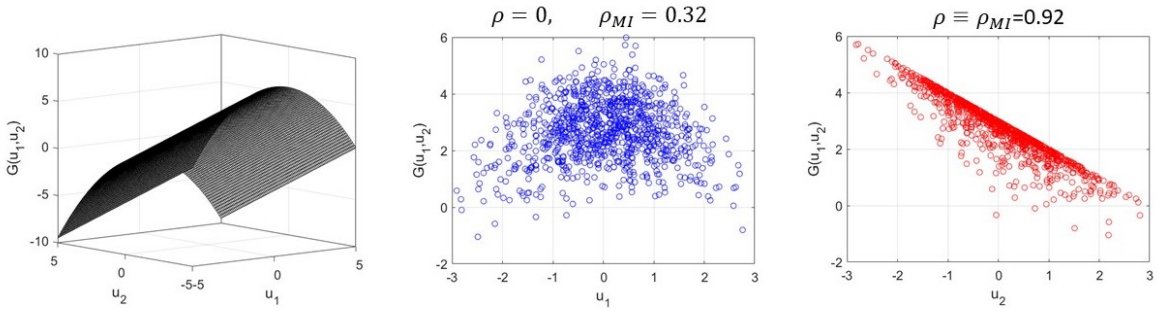


Figure 2. Sensitivity of the limit state function

6.2. Adaptive Kriging

The "reference" failure probability $P_f = 9.737 \times 10^{-3}$ has been evaluated as the average values of 50 MCS runs, each one with a coefficient of variation $\nu = 1\%$. The design point is found to be $\mathbf{u}^* = \{2.11, 1.66\}$; the FORM solution provides $P_{f,FORM} = 3.60 \times 10^{-3}$, with a relative error $e_{FORM} = 63.02\%$. For this toy example, aimed at benchmarking the procedure, the set of the candidate points $\mathcal{P}_c$ includes a uniform grid $u_1 \in [-5, 5]$, $u_2 \in [-5, 5]$, with a step $\Delta u = 0.10$.

For the Kriging model, an initial sampling plan of $m = 9$ points is chosen, distributed along an uniform grid $u_1 \in [-1, 1]$, $u_2 \in [-1, 1]$ with a step $\Delta u = 1$, see Fig.3. To analyze the capabilities of convergence of Kriging, the design point is not included in the initial training set $\mathcal{D}^{(9)}$. A first Kriging model $\hat{G}_{(0)}$ is evaluated and represented in Fig.3 (blue line), together with the target limit state function (black line) and the initial training points (blue points). In the right panel of Fig.3

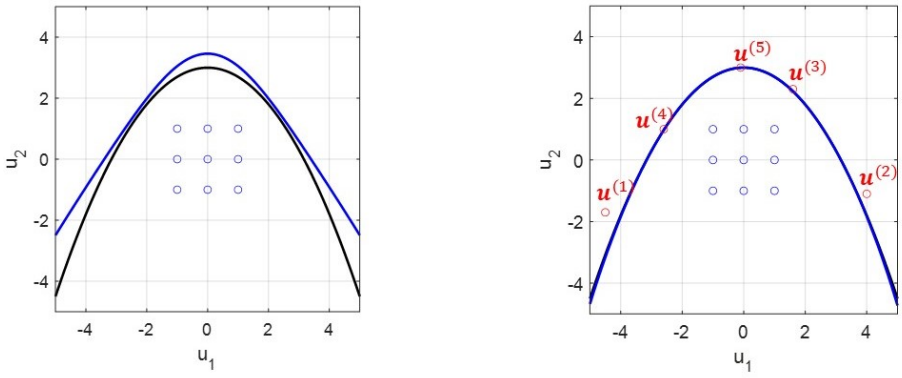


Figure 3. Adaptive Kriging, limit state

the first five points of the adaptive plan are shown (red dots), together with the corresponding Kriging model $\hat{G}_{(5)}$ (blue line).

To check the convergence properties, we represent in Fig.4 (on the left) the loss function of AK-MI, evaluated with respect to all the candidate points of the set $\mathcal{P}_c$. On the right panel of the Figure, it is represented the relative error of FORM (magenta line) and AK (blue line) with respect to the target failure probability P_f. The failure probabilities with Kriging model are estimated through MCS with a coefficient of variation $\nu = 5\%$. It is seen that after only two points of the adaptive sampling there is an error of approximately 10%, which is inside the range of the chosen coefficient of variation. The red line represents the average value of the probabilities estimated by the Kriging models $\hat{G}_{(k)}$ when more training points are added, which shows the convergence toward the target failure probability P_f.

7. Conclusions

In this paper, we have shown the application of information theory for probabilistic sensitivity analysis and surrogate modelling with adaptive sampling. One of the authors has recently proposed the adoption of the informational coefficient of correlation as a measure of dependence between the random variables, instead of the largely adopted linear coefficient of correlation.

In this paper two main contributions are introduced. First, it is shown that the informational coefficient of correlation ρ_{MI} can be used for probabilistic sensitivity analysis. One application shows how it outperforms the classical coefficient of correlation ρ. With reference to surrogate modelling with adaptive sampling, typically the epistemic uncertainty is modelled through the variance of the Gaussian process. Two novel learning functions for adaptive sampling are proposed. The first, called the $\mathcal{H}$- function, giving rise to the method AK-H (Adaptive Kriging - Entropy), describes the model error through the entropy; the second, called the $\mathcal{MI}$- function, giving rise to the method AL-MI (Active Learning - Mutual Information), describes the model error through the Mutual Information.

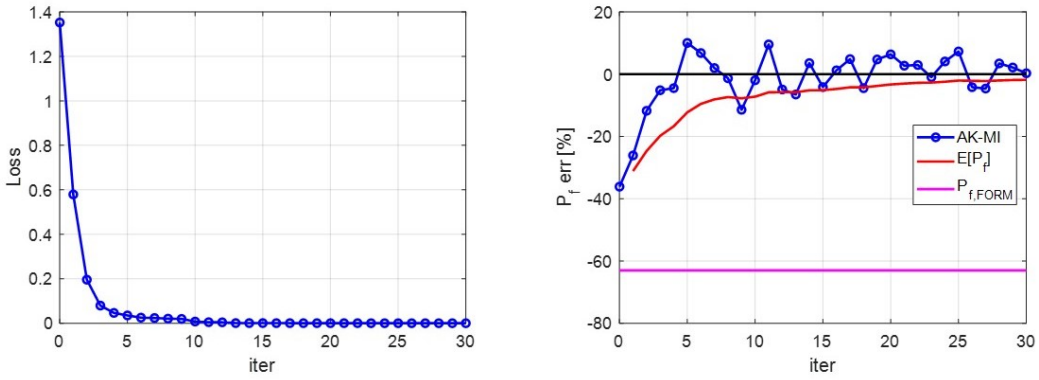


Figure 4. Loss function and error of the Kriging model

The $\mathcal{MI}$-function has the peculiarity that it allows the implementation of adaptive learning in any kind of surrogate modelling, which can differ from Kriging. This capability is attractive, given the high number of surrogate models which do not provide estimates of errors over their predictions. Further research has to be developed to check the computational performances of the proposed approaches, also in comparison with the other learning functions existing in the literature.

References

Alibrandi, U. A response surface method for stochastic dynamic analysis. *Reliab. Eng. Syst. Saf.*, 26:44–53, 2014.

Alibrandi, U. and L. V. Anderssen and K. M. Mosalam OpenAIUQ: data driven uncertainty quantification, risk analysis for risk-informed digital twins. In: *31st European Safety and Reliability Conference*, Angers, France, 2021

Alibrandi, U. and K. M. Mosalam Distribution with Independent Components for Uncertainty Quantification and Structural Reliability Analysis. In: *P13th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP13*, Seoul, South Korea, 2019

Alibrandi, U. and K. M. Mosalam Information Theory for data-driven Risk Analysis. In: *29th European Safety and Reliability Conference*, Hannover, Germany, 2020

Alibrandi, U. and K. M. Mosalam Information Theory and Machine Learning: Uncertainty Quantification and Risk Analysis revisited. *Reliability Engineering and System Safety*, (submitted)

Au, S. K. and J. L. Beck Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4):263–277, 2001.

Berveiller, M. and B. Sudret and M. Lemaire Stochastic finite elements: a non-intrusive approach by regression. *Eur. J. Comput. Mech.*, 15(1-3):81–92, 2006.

Bichon, B. J. and M. S. Eldred and L. P. Swiler and S. Mahadevan and J. M. MaFarland Efficient global reliability analysis for nonlinear implicit performance functions. *AIAA J*, 46(10):2459–2468, 2008.

Blatman, G. and B. Sudret An adaptive algorithm to build-up sparse polynomial chaos expansion for stochastic finite elements. *Prob. Eng. Mech.*, 25(2):183–197, 2010.

Bucher, C. and U. Burgound A fast and efficient response surface approach for structural reliability problems. *Structural Safety*, 7:57–66, 1990.

Bucher, C. and T. Most A comparison of approximate response functions in structural reliability analysis. *Probab. Eng. Mech.*, 23:154–163, 2008.

Choi, S. and R. Grandhi and R. Canfield Structural Reliability under non-Gaussian stochastic behavior. *Computers & Structures*, 82:1113–1121, 2004.

Chowdury, R. and B. N. Rao and A. M. Prasad High-dimensional model representation for structural reliability analysis. *J. Commun. Numer. Methods Eng.*, 25:301–307, 2005.

De Angelis, M. and E. Patelli and M. Beer Advanced line sampling for efficient robust reliability analysis. *Structural Safety*, 52:170–182, 2015.

Der Kiureghian, A. Analysis of structural reliability under parameter uncertainties. *Probabilistic Engineering Mechanics*, 23:351–358, 2008.

Der Kiureghian, A. and T. Dakessian Multiple design points in first- and second-order reliability. *Structural Safety*, 20:37–49, 1998.

Ditlevsen, O. and H. O. Madsen. Structural Reliability Methods. Wiley, 1996.

Duborg, V. and B. Sudret and F. Deheeger Metamodel-based importance sampling for structural reliability analysis. *Probab Eng Mech*, 33:47–57, 2013.

Dutfoy, A. and R. Lebrun Practical approach to dependence modelling using copulas. *Proc. IMech. E., Part O: J. Risk and Reliability*, 223:347–361, 2009.

Echard, B. and N. Gayton and M. Lemaire AK-MCS: an active learning reliability method combining Kriging and Monte Carlo simulation. *Struct Saf*, 33(2):145–154, 2011.

Echard, B. and N. Gayton and M. Lemaire and N. Relun A combined Importance Sampling and Kriging reliability method for small failure probabilities with time-demanding numerical models. *Reliab Eng Syst Saf*, 111:232–240, 2013.

Embrechts, A. and A. McNeil and D. Straumann Correlation and dependency in risk man-agement: properties and pitfalls. *Review of Economics and Statistics*, 86, 2000.

Faravelli, L. Response-surface approach for reliability analysis. *J. Eng. Mech.*, 115:2763–2781, 1989.

Forrester, A. I.J. and A. SObester and A. J. Keane Engineering design via surrogate modelling. John Wiley & Sons, 2006.

Gaspar, B. and A. P. Teixeira and C. Guedes Soares Assessment of the efficiency of Kriging surrogate models for structural reliability analysis. *Prob. Eng. Mech*, 37:24–34, 2014.

Huang, X. and J. Chen and H. Zhu Assessing small failure probabilities by AK–SS: An active learning method combining Kriging and Subset Simulation. *Structural Safety*, 59:86–95, 2016.

Kaymaz, I. Application of kriging method to structural reliability problem. *Struct. Saf.*, 27(2):133–151, 2005.

Li, G. and C. Rosenthal and H. Rabitz High-dimensional model representation. *J. Physical Chem. A*, 106:7765–7777, 2001.

Linfoot, E. An informational measure of correlation. *Information and Control*, 1:85–89, 1957.

LV, Z. and Z. Lu and P. Wang A new learning function for Kriging and its applications to solve reliability problems in engineering. *Computers and Mathematics with Applications*, 70:1182–1197, 2015.

Madsen, H. O., S. Krenk and N. C. Lind. Methods of Structural Safety. Prentice-Hall, 1986.

Melchers, R. E. Importance Sampling in structural systems. *Structural Safety*, 6(1):3–10, 1989.

Melchers, R. E. Structural Reliability, analysis and prediction. New York: John Wiley & Sons, 1999.

Rajashekhar, M. R. and B. R. Ellingwood A new look at the response surface approach for reliability analysis. *Structural Safety*, 12:205–220, 1993.

Rasmussen, C. E. and C. K.I. Williams Gaussian Processes for Machine Learning. MIT Press, 2006.

Rubinstein, R. and D. Krose Simulation and the Monte Carlo method. Wiley Series in Probability and Statistics: Wiley, 2008.

Sudret, B. and A. Der Kiureghian Comparison of finite element reliability methods. *Prob. Eng. Mech.*, 17:337–348, 2002.

Zhang, X. and L. Wang and J. D. Sørensen REIF: A novel active-learning function toward adaptive Kriging surrogate models for structural reliability analysis. *Reliability Engineering and System Safety*, 185:440–454, 2019.

Zheng, Y. and P. K. Das Improved response surface method and its application to stiffened plate reliability analysis. *Eng. Struct.*, 22:544–551, 2000.

An efficient approach for reliability-based optimization of linear dynamical structures subject to Gaussian excitation

Danko Jerez
Institute for Risk and Reliability, Leibniz Universität Hannover, 30167 Hannover, Germany, danko.jerez@irz.uni-hannover.de.

Hector Jensen
Department of Civil Engineering, Santa Maria University, 2390302 Valparaiso, Chile, hector.jensen@usm.cl.
International Joint Research Center for Engineering Reliability and Stochastic Mechanics, Tongji University, 200092 Shanghai, China.

Marcos Valdebenito
Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, 2562340 Viña del Mar, Chile, marcos.valdebenito@uai.cl

Mauricio Misraji
Department of Civil Engineering, Santa Maria University, 2390302 Valparaiso, Chile, mauricio.misraji@alumnos.usm.cl.

Michael Beer
Institute for Risk and Reliability, Leibniz Universität Hannover, 30167 Hannover, Germany, beer@irz.uni-hannover.de.
International Joint Research Center for Engineering Reliability and Stochastic Mechanics, Tongji University, 200092 Shanghai, China.
Institute for Risk and Uncertainty, School of Engineering, University of Liverpool, L69 7ZF Liverpool, United Kingdom.

Abstract. One pivotal task of structural engineering is to devise optimal structures according to a given criterion, while satisfying certain performance requirements that take into account all related uncertainties. This work proposes an efficient approach for reliability-based optimization of linear dynamical structures subject to stochastic Gaussian excitation, where the constraints are given in terms of first-excursion probabilities. Directional Importance Sampling (DIS) is considered for reliability assessment, which also estimates the reliability sensitivities as a by-product. The simulation technique is coupled with an efficient feasible-direction interior-point algorithm, in which every iteration provides a better design that is also feasible. The capabilities of the proposed approach are demonstrated by means of an illustrative example.

Keywords: reliability-based optimization, first-passage probabilities, gradient-based algorithms.

©2021 *by D. Jerez, H. Jensen, M. Valdebenito, M. Misraji, M. Beer.*

1. Introduction

Structural systems are usually devised to be optimum in terms of a certain criterion while satisfying given performance conditions (Haftka and Gürdal, 1992). In many applications, stochastic excitation models are considered and the corresponding requirements on the structural performance are expressed in terms of reliability measures (Suksuwan and Spence, 2018). In such cases, a reliability-based optimization (RBO) problem must be solved to obtain the optimum structural design (Jensen, 2005). These systems are associated with a high number of uncertain parameters (in the order of hundreds or thousands) and, therefore, simulation-based approaches (Schuëller and Pradlwarter, 2007) must be implemented to assess their reliability. As a result, these problems are highly involved from the numerical viewpoint.

Several approaches have been proposed in the literature to deal with the RBO of structural systems under stochastic excitation. Some of them include gradient-based schemes (Jensen et al., 2013), formulations in the so-called augmented reliability space (Ching and Hsieh, 2007) and stochastic optimization approaches (Taflanidis and Beck, 2008). Although the previous contributions have shown different levels of effectiveness, it is believed that there is still room for further developments in this area.

Attention is directed towards RBO problems involving deterministic linear structural systems under Gaussian excitation (Au and Beck, 2001). In this context, the present contribution proposes a framework that integrates a feasible-direction interior-point algorithm (Herskovits et al., 2011, Jensen et al., 2013) with Directional Importance Sampling (DIS) for reliability and reliability sensitivity evaluation (Misraji et al., 2020, Misraji et al., 2019). An illustrative example is considered to demonstrate the capabilities of the proposed approach.

The document is organized as follows. The class of problems of interest is formulated in Section 2. Section 3 discusses the optimization strategy. The main aspects of Directional Importance Sampling for reliability and reliability sensitivity estimation are discussed in Sections 4 and 5, respectively. An illustrative example is studied in Section 6. The paper closes with some final remarks.

2. Problem formulation

Consider the inequality-constrained nonlinear optimization problem

$$\begin{array}{ll} \min_{\mathbf{x}} & c(\mathbf{x}) \\ \text{s.t.} & g_j(\mathbf{x}) \leq 0,\ j = 1, \ldots, n_g \\ & r_j(\mathbf{x}) \leq 0,\ j = 1, \ldots, n_r \\ & x_i^l \leq x_i \leq x_i^u,\ i = 1, \ldots, n_x \end{array} \tag{1}$$

where $\mathbf{x} = \langle x_1, \ldots, x_{n_x} \rangle^T$ is the vector of n_x design or control variables associated with the structural system, $c(\mathbf{x})$ is a general cost function, $g_j(\mathbf{x}) \leq 0, j = 1, \ldots, n_g$ represent the n_g standard constraints, $r_j(\mathbf{x}) \leq 0, j = 1, \ldots, n_r$ correspond to the n_r constraints on the system reliability, and $x_i^l \leq x_i \leq x_i^u, i = 1, \ldots, n_x$ are the side constraints on the design variables. It is assumed that all the functions involved in (1) vary smoothly in terms of the design variables and that the evaluation of the reliability constraints involves considerable numerical efforts.

2.1. Reliability constraints

The reliability constraint functions are expressed as $r_j(\mathbf{x}) = P^*_{F_j} - P_{F_j}(\mathbf{x}), j = 1, \ldots, n_r$, where $P_{F_j}(\mathbf{x})$ is the failure probability function with maximum allowable value $P^*_{F_j}$. First-passage failure probabilities within a reference period T are considered. Thus, the failure events are given by

$$F_j(\mathbf{x}) = \{\boldsymbol{\theta} \in \Omega_\theta : d_j(\mathbf{x}, \boldsymbol{\theta}) \geq 1\}, \quad j = 1, \ldots, n_r \tag{2}$$

where $\boldsymbol{\theta} \in \Omega_\theta \subset \mathbb{R}^{n_\theta}$ is the vector of random variables and the normalized demand function is

$$d_j(\mathbf{x}, \boldsymbol{\theta}) = \max_{\ell=1,\ldots,n_j} \max_{t\in[0,T]} \frac{|h_{j,\ell}(t; \mathbf{x}, \boldsymbol{\theta})|}{h^*_{j,\ell}} \tag{3}$$

where $h_{j,\ell}(t; \mathbf{x}, \boldsymbol{\theta}), \ell = 1, \ldots, n_j$ are the response functions with maximum allowable thresholds $h^*_{j,\ell} > 0$. Thus, the normalized demand function d_j quantifies the maximum demand-to-capacity ratio observed during the reference period T across all the responses of interest. The corresponding failure probability function is given by $P_{F_j}(\mathbf{x}) = \int_{d_j(\mathbf{x},\boldsymbol{\theta})\geq 1} p(\boldsymbol{\theta}) d\boldsymbol{\theta}$, where $p(\boldsymbol{\theta})$ is the multivariate probability density function of the basic random variables. Since $\boldsymbol{\theta}$ is high-dimensional (in the order of hundreds or thousands), the previous integral is very challenging to evaluate. To this end, DIS is adopted here (Misraji et al., 2020), whose most salient features are given in Section 4.

3. Optimization strategy

A first-order feasible-direction interior-point algorithm is considered in the present implementation (Herskovits et al., 2011). Starting from a feasible design, each iteration identifies a feasible design with lower objective function value. The algorithm is based on the Karush-Kuhn-Tucker (KKT) first-order optimality conditions corresponding to the optimization, which can be stated as (Herskovits et al., 2011)

$$\begin{aligned}
&\nabla c(\mathbf{x}) + \boldsymbol{\nabla}\mathbf{g}(\mathbf{x})\boldsymbol{\lambda}_g + \boldsymbol{\nabla}\mathbf{r}(\mathbf{x})\boldsymbol{\lambda}_r = \mathbf{0} \\
&\mathbf{G}(\mathbf{x})\boldsymbol{\lambda}_g = \mathbf{0}, \qquad \mathbf{R}(\mathbf{x})\boldsymbol{\lambda}_r = \mathbf{0} \\
&g_j(\mathbf{x}) \leq 0, \quad j = 1, \ldots, n_g \\
&r_j(\mathbf{x}) \leq 0, \quad j = 1, \ldots, n_r \\
&\boldsymbol{\lambda}_g \geq \mathbf{0}, \quad \boldsymbol{\lambda}_r \geq \mathbf{0}
\end{aligned} \tag{4}$$

where $\boldsymbol{\lambda}_g \in \mathbb{R}^{n_g}$ and $\boldsymbol{\lambda}_r \in \mathbb{R}^{n_r}$ are the vectors of dual variables, $\boldsymbol{\nabla}\mathbf{g}(\mathbf{x}) \in \mathbb{R}^{n_x \times n_g}$ and $\boldsymbol{\nabla}\mathbf{r}(\mathbf{x}) \in \mathbb{R}^{n_x \times n_r}$ are the matrices of the derivatives of the standard and reliability constraint functions, respectively, given by $\boldsymbol{\nabla}\mathbf{g}(\mathbf{x}) = [\nabla g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_{n_g}(\mathbf{x})]$ and $\boldsymbol{\nabla}\mathbf{r}(\mathbf{x}) = [\nabla r_1(\mathbf{x}), r_2(\mathbf{x}), \ldots, r_{n_r}(\mathbf{x})]$, and $\mathbf{G}(\mathbf{x}) \in \mathbb{R}^{n_g \times n_g}$ and $\mathbf{R}(\mathbf{x}) \in \mathbb{R}^{n_r \times n_r}$ are diagonal matrices such that $G_{jj}(\mathbf{x}) = g_j(\mathbf{x}), j = 1, \ldots, n_g$ and $R_{jj}(\mathbf{x}) = r_j(\mathbf{x}), j = 1, \ldots, n_r$.

At each iteration, three main tasks are carried out. First, a feasible-descent direction is computed (Section 3.1). Then, the step length along the feasible-descent direction is determined based on an inexact line search procedure (Section 3.2). Finally, some auxiliary variables are updated (Section 3.3). Further implementation details can be found in (Herskovits et al., 2011, Jensen et al., 2013).

3.1. Feasible-descent direction

At the beginning of the k-th optimization cycle, $k = 0, 1, \ldots$, a feasible-descent direction is generated as $\mathbf{d}^k = \mathbf{d}_1^k + \rho^k \mathbf{d}_2^k$, where $\mathbf{d}_1^k$ is a descent direction of the objective function $c(\mathbf{x})$, $\mathbf{d}_2^k$ is a vector pointing towards the interior of the feasible domain and $\rho^k > 0$. The vectors $\mathbf{d}_1^k$ and $\mathbf{d}_2^k$ are obtained by solving

$$\begin{bmatrix} \mathbf{B}^k & \boldsymbol{\nabla}\mathbf{g}(\mathbf{x}^k) & \boldsymbol{\nabla}\mathbf{r}(\mathbf{x}^k) \\ \boldsymbol{\Lambda}_g^k \boldsymbol{\nabla}\mathbf{g}(\mathbf{x}^k)^T & \mathbf{G}(\mathbf{x}^k) & \mathbf{0} \\ \boldsymbol{\Lambda}_r^k \boldsymbol{\nabla}\mathbf{r}(\mathbf{x}^k)^T & \mathbf{0} & \mathbf{R}(\mathbf{x}^k) \end{bmatrix} \begin{Bmatrix} \mathbf{d}_1^k \\ \bar{\boldsymbol{\lambda}}_{g1}^{k+1} \\ \bar{\boldsymbol{\lambda}}_{r1}^{k+1} \end{Bmatrix} = - \begin{Bmatrix} \nabla c(\mathbf{x}^k) \\ \mathbf{0} \\ \mathbf{0} \end{Bmatrix} \tag{5}$$

$$\begin{bmatrix} \mathbf{B}^k & \boldsymbol{\nabla}\mathbf{g}(\mathbf{x}^k) & \boldsymbol{\nabla}\mathbf{r}(\mathbf{x}^k) \\ \boldsymbol{\Lambda}_g^k \boldsymbol{\nabla}\mathbf{g}(\mathbf{x}^k)^T & \mathbf{G}(\mathbf{x}^k) & \mathbf{0} \\ \boldsymbol{\Lambda}_r^k \boldsymbol{\nabla}\mathbf{r}(\mathbf{x}^k)^T & \mathbf{0} & \mathbf{R}(\mathbf{x}^k) \end{bmatrix} \begin{Bmatrix} \mathbf{d}_2^k \\ \bar{\boldsymbol{\lambda}}_{g2}^{k+1} \\ \bar{\boldsymbol{\lambda}}_{r2}^{k+1} \end{Bmatrix} = - \begin{Bmatrix} \mathbf{0} \\ \boldsymbol{\lambda}_g^k \\ \boldsymbol{\lambda}_r^k \end{Bmatrix} \tag{6}$$

where $\left(\mathbf{x}^k, \boldsymbol{\lambda}_g^k, \boldsymbol{\lambda}_r^k\right)$ is the starting point at the k-th iteration, $\boldsymbol{\Lambda}_g^k$ and $\boldsymbol{\Lambda}_r^k$ are diagonal matrices with coefficients $\Lambda_{gii}^k = \lambda_{gi}^k, i = 1, \ldots, n_g$ and $\Lambda_{rii}^k = \lambda_{ri}^k, i = 1, \ldots, n_r$, and $\mathbf{B}^k$ is a symmetric and positive definite matrix. An upper bound on ρ^k is established in order to ensure that $\mathbf{d}^k$ is a descent direction (Herskovits et al., 2011, Jensen et al., 2013), which is given by $\rho_{\text{limit}}^k = (\alpha - 1)\mathbf{d}_1^{kT} \nabla c(\mathbf{x}^k) / (\mathbf{d}_2^{kT} \nabla c(\mathbf{x}^k))$, with $\alpha \in (0, 1)$.

It is noted that the coefficient matrix of equations (5) and (6) involves the first-order derivatives of the reliability constraints. These quantities can be obtained as a by-product of DIS (Misraji et al., 2019). The main ideas of this approach are discussed in Section 5.

3.2. Line search

Once the search direction $\mathbf{d}^k$ is determined, the new candidate design is computed as $\mathbf{x}^{k+1} = \mathbf{x}^k + \tau_k \mathbf{d}^k$ where $\tau_k > 0$ is the step length along the feasible-descent direction. This quantity is obtained by means of a constrained line search procedure (Nocedal and Wright, 2006).

For each candidate step length τ_{cand}, the Armijo's criterion (Armijo, 1966) is first checked. This provides an upper bound for the step length. In addition, τ_{cand} must verify the Wolfe's criterion (Wolfe, 1969), which indicates a lower bound for the step length.

If Armijo's and Wolfe's criteria are simultaneously verified, then $\tau_k = \tau_{\text{cand}}$. Otherwise, a new candidate step length is obtained by solving the optimization problem

$$\begin{aligned} \max_{\tau} \quad & \tau \\ \text{s.t.} \quad & g_j(\mathbf{x}^k + \tau \mathbf{d}^k) \leq 0, \quad j = 1, \ldots, n_g \\ & \tilde{r}_j(\tau) \leq 0, \quad j = 1, \ldots, n_r \end{aligned} \tag{7}$$

where τ^L and τ^R are lower and upper limits for the step length, respectively, and $\tilde{r}_j(\tau)$ is an approximation of r_j along $\mathbf{d}^k$. The reader is referred to (Jensen et al., 2013) for further details.

3.3. Auxiliary variables updating

Once a new candidate design $\mathbf{x}^{k+1}$ has been determined, $\boldsymbol{\lambda}_g^{k+1}$ and $\boldsymbol{\lambda}_r^{k+1}$ are updated considering the scheme proposed in (Herskovits et al., 2011), while the positive definite matrix $\mathbf{B}^{k+1}$ is computed using a Broyden–Fletcher–Goldfarb–Shanno-type of formula (Nocedal and Wright, 2006).

4. Reliability estimation

The type of stochastic loading considered in this contribution corresponds to a Gaussian process. Thus, this loading is actually a linear function with respect to $\boldsymbol{\theta}$, which follows a standard normal distribution in n_θ dimensions. Furthermore, the type of structural systems considered in this contribution exhibits a linear behavior. Under these assumptions, it is straightforward to show that the response $h_{j,\ell}(t;\mathbf{x},\boldsymbol{\theta})$ is linear with respect to $\boldsymbol{\theta}$ at each time instant of analysis and for a given value of the design variables $\mathbf{x}$ (Au and Beck, 2001). Hence, the failure domain is bounded by a series of hyperplanes, as represented schematically in Figure 1.

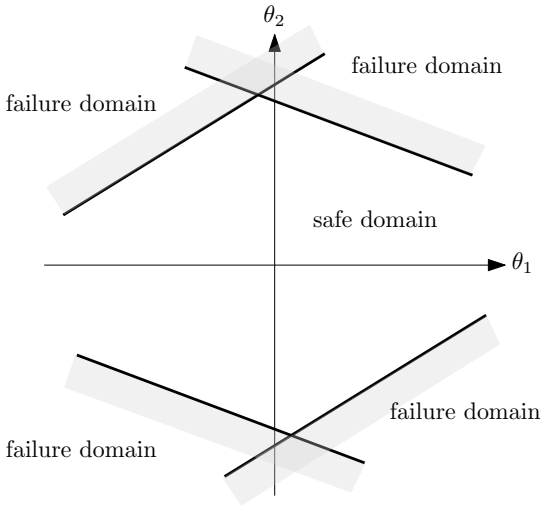


Figure 1. Schematic representation of failure domain

Taking into account the geometry of the failure domain, the probability integral can be formulated within the framework of Directional Importance Sampling, that is:

$$P_{F_j}(\mathbf{x}) = \int_{\boldsymbol{\zeta}\in\Omega_{\mathbf{Z}}} \int_{\psi_j^*(\mathbf{x},\boldsymbol{\zeta})}^{\infty} p_\Psi(\psi) \frac{p_{\mathbf{Z}}(\boldsymbol{\zeta})}{p_{\mathbf{Z}}^{\mathrm{IS}}(\boldsymbol{\zeta})} p_{\mathbf{Z}}^{\mathrm{IS}}(\boldsymbol{\zeta}) d\psi d\boldsymbol{\zeta} \tag{8}$$

where $\boldsymbol{\zeta}$ is a unit vector that points in the direction of $\boldsymbol{\theta}$, with associated probability distribution $p_{\mathbf{Z}}(u)$; $\Omega_{\mathbf{Z}}$ is the set of all points belonging to the unit hypersphere; ψ is the Euclidean norm of $\boldsymbol{\theta}$, with associated probability distribution $p_\Psi(\psi)$; $\psi_j^*(\mathbf{x},\boldsymbol{\zeta})$ is the value of ψ that fulfills the equation $d_j(\mathbf{x},\psi\boldsymbol{\zeta}) = 1$; and where $p_{\mathbf{Z}}^{\mathrm{IS}}(\boldsymbol{\zeta})$ is the importance sampling density function associated with the

direction vector. The importance sampling density function is equal to a weighted summation of the probability density function associated with $\boldsymbol{\zeta}$ conditioned on the occurrence of the failure event at each discrete time instant. It can be shown that explicit expressions associated with $p_{\mathbf{Z}}^{\mathrm{IS}}(\boldsymbol{\zeta})$ can be deduced by means of Bayes' theorem, as discussed in (Misraji et al., 2020). Thus, it is possible to evaluate the probability in equation (8) by means of simulation. Numerical experience indicates that a few hundreds of samples allow producing highly accurate estimates of the sought probability.

5. Reliability sensitivity estimation

The structure of equation (8) reveals that the failure probability P_{F_j} depends on the value of the design variable vector $\mathbf{x}$. This is clear from a physical viewpoint, as changes to this vector affect the response of the structure, which in turn affect the probability. Undoubtedly, assessing the rate of change of this probability with respect to changes in this design vector provides valuable information for decision making. Formally, such rate of change can be expressed in terms of the gradient, that is:

$$\frac{\partial P_{F_j}(\mathbf{x})}{\partial x_i} = \frac{\partial}{\partial x_i}\left(\int_{\boldsymbol{\zeta}\in\Omega_{\mathbf{Z}}}\int_{\psi_j^*(\mathbf{x},\boldsymbol{\zeta})}^{\infty} p_{\Psi}(\psi)\frac{p_{\mathbf{Z}}(\boldsymbol{\zeta})}{p_{\mathbf{Z}}^{\mathrm{IS}}(\boldsymbol{\zeta})}p_{\mathbf{Z}}^{\mathrm{IS}}(\boldsymbol{\zeta})d\psi d\boldsymbol{\zeta}\right),\ i=1,\ldots,n_x \tag{9}$$

The above integral can be simplified resorting to the Leibniz rule for differentiation, leading to an integral over the hypersurface that separates the safe and failure domains and that comprises the derivative of the normalized demand function $d_j(\mathbf{x},\psi\boldsymbol{\zeta})$ with respect to x_i. The integral over the hypersurface is solved straightforwardly within the framework of Directional Importance Sampling, as this simulation technique tracks (by definition) the aforementioned hypersurface. Moreover, the derivative of the normalized demand function can be easily calculated by considering the sensitivity of the spectral properties (natural frequencies and mode shapes of the structure), as discussed in (Lee and Jung, 1997).

In summary, the sensitivity of the failure probability as cast in equation (9) is calculated with all the information produced when calculating the failure probability itself plus the sensitivity of the response with respect to the design variables. In other words, the sensitivity of the probability becomes a by-product of the reliability analysis.

6. Illustrative example

6.1. Problem description

The illustrative example considered in this work addresses the optimum design of a two-degree-of-freedom linear system subject to Gaussian excitation. The system is shown in Figure 2. The mass of each floor is $m = 30\times10^3$ kg, k_1 and k_2 represent the stiffnesses of the first and second story, respectively, and $u_i, i = 1,2$ are the ground-relative displacements at the different floors. In addition, a 4% of the critical damping is considered at the modal level.

The structural system is subject to stochastic ground acceleration modeled as a modulated white noise process that passes through a Clough-Penzien type of filter. Specifically, $\ddot{u}_g(t) = \Omega_1^2 w_1(t) +$

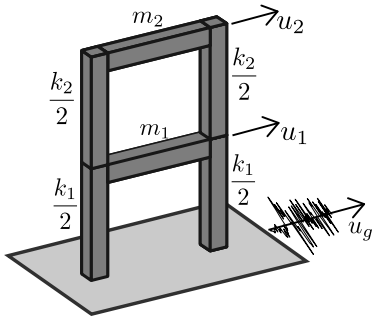


Figure 2. Two-degree-of-freedom system

$2\xi_1\Omega_1\dot{w}_1(t) - \Omega_2^2 w_2(t) - 2\xi_2\Omega_2\dot{w}_2(t)$, where $\Omega_1 = 15.6$ rad/s, $\xi_1 = 0.6$ are the parameters of the first filter, and $\Omega_2 = 1.0$ rad/s, $\xi_2 = 0.9$ are the parameters of the second filter. The white noise process has an intensity $S = 10^{-4}$ m^2/s^3, a duration $T = 15$ s and a time step $\Delta t = 0.01$ s. The time envelope function $h(t)$ is given by

$$h(t) = \begin{cases} (t/5)^2 & 0 \leq t \leq 5\text{s} \\ 1 & 5 < t \leq 10\text{s} \\ e^{-(t-10)^2} & t > 10\text{s} \end{cases} \tag{10}$$

The objective of this example is to minimize the structural cost with respect to the interstory stiffnesses under a reliability constraint associated with the top displacement, that is,

$$\begin{array}{ll} \max\limits_{\mathbf{x}=\langle x_1,x_2\rangle^T} & c(\mathbf{x}) \\ \text{s.t.} & P_F(\mathbf{x}) \leq P_F^* \\ & 0.5 \leq x_i \leq 1.5, \quad i = 1,2 \end{array} \tag{11}$$

where $c(\mathbf{x})$ is the cost function and $P_F(\mathbf{x})$ is the failure probability function with maximum allowable value $P_F^* = 10^{-3}$. The design variables are defined as the interstory stiffnesses normalized with respect to a reference value. Specifically, $x_i = k_i/\bar{k}, i = 1,2$, with $\bar{k} = 18 \times 10^6$ N/m. For illustration purposes, the cost function is assumed to be proportional to the normalized interstory stiffnesses. In particular, $c(\mathbf{x}) = x_1 + x_2$. Finally, the failure event is defined as $F = \{\boldsymbol{\theta} \in \Omega_\theta : d(\mathbf{x}, \boldsymbol{\theta}) > 1\}$, where $d(\mathbf{x}, \boldsymbol{\theta}) = \max_{t\in[0,T]} |u_2(t; \mathbf{x}, \boldsymbol{\theta})| / u_2^*$ and $u_2^* = 0.006$ m.

6.2. Results and discussion

In order to gain insight into the system under analysis, Figure 3 (left) shows the contours of the failure probability function in the design space. These curves have been smoothed for presentation purposes. The structure becomes more reliable for higher stiffness values and some interaction between the design variables is observed. This behavior is reasonable from a physical viewpoint, since the response of interest corresponds to the top displacement. In addition, a sketch of the feasible design space and some contours of the objective function are depicted in Figure 3 (right).

The optimization process is carried out considering the initial solution $\mathbf{x}^0 = \langle 1.4, 1.4\rangle^T$ and a total of 1000 samples for the implementation of DIS. The trajectory of the candidate solutions in the

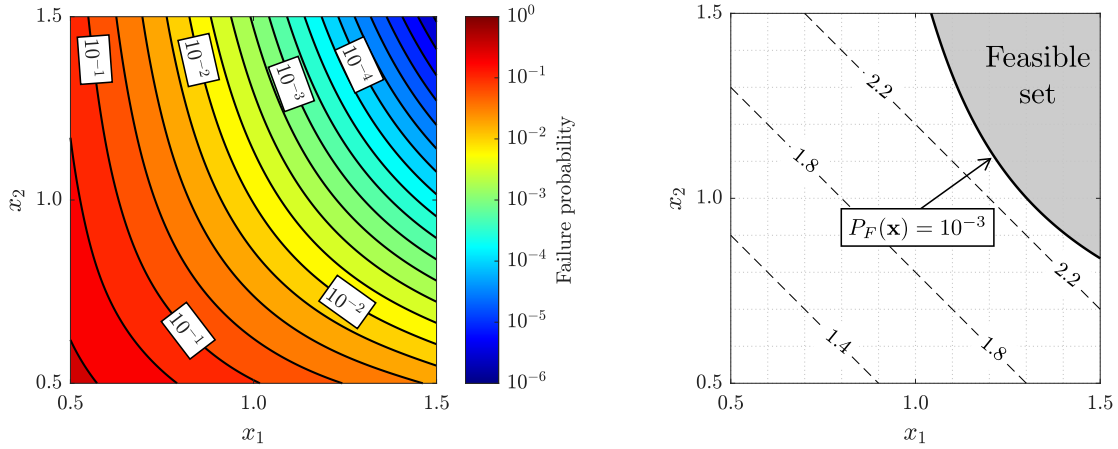


Figure 3. Left: Iso-probability curves. Right: Sketch of the feasible design space

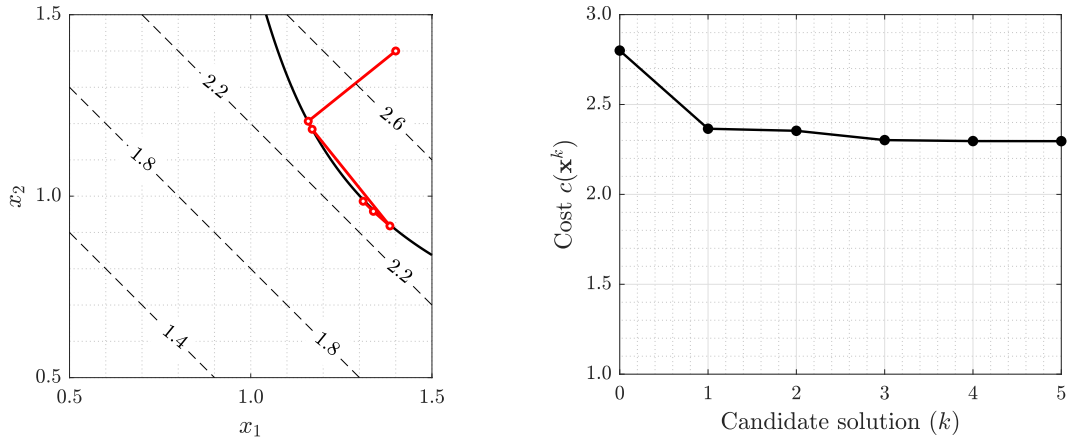


Figure 4. Candidate designs (left) and optimum costs (right) obtained during the optimization process

design space is shown in Figure 4 (left), where some contours of the cost function have been included with dashed lines. The initial search direction is orthogonal to the cost contours. Afterwards, the designs move nearly parallel to the reliability constraint. The corresponding objective function values are shown in Figure 4 (right). Most of the improvements of the objective function take place during the first steps of the optimization scheme. The final design is obtained after five iterations and is given by $\mathbf{x}^* = \langle 1.31, 0.98 \rangle^T$, with $P_F(\mathbf{x}^*) = 10^{-3}$ and $c(\mathbf{x}^*) = 2.29$.

7. Conclusions

This contribution has proposed an approach for reliability based design optimization of linear structures subject to dynamic Gaussian load. The proposed framework employs a first order optimization scheme which ensures that each candidate solution is feasible and better than the previous ones.

This scheme is integrated with Directional Importance Sampling, which is a simulation method which allows estimating first excursion probabilities and their gradients most efficiently.

The results presented in this contribution are encouraging, as it is possible to determine an optimal design solution with reduced numerical efforts. Nonetheless, the scope of application of the proposed approach remains to be verified in problems involving large scale structures. This issue is currently under investigation by the authors.

Acknowledgements

This research was supported in part by ANID (National Agency for Research and Development, Chile) under its program FONDECYT, grant numbers 1200087 and 1180271. Also, this research has been partially supported by ANID and DAAD (German Academic Exchange Service, Germany) under CONICYT-PFCHA/Doctorado Acuerdo Bilateral DAAD Becas Chile/2018-62180007.

References

Armijo, L.: 1966, 'Minimization of Functions having Lipschitz Continuous First Partial Derivatives'. *Pacific Journal of Mathematics* **16**(1), 1–3.

Au, S. and J. Beck: 2001, 'First excursion probabilities for linear systems by very efficient importance sampling'. *Probabilistic Engineering Mechanics* **16**(3), 193–207.

Ching, J. and Y. Hsieh: 2007, 'Approximate Reliability-Based Optimization Using a Three-Step Approach Based on Subset Simulation'. *Journal of Engineering Mechanics* **133**(4), 481–493.

Haftka, R. and Z. Gürdal: 1992, *Elements of Structural Optimization*. Dordrecht, The Netherlands: Kluwer, 3^{rd} edition.

Herskovits, J., W. Freire, M. Fo, and A. Canelas: 2011, 'A feasible directions method for nonsmooth convex optimization'. *Structural and Multidisciplinary Optimization* **44**(3), 363–377.

Jensen, H.: 2005, 'Structural optimization of linear dynamical systems under stochastic excitation: a moving reliability database approach'. *Computers Methods in Applied Mechanics and Engineering* **194**(12-16), 1757–1778.

Jensen, H., L. Becerra, and M. Valdebenito: 2013, 'On the Implementation of an Advanced Interior Point Algorithm for Stochastic Structural Optimization'. *Computers & Structures* **126**, 69–85.

Lee, I.-W. and G.-H. Jung: 1997, 'An efficient algebraic method for the computation of natural frequency and mode shape sensitivities - Part I. Distinct natural frequencies'. *Computers & Structures* **62**(3), 429–435.

Misraji, M., M. Valdebenito, H. Jensen, and C. Mayorga: 2020, 'Application of directional importance sampling for estimation of first excursion probabilities of linear structural systems subject to stochastic Gaussian loading'. *Mechanical Systems and Signal Processing* **139**, 106621.

Misraji, M., M. Valdebenito, C. Mayorga, and H. Jensen: 2019, 'Reliability Sensitivity Analysis for Linear Structures Subject to Dynamic Gaussian Excitation'. In: *29th European Safety and Reliability Conference (Esrel 2019)*. Hannover, Germany.

Nocedal, J. and S. Wright: 2006, *Numerical Optimization*. New York: Springer New York, 2nd edition.

Schuëller, G. and H. Pradlwarter: 2007, 'Benchmark Study on Reliability Estimation in Higher Dimensions of Structural Systems – An Overview'. *Structural Safety* **29**, 167–182.

Suksuwan, A. and S. Spence: 2018, 'Optimization of uncertain structures subject to stochastic wind loads under system-level first excursion constraints: A data-driven approach'. *Computers & Structures* **210**, 58 – 68.

Taflanidis, A. and J. Beck: 2008, 'An efficient framework for optimal robust stochastic system design using stochastic simulation'. *Computer Methods in Applied Mechanics and Engineering* **198**(1), 88–101.

Wolfe, P.: 1969, 'Convergence Conditions for Ascent Methods'. *SIAM Review* **11**(2), 226–235.

A Fully Decoupled Approach for a Class of Reliability-based Optimization Problems in Stochastic Linear Dynamics

Marcos A. Valdebenito
Faculty of Engineering and Sciences, Adolfo Ibanez University, 2562340 Viña del Mar, Chile, marcos.valdebenito@uai.cl.

Matthias G.R. Faes
Department of Mechanical Engineering, Technology campus De Nayer, KU Leuven, 2860 St.-Katelijne-Waver, Belgium, matthias.faes@kuleuven.be

Abstract. Reliability-based optimization (RBO) allows determining optimal structural designs while explicitly taking into account the effects of uncertainty on structural performance. Its practical implementation is quite challenging from a numerical viewpoint, as it demands solving simultaneously a reliability problem nested in an optimization procedure, that is, a double-loop problem. This contribution develops a most efficient approach for RBO that avoids the aforementioned double-loop implementation. This approach is applicable for a specific class of problems: the minimization of the failure probability of linear structural systems subject to Gaussian stochastic loading. The proposed approach is formulated within the framework of the operator norm theorem. In this way, the RBO problem is reduced to the solution of a single deterministic optimization problem followed by a single reliability analysis, avoiding a double-loop (or nested) implementation. The application and capabilities of the proposed approach are illustrated by means of an example, indicating that the involved numerical costs can be reduced drastically.

Keywords: Reliability-based optimization, Linear dynamics, First excursion probability, Gaussian load, Operator norm.

1. Introduction

Reliability-based optimization (RBO) is a methodology that can assist decision-making under uncertainty in structural mechanics. In fact, it provides the means for determining the best design solution with respect to a certain criterion while explicitly taking into account the effects of uncertainty (Schuëller and Jensen, 2008). The advantages of RBO over traditional deterministic approaches are evident. Nonetheless, its practical implementation is extremely challenging from a numerical viewpoint, as it demands assessing different design solutions and for each of them, it is necessary to perform uncertainty quantification. In other words, RBO corresponds to a so-called *double-loop problem*, where reliability analysis is embedded within an optimization algorithm. Therefore, different strategies for solving RBO problems have been developed to decrease numerical costs associated with its practical implementation. These strategies involve, for example, building surrogate models, application of stochastic search techniques, etc. For a detailed

© 2021 *by Marcos A. Valdebenito and Matthias G.R. Faes.*

overview about these strategies, it is referred existing overview papers in the literature, such as e.g. (Aoues and Chateauneuf, 2010, Enevoldsen and Sørensen, 1994, Schuëller and Jensen, 2008, Valdebenito and Schuëller, 2010).
This paper develops a framework for solving RBO problems for a specific class of applications: minimization of the first excursion probability of linear structures subject to dynamic loading of the Gaussian type. The proposed framework fully exploits the particular characteristics of the class of applications considered and is formulated in a *fully decoupled* fashion. This means that the RBO problem is solved by conducting a single deterministic optimization procedure followed by a single reliability estimation. The theoretical basis for formulating this fully decoupled framework lies on the operator norm theorem (see, e.g. (Tropp, 2004)), which has already been explored by the authors within the context of imprecise reliability analysis (Faes et al., 2020, Faes et al., 2021).

2. Formulation of the Problem

2.1. Stochastic Loading

Consider a zero-mean dynamic Gaussian loading f acting over a structural system. This Gaussian process is represented at discrete time instants $t_k = (k-1)\Delta t,\ k = 1, \ldots, n_T$, where Δt is the time step; n_T is the total number of time instants considered and $T = (n_T - 1)\Delta t$ is the load duration. Applying the Karhunen-Loève expansion (see, e.g. (Schenk et al., 2005, Stefanou, 2009)), the Gaussian loading is described as:

$$\boldsymbol{f}(\boldsymbol{z}) = \boldsymbol{\Psi}\boldsymbol{\Lambda}^{1/2}\boldsymbol{z} \tag{1}$$

where $\boldsymbol{f}$ is a realization of the stochastic loading, which is a $n_T \times 1$ vector, whose k-th component f_k represents the loading at time t_k; $\boldsymbol{\Psi}$ is a matrix of dimension $n_T \times n_{KL}$ containing the first n_{KL} eigenvectors of the covariance matrix $\boldsymbol{\Gamma}$ associated with the Gaussian process; $\boldsymbol{\Lambda}$ is a matrix whose diagonal contains the first n_{KL} eigenvalues of the covariance matrix $\boldsymbol{\Gamma}$; n_{KL} is the number of terms retained for the Karhunen-Loève expansion ($n_{KL} \leq n_T$, see, e.g. (Stefanou, 2009)); and $\boldsymbol{z}$ is a realization of a standard Gaussian random variable vector $\boldsymbol{Z}$ of dimension $n_{KL} \times 1$ and whose probability density function is denoted as $p_{\boldsymbol{Z}}(\boldsymbol{z})$.

2.2. Stochastic Response

The stochastic Gaussian loading described above acts over a linear elastic structure with classical damping. Some of the parameters of this structure (such as elements' cross sections) can be chosen by the designer and are termed as *design variables*. These design variables are collected in a vector $\boldsymbol{y}$ of dimension n_y and can alter the structural behavior of the system (for example, stiffer structural members may help in reducing displacements).
For practical design purposes, it is of interest monitoring n_R responses of interest of the structure. In view of linearity, these responses are calculated by means of a convolution integral. Taking into account the discrete time representation of the dynamic stochastic loading, the different responses of interest at different time instants can be calculated in discrete form as:

$$\boldsymbol{\eta}_i(\boldsymbol{y}, \boldsymbol{z}) = \boldsymbol{A}_i(\boldsymbol{y})\boldsymbol{z},\ i = 1, \ldots, n_R \tag{2}$$

where $\boldsymbol{\eta}_i$ is a vector of dimension $n_T \times 1$ containing the discrete-time representation of the i-th response along the duration T of the stochastic loading; and $\boldsymbol{A}_i$ is a matrix of dimension $n_T \times n_{KL}$ whose k-th row contains the discrete time representation of the convolution integral (Faes et al., 2020).
The responses of interest should remain below acceptable threshold levels $b_i,\ i = 1, \ldots, n_R$ within the duration T of the stochastic load in order to avoid an undesirable behavior. For verifying such condition, it is useful to define the so-called normalized response function $r(\boldsymbol{y}, \boldsymbol{z})$ (Au and Beck, 2001), which is defined as:

$$r(\boldsymbol{y}, \boldsymbol{z}) = ||\boldsymbol{A}(\boldsymbol{y})\boldsymbol{z}||_\infty \tag{3}$$

where $||\cdot||_\infty$ denotes infinity norm; and $\boldsymbol{A}$ is a matrix that collects all matrices $\boldsymbol{A}_i$ associated with the calculation of the i-th response (see eq. (2)), normalized by their respective threshold levels. Matrix $\boldsymbol{A}$ is defined as:

$$\boldsymbol{A}(\boldsymbol{y}) = \begin{bmatrix} b_1^{-1}\boldsymbol{A}_1(\boldsymbol{y}) \\ \vdots \\ b_{n_R}^{-1}\boldsymbol{A}_{n_R}(\boldsymbol{y}) \end{bmatrix} \tag{4}$$

and its dimension is $(n_R n_T) \times n_{KL}$.

2.3. First Excursion Probability

The chance that any of the responses of interest exceed their prescribed threshold within the duration of the stochastic loading is quantified by means of the following classical probability integral:

$$p_F(\boldsymbol{y}) = \int_{\boldsymbol{z} \in \mathcal{R}^{n_{KL}}} I_F(\boldsymbol{y}, \boldsymbol{z}) p_{\boldsymbol{Z}}(\boldsymbol{z}) d\boldsymbol{z} \tag{5}$$

where p_F denotes the failure probability; and $I_F(\cdot, \cdot)$ is the indicator function, which is equal to one in case $r(\boldsymbol{y}, \boldsymbol{z}) \geq 1$ and zero, otherwise. As the dimensionality of the failure probability integral is usually quite large (in the order of hundreds or thousands) and the indicator function is known point-wise for specific values of the design variable vector $\boldsymbol{y}$ and the stochastic excitation $\boldsymbol{z}$, eq. (5) must be usually evaluated by means of simulation, as discussed in (Schuëller and Pradlwarter, 2007). However, this is challenging from a numerical viewpoint, as it demands performing repeated dynamic analyses.

2.4. Reliability-based optimization

Recall that design variables $\boldsymbol{y}$ may be chosen by the designer in order to alter the dynamic performance of the structural system. A possible means for selecting these design variables is formulating a Reliability-based Optimization (RBO) problem. For example, one could carry minimization of failure probability under a constraint related with available resources (see, e.g. (Jensen et al., 2020)),

which leads to the following optimization problem.

$$\begin{array}{l} \min_{\boldsymbol{y}} p_F(\boldsymbol{y}) \\ \text{subject to} \\ c_l(\boldsymbol{y}) \leq 0,\ l = 1, \ldots, n_C \\ y_j^L \leq y_j \leq y_j^U,\ j = 1, \ldots, n_y \end{array} \tag{6}$$

where $c_l(\cdot)$ represents a deterministic constraint; y_j is the j-th element of $\boldsymbol{y}$; n_y denotes the dimension of the design variable vector; and y_j^L and y_j^U are the lower and upper bounds associated with y_j.

The solution of the optimization design problem in eq. (6) presents two challenges. First, the calculation of the failure probability for a given value of $\boldsymbol{y}$ by means of simulation demands performing repeated structural analyses for different realizations of the stochastic loading vector $\boldsymbol{z}$. Second, the calculation of the failure probability is embedded within the optimization algorithm, leading to the so-called double-loop implementation. These issues may impose a huge numerical cost. As a remedy to this situation, the application of the operator norm theorem is discussed in the remaining part of this work.

3. Operator Norm Theorem and its Application for Reliability-based Optimization

A close examination of eq. (3) reveals that the system's normalized response can be explained as the effect of the stochastic loading (represented in terms of $\boldsymbol{z}$) which acts over the system (represented in terms of the matrix $\boldsymbol{A}(\boldsymbol{y})$). In other words, the stochastic loading $\boldsymbol{z}$ is *stretched* by the system's matrix $\boldsymbol{A}(\boldsymbol{y})$. A bound on the amount of stretching exerted by $\boldsymbol{A}(\boldsymbol{y})$ is given by the operator norm theorem, that is (Tropp, 2004):

$$||\boldsymbol{A}(\boldsymbol{y})||_{\infty,2} = \inf\{c(\boldsymbol{y}) \geq 0 : ||\boldsymbol{A}(\boldsymbol{y})\boldsymbol{z}||_{\infty} \leq c(\boldsymbol{y})\, ||\boldsymbol{z}||_2\} \tag{7}$$

where the amount of stretching is calculated with respect to the Euclidean norm of $\boldsymbol{z}$, as this can be loosely interpreted as the energy content of the stochastic load (Faes et al., 2020, Faes et al., 2021). The expression above implies that the stochastic load vector can be *stretched* by matrix $\boldsymbol{A}$ by factor c at most. Such an idea is illustrated schematically in Figure 1, where it has been assumed for simplicity that $n_R = 1$ and $n_T = n_{KL} = 2$. It is observed that a given realization $\boldsymbol{z}^{(1)}$ associated with the stochastic load produces a dynamic response given by $\boldsymbol{A}(\boldsymbol{y})\boldsymbol{z}^{(1)}$. None of the components of that dynamic response may exceed the associated scalar $c\left|\left|\boldsymbol{z}^{(1)}\right|\right|$.

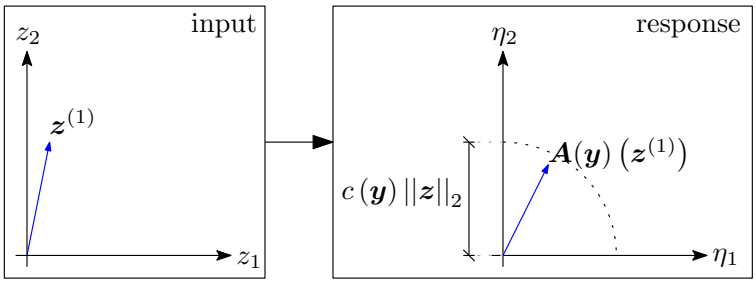


Figure 1. Schematic interpretation of operator norm theorem.

The solution to the operator norm problem in eq. (7) is equal to the row of matrix $\boldsymbol{A}$ with largest Euclidean norm, that is (Tropp, 2004):

$$||\boldsymbol{A}(\boldsymbol{y})||_{\infty,2} = \max_{i=1,\ldots,n_R,\ k=1,\ldots,n_T}\left(\sqrt{\boldsymbol{a}_{i_k}\boldsymbol{a}_{i_k}^T}\right) \tag{8}$$

where $\boldsymbol{a}_{i_k}$ represents the row of matrix $\boldsymbol{A}$ associated with the calculation of the i-th response at time instant t_k.

Note from the above discussion that the scalar c expresses the maximum amplification that the stochastic load may undergo when applied to the structure. As the normalized response function directly influences the calculation of the failure probability (see eq. (5)), this implies that minimizing the operator norm associated with $||\boldsymbol{A}(\boldsymbol{y})||_{\infty,2}$ offers a proxy which is equivalent to minimizing the failure probability. The quality of this proxy has been verified and validated, e.g. for estimating imprecise probabilities in (Faes et al., 2020, Faes et al., 2021). In view of this conclusion, the RBO problem formulated in eq. (6) can be replaced by the following problem:

$$\begin{aligned} &\min_{\boldsymbol{y}} ||\boldsymbol{A}(\boldsymbol{y})||_{\infty,2} \\ &\text{subject to} \\ &c_l(\boldsymbol{y}) \leq 0,\ l = 1,\ldots,n_C \\ &y_j^L \leq y_j \leq y_j^U,\ j = 1,\ldots,n_y \end{aligned} \tag{9}$$

This optimization problem in eq. (9) is actually deterministic, as it does not involve the stochastic load $\boldsymbol{z}$. Hence, once its optimum $\boldsymbol{y}^\star$ is found by means of any suitable optimization algorithm, it suffices to carry out a single reliability analysis for that optimum in order to determine the minimum failure probability $p_F\,(\boldsymbol{y}^\star)$ associated with the RBO problem in eq. (6).

As noted from the above description, the operator norm theorem framework offers the means for fully decoupling the solution of the RBO problem. That is, a nested problem is broken into a single deterministic optimization problem, followed by a single reliability problem. This is evidently most advantageous, as it decreases numerical costs substantially.

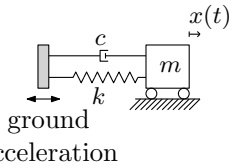


Figure 2. Single-degree-of-freedom (SDOF) oscillator subject to stochastic ground acceleration.

4. Test Example

Consider a single-degree-of-freedom (SDOF) oscillator with mass $m = 1$ [kg], nominal stiffness $k = 225$ [N/m] and classical damping $d = 5\%$, as depicted in Figure 2. The oscillator is subjected to a zero-mean stochastic Gaussian load of time duration $T = 20$ [s], discretized considering times steps of $\Delta t = 0.01$ [s]. This Gaussian load follows a modulated Clough-Penzien spectrum (see, e.g. (Zerva, 2009)) with: spectral intensity of 5×10^{-3} [kg m^2/s^3]; natural circular frequencies of 6π [rad/s] and 0.6π [rad/s] for the primary and secondary filters, respectively; damping ratios of 60%; and modulation function following the Shinozuka-Sato model with shape parameters $c_1 = 0.14$ and $c_2 = 0.16$ (Shinozuka and Sato, 1967).

The responses of interest are the maximum values of the relative displacement and absolute acceleration of the SDOF oscillator. The admissible threshold values for these two responses are 0.07 [m] and 7.5 [m/s^2]. The first excursion probability associated with these responses is calculated by means of Directional Importance Sampling (Misraji et al., 2020) considering a total of 1000 samples.

The objective of this example is determining the value of the stiffness y of the oscillator which minimizes its first excursion probability, under the constraint that y is a positive real value (that is, $y \in [0, \infty[$). Before solving the actual optimization problem, the failure probability is calculated over a discrete grid of values of the stiffness such that $y \in [70, 170]$ [N/m]. The estimates of both the failure probability and the associated operator norm for each of those discrete grid values is shown in Figure 3. It is noted that the minimum value of the failure probability is attained for a stiffness of about 105 [N/m]. The stiffness value for which the failure probability curve assumes its minimum matches with the value of the stiffness for which the minimum operator norm is attained, confirming the validity of applying the operator norm as a proxy for solving the RBO problem.

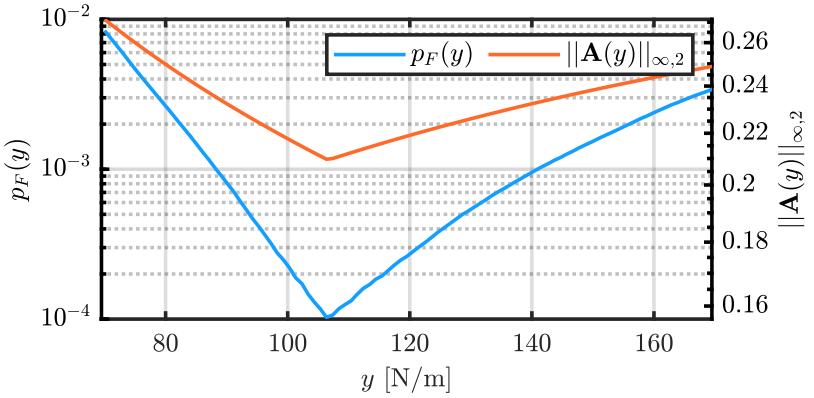


Figure 3. Failure probability and operator norm as a function of the stiffness y.

After these preliminary results, the actual RBO problem is solved using both a classical double-loop approach and the proposed, fully decoupled approach based on operator norm. The initial guess for the optimum is selected as $y^{(0)} = 90$ [N/m]. Then, the identified optimum results equal for both approaches (that is, double loop and proposed approach) and its numerical value is $y^* = 106.6$ [N/m]. However, the numerical costs are quite different. On one hand, for the double-loop implementation, the optimization solver demands a total of 52 function calls (that is, evaluations of the failure probability), each involving 1000 deterministic samples to compute the failure probability. In other words, a total of 52000 dynamic analyses are carried out. On the other hand, the solution by means of the proposed approach demanded 80 function calls (that is, calculation of the operator norm) followed by 1000 samples (dynamic analyses) for determining the reliability at the optimum. These results highlight the benefits of the proposed approach for RBO.

5. Conclusions

This paper has introduced an approach for solving a specific class of RBO problems. That is, minimization of the failure probability of linear systems subject to Gaussian stochastic loading. The approach is based on the operator norm and effectively decouples optimization from reliability assessment.

Although the results presented in this contribution are encouraging, it should be kept in mind that the proposed approach is applicable for a very specific types of problems. Therefore, the authors are working on the extension of this approach to nonlinear cases with the help of equivalent stochastic linearization.

Acknowledgements

Marcos Valdebenito acknowledges the support of ANID (National Agency for Research and Development, Chile) under its program FONDECYT, grant number 1180271. Matthias Faes gratefully

acknowledges the financial support of the Research Foundation Flanders (FWO) under grant number 12P3519N.

References

Aoues, Y. and A. Chateauneuf: 2010, 'Benchmark study of numerical methods for reliability-based design optimization'. *Structural and Multidisciplinary Optimization* **41**(2), 277–294.

Au, S. and J. Beck: 2001, 'Estimation of Small Failure Probabilities in High Dimensions by Subset Simulation'. *Probabilistic Engineering Mechanics* **16**(4), 263–277.

Enevoldsen, I. and J. Sørensen: 1994, 'Reliability-based Optimization in Structural Engineering'. *Structural Safety* **15**(3), 169–196.

Faes, M., M. Valdebenito, D. Moens, and M. Beer: 2020, 'Bounding the first excursion probability of linear structures subjected to imprecise stochastic loading'. *Computers & Structures* **239**, 106320.

Faes, M., M. Valdebenito, D. Moens, and M. Beer: 2021, 'Operator norm theory as an efficient tool to propagate hybrid uncertainties and calculate imprecise probabilities'. *Mechanical Systems and Signal Processing* **152**, 107482.

Jensen, H., D. Jerez, and M. Valdebenito: 2020, 'An adaptive scheme for reliability-based global design optimization: A Markov chain Monte Carlo approach'. *Mechanical Systems and Signal Processing* **143**, 106836.

Misraji, M., M. Valdebenito, H. Jensen, and C. Mayorga: 2020, 'Application of directional importance sampling for estimation of first excursion probabilities of linear structural systems subject to stochastic Gaussian loading'. *Mechanical Systems and Signal Processing* **139**, 106621.

Schenk, C., H. Pradlwarter, and G. Schuëller: 2005, 'Non-stationary response of large, non-linear finite element systems under stochastic loading'. *Computers & Structures* **83**(14), 1086–1102.

Schuëller, G. and H. Jensen: 2008, 'Computational Methods in Optimization Considering Uncertainties – An Overview'. *Computer Methods in Applied Mechanics and Engineering* **198**(1), 2–13.

Schuëller, G. and H. Pradlwarter: 2007, 'Benchmark Study on Reliability Estimation in Higher Dimensions of Structural Systems – An Overview'. *Structural Safety* **29**, 167–182.

Shinozuka, M. and Y. Sato: 1967, 'Simulation of nonstationary random process'. *Journal of the Engineering Mechanics Division* **93**(1), 11–40.

Stefanou, G.: 2009, 'The stochastic finite element method: Past, present and future'. *Computer Methods in Applied Mechanics and Engineering* **198**(9-12), 1031–1051.

Tropp, J.: 2004, 'Topics in Sparse Approximation'. Ph.D. thesis, The University of Texas at Austin.

Valdebenito, M. and G. Schuëller: 2010, 'A Survey on Approaches for Reliability-based Optimization'. *Structural and Multidisciplinary Optimization* **42**(5), 645–663.

Zerva, A.: 2009, *Spatial Variation of Seismic Ground Motions – Modeling and Engineering Applications*. CRC Press.

Estimation of fragilities by the modified intensity measures (IMs)

Matteo Ciano[1*], Massimiliano Gioffrè[1], Mircea Grigoriu[2]
[1] *Dept. of Civil and Environmental Engineering, Univ. of Perugia, Via G. Duranti 93, 06125 Perugia, Italy*
matteo.ciano@unipg.it, massimiliano.gioffre@unipg.it
[2] *Dept. of Civil and Environmental Engineering, Cornell University, Ithaca, NY 14853, USA*
mdg12@cornell.edu

Abstract. Fragility curves are commonly computed to estimate structural system performance. Seismic fragilities are the probability that the structural response of a system overcomes prefixed limit states for given seismic intensity measure. The most widespread procedure to estimate seismic fragility curves is based on scaling seismic accelerograms by a reference intensity measure (e.g. single/multiple ordinates of the pseudo-acceleration response spectrum or peak ground acceleration). Recently, it was shown that this methodology gives limited if any information on the structural seismic performance when the dependence between the intensity measure and the system demand parameter of interest (e.g. max inter-story displacement) is weak. However, the widespread method is currently applied in Performance-Based Earthquake Engineering because it is simple and it overcomes the problem of the limited number of natural recorded ground motions available for fragility analysis.

In this work a general approach to improve the accuracy in fragilities estimation when the dependence between the intensity measure and the demand parameter is weak and the widely used method does not give accurate results is presented. The proposed algorithm is based on a linear transformation of samples of a given intensity measure, which improves the correlation with a set of demand parameters. Fragility curves are obtained using the transformed intensity measure samples and compared with those estimated with the standard approach. The effectiveness of the proposed algorithm is demonstrated for a linear elementary oscillator and complex multi-degree of freedom structural system.

Keywords: Seismic intensity measures, Modified seismic intensity measures, Fragility analysis, Fragility cureves, Earthquake engineering, Complex mdof structural systems.

1. Introduction

In Performance-Based Earthquake Engineering (PBEE) the evaluation of structural efficiency by fragility analysis is formalized through a methodology with a probabilistic basis. The greatest interest for structural engineering is focused to two parameters in PBEE, that are the seismic intensity measure IM and the system demand parameter D, which reflect the seismic hazard and the structural response, respectively. These parameters are used to develop the fragility analysis and

* Corresponding author.

© 2021 *by M. Ciano, G. Gioffrè, M. Grigoriu.*

therefore build the fragility curve. Seismic fragilities quantify the probability that the structural response of a system overcomes specified limit states for given seismic intensity measure (IM). The intensity measure is an intermediate variable that connects the seismic fragility analysis with the structural analysis, and the demand parameter depends intensely on the chosen IM. Specific features are required of an ideal IM, it should be efficient and sufficient. The efficiency is inherent to the low dispersion that should be characterize the demand parameters and IM, while the sufficiency implies that D evaluated by a seismic ground motion acceleration with a IM value should be only dependent on the value of this IM, and not on other seismic ground motion features (i.e. fault mechanisms, magnitude, etc).

The common method used in PBEE to estimate the seismic fragility curves, recommended by Federal Emergency Management Agency (FEMA), is based on scaling the seismic accelerograms with a reference intensity measure IM (FEMA, 2012). This method is widely applied because it is simple and, together with Monte Carlo simulation, it overcomes the problem of the limited number of natural recorded ground motions available for fragility analysis. Generally, two groups of IMs are considered to define the seismic fragilities in fragility analysis: (i) functional of samples of seismic ground acceleration process $A(t)$, such as peak ground acceleration (PGA); (ii) functional of filtered version of samples of $A(t)$, e.g. single/multiple ordinates of pseudo-acceleration response spectrum $S_a(T)$ for different structural system periods T (O'connor and Ellingwood, 1992; Lopez Garcia and Soong, 2003; Sevieri et al., 2020). In particular, the IMs in the second group, that are widely used to define the fragilities, depend on the system demand parameter D (e.g. max inter-story displacement) on which the analysis is based (Kwong et al., 2014; Ebrahimian et al., 2015; Kwong et al., 2015a; Kwong et al., 2015b).

Recently, it was shown that the method in (FEMA, 2012) gives limited if any information on the structural seismic performance when the dependence between the IM and D is weak (Grigoriu, 2016). When the two variables (i.e. IM and D) are not correlated it is a clear violation of the efficiency condition that the IM must possess. The dependence between these variables plays a fundamental role on the accuracy in fragility analysis (Ciano et al., 2018a; Ciano et al., 2020a). In particular, this dependence depends on the seismic direction, the intensity measure and the demand parameter considered. For the same IM the dependence can vary significantly with different seismic directions, with which the fragility analysis is developed. The fragilities defined as function of $S_a(T)$ for nonlinear single degree of freedom (SDOF) system to a seismic acceleration process $A(t)$ have large uncertainties when D and $S_a(T)$ are weakly correlated. Furthermore, the fragilities defined as function of multiple ordinates of $S_a(T)$ have poor improvements in reference to the case of single ordinate of $S_a(T)$. It was demonstrated that the dependence between IMs and various system demand parameters, D, is weak for nonlinear system and also for complex multi-degree of freedom (MDOF) linear structures (Grigoriu, 2016; Ciano et al., 2018a; Ciano et al., 2020a). The method in (FEMA, 2012) based on scaled seismic ground motion acceleration is very useful, this is not questioned, but it needs the use of efficient IM to give accurate results. The dependence problem can be solved by defining the fragilities as functions of the parameters of the law of $A(t)$. This approach was introduced in (Kafali and Grigoriu, 2007) and used in (Radu, 2017; Radu and Grigoriu, 2018; Ciano et al., 2019) to estimate fragility surfaces. However, these surfaces require a

high computational cost, and they are dependent on moment magnitude and source to site distance.

This paper presents a general approach to improve the accuracy in fragilities estimation when the dependence between the intensity measure IM and the demand parameter D is weak and the widely used method in Performance-Based Earthquake Engineering (FEMA, 2012) does not give accurate results. This general approach is based on a modified version of the current intensity measure method. In particular, once an IM is chosen it is mapped in a suitable space where D and the IM are correlated. The proposed algorithm is based on a linear transformation of samples of a given IM, which improves the correlation with a set of demand parameters. Fragility curves are obtained using the transformed IM samples and compared with those estimated with the standard approach. Numerical results are reported for a simple linear SDOF system and for a real multi-degree of freedom structural system.

Finally, the effectiveness of this general approach is demonstrated for both linear SDOF and real complex MDOF system by comparing the fragility estimates obtained with a chosen seismic intensity measure and the its modified version.

2. Background

Let I be the set of demand parameters which yield a structural failure. The fragility can be defined as

$$P_f(\xi) = P(D \in I \subset \mathbb{R} | IM = \xi) = E[1(D \in I | IM = \xi)] = \int_I f_{D|IM}(x|\xi)\, dx \tag{1}$$

i.e. the probability that a structural system enters a damage state given a ground motion with scalar/vector intensity measure $IM = \xi$. The quantities $1(\cdot)$ and $f_{D|IM}(\cdot|\xi)$ indicate the indicator function and probability density function (PDF) of the conditional variable $D|(IM = \xi)$, respectively. The fragility in (1) is usually estimated from the structural response to scaled seismic time histories, $a(t)$ of the stochastic process $A(t)$ (Baker, 2010), and its accuracy depends on the scaling procedure, the sample size and the IMs properties.

When D and IM are strongly dependent, the random variable $D|IM$ has small variance, i.e. $f_{D|IM}(\cdot|\xi)$ is concentrated about its mean value. On the contrary, when D and IM are weakly dependent the random variable $D|IM$ has large variance. In the limit, $f_{D|IM}(\cdot|\xi)$ becomes a δ-function or D and $D|IM$ have the same PDF when D and $D|IM$ are perfectly correlated or independent, respectively (Grigoriu, 2016). In the latter case, the fragility in (1) does not depend on ξ. Within this context, especially if (1) is estimated by the commonly used method (FEMA, 2012), it is crucial to quantify the dependence between the demand parameter D and various IM definitions to implicitly determine whether or not fragilities, defined as function of the commonly used IMs, can provide useful information for PBEE of actual complex MDOF linear and nonlinear structural systems. If the dependence between D and IM is strong, the fragility $P_f(\xi)$ gives accurate information (Ciano et al., 2018a; Ciano et al., 2020a). The opposite holds when the dependence between D and IM is weak. It is worth noting that, the efficiency condition which the IM must respect is a consequence of the correlation between D and IM. The violation of this condition is determined when D has a large dispersion for given IM, i.e. D and IM are independent. Several

statistical tools can be used to investigate the dependence between the random variables D and IM, which includes correlation coefficients, copula models and multivariate extreme value theory (Grigoriu, 2016). In this work the correlation coefficient is used to quantify the dependence between D and IM.

2.1. Intensity measure (IM)

The IMs used in PBEE to scale seismic acceleration ground motion are divided into three broad categories:

1. non-structures-specific scalar IMs;
2. structures-specific scalar IMs;
3. vector-valued IMs.

A complete classification of the most known IMs in literature is reported in (Ebrahimian et al., 2015). In this work, two most common scalar version of IMs are considered to estimate the fragilities

$$IM_1 = PGA \tag{2}$$

$$IM_2 = S_a(T_1, \zeta_1) \tag{3}$$

where T_1 is the first period of the MDOF structural system with associated damping ratios ζ_1. The period T_1 is taken to in account since it is the fundamental one of the structure as is customary in literature (Ebrahimian et al., 2015).

2.2. Fragilities estimation

The seismic fragility analysis, as also recommended by FEMA, is commonly based on a four-steps algorithm consisting of scaling seismic accelerograms by a reference IM:

1. selection of a finite set of intensity measures $\{\xi_k\}$, $k = 1, \ldots, N$;
2. generation of n_s independent samples $a_i(t)$, $i = 1, \ldots, n_s$ of $A(t)$;
3. for each of the values $\{\xi_k\}$ scale the n_s acceleration records, $a_i(t)$, in order to have the intensity level $IM = \xi_k$, $\xi_k > 0$;
4. for each of the values $\{\xi_k\}$ estimate fragility as

$$\hat{P}_f(\xi_k) = \sum_{i=1}^{n_s} 1(d_{k,i} \in I)/n_s \tag{4}$$

where $1(\cdot)$ is the indicator function and $d_{k,i}$, $i = 1, \ldots, n_s$, are samples of the demand parameters computed for each intensity level ξ_k by linear/nonlinear dynamic analysis. The previous four-steps algorithm are used to construction fragility curves in reference to the generic intensity measures,

e.g. the Equations (2)-(3). The probability estimation that maximum system response exceeds a critical limit state in Equation (4) is common applied (O'connor and Ellingwood, 1992; Hwang and Huo, 1994; Song and Ellingwood, 1999; Schotanus et al., 2004).

3. Modified intensity measure (IM^*)

Let $X(t)$ be the response vector-valued process of an arbitrary MDOF structural system to the ground acceleration $A(t)$. The demand parameter is defined as

$$D^{(j)} = \max_{0 \le t \le \tau} |h(X(t))| \,, \quad j = 1, \ldots, m \tag{5}$$

where τ is the time length of $A(t)$, while $h(\cdot)$ is a function mapping the response $X(t)$ into the j-th system demand parameter (e.g. maximum displacement/acceleration) and m is the number of demand parameters of interest. The dependence between the random variables $D^{(j)}$, $j = 1, \ldots, m$, and IM is a measure to quantify the accuracy in the estimation of $P_f(\xi)$ (Ciano et al., 2020a). When this dependence is weak the widely used method in PBEE (FEMA, 2012) does not give accurate results. To overcome this issue it is proposed to replace samples of a chosen standard intensity measure IM with a suitable linear transformation that modifies the dependence with the selected demand parameters samples. In particular, for each n_s samples, $a_i(t)$, $i = 1, \ldots, n_s$, of the random process $A(t)$ the corresponding samples of IM and $D^{(j)}$ are evaluated, i.e. im_i and $d_i^{(j)}$, $i = 1, \ldots, n_s$. For each of these samples, the z-scores $z(d_i^{(j)})$ and $z(im_i)$, $i = 1, \ldots, n_s$, of the standardized random variables

$$z(D^{(j)}) = \frac{D^{(j)} - E[D^{(j)}]}{\sigma(D^{(j)})} \,, \quad j = 1, \ldots, m \tag{6}$$

$$z(IM) = \frac{IM - E[IM]}{\sigma(IM)} \tag{7}$$

are computed, respectively. It is worth noting that if the random variables defined in Equations (6) and (7) are linearly dependent, the correlation coefficient is equal to one and the scatter plots of pairs of samples $z(d_i^{(j)})$ and $z(im_i)$ describe straight line. Given the demand parameter $D^{(j)}$, for each pair of samples $z(d_i^{(j)})$ and $z(im_i)$ it is possible to evaluate the distance from perfect correlation

$$e_i^{(j)} = z(d_i^{(j)}) - z(im_i) \,, \quad i = 1, \ldots, n_s, \quad j = 1, \ldots, m \tag{8}$$

and estimate the average distance of the j demand parameters

$$\bar{e}_i = \frac{1}{m} \sum_{j=1}^{m} e_i^{(j)} \,, \quad i = 1, \ldots, n_s \,. \tag{9}$$

The samples obtained with Equation (9) can be considered as realizations of a vector $\bar{E}$. The linear transformation

$$im_i^* = [z(im_i) + \bar{e}_i]\sigma(IM) + E[IM] \,, \quad i = 1, \ldots, n_s \tag{10}$$

gives samples that can be considered the realizations of a new intensity measure IM^*.

The procedure reported above describes a general algorithm that can be used to transform any intensity measure IM (e.g. Equations (2) and (3)) into samples of its modified version, IM^*. It is worth noting that for $m = 1$ (i.e. one demand parameter or SDOF system), the correlation estimated from the n_s samples of D and IM^* is exactly one. When $m > 1$ it is not possible to have perfect correlation, but the correlation between $D^{(j)}$ and IM^* is significantly higher than the one estimated from samples of $D^{(j)}$ and IM.

It follows that the IM^* samples obtained with Equation (10) can be used to scale the ground acceleration records to build fragility curves that give more accurate information on the structural system performance when compared with the standard intensity measures.

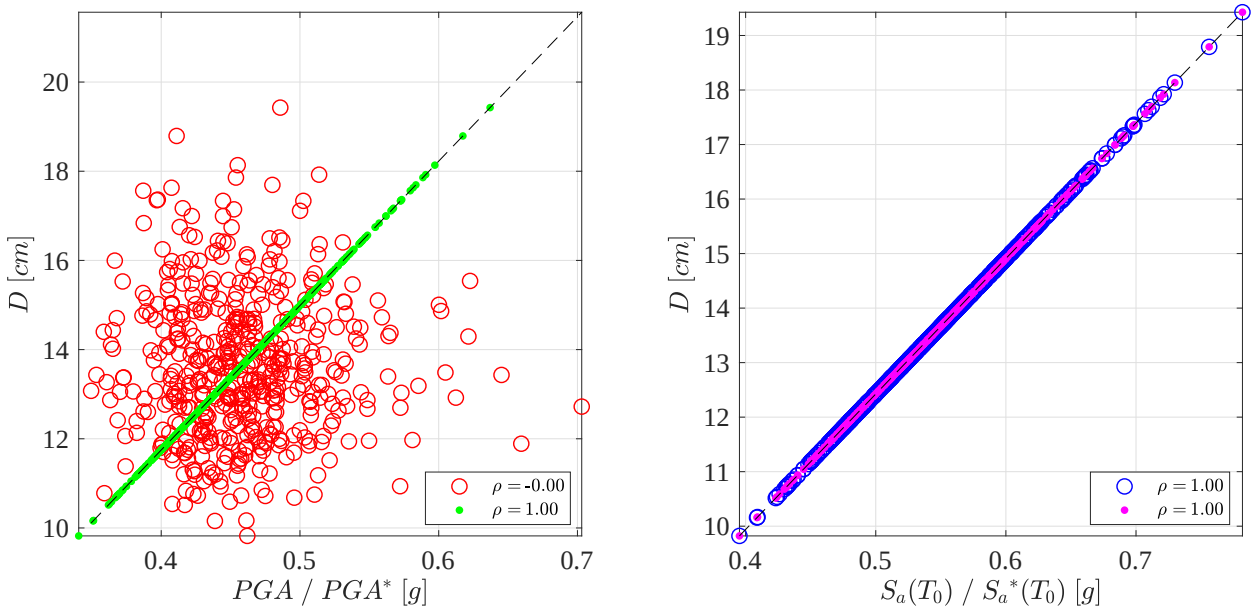


Figure 1. Scatter plots of $n_s = 500$ samples for linear SDOF system with $\omega_0 = 6.28$ rad/sec and $\zeta = 5\%$: (PGA, D) red circles, (PGA^*, D) green dots (left panel); ($S_a(T_0)$, D) blue circles, ($S_a{}^*(T_0)$, D) magenta dots (right panel).

4. Single degree of freedom application

In this Section the results of linear SDOF system about fragility estimation are reported. The Monte Carlo simulation four-steps algorithm in Section 2.2 is used. To aim this, two different intensity measures IM_l, $l = 1, 2$, i.e. Equations (2)-(3), and their modified version are considered. In this context, the first period T_1 and ζ_1 in (3) coincides with the natural period T_0 and the damping ration ζ of SDOF system. The approach proposed in (Cacciola, 2010) is used for the generation of

n_s independent samples of $A(t)$ as spectra-compatible stochastic process.

Let $X(t)$ be the response process of SDOF oscillator under the spectra-compatible stochastic process $A(t)$ that satisfies the differential equation

$$\ddot{X}(t) + 2\zeta\omega_0\dot{X}(t) + {\omega_0}^2 X(t) = -A(t) \tag{11}$$

where $\omega_0 = 2\pi/T_0$ is the natural pulsation in rad/sec. The numerical solution of Equation (11) are developed for $\omega_0 = 6.28$ rad/sec, $\zeta = 5\%$ to compute the demand parameter $D = D^{(1)}$, $m = 1$, as the maximum absolute relative displacement, by Equation (5).

Figure 1 shows the scatter plot of $n_s = 500$ samples of D against the intensity measures PGA and $S_a(T_0)$, in red and blue circles, in left and right panel, respectively, together with the estimated correlation coefficient ρ. These two results for linear SDOF system, as reported in literature (Kafali and Grigoriu, 2007), demonstrated that the PGA and $S_a(T_0)$ are unsatisfactory and satisfactory intensity measure since ρ estimated in relation to D is zero and one, respectively. Moreover, the PGA is a inefficient IM because it determines a high dispersion on D, as opposed to $S_a(T_0)$.

Samples of the modified intensity measure PGA^* and ${S_a}^*(T_0)$ are evaluated to improve the dependence with the demand parameter. First, the n_s samples of the D and IM_l, $l = 1, 2$, are standardized using Equations (6) and (7) and the distance from the perfect correlation is evaluated by Equation (8) with $m = 1$. Second, the n_s average distances are evaluated using Equation (9).

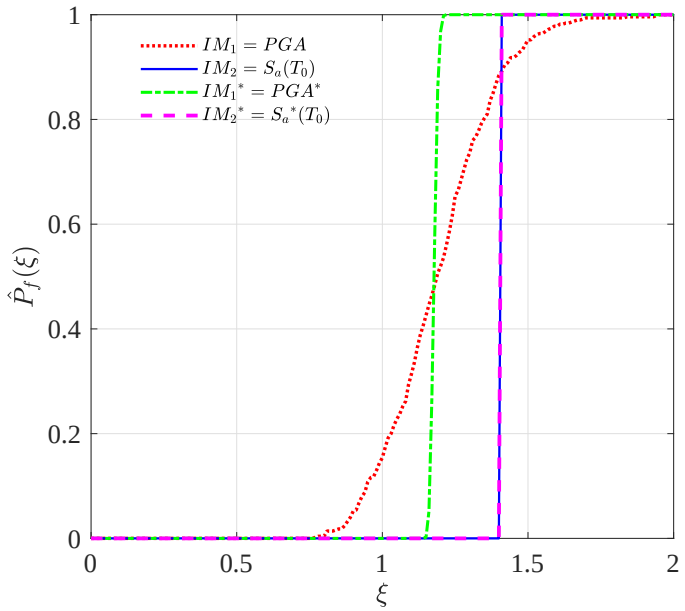


Figure 2. Fragilities against intensity level ξ considering different definitions of IMs and its modified version IMs^* for linear SDOF system with $\omega_0 = 6.28$ rad/sec, $\zeta = 5\%$ and limit state $\bar{D} = 35$ cm.

Finally, the samples of the modified intensity measure version $IM_l{}^*$, $l = 1, 2$, are computed by Equation (10). In particular, since $m = 1$ (SDOF system) the correlation between D and IM^* is exactly one and the samples computed using the Equation (9) coincide with the those by Equation (8). Continuing the reference to the Figure 1, in green and magenta dots the scatter plots of (PGA^*, D) and ($S_a{}^*(T_0)$, D) are shown on the left and right panel. Clearly, since D for the SDOF system is also equal to $S_a(T_0)/{\omega_0}^2$ the random variables $S_a(T_0)$ and $S_a{}^*(T_0)$, i.e. the results in blue circles and magenta dots, coincide.

Figure 2 reports fragility curves estimated by four-steps algorithm (Section 2.2) using the intensity measure definitions IM_l, $l = 1, 2$, and their modified version IM_l^*, $l = 1, 2$, for the system demand parameter D assuming limit state $\bar{D} = 35$ cm. In particular, the dotted red line regards to the PGA and it has poor information on the linear SDOF system performance since D and IM_1 are uncorrelated (left panel of Figure 1), while the best information are provided when ρ is unitary in order to the curve by PGA^*, dot dash green line, that becomes a $\delta-$function (see Section 2). This last aspect is also valid for the curves computed with $S_a(T_0)$ and $S_a{}^*(T_0)$, and them being equal (right panel of Figure 1) the continuous blue and dashed magenta lines coincide.

5. Complex multi-degree of freedom application

The procedure described in Section 2.2 is used to build fragility curves of a structural system of a school in Norcia (Italy) that was affected by the 2016 earthquake sequence. The fragility analysis is developed with standard intensity measures IM_l, $l = 1, 2$, (Equations (2) and (3)) and their

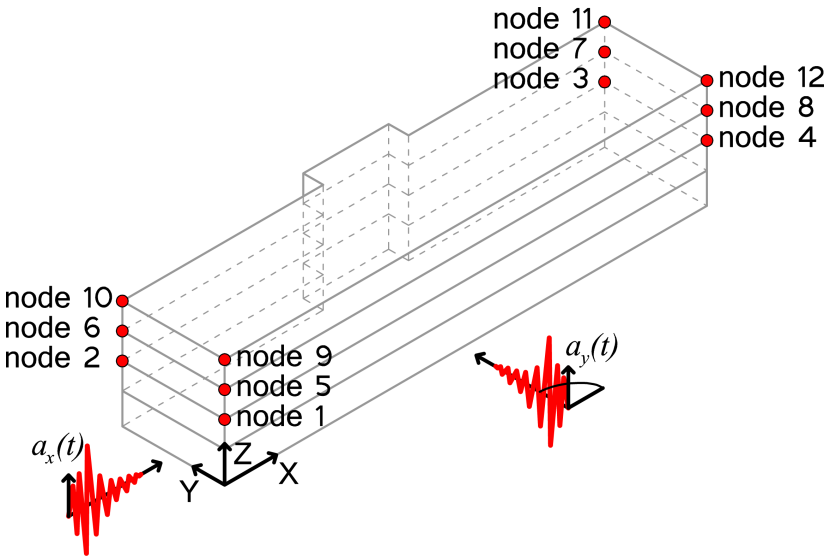


Figure 3. Seismic directions used in the analyses and reference node locations.

modified version that are computed with the proposed procedure in Section 3.

The structural system of the school building consists of a reinforced concrete frame (RCF) on four floors with a dissipative bracing system, made of Buckling-Restrained Axial Dampers (BRADs), and inverse beams foundation. The RCF and foundation are modeled with classical beam elements, while the BRAD elements are modeled with the constitutive law described in (Wen, 1976). The obtained numerical model details on the school building and the model modal parameters are reported in (Ciano et al., 2018a).

Linear analyses are performed to evaluate the structural response. In this work the maximum absolute displacement in the y-direction D_{dy} (Fig. 3) is considered as demand parameter in the fragility analysis.

The strong event recorded in Norcia on October 30th, 2016 is used to calibrate the non-stationary stochastic process $\boldsymbol{A}(t)$. Structural analyses are developed for each samples of the process $\boldsymbol{A}(t) = \{A_x(t); A_y(t)\}$, i.e. $\boldsymbol{A}(t)$ is a stochastic process with independent components, that acts in x and y global-directions of the finite element numerical model (FEM) (Figure 3). The non-stationary model described in (Grigoriu et al., 1988; Ciano et al., 2018b; Ciano et al., 2020b) is used to generate samples $a_x(t)$ and $a_y(t)$ of $A_x(t)$ and $A_y(t)$, respectively. The probabilistic model used for $\boldsymbol{A}(t)$ is

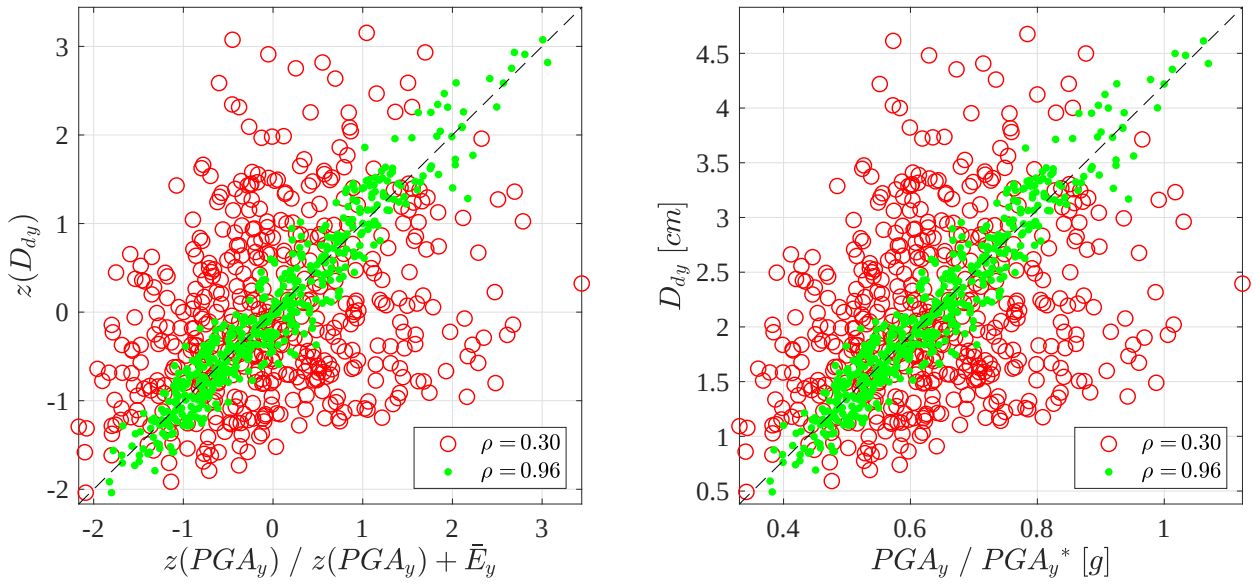


Figure 4. Scatter plots of $n_s = 500$ samples at node #10: $(z(PGA_y), z(D_{dy}))$ red circles, $(z(PGA_y) + \bar{E}_y, z(D_{dy}))$ green dots (left panel); (PGA_y, D_{dy}) red circles, $(PGA_y{}^*, D_{dy})$ green dots (right panel).

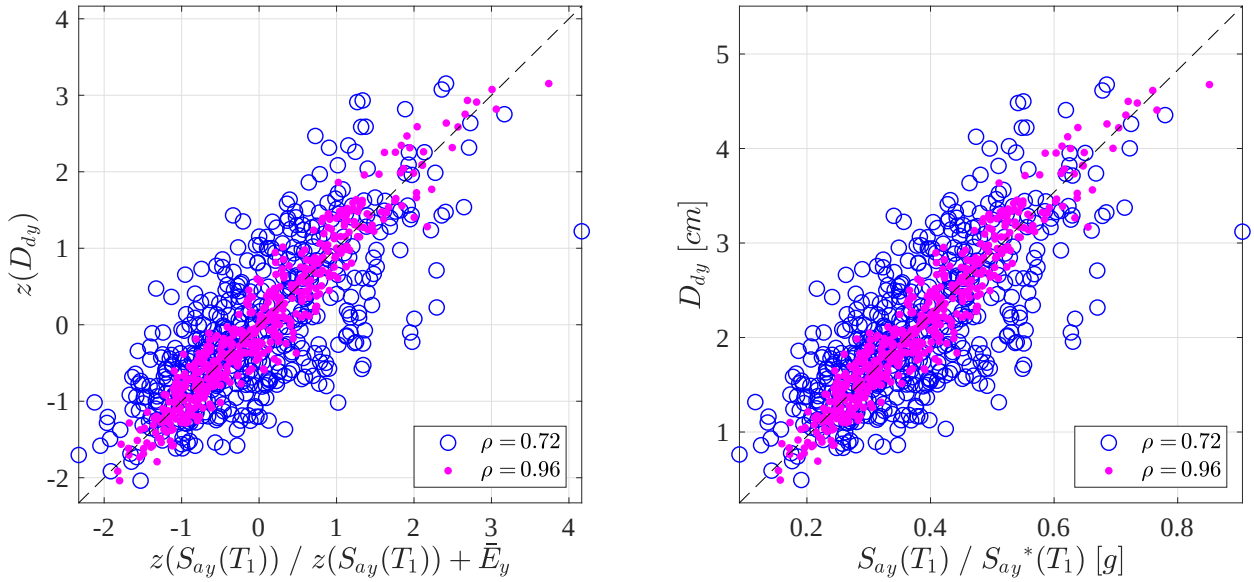


Figure 5. Scatter plots of $n_s = 500$ samples at node #10: $(z(S_{ay}(T_1)), z(D_{dy}))$ blue circles, $(z(S_{ay}(T_1)) + \bar{E}_y, z(D_{dy}))$ magenta dots (left panel); $(S_{ay}(T_1), D_{dy})$ blue circles, $(S_{ay}{}^*(T_1), D_{dy})$ magenta dots (right panel).

calibrated with actual time histories recorded at the school base during the October 30th, 2016 earthquake.

The demand parameter D_{dy} is obtained by a three-step Monte Carlo algorithm: (i) n_s ground motion acceleration time series $a_n(t)$ of the process $A_n(t)$, $n = x, y$, are generated using the method proposed in (Grigoriu et al., 1988); (ii) linear time domain numerical dynamic structural analyses are used to obtain response samples $\boldsymbol{x}(t)$ of $\boldsymbol{X}(t)$ assuming proportional damping ratio $\zeta = 5\%$; (iii) n_s samples of the random vector $D_{dy} = \{D_y^{(1)}, D_y^{(2)}, \ldots, D_y^{(m)}\}$ are computed using Equation (5). In this case study the maximum absolute displacement in y-direction at the $m = 12$ node positions shown in Figure 3 are selected as demand parameters.

The dependence between the random variables $D_y^{(j)}$, the two IMs in Equations (2) and (3) and their modified version are investigated. In particular, the results regards the max absolute displacement in y-direction $D_{dy} = D_y^{(j)}$, $j = 10$, (Fig. 3), and the two intensity measures IM_{ly}, $l = 1, 2$, estimated from samples of $A_y(t)$ are shown in Figures 4 and 5, respectively, together with the estimated correlation coefficient ρ. These results are computed for $n_s = 500$ samples. The results presented in the right panel of Figure 4 (red circles) demonstrate that the peak ground acceleration in the y-direction, $PGA_y = IM_{1y}$, represents an unsatisfactory intensity measure given the low value of the correlation coefficient, $\rho = 0.30$, between D_{dy} and IM_{1y}. The fragility estimated using this IM would provide poor information on the seismic structural performance. Referring to right panel of Figure 5 (see blue circles), the same holds for $IM_{2y} = S_{ay}(T_1)$ since the conditional random

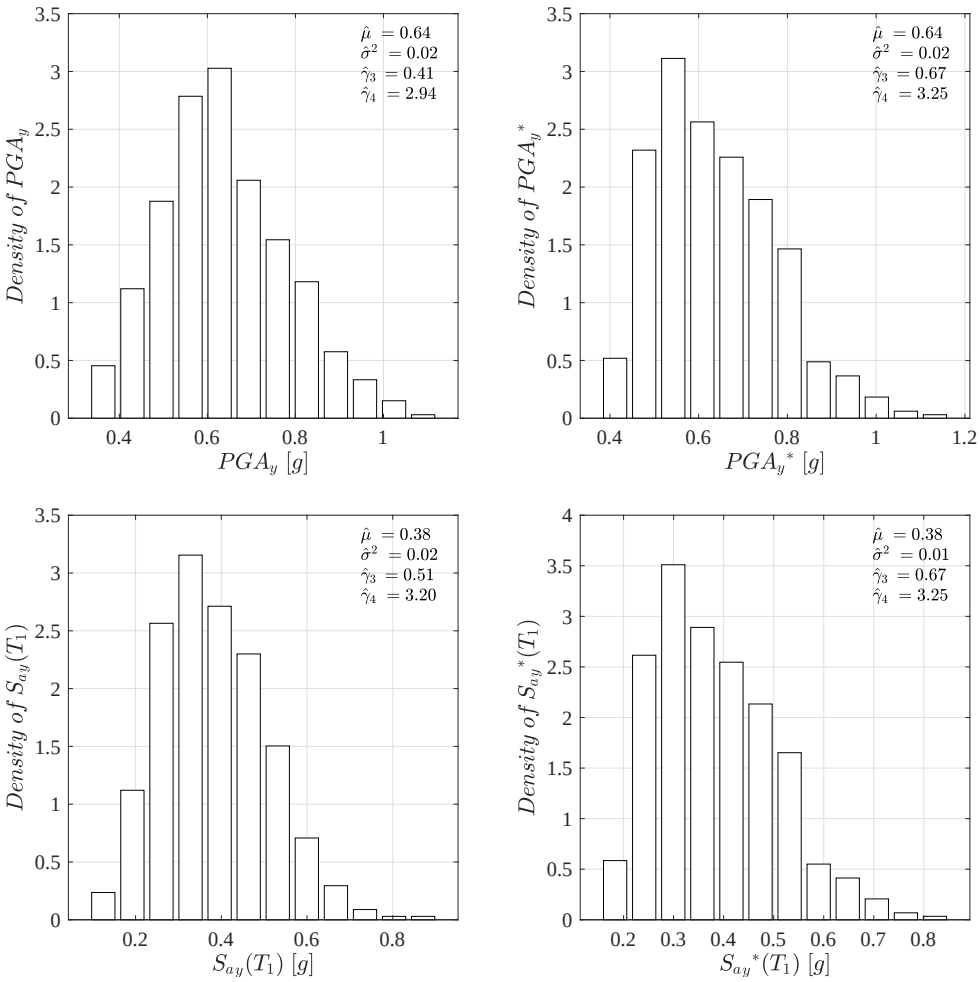


Figure 6. Estimated PDF for $n_s = 500$ samples of PGA (left top panel), PGA^* (right top panel), $S_a(T_1)$ (left bottom panel) and $S_a{}^*(T_1)$ (right bottom panel).

variable $D_{dy}|IM_{2y}$ has large variance (i.e. D_{dy} and IM_{2y} are low correlated). In general, when using the intensity measure in Equation (3), it was demonstrated that IM_{2n}, $n = x, y$, can be both strongly and weakly correlated with the demand parameter depending on the selected structural response of interest and the earthquake direction (Ciano et al., 2020a).

The initial correlation between the chosen m demand parameters and the IMs described above can be improved using the procedure presented in Section 3. Samples of the demand parameters and the two intensity measures in Figures 4-5 are first transformed into their standardized versions using Equations (6) and (7) and the distance from perfect correlation is evaluated by Equation (8) for each of the m demand parameters. Then, the n_s average distances are evaluated using Equation

(9) and used to correct the samples of $z(IM_{ly})$, $l = 1, 2$. The last step is to evaluate samples of the modified intensity measure IM^* using Equation (10). In particular, the left panels of Figures 4 and 5 report the scatter plots and the correlation coefficients before and after the correction in the standardized space, i.e. $(z(IM_{ly}), z(D_{dy}))$, and $(z(IM_{ly}) + \bar{E}_y, z(D_{dy}))$, $l = 1, 2$, respectively. The right panels report the scatter plots of the same samples (before and after the correction) linearly transformed back into their original space by Equation (10), i.e. (IM_{ly}, D_{dy}), and $(IM_{ly}{}^*, z(D_{dy}))$, $l = 1, 2$, respectively. In these figures the scatter plots with $IM_{1y} = PGA_y$ are red circles and green dots, before and after the correction, respectively, while those with $IM_{2y} = S_{ay}(T_1)$ are blue circles (before) and magenta dots (after).

The two figures demonstrate the ability of the proposed algorithm to improve the correlation between the selected demand parameters and intensity measures for the MDOF system case. This will result into an improved accuracy in the fragility curves.

Is interesting to note that the linear transformation of the intensity measure samples obtained in this case study does not significantly change the intensity measures first four statistical moments. Figure 6 shows the probability density functions (PDFs) and the first four statistical moments estimated from samples of the original intensity measures IM_{ly}, $l = 1, 2$ (left panels) and their modified versions $IM_{ly}{}^*$ (right panels). In particular, the top panels show the results for $IM_{1y} = PGA_y$, while the bottom panels refer to $IM_{2y} = S_{ay}(T_1)$.

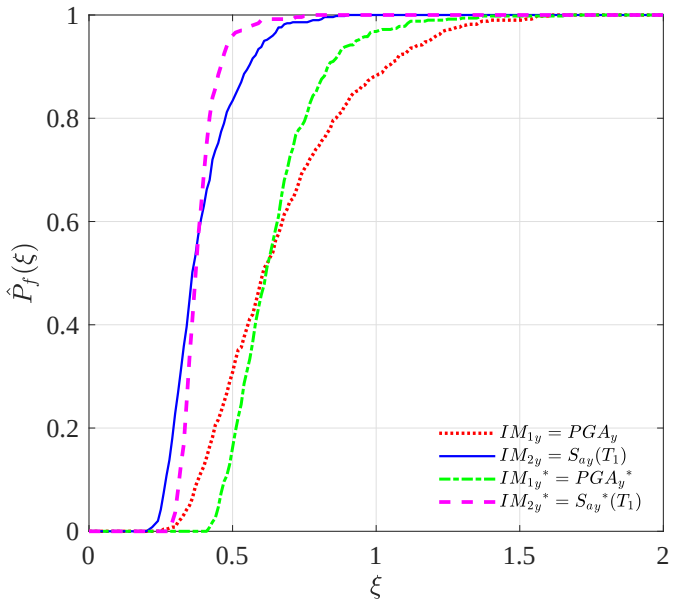


Figure 7. Fragilities against intensity level ξ for different definitions of IMs and its modified version IMs^* at node #10, maximum absolute displacement D_{dy} for limit state $\bar{D}_{d_y} = 2$ cm.

Figure 7 reports the obtained fragility curves of the demand parameter D_{dy} at node #10 for limit state $\bar{D}_{d_y} = 2$ cm. It is worth noting that the fragility curves obtained with the modified intensity measures (dash-dotted green and dashed magenta lines) are steeper than those estimated using the original intensity measures (dotted red and continuous blue lines). This is consistent with the lower variability (i.e. lower variance) of the conditional random variable $D|IM^*$. In general, for a prefixed value of ξ the fragility curves obtained with IM and the respective IM^* provide conflicting information on the structural seismic performance (Fig. 7).

6. Conclusion

In Performance-Based Earthquake Engineering the seismic fragilities are commonly estimated by widespread approach based on scaling the seismic accelerograms with a reference intensity measure (IM). The ideal intensity measure should be efficient and sufficient, this means that IM should determine a low dispersion in the structural system demand parameter D and enclose information, without dependence on other variables, for the evaluation of D, respectively. As a consequence, the application of widespread method requires a strong dependence between the demand parameter D of the structural system, on which the fragility analysis is based, and the IM used for scaling in order to have accurate fragilities. The approach provides limited if any information on the seismic performance of the structural system for weak dependence between D and IM.

This work presents a general approach to improve the accuracy in fragilities estimation when the dependence between the intensity measure IM and the demand parameter D is weak and the widely used method in Performance-Based Earthquake Engineering does not give accurate results. The general approach is based on a modified version of the current intensity measure method. In particular, samples of any chosen intensity measure IM are linearly transformed in order to improve the correlation with a set of selected demand parameters. The new samples can be considered as realizations of a modified intensity measure IM^* and used to scale the ground acceleration records to build fragility curves that give more accurate information on the structural system performance when compared with the standard intensity measures.

The effectiveness of the proposed method in improving the fragility curves accuracy was demonstrated on linear single degree of freedom system and an actual multi-degree of freedom structural system of a school building in Norcia (central Italy). The results demonstrate the effectiveness of the modified seismic intensity measure in fragility analysis. Work is in progress to investigate the accuracy of fragilities based on modified intensity measures in nonlinear structural systems.

References

Baker, J. W. Conditional mean spectrum: Tool for ground-motion selection. *Journal of Structural Engineering*, 137(3):322–331, 2010.

Cacciola, P. A stochastic approach for generating spectrum compatible fully nonstationary earthquakes. *Computers & Structures*, 88(15):889 – 901, 2010.

Ciano, M., Gioffrè, M., and Grigoriu, M. On the accuracy of seismic fragilities for actual non-linear MDOF systems. In *Proceedings of the 8th International Conference on Computational Stochastic Mechanics (CSM 8)*, pages 121–130, 2018a.

Ciano, M., Gioffré, M., and Grigoriu, M. Seismic performance detection by fragility analysis: a comparison between standard and alternative approaches. In *Proceedings of the XVIII ANIDIS Conference Seismic engineering in Italy: Ascoli Piceno, 15-19 September*, pages 226–234. Pisa University Press, 2019.

Ciano, M., Gioffrè, M., and Grigoriu, M. The role of intensity measures on the accuracy of seismic fragilities. *Probabilistic Engineering Mechanics*, 60:103041, 2020a.

Ciano, M., Gioffrè, M., Gusella, V., and Grigoriu, M. Modeling and simulation of non-stationary thunderstorm outflows. In *Conference of the Italian Association for Wind Engineering*, pages 222–233. Springer, 2018b.

Ciano, M., Gioffrè, M., Gusella, V., and Grigoriu, M. D. Non-stationary dynamic structural response to thunderstorm outflows. *Probabilistic Engineering Mechanics*, 62:103103, 2020b.

Ebrahimian, H., Jalayer, F., Lucchini, A., Mollaioli, F., and Manfredi, G. Preliminary ranking of alternative scalar and vector intensity measures of ground shaking. *Bulletin of Earthquake Engineering*, 13(10):2805–2840, 2015.

FEMA. Seismic performance assessment of buildings: Vol. 1–methodology. *Applied Technology Council for the Federal Emergency Management Agency, CA, USA*, 2012.

Grigoriu, M. Do seismic intensity measures (IMs) measure up? *Probabilistic Engineering Mechanics*, 46:80 – 93, 2016.

Grigoriu, M., Ruiz, S. E., and Rosenblueth, E. The mexico earthquake of september 19, 1985 — nonstationary models of seismic ground acceleration. *Earthquake Spectra*, 4(3):551–568, 1988.

Hwang, H. H. and Huo, J.-R. Generation of hazard-consistent fragility curves. *Soil Dynamics and Earthquake Engineering*, 13(5):345–354, 1994.

Kafali, C. and Grigoriu, M. Seismic fragility analysis: Application to simple linear and nonlinear systems. *Earthquake Engineering & Structural Dynamics*, 36(13):1885–1900, 2007.

Kwong, N. S., Chopra, A. K., and McGuire, R. K. A framework for the evaluation of ground motion selection and modification procedures. *Earthquake Engineering & Structural Dynamics*, 44(5):795–815, 2014.

Kwong, N. S., Chopra, A. K., and McGuire, R. K. Evaluation of ground motion selection and modification procedures using synthetic ground motions. *Earthquake Engineering & Structural Dynamics*, 44(11):1841–1861, 2015a.

Kwong, N. S., Chopra, A. K., and McGuire, R. K. A ground motion selection procedure for enforcing hazard consistency and estimating seismic demand hazard curves. *Earthquake Engineering & Structural Dynamics*, 44(14):2467–2487, 2015b.

Lopez Garcia, D. and Soong, T. T. Sliding fragility of block-type non-structural components. part 1: Unrestrained components. *Earthquake Engineering & Structural Dynamics*, 32(1):111–129, 2003.

O'connor, J. M. and Ellingwood, B. R. Site-dependent models of earthquake ground motion. *Earthquake Engineering & Structural Dynamics*, 21(7):573–589, 1992.

Radu, A. A framework for earthquake risk engineering. *Procedia engineering*, 199:3576–3581, 2017.

Radu, A. and Grigoriu, M. An earthquake-source-based metric for seismic fragility analysis. *Bulletin of Earthquake Engineering*, 16(9):3771–3789, 2018.

Schotanus, M., Franchin, P., Lupoi, A., and Pinto, P. Seismic fragility analysis of 3d structures. *Structural Safety*, 26(4):421–441, 2004.

Sevieri, G., De Falco, A., and Marmo, G. Shedding light on the effect of uncertainties in the seismic fragility analysis of existing concrete dams. *Infrastructures*, 5(3):22, 2020.

Song, J. and Ellingwood, B. R. Seismic reliability of special moment steel frames with welded connections: Ii. *Journal of Structural Engineering*, 125(4):372–384, 1999.

Wen, Y.-K. Method for random vibration of hysteretic systems. *Journal of the engineering mechanics division*, 102(2):249–263, 1976.

Analyses of resistance of openings to Pyroclastic Flows

Mauro Iacuaniello[1)], Andrea Montanino[1)], Daniela De Gregorio[2)], Giulio Zuccaro[1,2)]

[1)]*Department of Structures for Engineering and Architecture, University of Naples Federico II, Napoli, Italy. E-mail: mauro.iacuaniello@unina.it, andrea.montanino@unina.it, giulio.zuccaro@unina.it*
[2)] *PLINIVS - LUPT Study Centre, University of Naples Federico II, Via Toledo 402, 80134, Naples, Italy. E-mail: daniela.degregorio@unina.it*

Abstract: During an explosive volcanic eruption, constructions are hit by various inertial or surface actions often associated with high temperatures, that can cause fires and/or explosions that can affect the mechanical properties of structural and non-structural elements. Among the volcanic phenomena the pyroclastic flows have a devastating effect. They are a gas-solid mixture, which can flow slope down up to reach considerable distances from the point of emission, a speed that can easily exceed 100km/h (~ 30m/s). Their temperatures may be higher than 400°C and they cause casualties and problems to the structures. The action exerted by pyroclastic flows on buildings is a dynamic pressure accompanied by high temperatures. The experience of the eruption in Montserrat (British overseas territory in the Caribbean) in 1995 has shown that the openings of buildings are particularly vulnerable to the stresses caused by pyroclastic flows, even when the static nature of the building itself is not compromised, the risk associated with the passage of the flow in the internal environments following the breakthrough of the openings is therefore significant, increasing the risk of fire inside the building. In this work, the objective is to contribute to the understanding of resistance of the characteristic openings of buildings potentially at risk along the flow path, considering the dynamic pressures and temperature ranges associated with a specific scenario at Vesuvius and the Campi Flegrei, volcanoes (Campania Region, Italy) defined in the National Emergency Plans.

Keywords: Mitigation strategies; Volcanic risk; Pyroclastic flows.

1. Introduction

Explosive volcanic eruptions can cause casualties and economic losses unless appropriate mitigation measures can be taken. Inside the metropolitan area surrounding the city of Naples (Campania Region, Italy) there are two active volcanic systems, the Somma-Vesuvius and the Campi Flegrei, whose geological history allows to believe that they will be able to produce, in the future, some explosive eruptions. The hazard of both volcanoes, the high exposed value of the urban area, which counts about three thousand people, and the high vulnerability of the urban settlements make the Neapolitan territory one of the riskiest volcanic areas (Zuccaro and De Gregorio, 2011). Volcanic eruptions encompass different hazards as volcanic earthquake, ashfall, pyroclastic flows and subsequent floods and mudflows. The most dangerous phenomenon is the pyroclastic flows, which can occur with little warning, move at high speeds, and have enormous destructive power (Spence et al., 2010). Although the Emergency Plans provide for the preventive evacuation of the areas affected by the phenomenon, the protection of the openings represents the main objective as it happened during the volcanic eruption on the Caribbean island of Montserrat

© 2021 *by M. Iacuaniello, A. Montanino, D. De Gregorio, G. Zuccaro*

(British Overseas Territory), the openings and especially the unprotected ones were the first elements of the building envelope to be compromised and causing considerable damage to the structure itself. Therefore, such a strategy would allow for faster and cheaper recovery action.
Glass breakage, in the presence of pyroclastic flows, can occur both because of the acting pressures (3 - 7 kPa) and because of the high temperature that can occur by conductive and radiative exchange from the hot materials of the flows themselves. The temperature increase of the transparent surface will cause local expansion and breakage. The forecast temperature represents an important issue to be tackled since some of the windows' components are thermoplastic. Indeed if they reach their temperature of glass transition (Tg), they soften, and it is possible to hypothesize in the case of gaskets the leakage of the seal between the glass pane and the frame, whilst in the case of polyamide the loss of continuity between aluminium sections; so the probability of the failure of the window can be considered very high. To deal with these issues and consequently define the appropriate mitigation strategies, this study shows a different method to assess the vulnerability both for the pressure and for the temperature. Finally, the research includes developing methodologies for the modelling and evaluation of technical solutions, design guidelines, demonstration projects and prototype components for the realization of new industrial products for the building envelope aimed at volcanic risk mitigation and energy saving.

2. Volcanic Phenomena

The Emergency Plans, developed by the Italian Civil Protection Department, are a useful tool for mitigation of volcanic risk, which starts from the examination of the hazard of the volcano (through the choice of the reference event), the exposure of the elements affected by the eruption (people, buildings and infrastructure) and their vulnerability to the eruptive phenomena. Explosive eruptions are expected for Vesuvius, and the Campi Flegrei, characterized by the formation of a sustained eruptive column several kilometres high, the fallout of volcanic bombs and stone blocks near the crater and smaller particles (ash and lapilli) even several tens of kilometres downwind, as well as the formation of pyroclastic flows that would flow along the slopes of the volcano for several kilometres. For the Campi Flegrei, unlike what happens in volcanoes with central apparatus, such as Vesuvius, the area of the possible opening of the eruptive vent is huge; moreover, regarding fallout, it should be considered that, unlike Vesuvius, the city of Naples is downwind of the dominant winds and would therefore be involved. For Vesuvius, as scenario, it has hypothesized a Sub-Plinian event, corresponding to a volcanic explosivity index VEI =4), with a conditioned probability of occurrence preferably less than 30% (Marzocchi et al., 2004). On the other hand, for Campi Flegrei, as eruption scenario, a medium-size event was assumed to occur with a conditioned probability of occurrence rather than about 24% (Costa et al., 2009). The seismic events that characterize an eruptive phenomenon can be generally considered of low-medium intensity. Considering a sub-Plinian I eruption scenario, there would be a marked increase of low magnitude in the pre-eruptive phase and the occurrence of a low to medium magnitude. Some of these precursor events may exceed the size threshold that can cause damage, including severe, to buildings and infrastructure in the areas surrounding the volcano. The initial phase of a sub-Plinian eruption is characterized by the formation of a sustained eruptive column of gas-solid pyroclastic dispersal. It represents an excellent risk for existing buildings because pyroclastic solid material of various size falls to ground from the eruptive column dispersed by wind and the affected area has frequently an elliptical shape elongated in the direction of the wind. Grain size and thickness of ash fall deposits decrease with the eruptive vent's distance and is generally uniform over small areas.

The pyroclastic flows are the most dangerous volcanic phenomenon produced by a sub-Plinian eruption. They are generated by the gravitational collapse of the eruptive column at the end of the ash fall phase. The pyroclastic flows are a suspension of gas and solid particles of various sizes. Their hazard at Vesuvius had been studied by numerical modelling by Todesco et al. (2002) and Esposti Ongaro et al. (2002). In structural assesses to evaluate the vulnerability, the action due to pyroclastic flows could be considered a uniformly distributed static pressure (Petrazzuoli and Zuccaro, 2004), within a temperature range between 200 and 350 °C (Gurioli et al. 2008). In previous studies, (Esposti Ongaro et al. 2008; Neri et al. 2007), through a 4D model, the Vesuvius was schematized with its real geometric dimension, and the variable time is included. After 900 s since the pyroclastic flow's origin, a pressure of 1-3 kPa at 7.5 km from the vent, with a temperature of 250 °C was assessed.

The experience from the 1997 Montserrat eruption (UK Caribbean Islands) had shown that a building could withstand moderate pyroclastic flows pressure (1–5 kPa), whereas if one or more openings (windows and doors) fail, allowing hot gas and ash to pass in, the entire building is likely to be destroyed (Baxter et al. 2005). In this case, the contents of the construction and any timber structure are likely to catch fire; simultaneously, the principal structural walls and roofing will suffer a combination of internal and external pressures, which will cause partial or total failure (Spence et al. 2004a). In general, the first elements to collapse are the glass windows and the shutters. Nevertheless, a building's lateral resistance to pyroclastic flow strongly depends on the design criteria applied to resist ordinary load conditions: of course, an earthquake-resistant building has more considerable strength and stiffness capabilities than a non-seismic building. The structural behaviour of buildings under pyroclastic flows is not comparable to that induced by earthquakes, since the horizontal pressure is not a cyclic action. So, the structural response is less influenced by the ductility, like the capability to dissipate energy.

3. Exposure

3.1. Description of building data of Vesuvius and Campi Flegrei areas

In order to define the best mitigation strategies, it has been necessary to arrange some available data about the buildings in the areas that will be evacuated before the eruption (called ‘red zones’), gathered up by the P.LIN.V.S. Among these data, the useful ones for the purpose are the vertical structure, material of frame and shutters. Indeed, different openings are used in the different structures both in the Vesuvian and Phlegrean areas; there is a wide diffusion of buildings framed in reinforced concrete with thick infills panels and masonry structures with square blocks in brick or tuff. Besides, there is a widespread diffusion of aluminium and wood windows with UPVC shutters for both the areas (Fig.1, Fig.2). In this study, the most studied openings are the windows, which, together with the doors, represent a weak point of the building envelope during a pyroclastic flow event (Spence et al. 2004). since the dynamic pressure exceeds the characteristic resistance of them, increasing the vulnerability. So, the data about openings have been divided into three groups:Size of openings;

- Frame types;
- Shutters types.

Each of these characteristics is important in assessing the vulnerability and so in defining the adequate mitigation measures. Besides, the sizes of openings were recorded in three classes:

- Large windows, whose area is greater than 1,5 m^2;

- Typical windows, whose area range from 0.5 to 1.5 m^2;
- Small windows, whose area is less than 0.5 m^2.

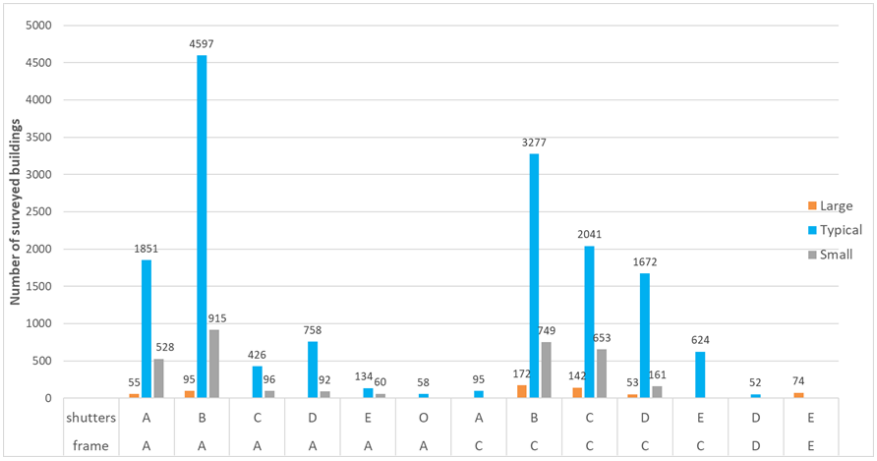


Figure 1. Number of windows in Vesuvius area divided by size and material

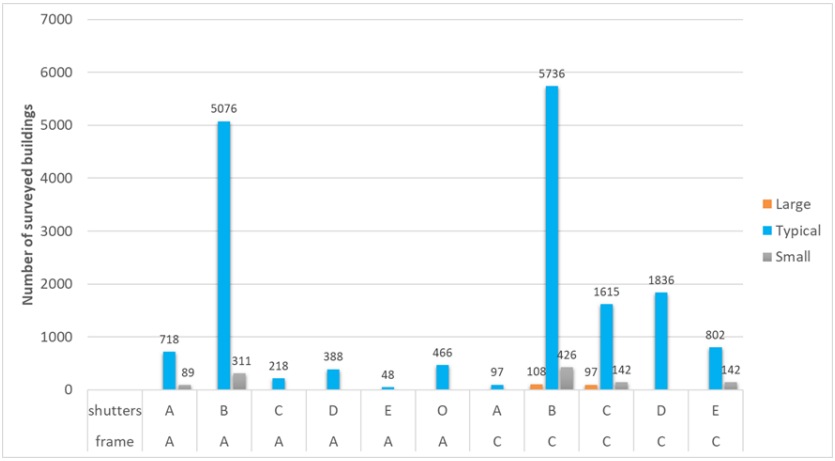


Figure 2. Number of windows in Phlegrean area divided by size and material.

3.2. ALUMINIUM WINDOWS

The EN AW-6060 alloy is the most widespread extrusion alloy on the European market, thanks to its high hot forming speed. The alloy allows the production of profiles with even complex sections, including cavities and multiple grooves, to bring the design of the extruded part as close as possible to that of the finished product and to minimize intermediate machining. The mechanical model used is Ramberg-Osgood (1) and the characteristics of the alloy are (Table 1):

Table 1 Physical and mechanical properties of aluminium EN-AW 6060	
Density	2700 [kg/m^3]
Elastic modulus	70000 [MPa]
Breaking voltage	160 [MPa]
Poisson's ratio	0.33
Specific heat capacity	900 [J/kgK]
Thermal conductivity	238 [W/mK]
Thermal expansion	3,7 e^{-7} [1/K]

$$\varepsilon = \frac{\sigma}{E} + K\left(\frac{\sigma}{\sigma_y}\right)^{\frac{1}{n}} \quad (1)$$

- σy is the yield strength of the material,
- σ is the value of the stress considered,
- E Young's modulus,
- n exponent of the hardening of the material.

The type of glass commonly used is composed of silica oxide and lime. As defined in the Instructions for the design, execution and control of constructions with structural glass elements, the latter can be considered a homogeneous, isotropic material with linear elastic behaviour at breakage, both tensile and compressive. The characteristics of this type of glass (Table 2) are:

Table 2 Physical and mechanical properties of soda lime glass	
Density	2400 [kg/m^3]
Elastic Modulus	71000 [MPa]
Ultimate Tensile Strength	41 [MPa]
Compressive strength	300 [MPa]
Poisson's ratio	0.33
Specific Heat Capacity	800 [J/kgK]
Thermal Conductivity	1 [W/mK]

The technology hypothesized, as the most common, is that of insulating glass, which indicates the set of two or more sheets of equal or variable thickness, separated by a cavity, usually of air. For the analyses, two panes of the same thickness, i.e. (4/5/6) mm. The geometric model used is thermal break window and door frame (Fig. 3).

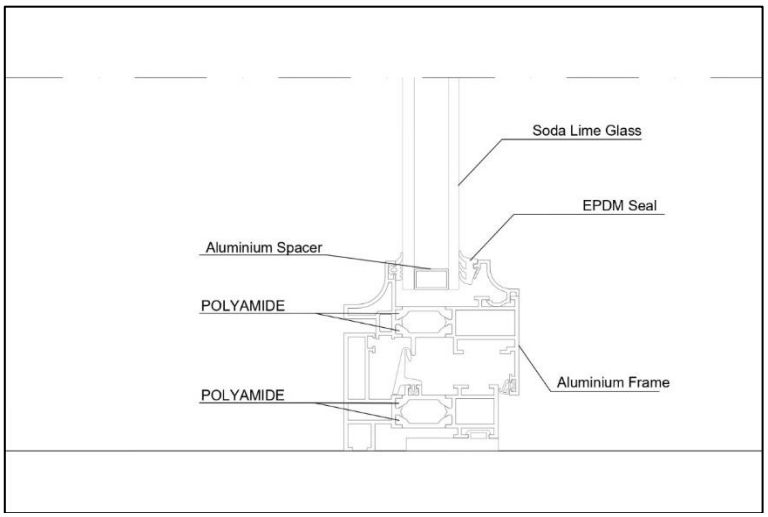


Figure 3. Thermal break aluminum window section.

3.3. TIMBER WINDOWS

The wood is the other material widespread in the Campi Flegrei and Vesuvian area for the construction of windows and doors. The choice of this material depends mainly on the good thermal insulation characteristics compared to UPVC and aluminium windows and doors. In fact, this choice is economically disadvantageous because wood is certainly a more delicate material compared to UPVC and aluminium, as it requires regular maintenance, and because the price of wooden windows and doors is still higher than that of aluminium and PVC.
There are several species of wood, belonging to the broadleaf and conifer families, used for the construction of windows and doors:

- chestnut,
- fir,
- pine.

Besides a first hypothesis was to consider the material as a homogeneous and isotropic, whose behaviour has been hypothesized linear elastic until breakage. Therefore, the characteristics of the two wood species (Tab. 4, Tab. 5).

Table 4 Physical and mechanical properties of Pine	
Density	532 [kg/m^3]
Elastic Modulus	13700 [MPa]
Tensile Strength	85 [MPa]
Compressive Strength	45 [MPa]
Poisson’s ratio	0.33

Table 5 Physical and mechanical properties of Chesnut	
Density	630[kg/m^3]
Elastic Modulus	114000 [MPa]
Tensile Strength	95 [MPa]
Compressive Strength	51 [MPa]
Poisson's ratio	0.30

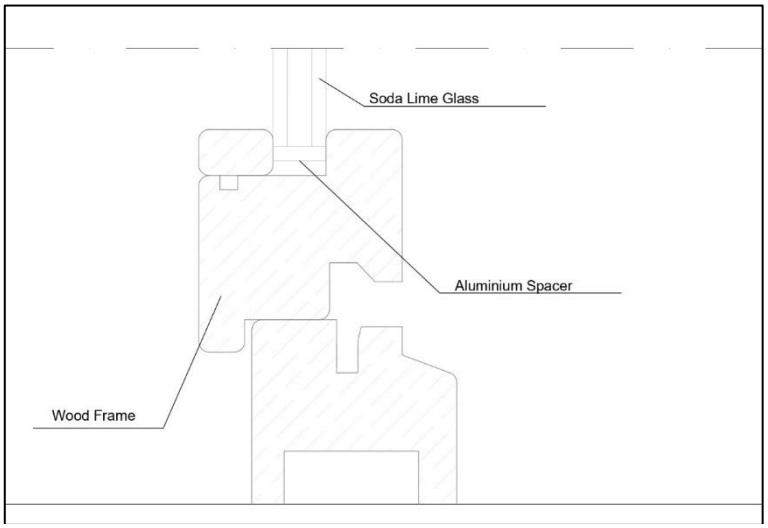


Figure 4 Wood Windows.

3.4. UPVC WINDOWS

Another common material to produce windows and doors is UPVC. UPVC is a material composed of macromolecules which are in turn formed by hydrogen, carbon and chlorine atoms. A UPVC window and door frame has very similar characteristics to those of wood. UPVC has a low mechanical resistance; to overcome this, multi-chamber profiles are extruded and reinforced with the help of metal elements (Fig.5). These windows and doors have the main characteristic of resisting very well to the aggressions of atmospheric agents, are very light and offer good thermal insulation (Mottura and Pennisi, 2014). UPVC is a thermoplastic so its mechanical characteristics (Tab.6) are highly dependent on the glass transition temperature (T_G); that is the temperature, below which the physical properties of plastics change to those of a glassy or crystalline state. Above T_g they behave like rubbery materials. Below the T_g a plastic's molecules have relatively little mobility (Ebnesajjad, 2016) so the material behaves as a fragile and rigid material. UPVC has a glass transition temperature of 80°C. Thus, considering the temperatures expected in the red areas, UPVC windows and doors are not suitable to withstand the pressures applied by the pyroclastic flows.

Table 6 Physical and mechanical properties of UPVC	
Density	1400 [kg/m^3]
Elastic Modulus	3700 [MPa]
Tensile Strength	47 [MPa]
Coefficient of thermal expansion	0,8e^{-4} [1/K]
Poisson's ratio	0.40

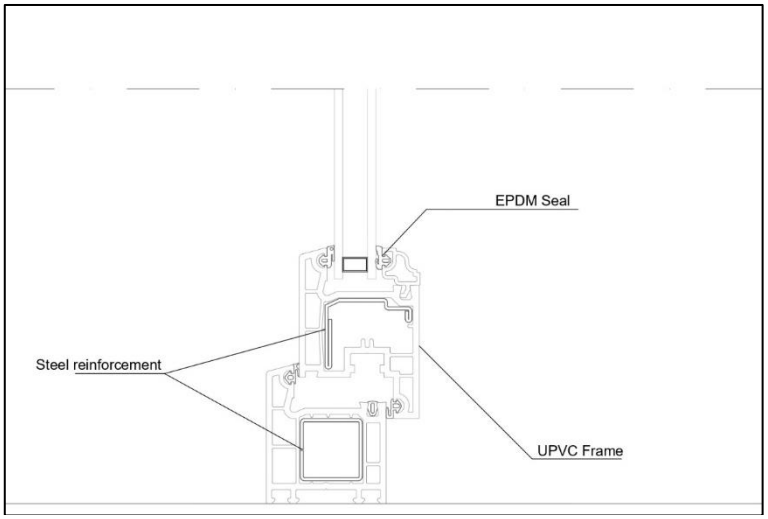


Figure 5. UPVC windows.

4. Vulnerability Assessment

To assess the vulnerability of the frames to the pressures expected in both areas, a two-dimensional stationary linear static model (2) has been set up for both aluminium and wooden frames.

$$0 = \nabla \cdot S + F_V \tag{2}$$

where:

- $\nabla \cdot S$ is stress tensor (Pa),
- F is the applied force (N/mm)

Since the resistance of openings to dynamic pressure depends on several factors of which the most important are the size; therefore, three different heights have been considered for each group of windows, i.e. for the large openings a height of 2.4 m had been considered, for the Typical a height of 1.2 m and for the small a height of 0.8 m has been considered; and for each window size the different thicknesses of the glass have been considered. Furthermore, it has been considered a fixed constraint at the base of the wall on which the window is placed (Fig.3). Additionally, it has been considered half section in order to reduce the computational time, assuming a condition of symmetry in the upper part of the window.

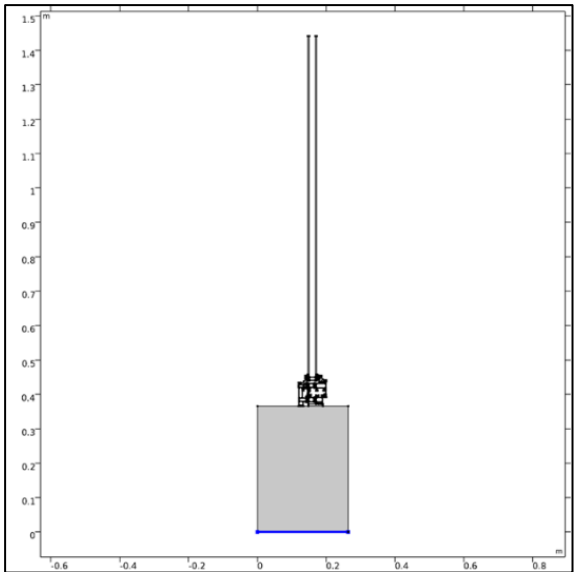


Figure 6. Fixed constraint condition-

Finally, a uniformly distributed load applied (3) on the external front has been assumed, in favor of opening, which is linearly variable according to a parameter that has been imposed through a range function.

$$\boldsymbol{S} \cdot \mathbf{n} = F_A \tag{3}$$

$$\boldsymbol{F_A} = \frac{F_L}{d} \tag{4}$$

From these first mechanical analyses, the glass is the first element of the fixed system to fail (Tab.7). In particular, the glass of 4 mm of large dimensions, therefore with dimensions equal to 2.4 m of height and 0.6 of width, is the most vulnerable because the calculated breaking pressure is equal to 0.6 kPa. This situation is not entirely similar for wooden frames, as the glass of this type is placed inside the frame without the aid of gaskets, so the glass is perfectly embedded in the frame itself. Although the glass in the case of wooden frames may be suitable to withstand the expected pressures, the problem lies in the resistance of the glass to temperature variation.

Table 7 Glass breakage load aluminium windows

Frame	**Glass**								
Aluminium	SODA LIME GLASS								
	Large			Typical			Small		
	4mm	5mm	6mm	4mm	5mm	6mm	4mm	5mm	6mm
	0,6 kPa	1 kPa	1,3 kPa	2 kPa	3 kPa	4,3 kPa	5,3 kPa	8 kPa	10 kPa

Table 8 Glass breakage load timber windows									
Frame	**Glass**								
Wood	SODA LIME GLASS								
	Large			Typical			Small		
	4mm	5mm	6mm	4mm	5mm	6mm	4mm	5mm	6mm
	2 kPa	2.3 kPa	3.3 kPa	4.6 kPa	6.6 kPa	9.6 kPa	8.6 kPa	14.6 kPa	22.6 kPa

4.1. MITIGATION STRATEGY

An initial mitigation strategy, suitable for making glass capable of withstanding high pressures, could be the use of laminated glass technology.
Laminated glass for architectural glazing applications consists of two layers of glass bonded to a thin thermoplastic interlayer. The interlayer is responsible for keeping the glass fragments together after breaking, while increasing the residual strength of the material. There are different types of interlayer depending on the material used, the most used is Polyvinyl butyral (PVB) even though it has low mechanical qualities. Currently a new type of interlayer produced by Dupont® is also used, the
Sentry Glass Plus (SGP) which, unlike the previous one, presents a resistance relatively high. This hypothesis of mitigation is not suitable for the pyroclastic flow, since these interlayers are thermoplastic, and their behavior depends on the temperature of exposure. Indeed, above the 80°C the PVB film starts to separate from the glass (delamination); besides, the SGP structural interlayer is also characterized by a melting temperature of 94°C.

5. Thermal Analysis

5.1. THERMAL SHOCK

Once the mechanical vulnerability had been defined, a first thermal analysis of the transparent elements was carried out, in particular the exposure time at three different temperatures (100/200/300) °C was calculated, until the critical temperature that causes the glass to break due to the thermal shock, considering both the three different thicknesses and the three different dimensions previously defined. Thermal shock occurs when a thermal gradient causes different parts of an object to expand in different quantities. This differential expansion can also be understood in terms of stress or deformation (5) At some point, this stress may exceed the strength of the material, causing a crack to form. If nothing prevents this crack from propagating through the material, the glazing will lose its structural integrity. Glass objects are particularly vulnerable to failure due to thermal shock, due to their low strength and low thermal conductivity. If the glass is then suddenly exposed to extreme heat, the shock will cause the glass to break.

$$\Delta T = \frac{(\sigma_{TS} * (1-\upsilon))}{E * \alpha} \tag{5}$$

where:

- σ_{TS} is the yield strength of the material,
- ν Poisson's ratio,
- E Elastic modulus,

- α coefficient of thermal expansions.

In the case of soda lime glass, the critical temperature is 52 °C.
Once the critical temperature is defined, it has been necessary to assess the time to reach it through the heat transfer equation (6)

$$d_z \rho C_p \frac{\delta T}{\delta t} + d_z \rho C_p \mathbf{u} \cdot \nabla T + \nabla \cdot q = d_z Q \quad (6)$$

where:

- ρ is the density (kg/m^3),
- Cp is the specific heat (J/(kg*K)),
- T is the temperature (K),
- u is the speed vector of motion (m/s),
- Q is the heat source (W/m^3),
- dz is thickness of domain in the out-of-plane direction (m),
- q is the conduction heat flow (W/m^2).

To model this problem properly, a time-dependent study was used in the 60 second interval using the range function (0.1.60) s. Moreover, to model the sudden temperature rise, a ramp function (Fig.7) was applied in the Temperature node, using the following expression:

$$20 + x * ramp(t) \quad (7)$$

where variable x is equal respectively to 80, 180, 280 and 380°C so that the glass is subject to three different maximum temperatures: 100°C, 200°C and 300°C.

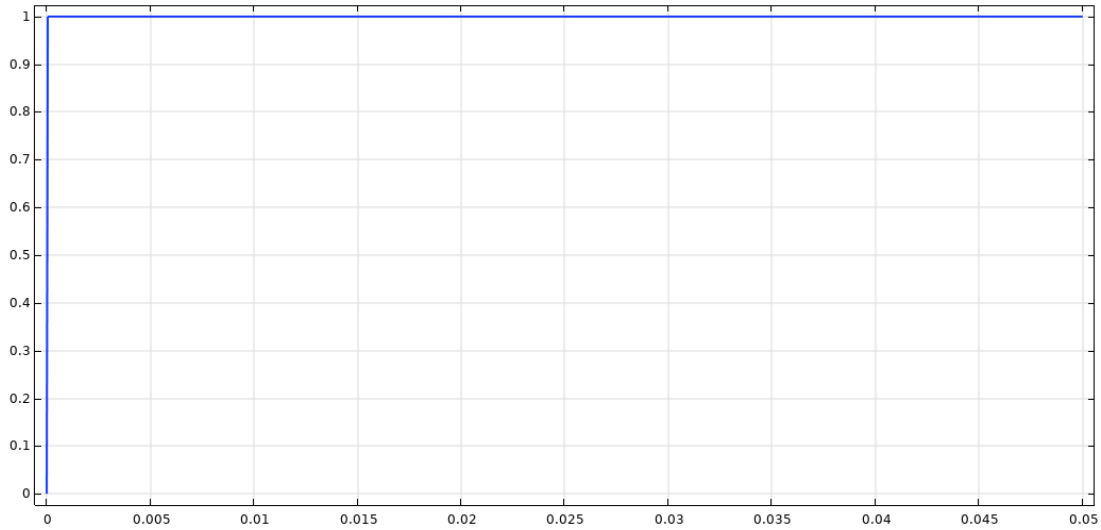


Figure 7. Ramp Function.

From this kind of analysis considering an applied temperature of 100 °C, soda-lime glass reaches the critical temperature in the 5-second time interval. (Fig 8). Therefore, soda-lime glass is totally vulnerable to the temperatures expected in the red areas.

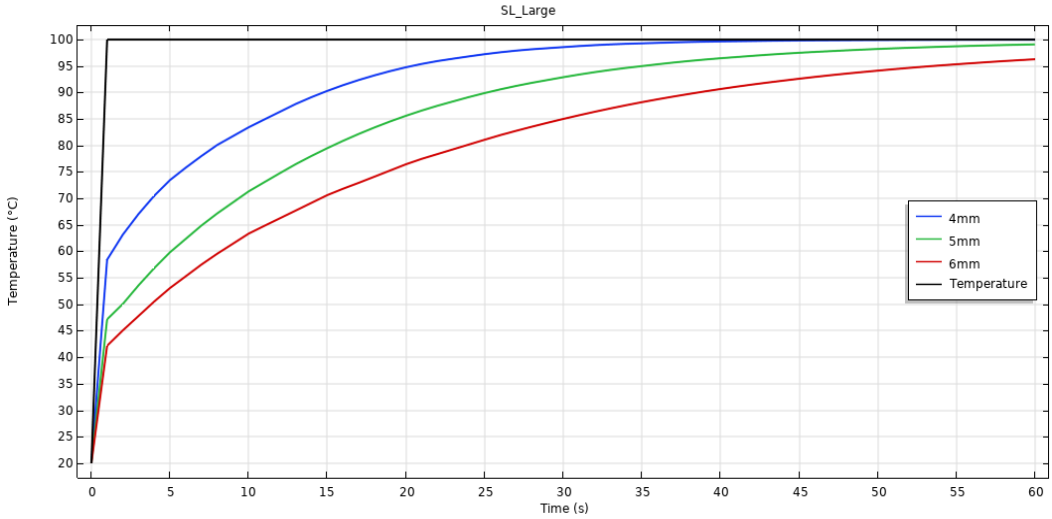


Figure 8. Heat Transfer for large glass pane.

5.2. Aluminium Resistance

In order to analyse the resistance of the aluminium frame, the stress due to the agent temperature (8) was calculated, which was applied, for a time interval of 1200 s, through the ramp function (6). The first results show that the aluminium can withstand, as there is no breaking-strength of 160 MPa, for the expected temperatures. On the other hand, the time of reaching the glass transition temperature of the polyamide that constitutes the thermal break of the frame was also evaluated, since once this temperature of 50°C has been reached, the material softens and therefore the solution of continuity between the two aluminium sections constituting the frame and counter frame is lost and therefore the breakage of the frame. In the case of the aluminium frame with thermal break, the polyamide reaches this temperature in a range of 490s in the case of 100°C, while in the case of 200°C and 300°C the time is reduced to 50s and 20s respectively (Fig.10).

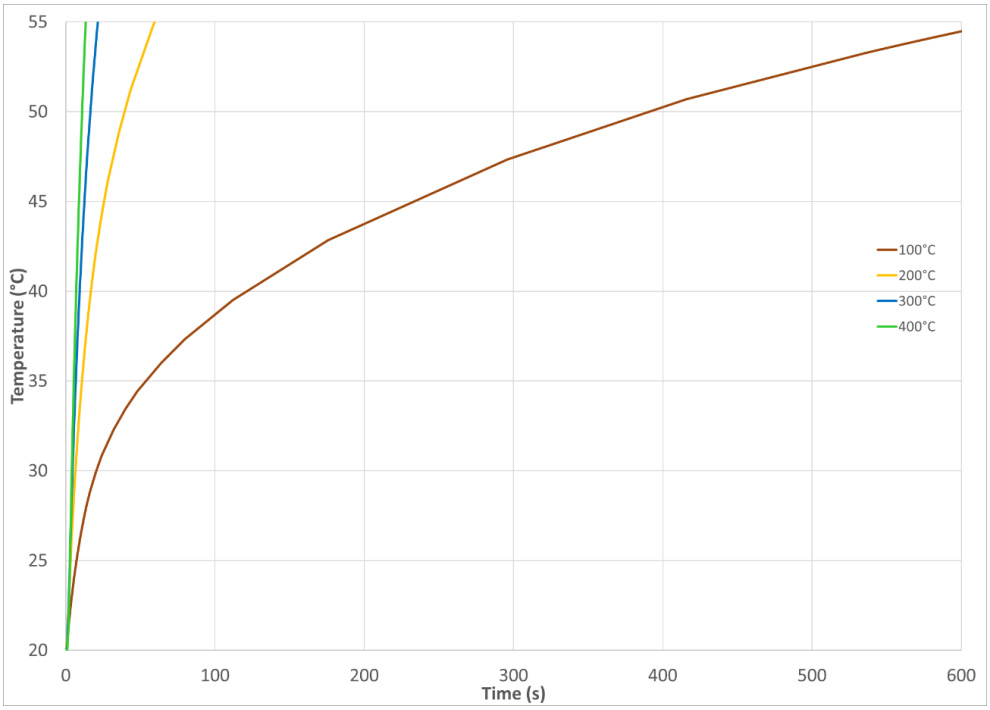


Figure 9. Temperature of Polyamide of Unprotected Section.

$$\varepsilon_{th} = \alpha\,(T)\,(T - T_{ref}) \tag{8}$$

5.3. Mitigation Strategy

A mitigation strategy for this problem is to replace glass with one that has good thermal shock resistance, i.e. tempered glass. This is a type of safety glass treated with controlled thermal or chemical processes to increase its resistance compared to normal glass. When (5) is applied, the critical temperature calculated, considering the characteristics of such glass (Table 9), is 355 °C. Therefore, this option could be suitable for urban settlements present at distances very far from the vent, e.g. in the case of the Vesuvius area it is about 7-8 km. While to resolve the issue of the polyamide it has been hypothesized to apply the wood-aluminium frame technology in reverse. In fact, it has been placed on the external front of the fixed and mobile frame of the pine wood sections with a thickness of 2 cm with the function of shielding (Fig.10). From the thermal analysis carried out (5), an average temperature of 44°C is recorded for the polyamide elements, considering a time interval of 1200s, only in the case of an agent temperature of 300°C (Fig.11) and therefore the continuity of the sections is preserved.

Table 9 Physical and mechanical properties of Tempered Glass	
Density	2400 [kg/m^3]
Elastic Modulus	70000 [MPa]
Tensile Strength	180 [MPa]
Coefficient of thermal expansion	9 e^{-6} [1/K]
Poisson's ratio	0.23

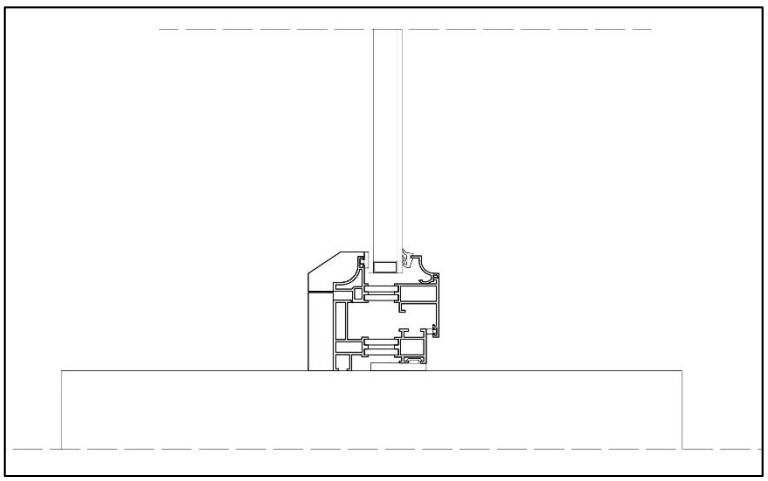


Figure 10. Hypothetical aluminum - wood section.

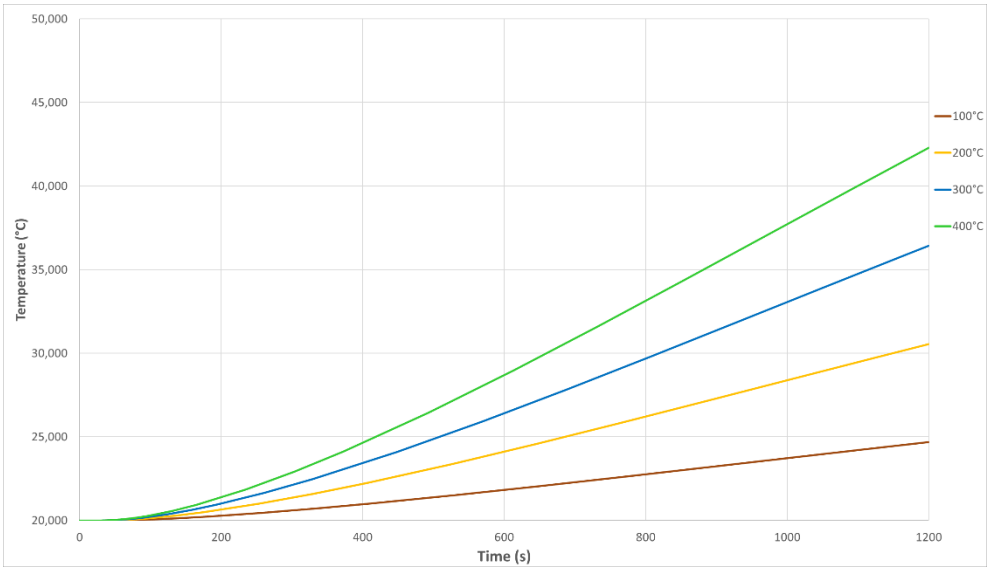


Figure 11. Temperature of Polyamide of Protected Section.

6. Conclusion

The research on mitigation technologies for the reduction of volcanic risk to buildings represents a key field of investigation that is directly connected with the study of volcanic hazard and emergency management. In this context, the definition of possible interventions to reduce the expected damage due to possible eruptions represents a key issue. Through this study, it has been demonstrated that it is not only glass that represents an element of the weakness of the framed system, in fact considering the thermal problem also the other constituent elements such as gaskets and thermal break are a problem to be considered as thermoplastics. Moreover, from these first analyses, it is possible to state that in order to solve the mechanical problem, instead of using laminated glass technology, it is necessary to protect the window frame, the shutters could accomplish this task, and in this case, it is necessary to carry out fluid-dynamic evaluations to identify any infiltration problems or through a sandwich panel, which can withstand both to the expected pressure and the expected temperature; meanwhile, this solution can satisfy the problem of energy saving. Even though the next event is expected to have an intensity lower than a sub-Plinian I, and in this case all the analyses are here designed for a sub-Plinian eruption, would respond more effectively to eruptive phenomena. Nevertheless, in the case of Vesuvius and Campi Flegrei, the main obstacle to the implementation of such measures is the extent of the potentially affected areas, which raises important issues about the economic sustainability of mitigation interventions. Local authorities and private citizens need to be fully aware of the potential damages following an eruption and of the cost-effectiveness of possible mitigation interventions. Considering the economic, political, and social "weight" of volcanic risk in densely populated areas, a valid evaluation method of the effectiveness of mitigation solutions can give scientific support to strategic choices and emergency plans. The approach discussed in the present paper, although about the Vesuvius and Campi Flegrei areas in the specific parameters analyzed (hazard characterization, building typologies, construction technologies), represents a methodology that could also be adopted in other contexts. Further work is being developed to understand the behaviour of shutters considering a fluid dynamic problem and a 3D model of the window to analyze the contribution of further elements such as the hinges and the actual locking system, finally a model in which the mechanical and thermal problem is analyzed simultaneously.

References

Baxter, P. J., et al. (2008). "Emergency planning and mitigation at Vesuvius: A new evidence-based approach." J. Volcanol. Geotherm. Res., 178(3), 454–473.

Baxter, P. J., Boyd, R., Cole, P., Neri, A., Spence, R. J. S., and Zuccaro, G. (2005). "The impacts of pyroclastic surges on buildings at the eruption of the Soufrière Hills volcano,Montserrat." Bull. Volcanol., 67(4), 292–313.

Baxter, P. J., Woo, G., and Pomonis, A. (1998). "Preliminary assessment of volcanic risk on Montserrat." Montserrat Volcano Observatory, Montserrat, West Indies.

Carrot, C., Bendaoud, A., Pillon. C., (2015). *Polyvinyl Butyral from:* Handbook of Thermoplastics CRC Press

Convertito, V., and Zollo, A. (2011). "Assessment of pre-crisis and syncrisis seismic hazard at Campi Flegrei and Mt. Vesuvius volcanoes, Campania, southern Italy." Bull. Volcanol., 73(6), 767–783.

Ebnesajjad, S. (2016) "Expanded PTFE applications handbook: Technology, manufacturing and applications." William Andrew Publishing.

Esposti Ongaro, T., Neri, A., Todesco, M., and Macedonio, G. (2001). "Pyroclastic flow hazard assessment at Vesuvius by using numerical modelling. 2: Analysis of flow variables". *Bull. Volcanol., 64, 78-191.*

Gurioli L., Sulpizio R., Cioni R., Sbrana A., Santacroce R., Luperini W., Andronico D., 2010, Pyroclastic flow hazard assessment at Somma -Vesuvius based on the geological record. Bull Volcanol 72:1021-1038, Springer-Verlag.

International Volcanic Health Hazard Network (IVHHN), Cities and Volcanoes Commission (CaV), GNS Science, United States Geological Survey (USGS). (2009). Guidelines on preparedness before, during
and after an ashfall, IVHHN, Durham, UK.

Mottura, G., and Pennisi, A. (2014). “Serramenti e schermature solari”, Maggioli Editore, ISBN: 978-8891600929

Neri, A., et al. (2008). “Developing an event tree for probabilistic hazard and risk assessment at Vesuvius.” J. Volcanol. Geotherm. Res., 178(3), 397–415.

Neri, A., Esposti Ongaro, T., Macedonio, G., and Gidaspow, D. (2003). “Multiparticle simulation of collapsing volcanic columns and pyroclastic flows.” J. Geophys. Res., 108(B4), 2202–2225.

Petrazzuoli, S. M., and Zuccaro, G. (2004). “Structural resistance of reinforced concrete buildings under pyroclastic flows: A study of the Vesuvian area.” J. Volcanol. Geotherm. Res., 133(1-4), 353–364.

Spence R.J.S., Baxter P.J., Zuccaro G. (2004a). Building vulnerability and human casualty estimation for a pyroclastic flow: a model and its application to Vesuvius. *J. Volcanol. Geoth. Res.* 133 (2004):321–343, 2004.

Spence R., Zuccaro G., Petrazzuoli S., Baxter P.J., (2004b). The resistance of buildings to pyroclastic flows: theoretical and experimental studies in relation to Vesuvius. *ASCE Nat Hazards* Rev 5:48–50. ISSN: 1527-6988, 2004.

Todesco, M., and Longo, A. (2002). “Pyroclastic flow hazard assessment at Vesuvius by usinf numerical modelling. 1: Large scale dynamics”. *Bull. Volcanol., 64, 155-177.*

Zuccaro G (2004). Final report SAVE. Project Link to the Web Page of the ‘‘Save’’ Project in ‘‘Atlantis’’. http://gndt.ingv.it/Att_scient/Prodotti_attesi_2004/Dolce_Zuccaro/Mappe/ Start2.htm. Roma

Zuccaro G., Albanese V., Cacace F., Mercuri C. and Papa F., (2008a). Seismic Vulnerability Evaluations within the Structural and Functional Survey Activities of the COM bases in Italy, AIP Conference, 2008.

Zuccaro G., Cacace F., Spence R.J.S., Baxter P.J., (2008b). Impact of explosive eruption scenarios at Vesuvius", J Volcanol Geoth Res 178(2008):416–453, 2008.

Zuccaro G. and Cacace F. (2010). Seismic impact scenarios in the volcanic areas in Campania. In: F.M. Mazzolani. Urban habitat construction under catastrophic events. p. 725-730, CRC Press, ISBN: 9780415606851.

Zuccaro G., De Gregorio D. (2011). Valutazioni di impatto in aree vulcaniche - Vesuvio e Campi Flegrei ", In: Lirer L., Il vulcanismo esplosivo, p. 56-73, CNR press, ISBN: 979-12-200-1041-2

Zuccaro G., Cacace F., De Gregorio D. (2012). Buildings inventory for seismic vulnerability assessment at National and regional scale", In: 15th World Conference on Earthquake Engineering, 15th WCEE 2012, Lisbon (Portugal), 24–28 September 2012. Paper no 2829, 2012.

Zuccaro G. and De Gregorio D., (2013). Time and Space dependency in impact damage evaluation of a sub-Plinian eruption at Mount Vesuvius. Natural Hazard (2013) 68:1399–1423. DOI: 10.1007/s11069-013-0571-8, 2013.

Zuccaro G., Leone M. F., Del Cogliano D., Sgroi A. (2013). Economic impact of explosive volcanic eruptions: A simulation-based assessment model applied to Campania region volcanoes. JOURNAL OF VOLCANOLOGY AND GEOTHERMAL RESEARCH, vol. 266, p. 1-15, ISSN: 0377-0273, doi: 10.1016/j.jvolgeores.2013.09.002.

Zuccaro G., De Gregorio D., Baxter P. (2014). Human and Structural Vulnerability to Volcanic Processes. In: P. Papale. (a cura di): J. Shroder, Volcanic Hazards, Risks and Disasters. vol. Chapter 10, p. 261-288, Elsevier, ISBN: 9780123964533.

Towards a multiphase finite element code for the evaluation of volcanic risk

Andrea Montanino[1], Alessandro Franci[2], Giulio Zuccaro[1]

[1] *Department of Structures for engineering and Architecture (DiST)*
University of Naples "Federico II" - Naples - ITALY
{*andrea.montanino, zuccaro*}*@unina.it*

[2] *International Centre for Numerical Methods in Engineering*
Universitat Politècnica de Catalunya - Barcelona - SPAIN
falessndro@cimne.upc.edu

Abstract. The evaluation of volcanic risk is of great relevance for the institutions of Civil Protection, which are required to design emergency plans and to take all the necessary countermeasures against this type of catastrophic natural hazard. The assessment of volcanic risk can be performed basing on the observation and statistical interpretation of historical data, or using numerical instruments for the analysis of possible scenarios. In this work, we focus on the latter strategy and develop a novel Finite Element formulation for the description of multiphase flows, with the final aim of simulating pyroclastic density currents at the scale of urban settlements. Pyroclastic flows are here modeled numerically by solving the conservation laws of mass, linear momentum, and energy, under the hypotheses of kinetic and thermal equilibrium. In order to avoid unphysical oscillations of solution variables, the Finite Element formulation is stabilized with a recent implementation of the Variational Multiscale method for multiphase flows. Benchmark problems in the field of compressible monophase and multiphase flows are here analyzed to prove accuracy and robustness of the proposed formulation. A realistic scale case of pyroclastic gravity current hitting some buildings is also presented. The test shows the potential of the proposed approach to evaluate quantities of interest for the evaluation of the hazard deriving from volcanic events, such as the generated pressure on civil constructions.

Keywords: Volcanic risk assessment, multiphase flows, Variational Multiscale method

1. Introduction

The evaluation of the volcanic risk is of primary importance for the departments of Civil Protection, since on these predictions are based the design of the emergency plans (Baxter et al., 2005) and mitigation strategies (Zuccaro and De Gregorio, 2019), including the systems of building protection, such as panels and resistant fixtures (Zuccaro and Leone, 2012; Iacuaniello et al., 2020).

For the evaluation of the volcanic risk, it is crucial to have an estimation of those quantities that could be dangerous for urban settlements, such as the pressure on buildings and the temperature of the pyroclastic flow. This could be done basing on observations, historical reports (Zuccaro et al., 2008), and/or numerical simulations (Neri et al., 2003; Esposti Ongaro et al., 2007).

© 2021 *by A. Montanino, A. Franci, G. Zuccaro.*

The aim of this work is to propose a reliable numerical model for the prediction of the effect of this natural hazard on civil constructions. In particular, the objective is to evaluate the pressure and temperature fields acting on civil buildings for different geometrical characteristics of the urban settlements and intensities of the volcanic eruption event.

The pyroclastic flow is here modeled by solving the *mixture* equations of conservation of the mass, linear momentum, and energy, following the formulation proposed by Pelanti and LeVeque (2006) and developed for the case of thermal and kinetic equilibrium by Cerminara et al. (2016). Such formulation holds at a distance from the volcano source that depends on the characteristics of the pyroclastic flow.

The resultant system of equations has the same nature as that of monophase compressible flows. It is well-known (Bochev et al., 2004; Tezduyar and Sathe, 2003) that the Eulerian Finite Element methods suffer from numerical instability when applied to advection-diffusion problems with a dominant convective term. For this reason, the numerical method needs to be stabilized. Among different stabilizing tools proposed in the literature, here we adopt the Variational Multiscale Method (Hughes et al., 1998; Bazilevs et al., 2007; Franca et al., 1992). The basis of this stabilization strategy lays in the description of the unknown variables as the sum of a part that can be represented at the scale of finite elements and in a remainder (subgrid scale) that cannot be approximated by finite element shape functions. This stabilization technique is well-assessed in the framework of monophase compressible flows (Koobus and Farhat, 2004; Rispoli and Saavedra, 2006; Bayona Roa et al., 2016), whereas it has been much less used for multi-phase compressible cases (Montanino et al., 2021).

To show the applicability of the proposed formulation to the simulation of pyroclastic gravity currents, we present the solution of well-known benchmark problems for monophase and multi-phase flows, showing accuracy and convergent behavior of the method. Furthermore, the impact of pyroclastic flow on an urban settlement is also analyzed considering realistic flow characteristics, material properties, and scale of the problem.

The work is structured as follows: in Section 2 we present the governing equations of the problem and introduce the hypotheses of kinetic and thermal equilibrium; in Section 3 we derive the weak form of the governing equations and introduce the VMS stabilization; in Section 4 we present the application of the method to some benchmark problems, and finally, in Section 5, we draw some conclusions and discuss the future perspectives for this work.

2. Governing equations

The flow of a dilute multiphase flow (with solid concentration $\varepsilon_s < 10^{-2}$) is described by the equations of conservation of mass, linear momentum, and total energy of each phase (Pelanti and LeVeque, 2006)

$$\frac{\partial}{\partial t}(\varepsilon_g \rho_g) + \frac{\partial}{\partial x_i}(\varepsilon_g \rho_g u_{g,i}) = 0 \tag{1a}$$

$$\frac{\partial}{\partial t}(\varepsilon_g \rho_g u_{g,i}) + \frac{\partial}{\partial x_j}(\varepsilon_g \rho_g u_{g,i} u_{g,j} + p) = \frac{\partial \tau_{g,ij}}{\partial x_j} + \varepsilon_g \rho_g b_{g,i} - D(u_{g,i} - u_{s,i}) \tag{1b}$$

$$\frac{\partial}{\partial t}(\varepsilon_g\rho_g e_g) + \frac{\partial}{\partial x_i}((\varepsilon_g\rho_g e_g + p)u_{g,i}) = u_{g,i}\frac{\partial}{\partial x_j}\tau_{g,ij} - \frac{\partial}{\partial x_i}q_{g,i} + \varepsilon_g\rho_g b_{g,i}u_{g,i} + \varepsilon_g\rho_g r_g - D(u_{g,i} - u_{s,i})u_{g,i} - Q(T_g - T_s) \tag{1c}$$

$$\frac{\partial}{\partial t}(\varepsilon_s\rho_s) + \frac{\partial}{\partial x_i}(\varepsilon_s\rho_s u_{s,i}) = 0 \tag{1d}$$

$$\frac{\partial}{\partial t}(\varepsilon_s\rho_s u_{s,i}) + \frac{\partial}{\partial x_j}(\varepsilon_s\rho_s u_{s,i}u_{s,j}) = \frac{\partial\tau_{s,ij}}{\partial x_j} + \varepsilon_s\rho_s b_{s,i} + D(u_{g,i} - u_{s,i}) \tag{1e}$$

$$\frac{\partial}{\partial t}(\varepsilon_s\rho_s e_s) + \frac{\partial}{\partial x_i}(\varepsilon_s\rho_s e_s u_{s,i}) = u_{s,i}\frac{\partial}{\partial x_j}\tau_{s,ij} - \frac{\partial}{\partial x_i}q_{s,i} + \varepsilon_s\rho_s b_{s,i}u_{s,i} + \varepsilon_s\rho_s r_s + D(u_{g,i} - u_{s,i})u_{s,i} + Q(T_g - T_s) \tag{1f}$$

where ε_g and ε_s are the volumetric fractions of the gas and the solid phases being $\varepsilon_g + \varepsilon_s = 1$, ρ_g and ρ_s are the densities, $\boldsymbol{u}_g$ and $\boldsymbol{u}_s$ are the velocity fields, e_g and e_s are the specific energies, where in particular

$$e_g = c_v T_g + \frac{1}{2}u_{g,i}u_{g,i} \quad \text{and} \quad e_s = c_s T_s + \frac{1}{2}u_{s,i}u_{s,i} \tag{2}$$

where c_v and c_s are the specific heat at constant volume of the gas, and the one of the solid material, respectively, T_g and T_s are the gas and solid temperatures. Moreover, $p = \varepsilon_s p_g$ is the total pressure, being p_g the pressure of the gas phase; $\boldsymbol{b}_g$ and $\boldsymbol{b}_s$ are body forces of the two phases, and r_g and r_s are heat sources. Finally, $\boldsymbol{\tau}_g$ and $\boldsymbol{\tau}_s$ are the stress tensors of each mixture phase, analogously $\boldsymbol{q}_g$ and $\boldsymbol{q}_s$ are the heat flows of each phase. The momentum and energy exchanges between the solid and the gas phases are modeled through the terms $Q(T_g - T_s)$ and $D(\boldsymbol{u}_g - \boldsymbol{u}_s)$, that vanish when the velocity and temperature fields of the two phases are the same.

The formulation is completed by the equations of state. For the gas phase, this reads

$$p_g = \rho_g R T_g \tag{3}$$

being R the specific constant of the gas; for the solid phase, one must simply state the incompressibility of the solid material, *i.e.* $\rho_s = const.$

When a solid particle of radius σ and mass m enters in a gas flow with dynamic viscosity μ_g, the difference of velocities between the particle and the flow reduces and tends asymptotically to zero (Marble, 1970). This process is exhausted in a characteristic time

$$\tau_U = \frac{m}{6\pi\sigma\mu_g} \tag{4}$$

A similar phenomenon occurs for the temperature difference, that tends to zero after the following characteristic time

$$\tau_T = \frac{mc_p}{4\pi\sigma\lambda} \tag{5}$$

where c_p is the specific heat of the mixture and λ is the thermal diffusivity of the gas phase.

It follows that, after a time t from the volcanic event such that $t > \tau_U$ and $t > \tau_T$, the temperature and velocity fields of the two phases are uniform, therefore we can set $\boldsymbol{u}_g = \boldsymbol{u}_s = \boldsymbol{u}$, and $T_g = T_s = T$. This condition is known as the hypothesis of thermal and kinetic equilibrium (Cerminara et al., 2016). Using this assumption and summing Equations (1a) and (1d), Equations (1b) and (1e), and Equations (1c) and (1f), we obtain

$$\frac{\partial}{\partial t}\rho + \frac{\partial}{\partial x_i}(\rho u_i) = 0 \tag{6a}$$

$$\frac{\partial}{\partial t}(\rho u_i) + \frac{\partial}{\partial x_j}(\rho u_i u_j + p) = \frac{\partial \tau_{ij}}{\partial x_j} + \rho b_i \tag{6b}$$

$$\frac{\partial}{\partial t}E + \frac{\partial}{\partial x_i}((E+p)u_i) = u_i \frac{\partial}{\partial x_j}\tau_{ij} - \frac{\partial}{\partial x_i}q_i + \rho b_i u_i + \rho r \tag{6c}$$

We also retain Equation (1d) to control the conservation of the solid phase mass, and we rewrite it in terms of the new uniform velocity field as

$$\frac{\partial}{\partial t}(\varepsilon_s \rho_s) + \frac{\partial}{\partial x_i}(\varepsilon_s \rho_s u_i) = 0 \tag{7}$$

In Equation (6) $\rho = \varepsilon_g \rho_g + \varepsilon_s \rho_s$ is the total density of the mixture, and $E = \varepsilon_g \rho_g e_g + \varepsilon_s \rho_s e_s$ is the total mixture energy; the stress tensor and the heat flow vector are written as linear terms with a new mixture viscosity μ_m and mixture thermal diffusivity λ_m, following (Marble, 1970), as

$$\mu_m = \frac{\mu_g}{1+\kappa} \qquad \text{and} \qquad \lambda_m = \frac{\lambda_g}{1+\kappa} \tag{8}$$

being $\kappa = \varepsilon_s \rho_s / {}_g \rho_g$.

3. Weak form and stabilization

In this section, we introduce the weak form of Problem (6)-(7). First, for convenience, we reformulate the problem in terms of a set of *conservative* variables as

$$\frac{\partial}{\partial t}U_i + \frac{\partial}{\partial x_j}F_{ij} = \frac{\partial}{\partial x_i}G_{ij} + S_{ij}U_j \tag{9}$$

where $\boldsymbol{U}$ is the vectcor of conservative variables, $\boldsymbol{F}$ and $\boldsymbol{G}$ represent the convective and the dissipative flux, respectively, and $\boldsymbol{S}$ is the matrix of the source terms.

For the case of multiphase flows, the set of conservative variables is

$$\boldsymbol{U} = [\rho, P_s, \boldsymbol{M}, E]^T \tag{10}$$

where $P_s = \varepsilon_s \rho_s$ is the contribution of the solid phase to the total mixture density, and $\boldsymbol{M} = \rho \boldsymbol{u}$ is the total linear momentum of the mixture; consequently, the convective and diffusive fluxes are

$$\mathbf{F} = \begin{bmatrix} M_j \\ \dfrac{P_s}{\rho} M_j \\ M_i \dfrac{M_j}{\rho} + p\delta_{ij} \\ \dfrac{M_j}{\rho}(E+p) \end{bmatrix}, \qquad \mathbf{G} = \begin{bmatrix} 0_j \\ 0_j \\ \tau_{ij} \\ u_i \tau_{ij} + q_j \end{bmatrix} \tag{11}$$

where i, j range from 1 to d being d the space dimensions, and δ_{ij} is the Kronecker delta. Moreover, the pressure p can be expressed in terms of the conservative variable, using the equation of state and the definition of the total energy.

Finally, the source term matrix is

$$\mathbf{S} = \begin{bmatrix} 0 & 0 & \mathbf{0}^T & 0 \\ 0 & 0 & \mathbf{0}^T & 0 \\ \mathbf{b} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ r & 0 & \mathbf{b}^T & 0 \end{bmatrix} \tag{12}$$

We obtain the weak form of Equation (9) by pre-multiplying it by a test function vector $\boldsymbol{V}$ and integrating over the domain Ω as

$$\left(V_i, \frac{\partial}{\partial t} U_i\right) + \left(V_i, A_{ijk} \frac{\partial}{\partial x_j} U_k\right) = \left(V_i, \frac{\partial}{\partial x_j} G_{ij}\right) + \left(V_i, S_{ij} U_j\right) \tag{13}$$

which can be rewritten integrating by part the diffusive term, and incorporating the source term into the non linear, convective term

$$\left(V_i, \frac{\partial}{\partial t} U_i\right) + \left(V_i, A_{ijk} \frac{\partial}{\partial x_j} U_k - S_{ij} U_j\right) + \left(\frac{\partial}{\partial x_j} V_i, G_{ij}\right) \tag{14}$$

being

$$A_{ijk} = \frac{\partial F_{ij}}{\partial U_k} \tag{15}$$

the Euler-Jacobian matrix.

A Finite Element discretization of Equation (14) leads to unphysical oscillations of variables, due to the dominance of the convective part. This kind of instability has been the object of diffuse literature (Codina, 2002, 2000) and has been tackled with different approaches. Here, we use the Variational Multiscale (VMS) technique (Hughes et al., 1998) as the stabilizing tool. In the VMS spirit, the vectors of conservative variables and the test function are decomposed into two parts, a first one ($\boldsymbol{U}^h$) that can be represented at the scale of Finite Elements, and the second one ($\tilde{\boldsymbol{U}}$) that is representative of the fluctuations occurring at a finer scale.

$$\boldsymbol{U} = \boldsymbol{U}^h + \tilde{\boldsymbol{U}}\,, \qquad \boldsymbol{V} = \boldsymbol{V}^h + \tilde{\boldsymbol{V}} \tag{16}$$

Following the formulation proposed in (Bayona Roa et al., 2016) for compressible monophase flows, and adapted in (Montanino et al., 2021) for multiphase flows, we discretize the following weak form

$$\left(V_i^h, \frac{\partial U_i}{\partial t} + L_{ik}U_k\right) - \left(L_k^*, \tau_{kj}R_j^h\right) + \left(\frac{\partial V_i^h}{\partial x_j}, G_{ij}(\mathbf{U}_h)\right) = 0 \tag{17}$$

where the second term is the stabilizing part, in accordance to the the prescription of the VMS. In Equation (17)

$$L_k^* = \frac{\partial V_i^h}{\partial x_j}A_{ijk} + V_i B_{ijkm}\frac{\partial U_m^h}{\partial x_j} - V_i S_{ik} \tag{18}$$

is the adjoint term, $\boldsymbol{R}^h$ is the residual of the Finite Element scale, and $\boldsymbol{\tau}$ is the diagonal matrix of the stabilizing terms, being

$$\boldsymbol{\tau}^{-1} = \begin{bmatrix} \tau_\rho^{-1} & 0 & 0 & 0 \\ 0 & \tau_\rho^{-1} & 0 & 0 \\ 0 & 0 & \tau_M^{-1} & 0 \\ 0 & 0 & 0 & \tau_E^{-1} \end{bmatrix} \tag{19}$$

with

$$\tau_\rho^{-1} = c_2\frac{u + c_m}{h}, \quad \tau_M^{-1} = c_1\frac{\nu}{h^2} + c_2\frac{u + c_m}{h}, \quad \tau_E^{-1} = c_1\frac{\lambda_m}{\rho c_{p.m} h^2} + c_2\frac{u + c}{h} \tag{20}$$

where c_1 and c_2 are algorithmic constant, set to $c_1 = 4$ and $c_2 = 2$; c_m is the mixture speed of sound (Pelanti and LeVeque, 2006), c_p is the mixture specific heat at constant pressure (Marble, 1970), and finally h is the characteristic length of the element.

The numerical solution of Equation (17) is got after discretization in space and time. In particular, linear finite elements are used for spatial discretization, and an explicit first-order Euler scheme is used as time integration scheme. For details about the discrete form of the problem we refer to (Montanino et al., 2021).

The overall implementation of the numerical formulation is done in Kratos Multiphysics (Dadvand et al., 2010), an open-source coding environment suited for parallelization using FEM.

4. Application

In this section, we apply the formulation developed above to two different cases. First, we analyze the Sod shock tube, a well-known benchmark case for compressible flows, to assess the accuracy of the formulation and discuss its convergent behavior. Then, the method is applied to the case of a pyroclastic current hitting two buildings. This test is approached considering realistic material and geometrical data.

For both cases, the parameters adopted are $\rho_s = 1800$ kg/m^3, $c_v = 722$ W/Kg K, $c_p = 1010$ W/Kg K, $c_s = 1300$ W/kg K.

4.1. Sod shock tube

The Sod tube is a common benchmark for testing the performances of different numerical methods for monophase compressible flows (Cremonesi and Frangi, 2016; Fu, 2019).

It consists of a two-dimensional rectangular domain with the length ($L = 1$ m) much higher than the height in order to enforce a monodimensional flow. A membrane separates the domain into two parts that are filled with a gas at different initial conditions. In the left part, the gas density is $\rho_g = 1$ kg/m^3 and its pressure $p = 1$ Pa, while, in the right side, the gas density is $\rho_g = 0.125$ kg/m^3 and its pressure is $p = 0.1$ Pa. The initial velocity for all the domain is $u = 0$ m/s. To reproduce the standard monophase case typically solved in the literature, we assume here that there is no solid concentration in the domain ($\varepsilon_s = 0$). At $t = 0$ s, the membrane is removed, and the jumps in pressure and density between the two parts of the domain give rise to a shock wave propagating toward the right side of the domain, and to a refracting wave moving toward the left side. Both waves move at the speed of sound, which depends on the characteristics of the gas. For the monophase case, the test has an analytical solution that can be obtained through an iterative procedure, such that presented in (Toro, 2013).

The velocity analytical solution at $t = 0.2$ s is represented in Figure 1 together with the numerical solutions obtained for different discretizations. The graph shows a good agreement between the numerical and the analytical solution, and a clear convergence to the analytical solution for reducing mesh sizes.

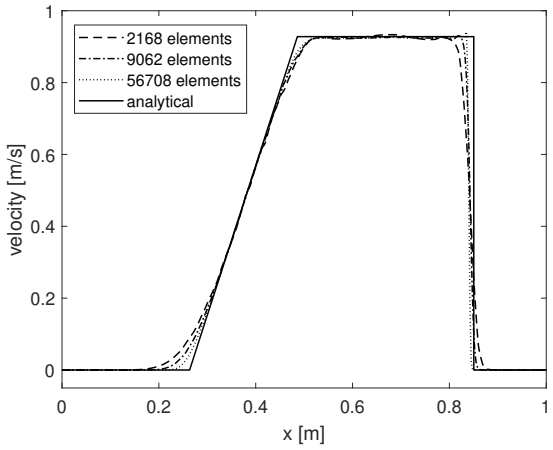


Figure 1. Monophase shock tube - velocity distribution at $t = 0.2$ s.

Here, we propose a modification of the standard monophase Sod tube problem to assess the accuracy of the multiphase formulation. In particular, we consider the same values for the gas density ρ_g, pressure p, and velocity as the monophase case, but we include in all the domain a

concentration of solid particles. Three different solid concentrations are considered in three different analyses, namely $\varepsilon_s = 10^{-5}$, $\varepsilon_s = 10^{-4}$, and $\varepsilon_s = 10^{-3}$.

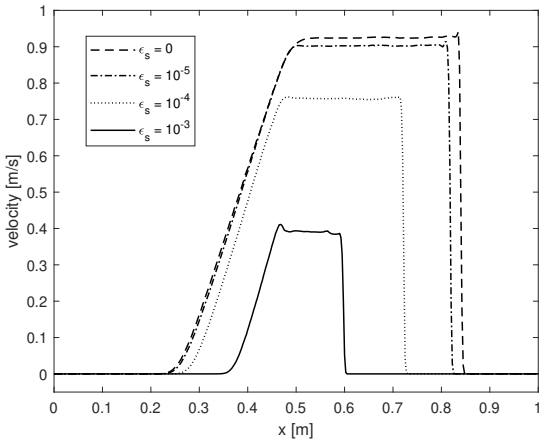


Figure 2. Multiphase shock tube - velocity distribution at $t = 0.2$ s.

In Figure 2, we show the velocity distribution over the x-axis for the three different cases. For comparison purposes, the monophase case solution obtained with a discretization of 56708 finite elements is also plotted in the same graph.

We observe that both on the left and the right sides, the initial velocity is preserved in a larger zone for increasing values of the solid phase concentration. This is due to the fact that the speed of sound in the mixture is lower than the one of the monophase case.

In Figures 3, we show the convergence of the velocity field at $t = 0.2$ considering different discretizations and the highest solid phase concentrations ($\varepsilon_s = 10^{-4}$ and $\varepsilon_s = 10^{-3}$). Although no analytical solution is known for this case, the convergence of the solution is assessed.

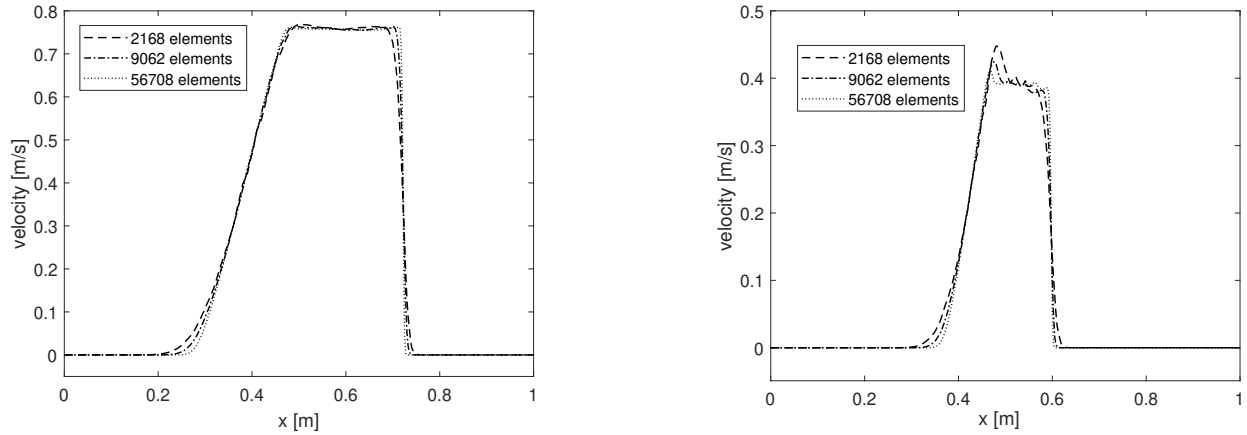


Figure 3. Multiphase shock tube - velocity distribution at $t = 0.2$ s for the cases $\varepsilon = 10^{-4}(left)$ and $\varepsilon = 10^{-3}$ (right).

4.2. Pyroclastic gravity current

In this section, we consider hypothetical scenarios of a pyroclastic gravity current hitting two buildings. We analyze the phenomena at a distance from the volcanic eruption such that the hypotheses of kinetic and thermal equilibrium can be assumed. For this case, the air viscosity is set to $\mu = 10^{-5}$ kg/m s and the thermal conductivity is $\lambda_g = 0.026$ W/m K.

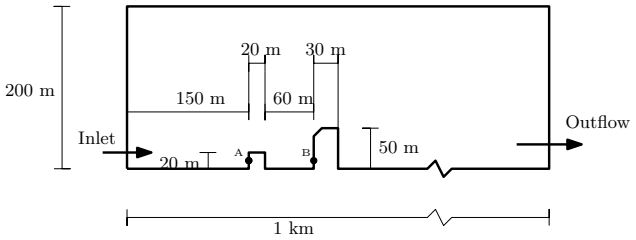


Figure 4. Real scale application: Computational domain used for the simulation.

The computational domain is depicted in Figure 4. We assume as initial condition still air at a density $\rho_g = 1.225$ kg/m^3 and temperature $T = 300$ K.

Concerning the boundary conditions, the inlet is on the left side of the domain and is described by

$$\begin{cases} \varepsilon_s(x_1 = 0, x_2, t) = \varepsilon_{s,max}\alpha(x_2)\beta(t) \\ u_1(x_1 = 0, x_2, t) = u_{1,max}\alpha(x_2)\beta(t) \\ u_2(x_1 = 0, x_2, t) = 0\,\text{m/s} \\ T(x_1 = 0, x_2, t) = 300\ \text{K} \end{cases} \tag{21}$$

where $\varepsilon_{s,max}$ and $u_{1,max}$ are the maximum values assumed by the solid concentration and of the mixture velocity at the inlet.

The functions $\alpha(x_2)$ and $\beta(t)$ are shown in Figure 5; $\alpha(x_2)$ represents the spatial distribution of the inlet solid concentration and the mixture velocity, and it decreases with the altitude from the ground level; $\beta(t)$ has a periodic behavior and simulates the different waves of pyroclastic material.

We perform two different simulations, considering a solid concentration of $\varepsilon_{s,max} = 5 \cdot 10^{-3}$ in the first case, and $\varepsilon_{s,max} = 1 \cdot 10^{-3}$ in the second case. For both analyses, we consider $u_{1,max} = 60$ m/s.

In Figure 6, we report the results in terms of pressure difference with respect to the atmospheric conditions measured at points A and B. We observe that, as the solid concentration reaches the first building (after about 10 seconds from the beginning of the simulation), for both the cases $\varepsilon_{s,max} = 5 \cdot 10^{-3}$ and $\varepsilon_{s,max} = 1 \cdot 10^{-3}$ the pressure at point A increases with respect to the atmospheric conditions, whereas the point B does not experience any pressure variation. After hitting the first obstacle with a different impact velocity depending on the solid concentration, the pyroclastic current reaches the second building (point B) giving rise to an increase of pressure on it. As expected, for higher values of solid concentration, the pressure jump is higher and occurs before.

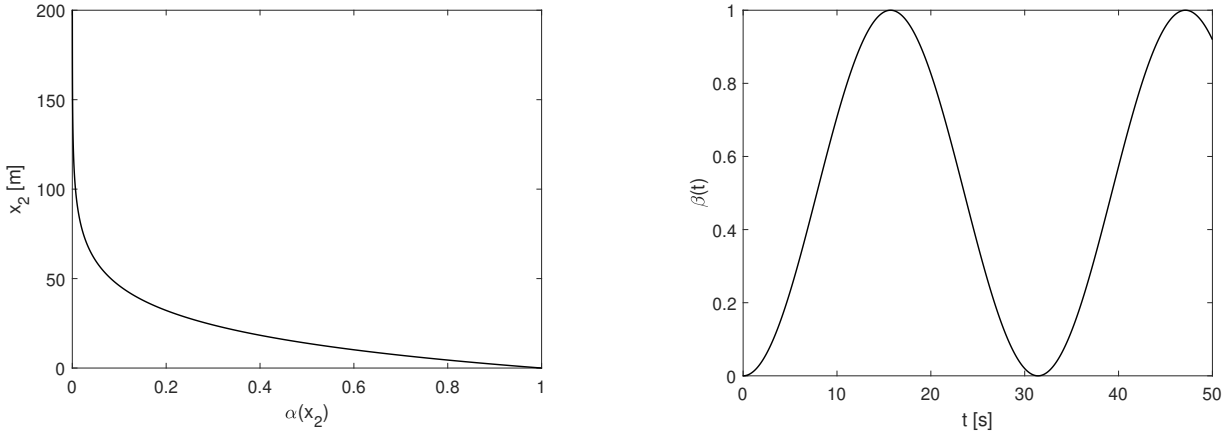


Figure 5. Real scale application: function $\alpha(x_2)$ and $\beta(t)$ for the description of the inlet

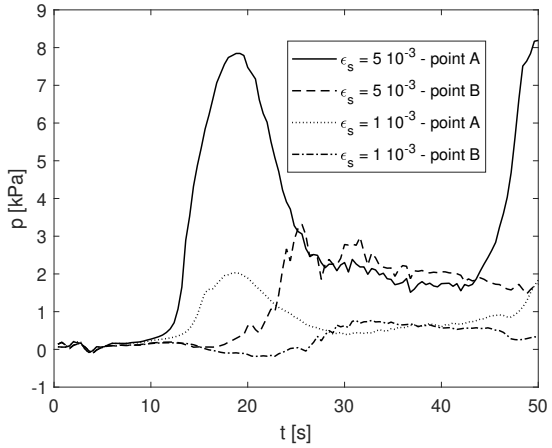


Figure 6. Real scale application: pressure at points A and B for two different intensities of solid concentration.

To further analyze the dynamics of the problem, in Figures 7 we show the solid phase concentration distribution at some significant instants of the simulation. In particular, in Figures 7a and 7b ($t = 14$ s), the simulation with $\varepsilon_{s,max} = 5 \cdot 10^{-3}$ reaches the first obstacle, whereas the more diluted one is still far. This is consequence of the minor kinetic energy possessed by the flow with lower solid phase concentration.

Figures 7c and 7d (at $t = 21$ s) show that while the denser flow is already hitting the second building, the other one has just overcome the first building.

Finally, in Figures 7e and 7f (at $t = 28$ s) the denser mixture reaches the ground behind the second building and has already filled the gap between the two buildings, while the more diluted distribution is still collapsing onto the second building.

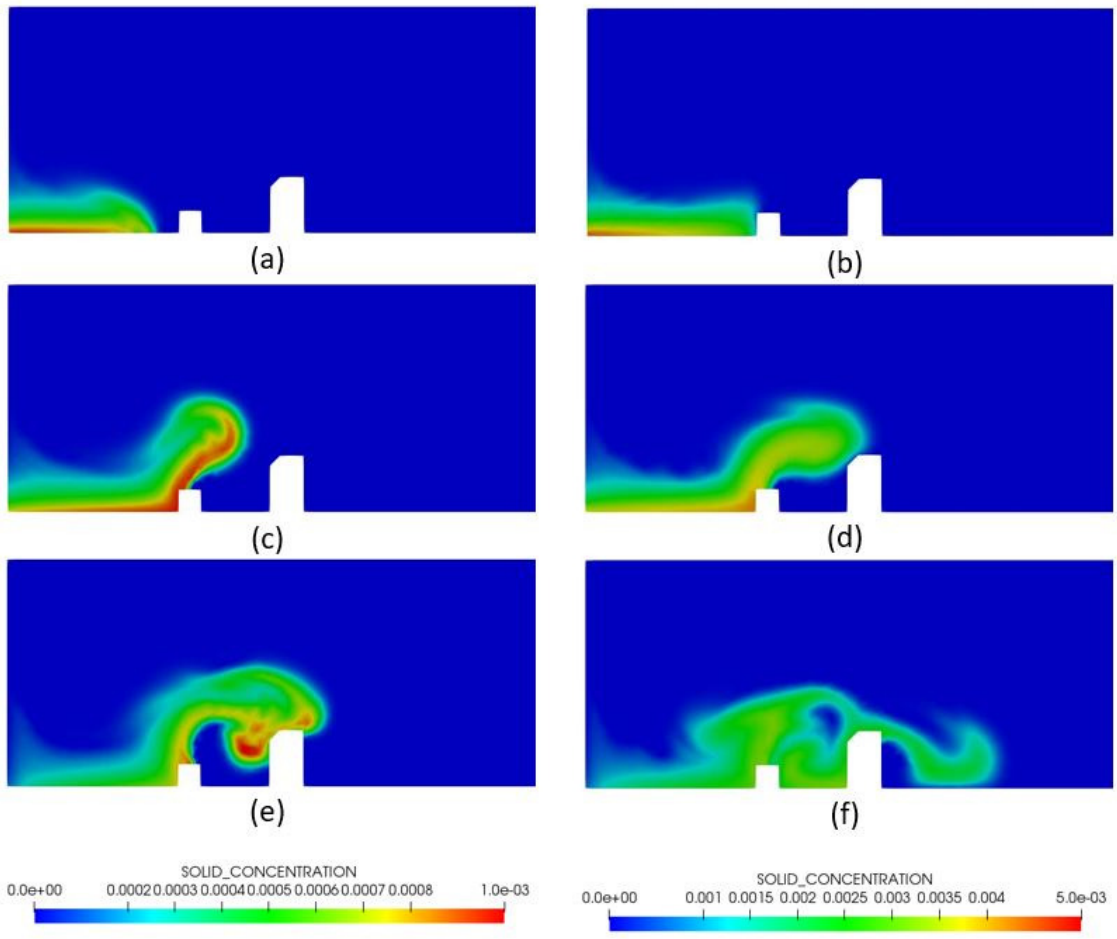


Figure 7. Real scale application: solid phase distribution at different instants.
$t = 14$ s: (a) $\varepsilon_{s,max} = 1 \cdot 10^{-3}$; (b) $\varepsilon_{s,max} = 5 \cdot 10^{-3}$;
$t = 21$ s: (c) $\varepsilon_{s,max} = 1 \cdot 10^{-3}$; (d) $\varepsilon_{s,max} = 5 \cdot 10^{-3}$
$t = 28$ s: (e) $\varepsilon_{s,max} = 1 \cdot 10^{-3}$; (f) $\varepsilon_{s,max} = 5 \cdot 10^{-3}$

The proposed simulation, with realistic values for the geometry of a urban settlement and for the gas and solid properties, is able to describe effectively the barrier effect that occurs between buildings during a pyroclastic event, showing the reduced values of the pressure on building less exposed to the the mixture current.

5. Conclusion

In this work, we have presented a stabilized finite element formulation for the description of pyroclastic gravity currents generated by explosive volcanic eruptions. In particular, we adopted a

Variational Multiscale Method, recently extended to multiphase compressible flows, to stabilize the finite element formulation and thus to avoid spurious oscillations of the solution variables.

The formulation is here tested and assessed against well-known benchmarks for compressible flow problems and applied to a real scale application considering a pyroclastic current hitting some buildings.

In all examples, the developed finite element formulation has shown robustness, convergence and accuracy, and it has demonstrated its potential for the evaluation of the effects of volcanic hazard on civil constructions placed at a certain distance from the volcano.

Future developments of this work foresee parametric analyses to assess the influence of different geometrical and physical quantities on the quantification of the hazard, as well as the extension to three-dimensional cases and the application to the real topography of the urban agglomerates around the Vesuvius volcano.

References

Baxter, P. J., Boyle, R., Cole, P., Neri, A., Spence, R., and Zuccaro, G. (2005). The impacts of pyroclastic surges on buildings at the eruption of the soufrière hills volcano, montserrat. *Bulletin of volcanology*, 67(4):292–313.

Bayona Roa, C. A., Baiges, J., and Codina, R. (2016). Variational multi-scale finite element approximation of the compressible navier-stokes equations. *International Journal of Numerical Methods for Heat & Fluid Flow*.

Bazilevs, Y., Calo, V., Cottrell, J., Hughes, T., Reali, A., and Scovazzi, G. (2007). Variational multiscale residual-based turbulence modeling for large eddy simulation of incompressible flows. *Computer methods in applied mechanics and engineering*, 197(1-4):173–201.

Bochev, P. B., Gunzburger, M. D., and Shadid, J. N. (2004). Stability of the supg finite element method for transient advection–diffusion problems. *Computer methods in applied mechanics and engineering*, 193(23-26):2301–2323.

Cerminara, M., Esposti Ongaro, T., and Berselli, L. C. (2016). Ashee-1.0: a compressible, equilibrium–eulerian model for volcanic ash plumes. *Geoscientific Model Development*, 9(2):697–730.

Codina, R. (2000). Stabilization of incompressibility and convection through orthogonal sub-scales in finite element methods. *Computer methods in applied mechanics and engineering*, 190(13-14):1579–1599.

Codina, R. (2002). Stabilized finite element approximation of transient incompressible flows using orthogonal subscales. *Computer methods in applied mechanics and engineering*, 191(39-40):4295–4321.

Cremonesi, M. and Frangi, A. (2016). A lagrangian finite element method for 3d compressible flow applications. *Computer Methods in Applied Mechanics and Engineering*, 311:374–392.

Dadvand, P., Rossi, R., and Oñate, E. (2010). An object-oriented environment for developing finite element codes for multi-disciplinary applications. *Archives of computational methods in engineering*, 17(3):253–297.

Esposti Ongaro, T., Cavazzoni, C., Erbacci, G., Neri, A., and Salvetti, M.-V. (2007). A parallel multiphase flow code for the 3d simulation of explosive volcanic eruptions. *Parallel Computing*, 33(7-8):541–560.

Franca, L. P., Frey, S. L., and Hughes, T. J. (1992). Stabilized finite element methods: I. application to the advective-diffusive model. *Computer Methods in Applied Mechanics and Engineering*, 95(2):253–276.

Fu, L. (2019). A low-dissipation finite-volume method based on a new teno shock-capturing scheme. *Computer Physics Communications*, 235:25–39.

Hughes, T. J., Feijóo, G. R., Mazzei, L., and Quincy, J.-B. (1998). The variational multiscale method—a paradigm for computational mechanics. *Computer methods in applied mechanics and engineering*, 166(1-2):3–24.

Iacuaniello, M., Montanino, A., De Gregorio, D., and Zuccaro, G. (2020). Pyroclastic flow mitigation strategies: a new perspective for the red area. *Natural Hazards and Earth System Sciences Discussions*, pages 1–28.

Koobus, B. and Farhat, C. (2004). A variational multiscale method for the large eddy simulation of compressible turbulent flows on unstructured meshes—-application to vortex shedding. *Computer Methods in Applied Mechanics and Engineering*, 193(15-16):1367–1383.

Marble, F. E. (1970). Dynamics of dusty gases. *Annual Review of Fluid Mechanics*, 2(1):397–446.

Montanino, A., Franci, A., Rossi, R., and Zuccaro, G. (2021). Finite element formulation for compressible multiphase flows and its application to pyroclastic gravity currents. *submitted to Computational Mechanics*.

Neri, A., Esposti Ongaro, T., Macedonio, G., and Gidaspow, D. (2003). Multiparticle simulation of collapsing volcanic columns and pyroclastic flow. *Journal of Geophysical Research: Solid Earth*, 108(B4).

Pelanti, M. and LeVeque, R. J. (2006). High-resolution finite volume methods for dusty gas jets and plumes. *SIAM Journal on Scientific Computing*, 28(4):1335–1360.

Rispoli, F. and Saavedra, G. R. (2006). A stabilized finite element method based on sgs models for compressible flows. *Computer methods in applied mechanics and engineering*, 196(1-3):652–664.

Tezduyar, T. and Sathe, S. (2003). Stabilization parameters in supg and pspg formulations. *Journal of computational and applied mechanics*, 4(1):71–88.

Toro, E. F. (2013). *Riemann solvers and numerical methods for fluid dynamics: a practical introduction*. Springer Science & Business Media.

Zuccaro, G., Cacace, F., Spence, R., and Baxter, P. (2008). Impact of explosive eruption scenarios at vesuvius. *Journal of Volcanology and Geothermal Research*, 178(3):416–453.

Zuccaro, G. and De Gregorio, D. (2019). Impact assessments in volcanic areas-the vesuvius and campi flegrei cases studies. *Annals of Geophysics*, 62(1):02.

Zuccaro, G. and Leone, M. F. (2012). Building technologies for the mitigation of volcanic risk: Vesuvius and campi flegrei. *Natural hazards review*, 13(3):221–232.

Design Criteria for Temporary Structures in Uncertain Extreme Natural Hazard Load Environments

E. Saleh[1)], and J. Mohammadi[2)]

[1)]*Civil Engineering Department, Hashemite University, Zarqa, Jordan, esaleh@hawk.iit.edu*
[2)]*Civil, Architectural and Environmental Engineering, Illinois Institute of Technology, Chicago 60616, United States, mohammadi@iit.edu*

Abstract: In design of temporary structures, designers often face the question:"What is an appropriate design load level for the structure, given that future loads are uncertain?" This is especially the case in situations where the structure is subject to potential natural hazards such as hurricanes, with potentials for extreme wind pressures. The uncertainty in predicting loads may result in inadequate designs causing severe losses to contractors and erectors. A proper design with considerations for the uncertainties inherent in natural-hazard loads, along with information on extreme load events during construction, will minimize potential losses and enhance the system reliability. In this paper, a three-stage decision-based programming model is introduced to determine design loads resulting from probable extreme hazard events for temporary structures such as hurricanes. This model evaluates the expected cost of implementing a contingency plan against extreme wind loads when meteorological data indicate potentials for occurrence of hurricane events exist. An example of a steel structure under construction is presented to illustrate the proposed methodology. In this example, the significance of having a contingency plan, to safeguard against hurricanes, is presented and shown to improve the profitability of the project, and to impact the decision-making process at the time of bidding.

Keywords: temporary structure; wind load; uncertainty; optimization.

1. Introduction

Hurricanes are frequent events in the southern and southeastern regions of the United States. In many such events, the potential for landfalls exists with severe property damage and losses. The annual frequency of occurrence in the eastern US is estimated to be 2.1 hurricanes per year (Miller, 1999). Hurricanes are classified based on the Saffir–Simpson Scale, which is driven by the hurricane's wind speed (Li, 2011). Klotzbach et al. (2009) classify hurricanes into three categories; (1) Tropical storm, (2) Hurricane, and (3) Intense hurricane. A hurricane has categories 1 or 2 and an intense hurricane has categories 3, 4, or 5 on the Saffir–Simpson Scale.

Forecasting of hurricane losses provides a basis for estimating the adequate catastrophe cover such as insurance premiums and reinsurance. The premium is usually evaluated as the expected annual loss (EAL) for each category of structures in a certain region; while the probable maximum loss (PML) dictates if reinsurance is required or not. The forecasted loss can also be represented by the exceedance probability curve (EP-curve). Such a curve describes the probability of exceeding a certain loss value given the occurrence of

© 2021 *by E. Saleh, and J. Mohammadi*

a hurricane. Understanding the risk of an upcoming hurricane during construction may help contractors determine the cost associated with the construction process and recovery, if any.

Three principles govern the risk of a temporary structure or a structure during construction to hurricanes or extreme wind events. These are (1) hazard; (2) vulnerability; and (3) exposure. The hazard component identifies the frequency and severity of the hurricane by generating the distribution of hurricane wind speed at the desired location. The vulnerability and exposure components take these expected wind speed values along with specific information pertaining to the susceptibility of the structure to the hazard to evaluate the loss or damage generated at the different levels of the probability of hurricane occurrences. This information on hazard, exposure, and vulnerability can be used in evaluating the insurance losses following hurricane events. This is often done by producing an exceedance probability curve for a given structure or structure type (Hamid, 2007). However, these models are mainly applicable to permanent structures that represent the major stock of buildings in the insurance inventory. Besides, extending these models to temporary structures may not be an attractive option for some contractors since they often require a large computational and knowledge base. Nevertheless, in some projects where the profitability is highly dependent on whether a hurricane occurs or not during the construction period, the use of models similar to those in permanent structures may be justified. In such cases, this could affect the decision of the contractor when participating in the bid process.

Extreme events such as hurricanes or earthquakes may need a robust planning process to account for their possibilities to attain a reliable and economic design. A recourse decision after observing of uncertainties may be possible, which represents an opportunity for contractors to reach a more optimal design. Stochastic optimization provides one of the tools for optimizing the design while considering the effect of uncertainties. Stochastic optimizations have been used in many fields such as constructing models for solving capacity allocation under uncertainty (Chen et al., 2015; Chen et al., 2002; Andreatta et al., 2011), air traffic flow management (Triki et al.,2005), and in applications for disaster management (Barbarosoğlu et al., 2004; Bozorgi et al., 2013; Noyan et al., 2012).

In stochastic programming, several design parameters are defined as random variables, with data sampling methods used to determine their probability distributions. However, enumeration of all possible scenarios, in a data-sampling scheme, is sometimes unattainable because of lack of basic information upon which a scenario can be modeled. Before starting the scenario generation, the stochastic optimization problem should be properly defined. The information needed for proper definition includes (1) the number of stages that decisions are required; (2) the time between stages; (3) type of information that will be available at the time that a decision should be made; and (4) time needed for taking a response action (for example, the forecasting system for hurricanes will require 3-4 days for taking actions).

In this study, the effect of unpredictable actions such as hurricanes on a temporary bracings design plan for structures under construction or erection is discussed. An example of a steel structure under construction is given to illustrate the proposed methodology.

2. Proposed Hurricane Catastrophe Model

Estimating the loss generated from hurricanes for temporary structures has its peculiarities. For instance, the estimated loss could be affected by (1) the dynamic nature of temporary structures such as the change in the resistance as a function of the construction sequence; (2) the construction period and the fact that it may not

completely fall within the hurricane season; (3) parties involved – since, in contrast to permanent structures where only the owner and the insurer share the risk, temporary structure failures may be shared by several parties (e.g., the contractor, subcontractor, erector, owner and the insurance company, depending on the contract's nature); (4) potential for the alteration of the original design after forecasting a hurricane; and (5) major components (such as cranes, scaffold and falseworks), their installation demands and usage. To clarify the latter item, it is noted that each time one of these components is erected, it is treated as a new system under a new wind load effect (triggered by the locality) and with different usage, configurations, and users (Shapiro, 2004).

2.1 ESTIMATED LOSS

Using the Poisson distribution to model hurricane events with an average occurrence rate λ, and considering the cumulative distribution function of losses given the occurrence of a hurricane as $F(l)$, then the exceedance probability (EP) of a given loss, l , can be evaluated using the following series of computation

$$EP = \Sigma P(l|X = i)P(X = i) \quad (1)$$

The summation considers the fact that theoretically during a given year, any number of events would be possible. $P(X=i)$ is from Poisson distribution and indicates the probability that there will be i events in T=1 year. In general, for i events (and assuming events are statistically independent of one another),

$$P(l|X = i) = 1 - (1 - p)^i \quad (2)$$

Where $p = 1 - F(l)$, in which p, and $F(l)$ are the exceedance probability and the cumulative distribution function of the loss for i=1, respectively. Using this relation in the equation for EP and substituting for $P(X=i)$ from the Poisson equation, it can be shown that the resulting series has a limit, according to the Maclaurin series for the exponential function, in the following form (noting that T= 1 year).

$$EP(l, T = 1year) = 1 - e^{-\lambda(1-F(l))} \quad (3)$$

The uncertainties in loss estimations of temporary structures can be considered to be a product of two components – namely, the uncertainty in the wind speed distribution and uncertainty in the construction sequence when the hurricane hits. These uncertainties are related to the wait-and-see phase (before uncertainties about the wind speed are realized, typically at the bidding phase) since it cannot be known unless the hurricane makes landfall. However, at the beginning of the project, the design wind speed for each construction sequence is known and belong to the here-and-now phase (see Fig.1). Considering these uncertainties, Eqs. 1,2 and 3 can be adjusted such that they become valid for temporary structure as follows:

$$EP(l, t = 1year) = 1 - \sum_{s_i} \sum_{n=0}^{\infty} F^n(l|s_i) \frac{(\lambda)^n}{n!} e^{-\lambda} P(s_i)$$

$$EP(l, t = 1year) = 1 - \sum_{s_i} P(s_i) e^{-\lambda(1-F(l|s_i))} \quad (4)$$

Where $P(s_i)$ represents the probability of being in construction sequence i, $F(l|s_i)$ is the cumulative density function of loss given erection sequence s_i, and n is the number of events per year.

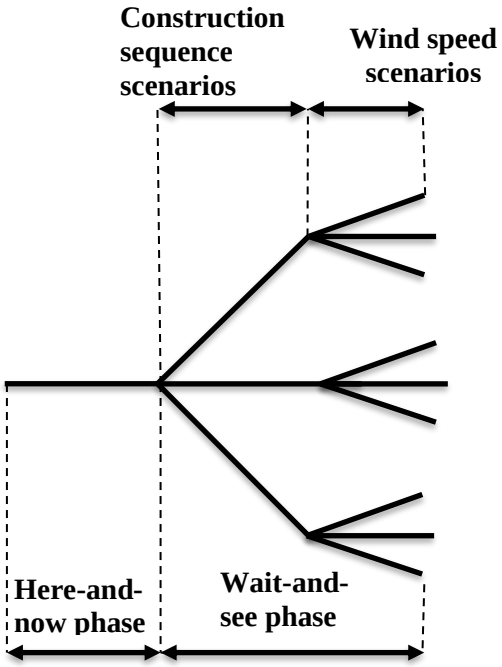


Figure 1: Scenario tree

The information needed to estimate the loss is (1) the probability of being in a certain erection sequence; (2) the mean frequency of occurrence of hurricanes and the wind speed probability distribution under such an event in the region; and (3) the distribution of losses given a certain wind speed distribution and an erection sequence.

If the construction period does not cover the full hurricane season (between July 1 and October 31), Eq.4 can be adjusted by reducing the mean frequency of occurrence based on the length of time when the construction falls within the hurricane season. Alternatively, a new mean frequency of occurrence of a hurricane for months of construction may be derived from hurricane wind data.

Dividing the construction into sequences requires identifying the stopping points of each phase of construction. For example, if an erector could put up the entire structure in a day, the construction could be thought of as one sequence. However, often the erection takes a longer time requiring the erector to decide upon stages of construction or a stopping point where the construction could be thought of as a separate structure with its design wind speed. Identifying the stopping points does not follow any set of rules. For example, the erector may decide on dividing the construction into sequences based on the time taken to finish each stage (say for instance, if erecting two bays and two tiers can be finished per day). Then, the stages of construction could easily be identified. It is noted that, the probability of being in a certain construction sequence could be evaluated as the portion of time spent in that sequence, or from past performances of the contractor in similar situations.

2.1.1 Optimization of the design of temporary bracings

In structures under construction, the number of temporary bracings needed to ensure stability depends on the uncertain demand (e.g., wind load). Thus, the design of these bracings requires forecasting wind loads to

provide the necessary capacity and resiliency. In the following paragraph, a stochastic formulation for optimizing the expected cost of construction during the hurricane season is presented.

2.1.1.1 Model Formulation The problem that faces the contractor can be summarized as follows:

- At the beginning of the project, the erector needs to decide on a certain design wind speed, x_1. This is the decision stage and can also be denoted as [here-and-now] decision that implies a cost $A(x_1)$.

- After forecasting a hurricane, the erector may decide on the installation of additional temporary bracings x_2. This new decision is made after forecasting and implies an additional cost $B(x_1, x_2, \xi_1, \xi_2)$ that depends on the initially decided design wind speed x_1, decided alterations x_2, stage of construction ξ_1, and the forecasted wind (or hurricane) ξ_2.

- Depending on the stage of construction ξ_1 and the forecasted wind (or hurricane) ξ_2 the demand is $d(\xi_1, \xi_2)$.

- Since the installation of temporary bracings was based on the forecasting of an incoming hurricane and not the actual wind speed, an additional cost of failure, $D(x_1, x_2, \xi_1, \xi_2)$ may apply that depends on the amount of alterations in the design wind speed (i.e., after the installation of additional temporary bracings) and the actual demand $d(\xi_1, \xi_2)$.

- Consequently, the total cost for each [here-and-now] option $C(x_1, x_2, \xi_1, \xi_2) = A(x_1) + B(x_1, x_2, \xi_1, \xi_2) + D(x_1, x_2, \xi_1, \xi_2)$ is a random variable. Therefore, choosing an optimal [here-and-now] decision (i.e., design wind speed at the beginning of the construction) is a choice among distributions, not numbers.

With the above assumptions, the problem is formulated in a three-stage expected cost stochastic programming formulation (Birge, and Louveaux, Francois, 2011; Wallace and Ziemba, 2005) as follows assuming a linear variation of cost with the decision variables:

$$\min \mathbb{E}_{\xi_1,\xi_2}[c_0(\xi_1)x_1(\xi_1) + c_1(\xi_1)x_2(x_1, \xi_1, \xi_2) + c_2(\xi_1)w(\xi_1, \xi_2, x_1, x_2)] \quad (5)$$

$$s.t$$

$$x_1 + x_2(x_1, \xi_1, \xi_2) \geq k(\xi_1, \xi_2) \quad (6)$$

$$x_2(x_1, \xi_1, \xi_2) \leq \big(e - \Pi(\xi_1, \xi_2)\big)M \quad (7)$$

$$x_1(\xi_1), x_2(\xi_1, x_1) \geq 0 \quad (8)$$

Where:
$x_1(\xi_1)$ is the first stage decision (i.e., initially decided design wind speed) that must be made before the future is realized. This first stage decision could be allowed to vary with the construction sequence ξ_1.

$x_2(x_1, \xi_1, \xi_2)$ is the second stage decision (i.e., amount of alteration in design wind speed) that could adapt to the additional information available at the end of the first stage that is a function of the construction sequence ξ_1, first stage decision x_1, and information about a forecasted hurricane ξ_2.

$w(\xi_1, \xi_2, x_1, x_2)$ is a measure of the deficiency of meeting the demand that is a function of the construction sequence ξ_1, first stage decision x_1, second stage decision x_2, and information about a forecasted hurricane ξ_2.

$c_0(\xi_1)$ = cost coefficient that converts first stage design variable x_1 to monetary value. To consider the dynamic nature of construction where the cost of temporary bracings does not apply until the start of the new construction sequence, this coefficient is assumed to be a function of the construction sequence ξ_1.

$c_1(\xi_1)$ = cost coefficient that converts second stage design variable x_2 to the monetary value that is a function of the construction sequence.

$c_2(\xi_1)$ = cost coefficient that converts the deficiency of meeting the actual demand that is a function of the construction sequence.

$k(\xi_1, \xi_2)$ = the forecasted demand at the second stage (i.e., after forecasting a hurricane) that is a function of construction sequence and forecasted wind speed.

e is a vector of ones (i.e., $e_s = 1$ for all construction sequence, s)

$\Pi(\xi_1, \xi_2)$ is equal to 1, if the contractor cannot fulfill the required design wind speed at a certain construction sequence in the time limit of response (generally the response time for a hurricane is between 3 to 4 days) and $\Pi(\xi_1, \xi_2)$=0 if he can install the additional temporary bracings in the time limit for response. This value is assumed to be dependent on the construction sequence, and forecasted wind speed.

M is a large positive constant (e.g., 1×10^9). There is only one restriction on M, i.e., M should guarantee that some feasible x_2 always exists, such that the time constraint for taking a response action can be met.

Expression (5) describes the erector's objective in making a decision that is considered optimal on average by minimizing the expected value $\mathbb{E}_{\xi_1,\xi_2}[.]$. Note that the variable $x_2(x_1, \xi_1, \xi_2)$ defines a policy or rule to be followed, if a certain uncertainty is detected. For example, if a hurricane was forecasted, x_2 takes some value depending on the information available at that time about the stage of construction, and the wind speed distribution.

The condition denoted by Expression (6) represents a minimum capacity that will need to be considered for the structure at different construction sequences, once a hurricane is forecasted. The parameter $k(\xi_1, \xi_2)$ is equated to x_1, if the contractor is willing to accept the failure of the structure at all or some stages of construction. The condition denoted by Expression (7) considers the ability of the contractor to put the necessary temporary bracings within the appropriate time. This condition indicates that if time limitation exists, the original design wind speed at that particular construction sequence in response to the forecasted hurricane cannot be altered.

For practical purposes, the above formulation is approximated by a scenario tree (see Fig. 1). Using the generated scenarios, the formulation in (5–8) could be rewritten as follows:

$$\min_{x_1 \geq x_1^l} c_0 x_1 + \sum_{i_1 \in \Omega_1} p_{i_2} \, \mathrm{Q}_1(\mathrm{x}, \xi) \tag{9}$$

Where:

$$Q_1(x,\xi) = \sum_{i_1 \in \Omega_1} p_{i_1|i_2} \left\{ \begin{array}{c} \min\left[c_1\left(\xi_1^{\,i_1}\right) x_2\left(x_1, \xi_1^{\,i_1}, \xi_2^{\,i_2}\right) + c_2\left(\xi_1^{\,i_1}\right) w\left(\xi_1^{\,i_1}, \xi_2^{\,i_2}, x_1, x_2\right)\right] \\ s.t \\ x_1 + x_2\left(x_1, \xi_1^{\,i_1}, \xi_2^{\,i_2}\right) \geq k\left(\xi_1^{\,i_1}, \xi_2^{\,i_2}\right) \\ x_2\left(x_1, \xi_1^{\,i_1}, \xi_2^{\,i_2},\right) \leq \left(e - \Pi\left(\xi_1^{\,i_1}, \xi_2^{\,i_2}\right)\right) M \\ x_2\left(x_1, \xi_1^{\,i_1}, \xi_2^{\,i_2},\right) \geq 0 \end{array} \right\}$$

$i_1 \; and \; i_2$ are the scenario generated of the construction sequence, and wind speed, respectively.

p_{i_1} and p_{i_2} are the probability of being in the construction sequence i_1, and the probability of realizing a wind speed i_2, respectively.

$\xi_1^{\,i_1}$ is a realization of a construction sequence such that $i_1 \in \Omega_1$, where Ω_1 is the space construction sequence scenarios.

$\xi_2^{\,i_2}$ is a realization of a wind speed forecast such that $i_2 \in \Omega_2$, where Ω_2 is the space of wind speed scenarios.

x_1^l is the lower bound of first stage design wind speed that could be based on the daily expected wind speed.

$p_{i_1|i_2}$ is the conditional probability of realizing construction sequence i_1 at realized wind speed i_2.

Example. Steel frame under construction during the hurricane season.

Situation:
The contractor needs to decide on a design wind speed value before starting the construction. The construction falls in the hurricane season. Of course, the contractor does not know whether or not a hurricane will hit during the construction period. Therefore, the contractor wants to quantify the risk of construction to identify the need for catastrophe risk transfer measures and to quantify the feasibility of the project.

In this example, the construction is decomposed into 16 construction sequences as shown in Fig. 2. The sequences start with the first-floor bays then followed by the tier on the second floor. For example, SEQ12 represents the first sequence at the second tier. The construction sequence starts by constructing an initial braced box then going outward. The cost is assumed to be a function of the design wind speed multiplied by a cost coefficient that converts the design wind speed into a monetary value. The contractor or erector may consider modeling the cost as a function of the return period, this modeling choice is also acceptable.

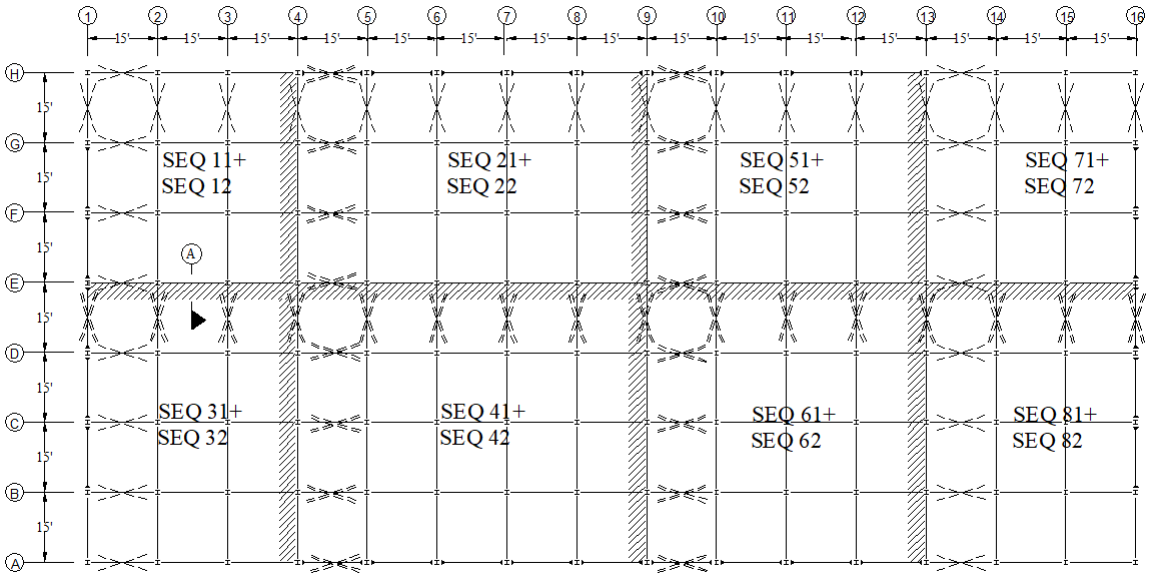


Figure 2. The sequencing plan of construction

The process the contractor would follow is:

- Decide on the current design wind speed.

- Add additional temporary bracings if a hurricane was forecasted, this additional quantity is a function of the initial design wind speed and the predefined estimate of required hurricane design wind speed.

Economic information:

- The cost coefficient of initial temporary bracings, $c_0 = 100$

- The cost coefficient of the additional temporary bracings needed once a hurricane is forecasted can be higher than the initial unit cost if the surge in the prices once a hurricane is forecasted is considered. In this example, we use $c_1 = 150$

- The cost of failure if the demand was not met varies with the construction sequence is also needed,
- The cost of failure of each sequence is assumed to be equal to 10,000.

For simplicity in the optimization, $w(\xi_1, \xi_2, x_1, x_2) = \max[d(\xi_2) - (x_1 + x_2), 0]$. This condition implies that the cost of failure only applies if the demand $d(\xi_2)$ was higher than the new altered design wind speed of the sequence. The probability of being in any construction sequence is assumed to be equally likely (see Fig. 3). An alternative model may consider the number of bays constructed in each sequence relative to the

total number of bays or the plan area of the sequence to the total plan area as a measure of the probability of being in that particular construction sequence.

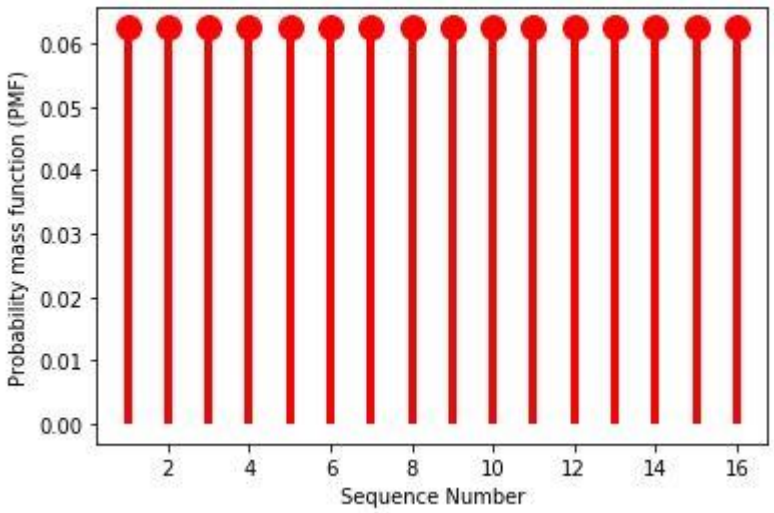


Figure 3: Uniform distribution for the probability of being in a particular construction sequence

The conditional probability is evaluated assuming independence between wind speed and construction sequence, that is $p_{i1|i2} = p_{i1}$. However, if construction is to be partially conducted in the hurricane season, it may be necessary to consider the interdependence between the wind speed and the construction sequence, because the forecasting of the hurricane may delay or alter a particular construction sequence.

Our construction is assumed to be located in Miami-Dade County (in the State of Florida, in US), where the associated hurricane probabilities are as follows (Klotzbach et al., 2009):

– Probability of tropical storms bringing a wind speed that exceed 40 mph = 33.2%

– Probability of hurricanes bringing a wind gust that exceed 75 mph to the county =11.7%

– Probability of intense hurricanes bringing a wind gust that exceeds 115 mph to the county =4.5%.

The above probabilities are equivalent to the probability of forecasting a certain event that requires a response plan. A distribution model based on these probabilities is shown in Fig. 4. This probability distribution is constructed based on the assumption that wind speeds are equally likely to occur within the aforementioned ranges. For example, wind speed in the range between 115 mph to 300 mph is equally likely with a probability density function of $4.5/(300 - 115)$. However, hurricanes cannot be predicted very accurately; and several past hurricanes (e.g., hurricane Katrina) surprised observers with their change in intensity or their inaccurate forecast (Hayden, 2006; Hayden 2019). To account for imperfect forecasting, the probabilities of the wind speed scenarios were taken to be equivalent to the probability of an inaccurate forecast that is assumed to follow a normal distribution centered around the forecasted value ξ_2 and with a standard deviation of 20 mph (Hayden, 2006; Hayden 2019) (i.e., $i_2 \sim N(\xi_2, 50)$). The normal distribution was chosen because the reported data indicates that the shift in hurricane intensity could lead to either a more

intense or less intense hurricane than what has been forecasted. For example, a category 5 (wind speed in the range of 155-250 mph) hurricane may end up as a category 3 hurricane (wind speed in the range of 110-130 mph) (Hayden, 2006; Hayden 2019).

Cost coefficients could be evaluated by calculating the increase in cost per one mph increase in design wind speed. The parameter c_2 represents the cost of failure per deficiency in satisfying one mph of wind speed. For example, in SEQ82, $c_2 = 8{,}500$ means that there is a potential loss of \$8,500 if the demand was higher than the design wind speed by 1 mph. However, this failure cost is limited to the cost of the construction sequence (i.e., as seen in Fig. 5, beyond certain wind speed (within Category 5), the failure cost is considered independent of the wind speed). Table 1. summarizes the information needed for each construction sequence.

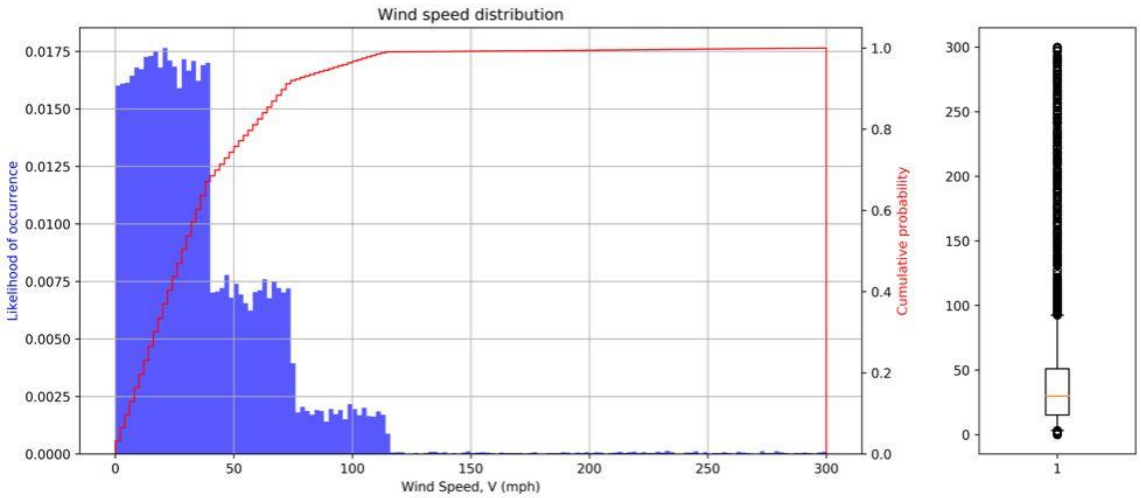


Figure 4: Distribution of 100,000 random samples of wind speed.

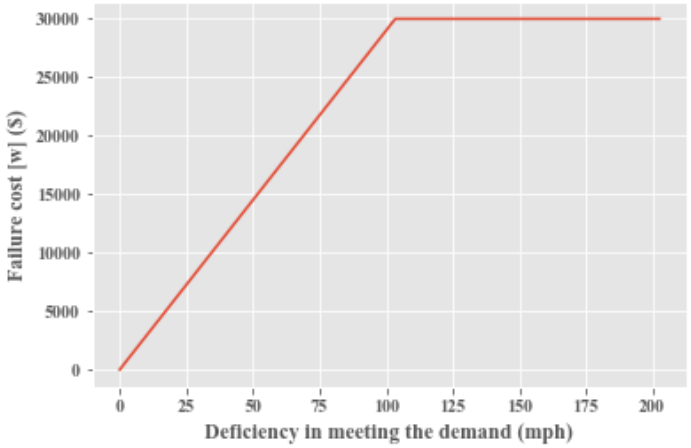


Figure 5. Failure cost as a function of the deficiency in meeting the realized wind speed

Table 1: Defined parameters for the optimization problem

Sequence ID	p_{i1}	$c_0(\$/mph)$	$c_1(\$/mph)$	$c_2(\$/mph)$
SEQ11	1/16	10	15	100
SEQ12	1/16	11	16	150
SEQ21	1/16	12	17	200
SEQ22	1/16	13	18	250
SEQ31	1/16	14	19	300
SEQ32	1/16	15	20	350
SEQ41	1/16	16	21	400
SEQ42	1/16	17	22	450
SEQ51	1/16	18	23	500
SEQ52	1/16	19	24	550
SEQ61	1/16	20	25	600
SEQ62	1/16	21	26	650
SEQ71	1/16	22	27	700
SEQ72	1/16	23	28	750
SEQ81	1/16	24	29	800
SEQ82	1/16	25	30	850

If a multiplicative relationship between cost and design wind speed cannot be properly prescribed, various scenario generation in the optimization problem can be considered to include different alternative design wind speeds and their associated costs. Accordingly, the optimization is performed to arrive at the best alternative by assuming no time limitation at any construction sequence and a relatively reasonable wind speed forecast. With this scheme, we also accept that the failure of the structure will occur if the cost estimation was in favor of the failure.

Optimization result

Assigning a lower bound at the first stage based on expected daily wind equal to 40 mph, the optimization model yields an expected cost of temporary bracings = $11,200. It is noted that the optimization model finds that the lower bound of the first stage wind speed (i.e., 40 mph) is the optimum decision at all stages. This result is understandable because of the low failure cost and the use of a "neutral" risk measure (i.e., the expectation). If the failure cost is greater or a more conservative risk measure is used, a change in the decision is expected.

3. Concluding Remarks

Temporary structures may be more vulnerable to a higher wind speed than the permanent ones due to their lower stability, lower resistance and the high drag coefficient that is especially prevalent in temporary structures with lattice elements. To account for the higher wind effects, which are often expected in hurricanes, temporary bracings are usually installed until the structure can resist the applied loads.

For the construction period that falls into the hurricane season, it is sometimes necessary to account for hurricane loads in the second stage after the forecasting of a hurricane. However, the cost of a contingency plan should be evaluated to ensure that the risk on the contractor and erector is not excessive. This study

presents a three-stage stochastic optimization formulation that can help in choosing the appropriate design wind load, given that the wind load is uncertain. The formulation presented can also be used in evaluating the expected cost of the contingency plan against hurricane wind loads. This optimization model can be refined by considering measures of risk other than those expected if the contractor feels that a more conservative decision-making process is required. Some measures of risk that can be used are (1) conditional value at risk (C-VAR); (2) value at risk (VAR); (3) chance constraints; and (4) worst-case scenario optimization, which seems only reasonable in highly sensitive applications such as in the case of nuclear power plant temporary works.

References

American Institute of Steel Construction. Code of standard practice for steel buildings and bridges. American Institute of Steel Construction, 2000.

Andreatta, Giovanni, Paolo Dell'Olmo, and Guglielmo Lulli.An aggregate stochastic programming model for air traffic flow management. European Journal of Operational Research, 215(3):697–704, 2011.

ASCE. ASCE/SEI 7 minimum design loads for buildings and other structures,2016.

Birge, J. R., & Louveaux, F. (2011). *Introduction to stochastic programming*. Springer Science & Business Media.

Bozorgi-Amiri, Ali, Mohamad Saeed Jabalameli, and SMJ Mirzapour Al-eHashem. A multi-objective robust stochastic programming model for disaster relief logistics under uncertainty. OR spectrum, 35(4):905–933, 2013.

Chen, Yin-Yann and Hsiao-Yao Fan. An application of stochastic programming in solving capacity allocation and migration planning problem under uncertainty. Mathematical Problems in Engineering, 2015, 2015.

Chen, Zhi-Long, Shanling Li, and Devanath Tirupati. A scenario-based stochastic programming approach for technology and capacity planning. Computers &Operations Research, 29(7):781–806, 2002.

Hamid, Shahid S, Jean-Paul Pinelli, Shu-Ching Chen, and Kurt Gurley. Catastrophe model-based assessment of hurricane risk and estimates of potential insured losses for the state of Florida. Natural Hazards Review, 12(4):171–176, 2011.

Hamid, Shahid. Actuarial model for estimating insured hurricane losses. International Hurricane Research Center. Florida International University, Miami, Florida, USA, 2007.

Hayden, Thomas. Super storms: No end in sight.

Hayden, Thomas. Super storms: no end in sight. National geographic, 210(2):66–77, 2006.

Howard I Shapiro, Jay P Shapiro, Lawrence K Shapiro, and HI Shapiro. Cranesand derricks. McGraw-Hill New York, 2004.

Klotzbach, P, W Gray, U Shama, L Harman, and D Fitch. United states landfalling hurricane probability project. Available at: www. e-transit.org/hurricane/welcome. html (accessed December 20, 2017). [Google Scholar],2014.

Li, Chun-yang , Xun Chen, Xiao-shan Yi, and Jun-yong Tao. Interval-valued reliability analysis of multi-state systems. IEEE Transactions on Reliability,60(1):323–330, 2011.

Miller, David. Uncertainty in hurricane risk modeling and implications for securitization. Causal Actuarial Society: Discussion Paper Program, page 117,1999.

Noyan, Nilay. Risk-averse two-stage stochastic programming with an application to disaster management. Computers & Operations Research, 39(3):541–559,2012.

Ruby, David I and Jeffrey Gasparott. Aisc live webinar ASCE 37: Design loads on structures during construction/presentation. AISC, 2019.

Triki, Chefi, Patrizia Beraldi, and George Gross. Optimal capacity allocation in multi-auction electricity markets under uncertainty. Computers & operations research, 32(2):201–217, 2005.

Twarek, Andrew. AISC live webinar introduction to the steel construction process: The team behind the building. AISC, 2019.

Wallace, S. W., & Ziemba, W. T. (Eds.). (2005). *Applications of stochastic programming*. Society for Industrial and Applied Mathematics.

Reliability Analysis of Randomly Excited Structures with Interval Stiffness Parameters via Sensitivity Analysis

F. Giunta[1], G. Muscolino[1], and A. Sofi[2]

[1]*Department of Engineering, University of Messina, Messina, Italy, filippogiunta91@gmail.com, gmuscolino@unime.it*
[2]*Department of Architecture and Territory, "Mediterranea" University of Reggio Calabria, Reggio Calabria, Italy, alba.sofi@unirc.it*

Abstract: The present study focuses on reliability analysis of linear discretized structures with interval stiffness parameters subjected to stationary Gaussian multi-correlated random excitation. The reliability function for the *extreme value* stress process is evaluated in the framework of the *first-passage* theory. Such a function turns out to have an interval nature due to the uncertainty affecting the stiffness parameters. The aim of the analysis is the evaluation of the bounds of the *interval reliability function* which provides a range of structural performance. The range of stress-related quantities may be significantly overestimated as a consequence of the so-called *dependency phenomenon*. To reduce overestimation, a sensitivity-based procedure is proposed in the context of the *Improved Interval Analysis* via *Extra Unitary Interval*. The main advantage of this procedure is the capability of providing appropriate combinations of the endpoints of the uncertain parameters which yield accurate estimates of the bounds of the *interval reliability function* for the *extreme value* stress process.

A wind-excited steel telecommunication antenna mast with uncertain Young's modulus is analyzed to demonstrate the accuracy and efficiency of the proposed method.

Keywords: Interval reliability function; Interval analysis, Sensitivity analysis.

1. Introduction

The actual values of parameters involved in any engineering design are affected by several sources of uncertainty ensuing from manufacturing inaccuracies, model or measurement errors etc. (see e.g., Ayyub and Klir, 2006; Der Kiureghian, 2008; Corotis, 2015). Such uncertainties have been traditionally taken into account in structural reliability analysis by applying well-established probabilistic approaches. Furthermore, failure probabilities are highly sensitive to the assumed probabilistic distribution of input parameters (Ben-Haim, 1994; Elishakoff, 1995). This entails that the outcomes of the classical probabilistic reliability analysis may be considered accurate only when sufficient information is available to define the probability density function of the uncertain parameters. Nowadays, it is largely recognized that, if only vague, incomplete or fragmentary data are available, the non-probabilistic approaches (see e.g., Moens and Vandepitte, 2005; Elishakoff and Ohsaki, 2010) are more appropriate to obtain reliable predictions of the safety level. This issue has been first addressed by Ben-Haim (Ben-Haim, 1994) who presented a non-probabilistic concept of reliability in the framework of the convex model of uncertainty. Besides uncertainty affecting the structural parameters, the inherent random nature of environmental loads, such as

© 2021 *by F. Giunta, G. Muscolino, A. Sofi*

earthquake ground motion, sea waves or gusty winds (see e.g., Lutes and Sarkani, 1997; Roberts and Spanos, 2003; Li and Chen, 2009) has to be taken into account in the context of structural safety assessment. Studies relying on the traditional probabilistic uncertainty model have shown that reliability analysis becomes quite challenging when uncertainties affecting geometrical and/or mechanical properties and the random nature of excitations are simultaneously taken into account (see e.g., Gupta and Manohar, 2006; Goller et al., 2013).

To the best of the authors' knowledge, in literature only a few studies have focused on reliability analysis of structures subjected to random excitation with uncertain parameters described using non-probabilistic models. Among these, the most popular approach is the interval model which describes the uncertain parameters as interval variables with given lower bound and upper bound (Moore, 1966; Moore et al., 2009). Under this assumption, the statistics of structural response in the presence of random excitation have an interval nature and the measure of structural performance is provided by lower and upper values of the *interval reliability function*. In this framework, recently, Muscolino et al. (Muscolino et al., 2015; 2016a; 2016b) addressed the reliability analysis of linear discretized structures with interval stiffness properties subjected to stationary Gaussian random excitation by interval extension of the formulation of the *first-passage problem* (see e.g., Lutes and Sarkani, 1997). A selected displacement component has been assumed as critical response quantity and the uncertainties have been described by means of the *Improved Interval Analysis* via *Extra Unitary Interval* (*IIA* via *EUI*) (Muscolino and Sofi, 2012). The *IIA* via *EUI* has been introduced to limit the conservatism affecting computations based on the *Classical Interval Analysis* (*CIA*) (Moore, 1966; Moore et al., 2009). Such conservatism is caused by the so-called *dependency phenomenon* which is related to the inability of the *CIA* to treat multiple occurrences of the same interval variables in a mathematical expression as dependent ones.

In this paper, it is assumed that the structure fails as soon as a selected stress process at a critical location firstly exceeds a prescribed safe domain. In this case, the evaluation of the bounds of the *interval reliability function* is much more challenging due to the high overestimation of the range of stress-related interval functions which typically affects structural analysis based on the rules of the *CIA*. Indeed, stresses are more sensitive to the *dependency phenomenon* than displacements since their expression involves multiple occurrences of the same interval variables. Recently, this issue has been successfully addressed by Sofi et al. (2020) who proposed a sensitivity-based procedure for estimating the bounds of the *interval reliability function* for structures with interval axial stiffness. The main purpose of the present study is the extension of such sensitivity-based procedure to general finite element models involving uncertain stiffness properties. The formulation is developed in the context of the *first-passage* theory, under the Poisson assumption of independent up-crossings of a prescribed threshold (Lutes and Sarkani, 1997; Li and Chen, 2009). The issue of overestimation is tackled by describing interval uncertainties affecting the stiffness matrix of the structure through the *IIA* via *EUI* (Muscolino and Sofi, 2012). In order to ensure a conservative design, the worst case scenario corresponding to the lower bound of the *interval reliability function* has to be considered. The proposed method is able to identify through sensitivity analysis the set of uncertain parameters which lead to this scenario. Then, the lower bound of the *interval reliability function* is evaluated by performing a stochastic analysis of the randomly excited structure.

For validation purpose, a steel telecommunication antenna mast is selected as case-study. The accuracy of the proposed sensitivity-based procedure is demonstrated through appropriate comparisons with a combinatorial procedure, known as *vertex method* (Dong and Shah, 1987), which yields the exact bounds of the solution in the case of monotonic problems.

2. Problem Formulation

2.1. EQUATIONS OF MOTION

Let us consider a linear-elastic structure subjected to a stationary multi-correlated Gaussian stochastic process $\mathbf{F}(t)$. The structure is discretized into $N^{(e)}$ *finite elements* (*FE*s) resulting into a n-DOFs model. Young's modulus of the i-th *FE* is assumed uncertain and is described as an interval variable by means of the *Improved Interval Analysis* (*IIA*) via *Extra Unitary Interval* (*EUI*) (Muscolino and Sofi, 2012), i.e.

$$E^{(i)}(\alpha_i^I) = E_0^{(i)}\left(1+\alpha_i^I\right) = E_0^{(i)}\left(1+\Delta\alpha_i\hat{e}_i^I\right), \quad i=1,2,\ldots,r \tag{1}$$

where $r \le N^{(e)}$ denotes the number of *FE*s with uncertain stiffness; $\alpha_i^I = \Delta\alpha_i\hat{e}_i^I$ is a symmetric interval variable denoting the dimensionless fluctuation of Young's modulus around the nominal value $E_0^{(i)}$, with $\Delta\alpha_i$ the associated deviation amplitude, and $\hat{e}_i^I \triangleq [-1,+1]$ a particular unitary interval, called *EUI* (Muscolino and Sofi, 2012), which does not obey the rules of the *Classical Interval Analysis* (*CIA*) (Moore, 1966; Moore et al., 2009).

Taking into account Eq.(1), the elastic matrix of the i-th *FE* can be expressed as:

$$\mathbf{E}^{(i)}(\alpha_i^I) = \left(1+\Delta\alpha_i\hat{e}_i^I\right)\mathbf{E}_0^{(i)} \tag{2}$$

where $\mathbf{E}_0^{(i)}$ is the element elastic matrix with nominal Young's modulus $E_0^{(i)}$.

Let $\mathbf{x} = [x_1 \quad x_2 \quad x_3]^{\mathrm{T}}$ indicate the position vector of a generic point referred to a Cartesian coordinate system $O(x_1, x_2, x_3)$. Following the standard displacement-based *FE* formulation, the interval displacement field within the i-th *FE* can be approximated as follows (Sofi and Romeo, 2016; Sofi et al., 2019):

$$\mathbf{u}^{(i)}(\boldsymbol{\alpha}^I, \mathbf{x}, t) = \mathbf{N}^{(i)}(\mathbf{x})\,\mathbf{L}^{(i)}\,\mathbf{U}(\boldsymbol{\alpha}^I, t) \tag{3}$$

where $\mathbf{N}^{(i)}(\mathbf{x})$ denotes the shape-function matrix; and $\mathbf{L}^{(i)}$ a Boolean matrix defined so as to take into account the boundary conditions. In the previous equation, $\mathbf{U}(\boldsymbol{\alpha}^I, t)$ is the vector of global nodal displacements, which is an interval vector, depending both on time t and on the interval fluctuations collected into the vector $\boldsymbol{\alpha}^I = \left[\underline{\boldsymbol{\alpha}}, \overline{\boldsymbol{\alpha}}\right] \in \mathbb{IR}^r$, where the symbols $\underline{\boldsymbol{\alpha}}$ and $\overline{\boldsymbol{\alpha}}$ denote the vectors gathering the *lower bound* (*LB*) and *upper bound* (*UB*) of the symmetric interval parameters α_i^I, respectively. Due to the symmetry of α_i^I, the midpoint value vector of $\boldsymbol{\alpha}^I$ is null, while $\underline{\boldsymbol{\alpha}} \equiv -\Delta\boldsymbol{\alpha}$ and $\overline{\boldsymbol{\alpha}} \equiv \Delta\boldsymbol{\alpha}$, where the vector $\Delta\boldsymbol{\alpha}$ collects the deviation amplitudes $\Delta\alpha_i$ of α_i^I.

The strain-displacement equations and the linear-elastic constitutive equations yield the following expressions of the interval strain and stress fields within the i-th *FE*:

$$\begin{aligned} &\boldsymbol{\varepsilon}^{(i)}(\boldsymbol{\alpha}^I, \mathbf{x}, t) = \mathbf{B}^{(i)}(\mathbf{x})\,\mathbf{L}^{(i)}\,\mathbf{U}(\boldsymbol{\alpha}^I, t); \\ &\boldsymbol{\sigma}^{(i)}(\boldsymbol{\alpha}^I, \mathbf{x}, t) = \mathbf{E}^{(i)}(\alpha_i^I)\,\boldsymbol{\varepsilon}^{(i)}(\boldsymbol{\alpha}^I, \mathbf{x}, t) = \left(1+\Delta\alpha_i\hat{e}_i^I\right)\mathbf{E}_0^{(i)}\,\mathbf{B}^{(i)}(\mathbf{x})\,\mathbf{L}^{(i)}\,\mathbf{U}(\boldsymbol{\alpha}^I, t) \end{aligned} \tag{4a,b}$$

where $\mathbf{B}^{(i)}(\mathbf{x})$ is the strain-displacement matrix and the definition of the interval elastic matrix $\mathbf{E}^{(i)}(\alpha_i^I)$ in Eq. (2) has been taken into account. The stiffness matrix of the i-th *FE* is an interval matrix, formally analogous to the one pertaining to the deterministic *FE*, i.e.:

$$\mathbf{k}^{(i)}(\alpha_i^I) = \int_{V^{(i)}} \mathbf{B}^{(i)\mathrm{T}}(\mathbf{x})\,\mathbf{E}^{(i)}(\alpha_i^I)\,\mathbf{B}^{(i)}(\mathbf{x})\,\mathrm{d}V^{(i)} = \left(1+\Delta\alpha_i\hat{e}_i^I\right)\mathbf{k}_0^{(i)} \tag{5}$$

where $V^{(i)}$ is the volume of the i-th *FE*; $\mathbf{k}_0^{(i)} = \mathbf{k}^{(i)}(\alpha_i)\big|_{\alpha_i=0}$ is the nominal stiffness matrix and the interval elastic matrix is given by Eq. (2). In Eq. (5), the apex T denotes the transpose matrix operator.

Assuming for the sake of simplicity that all the quantities are referred to the global coordinate system, the standard assembly procedure yields the following interval global equations of motion:

$$\mathbf{M}\ddot{\mathbf{U}}(\boldsymbol{\alpha}^I,t) + \mathbf{C}(\boldsymbol{\alpha}^I)\dot{\mathbf{U}}(\boldsymbol{\alpha}^I,t) + \mathbf{K}(\boldsymbol{\alpha}^I)\mathbf{U}(\boldsymbol{\alpha}^I,t) = \mathbf{F}(t) \tag{6}$$

where a dot over a variable denotes differentiation with respect to time t; $\mathbf{M}$ is the mass matrix, and $\mathbf{K}(\boldsymbol{\alpha}^I)$ is the interval global stiffness defined as:

$$\mathbf{K}(\boldsymbol{\alpha}^I) = \mathbf{K}_0 + \sum_{i=1}^{r}\mathbf{K}_i\Delta\alpha_i\hat{e}_i^I; \quad \mathbf{K}_i = \left.\frac{\partial\mathbf{K}(\boldsymbol{\alpha}^I)}{\partial\alpha_i}\right|_{\boldsymbol{\alpha}=\mathbf{0}} = \mathbf{L}^{(i)\mathrm{T}}\mathbf{k}_0^{(i)}\mathbf{L}^{(i)} \tag{7a,b}$$

where $\mathbf{K}_0 = \mathbf{K}(\boldsymbol{\alpha})\big|_{\boldsymbol{\alpha}=\mathbf{0}}$ and $\mathbf{K}_i$ are the nominal and deviation stiffness matrices, respectively. By adopting the Rayleigh model, the global damping matrix $\mathbf{C}(\boldsymbol{\alpha}^I)$ turns out to be an interval matrix as well, defined as:

$$\mathbf{C}(\boldsymbol{\alpha}^I) = c_0\mathbf{M} + c_1\mathbf{K}(\boldsymbol{\alpha}^I) = \mathbf{C}_0 + c_0\mathbf{M} + c_1\sum_{i=1}^{r}\mathbf{K}_i\Delta\alpha_i\hat{e}_i^I \tag{8}$$

where $\mathbf{C}_0 = c_0\mathbf{M} + c_1\mathbf{K}_0$ is the nominal damping matrix; c_0 and c_1 denote the Rayleigh damping constants, herein evaluated setting the uncertain parameters equal to their nominal values.

2.2. INTERVAL GAUSSIAN STOCHASTIC STATIONARY RESPONSE PROCESS

The external load vector $\mathbf{F}(t)$ in Eq. (6) can be expressed as sum of the mean-value $\boldsymbol{\mu}_\mathbf{F} = \mathrm{E}\langle\mathbf{F}(t)\rangle$, with $\mathrm{E}\langle\bullet\rangle$ denoting the stochastic average operator, plus a zero-mean random fluctuating component $\tilde{\mathbf{X}}_\mathbf{F}(t)$, i.e. $\mathbf{F}(t) = \boldsymbol{\mu}_\mathbf{F} + \tilde{\mathbf{X}}_\mathbf{F}(t)$. Thus, in the frequency domain, the full probabilistic characterization of the external load vector $\mathbf{F}(t)$ requires the knowledge of the mean-value vector, $\boldsymbol{\mu}_\mathbf{F} = \mathrm{E}\langle\mathbf{F}(t)\rangle$, and of the one-sided *Power Spectral Density* (*PSD*) function matrix $\mathbf{G}_{\tilde{\mathbf{X}}_\mathbf{F}\tilde{\mathbf{X}}_\mathbf{F}}(\omega)$ of the fluctuating component $\tilde{\mathbf{X}}_\mathbf{F}(t)$.

The interval stationary Gaussian stochastic response process $\mathbf{U}(\boldsymbol{\alpha}^I,t)$, ruled by the equations of motion in Eq.(6), is completely characterized in the frequency domain by the interval mean-value vector:

$$\boldsymbol{\mu}_\mathbf{U}(\boldsymbol{\alpha}^I) = \mathrm{E}\left\langle\mathbf{U}^I(t)\right\rangle = \mathbf{K}^{-1}(\boldsymbol{\alpha}^I)\boldsymbol{\mu}_\mathbf{F} \tag{9}$$

and by the interval one-sided *PSD* function matrix, $\mathbf{G}_{\mathbf{UU}}(\boldsymbol{\alpha}^I,\omega)$, defined as follows:

$$\mathbf{G}_{\mathbf{UU}}(\boldsymbol{\alpha}^I,\omega) = \mathbf{H}^*(\boldsymbol{\alpha}^I,\omega)\mathbf{G}_{\tilde{\mathbf{X}}_\mathbf{F}\tilde{\mathbf{X}}_\mathbf{F}}(\omega)\mathbf{H}^{\mathrm{T}}(\boldsymbol{\alpha}^I,\omega) \tag{10}$$

where the asterisk means complex conjugate, and $\mathbf{H}(\boldsymbol{\alpha}^I,\omega)$ is the interval *Frequency Response Function* (*FRF*) matrix given by:

$$\mathbf{H}(\boldsymbol{\alpha}^I,\omega) = \left[-\omega^2\mathbf{M} + \mathrm{j}\omega\mathbf{C}(\boldsymbol{\alpha}^I) + \mathbf{K}(\boldsymbol{\alpha}^I)\right]^{-1} \tag{11}$$

with $\mathrm{j} = \sqrt{-1}$ denoting the imaginary unit.

Attention is herein focused on the j-th component of the interval stationary Gaussian random stress vector process $\boldsymbol{\sigma}^{(h)}(\boldsymbol{\alpha}^I,\mathbf{x},t)$ (see Eq.4b) at a given position $\mathbf{x}$ within the h-th *FE*:

$$Y_j^{(h)}(\boldsymbol{\alpha}^I,t) \equiv \sigma_j^{(h)}(\boldsymbol{\alpha}^I,\mathbf{x},t) = \left(1+\Delta\alpha_h\hat{e}_h^I\right)\mathbf{r}_j^{(h)\mathrm{T}}(\mathbf{x})\mathbf{U}(\boldsymbol{\alpha}^I,t) \tag{12}$$

where $\mathbf{r}_j^{(h)\mathrm{T}}(\mathbf{x})$ is the j-th row of the $n\times n$ matrix $\mathbf{E}_0^{(h)}\mathbf{B}^{(h)}(\mathbf{x})\mathbf{L}^{(h)}$. To simplify the notation, the dependence of $\mathbf{r}_j^{(h)}$ on $\mathbf{x}$ is hereinafter omitted since the stress is evaluated at a given position. By inspection of Eq.(12), it is readily inferred that the interval variable $\alpha_h^I = \Delta\alpha_h \hat{e}_h^I$ occurs more than once. It follows that quantities related to interval stresses are more vulnerable to the *dependency phenomenon* than displacements.

The interval stationary Gaussian stress random process in Eq.(12) can be expressed as sum of the interval mean-value, $\mu_{Y_j^{(h)}}(\boldsymbol{\alpha}^I)$, plus a zero-mean random fluctuation, $\tilde{Y}_j^{(h)}(\boldsymbol{\alpha}^I,t)$, i.e. $Y_j^{(h)}(\boldsymbol{\alpha}^I,t) = \mu_{Y_j^{(h)}}(\boldsymbol{\alpha}^I) + \tilde{Y}_j^{(h)}(\boldsymbol{\alpha}^I,t)$. Its complete probabilistic characterization in the frequency domain thus requires the knowledge of the interval mean-value, $\mu_{Y_j^{(h)}}(\boldsymbol{\alpha}^I)$, and of the interval one-sided *PSD* function, $G_{\tilde{Y}_j^{(h)}\tilde{Y}_j^{(h)}}(\boldsymbol{\alpha}^I,\omega) \equiv G_{Y_j^{(h)}Y_j^{(h)}}(\boldsymbol{\alpha}^I,\omega)$, of the zero-mean random fluctuation process $\tilde{Y}_j^{(h)}(\boldsymbol{\alpha}^I,t)$. The interval mean-value $\mu_{Y_j^{(h)}}(\boldsymbol{\alpha}^I)$ can be evaluated by applying the stochastic average operator to both sides of Eq.(12), i.e.:

$$\mu_{Y_j^{(h)}}(\boldsymbol{\alpha}^I) = \mathrm{E}\left\langle Y_j^{(h)}(\boldsymbol{\alpha}^I,t)\right\rangle = \left(1+\Delta\alpha_h\hat{e}_h^I\right)\mathbf{r}_j^{(h)\mathrm{T}}\boldsymbol{\mu}_{\mathbf{U}}(\boldsymbol{\alpha}^I) \tag{13}$$

where $\boldsymbol{\mu}_{\mathbf{U}}(\boldsymbol{\alpha}^I)$ is the interval mean-value of the displacement vector given in Eq. (9).

Based on Eqs. (10) and (12), the interval one-sided *PSD* function $G_{Y_j^{(h)}Y_j^{(h)}}(\boldsymbol{\alpha}^I,\omega)$ of the interval stress random process $Y_j^{(h)}(\boldsymbol{\alpha}^I,t)$ takes the following form:

$$G_{Y_j^{(h)}Y_j^{(h)}}(\boldsymbol{\alpha}^I,\omega) = \left(1+\Delta\alpha_h\hat{e}_h^I\right)^2 \mathbf{r}_j^{(h)\mathrm{T}}\,\mathbf{H}^*(\boldsymbol{\alpha}^I,\omega)\mathbf{G}_{\tilde{\mathbf{X}}_\mathbf{F}\tilde{\mathbf{X}}_\mathbf{F}}(\omega)\mathbf{H}^{\mathrm{T}}(\boldsymbol{\alpha}^I,\omega)\mathbf{r}_j^{(h)}. \tag{14}$$

Notice that the same interval variables occur many times in the previous equation. This implies that the statistics of the interval stress random process may be affected by very serious overestimation due to the *dependency phenomenon*.

3. Bounds of the Interval Reliability Function

3.1. INTERVAL RELIABILITY FUNCTION

Failure or unsatisfactory performance of a structural system is herein identified with the *first-passage failure*, which occurs when the *extreme value* random process for some response measure (e.g., displacement, strain or stress) firstly exceeds a prescribed safe domain within a specified time interval $[0,T]$. Specifically, it is assumed that the structure fails in a *first-passage* sense if the j-th component of the interval stationary Gaussian random stress vector process $\boldsymbol{\sigma}^{(h)}(\boldsymbol{\alpha}^I,\mathbf{x},t)$ (see Eq.4b) at a given position $\mathbf{x}$ within the h-th *FE*, i.e. $Y_j^{(h)}(\boldsymbol{\alpha}^I,t)$ (see Eq. (12)), reaches a prescribed threshold.

The *extreme value* random process of $Y_j^{(h)}(\boldsymbol{\alpha}^I,t)$, over the time interval $[0,T]$, has an interval nature and is mathematically defined as:

$$Y_{j,\max}^{(h)}(\boldsymbol{\alpha}^I,T) = \max_{0\le t\le T}\left|Y_j^{(h)}(\boldsymbol{\alpha}^I,t)\right| \tag{15}$$

where the symbol $|\bullet|$ denotes absolute value. The probability that $Y_{j,\max}^{(h)}(\boldsymbol{\alpha}^I,T)$ is equal to or less than the critical level $b>0$ within the time interval $[0,T]$ is defined by the *cumulative distribution function* (*CDF*), also called *reliability function*, which has an interval nature as well:

$$L_{Y_{j,\max}^{(h)}}(\boldsymbol{\alpha}^I,b,T)=\Pr\left[Y_{j,\max}^{(h)}(\boldsymbol{\alpha}^I,T)\le b\right]=\left[\underline{L}_{Y_{j,\max}^{(h)}}(b,T),\quad \overline{L}_{Y_{j,\max}^{(h)}}(b,T)\right]. \tag{16}$$

The *LB* (or right bound) and *UB* (or left bound) of the interval *CDF* define a probability box (p-box) (Ferson et al., 2003) representative of the range of structural performance under prescribed variations of the uncertain parameters within their respective intervals.

If the Poisson assumption of independent up-crossings of a prescribed threshold is applied, then the *interval CDF* for unit initial probability can be expressed as (see e.g., Lutes and Sarkani, 1997):

$$L_{Y_{j,\max}^{(h)}}(\boldsymbol{\alpha}^I,b,T)=\Pr\left[Y_{j,\max}^{(h)}(\boldsymbol{\alpha}^I,T)\le b\right]\approx\exp\left\{-T\,\nu_{Y_j^{(h)}}^{+}(\boldsymbol{\alpha}^I)\exp\left[-\frac{\left(b-\left|\mu_{Y_j^{(h)}}(\boldsymbol{\alpha}^I)\right|\right)^2}{2\lambda_{0,Y_j^{(h)}}(\boldsymbol{\alpha}^I)}\right]\right\} \tag{17}$$

where

$$\nu_{Y_j^{(h)}}^{+}(\boldsymbol{\alpha}^I)=\frac{1}{2\,\pi}\sqrt{\frac{\lambda_{2,Y_j^{(h)}}(\boldsymbol{\alpha}^I)}{\lambda_{0,Y_j^{(h)}}(\boldsymbol{\alpha}^I)}} \tag{18}$$

is the mean up-crossing rate at level $\left|\mu_{Y_j^{(h)}}(\boldsymbol{\alpha}^I)\right|$; $\lambda_{0,Y_j^{(h)}}(\boldsymbol{\alpha}^I)$ and $\lambda_{2,Y_j^{(h)}}(\boldsymbol{\alpha}^I)$ are the interval spectral moments of zero- and second-order, respectively, of the interval stress random process, $Y_j^{(h)}(\boldsymbol{\alpha}^I,t)$, given by

$$\lambda_{\ell,Y_j^{(h)}}(\boldsymbol{\alpha}^I)=\int_0^\infty \omega^\ell\, G_{Y_j^{(h)}Y_j^{(h)}}(\boldsymbol{\alpha}^I,\omega)\,\mathrm{d}\omega,\quad(\ell=0,2) \tag{19}$$

where $G_{Y_j^{(h)}Y_j^{(h)}}(\boldsymbol{\alpha}^I,\omega)$ is the one-sided interval *PSD* function of $Y_j^{(h)}(\boldsymbol{\alpha}^I,t)$ defined in Eq.(14).

3.2. PROPOSED SENSITIVITY-BASED PROCEDURE

In the context of the *first-passage* theory, the aim of interval reliability analysis is the evaluation of the *LB* and *UB* of the *interval reliability function* defined by Eq. (17). Since interval stress-related quantities are more affected by overestimation than displacements, the bounds of the *interval reliability function* of the *extreme value* stress random process $Y_{j,\max}^{(h)}(\boldsymbol{\alpha}^I,T)$ are herein evaluated by applying a sensitivity-based procedure.

The key idea of this procedure is to perform a preliminary sensitivity analysis to identify suitable combinations of the values of the interval parameters which provide accurate estimates of the *LB* and *UB* of the interval *CDF*. This approach requires the evaluation of the sensitivity of the *CDF* of the random process $Y_{j,\max}^{(h)}(\boldsymbol{\alpha}^I,T)$ with respect to the uncertain parameters $\alpha_i\in\alpha_i^I$ $(i=1,2,\ldots,r)$ by direct differentiation of Eq. (17) which leads to (Sofi et al., 2020):

$$S_{L_{Y_{j,\max}^{(h)}},i}(b,T)=\left.\frac{\partial L_{Y_{j,\max}^{(h)}}(\boldsymbol{\alpha},b,T)}{\partial\alpha_i}\right|_{\boldsymbol{\alpha}=\mathbf{0}}$$
$$=C_{Y_j^{(h)}}(b,T)\left\{2\left(b-\left|\mu_{Y_j^{(h)}}^{(0)}\right|\right)\frac{\left|\mu_{Y_j^{(h)}}^{(0)}\right|}{\mu_{Y_j^{(h)}}^{(0)}}S_{\mu_{Y_j^{(h)}},i}+\left[\frac{\left(b-\left|\mu_{Y_j^{(h)}}^{(0)}\right|\right)^2}{\lambda_{0,Y_j^{(h)}}^{(0)}}-1\right]S_{\lambda_{0,Y_j^{(h)}},i}+\frac{\lambda_{0,Y_j^{(h)}}^{(0)}}{\lambda_{2,Y_j^{(h)}}^{(0)}}S_{\lambda_{2,Y_j^{(h)}},i}\right\} \tag{20}$$

where $\mu_{Y_j^{(h)}}^{(0)}$ and $\lambda_{\ell,Y_j^{(h)}}^{(0)}$ $(\ell=0,2)$ denote the nominal mean-value and spectral moments of the selected stress process, given by Eqs. (13) and (19) with $\boldsymbol{\alpha}=\mathbf{0}$; the function $C_{Y_j^{(h)}}(b,T)$ is defined as follows:

$$C_{Y_j^{(h)}}(b,T)=-\frac{T}{4\pi}\frac{\lambda_{2,Y_j^{(h)}}^{(0)}}{\lambda_{0,Y_j^{(h)}}^{(0)}}\frac{L_{Y_{j,\max}^{(h)}}^{(0)}(b,T)}{\sqrt{\lambda_{0,Y_j^{(h)}}^{(0)}\,\lambda_{2,Y_j^{(h)}}^{(0)}}}\exp\left[-\frac{\left(b-\left|\mu_{Y_j^{(h)}}^{(0)}\right|\right)^2}{2\lambda_{0,Y_j^{(h)}}^{(0)}}\right] \tag{21}$$

with $L_{Y_{j,\max}^{(h)}}^{(0)}(b,T)=L_{Y_{j,\max}^{(h)}}(\boldsymbol{\alpha},b,T)\Big|_{\boldsymbol{\alpha}=\mathbf{0}}$ denoting the nominal *CDF*. Furthermore, in Eq.(20) $S_{\mu_{Y_j^{(h)}},i}$ is the sensitivity of the mean-value $\mu_{Y_j^{(h)}}(\boldsymbol{\alpha})$ of the interval stress random process (see Eq. (13)) with respect to the uncertain parameter α_i, which can be evaluated as:

$$S_{\mu_{Y_j^{(h)}},i}=\begin{cases}\left.\dfrac{\partial\mu_{Y_j^{(h)}}(\boldsymbol{\alpha})}{\partial\alpha_i}\right|_{\boldsymbol{\alpha}=\mathbf{0}}=\mu_{Y_j^{(h)}}^{(0)}-\mathbf{r}_j^{(h)\mathrm{T}}\mathbf{K}_0^{-1}\mathbf{K}_h\mathbf{K}_0^{-1}\boldsymbol{\mu}_{\mathbf{F}}, & \text{if } i=h\\ \left.\dfrac{\partial\mu_{Y_j^{(h)}}(\boldsymbol{\alpha})}{\partial\alpha_i}\right|_{\boldsymbol{\alpha}=\mathbf{0}}=-\mathbf{r}_j^{(h)\mathrm{T}}\mathbf{K}_0^{-1}\mathbf{K}_i\mathbf{K}_0^{-1}\boldsymbol{\mu}_{\mathbf{F}}, & \text{if } i\neq h\end{cases} \tag{22a,b}$$

where $\mathbf{K}_0$ is the nominal stiffness matrix; $\mathbf{K}_i$ is given by Eq. (7b); and $\mathbf{K}_0^{-1}\mathbf{K}_i\mathbf{K}_0^{-1}\boldsymbol{\mu}_{\mathbf{F}}$ is the $i-$th sensitivity of $\boldsymbol{\mu}_{\mathbf{U}}(\boldsymbol{\alpha})$ (derived by Eq. (9)). Finally, in Eq.(20), $S_{\lambda_{\ell,Y_j^{(h)}},i}$ denotes the sensitivity of the spectral moment of order ℓ of the interval stress random process $Y_j^{(h)}(\boldsymbol{\alpha}^I,t)$ with respect to the $i-$th parameter α_i:

$$S_{\lambda_{\ell,Y_j^{(h)}},i}=\left.\frac{\partial\lambda_{\ell,Y_j^{(h)}}(\boldsymbol{\alpha})}{\partial\alpha_i}\right|_{\boldsymbol{\alpha}=\mathbf{0}}=\int_0^\infty\omega^\ell S_{G_{Y_j^{(h)}Y_j^{(h)},i}}(\omega)\,\mathrm{d}\omega,\quad(\ell=0,2). \tag{23}$$

In the previous expression, $S_{G_{Y_j^{(h)}Y_j^{(h)},i}}(\omega)$ is the $i-$th sensitivity of the one-sided *PSD* function $G_{Y_j^{(h)}Y_j^{(h)}}(\boldsymbol{\alpha}^I,\omega)$ (see Eq. (14)) which can be evaluated as:

$$S_{G_{Y_j^{(h)}Y_j^{(h)}},i}(\omega)=\begin{cases}\left.\dfrac{\partial G_{Y_j^{(h)}Y_j^{(h)}}(\boldsymbol{\alpha},\omega)}{\partial\alpha_i}\right|_{\boldsymbol{\alpha}=\mathbf{0}}=2G^{(0)}_{Y_j^{(h)}Y_j^{(h)}}(\omega)+\mathbf{r}_j^{(h)\mathrm{T}}\mathbf{P}_h(\omega)\mathbf{r}_j^{(h)}, & \text{if } i=h;\\ \left.\dfrac{\partial G_{Y_j^{(h)}Y_j^{(h)}}(\boldsymbol{\alpha},\omega)}{\partial\alpha_i}\right|_{\boldsymbol{\alpha}=\mathbf{0}}=\mathbf{r}_j^{(h)\mathrm{T}}\mathbf{P}_i(\omega)\mathbf{r}_j^{(h)}, & \text{if } i\neq h\end{cases} \tag{24a,b}$$

where $G^{(0)}_{Y_j^{(h)}Y_j^{(h)}}(\omega)$ is the nominal one-sided *PSD* function of the selected stress process, given by Eq. (14) with $\boldsymbol{\alpha}=\mathbf{0}$. Furthermore, in the previous equations the matrix $\mathbf{P}_i(\omega)$ is defined as:

$$\mathbf{P}_i(\omega)=\mathbf{S}_i^*(\omega)\mathbf{G}_{\tilde{\mathbf{X}}_\mathbf{F}\tilde{\mathbf{X}}_\mathbf{F}}(\omega)\mathbf{H}_0^{\mathrm{T}}(\omega)+\mathbf{H}_0^*(\omega)\mathbf{G}_{\tilde{\mathbf{X}}_\mathbf{F}\tilde{\mathbf{X}}_\mathbf{F}}(\omega)\mathbf{S}_i^{\mathrm{T}}(\omega) \tag{25}$$

with

$$\mathbf{S}_i(\omega)=\left.\frac{\partial\mathbf{H}(\boldsymbol{\alpha},\omega)}{\partial\alpha_i}\right|_{\boldsymbol{\alpha}=\mathbf{0}}=-(1+\mathrm{j}c_1\omega)\,\mathbf{H}_0(\omega)\,\mathbf{K}_i\,\mathbf{H}_0(\omega) \tag{26}$$

where $\mathbf{K}_i$ is given by Eq.(7b), and $\mathbf{H}_0(\omega)$ is the nominal *FRF* matrix given as:

$$\mathbf{H}_0(\omega)=\left[-\omega^2\mathbf{M}+\mathrm{j}\omega\mathbf{C}_0+\mathbf{K}_0\right]^{-1}. \tag{27}$$

The knowledge of the sensitivity $S_{L_{Y_{j,\max}^{(h)}},i}$ defined in Eq.(20) allows us to predict the influence of a small variation of the $i-$th uncertain parameter α_i on the *interval reliability function* $L_{Y_{j,\max}^{(h)}}(\boldsymbol{\alpha}^I,b,T)$. Specifically, within a small range around $\boldsymbol{\alpha}=\mathbf{0}$, $L_{Y_{j,\max}^{(h)}}(\boldsymbol{\alpha}^I,b,T)$ is a monotonic increasing or decreasing function of α_i depending on whether $S_{L_{Y_{j,\max}^{(h)}},i}>0$ or $S_{L_{Y_{j,\max}^{(h)}},i}<0$, and its bounds, therefore, correspond to suitable combinations of the endpoints of the uncertain parameters. Relying on the monotonic increasing or decreasing behaviour predicted by studying the sign of sensitivities, the combinations of the extreme values of the uncertain parameters, which yield accurate estimates of the *LB* and *UB* of the *interval reliability function* $L_{Y_{j,\max}^{(h)}}(\boldsymbol{\alpha}^I,b,T)$, denoted as $\alpha^{(\mathrm{LB})}_{Y_{j,\max}^{(h)},i}$ and $\alpha^{(\mathrm{UB})}_{Y_{j,\max}^{(h)},i}$, $(i=1,2,\ldots,r)$, can be identified as follows:

$$\begin{aligned}&\text{if } S_{L_{Y_{j,\max}^{(h)}},i}>0, \text{ then } \alpha^{(\mathrm{UB})}_{Y_{j,\max}^{(h)},i}=\overline{\alpha}_i,\quad \alpha^{(\mathrm{LB})}_{Y_{j,\max}^{(h)},i}=\underline{\alpha}_i;\\ &\text{if } S_{L_{Y_{j,\max}^{(h)}},i}<0, \text{ then } \alpha^{(\mathrm{UB})}_{Y_{j,\max}^{(h)},i}=\underline{\alpha}_i,\quad \alpha^{(\mathrm{LB})}_{Y_{j,\max}^{(h)},i}=\overline{\alpha}_i.\end{aligned} \tag{28a,b}$$

Such combinations can be collected into the following two vectors:

$$\begin{aligned}\boldsymbol{\alpha}^{(\mathrm{LB})}_{Y_{j,\max}^{(h)}}&=\left[\alpha^{(\mathrm{LB})}_{Y_{j,\max}^{(h)},1}\quad\alpha^{(\mathrm{LB})}_{Y_{j,\max}^{(h)},2}\quad\ldots\quad\alpha^{(\mathrm{LB})}_{Y_{j,\max}^{(h)},r}\right]^{\mathrm{T}};\\ \boldsymbol{\alpha}^{(\mathrm{UB})}_{Y_{j,\max}^{(h)}}&=\left[\alpha^{(\mathrm{UB})}_{Y_{j,\max}^{(h)},1}\quad\alpha^{(\mathrm{UB})}_{Y_{j,\max}^{(h)},2}\quad\ldots\quad\alpha^{(\mathrm{UB})}_{Y_{j,\max}^{(h)},r}\right]^{\mathrm{T}}.\end{aligned} \tag{29a,b}$$

Finally, the *LB* and *UB* of the *interval reliability function* for the *extreme value* stress random process $Y_{j,\max}^{(h)}(\boldsymbol{\alpha}^I,T)$ can be obtained by evaluating Eq.(17) for $\boldsymbol{\alpha}=\boldsymbol{\alpha}^{(\mathrm{LB})}_{Y_{j,\max}^{(h)}}$ and $\boldsymbol{\alpha}=\boldsymbol{\alpha}^{(\mathrm{UB})}_{Y_{j,\max}^{(h)}}$, respectively:

$$\underline{L}_{Y_{j,\max}^{(h)}}(b,T) = L_{Y_{j,\max}^{(h)}}(\boldsymbol{\alpha},b,T)\Big|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_{Y_{j,\max}^{(h)}}^{(\mathrm{LB})}}; \qquad \bar{L}_{Y_{j,\max}^{(h)}}(b,T) = L_{Y_{j,\max}^{(h)}}(\boldsymbol{\alpha},b,T)\Big|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_{Y_{j,\max}^{(h)}}^{(\mathrm{UB})}}. \tag{30a,b}$$

Summarizing, the sensitivity-based procedure requires only two stochastic analyses of the structure for assigned values of the uncertain parameters given by Eqs (29a,b) in order to evaluate the mean-value and spectral moments of zero- and second-order of the interval random stress process $Y_j^{(h)}(\boldsymbol{\alpha}^I,t)$ entering the definition of the *CDF* (see Eq. (17)).

Under the assumption that the spectral moments are monotonic functions of the interval parameters, the proposed procedure requires preliminarily the evaluation of the interval global stiffness matrix (see Eqs.(7)), and of the interval damping matrix (see Eq.(8)). Then, the following sensitivities for the *extreme value* stress random process $Y_{j,\max}^{(h)}(\boldsymbol{\alpha}^I,T)$, with respect to the uncertain parameters $\alpha_i \in \alpha_i^I$ $(i=1,2,\ldots,r)$ are required: i) sensitivities of mean-value (see Eqs.(22)), ii) sensitivities of spectral moments (see Eq.(23)); iii) sensitivities of the *CDF* (see Eq. (20)). Finally, once the sensitivities of the *CDF* are determined, the vector $\boldsymbol{\alpha}_{Y_{j,\max}^{(h)}}^{(\mathrm{LB})}$ (see Eq. (29a)) has to be defined in order to evaluate the *LB* of the *interval CDF* (see Eq. (30a)) which defines the worst case scenario for a conservative design.

The knowledge of the sensitivities of the *interval CDF* of the *extreme value* stress random process $Y_{j,\max}^{(h)}(\boldsymbol{\alpha}^I,T)$ in Eq. (20) can also be exploited to enhance the computational efficiency of the proposed procedure. As known, sensitivity analysis allows us to identify the most influential parameters on the response quantity of interest. To this aim, the so-called *function of sensitivity* of the *interval CDF*, $L_{Y_{j,\max}^{(h)}}(\boldsymbol{\alpha}^I,b,T)$, is evaluated:

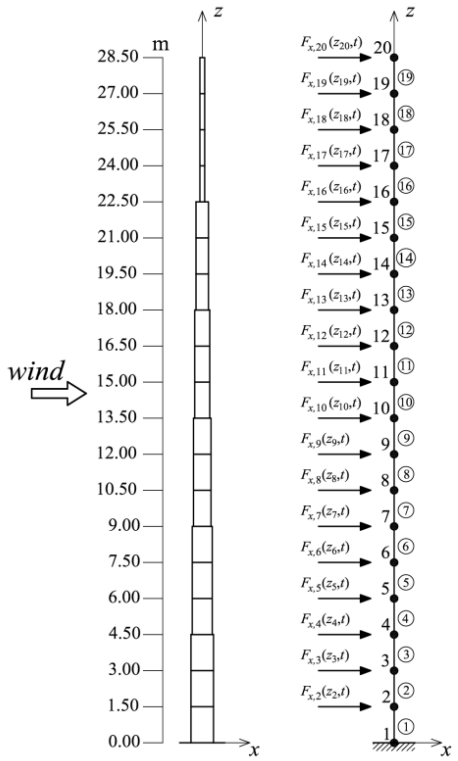
$$\varphi_{i,L_{Y_{j,\max}^{(h)}}}(b,T)(\%) = \frac{S_{L_{Y_{j,\max}^{(h)}},i}(b,T)}{L_{Y_{j,\max}^{(0)}}^{(0)}(b,T)}\Delta\alpha_i \times 100, \tag{31}$$

where $S_{L_{Y_{j,\max}^{(h)}},i}(b,T)$ is the $i-\mathrm{th}$ sensitivity of the *interval CDF*, $L_{Y_{j,\max}^{(h)}}(\boldsymbol{\alpha}^I,b,T)$, defined in Eq. (20); $L_{Y_{j,\max}^{(h)}}^{(0)}(b,T)$ is the *CDF* pertaining to the nominal system. The *function of sensitivity* represents a percentage measure of the influence of the generic interval variable α_i^I on the *CDF* of the selected *extreme value* stress process. This implies that the crucial uncertain parameters are those characterized by higher values of the *function of sensitivity*. The least influential parameters can be reasonably assumed deterministic and set equal to their nominal values.

4. Numerical Application

The effectiveness of the proposed procedure is assessed by performing *first-passage* reliability analysis of a steel telecommunication antenna mast with interval stiffness uncertainties subjected to wind excitation modelled as a stationary Gaussian multi-correlated random process.

The steel telecommunication antenna mast, 28.50 m high, depicted in Figure1a, is discretized into $N^{(e)}=19$ two-node Euler-Bernoulli beam *FE*s resulting in a $n=38$ DOFs system (Fig. 1b). A lumped mass model is assumed. The properties of the *FE* model of the antenna are listed in Table I, where the masses lumped at nodes and the tributary areas A_i entering the definition of wind loads $F_{x,i}(z_i,t)$ (see Eq. (32)) are also reported. The nominal Young's modulus is set equal to E_0=210 GPa for the whole structure. The values $c_0=0.149575\,\mathrm{s}^{-1}$ and $c_1=0.00316358\,\mathrm{s}$ are assumed for the Rayleigh damping constants in Eq. (8) in such a way that the modal damping ratio of the first and third modes of the nominal structure is $\zeta_0=0.01$. The fundamental period of the nominal structure is $T_0=0.723\,\mathrm{s}$.



a) b)

Figure 1. Steel telecommunication antenna mast under wind excitation: a) 2D model; b) *FE* model.

Table I. Properties of the telecommunication antenna mast					
Node	**Height [m]**	**Outer diameter [m]**	**Thickness [m]**	**Lumped mass [t]**	**Tributary area [m^2]**
2	1.50	0.9148	0.008	0.3050	1.3722
3	3.00	0.9148	0.008	0.3050	1.3722
4	4.50	0.9148	0.008	0.2754	1.2957
5	6.00	0.8128	0.0071	0.2458	1.2192
6	7.50	0.8128	0.0071	0.2458	1.2192
7	9.00	0.8128	0.0071	0.2211	1.143
8	10.50	0.7112	0.0063	0.1964	1.0668
9	12.00	0.7112	0.0063	0.1964	1.066
10	13.50	0.7112	0.0063	0.1760	0.9906
11	15.00	0.6096	0.0056	0.1530	0.9144
12	16.50	0.6096	0.0056	0.1505	0.9144
13	18.00	0.6096	0.0056	0.1288	0.8382
14	19.50	0.508	0.005	0.2671	0.762
15	21.00	0.508	0.005	0.1071	0.762
16	22.50	0.508	0.005	0.1650	0.526275
17	24.00	0.1937	0.0045	0.0429	0.29055
18	25.50	0.1937	0.0045	0.1329	0.29055
19	27.00	0.1937	0.0045	0.0429	0.29055
20	28.50	0.1937	0.0045	0.0214	0.145275

Young's modulus of the material is assumed to be affected by uncertainty. In a first stage, in order to allow comparisons with the *vertex method* (Dong and Shah, 1987), only the first $r=12$ *FE*s (see Fig. 1b) are assumed to be characterized by interval Young's moduli i.e. $E^{(i)I}=E_0\left(1+\Delta\alpha\hat{e}_i^I\right)$, $(i=1,2,\ldots,r=12)$, where the same deviation amplitude is considered. Under this assumption, the application of the *vertex method* involves $2^r=2^{12}$ stochastic analyses of the structure while the proposed method requires to evaluate the *reliability function* only twice, regardless of the number of uncertain parameters. Indeed, the *vertex method* requires to evaluate the *reliability function* for all the 2^r possible combinations of the

endpoints of the r uncertain parameters. Then, for each barrier level b, the *LB* and *UB* of the *interval reliability function* are obtained as the minimum and maximum among the 2^r values pertaining to the *vertex analysis*.

The maximum interval axial stress, $Y_1^{(1)}(\boldsymbol{\alpha}^I,t)$, at the antenna base section, i.e. at node 1 of *FE* 1, is assumed as the response quantity responsible of structural failure. To predict the range of structural performance, the bounds of the *interval reliability function*, $L_{Y_{1,\max}^{(1)}}(\boldsymbol{\alpha}^I,b,T)$, (see Eq. (17)) are evaluated. The observation time is assumed equal to $T=1000T_0$, T_0 being the fundamental period of the structure with nominal Young's moduli.

The along wind force (x-direction) exerted on the i-th node at height z_i of the discretized structure is defined by the well-known formula (Simiu and Scanlan, 1996):

$$F_{x,i}(z_i,t)=\frac{1}{2}\rho C_D A_i w_s^2(z_i)+\rho C_D A_i \tilde{W}(z_i,t) w_s(z_i) \tag{32}$$

where the contribution of the nodal velocities of the structure and the square of the fluctuating component of wind speed have been neglected. In Eq. (32), $\rho=1.25$ kg/m^3 is the air density; C_D is the drag coefficient; A_i is the tributary area of the i-th node; $w_s(z)=w_{s,10}\left(z/10\right)^{\gamma}$ is the mean wind velocity which is assumed to vary with the elevation z following a power law; $w_{s,10}$ is the mean wind speed measured at height $z=10$ m above ground and γ is a coefficient depending on surface roughness, herein taken equal to $w_{s,10}=25$ m/s and $\gamma=0.3$, respectively. Furthermore, $\tilde{W}(z,t)$ denotes the fluctuating component of wind velocity field, which is modelled as a zero-mean stationary Gaussian random field, fully described from a probabilistic point of view by the one-sided *PSD* function proposed by Davenport (1961):

$$G_{\tilde{W}\tilde{W}}(\omega)=4K_0 w_{s,10}^2\frac{\chi^2}{\omega\left(1+\chi^2\right)^{4/3}} \tag{33}$$

where K_0 is the non-dimensional roughness coefficient, herein set equal to $K_0=0.03$, and $\chi=b_1\omega/(\pi w_{s,10})$ with $b_1=600$ m. The vector $\tilde{\mathbf{W}}(t)$ collecting wind velocity fluctuations at the wind-exposed nodes located at different heights z_i is characterized from a probabilistic point of view by the one-sided *PSD* function matrix $\mathbf{G}_{\tilde{\mathbf{W}}\tilde{\mathbf{W}}}(\omega)$. If the imaginary part (*q*-spectrum) is neglected, the one-sided cross-*PSD* components of $\mathbf{G}_{\tilde{\mathbf{W}}\tilde{\mathbf{W}}}(\omega)$ can be expressed as follows (Simiu and Scanlan, 1996):

$$G_{\tilde{W}_i\tilde{W}_j}(z_i,z_j;\omega)=G_{\tilde{W}\tilde{W}}(\omega)\,f_{ij}(\omega)=G_{\tilde{W}\tilde{W}}(\omega)\mathrm{Exp}\left\{-\frac{|\omega|\sqrt{C_z^2(z_i-z_j)^2}}{\pi\left[w_s(z_i)+w_s(z_j)\right]}\right\} \tag{34}$$

where $f_{ij}(\omega)$ is the so-called *coherence function* with $C_z=10$ denoting an appropriate decay coefficient.

In Figure 2, the proposed *LB* and *UB* of $L_{Y_{1,\max}^{(1)}}(\boldsymbol{\alpha}^I,b,T)$ are compared with those provided by the *vertex method*. Two different deviation amplitudes of the uncertain parameters, $\Delta\alpha=0.10$ and $\Delta\alpha=0.20$, are considered. The nominal *CDF*, $L_{Y_{1,\max}^{(1)}}^{(0)}(b,T)$, is also reported. An excellent agreement between the proposed sensitivity-based procedure and the *vertex method* can be observed. As expected, the p-box describing structural performance becomes wider as the degree of uncertainty increases. The deviation of the left and right bounds of the p-box from the nominal solution proves that assuming the nominal value of Young's

moduli may lead to inaccurate predictions of the structural safety level. As already mentioned, a conservative design should rely on the *LB* of the *interval reliability function*.

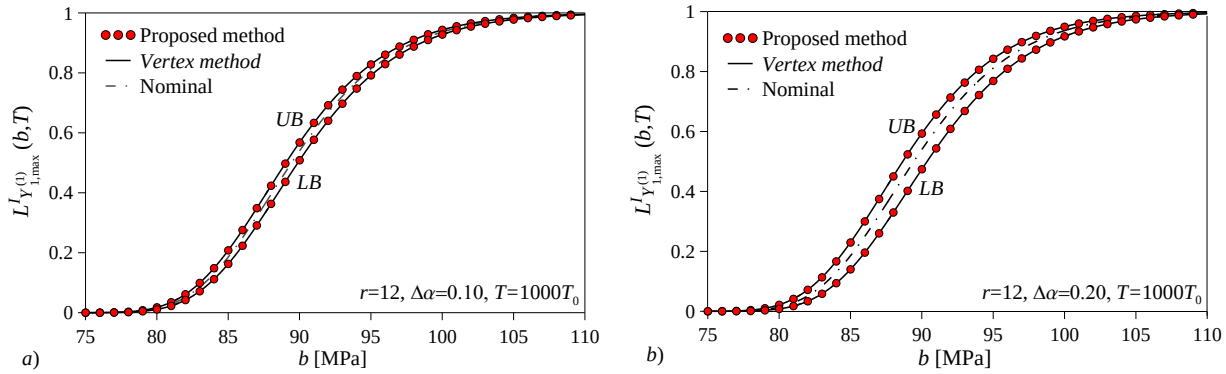


Figure 2. Bounds of the *interval CDF* of the *extreme value* axial stress process $Y_{1,\max}^{(1)I}(T)$ of the telecommunication antenna with uncertain Young's moduli $E^{(i)I} = E_0\left(1+\Delta\alpha\hat{e}_i^I\right)$, $(i=1,2,\ldots,r=12)$: comparison between the proposed procedure and the *vertex method* for *a*) $\Delta\alpha = 0.10$ and *b*) $\Delta\alpha = 0.20$.

In order to predict the influence of a small change of Young's moduli on the performance of the telecommunication antenna mast, the *function of sensitivity* of the *CDF* $L_{Y_{1,\max}^{(1)}}(\boldsymbol{\alpha}^I, b, T)$ is evaluated (see Eq. (31)) under the assumption that all Young's moduli are described by intervals. In Figure 3, the *functions of sensitivity* $\varphi_{i,L_{Y_{1,\max}^{(1)}}}(b,T)$ of $L_{Y_{1,\max}^{(1)}}(\boldsymbol{\alpha}^I, b, T)$ with respect to the fluctuations $\alpha_i^I = \Delta\alpha\hat{e}_i^I$ of a selected number of interval Young's moduli $E_i^I = E_0(1+\alpha_i^I) = E_0(1+\Delta\alpha\hat{e}_i^I)$, $(i=1,2,\ldots,8,12,13,14,16,17)$ versus the deterministic barrier level b are plotted ($\Delta\alpha = 0.10$).

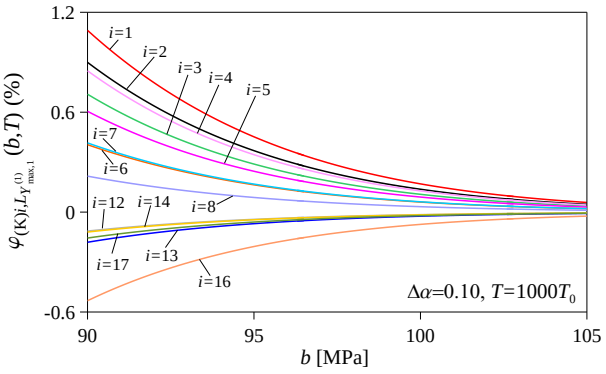


Figure 3. Functions of sensitivity of the *interval reliability function* of the *extreme value* axial stress process $Y_{1,\max}^{(1)I}(T)$ of the telecommunication antenna with respect to the fluctuations of Young's moduli $E^{(i)I} = E_0\left(1+\Delta\alpha\hat{e}_i^I\right)$, $(i=1,2,\ldots,8,12,13,14,16,17)$, versus the deterministic barrier level b ($\Delta\alpha = 0.10$, $T = 1000T_0$).

For the sake of clarity, the *functions of sensitivity* with respect to the fluctuations of the remaining Young's moduli are omitted. It is observed that the most influential Young's moduli are those of *FE*s 1 and 16 for any value of the barrier level. A close inspection of Figure 3 also shows that a small increase of Young's moduli of the first eight *FE*s would produce an increment of structural reliability since the pertinent *functions of sensitivity* are positive. Conversely, a small increase of Young's moduli of *FE*s 12,13,14,16,17 would lead to a lower safety level.

Finally, in Figure 4, the bounds of the *interval reliability function* $L^I_{Y^{(1)}_{1,\max}}(b,T)$ evaluated by applying the proposed sensitivity-based approach considering Young's moduli of all the $r = N^{(e)} = 19$ *FE*s as uncertain (full) are contrasted with the ones computed retaining only the $r = 13$ most influential uncertain parameters (reduced) identified by sensitivity analysis. For comparison purpose, the bounds pertaining to the structure with the first $r = 6$ most influential uncertain Young's moduli $E^{(i)I} = E_0\left(1+\Delta\alpha\hat{e}_i^I\right)$, $(i = 1,2,3,4,5,16)$, are also plotted. Also in this case, two different deviation amplitudes of the uncertain parameters, $\Delta\alpha = 0.10$ and $\Delta\alpha = 0.20$, are considered. It can be observed that the left and right bounds of the p-box describing structural performance obtained considering only the first $r = 6$ uncertain parameters are enclosed by the bounds pertaining to full uncertainty. This entails that some of the neglected parameters play a crucial role on the prediction of the safety level. Conversely, the region of the *interval CDF* predicted retaining $r = 13$ uncertain Young's moduli is almost coincident with the one obtained performing full uncertainty analysis.

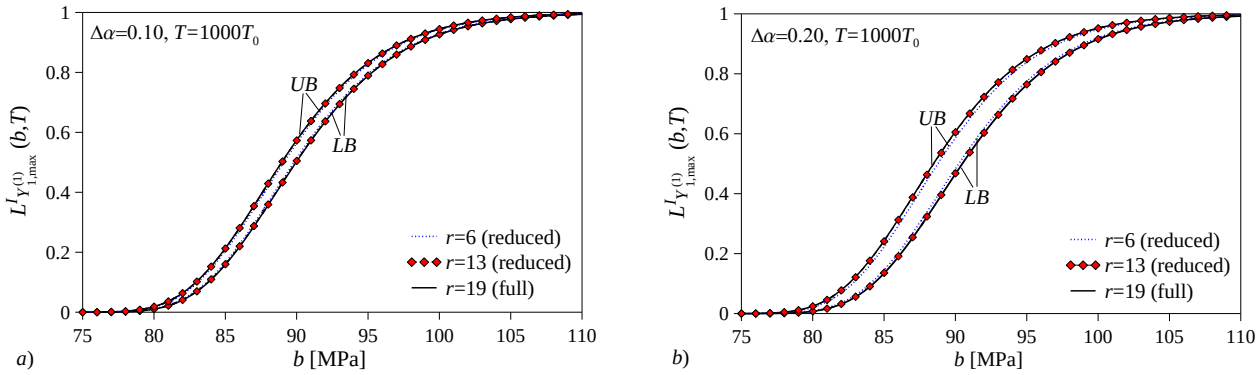


Figure 4. Bounds of the *interval CDF* of the *extreme value* axial stress process $Y^{(1)I}_{1,\max}(T)$ of the telecommunication antenna provided by the proposed procedure considering Young's moduli of all the 19 *FE*s as uncertain (full) and retaining only the first $r = 6$ and $r = 13$ most influential uncertain parameters (reduced): a) $\Delta\alpha = 0.10$ and b) $\Delta\alpha = 0.20$ ($T = 1000T_0$).

5. Summary and Conclusions

A sensitivity-based procedure for reliability analysis of finite element modeled structures with interval stiffness subjected to stationary Gaussian multi-correlated random excitation is presented. The formulation is developed in the context of the *first-passage* theory under the Poisson assumption of independent up-crossings of a prescribed threshold. The presented procedure basically consists in identifying suitable

combinations of the endpoints of the uncertain structural parameters which yield accurate estimates of the bounds of the *interval reliability function* of the selected *extreme value* stress random process. This task is pursued by performing a preliminary sensitivity analysis of the reliability function.

The main features of the proposed sensitivity-based procedure may be summarized as follows: i) the bounds of the *interval reliability function* of the selected stress process are obtained by performing only two stochastic analyses of the structure wherein the uncertain parameters are set equal to the pertinent combinations identified by sensitivity analysis; ii) the presented procedure yields results in excellent agreement with the ones provided by the *vertex method* in spite of the much higher computational efficiency; iii) reliability analysis of arbitrary finite element modeled structures involving stiffness uncertainties can be performed; iv) sensitivity analysis provides a deep insight into the impact of stiffness fluctuations on structural performance allowing the identification of the least influential parameters which may be set equal to the nominal values to enhance the computational efficiency.

References

Ayyub, B. M. and G.J. Klir. *Uncertainty modeling and analysis in engineering and the sciences*. Chapman & Hall/CRC, Taylor & Francis Group, Boca Raton, 2006.

Ben-Haim, Y. A non-probabilistic concept of reliability. *Structural Safety*, 14(4): 227-245, 1994.

Corotis, R.B. An overview of uncertainty concepts related to mechanical and civil engineering. *ASCE-ASME Journal of Risk Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, 1(4): 040801, 2015.

Davenport, A. G. The spectrum of horizontal gustiness near the ground in high winds. *Quarterly Journal of the Royal Meteorological Society*, 87: 194–211, 1961.

Der Kiureghian, A. Analysis of structural reliability under parameter uncertainties. *Probabilistic Engineering Mechanics*, 23: 351–358, 2008.

Dong, W. and H. Shah. Vertex method for computing functions of fuzzy variables. *Fuzzy Sets and Systems*, 24: 65–78, 1987.

Elishakoff, I. and M. Ohsaki. *Optimization and Anti-Optimization of Structures under Uncertainty*. Imperial College Press, London, 2010.

Elishakoff, I. Essay on uncertainties in elastic and viscoelastic structures: From A. M. Freudenthal's criticisms to modern convex modeling. *Computers and Structures*, 56(6): 871–895, 1995.

Ferson, S., V. Kreinovich, L. Ginzburg, D.S. Myers, and K. Sentz. Constructing probability boxes and Dempster-Shafer structures. Sandia National Laboratories SAND2002–4015, 2003.

Goller, B., H. J. Pradlwarter, and G. I. Schüeller. Reliability assessment in structural dynamics. *Journal of Sound and Vibrations*, 332: 2488–2499, 2013.

Gupta, S. and C. S. Manohar. Reliability analysis of randomly vibrating structures with parameter uncertainties. *Journal of Sound and Vibrations*, 297: 1000–1024, 2006.

Li J. and J.B. Chen. *Stochastic Dynamics of Structures*. John Wiley & Sons, Singapore, 2009.

Lutes L. D. and S. Sarkani. *Stochastic Analysis of Structural and Mechanical Vibrations*. Prentice-Hall, Upper Saddle River, 1997.

Moens, D. and D. Vandepitte. A survey of non-probabilistic uncertainty treatment in finite element analysis. *Computers Methods in Applied Mechanics and Engineering*, 194: 1527-1555, 2005.

Moore, R. E. *Interval Analysis*. Prentice-Hall, Englewood Cliffs, 1966.

Moore, R. E., R. B. Kearfott, and M. J. Cloud. *Introduction to Interval Analysis*. SIAM, Philadelphia, 2009.

Muscolino, G. and A. Sofi. Stochastic analysis of structures with uncertain-but-bounded parameters via improved interval analysis. *Probabilistic Engineering Mechanics*, 28: 152-163, 2012.

Muscolino, G., R. Santoro, and A. Sofi. Explicit reliability sensitivities of linear structures with interval uncertainties under stationary stochastic excitations. *Structural Safety*, 52: 219–232, 2015.

Muscolino, G., R. Santoro, and A. Sofi. Interval fractile levels for stationary stochastic response of linear structures with uncertainties. *ASCE-ASME Journal of Risk Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, 2(1): 011004, 2016a.

Muscolino, G., R. Santoro, and A. Sofi. Reliability analysis of structures with interval uncertainties under stationary stochastic excitations. *Computers Methods in Applied Mechanics and Engineering*, 300: 47–69, 2016b.

Roberts J.B. and P.D. Spanos. *Random vibration and statistical linearization*. Dover, 2003.

Simiu, E. and R. Scanlan. *Wind Effects on Structures*. John Wiley & Sons, New York, 1996.

Sofi, A. and E. Romeo. A novel Interval Finite Element Method based on the improved interval analysis. *Computers Methods in Applied Mechanics and Engineering*, 311: 671–697, 2016.

Sofi, A., E. Romeo, O. Barrera, and A. Cocks. An interval finite element method for the analysis of structures with spatially varying uncertainties. *Advances in Engineering Software*, 128: 1–19, 2019.

Sofi, A., G. Muscolino, and F. Giunta. A sensitivity-based approach for reliability analysis of randomly excited structures with interval axial stiffness. *ASCE-ASME Journal of Risk Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, 6: 041008, 2020.

How to Gauge the Quality of a Testing Method When Ground Truth Is Known with Uncertainty

Nicholas Gray[1], Scott Ferson[1], and Vladik Kreinovich[2]
[1]*Institute for Risk and Uncertainty*
University of Liverpool
Liverpool L69 7ZF, UK
Nicholas.Gray@liverpool.ac.uk, sandp8@gmail.com
[2]*Department of Computer Science*
University of Texas at El Paso
El Paso, TX 79968, USA, vladik@utep.edu

Abstract. The quality of a testing method is usually measured by using sensitivity, specificity, and/or precision. To compute each of these three characteristics, we need to know the ground truth, i.e., we need to know which objects actually have the tested property. In many applications (e.g., in medical diagnostics), the information about the objects comes from experts, and this information comes with uncertainty. In this paper, we show how to take this uncertainty into account when gauging the quality of testing methods.

Keywords: sensitivity, specificity, precision, unknown ground truth

1. Formulation of the Problem

Formulation of the problem: in brief. In many practical situations, algorithms help us recognize the situation. Let us give a few examples.

- In medicine, algorithms use symptoms and measurement results to provide a diagnosis.
- In engineering, algorithms take the results of measurements and observations and, based on these results, decide whether, e.g., a road is expected to seriously deteriorate in the nearest future (and thus, repairs are needed now), or it can stay in working condition until the next year's testing.
- In military applications, algorithms help us decide whether a radar signal indicates an incoming enemy plane or an innocent flock of birds, etc.

In many applications, the available algorithms are not perfect: sometimes, they lead to a wrong result:

- a medical system can misdiagnose,
- a military system can mistakenly classify an innocent object as an enemy attack, etc.

© 2021 *by N. Gray, S. Ferson, and V. Kreinovich.*

In situations when we eventually learn the ground truth, we can gauge the quality of a testing method by comparing its results with the ground truth. Based on the results of this comparison, we can estimate how good is the testing method.

The challenge is that in many application areas, we do not always know the ground truth. For example, in medical diagnostics, the ground truth is supposed to come from medical doctors. However, in many cases, the doctors themselves are not 100% confident in their diagnoses. The existing techniques for gauging the quality of testing methods:

- either ignore such uncertain diagnoses altogether,
- or, vice versa, ignore the corresponding uncertainty and treat all the diagnoses as the ground truth.

To get a better understanding of the quality of different testing methods, it is therefore desirable to explicitly take the experts' uncertainty into account. This is what we do in this paper.

Before we describe how we propose to do it, let us first describe the problem in more detail.

Quality of testing methods. For many properties – e.g., for different diseases – we have different testing methods. These methods are rarely perfect. For example, for medical tests:

- sometimes, the test missed a disease, and
- sometimes, the test return an alarming result even when the patient does not have the corresponding disease.

To gauge the quality of a testing method – and to compare the quality of different testing methods – several characteristics are used. The most widely used are sensitivity, specificity, and precision; see, e.g., (Yerushalmy, 1947; Altman and Bland, 1994; Boyko, 1994; Loong, 2003; Macmillan and Creelman, 2004; Parikh et al., 2008; Powers, 2011; Sheskin, 2011; Ting, Sammut, and Webb, 2011; Tharwat, 2018). In order to describe these characteristics, let us introduce the corresponding notations.

Notations and comments.

- Let P denote the set of all the objects from the tested sample that *actually have* the tested property (e.g., the set of all the people in the sample who actually have the tested disease).
- Let N denote the set of all the objects from the tested sample that *do not have* the tested property (e.g., the set of all the people in the sample who do not have the tested disease).
- Let S_+ denote the set of all the objects for which the test *concluded that they have* the tested property (e.g., the set of all the people who the test classified as having the tested disease).
- Let S_- denote the set of all the objects for which the test *concluded that they do not have* the tested property (e.g., the set of all the people who the test classified as having the tested disease).

A perfect test should classify all the objects that actually have this property – and only these objects – as having the tested property. So, for a perfect test, we should have $P = S_+$, and, correspondingly, $N = S_-$. In reality, as we have mentioned, tests are not perfect, so we may have misclassified objects. The usual characteristics for gauging the quality of a testing method use the numbers of objects with or without the tested property that were classified correctly or incorrectly. In general, the number of elements in a set S will be denoted by $|S|$.

Let us now describe the usual characteristics.

Sensitivity. The first of the three characteristics is *sensitivity*. It is also known as *recall* or *True Positive Rate* – TPR for short. In the formulas in this paper, we will use the abbreviation TPR to describe sensitivity.

Sensitivity is defined as the proportion, among all the objects with the tested property, of the ones that were correctly classified by the test: e.g., the proportion of sick people for which the test recognized the disease.

In terms of our notations, the set of objects that have the tested property is P. The number of elements in this set is $|P|$. Among these objects, the set of all objects that have been correctly classified by the testing method is the intersection $P \cap S_+$ of the set P and the set S_+ of all the objects that were classified by the testing method as having the property. The number of such objects is equal to $|P \cap S_+|$. Thus, the sensitivity is equal to

$$\text{TPR} = \frac{|P \cap S_+|}{|P|}. \tag{1}$$

Specificity. The second of the three most used characteristics is *specificity*. It is also known as *True Negative Rate* – TNR, for short. In the formulas in this paper, we will use the abbreviation TNR to describe specificity.

Specificity is defined as the proportion, among the objects that do not have the tested property, of the ones that were correctly classified by the test: e.g., the proportion of healthy people that this test classified as healthy.

In terms of our notations, the set of objects that do not have the tested property is N. The number of elements in this set is $|N|$. Among these objects, the set of all objects that have been correctly classified by the method is the intersection $N \cap S_-$ of the set N and the set S_- of all the objects that were classified by the testing method as not having the property. The number of such objects is equal to $|N \cap S_-|$. Thus, the specificity is equal to

$$\text{TNR} = \frac{|N \cap S_-|}{|N|}. \tag{2}$$

Precision. The final of the three characteristics is *precision*. It is also known as Positive Predictive Value – PPV, for short. In the formulas in this paper, we will use the abbreviation PPV to describe precision.

Precision is defined as the proportion, among object that the test classified as having the tested property, of the objects who actually have this property – e.g., the proportion of sick people among those that the test classified as sick.

In terms of our notations, the set of objects that were classified as having the property is S_+. The number of elements in this set is $|S_+|$. Among these objects, the set of all objects that actually have the tested property is the intersection $P \cap S_+$ of the set S_+ and the set P of all the objects that actually have the tested property. The number of such objects is equal to $|P \cap S_+|$. Thus, the precision is equal to

$$\text{PPV} = \frac{|P \cap S_+|}{|S_+|}. \tag{3}$$

How can we use these characteristics to compare different testing methods. For each of the three characteristics, the larger the value of the characteristic, the better – and in the prefect case, all three characteristics are equal to 1. From this viewpoint, a reasonable way to compare different testing methods is to compare the values of one or more of the three characteristics: if for one of the methods, the corresponding value is larger, this means that, from the viewpoint of this characteristic, this method is better.

Comment. Of course, to make a definite conclusion about which testing method is better, we need to take into account that the values of each characteristic come from a finite sample and are, thus, only an approximate representation of the actual quality of a testing method. For the same method, for different random samples, we can get slightly larger or slightly smaller values of the corresponding characteristic.

So, strictly speaking, to make a definite conclusion that one of the testing methods is better, we need to check that the difference between the values of the characteristic is statistically significant. There are known statistical procedures for checking this.

This is especially important to take into account when the sample sizes are small. When the sample sizes are large, the corresponding randomness becomes very small.

Important problem: often, we do not know the "ground truth". The formulas for computing the above three characteristics assume that we know know the "ground truth", i.e., that we know exactly:

- which objects have the tested property and
- which objects do not have this property.

In the above example, which patients have the tested disease.

In practice, however, this information often comes from experts – e.g., from medical doctors – and experts are often not 100% sure about their statements and their diagnoses.

How can we take this expert uncertainty into account when gauging the quality of a test?

What we do in this paper. This is the problem that we address in this paper: we sow how to take the expert's uncertainty into account when estimating the above characteristics of the testing method.

2. How to Describe Expert's Uncertainty

Need for subjective probability. For each object i, an expert makes:

– either a statement that the object has the tested property,

– or a statement that the object does not have the tested property.

In both cases, the expert is usually not absolutely confident in his/her statement.

Since the whole procedure is based on statistics, it is reasonable to try to gauge the expert's degree of certainty c_i in his or her statement by a probability value. Once we know this degree of certainty, then:

– If the expert believes that the object i most probably has the tested property, then the probability p_i that this object *has* the tested property is equal to $p_i = c_i$. Correspondingly, the probability that the object i *does not have* the tested property is equal to $1 - c_i$.

– If the expert believes that the object i most probably does not have the tested property, then the probability that this object *does not have* the tested property is equal to c_i. Correspondingly, the probability that the object i *has* the tested property is equal to $p_i = 1 - c_i$.

Comment. Probability values describing expert's degree of confidence are known as *subjective probabilities* – to distinguish them from usual (*objective*) probabilities, that describe the frequency with which certain events occur. For example, the fact that the probability 1/2 of the coin falling heads means that, in general, the coin will fall heads in half of the cases.

How do we get subjective probabilities? Where can we get the subjective probabilities from? A natural idea is to ask the experts. Sometimes, they are able to gauge their own degrees of certainty by providing the corresponding number.

What is the expert cannot provide such probabilities – but we have a record of the expert's past estimates. In many cases, the expert cannot meaningfully provide the corresponding subjective probabilities. How can we then gauge the expert's uncertainty?

One possible approach is to use the above analogy between subjective and objective probabilities. If we have a record of past estimates of the same expert, estimates for which we actually know the ground truth, then, for this expert, we can estimate our degree of confidence c_i in this expert's statement as the proportion of cases in which the expert turned out to be right. For example, if in the past, the medical doctor was right 80% of the time, we take $c_i = 0.8$.

Sometimes, we cannot do this for each individual expert, but we can estimate the overall subjective probability c of experts. The confidence c is usually close to 1, to it makes sense to represent it as $c = 1 - \varepsilon$ for some small $\varepsilon > 0$. In this case, we take $p_i = c = 1 - \varepsilon$ is the experts believe that the i-th object has the tested property, and $p_i = 1 - c = \varepsilon$ if they don't.

What can we do in all other cases? But what if an expert cannot estimate his/her degree of confidence by a number, and we do not have the record of this expert's past estimates. How can we then estimate the expert's degree of confidence?

To do that, we can use standard techniques from decision theory; see, e.g., (Fishburn, 1969; Luce and Raiffa, 1989; Raiffa, 1997; Nguyen, Kosheleva, and Kreinovich, 2009; Kreinovich, 2014). According to decision theory, to estimate the expert's certainty in a statement S, we can ask this expert to compare, for different values p from the interval $[0, 1]$, the following two alternatives:

– getting a certain reward (e.g., \$100) with probability p, or

– getting the exact same reward if the statement S turned out to be true.

Clearly:

– If the expert prefers the first alternative, this means that his/her subjective probability of S is smaller than p.

– If the expert prefers the second alternative, this means that his/her subjective probability of S is larger than p.

We can use following bisection procedure to find the corresponding subjective probability. In the beginning, all we know about the subjective probability p is that it is located somewhere in the interval $[\underline{p}, \overline{p}] = [0, 1]$. At each stage of this process, we will decrease the size of this interval by half. This can be done as follows.

Suppose that at some stage, we have an interval $[\underline{p}, \overline{p}]$. Then, on the next stage, we compute the midpoint

$$p_m = \frac{\underline{p} + \overline{p}}{2} \tag{4}$$

and ask the expert to compare the alternative "reward with probability p_m" with the alternative "reward if S is true".

– If the expert prefers the alternative "reward with probability p_m", this means that his/her subjective probability is smaller than p_m. Since we already know that the subjective probability p is in the interval $[\underline{p}, \overline{p}]$ and is, thus, larger than or equal to $\underline{p}$, we can thus conclude that p is in the interval $[\underline{p}, p_m]$.

– If the expert prefers the alternative "reward is S is true", this means that his/her subjective probability is larger than p_m. Since we already know that the subjective probability p is in the interval $[\underline{p}, \overline{p}]$ and is, thus, smaller than or equal to $\overline{p}$, we can thus conclude that p is in the interval $[\underline{p}_m, \overline{p}]$.

In both cases, we get an interval of half-size that contains the actual subjective probability. We start with an interval of width 1. In the first step, we decrease the width of the interval to 1/2, in 2 steps to 1/4, ..., and in k steps, we get an interval of width 2^{-k}. If we take a midpoint of this interval, this midpoint represents the subjective probability with accuracy $2^{-(k+1)}$.

This way, after a small number of iterations, we get the subjective probability with a reasonably high accuracy. In particular:

– in 3 steps – i.e., by asking 3 questions to the expert – we estimate the subjective probability with accuracy $2^{-4} = \frac{1}{16} < 10\%$;

– in 6 steps – i.e., by asking 6 questions to the expert – we estimate the subjective probability with accuracy $2^{-7} = \frac{1}{128} < 1\%$; and

– in 9 steps – i.e., by asking 9 questions to the expert – we estimate the subjective probability with accuracy $2^{-10} = \dfrac{1}{1024} < 0.1\%$.

Summarizing. By using one of the above methods, we can estimate, for each object i, the expert's degree of confidence c_i in his or her statement about this object. Depending on whether this statement was positive or negative, we can then estimate the expert's subjective probability p_i that the i-th object has the tested property:

– if the expert believes that most probably the object has the tested property, then we take $p_i = c_i$, and

– if the expert believes that most probably the object does not have the tested property, then we take $p_i = 1 - c_i$.

In some cases, instead of individual values p_i for each i, we only know the overall degree of confidence $c = 1 - \varepsilon$ in the expert's statement. In this case:

– if the expert believes that most probably the object has the tested property, then we take $p_i = c = 1 - \varepsilon$, and

– if the expert believes that most probably the object does not have the tested property, then we take $p_i = 1 - c = \varepsilon$.

Let us now show how we can use these subjective probabilities p_i.

3. How to Take Expert's Uncertainty into Account: General Analysis

Notations. Let us first introduce some additional notations.

Let us denote:

– by E_+ the set of all the objects that, according to the experts, most probably have the desired property, and

– by E_- the set of all the objects that, according to the experts, most probably do not have the desired property.

In general, due to the expert uncertainty:

– the set E_+ may be different from the set of P of the objects that actually have the tested property, and

– the set E_- may be different from the set of N of the objects that actually do not have the tested property.

Let n denote the overall number of tested objects. In terms of our previous notations,

$$n = |P| + |N| = |E_+| + |E_-| = |S_-| + |S_+|. \tag{5}$$

Let us enumerate these objects by numbers from 1 to n. In these notation, all the sets that considered earlier – namely, the sets P, N, S_-, and S_+ and their intersections – become subsets of the sample $\{1, \ldots, n\}$.

Let $\chi_P(i)$ denote the characteristic function of the set P of all the objects that actually have the tested property, i.e.:

- if the object i has the tested property, then $\chi_P(i) = 1$, and
- if the object i does not have the tested property, then $\chi_P(i) = 0$.

General analysis of the problem. We consider situations in which we do not know for sure whether the i-th object has the tested property or not. All we know, based on the expert's estimate, is that this happens with probability p_i. In other words, the value $\chi_P(i)$ is a random variable:

- with probability p_i, we have $\chi_P(i) = 1$, and
- with the remaining probability $1 - p_i$, we have $\chi_P(i) = 0$.

In statistics, for each random variable η, a reasonable idea is to compute its mean $E[\eta]$ and its variance $V[\eta]$– and, as we will see later, this is useful in our case as well. For the random variable $\chi_P(i)$, we have

$$E\left[\chi_P(i)\right] = p_i \cdot 1 + (1 - p_i) \cdot 0 = p_i \tag{6}$$

and

$$V\left[\chi_P(i)\right] = E\left[\left(\chi_P(i) - E\left[\chi_P(i)\right]\right)^2\right] = p_i \cdot (1 - p_i)^2 + (1 - p_i) \cdot (0 - p_i)^2 =$$

$$p_i \cdot (1 - p_i)^2 + (1 - p_i) \cdot p_i^2 = p_i \cdot (1 - p_i) \cdot [(1 - p_i) + p_i] = p_i \cdot (1 - p_i). \tag{7}$$

We can use these results to answer, e.g., a question of how many objects actually have the desired property, i.e., what is the number of elements $|P|$ in the set P. This number can be obtained if we consider all the elements from the sample $\{1, \ldots, n\}$ one by one, and add 1 every time we have an element from the set P, i.e., every time when $\chi_P(i) = 1$. If the element i does not belong to the set P (i.e., when $\chi_P(i) = 0$), then we do not add anything – which is also equivalent to adding $\chi_P(i)$. So, we can describe the above procedure as simply adding all the values $\chi_P(i)$ corresponding to all n objects. As a result, we get the value

$$|P| = \sum_{i=1}^{n} \chi_P(i). \tag{8}$$

To be able to get a good estimate of the test's quality, we need to test a sufficiently large number of objects. Thus we can conclude that the number n is large. So, the above sum (8) is the sum of a large number of small independent random variables.

It is reasonable to assume that the estimates corresponding to different objects – and often produced by different experts – are statistically independent. It is known that, due to the Central Limit Theorem, the distribution of such sums is close to Gaussian; see, e.g., (Sheskin, 2011). Thus, it is reasonable to assume that PPV is normally distributed. Its mean is equal to the sum of the means, i.e.,

$$E[|P|] = \sum_{i=1}^{n} p_i. \tag{9}$$

For the sum of independent random variables, the variance is equal to the sum of the variables, so we have

$$V[|P|] = \sum_{i=1}^{n} p_i \cdot (1 - p_i). \tag{10}$$

Now, we are ready to analyze how the expert's uncertainty affect the values of the three characteristics. We will start with the case of precision, which turns out to be the easiest to analyze.

4. Estimating Precision: Analysis of the Problem

General case. According to the formula (3), precision PPV is defined as the ratio of $|P \cap S_+|$ to $|S_+|$. The set S_+ of all the objects that the test classifies as having the tested property does not depend on expert estimates. The only thing that, in this formula, depends on the expert estimates, is the value $|P \cap S_+|$ – since it depends on which objects actually have this property or not, and our only information about this comes from the experts.

Similarly to the previous section, we can conclude that

$$|P \cap S_+| = \sum_{i \in S_+} \chi_P(i). \tag{11}$$

Thus, we have

$$E\left[|P \cap S_+|\right] = \sum_{i \in S_+} p_i \tag{12}$$

and

$$V\left[|P \cap S_+|\right] = \sum_{i \in S_+} p_i \cdot (1 - p_i). \tag{13}$$

When we divide a random variable by a constant (in this case, by $|S_+|$), then the mean value divides by the same constant, while the variance divides by the square of this constant. So, we have

$$E[\mathrm{PPV}] = \frac{1}{|S_+|} \cdot \sum_{i \in S_+} p_i \tag{14}$$

and

$$V[\mathrm{PPV}] = \frac{1}{|S_+|^2} \cdot \sum_{i \in S_+} p_i \cdot (1 - p_i). \tag{15}$$

Case when we only know the overall degree of confidence $c = 1-\varepsilon$ in expert statements. In this case, we have $p_i = 1 - \varepsilon$ if $i \in E_+$ and $p_i = \varepsilon$ if $i \in E_-$. Thus:

$$\sum_{i \in S_+} p_i = \sum_{i \in S_+ \cap E_+} (1-\varepsilon) + \sum_{i \in S_+ \cap E_-} \varepsilon = |S_+ \cap E_+| \cdot (1-\varepsilon) + |S_+ \cap E_-| \cdot \varepsilon. \tag{16}$$

Here, $|S_+ \cap E_-| = |S_+| - |S_+ \cap E_-|$, so

$$\sum_{i \in S_+} p_i = |S_+ \cap E_+| \cdot (1-2\varepsilon) + |S_+| \cdot \varepsilon. \tag{17}$$

Therefore,

$$E[\mathrm{PPV}] = (1-2\varepsilon) \cdot \frac{|S_+ \cap E_+|}{|S_+|} + \varepsilon. \tag{18}$$

Similarly, we have $p_i \cdot (1-p_i) = \varepsilon \cdot (1-\varepsilon)$, so $\sum\limits_{i \in S_+} p_i \cdot (1-p_i) = |S_+| \cdot \varepsilon \cdot (1-\varepsilon)$, and the formula (15) takes the form

$$V[\mathrm{PPV}] = \frac{\varepsilon \cdot (1-\varepsilon)}{|S_+|}. \tag{19}$$

Based on precision, when can we say that one testing method is better than the other one? To compare two different methods, with means $E[\mathrm{PPV}_1]$ and $E[\mathrm{PPV}_2]$ and variances $V[\mathrm{PPV}_1]$ and $V[\mathrm{PPV}_2]$, we can use the usual technique for comparing two random variables, and conclude that the difference between PPV_1 and PPV_2 if

$$E[\mathrm{PPV}_1] - E[\mathrm{PPV}_2] \geq t \cdot \sqrt{V[\mathrm{PPV}_1] + V[\mathrm{PPV}_2]}, \tag{20}$$

for an appropriate t (the value t depends on the desired confidence level).

5. Estimating Precision: Results

General conclusion. If we take into account expert uncertainty, then PPV – as well as two other characteristics – becomes a random variable.

General case. If we know, for each object i, the subjective probability p_i that this object has the tested property, then the mean and variance of PPV can be determined by using formulas (14) and (15).

Case when we only know the overall degree of confidence $c = 1-\varepsilon$ in expert statements. In this case, we can estimate the mean nd variance of PPV by using formulas (18) and (19).

Comments.

– The formula (18) can be reformulated as follows: we take the value that we would have obtained if we did not take expert uncertainty into account, multiply it by $1 - 2\varepsilon$, and add ε to the resulting product.

– Strictly speaking, to the variance values estimated by using formulas (15) and (19), we should add the variances caused by the fact that we are estimate PPV based on the finite sample. Since the expert uncertainty and the uncertainty caused by the finiteness of the sample are independent, to get the overall variance, we should simply add the new expression to the variance to the known expressions corresponding to sample finiteness.

How do we decide which testing method is better. To decide whether one testing method is statistically significantly better than another one, we use the formula (20) with an appropriate value t.

6. Estimating Sensitivity: Analysis of the Problem

General case. In terms of the values $\chi_P(i)$, the actual value of the sensitivity is

$$\mathrm{TPR} = \frac{\sum\limits_{i \in S_+} \chi_P(i)}{\sum\limits_{i=1}^{n} \chi_P(i)}. \tag{21}$$

Here,

$$\sum_{i=1}^{n} \chi_P(i) = \Sigma_+ + \Sigma_-, \tag{22}$$

where we denoted $\Sigma_+ \stackrel{\text{def}}{=} \sum\limits_{i \in S_+} \chi_P(i)$ and $\Sigma_- \stackrel{\text{def}}{=} \sum\limits_{j \in S_-} \chi_P(j)$. Since these two sums contain different random variables $\chi_P(i)$ and $\chi_P(j)$, and we assumed that all the variables $\chi_P(i)$ an $\chi_P(j)$ are independent, the sums Σ_+ and Σ_- are independent as well. Thus

$$\mathrm{TPR} = \frac{\Sigma_+}{\Sigma_+ + \Sigma_-}. \tag{23}$$

Here, similarly to the case of precision, we can conclude that both Σ_+ and Σ_- are independent (approximately) Gaussian random variables, with means

$$E[\Sigma_+] = \sum_{i \in S_+} p_i \text{ and } E[\Sigma_-] = \sum_{j \in S_-} p_j \tag{24}$$

and variances

$$V[\Sigma_+] = \sum_{i \in S_+} p_i \cdot (1 - p_i) \text{ and } V[\Sigma_-] = \sum_{j \in S_-} p_j \cdot (1 - p_j). \tag{25}$$

We can thus find the mean and standard deviation of TPR if we simulate two Gaussian random variables Σ_+ and Σ_-, then compute the ratio (23) for each simulation, and compute the mean and average of these simulation results.

Case when we only know the overall degree of confidence $c = 1-\varepsilon$ in expert statements. In this case, the formulas (24) and (25) take the following form:

$$E[\Sigma_+] = (1-\varepsilon)\cdot|S_+ \cap E_+| + \varepsilon\cdot|S_+ \cap E_-| \text{ and } E[\Sigma_-] = (1-\varepsilon)\cdot|S_- \cap E_+| + \varepsilon\cdot|S_- \cap E_-|; \quad (26)$$

$$V[\Sigma_+] = \varepsilon\cdot(1-\varepsilon)\cdot|S_+| \text{ and } V[\Sigma_-] = \varepsilon\cdot(1-\varepsilon)\cdot|S_-|. \quad (27)$$

How do we decide which testing method is better. For each characteristic X, we can say – similarly to the formula (20) – that the first testing method is better if

$$E[X_1] - E[X_2] \geq t\cdot\sqrt{V[X_1] + V[X_2]}, \quad (28)$$

for an appropriate t; the value t depends on the desired confidence level.

7. Estimating Sensitivity: Resulting Algorithm

How to estimate sensitivity of a testing method based on the testing results. First, depending on whether we know all the values p_i for each i or only one value ε, we use either the formulas (24)–(25) or the formulas (26)–(27) to find the values of the mean and variance of Σ_+ and Σ_-.

Then, several (K) times we run a usual random number generator for normally distributed random variables to get N simulated values $\Sigma_+^{(k)}$ and $\Sigma_-^{(k)}$. Based on these simulated values, we use the formula (23) to estimate the simulated values of TPR as

$$\mathrm{TPR}^{(k)} = \frac{\Sigma_+^{(k)}}{\Sigma_+^{(k)} + \Sigma_-^{(k)}}. \quad (29)$$

Finally, based on these simulated values, we estimate the mean and variance of TRP in the usual way, as:

$$E[\mathrm{TPR}] = \frac{1}{K}\cdot\sum_{k=1}^{K}\mathrm{TPR}^{(k)} \text{ and } V[\mathrm{TPR}] = \frac{1}{K-1}\cdot\sum_{k=1}^{K}\Big(\mathrm{TPR}^{(k)} - E[\mathrm{TPR}]\Big)^2. \quad (30)$$

How do we decide which testing method is better. We say that the first testing method is better if

$$E[\mathrm{TPR}_1] - E[\mathrm{TPR}_2] \geq t\cdot\sqrt{V[\mathrm{TPR}_1] + V[\mathrm{TPR}_2]}, \quad (31)$$

for an appropriate t; the value t depends on the desired confidence level.

8. Estimating Specificity: Analysis of the Problem and the Resulting Algorithm

Analysis of the problem. In terms of the values $\chi_P(i)$, the actual value of the sensitivity is

$$\mathrm{TNR} = \frac{\sum\limits_{i \in S_-} (1 - \chi_P(i))}{\sum\limits_{j=1}^{n} (1 - \chi_P(j))} = \frac{|S_-| - \sum\limits_{j \in S_-} \chi_P(j)}{n - \sum\limits_{i=1}^{n} \chi_P(i)} = \frac{|S_-| - \Sigma_-}{n - \Sigma_+ - \Sigma_-}. \tag{32}$$

We already know that Σ_+ and Σ_- can be viewed as independent normally distributed random variables, with known means and variances. Thus, we arrive at the following algorithm.

How to estimate sensitivity of a testing method based on the testing results. First, depending on whether we know all the values p_i for each i or only one value ε, we use either the formulas (24)–(25) or the formulas (26)–(27) to find the values of the mean and variance of Σ_+ and Σ_-.

Then, several (K) times we run a usual random number generator for normally distributed random variables to get N simulated values $\Sigma_+^{(k)}$ and $\Sigma_-^{(k)}$.

Up to now, we perform the same computation steps as when estimating sensitivity. Now, the computations differ. To estimate specificity, based on the simulated values $\Sigma_+^{(k)}$ and $\Sigma_-^{(k)}$, we use the formula (32) to estimate the simulated values of TNR as

$$\mathrm{TNR}^{(k)} = \frac{|S_-| - \Sigma_-^{(k)}}{n - \Sigma_+^{(k)} - \Sigma_-^{(k)}}. \tag{33}$$

Finally, based on these simulated values, we estimate the mean and variance of TNP in the usual way, as:

$$E[\mathrm{TNR}] = \frac{1}{K} \cdot \sum_{k=1}^{K} \mathrm{TNR}^{(k)} \text{ and } V[\mathrm{TNR}] = \frac{1}{K-1} \cdot \sum_{k=1}^{K} \Big(\mathrm{TNR}^{(k)} - E[\mathrm{TNR}]\Big)^2. \tag{34}$$

How do we decide which testing method is better. We say that the first testing method is better if

$$E[\mathrm{TNR}_1] - E[\mathrm{TNR}_2] \geq t \cdot \sqrt{V[\mathrm{TNR}_1] + V[\mathrm{TNR}_2]}, \tag{35}$$

for an appropriate t; the value t depends on the desired confidence level.

9. Conclusion

A usual way of gauging the quality of a testing method is to compare the results of this method with ground truth. However, in many practical situations, we do not always know the ground truth. For example, we may want to gauge the quality of a medical diagnostic system, but for some patients, the medical doctors are not 100% sure what is the correct diagnosis. Usually, we either ignore such

cases – or simply ignore the uncertainty and consider the most probable diagnosis as the ground truth. To get a more accurate description of a quality of a testing method, it is desirable to explicitly take into account the degree of expert's certainty.

In this paper, we provide methods that explicitly take into account these degrees of certainty when estimating the quality of a testing method.

Acknowledgements

This work was supported in part by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science) and HRD-1242122 (Cyber-ShARE Center of Excellence). It was also supported by the program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478.

The authors are thankful to all the participants of the First International Conference on Data and Information Fusion DIR'2019 (Santa Fe, New Mexico, August 21–23, 2019) for valuable discussions, and to the anonymous referees for useful advice.

References

Altman, D. G., and J. M. Bland, "Diagnostic tests. 1: Sensitivity and specificity", *BMJ*, 308(6943): 1552, 1994.

Boyko, E. J., "Ruling out or ruling in disease with the most sensitive or specific diagnostic test: short cut or wrong turn?", *Medical Decision Making*, 14(2): 175–179, 1994.

Fishburn, P. C., *Utility Theory for Decision Making*, John Wiley & Sons Inc., New York, 1969.

Kreinovich, V., Decision making under interval uncertainty (and beyond)", In: Guo, P., and W. Pedrycz (eds.), *Human-Centric Decision-Making Models for Social Sciences*, Springer Verlag, 2014, pp. 163–193.

Loong, T. W., "Understanding sensitivity and specificity with the right side of the brain", *BMJ*, 327(7417): 716–719, 2003.

Luce, R. D., and R. Raiffa, *Games and Decisions: Introduction and Critical Survey*, Dover, New York, 1989.

Macmillan, N. A., and C. D. Creelman, *Detection Theory: A User's Guide*, Psychology Press, Hove, East Sussex, United Kingdom, 2004.

Nguyen, H. T., O. Kosheleva, and V. Kreinovich, "Decision making beyond Arrow's 'impossibility theorem', with the analysis of effects of collusion and mutual attraction", *International Journal of Intelligent Systems*, 24(1): 27–47, 2009.

Parikh, R., A. Mathai, S. Parikh, G. Chandra Sekhar, and R. Thomas, "Understanding and using sensitivity, specificity and predictive values", *Indian Journal of Ophthalmology*, 56(1): 45–50, 2008.

Powers, D. M. W., "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation", *Journal of Machine Learning Technologies*, 2(1): 37–63, 2011.

Raiffa, H., *Decision Analysis*, McGraw-Hill, Columbus, Ohio, 1997.

Sheskin, D. J., *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011.

Tharwat, A., "Classification assessment methods", *Applied Computing and Informatics*, 2018 doi:10.1016/j.aci.2018.08.003.

Ting, K. M., C. Sammut, and G. I. Webb, (eds.), *Encyclopedia of Machine Learning*, Springer, Cham, Switzerland, 2011.

Yerushalmy, J., "Statistical problems in assessing methods of medical diagnosis with special reference to X-ray techniques", *Public Health Reports*, 62(2): 1432–1439, 1947.

Random set solutions to partial differential equations

Jelena Karakašević and Michael Oberguggenberger
Unit of Engineering Mathematics, University of Innsbruck, 6020 Innsbruck, Austria;
jelena.karakasevic.mail@gmail.com, michael.oberguggenberger@uibk.ac.at

Abstract. The past decades have seen increasing interest in modelling uncertainty by heterogeneous methods, combining probability and interval analysis, especially for assessing parameter uncertainty in engineering models. A unifying mathematical framework admitting the combination of a wide range of such methods is the theory of random sets, describing input and output of a structural model by set-valued random variables. The purpose of this paper is to highlight the mathematics behind this approach. The modelling and computational implications are discussed and demonstrated with the help of a numerical example from elastostatics.

Keywords: Partial differential equations; random fields; interval parameters; random sets

1. Introduction

In many branches of engineering, such as linear elasticity or wave propagation, the quantities of interest are solutions to equilibrium equations or to equations of motion. These equations contain coefficients that—as a rule—depend on further parameters, such as material constants or geometric dimensions. In addition, the parameters may vary in space and time. A major task in engineering design is the analysis of the uncertainty of the computed output. This uncertainty traces back to uncertainty in the input parameters, which may be due to many sources: random fluctuations, lack of information, conflicting assessments, random measurement errors, systematic measurement errors, fluctuations due to spatial inhomogeneity, errors made by assigning parameter status to state variables, model insufficiency, just to name a few. (Uncertainty in the model itself is important, but will be left aside in this paper.)

Available information on parameter variability may consist in frequency distributions obtained from large samples, values from small samples or single measurements, interval bounds, experts' point estimates, or educated guesses from experience. These different types of uncertainty have often been classified into uncertainties of aleatoric or of epistemic type. Accordingly, an increasing desire has risen in the engineering community to introduce methods of quantifying the uncertainty beyond probability, such as interval analysis, set-valued models, fuzzy sets, evidence theory, random sets, sets of probability measures, imprecise probability, lower and upper previsions, info-gap-analysis, etc. This paper focusses on combinations of probability and interval analysis. This includes random variables, random fields, intervals, interval fields, random variables with interval parameters and random fields with interval parameters (such as mean, variance, correlation length).

All these types and their combinations can be accommodated in the framework of *random sets*. Random sets are set-valued random variables, whereby the values may be sets of real numbers,

© 2021 *by authors.*

subsets of finite dimensional vector spaces, or subsets of function spaces. For example, random variables are single-valued random sets and intervals are random sets on a one-point probability space. The purpose of this paper is to highlight some important mathematical modelling questions that arise when using random sets. The considerations will be exemplified with the help of a static model, the equilibrium equation $-\operatorname{div}(a(x)\operatorname{grad} u(x)) = f(x)$ for the transversal displacement $u(x)$ of a membrane. The coefficient $a(x)$ contains the uncertain material properties of the membrane, while $f(x)$ is the load, both of which may vary with respect to location x. When $a(x)$ is a random field, a random field with interval parameters, or a (finite) interval field, the resulting solution $u(x)$ is no longer single-valued, but rather a random set with values in a function space. (The values at fixed points x are real-valued random sets.) This situation has been investigated from several viewpoints, including computational aspects, in the literature; see e.g. (Do, Gao, Song and Beer, 2016; Faes and Moens, 2017; Muscolino, Santoro and Sofi, 2016; Oberguggenberger, 2015; Oberguggenberger and Wurzer, 2019; Schmelzer, Oberguggenberger and Adam, 2010; Sofi and Romeo, 2018). Papers laying special emphasis on comparing stochastic and interval methods are (Gao, Song and Tin-Loi, 2010; Jiang, Zheng and Han, 2018; Muscolino, Sofi and Zingales, 2013; Sofi, 2015; Sofi, Muscolino and Giunta, 2020).

In order to establish that the values of the solution are intervals, random variables, or random sets, the main issue is the continuous dependence of the solution on the parameters of the random field. The arguments are rather straightforward in the case of intervals or random variables/fields, but involve some subtleties when the latter are combined, resulting in general random sets.

The plan of the paper is as follows. Section 2 will be devoted to the required mathematical set-up. These facts are known, but it appears useful to elaborate them in some detail for the later applications. Section 3 singles out the principal mathematical argument that allows one to prove that the output quantity is a random set. The combination of arguments giving a full proof of this result seems to be new. Section 4 details the considerations for the membrane equation and is a new application of the findings of Section 3. Section 5 contains a comparison with the parametric approach and numerical aspects. Detailed proofs can be found in the dissertation of the first author (Nedeljković, 2020).

2. Mathematical preliminaries

This section specifies some mathematical notions which are required for a rigorous treatment of random sets. Recall that a subset S of a finite dimensional vector space such as n-dimensional Euclidean space $\mathbb{R}^d$ is *open*, if for each point x in S, each ball of the form $\{y \in \mathbb{R}^d : \|y - x\| < \eta\}$ with center in x and sufficiently small radius η is also contained in S. The set S is *closed*, if with each converging sequence whose terms belong to S, the limit also belongs to S. The subset S is *bounded*, if the norms $\|x\|$ of the elements of S are all smaller than a fixed constant. Further, finite dimensional Euclidean spaces are *separable*, that is, they contain a countable dense subset (a sequence of points such that every ball in $\mathbb{R}^d$ contains at least one element of the sequence), and they are *complete* (that is, if x_k is a sequence such that the sum of its norms $\sum_{k=1}^{\infty} \|x_k\|$ converges, then the sum $\sum_{k=1}^{\infty} x_k$ also converges).

All these notions can be literally generalized to infinite dimensional normed spaces. A typical example is the space of continuous functions on a bounded, closed interval I, denoted by $\mathcal{C}(I)$, with norm $\|f\| = \max_{x\in I} |f(x)|$. This space is separable and complete.

2.1. Random variables

A real-valued random variable a is characterized by its distribution function $F_a(x)$ which defines the probabilities P of basic events of the form $\{a \leq x\}$ via $P(a \leq x) = F_a(x)$. This point of view suffices for most purposes. However, in the sequel *families of random variables* will be needed, and it will be necessary to specify the domains of the random variables under consideration. The usual mathematical setting is a probability space (Ω, Σ, P) where Ω is a set (the collection of all elementary events), Σ is the family of measurable subsets of Ω, and P is a probability measure (that assigns to every set belonging to Σ a number between 0 and 1, satisfying the usual axioms). In this point of view, a random variable is a mapping from Ω into the real line such that the events $\{\omega \in \Omega : a(\omega) \leq x\}$ are measurable (i.e., belong to Σ). Thus a probability $P(\{\omega \in \Omega : a(\omega) \leq x\})$ can be assigned to such events, denoted again by $P(a \leq x)$, and these probabilities in turn define the distribution function of the random variable a. From the probabilities $P(a \leq x)$, the probabilities of open and closed intervals, their complements, their countable unions and intersections etc. can be computed. The family of subsets of the real line which can be obtained by indefinite repetition of these set theoretic operations defines the so-called *Borel σ-algebra* $\mathcal{B}$. A random variable has the property that all sets of the form $\{\omega \in \Omega : a(\omega) \in B\}$, where B is any Borel set, are measurable. Accordingly, random variables are often referred to as measurable functions.

If the probability space is finite, say $\Omega = \{1, \ldots, m\}$, one may take $\Sigma = \Omega$, and every real-valued function on Ω is measurable. A typical infinite probability space is standard Gaussian space with $\Omega = \mathbb{R}$, Σ the Borel measurable subsets of $\mathbb{R}$, and the probability of an event B given by the Gaussian integral $P(B) = (2\pi)^{-1/2} \int_B \mathrm{e}^{-\omega^2/2} \mathrm{d}\omega$; the probability measure P is given through the Gaussian density $\mathrm{e}^{-\omega^2/2}/\sqrt{2\pi}$. On standard Gaussian space, every continuous real-valued function on Ω is a random variable.

These notions can be easily extended to finite dimensional random variables, i.e., with values in $\mathbb{R}^d$ (then the joint probability distributions have to be specified). Measurable functions with values in an infinite dimensional normed space can be defined in the same way.

2.2. Random fields

A random field is a process that assigns a random variable $q(x)$ to every point x in a region D in space. To define the probabilistic properties of the field, the joint distributions of the values at any finite number of points $q(x_1), \ldots q(x_n)$ should be specified. If the random field is Gaussian, it is completely specified by the mean value $\mu_q = \mathrm{E}(q(x))$ and the second moments, i.e., the covariance $\mathrm{COV}(q(x), q(y))$ for any two points x, y. If it is (weakly) homogeneous and isotropic, the covariance depends only on the distance $\rho = \|x - y\|$ of the points and is of the form $\mathrm{COV}(q(x), q(y)) = C(x, y) = \sigma^2 c(\rho)$ with the field variance σ^2 and the autocorrelation function

$c(\rho)$. A typical autocorrelation function is of the form

$$c(\rho) = \exp\Big(-\frac{|\rho|}{\ell}\Big), \tag{1}$$

where ℓ is the so-called correlation length. The standard method of simulation of a random field is based on the *Karhunen-Loève expansion*. It is obtained by solving the eigenvalue problem

$$\int_D C(x,y)\varphi_k(y)\,\mathrm{d}y = c_k\varphi_k(x) \tag{2}$$

which has a sequence of positive eigenvalues c_k and orthonormal eigenfunctions $\varphi_k(x)$ (orthonormality in mean square). Then

$$q(x) = \sum_{k=1}^{\infty} \sqrt{c_k}\xi_k\varphi_k(x) \tag{3}$$

where the ξ_k are uncorrelated random variables with unit variance. If the process is Gaussian, the ξ_k are independent and distributed according to $\mathcal{N}(0,1)$.

For the numerical simulation, the spatial region is discretized by a grid and the φ_k are taken, e.g., piecewise constant on the grid elements. The eigenvalue problem becomes a matrix eigenvalue problem, and the series (3), with approximate eigenvalues and eigenfunctions, truncated after a finite number M of terms, can be used for Monte Carlo simulations of the field trajectories (i.e., realizations of the field).

If the random field $q(x)$ is the input, say, in a static structural model, the model output $u(x)$ is also a random field. The required number of terms M in the expansion depends crucially on the correlation length and the smoothness of the input-output function $q \to u$. The model output in turn depends, among others, on the field parameters μ, σ, ℓ. When applying random sets subsequently, information about the type of dependence on these parameters (e.g., continuity) is needed. This is aided by the fact that a mean zero random field on $D = \mathbb{R}$ with autocorrelation function (1) and variance σ^2 can be obtained as an Ornstein-Uhlenbeck process, namely as solution to the Langevin stochastic differential equation

$$\mathrm{d}q = -\frac{1}{\ell}q\,\mathrm{d}x + \sqrt{\frac{2}{\ell}}\,\sigma\,\mathrm{d}W, \quad q|_{x=0} \sim \mathcal{N}(0,\sigma^2), \tag{4}$$

where $W = W(x,\omega)$ denotes Wiener process on the real line. In this case, realizations of the random field can be generated by simulating solutions to the Langevin equation (Arnold, 1974).

2.3. Random sets

In general, a *random set* is a set-valued random variable satisfying certain measurability conditions, to be detailed below. The simplest case arises when the underlying probability space is finite. In this case, one speaks of *finite random sets* or *Dempster-Shafer structures*. Such a structure is given by finitely many closed subsets $A_i, i = 1,\dots,m$ of a target space $\mathcal{A}$, usually Euclidean space $\mathbb{R}^d$, called the *focal elements*, each of which comes with a *probability weight* $p_i = p(A_i)$, $\sum p_i = 1$.

Viewed as a random set, a Dempster-Shafer structure is given by an m-point probability space $\Omega = \{1, 2, \ldots, m\}$ with probability masses $\{p_1, p_2, \ldots, p_m\}$; the assignment $i \to A_i$ is the defining set-valued random variable A.

However, random sets generated by infinite families of random variables will be needed in the context of random fields. This leads to the notion of infinite random sets. In this case, the random sets will be defined on a general probability space (Ω, Σ, P). Also, the target space can be infinite-dimensional, for example, a function space containing the system outputs.

In the sequel, the target space will be either Euclidean space $\mathbb{R}^d$ or more generally a normed, complete and separable space $\mathcal{A}$. Given a set-valued map $\omega \to A(\omega)$ where each $A(\omega)$ is a subset of the target space $\mathcal{A}$, upper and lower inverses of subsets B of the target space $\mathcal{A}$ are defined by

$$A^-(B) = \{\omega \in \Omega : A(\omega) \cap B \neq \emptyset\}, \qquad A_-(B) = \{\omega \in \Omega : A(\omega) \subset B\}. \tag{5}$$

The requirements that $\omega \to A(\omega)$ is a *random set* are

- each $A(\omega)$ is closed;
- the upper inverses $A^-(B)$ are measurable subsets of Ω for every closed subset B of the target space $\mathcal{A}$.

Recall that a *measurable selection* of a random set A is a random variable a such that $a(\omega) \in A(\omega)$ for all $\omega \in \Omega$. An important tool is the *fundamental measurability theorem*, see e.g. (Molchanov, 2005), that states the equivalence of the measurability property of $A^-(B)$ for open, for closed, and for Borel subsets B of the target space as well as the equivalence with the existence of a *Castaing representation*. A Castaing representation is a sequence of measurable selections $a_k(\omega) \in A(\omega)$ such that for all ω, $\{a_k(\omega) : k = 1, 2, 3, \ldots\}$ is dense in $A(\omega)$. (To be precise, the fundamental measurability theorem requires *completeness of the probability space* (Ω, Σ, P) in the sense that every subset of a set of probability zero belongs to Σ.)

The lower inverses of events B can be written as $A_-(B) = \{\omega \in \Omega : A(\omega) \cap B^{\mathsf{c}} \neq \emptyset\}^{\mathsf{c}}$, where B^{c} denotes the complement of B, and hence are measurable due to the fundamental measurability theorem and the measurability of $A^-(B)$. Thanks to the required measurability properties, it is legitimate to introduce upper and lower probabilities

$$\overline{P}(B) = P(A^-(B)) = P(\{\omega \in \Omega : A(\omega) \cap B \neq \emptyset\}), \quad \underline{P}(B) = P(A_-(B)) = P(\{\omega \in \Omega : A(\omega) \subset B\})$$

for any Borel set $B \subset \mathcal{A}$.

The concept of random sets encompasses many other well-established methods of uncertainty modelling. Clearly, every interval defines a random set (as a random variable on a one-point probability space). Similarly, every random variable a on a probability space (Ω, Σ, P) defines a (single-valued) random set $\omega \to A(\omega) = \{a(\omega)\}$. The upper and lower probability of an event B coincide with its probability $P(B)$. Further, every normalized fuzzy number can be viewed as a random set. Here the probability space is the interval $\Omega = [0, 1]$, equipped with the uniform probability distribution, and the focal elements $A(\omega)$ are just the ω-level sets. It is not difficult to prove that the possibility measure $\pi(B)$ of a subset B of the real line coincides with its upper probability (Goodman and Nguyen, 2002). Indeed, and rather obviously,

$$\pi(B) = \sup_{\omega \in [0,1]} \{A(\omega) \cap B \neq \emptyset\} = P(\{\omega \in \Omega : A(\omega) \cap B \neq \emptyset\}) = \overline{P}(B).$$

2.4. Random sets generated by families of random variables

To set the stage, consider random variables a_λ, defined on the same probability space (Ω, Σ, P), and depending on a set of parameters $\lambda \in \Lambda$. Suppose the random variables have values in a target space $\mathcal{A}$. Then one can define a set-valued map $\omega \to A(\omega) = \{a_\lambda(\omega) : \lambda \in \Lambda\}$. The question is whether this assignment defines a random set. Before answering this question, consider the example of an imprecise Gaussian family

$$a_\lambda(\omega) = \mu + \sigma\Phi^{-1}(\omega) \tag{6}$$

with interval parameters $\lambda = (\mu, \sigma) \in \Lambda = [\underline{\mu}, \overline{\mu}] \times [\underline{\sigma}, \overline{\sigma}]$. Here the underlying probability space is the unit interval $\Omega = (0, 1)$ with the uniform distribution, and Φ is the standard normal distribution function. It follows that each random variable a_λ is distributed according to $\mathcal{N}(\mu, \sigma^2)$. The corresponding subset of the real line is

$$A(\omega) = [\underline{a}(\omega), \overline{a}(\omega)] \quad \text{with} \quad \underline{a}(\omega) = \min\left(a_\lambda(\omega) : \lambda \in \Lambda\right), \quad \overline{a}(\omega) = \max\left(a_\lambda(\omega) : \lambda \in \Lambda\right).$$

Due to the fact that the parameter set Λ is a (two-dimensional) closed and bounded interval and the fact that the assignment $(\mu, \sigma) \to \mu + \sigma\Phi^{-1}(\omega)$ is continuous at fixed ω, it follows that each $A(\omega)$ is a closed and bounded interval, indeed. Some members of the family and the resulting random intervals are depicted in Figure 1. Anticipating that $\omega \to A(\omega)$ is a random set, one may define the

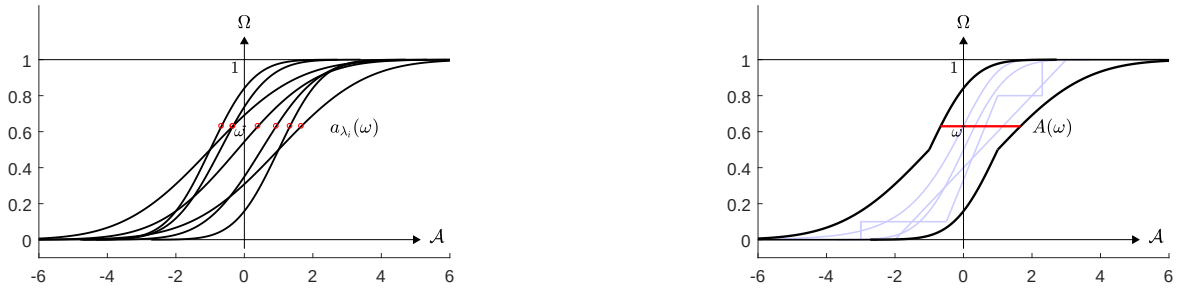


Figure 1. Some members of the family a_λ (left), random set with focal element (right). The parameters $\lambda = (\mu, \sigma)$ were taken as $-1 \le \mu \le 1$, $1 \le \sigma \le 2$.

upper and lower distribution functions

$$\underline{F}(b) = P(\omega : \overline{a}(\omega) \le b), \quad \overline{F}(b) = P(\omega : \underline{a}(\omega) \le b).$$

The upper and lower distribution functions enclose a so-called *probability box* or p-box (Ferson, Ginzburg, Kreinovich, Myers and Sentz, 2003). Note that the p-box contains many probability distribution functions other than the one coming from the Gaussian family. Anticipating also that the functions $\underline{a}$, $\overline{a}$ bounding the random intervals are random variables, one obtains simply that $\underline{F}$ is the distribution function of $\overline{a}$ and $\overline{F}$ is the distribution function of $\underline{a}$.

Here is the argument why $\omega \to A(\omega)$ is a random set. First, one constructs a Castaing representation by taking a countable dense subset of parameter values $\lambda_k = (\mu_k, \sigma_k)$ in $\Lambda = [\underline{\mu}, \overline{\mu}] \times [\underline{\sigma}, \overline{\sigma}]$, for example, all points with rational coordinates. Again from the continuity of the assignment $\lambda \to$

$a_\lambda(\omega)$ at fixed ω, i.e., $(\mu, \sigma) \to \mu + \sigma\Phi^{-1}(\omega)$, see (6), it follows that the set $\{a_{\lambda_k}(\omega) : k = 1, 2, 3, \ldots\}$ is dense in $A(\omega)$ for every fixed ω. Let B be an open set. Then

$$\begin{aligned} A^-(B) &= \{\omega \in \Omega : A(\omega) \cap B \neq \emptyset\} = \{\omega \in \Omega : \text{ there is } k \text{ such that } a_{\lambda_k}(\omega) \in B\} \\ &= \bigcup_k \{\omega \in \Omega : a_{\lambda_k}(\omega) \in B\} \end{aligned} \tag{7}$$

is measurable as a countable union of measurable sets. (The individual sets are measurable, because the maps $\omega \to a_\lambda(\omega)$ are measurable at each fixed λ.) The fundamental measurability theorem implies that $A^-(B)$ is measurable for every Borel set B, in particular, for every closed set B. This is the required property for $\omega \to A(\omega)$ to be a random set.

Using again that $\{a_{\lambda_k}(\omega) : k = 1, 2, 3, \ldots\}$ is dense in $A(\omega)$, it follows that $\underline{a}(\omega) = \inf\{a_{\lambda_k}(\omega) : k = 1, 2, 3, \ldots\}$ and $\overline{a}(\omega) = \sup\{a_{\lambda_k}(\omega) : k = 1, 2, 3, \ldots\}$ are measurable as infima and suprema of countably many random variables.

It is seen that the required arguments are quite elaborate. The underlying structure will be worked out in the following sections. Further details on interpretations and applications of random sets can be found, e.g., in the monographs (Bernardini and Tonon, 2010; Molchanov, 2005; Nguyen, 2006) and the survey papers (Beer, Ferson and Kreinovich, 2013; Oberguggenberger, 2013).

3. Random sets generated by parametrized random fields

This central section lays down the basic structure that will allow one to prove that set-valued maps obtained as solutions to partial differential equations whose coefficients are given as intervals, random fields, or parametrized random fields, are indeed random sets.

3.1. The general measurability result

For the present purpose, the most general situation to be considered is the following.

- (Ω, Σ, P) is a complete probability space;
- the target space $\mathcal{A}$ is a normed, complete and separable space;
- Λ is a bounded and closed subset of some Euclidean space;
- $a_\lambda(\omega)$, $\lambda \in \Lambda$ is a family of random variables with values in $\mathcal{A}$, that is, the maps $\omega \to a_\lambda(\omega)$ are measurable for each λ;
- the random variables depend continuously on λ, that is, the maps $\Lambda \to \mathcal{A} : \lambda \to a_\lambda(\omega)$ are continuous for each ω.

Then the following assertions hold: The set-valued map

$$\omega \to A(\omega) = \{a_\lambda(\omega) : \lambda \in \Lambda\}$$

defines a random set. All focal elements $A(\omega)$ are bounded and closed subsets of $\mathcal{A}$. In addition, if Λ is a (multi-dimensional) interval and $\mathcal{A} = \mathbb{R}$, the sets $A(\omega)$ are intervals.

Proof of assertion. First, it is clear from the continuity assumption that all $A(\omega)$ are bounded and closed subsets of $\mathcal{A}$. The assertion that they are intervals under the additional specializing hypothesis follows from the continuity as well. To prove the random set property, one has to show that the upper inverses $A^-(B)$ are measurable for every open subset B of $\mathcal{A}$. This follows exactly by the argument given in equation (7) and applying the fundamental measurability theorem.

Example. Consider a function $u = u(\lambda, a)$ that depends on an interval parameter λ, say $\lambda \in [\underline{\lambda}, \overline{\lambda}]$, and a fixed random variable a. Such a situation arises, for example, when u is the response of a system with mixed interval and random uncertainty. Both λ and a can be multi-dimensional, so this covers the case of discretized interval and random fields. Assume that the dependence on λ and a is continuous. The functions $a_\lambda(\omega) = u(\lambda, a(\omega))$ form a family of random variables exactly of the type discussed here. Collecting the responses in $A(\omega) = \{u(\lambda, a(\omega)) : \lambda \in \Lambda\}$ therefore leads to a random set, actually a random interval $\omega \to [\underline{a}(\omega), \overline{a}(\omega)]$ if the response is one-dimensional.

3.2. Continuous dependence of random fields on their parameters

A more involved instance of a random set is obtained by taking a random field on the real line ($x \in \mathbb{R}$) whose parameters are intervals. For example, consider mean zero random fields $q(x, \omega, \ell)$ with fixed variance σ^2, but correlation length ℓ varying in an interval $[\underline{\ell}, \overline{\ell}]$. A common probability space for these random fields with different correlation lengths is needed. A convenient way of generating such random fields, in the case of the autocorrelation function (1), is by means of the Langevin equation (4). The common probability space is Wiener space Ω_W (the space of trajectories of Wiener process with the induced probability measure). Due to the fact that the random fields $q(x, \omega, \ell)$ are solutions to (4), the following properties hold; see e.g. (Schmelzer, 2010; Schmelzer, 2013):

- For fixed $x \in \mathbb{R}$ and $\ell \in [\underline{\ell}, \overline{\ell}]$, the map $\omega \to q(x, \omega, \ell)$ is measurable.
- For fixed $x \in \mathbb{R}$ and $\omega \in \Omega_W$, the map $\ell \to q(x, \omega, \ell)$ is continuous.

Thus at every chosen point x, the assignment $\omega \to Q(x, \omega) = \{q(x, \omega, \ell) : \ell \in [\underline{\ell}, \overline{\ell}]\}$ defines a (scalar) random set, as can be seen by the same arguments as in the example above. The trajectories of the random field are interval-valued curves, given by $x \to [\min(q(x, \omega, \ell), \max(q(x, \omega, \ell)]$, where the minimum and maximum are obtained by letting ℓ run through $[\underline{\ell}, \overline{\ell}]$.

Alternatively, the Karhunen-Loève expansion (3) can be used to place the random fields $q(x, \omega, \ell)$ in a common probability space Ω, not depending on ℓ. Formally, this can be done by considering an infinite product of standard Gaussian spaces with elements $\omega = (\omega_1, \omega_2, \omega_3, \ldots)$. The required sequence of Gaussian variables ξ_k is then given by $\xi_k(\omega) = \omega_k$, and the Karhunen-Loève expansion reads

$$q(x, \omega, \ell) = \sum_{k=1}^{\infty} \sqrt{c_{k,\ell}}\, \xi_k(\omega) \varphi_{k,\ell}(x). \tag{8}$$

In general, there is no way to assert that the eigenvalues $c_{k,\ell}$ and eigenfunctions $\varphi_{k,\ell}(x)$ depend continuously on the correlation length ℓ, even after discretization and reduction to a matrix eigenvalue problem. Thus one has to analyze the situation in each case individually. The case of the autocovariance function (1) on a finite interval can be treated explicitly, thanks to the results in

(Ghanem and Spanos, 1991). Without restriction of generality, one may assume that the finite interval is $D = [-1, 1]$. Then the eigenvalue problem (2) becomes

$$\int_{-1}^{1} \mathrm{e}^{-|x-y|/\ell} \varphi(y)\, \mathrm{d}y = c\varphi(x). \tag{9}$$

Following (Ghanem and Spanos, 1991), problem (9) can be solved in the following way. First, the equations

$$\frac{1}{\ell} - \alpha \tan \alpha = 0, \quad \text{respectively} \quad \alpha + \frac{1}{\ell} \tan \alpha = 0 \tag{10}$$

have two sequences of positive solutions denoted by $\alpha_k = \alpha_k(\ell)$ and $\alpha_k^* = \alpha_k^*(\ell)$, for even and odd k, respectively. The resulting eigenvalues and eigenfunctions are

$$c_{k,\ell} = \frac{2\ell}{1 + \ell^2 \alpha_k^2(\ell)},\ c_{k,\ell}^* = \frac{2\ell}{1 + \ell^2 \alpha_k^{*2}(\ell)},\ \varphi_{k,\ell}(x) = \frac{\cos \alpha_k(\ell) x}{\sqrt{1 + \frac{\sin 2\alpha_k(\ell)}{2\alpha_k(\ell)}}},\ \varphi_{k,\ell}^*(x) = \frac{\sin \alpha_k^*(\ell) x}{\sqrt{1 - \frac{\sin 2\alpha_k^*(\ell)}{2\alpha_k^*(\ell)}}} \tag{11}$$

for even and odd k, respectively. Therefore, the one-dimensional random field $q(x, \omega, \ell)$ with covariance function given by the equation (1) has the Karhunen-Loève expansion

$$q(x, \omega, \ell) = \lim_{M \to \infty} \sum_{k=1}^{M} \Big(\sqrt{c_{k,\ell}} \xi_k(\omega) \varphi_{k,\ell}(x) + \sqrt{c_{k,\ell}^*} \xi_k^*(\omega) \varphi_{k,\ell}^*(x) \Big). \tag{12}$$

For the truncated Karhunen-Loève expansion, one may take a $2M$-dimensional standard Gaussian space as probability space Ω, on which the random variables ξ_k and ξ_k^* are defined. It is clear that the solutions w and w^* of the deterministic equations (10) depend continuously on ℓ. The same is true for the eigenvalues as well as eigenfunctions shown in (11). Thus truncating the Karhunen-Loève expansion (12) at a finite M, the resulting random field depends continuously on the correlation length ℓ at fixed ω.

3.3. Random fields as random variables valued in function spaces

Consider the random field $q(x, \omega, \ell)$ as in Subsection 3.2. The *paths or trajectories (in space)* of the random field are the maps $x \to q(x, \omega, \ell)$, with fixed ω and ℓ. It is well-known from the theory of stochastic differential equations that the paths are continuous (Arnold, 1974). This also follows from the Kolmogorov-Chentsov theorem (Kallenberg, 1997), because the autocovariance function (1) satisfies the hypotheses of this theorem. However, for our purposes, more is needed.

First, jointly continuous dependence on x and ℓ is required. This is easy to achieve, because one can consider the *paths in space and parameter set* as the maps $(x, \ell) \to q(x, \omega, \ell)$, with fixed ω and apply the previous arguments. Indeed, when using a truncated Karhunen-Loève expansion, the argument was just given after (12). When the the Langevin equation is used, one may resort to the results of (Schmelzer, 2010; Schmelzer, 2013) which assert the continuous dependence, jointly in x and ℓ.

Second, the random fields have to be interpreted as random variables in the space of continuous functions. This is because they enter as variable coefficients in the partial differential equations to be

solved, and the solution depends on the whole trajectories, not only the values of the random fields at individual points x. Let x belong to an interval I and ℓ to an interval L, both closed and bounded. Continuity of the paths in x and ℓ can be stated by saying that the map $q(\omega) : (x, \ell) \to q(x, \omega, \ell)$ belongs to the space of continuous function $\mathcal{C}(I \times L)$ for fixed ω. Thus the random field induces a map $\Omega \to \mathcal{C}(I \times L)$, $\omega \to q(\omega)$. Standard arguments from the theory of stochastic processes show that this map is measurable, that is, a random variable with values in $\mathcal{C}(I \times L)$. For details of the proof, see (Nedeljković, 2020, Section 1.2). The argument also shows that

- for fixed $\ell \in L$, the map $\Omega \to \mathcal{C}(I)$, $\omega \to q(\omega, \ell)$ is measurable;
- for fixed $\omega \in \Omega$, the map $L \to \mathcal{C}(I)$, $\ell \to q(\omega, \ell)$ is continuous.

Note that the same random field q may be considered in three different ways as a measurable function of ω: as an element of the function space $\mathcal{C}(I \times L)$, denoted by $q(\omega)$, as an element of the function space $\mathcal{C}(I)$ at fixed correlation length ℓ, denoted by $q(\omega, \ell)$, and as a real number at each fixed x and ℓ, denoted by $q(x, \omega, \ell)$. Whenever its values are collected in a random set, the corresponding capital letter Q will be used.

3.4. The mean values of a parametrized random field as an interval field

Let $\omega \to A(\omega)$ be a random set whose focal elements $A(\omega) = [\underline{a}(\omega), \overline{a}(\omega)]$ are closed and bounded intervals. Let $\{a_k(\omega)$, k = 1,2,3,. . . $\}$ be a Castaing representation. The *Aumann expectation* $\mathrm{E}(A)$ is defined as the closure of the set of numbers $\{\mathrm{E}(a_k) : k = 1, 2, 3, \ldots\}$, see (Molchanov, 2005). It is clear that in the considered case, the Aumann expectation is simply the interval $\mathrm{E}(A) = [\mathrm{E}(\underline{a}), \mathrm{E}(\overline{a})]$. Next, consider a random set generated by a parametrized random field $\omega \to Q(x, \omega) = \{q(x, \omega, \ell) : \ell \in [\underline{\ell}, \overline{\ell}]\}$ as in Subsection 3.2. Thanks to the continuous dependence on the parameter ℓ discussed in Subsection 3.2 each $Q(x, \omega)$ is a closed and bounded interval at fixed x. One may form the Aumann expectation $\mathrm{E}(Q(x))$, which is a deterministic interval at fixed x. Thus the assignment $x \to \mathrm{E}(Q(x))$ defines an interval field.

4. Application to elastostatics

This section is devoted to a simple scalar partial differential equation with random field coefficients, in order to demonstrate how the random field methods can be applied in practice. The problem is

$$\begin{aligned} -\operatorname{div}(a(x)\operatorname{grad} u(x)) &= f(x), \quad x \text{ in } D \\ u(x) &= 0, \qquad x \text{ in } \partial D \end{aligned} \tag{13}$$

where D is a bounded open domain in $\mathbb{R}^d$ and ∂D is its boundary. For example, in two space dimensions u might be the transversal displacement of a non-uniform membrane under transversal load f. The spatially varying elastic properties are subsumed in the coefficient $a(x)$. Alternatively, in three space dimensions, $u(x)$ could be the pressure of a fluid in a porous medium with permeability $a(x)$ and source rate $f(x)$ (Matthies, 2008).

In the foundations of the finite element method, problem (13) is commonly solved by variational methods. Suitable function spaces are defined as follows: $L^2(D)$ is the space of square integrable functions; $H^1(D)$ is the space of square integrable functions all whose first order partial derivatives are square integrable as well; $H_0^1(D)$ is the subspace of functions that vanish on the boundary ∂D. Both $L^2(D)$ and $H_0^1(D)$ are normed spaces; the square of the norms are given by

$$\|f\|^2_{L^2(D)} = \int_D f(x)\,\mathrm{d}x, \quad \|u\|^2_{H^1(D)} = \|u\|^2_{L^2(D)} + \|\mathrm{grad\,u}\|^2_{L^2(D)}.$$

The spaces are complete and separable. The variational formulation of equation (13) then is:

Find $u \in H_0^1(D)$ such that

$$\int_D a(x)\,\mathrm{grad}\,u(x)\cdot\mathrm{grad}\,v(x)\,\mathrm{d}x + \int_D f(x)v(x)\,\mathrm{d}x = 0$$

for all $v \in H_0^1(D)$.

The deterministic case. Assume that a is a continuous function on D which is bounded from above and below, that is, $a \in \mathcal{C}(D)$ and $\alpha \le a(x) \le \beta$ for some constants $\alpha, \beta > 0$ and all x in D. Further, assume that f is square integrable on D. Under these conditions, it is well-known that the variational problem has a unique solution u in $H_0^1(D)$, (Nečas, 1967; Trèves, 1975). For the purpose of this example, the source term f will be fixed. The solution u depends on the coefficient function a; this dependence will be made explicit in the notation $u = u_a$. Denote the set of functions which are continuous on D up the boundary with values between α and β by $\mathcal{C}(D;\alpha,\beta)$. It is quipped with the norm $\|a\| = \max_{x\in D}|a(x)|$. The crucial assertion is that the map

$$\mathcal{C}(D;\alpha,\beta) \to H_0^1(D) : a \to u_a$$

is continuous. This non-trivial result can be proved by invoking the variational formulation, as it was done in (Nedeljković, 2020). A different proof, for more general partial differential equations, can be found in (Nečas, 1967).

The random field case. To start with, let $a(x,\omega)$, $\omega \in \Omega$, be a random field with continuous paths lying between α and β. As noted in Subsection 3.3, it can be viewed as a random variable with values in $\mathcal{C}(D;\alpha,\beta)$. Due to the continuity of the above map, the assignment $\omega \to u_{a(\omega)}$ is measurable, thus the solution u_a is a random variable with values in $H_0^1(D)$. Intuitively, this means that u_a is a random field whose paths belong to $H_0^1(D)$. In order to interpret it as a classical random field, one should be able to assign values $u_{a(\omega)}(x)$ at each point x in D. In one space dimension, this is easy because $H_0^1(D)$ is continuously imbedded in $\mathcal{C}(D)$, and so the pointwise evaluations

$$\mathcal{C}(D;\alpha,\beta) \to \mathbb{R} : a \to u_a(x)$$

are continuous as well. Thus the maps $\omega \to u_{a(\omega)}(x)$ are random variables for each x, and the paths $x \to u_{a(\omega)}(x)$ are continuous functions. In space dimensions two and three, more theory is needed. First, the input random field $a(x,\omega)$ should have Lipschitz continuous paths, that is, satisfying an inequality of the form $|a(x,\omega) - a(y,\omega)| \le C|x-y|$ for some constant C and all x, y. (This is not

true of the random field generated by the Langevin equation, but it is true of the random field generated by the truncated Karhunen-Loève expansion). Then one can invoke regularity theory saying that the solution to problem (13) is in $H^2(D')$ in any open subregion D' of D and depends continuously on a in this space (Nečas, 1967). Again, $H^2(D')$ is continuously imbedded in $\mathcal{C}(D')$, and so one obtains a classical random field $u_{a(\omega)}(x)$, defined for all x in the interior of D.

The random set case. Let $a(x)$ be a parametrized random field. More specifically, it will be constructed as follows. In two space dimensions, take a closed and bounded interval I such that $D \subset I \times I$ and a random field of the form $q(x, \omega, \ell)$ as in Subsection 3.2 and set

$$a(x, \omega, \ell) = \mu(x) + \chi(q(x_1, \omega, \ell)q(x_2, \omega, \ell)), \quad \text{for} \quad x = (x_1, x_2) \text{ in } D. \tag{14}$$

Here $\mu(x) > 0$ is the mean field (assumed fixed and deterministic) and χ is a cut-off function; this is needed because the individual realizations of a Gaussian random field can be arbitrarily large with non-zero probability. In the end, the cut-off has to be chosen so as to guarantee that $\alpha \leq a(x, \omega, \ell) \leq \beta$.

One could also use different correlation lengths in the two variables and/or multiply the fields with different standard deviations σ_1, σ_2. The construction in one or three space dimensions is analogous.

When ℓ varies in an interval L, a set-valued solution to (13) can be defined as $U(\omega) = \{u_{a(\omega,\ell)} : \ell \in L\}$. Combining the continuity of the map $a \to u_a$ with the statement at the end of Subsection 3.3, one obtains that

- for fixed $\ell \in L$, the map $\Omega \to H_0^1(D)$, $\omega \to u_{a(\omega,\ell)}$ is measurable;
- for fixed $\omega \in \Omega$, the map $L \to H_0^1(D)$, $\ell \to u_{a(\omega,\ell)}$ is continuous.

It follows from Subsection 3.1 that $U(\omega)$ is a random set in the target space $H_0^1(D)$. Again, in one space dimension one may form the set-valued functions $U(\omega, x) = \{u_{a(\omega,\ell)}(x) : \ell \in L\}$ at each x in D. By the same argument, the $U(\omega, x)$ are random sets in $\mathbb{R}$; they form the point evaluations of the random set solution $U(\omega)$ at the points x in D. In space dimensions two and three, a modified argument as in the random field case has to be applied.

Numerical example. The purpose of this subsection is to illustrate the method in a simple example, the displacement of an L-shaped membrane with non-uniform, random elastic properties. This has the character of a text book example, and so all units are taken dimensionless. The L-shaped domain D can be seen in Figure 2 (left). The equation for the displacement is (13). A constant load $f(x) \equiv 1$ is assumed, and the random field $a(x)$ is taken of the form (14) with constant mean $\mu(x) \equiv 1$, autocovariance function of the form (1), and field variance identically equal to one as well. In this elliptic problem, it suffices to guarantee that the realizations of the random field $a(x, \omega, \ell)$ stay above zero. Thus a lower cut-off function χ was chosen at the value 0.1. To produce the random set solution, the correlation length was taken from the interval $0.5 \leq \ell \leq 1.5$.

Numerically, the L-shaped region was subdivided in a grid of size 18×18, and problem (13) was solved by the finite element method using a Matlab-program. To generate the random field, the Karhunen-Loève expansion truncated at $M = 130$ was used. The value of $a(x, \omega, \ell)$ entered in each finite element was chosen as the average of the values at its four nodes. A Monte Carlo sample size

of 500 was used. At each realization of the underlying $2M$-dimensional Gaussian variable, see (12), the finite element program was executed with eleven values of ℓ between 0.5 and 1.5, stepsize 0.1.

As outlined above, this procedure results in an approximation to the random set solution $U(\omega)$. The solution is a random set in $H_0^1(D)$ and in $H^2(D')$ in any subregion, actually everywhere except near the sharp inside corner (Grisvard, 1985). Consequently, each horizontal or vertical slice through the solution is a random set in the space of continuous functions. At fixed $x = (x_1, x_2)$, each realization of the solution is an interval. The mean (Aumann expectation) is an interval field. Figure 2 (left) shows a realization of the displacement random field at fixed correlation length $\ell = 0.5$. Realizations of the full random set solution can be visualized by plotting slices at fixed x_1 or x_2. Such a cut at $x_2 = 0.4444$ is shown in Figure 2 (center). Above each fixed x_1, the realization is an interval. Finally, the mean value field can also be depicted by slices. Figure 2 (right) shows a slice of this interval field at $x_2 = 0.4444$. The upper and lower bounding curves of the mean displacement are shown together with the mean solutions at correlation lengths 0.5, 0.6, ..., 1.5.

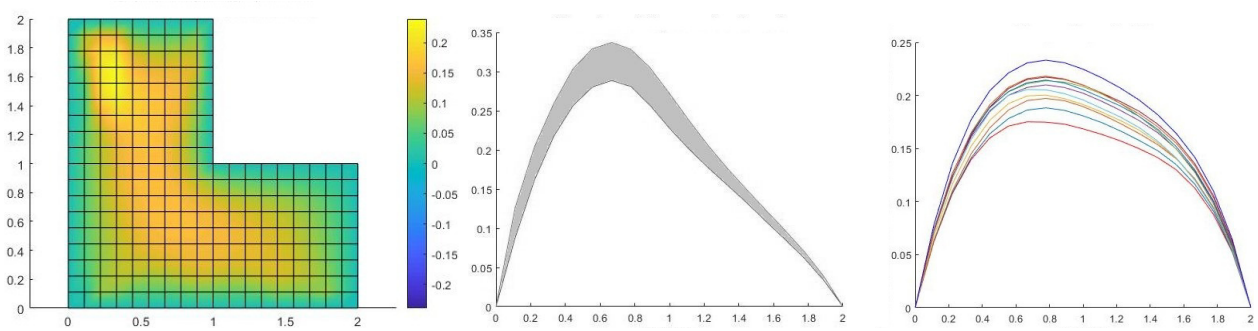


Figure 2. Some aspects of the random set solution. Trajectory of displacement random field at fixed correlation length $\ell = 0.5$ (left); slice at $x_2 = 0.4444$ through a realization of the full random set solution (center); slice at $x_2 = 0.4444$ through the mean value interval field, showing individual mean fields for varying correlation lengths inside.

5. Further results, numerical aspects

The parametric point of view. Given a parametrized family $a_\lambda, \lambda \in \Lambda$, of random variables or random fields, and a map $a_\lambda \to u_{a_\lambda}$ describing the response of a system, the approach favored in this paper has been to first construct the random set

$$U(\omega) = \{u_{a_\lambda(\omega)} : \lambda \in \Lambda\} \tag{15}$$

and then to define lower and upper probabilities of events B by

$$\underline{P}(B) = P(U_-(B)), \quad \overline{P}(B) = P(U^-(B)).$$

An alternative approach, the *parametric viewpoint*, would consist in skipping the random set step and defining lower and upper probabilities induced by the family $a_\lambda, \lambda \in \Lambda$, directly:

$$P_{\text{low}}(B) = \inf_{\lambda\in\Lambda} P(u_{a_\lambda}), \quad P_{\text{upp}}(B) = \sup_{\lambda\in\Lambda} P(u_{a_\lambda}).$$

For the events $B = (-\infty, b]$ one obtains again upper and lower distribution functions and a p-box as in Figure 1, which contains, however, only the cumulative distribution functions of the underlying family (and hence is often referred to as a *parametric p-box*). This approach gives tighter bounds (Fetz and Oberguggenberger, 2016):

$$\underline{P}(B) \leq P_{\text{low}}(B) \leq P_{\text{upp}}(B) \leq \overline{P}(B)$$

but deprives one of using the concepts of the theory of random sets. The numerical computation of the two types of upper and lower probabilities are different, but of comparable effort.

Numerical aspects. As is well-known, the computational effort when implementing heterogeneous uncertainty models is high: it involves both Monte Carlo sampling and the computation of interval bounds. Suppose the quantity of interest, given through a model response function $a_\lambda \to u_{a_\lambda}$ as above, is one-dimensional. (This could be the model response at a single point x or the value of a performance criterion.) Further, suppose the lower and upper distribution functions $\underline{F}(b)$, $\overline{F}(b)$ of this quantity are sought.

In both approaches, the parameter set Λ has to be discretized in a grid $\lambda_1, \ldots, \lambda_M$. In the random set case, one first produces a Monte Carlo sample $\omega_1, \ldots, \omega_N$ on the underlying probability space, then computes the bounding functions $\underline{u}(\omega_k)$, $\overline{u}(\omega_k)$ of the sets formed by the u_{a_λ} over all λ_i at fixed ω_k and evaluates the sample of bounding functions statistically. In the parametric point of view, one starts with a Monte Carlo sample of size N of each random variable a_{λ_i}, propagates it through the model function u, evaluates the result statistically at each λ_i and then finds the upper lower bounds of the resulting probabilities. A more detailed exposition of the two procedures can be found in (Fetz and Oberguggenberger, 2016).

In these crude forms, both algorithms require $M \cdot N$ evaluations of the (generally expensive) model function u. Accordingly, many approaches have been devised to reduce computational cost. In the random set case, a stochastic response surface, like a polynomial chaos expansion (Ghanem and Spanos, 1991), reduces the number M to a level required for the accuracy of the response surface. Evaluating a Monte Carlo sample of size N, given the response surface, is cheap (Oberguggenberger, 2015). For further sampling concepts for random sets, see (Alvarez, 2006). Other methods include approximations by inner and outer bounds (Alvarez, 2009; Alvarez, Hurtado and Ramírez, 2017; Tonon, 2004). In the parametric point of view, a prominent method is reweighting a single importance Monte Carlo sample, which means $M = 1$. Still N model evaluations are required (Fetz, 2019; Troffaes, 2018; Zhang and Shields, 2018). (This can of course be aided by a deterministic response surface.) Very efficient methods appear to be *advanced line sampling* (de Angelis, Patelli and Beer, 2015) that intertwines the two required loops and can reduce M and N simultaneously, and the method of *probability plots*, an enhancement of the first order reliability method (FORM) proposed by (Hurtado, Alvarez and Paredes, 2017).

6. Conclusion

The paper addressed mathematical and modelling issues when heterogeneous uncertainty is entered in engineering models. It was demonstrated that the framework of random sets is suitable for this task, both from the theoretical and computational viewpoint. It was shown how the theoretical

measurability questions can be addressed, and that the continuous dependence of the solution and the involved random and interval quantities on their parameters is important. This gives a solid basis for the commonly employed numerical methods. As was shown in (Nedeljković, 2020), the methods can also be applied to dynamic problem, such as wave or transport equations.

Acknowledgements

This work was supported by the FWF doctoral program Computational Interdisciplinary Modelling W1227 of the Austrian Science Fund as well as a doctoral scholarship of the University of Innsbruck.

References

Alvarez, D. A. On the calculation of the bounds of probability of events using infinite random sets. *International Journal of Approximate Reasoning*, 43:241–267, 2006.

Alvarez, D. A. A Monte Carlo-based method for the estimation of lower and upper probabilities of events using infinite random sets of indexable type. *Fuzzy Sets and Systems*, 160:384–401, 2009.

Alvarez, D. A., Hurtado, J. A. and J. Ramírez. Tighter bounds on the probability of failure than those provided by random set theory. *Computers & Structures*, 189:101–113, 2017.

Arnold, L. (1974). *Stochastic differential equations: theory and applications.* New York-London-Sydney: Wiley-Interscience.

Beer, M., Ferson, S. and V. Kreinovich. Imprecise probabilities in engineering analysis. *Mechanical Systems and Signal Processing*, 37:4–29, 2013.

Bernardini, A. and F. Tonon (2010). *Bounding Uncertainty in Civil Engineering - Theoretical Background.* Berlin: Springer.

de Angelis, M., Patelli, E. and M. Beer. Advanced Line Sampling for efficient robust reliability analysis. *Structural Safety*, 52:170–182, 2015.

Do, M. D., Gao, W., Song, C. and M. Beer. Interval spectral stochastic finite element analysis of structures with aggregation of random field and bounded parameters. *International Journal for Numerical Methods in Engineering*, 108:1198–1229, 2016.

Faes, M. and D. Moens. Identification and quantification of spatial interval uncertainty in numerical models. *Computers and Structures*, 192:16–33, 2017.

Ferson, S., Ginzburg, L., Kreinovich, V., Myers, D. and K. Sentz (2003). Constructing probability boxes and Dempster-Shafer structures. Technical Report SAND2002-4015. Albuqueque and Livermore: Sandia National Laboratories.

Fetz, T. Improving the convergence of iterative importance sampling for computing upper and lower expectations. *Proceedings of Machine Learning Research*, 103:185–193, 2019.

Fetz, T. and M. Oberguggenberger. Imprecise random variables, random sets, and Monte Carlo simulation. *International Journal of Approximate Reasoning*, 78:252–264, 2016.

Gao, W., Song C. and F. Tin-Loi. Probabilistic interval analysis for structures with uncertainty. *Structural Safety*, 32:191–199, 2010.

Ghanem, R. and P. Spanos (1991). *Stochastic Finite Elements: A Spectral Approach.* New York: Springer.

Goodman, I. R. and H. T. Nguyen. Fuzziness and randomness. In C. Bertoluzza, M. Gil and D. A. Ralescu, editors, *Statistical modeling, analysis and management of fuzzy data*, pages 3–21. Heidelberg: Physica-Verlag, 2002.

Grisvard, P. (1985). *Elliptic problems in nonsmooth domains.* Boston: Pitman Advanced Publishing Program.

Hurtado, J. A., Alvarez, D. A. and J. A. Paredes. Interval reliability analysis under the specification of statistical information on the input variables. *Structural Safety*, 65:35–48, 2017.

Jiang C., Zheng, J. and X. Han. Probability-interval hybrid uncertainty analysis for structures with both aleatory and epistemic uncertainties: a review. *Structural and Multidisciplinary Optimization*, 57:2485–2502, 2018.

Kallenberg, O. (1997). *Foundations of Modern Probability*. New York: Springer.

Matthies, H. G. Stochastic finite elements: computational approaches to stochastic partial differential equations. *ZAMM. Zeitschrift für Angewandte Mathematik und Mechanik*, 88(11):849–873, 2008.

Molchanov, I. (2005). *Theory of Random Sets*. Berlin: Springer-Verlag.

Muscolino, G., Santoro, R. and A. Sofi. Reliability analysis of structures with interval uncertainties under stationary stochastic excitations. *Computer Methods in Applied Mechanics and Engineering*, 300:47–69, 2016.

Muscolino, G., Sofi, A. and M. Zingales. One-dimensional heterogeneous solids with uncertain elastic modulus in presence of long-range interactions: Interval versus stochastic analysis. *Computers and Structures*, 122:217–229, 2013.

Nečas, J. (1967). *Les méthodes directes en théorie des équations elliptiques*. Paris: Masson et Cie, Prague: Academia.

Nedeljković, J. *Partial differential equations with random set and random field coefficients*. PhD thesis, Universität Innsbruck, Austria, 2020. PhD Technische Wissenschaften.

Nguyen, H. T. (2006). *An Introduction to Random Sets*. Boca Raton: Chapman & Hall.

Oberguggenberger, M. Combined methods in nondeterministic mechanics. In I. Elishakoff and C. Soize, editors, *Nondeterministic Mechanics*, pages 263–356. Wien: Springer, 2013.

Oberguggenberger, M. Analysis and computation with hybrid random set stochastic models. *Structural Safety*, 52:233–243, 2015.

Oberguggenberger, M. and L. Wurzer. Random set solutions to stochastic wave equations. *Proceedings of Machine Learning Research*, 103:314–323, 2019.

Schmelzer, B. On solutions to stochastic differential equations with parameters modeled by random sets. *International Journal of Approximate Reasoning*, 51:1159–1171, 2010.

Schmelzer, B. Set-valued assessments of solutions to stochastic differential equations with random set parameters. *Journal of Mathematical Analysis and Applications*, 400:425–438, 2013.

Schmelzer, B., Oberguggenberger, M. and C. Adam. Efficiency of tuned mass dampers with uncertain parameters on the performance of structures under stochastic excitation. *Journal of Risk and Reliability*, 224:297–308, 2010.

Sofi, A. Structural response variability under spatially dependent uncertainty: Stochastic versus interval model. *Probabilistic Engineering Mechanics*, 42:78–86, 2015.

Sofi, A., Muscolino, G., Giunta, F. Propagation of uncertain structural properties described by imprecise Probability Density Functions via response surface method. *Probabilistic Engineering Mechanics*, 60:103020, 2020.

Sofi, A. and E. Romeo. A unified response surface framework for the interval and stochastic finite element analysis of structures with uncertain parameters. *Probabilistic Engineering Mechanics*, 54:25–36, 2018.

Tonon, F. Using random set theory to propagate epistemic uncertainty through a mechanical system. *Reliability Engineering & System Safety*, 85(1):169–181, 2004.

Trèves, F. (1975). *Basic linear partial differential equations*. New York: Academic Press.

Troffaes, M. C. M. Imprecise Monte Carlo simulation and iterative importance sampling for the estimation of lower previsions. *Mechanical Systems and Signal Processing*, 101:31–48, 2018.

Zhang, J. and M. D. Shields. On the quantification and efficient propagation of imprecise probabilities resulting from small datasets. *Mechanical Systems and Signal Processing*, 98:465–483, 2018.

On the Solution of Forward and Inverse Problems in Possibilistic Uncertainty Quantification for Dynamical Systems

Dominik Hose and Michael Hanss
Institute of Engineering and Computational Mechanics, University of Stuttgart
Pfaffenwaldring 9, 70569 Stuttgart, Germany
{*dominik.hose, michael.hanss*}*@itm.uni-stuttgart.de*

Abstract. In this contribution, we adress an apparent lack of methods for the robust analysis of dynamical systems when neither a precise statistical nor an entirely epistemic description of the present uncertainties is possible. Relying on recent results of possibilistic calculus, we revisit standard prediction and filtering problems and show how these may be solved in a numerically exact way.

Keywords: State Estimation, Filter Design, Possibility Theory, Uncertainty Quantification, Imprecise Probabilities

1. Introduction

In the analysis and control of dynamical systems (Skogestad and Postlethwaite, 2007), the forward problem usually involves a prediction about the future system behavior, while the inverse problem aims at finding the feasible configurations of past system states that are in agreement with the system dynamics and the obtained measurements. Therein, the inclusion of non-probabilistic models of uncertainty is becoming increasingly more important.

The available techniques for robust system analysis can be divided into two approaches: Stochastic approaches, on the one hand, look back on a long tradition in the quantification of aleatory uncertainty in dynamical systems with popular results, such as the Kalman filter (Kalman, 1960). However, most of the available techniques share a common denominator; when considering the various forms of noise, they rely on precise probability distributions for their description. On the other hand, purely set-based approaches, such as tube-based model predictive control (Langson et al., 2004), usually do not take into account any stochastic information about the disturbances. Instead, they consider only epistemic uncertainty by finding sets to which the disturbances are confined. The simultaneous consideration of both aleatory and epistemic uncertainties is a topic which is seldomly adressed.

Nonetheless, often a mix of statistical variability and an epistemic lack of knowledge is present in models of real-world processes and there is a need for sensible uncertainty quantification frameworks which can address both these types of uncertainty. In this contribution, we demonstrate how possibility theory (Dubois and Prade, 1988) originating from fuzzy set theory is one such framework, and we derive tools for a possibilistic analysis of linear time-invariant (LTI) systems. Possibilistic calculus is based on the idea of descriptions of imprecise probabilities, and hence possibilities

© 2021 *by D. Hose and M. Hanss.*

may be interpreted intuitively by viewing them as upper probabilities. This allows for a more general description of measurement error and process noise due to not being restricted to unique probability distributions. Essentially, this calculus consists of the analysis of nested confidence intervals for imprecisely defined probability distributions, for which we present intuitive techniques for prediction and estimation. Special emphasis is put on the preservation of possibility-probability consistency in all calculations, which is enabled by recent theoretical results.

2. Possibility Spaces

We briefly recount the basic terminology of possibility theory as formalized in (Dubois and Prade, 1988). A set-valued function $\Pi : 2^\Omega \to [0,1]$ defined on the power set 2^Ω of the universe of discourse Ω may be called a *possibility measure* if it satisfies three conditions similar to those of a probability measure. Explicitly, $\Pi(\emptyset) = 0$, $\Pi(\Omega) = 1$ and $\Pi(\bigcup_k U_k) = \sup_k \Pi(U_k)$ for countable collections of sets $U_k \subseteq \Omega$ have to be fulfilled. The dual *necessity measure* is defined by $\mathrm{N}(U) = 1 - \Pi(\Omega \setminus U)$ for all $U \subseteq \Omega$.

Let $\widetilde{X} : \Omega \to \mathcal{X}$ be an ($\mathcal{X}$-valued) *uncertain (random, fuzzy, etc.) variable*. Below, $\mathcal{S} = \mathcal{S}(\mathcal{X})$ denotes a σ-field on $\mathcal{X}$. In practical applications, typically $\mathcal{X} \subseteq \mathbb{R}^N$, and $\mathcal{S}$ is the corresponding Borel σ-algebra. The uncertain variable $\widetilde{X}$ possesses a *possibility distribution* $\Pi_{\widetilde{X}}(U) = \Pi(\widetilde{X}^{-1}(U)) = \Pi(\{\omega \in \Omega : \widetilde{X}(\omega) \in U\})$ for all $U \in \mathcal{S}$ and a function $\pi_{\widetilde{X}} : \mathcal{X} \to [0,1]$ induces a possibility distribution via $\Pi_{\widetilde{X}}(U) = \sup_{x \in U} \pi_{\widetilde{X}}(x)$ if it is a measurable function satisfying $\sup_{x \in \mathcal{X}} \pi_{\widetilde{X}}(x) = 1$. It is then called a *possibility density*. The *superlevel sets/α-cuts* are defined as $\mathcal{C}^\alpha_{\Pi_{\widetilde{X}}} = \{x \in \mathcal{X} : \pi_{\widetilde{X}}(x) > \alpha\}$.

It is well-known that possibility theory offers a general framework for the analysis of *imprecise probabilities* (Augustin et al., 2014), i.e. for the consideration of partially defined probability distributions on the measure space $(\mathcal{X}, \mathcal{S})$. The fundamental principle of *probability-possibility consistency* was first introduced in (Dubois and Prade, 1992). It states that a probability measure $\mathrm{P}_{\widetilde{X}}$ and a possibility measure $\Pi_{\widetilde{X}}$ on $(\mathcal{X}, \mathcal{S})$ are consistent if the probability of any event is bounded from above by its possibility, i.e. if

$$\mathrm{P}_{\widetilde{X}}(U) \leq \Pi_{\widetilde{X}}(U) \qquad \forall\, U \in \mathcal{S} \tag{1}$$

and consequently from below by the necessity, $\mathrm{N}_{\widetilde{X}}(U) \leq \mathrm{P}_{\widetilde{X}}(U)$. In short, we write $\mathrm{P}_{\widetilde{X}} \preceq \Pi_{\widetilde{X}}$. Notice that the necessity – and hence the lower probability – of any α-cut is always bounded according to $\mathrm{N}_{\widetilde{X}}(\mathcal{C}^\alpha_{\widetilde{X}}) \geq 1 - \alpha$ for all $\alpha \in [0,1]$. In order to show consistency, one does not need to check consistency for all possible events as in Eq.(1). Instead, it is shown in (Couso et al., 2001) that it suffices to check the condition

$$\mathrm{P}_{\widetilde{X}}\left(\mathcal{C}^\alpha_{\Pi_{\widetilde{X}}}\right) \geq 1 - \alpha \qquad \forall\, \alpha \in [0,1]. \tag{2}$$

Furthermore, consistency is not a one-to-one relationship. In particular, given a possibility measure $\Pi_{\widetilde{X}}$, the credal set

$$\mathfrak{C}\left(\Pi_{\widetilde{X}}\right) = \left\{\mathrm{P}_{\widetilde{X}} : \mathrm{P}_{\widetilde{X}} \preceq \Pi_{\widetilde{X}}\right\} \tag{3}$$

contains a potentially infinite number of elements, i.e. of consistent probability distributions. One important observation is that the credal set of a possibility distribution $\Pi^1_{\widetilde{X}}$ is included in that of a second one $\Pi^2_{\widetilde{X}}$, i.e. $\mathfrak{C}(\Pi^1_{\widetilde{X}}) \subseteq \mathfrak{C}(\Pi^2_{\widetilde{X}})$ if and only if $\pi^1_{\widetilde{X}}(x) \leq \pi^2_{\widetilde{X}}(x)$ for all $x \in \mathcal{X}$. Then, $\Pi^1_{\widetilde{X}}$ is said to be *more specific* than $\Pi^2_{\widetilde{X}}$.

For an efficient analysis, we propose a quadruple description of convex fuzzy variables $\widetilde{\mathbf{X}} = (\mathbf{A}, \mathbf{b}, \mathbf{C}, \mathbf{d})$, an N_X^{ext}-dimensional generalization of the well-known one-dimensional fuzzy numbers. Therein, the matrix $\mathbf{A} = \mathbf{A}(\alpha) \in \mathbb{R}^{M_X^{\text{int}} \times N_X^{\text{int}}}$ and the vector $\mathbf{b} = \mathbf{b}(\alpha) \in \mathbb{R}^{M_X^{\text{int}}}$ are associated with an α-dependent internal halfspace. The matrix $\mathbf{C} \in \mathbb{R}^{N_X^{\text{ext}} \times N_X^{\text{int}}}$ and the vector $\mathbf{d} \in \mathbb{R}^{N_X^{\text{ext}}}$ provide the external description. The convex α-cuts of $\widetilde{\mathbf{X}}$, i.e. the superlevel sets of its possibility density, are given by

$$\mathcal{C}^{\alpha}_{\widetilde{\mathbf{X}}} = \left\{ \mathbf{x} = \mathbf{C} \cdot \xi + \mathbf{d} \ : \ \xi \in \mathbb{R}^{N_X^{\text{int}}} \wedge \mathbf{A}(\alpha) \cdot \xi \leq \mathbf{b}(\alpha) \right\} \qquad \forall\, \alpha \in [0,1]. \tag{4}$$

See Fig. 1 vor a visual representation of the α-cuts of a 2-dimensional convex fuzzy variable.

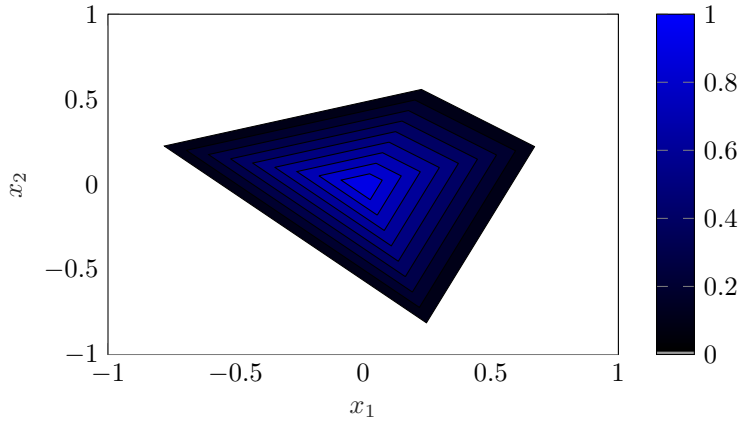


Figure 1. Convex α-cuts of a 2-dimensional convex fuzzy variable.

The possibility distribution may be reconstructed by the Decomposition Theorem

$$\pi_{\widetilde{\mathbf{X}}}(\mathbf{x}) = \sup \left\{ \alpha \in [0,1] : \mathbf{x} \in \mathcal{C}^{\alpha}_{\widetilde{\mathbf{X}}} \right\} \qquad \forall\, \mathbf{x} \in \mathbb{R}^{N_X^{\text{ext}}} \tag{5}$$

which is e.g. found in (Hanss, 2005). For a variety of reasons, this representation is well suited for the uncertainty analysis of linear time-invariant systems: Firstly, through the division in an internal and an external description, it is possible to account for interdependency of multivariate uncertain variables and, thus, to avoid certain effects of overestimation. Secondly, the rules of possibilistic calculus may be evaluated very efficiently for a selection of operations. And thirdly, it naturally fits into optimization frameworks, given the convexity of the α-cuts. The benefits are also highlighted below.

3. Possibilistic Calculus

In the following, selected elements of possibilistic calculus are presented. The guiding principle behind the specific implementations of these operations is the *preservation of consistency*. That is, if a possibility distribution and a probability distribution were consistent before performing an operation, such as the the aggregation of several distributions or the propagation through a model, then the resulting distributions ought to be consistent as well. This ensures that possibilities can truly be seen as upper probabilities. Special emphasis is put on the implementation for convex fuzzy variables.

3.1. Aggregation

Often only the marginal distributions $\pi_{\widetilde{X}_1}, \dots, \pi_{\widetilde{X}_N}$ on the distinct domains $\mathcal{X}_1, \dots, \mathcal{X}_N$ are available. Yet, e.g. the propagation and the inversion operations discussed below require the joint distribution which has to be constructed by a suitable aggregation operator. Consistency-preserving aggregation operators are presented e.g. in (Destercke et al., 2009) and, more recently, in (Hose and Hanss, 2019b). In the latter, it is shown that if the probability distributions in the respective credal sets are assumed to be independent, then a consistency-preserving joint possibility distribution is given by

$$\pi^{\text{ind}}_{\widetilde{X}_1,\dots,\widetilde{X}_N}(x_1,\dots,x_N) = \min_{i=1,\dots,N} 1 - \Big(1 - \pi_{\widetilde{X}_i}(x_i)\Big)^N \quad \forall\, (x_1,\dots,x_N) \in \mathcal{X}_1 \times \dots \times \mathcal{X}_N. \tag{6}$$

From the proof of Theorem 28 in (Hose and Hanss, 2019b), we infer that Eq. (6) is equivalent to

$$\mathcal{C}^{\alpha}_{\Pi^{\text{ind}}_{\widetilde{X}_1,\dots,\widetilde{X}_n}} = \mathcal{C}^{1-\sqrt[N]{1-\alpha}}_{\Pi_{\widetilde{X}_1}} \times \dots \times \mathcal{C}^{1-\sqrt[N]{1-\alpha}}_{\Pi_{\widetilde{X}_n}} \qquad \forall\, \alpha \in [0,1]. \tag{7}$$

Hence, if $\widetilde{X}_i = (\mathbf{A}_{\widetilde{X}_i}, \mathbf{b}_{\widetilde{X}_i}, \mathbf{C}_{\widetilde{X}_i}, \mathbf{d}_{\widetilde{X}_i})$ for $i = 1, \dots, N$ are N given convex fuzzy variables, then the joint fuzzy vector $\widetilde{\mathbf{X}} = (\mathbf{A}_{\widetilde{\mathbf{X}}}, \mathbf{b}_{\widetilde{\mathbf{X}}}, \mathbf{C}_{\widetilde{\mathbf{X}}}, \mathbf{d}_{\widetilde{\mathbf{X}}})$ is also a convex fuzzy variable with the matrices

$$\mathbf{A}_{\widetilde{\mathbf{X}}} = \begin{bmatrix} \mathbf{A}_{\widetilde{X}_1} \circ \Gamma^N & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{A}_{\widetilde{X}_N} \circ \Gamma^N \end{bmatrix},\ \mathbf{b}_{\widetilde{\mathbf{X}}} = \begin{bmatrix} \mathbf{b}_{\widetilde{X}_1} \circ \Gamma^N \\ \vdots \\ \mathbf{b}_{\widetilde{X}_N} \circ \Gamma^N \end{bmatrix},\ \mathbf{C}_{\widetilde{\mathbf{X}}} = \begin{bmatrix} \mathbf{C}_{\widetilde{X}_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{C}_{\widetilde{X}_N} \end{bmatrix},\ \mathbf{d}_{\widetilde{\mathbf{X}}} = \begin{bmatrix} \mathbf{d}_{\widetilde{X}_1} \\ \vdots \\ \mathbf{d}_{\widetilde{X}_N} \end{bmatrix}$$

where $\Gamma^N : \alpha \mapsto 1 - \sqrt[N]{1-\alpha}$ is a (monotone) rescaling function.

3.2. Propagation

The pushforward distribution of a fuzzy variable $\widetilde{Y} = \phi(\widetilde{X})$ under a $(\mathcal{X}, \mathcal{Y})$-measurable function $\phi : \mathcal{X} \to \mathcal{Y}$ is given by the extension principle (Zadeh, 1975)

$$\pi_{\widetilde{Y}}(y) = \sup_{x \in \mathcal{X}: y = \phi(x)} \pi_{\widetilde{X}}(x) \qquad \forall\, y \in \mathcal{Y}. \tag{8}$$

The preservation of consistency is discussed e.g. in (Baudrit et al., 2006) and (Hose and Hanss, 2019b). Of course, if only the marginals $\widetilde{X}_1, \ldots, \widetilde{X}_N$ are known, then they need to be aggregated into the joint vector $\widetilde{\mathbf{X}}$ beforehand. E.g. in (Baudrit et al., 2006) it is shown that Zadeh's non-interactive aggregation, i.e. the min-intersection of the possibility densities, which appears in the original formulation of the extension principle, does not preserve consistency.

The connection to the propagation of intervals is discussed in (Hanss, 2005). In particular, it can be shown that Eq. (8) is equivalent to

$$\mathcal{C}^{\alpha}_{\Pi_{\widetilde{Y}}} = \phi\left(\mathcal{C}^{\alpha}_{\Pi_{\widetilde{X}}}\right) \qquad \forall\, \alpha \in [0,1]. \tag{9}$$

Hence, if $\widetilde{\mathbf{X}} = (\mathbf{A}, \mathbf{b}, \mathbf{C}, \mathbf{d})$ is a given convex fuzzy variable and ϕ is a linear map represented by the matrix $\mathbf{M} \in \mathbb{R}^{N_{\mathcal{Y}} \times N_{\mathcal{X}}}$, then $\widetilde{\mathbf{Y}} = (\mathbf{A}, \mathbf{b}, \mathbf{MC}, \mathbf{Md})$ is also a convex fuzzy variable. E.g. the sum $\widetilde{Y} = \sum_{i=1}^{N} \widetilde{X}_i$ may then be obtained by propagation of $\widetilde{\mathbf{X}}$ under the matrix $\mathbf{M} = [1, \ldots, 1]$.

Marginalization, i.e. the computation of the marginal distributions $\pi_{\widetilde{X}_1}, \ldots, \pi_{\widetilde{X}_N}$ from the joint distribution $\pi_{\widetilde{X}_1,\ldots,\widetilde{X}_N}$ is just a special case of propagation under $\phi : \mathbf{x} \mapsto x_i$ for $i = 1, \ldots N$. Hence, it preserves consistency as well.

3.3. Inversion

The preservation of consistency for inversion under a surjective and $(\mathcal{X}, \mathcal{Y})$-measurable function $\phi : \mathcal{X} \to \mathcal{Y}$ is investigated in (Hose and Hanss, 2019a). Suppose the possibility distribution of $\widetilde{Y}$ is known and one wishes to infer the distribution of $\widetilde{X}$. Generally, there exists a (possibly infinite) number of possibility distributions $\Pi_{\widetilde{X}}$ yielding the pushforward distribution $\Pi_{\widetilde{Y}}$ under ϕ. These extensions may be gathered in the *set of inverse possibility distributions*

$$\mathcal{I}^{\phi}_{\Pi_{\widetilde{Y}}} = \left\{\Pi_{\widetilde{X}} : \Pi_{\widetilde{Y}}(V) = \Pi_{\widetilde{X}}\left(\phi^{-1}(V)\right)\ \forall\, V \in \mathcal{S}(\mathcal{Y})\right\}. \tag{10}$$

However, it is possible to account for this whole set by just one possibility distribution. The *minimum specific inverse possibility distribution* is given by

$$\pi^{\mathrm{inv}}_{\widetilde{X}}(x) = \Pi_{\widetilde{Y}}(\phi(x)) \qquad \forall\, x \in \mathcal{X} \tag{11}$$

This distribution is the least specific possibility distribution in the corresponding set of inverse possibility distributions and therefore a suitable representation thereof. Analogously, one can define a set of inverse probability distributions $\mathcal{I}^{\phi}_{\mathrm{P}_{\widetilde{Y}}}$ given a probability distribution $\mathrm{P}_{\widetilde{Y}}$. The perhaps most important result is that $\Pi^{\mathrm{inv}}_{\widetilde{X}}$ is consistent with all probability distributions $\mathrm{P}_{\widetilde{X}} \in \mathcal{I}^{\phi}_{\mathrm{P}_{\widetilde{Y}}}$ from the set of inverse probability distributions $\mathcal{I}^{\phi}_{\mathrm{P}_{\widetilde{Y}}}$ of all probability distributions $\mathrm{P}_{\widetilde{Y}} \in \mathcal{C}\left(\Pi_{\widetilde{Y}}\right)$ which are consistent with $\Pi_{\widetilde{Y}}$. Analogously to the forward propagation, we infer that Eq. (11) is equivalent to

$$\mathcal{C}^{\alpha}_{\Pi^{\mathrm{inv}}_{\widetilde{X}}} = \phi^{-1}\left(\mathcal{C}^{\alpha}_{\Pi_{\widetilde{Y}}}\right) \qquad \forall\, \alpha \in [0,1]. \tag{12}$$

Hence, if $\widetilde{\mathbf{Y}} = \left(\mathbf{A}_{\widetilde{\mathbf{Y}}}, \mathbf{b}_{\widetilde{\mathbf{Y}}}, \mathbf{C}_{\widetilde{\mathbf{Y}}}, \mathbf{d}_{\widetilde{\mathbf{Y}}}\right)$ is a given convex fuzzy variable and ϕ is a linear map, i.e. it corresponds to a matrix $\mathbf{M} \in \mathbb{R}^{N_{\mathbf{Y}} \times N_{\mathbf{X}}}$, then one obtains the convex fuzzy variable $\widetilde{\mathbf{X}} =$

$\left(\mathbf{A}_{\widetilde{\mathbf{X}}}, \mathbf{b}_{\widetilde{\mathbf{X}}}, \mathbf{C}_{\widetilde{\mathbf{X}}}, \mathbf{d}_{\widetilde{\mathbf{X}}}\right)$ with the matrices

$$\mathbf{A}_{\widetilde{\mathbf{X}}} = \begin{bmatrix} \mathbf{A}_{\widetilde{\mathbf{Y}}} & \mathbf{0} \\ -\mathbf{C}_{\widetilde{\mathbf{Y}}} & \mathbf{M} \\ \mathbf{C}_{\widetilde{\mathbf{Y}}} & -\mathbf{M} \end{bmatrix}, \quad \mathbf{b}_{\widetilde{\mathbf{X}}} = \begin{bmatrix} \mathbf{b}_{\widetilde{\mathbf{Y}}} \\ \mathbf{d}_{\widetilde{\mathbf{Y}}} \\ \mathbf{d}_{\widetilde{\mathbf{Y}}} \end{bmatrix}, \quad \mathbf{C}_{\widetilde{\mathbf{X}}} = \begin{bmatrix} \mathbf{0} & \mathbb{I} \end{bmatrix} \quad \text{and} \quad \mathbf{d}_{\widetilde{\mathbf{X}}} = \mathbf{0}$$

from the minimum specific inverse possibility distribution. This operation augments the vector of interior variables to $\xi_{\widetilde{\mathbf{X}}} = \begin{bmatrix} \xi_{\widetilde{\mathbf{Y}}}^{\mathrm{T}} & \mathbf{x}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ where $\xi_{\widetilde{\mathbf{Y}}}^{\mathrm{T}}$ are the interior variables of $\widetilde{\mathbf{Y}}$. Obviously, the lower two rows of the internal halfspace representation correspond to the equality constraint $\mathbf{C}_{\widetilde{\mathbf{Y}}}\xi_{\widetilde{\mathbf{Y}}} + \mathbf{d}_{\widetilde{\mathbf{Y}}} = \mathbf{M}\mathbf{x}$ which might as well be considered explicitly.

3.4. Confidence Sets

For the inverse problem, it is furthermore helpful to study the relation between α-cuts and confidence sets. Assume an uncertain variable $\widetilde{Y} = \psi(\vartheta, \widetilde{X})$ be a function of the uncertain variable $\widetilde{X}$ with known possibility distribution and of an unknown parameter $\vartheta \in \Theta$. Adopting a frequentist point of view, ϑ is not an uncertain variable, and hence, probabilistic or possibilistic expressions about it require discretion. Nevertheless, with the assumed knowledge, we can construct a confidence set $\mathcal{K}_{\vartheta}^{\alpha} = \mathcal{K}_{\vartheta}^{\alpha}(\widetilde{Y}) = \mathcal{K}_{\vartheta}^{\alpha}(\phi(\vartheta, \widetilde{X}))$ of ϑ for a given confidence level $\alpha \in [0, 1]$ through

$$\mathcal{K}_{\vartheta}^{\alpha} = \left\{ \theta \in \Theta : \widetilde{Y} \in \psi\left(\theta, \mathcal{C}_{\widetilde{X}}^{\alpha}\right) \right\}. \tag{13}$$

Proposition 1. *It holds that* $\mathrm{P}\left(\vartheta \in \mathcal{K}_{\vartheta}^{\alpha}\right) \geq 1 - \alpha$ *for all* $\vartheta \in \Theta$.

Proof. Let $\vartheta \in \Theta$ be arbitrary but fixed. From $\widetilde{Y} = \psi\left(\vartheta, \widetilde{X}\right)$, it follows that

$$\mathcal{K}_{\vartheta}^{\alpha} = \left\{ \theta \in \Theta : \psi\left(\vartheta, \widetilde{X}\right) \in \psi\left(\theta, \mathcal{C}_{\widetilde{X}}^{\alpha}\right) \right\}. \tag{14}$$

The set membership definition for $\vartheta \in \mathcal{K}_{\vartheta}^{\alpha}$, i.e. $\psi(\vartheta, \widetilde{X}) \in \psi(\vartheta, \mathcal{C}_{\widetilde{X}}^{\alpha})$, is trivially fulfilled if $\widetilde{X} \in \mathcal{C}_{\widetilde{X}}^{\alpha}$. Hence, the probability of the former being fulfilled is greater than the probability of the latter being fulfilled, yielding

$$\mathrm{P}\left(\vartheta \in \mathcal{K}_{\vartheta}^{\alpha}\right) \geq \mathrm{P}\left(\widetilde{X} \in \mathcal{C}_{\widetilde{X}}^{\alpha}\right) \geq 1 - \alpha \tag{15}$$

which concludes the proof. □

Notice that the proof of Proposition 1 requires the model ψ to be 'correct'. If ψ is not surjective, then it is possible that the confidence levels may be empty above a certain level γ^*. The actual value of γ^* can be computed by checking the existence of feasible solutions, e.g. within a bisection algorithm. Due to this result, we argue that – in possibility theory – α-cuts and confidence sets are similar concepts for uncertain variables and unknown parameters, respectively. Having observed $\widetilde{Y} = y$, we can even define a corresponding confidence density $\gamma : \Theta \to [0, 1]$ which is computed from the possibility density of $\widetilde{X}$ through

$$\gamma_{\hat{\vartheta}}(\vartheta) = \sup_{x \in \mathcal{X} : y = \psi(\vartheta, x)} \pi_{\widetilde{X}}(x) \tag{16}$$

Furthermore, confidence sets can be computed very efficiently from convex fuzzy variables. Let $\widetilde{\mathbf{X}} = \left(\mathbf{A}_{\widetilde{\mathbf{X}}}, \mathbf{b}_{\widetilde{\mathbf{X}}}, \mathbf{C}_{\widetilde{\mathbf{X}}}, \mathbf{d}_{\widetilde{\mathbf{X}}}\right)$ be a given convex fuzzy variable and let $\psi : (\vartheta, \mathbf{x}) \mapsto \mathbf{M}\mathbf{x} + \mathbf{N}\vartheta$ be an affine map for suitably sized matrices $\mathbf{M}$ and $\mathbf{N}$. Then, an estimator of ϑ can also be expressed in quadruple form $\hat{\vartheta} = \left(\mathbf{A}_{\hat{\vartheta}}, \mathbf{b}_{\hat{\vartheta}}, \mathbf{C}_{\hat{\vartheta}}, \mathbf{d}_{\hat{\vartheta}}\right)$ where

$$\mathbf{A}_{\hat{\vartheta}} = \begin{bmatrix} \mathbf{A}_{\widetilde{\mathbf{X}}} & \mathbf{0} \\ \mathbf{M}\mathbf{C}_{\widetilde{\mathbf{X}}} & \mathbf{N} \\ -\mathbf{M}\mathbf{C}_{\widetilde{\mathbf{X}}} & -\mathbf{N} \end{bmatrix}, \quad \mathbf{b}_{\hat{\vartheta}} = \begin{bmatrix} \mathbf{b}_{\widetilde{\mathbf{X}}} \\ \mathbf{y} - \mathbf{M}\mathbf{d}_{\widetilde{\mathbf{X}}} \\ \mathbf{y} - \mathbf{M}\mathbf{d}_{\widetilde{\mathbf{X}}} \end{bmatrix}, \quad \mathbf{C}_{\hat{\vartheta}} = \begin{bmatrix} \mathbf{0} & \mathbb{I} \end{bmatrix} \quad \text{and} \quad \mathbf{d}_{\hat{\vartheta}} = \mathbf{0}$$

Therein, the lower two rows of the internal halfspace representation correspond to the equality constraint $\mathbf{M}\mathbf{C}_{\widetilde{\mathbf{X}}}\xi_{\widetilde{\mathbf{X}}} + \mathbf{M}\mathbf{d}_{\widetilde{\mathbf{X}}} + \mathbf{N}\theta = \mathbf{y}$ which might as well be considered explicitly.

Again, even though similar in appearance, a confidence density is not a possibility density and a confidence set is not an α-cut from a frequentist point of view. Likewise, even though the computation is similar, the underlying concepts used for the inversion and for the confidence sets are fundamentally different. The inversion intends to find an actual possibility distribution for an input variable from the possibility distribution of a dependent output variable, whereas the confidence distribution indicates sets in which an unknown but fixed parameter is likely to reside.

In a subjective setting, however, it seems reasonable to view $1 - \alpha$ as the supremum acceptable buying price to pay for the gamble $f_\alpha : \theta \mapsto \mathcal{I}_{\mathcal{K}_{\hat{\vartheta}}^\alpha}(\theta)$, where $\mathcal{I}_{\mathcal{K}_{\hat{\vartheta}}^\alpha}$ is the indicator function of the confidence set $\mathcal{K}_{\hat{\vartheta}}^\alpha$ for all $\alpha \in [0, 1]$ and, subsequently, to define a coherent lower prevision (Augustin et al., 2014) therefrom.

4. Possibilistic System Analysis

The elements of possibilistic calculus in combination with convex fuzzy variables as proposed and described above may be used as a basis for the analysis of LTI systems with possibilistic error descriptions via convex fuzzy variables. For simplicity, we assume the following time-discrete state-space description, omitting any control input or other, which could, however, easily be included. The system dynamics is modeled by

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{v}_k \tag{17}$$

and the observation model is

$$\mathbf{y}_k = \mathbf{H}\mathbf{x}_k + \mathbf{w}_k. \tag{18}$$

Therein, the process noise $\mathbf{v}_k$ and the measurement error $\mathbf{w}_k$ are assumed to be independent and identically distributed (iid) realizations of the uncertain variables $\widetilde{\mathbf{V}}$ and $\widetilde{\mathbf{W}}$ which are modelled as convex fuzzy variables.

4.1. Prediction

For the purpose of predicting the future system behavior, we assume that information about the current system state $\mathbf{x}_k$ is present in the form of a convex fuzzy variable $\widetilde{\mathbf{X}}_k$. This knowledge

facilitates the prediction of the system states $\mathbf{x}_{k+1}, \mathbf{x}_{k+2}, \ldots$ and of the system output $\mathbf{y}_k, \mathbf{y}_{k+1}, \ldots$ at future points in time.

4.1.1. *State Prediction*

State prediction is mostly concerned with the evaluation of the system dynamics (17) and corresponds to a simple aggregation and propagation. In order to predict $\mathbf{x}_{k+1}$ from $\widetilde{\mathbf{X}}_k$, we consider the joint distribution of the vector $\left[\widetilde{\mathbf{X}}_k^{\mathrm{T}} \;\; \widetilde{\mathbf{V}}^{\mathrm{T}}\right]^{\mathrm{T}}$. Propagation under $\mathbf{M} = \left[\mathbf{F} \;\; \mathbb{I}\right]$ yields the convex fuzzy variable $\widetilde{\mathbf{X}}_{k+1}$ with the matrices

$$\mathbf{A}_{\widetilde{\mathbf{X}}_{k+1}} = \begin{bmatrix} \mathbf{A}_{\widetilde{\mathbf{X}}_k} \circ \Gamma^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{\widetilde{\mathbf{V}}} \circ \Gamma^2 \end{bmatrix}, \qquad \mathbf{b}_{\widetilde{\mathbf{X}}_{k+1}} = \begin{bmatrix} \mathbf{b}_{\widetilde{\mathbf{X}}_k} \circ \Gamma^2 \\ \mathbf{b}_{\widetilde{\mathbf{V}}} \circ \Gamma^2 \end{bmatrix},$$

$$\mathbf{C}_{\widetilde{\mathbf{X}}_{k+1}} = \left[\mathbf{F}\mathbf{C}_{\widetilde{\mathbf{X}}_k} \;\; \mathbf{C}_{\widetilde{\mathbf{V}}}\right] \qquad \text{and} \qquad \mathbf{d}_{\widetilde{\mathbf{X}}_{k+1}} = \mathbf{F}\mathbf{d}_{\widetilde{\mathbf{X}}_k} + \mathbf{d}_{\widetilde{\mathbf{V}}}.$$

This step can also be applied recursively in order to predict $\widetilde{\mathbf{X}}_{k+2}$, $\widetilde{\mathbf{X}}_{k+3}$, etc.

Proposition 2. *The system state* $\widetilde{\mathbf{X}}_{k+1} = \mathbf{F}\widetilde{\mathbf{X}}_k + \widetilde{\mathbf{V}}$ *will fall into the* α*-cut* $\mathcal{C}^{\alpha}_{\widetilde{\mathbf{X}}_{k+1}}$ *with probability*

$$\mathrm{P}\left(\widetilde{\mathbf{X}}_{k+1} \in \mathcal{C}^{\alpha}_{\widetilde{\mathbf{X}}_{k+1}}\right) \geq 1 - \alpha \qquad \forall\, \alpha \in [0,1], \tag{19}$$

for all independent and consistent probability distributions $\mathrm{P}_{\widetilde{\mathbf{X}}_k} \preceq \Pi_{\widetilde{\mathbf{X}}_k}$, *and* $\mathrm{P}_{\widetilde{\mathbf{V}}} \preceq \Pi_{\widetilde{\mathbf{V}}}$.

Proof. The proposition follows directly from the consistency-preserving properties of the aggregation and propagation. □

4.1.2. *Output Prediction*

Information about the output $\mathbf{y}_k$ may be obtained by additionally considering the measurement model (18). By propagation of the joint vector $\left[\widetilde{\mathbf{X}}_k^{\mathrm{T}} \;\; \widetilde{\mathbf{W}}^{\mathrm{T}}\right]^{\mathrm{T}}$ under $\mathbf{M} = \left[\mathbf{H} \;\; \mathbb{I}\right]$ one obtains the convex fuzzy variable $\widetilde{\mathbf{Y}}_k$ with the matrices

$$\mathbf{A}_{\widetilde{\mathbf{Y}}_k} = \begin{bmatrix} \mathbf{A}_{\widetilde{\mathbf{X}}_k} \circ \Gamma^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{\widetilde{\mathbf{W}}} \circ \Gamma^2 \end{bmatrix}, \qquad \mathbf{b}_{\widetilde{\mathbf{Y}}_k} = \begin{bmatrix} \mathbf{b}_{\widetilde{\mathbf{X}}_k} \circ \Gamma^2 \\ \mathbf{b}_{\widetilde{\mathbf{W}}} \circ \Gamma^2 \end{bmatrix},$$

$$\mathbf{C}_{\widetilde{\mathbf{Y}}_k} = \left[\mathbf{H}\mathbf{C}_{\widetilde{\mathbf{X}}_k} \;\; \mathbf{H}\mathbf{C}_{\widetilde{\mathbf{W}}}\right] \qquad \text{and} \qquad \mathbf{d}_{\widetilde{\mathbf{Y}}_k} = \mathbf{H}\left(\mathbf{d}_{\widetilde{\mathbf{X}}_k} + \mathbf{d}_{\widetilde{\mathbf{W}}}\right).$$

Proposition 3. *The measurement* $\widetilde{\mathbf{Y}}_k = \mathbf{H}\widetilde{\mathbf{X}}_k + \widetilde{\mathbf{W}}$ *will fall into the* α*-cut* $\mathcal{C}^{\alpha}_{\widetilde{\mathbf{Y}}_k}$ *with probability*

$$\mathrm{P}\left(\widetilde{\mathbf{Y}}_k \in \mathcal{C}^{\alpha}_{\widetilde{\mathbf{Y}}_k}\right) \geq 1 - \alpha \qquad \forall\, \alpha \in [0,1] \tag{20}$$

for all independent and consistent probability distributions $\mathrm{P}_{\widetilde{\mathbf{X}}_k} \preceq \Pi_{\widetilde{\mathbf{X}}_k}$, *and* $\mathrm{P}_{\widetilde{\mathbf{W}}} \preceq \Pi_{\widetilde{\mathbf{W}}}$.

Proof. The proposition follows directly from the consistency-preserving properties of the aggregation and propagation. □

4.2. Estimation

The elicitation of information about the system state is performed in a similar manner by projecting the measurements and possible perturbations onto the state space. Considering the involved process noise and measurement error, it is possible to construct confidence sets of the actual system states $\mathbf{x}_k$ for further analysis. Adopting a frequentist point of view, the $\mathbf{x}_k$ are not uncertain variables as they have already been realized.

To this end, we assume that the measurements are available in the vector $\widetilde{\mathbf{P}} = \left[\widetilde{\mathbf{Y}}_1^{\mathrm{T}} \ \dots \ \widetilde{\mathbf{Y}}_k^{\mathrm{T}}\right]^{\mathrm{T}}$. Evidently, these measurements only depend on the uncertain variables $\widetilde{\mathbf{Q}} = \left[\widetilde{\mathbf{V}}^{\mathrm{T}} \ \dots \ \widetilde{\mathbf{V}}^{\mathrm{T}}\right]^{\mathrm{T}}$, containing $k-1$ instances of the process noise $\widetilde{\mathbf{V}}$, and $\widetilde{\mathbf{R}} = \left[\widetilde{\mathbf{W}}^{\mathrm{T}} \ \dots \ \widetilde{\mathbf{W}}^{\mathrm{T}}\right]^{\mathrm{T}}$, containing k instances of the measurement error $\widetilde{\mathbf{W}}$. If we denote by $\vartheta = \left[\mathbf{x}_1^{\mathrm{T}} \ \dots \ \mathbf{x}_k^{\mathrm{T}}\right]^{\mathrm{T}}$ the vector of past system states which are to be estimated, then Eqs. (17) and (18) can be written in matrix-vector form as

$$\begin{bmatrix} \mathbf{0} \\ \widetilde{\mathbf{P}} \end{bmatrix} = \underbrace{\left[\begin{array}{cccc} -\mathbf{F} & \mathbb{I} & & \mathbf{0} \\ & \ddots & \ddots & \\ \mathbf{0} & & -\mathbf{F} & \mathbb{I} \\ \hline -\mathbf{H} & & \mathbf{0} & \\ & \ddots & & \\ \mathbf{0} & & -\mathbf{H} & \end{array}\right]}_{=:\mathbf{N}} \vartheta + \begin{bmatrix} \widetilde{\mathbf{Q}} \\ \widetilde{\mathbf{R}} \end{bmatrix}. \tag{21}$$

For $\mathbf{t} = \left[\mathbf{0}^{\mathrm{T}} \ \ \mathbf{p}^{\mathrm{T}}\right]^{\mathrm{T}}$, where $\mathbf{p}$ is a realization of $\widetilde{\mathbf{P}}$, and the aggregated joint vector $\widetilde{\mathbf{S}} = \left[\widetilde{\mathbf{Q}}^{\mathrm{T}} \ \ \widetilde{\mathbf{R}}^{\mathrm{T}}\right]^{\mathrm{T}}$, it is possible to estimate ϑ by $\hat{\vartheta}$ with the matrices

$$\mathbf{A}_{\hat{\vartheta}} = \begin{bmatrix} \mathbf{A}_{\widetilde{\mathbf{S}}} & \mathbf{0} \\ \mathbf{C}_{\widetilde{\mathbf{S}}} & \mathbf{N} \\ -\mathbf{C}_{\widetilde{\mathbf{S}}} & -\mathbf{N} \end{bmatrix}, \quad \mathbf{b}_{\hat{\vartheta}} = \begin{bmatrix} \mathbf{b}_{\widetilde{\mathbf{S}}} \\ \mathbf{t} - \mathbf{d}_{\widetilde{\mathbf{S}}} \\ \mathbf{t} - \mathbf{d}_{\widetilde{\mathbf{S}}} \end{bmatrix}, \quad \mathbf{C}_{\hat{\vartheta}} = \begin{bmatrix} \mathbf{0} & \mathbb{I} \end{bmatrix} \quad \text{and} \quad \mathbf{d}_{\hat{\vartheta}} = \mathbf{0}.$$

The confidence sets $\mathcal{K}^{\alpha}_{\hat{\mathbf{x}}_i}$ of $\mathbf{x}_i$ for $i = 1, \dots, k$ can be obtained by marginalization of $\hat{\vartheta}$.

Proposition 4. *The (non-empty) α-cuts $\mathcal{K}^{\alpha}_{\hat{\vartheta}}$ are confidence sets of ϑ, i.e.*

$$\mathrm{P}\left(\vartheta \in \mathcal{K}^{\alpha}_{\hat{\vartheta}}\right) \geq 1 - \alpha \qquad \forall\, \alpha \in [0, \alpha^*] \tag{22}$$

for all independent and consistent probability distributions $\mathrm{P}_{\widetilde{\mathbf{V}}} \preceq \Pi_{\widetilde{\mathbf{V}}}$, and $\mathrm{P}_{\widetilde{\mathbf{W}}} \preceq \Pi_{\widetilde{\mathbf{W}}}$.

Proof. The proposition follows from the consistency-preserving properties of the aggregation, and Proposition 1. □

5. Numerical Example

Consider the LTI system with the system and measurement matrices

$$\mathbf{F} = \begin{bmatrix} \frac{1}{6} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{6} \end{bmatrix} \qquad \text{and} \qquad \mathbf{H} = \begin{bmatrix} 1 & 0 \end{bmatrix}. \tag{23}$$

The initial conditions are chosen as $\mathbf{x}_0 = \begin{bmatrix} 0 & 40 \end{bmatrix}^{\mathrm{T}}$, i.e. they are assumed to be known precisely, even though it would suffice to only know a possibility distribution. Furthermore, it is assumed that the process noise $\widetilde{V}_1$ and $\widetilde{V}_2$ each are distributed according to some probability distribution which is unimodal, symmetric about zero and has a support bounded by $|\widetilde{V}_i| \leq v^{\max} = \frac{1}{10}$ for $i \in \{1, 2\}$. Dubois et al. show in (Dubois et al., 2004) that a triangular fuzzy number with possibility density $\pi_{\widetilde{V}_i}(v) = 1 - \frac{|v|}{v^{\max}}$ for $|v| \leq v^{\max}$ and zero outside induces a maximally specific possibility distribution which is consistent with all admissible probability distributions of $\widetilde{V}_i$. See Figure 2 for a visualization of the α-cuts of $\widetilde{V}_1$ and $\widetilde{V}_2$. Considering their independent aggregation, their joint vector $\widetilde{\mathbf{V}}$ is described by

$$\mathbf{A}_{\widetilde{\mathbf{V}}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \mathbf{b}_{\widetilde{\mathbf{V}}} = v^{\max}\sqrt{1-\alpha} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{C}_{\widetilde{\mathbf{V}}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } \quad \mathbf{d}_{\widetilde{\mathbf{V}}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \forall\, \alpha \in [0, 1].$$

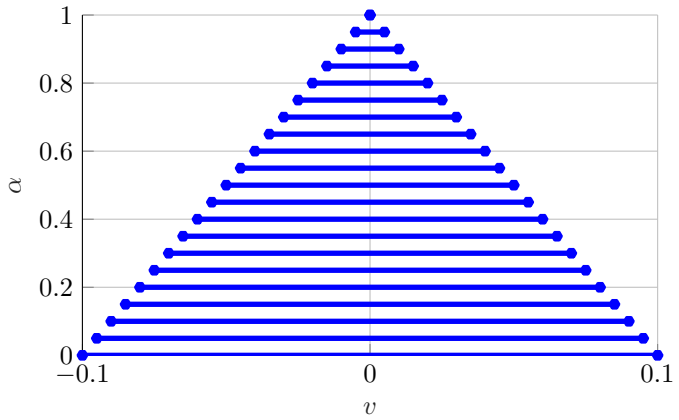


Figure 2. α-cuts of the triangular fuzzy numbers $\widetilde{V}_1$ and $\widetilde{V}_2$.

The measurement noise $\widetilde{W}$, is assumed to follow a probability distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$. Further information about additional moments is not provided. In (Baudrit and

Dubois, 2006), it is argued that, by means of the Chebychev inequality, a suitable representation of such knowledge is given by the possibility density

$$\pi_{\widetilde{W}}(w) = \min\left(1, \frac{\sigma^2}{(\mu - w)^2}\right) \qquad \forall\, w \in \mathbb{R}. \tag{24}$$

which can be expressed by

$$\mathbf{A}_{\widetilde{W}} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathbf{b}_{\widetilde{W}} = \frac{\sigma}{\sqrt{\alpha}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{C}_{\widetilde{W}} = 1 \quad \text{and} \quad \mathbf{d}_{\widetilde{W}} = \mu \qquad \forall\, \alpha \in (0, 1].$$

See Figure 3 for a visualization of the α-cuts of $\widetilde{W}$. The support, i.e. the α-cut for $\alpha = 0$, is simply $\mathcal{C}^0_{\widetilde{W}} = \mathbb{R}$.

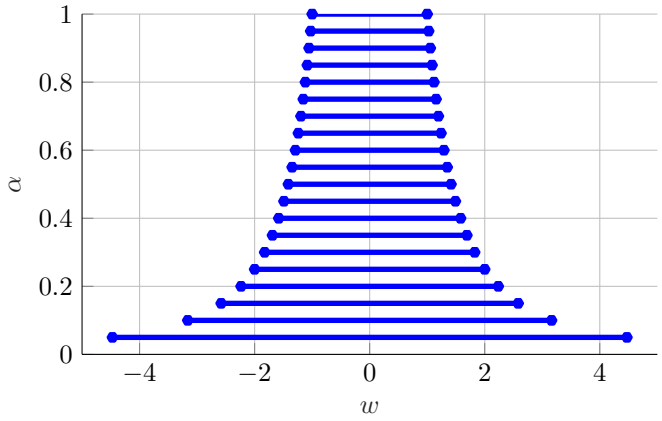


Figure 3. α-cuts of the Chebychev possibility distribution of $\widetilde{W}$.

In the reference simulation, the 'experimental' data y_k^{exp} were generated with consistent probability distributions for $\widetilde{\mathbf{V}}$ and $\widetilde{W}$. In particular, the system noise was generated from uniform probability distributions on the given support $[-\frac{1}{10}, \frac{1}{10}]$ and the measurement noise from a mixture with equal weights of two normal distributions with means $\pm\sqrt{\frac{3}{4}}$ and variance $\frac{1}{4}$ as shown in Figure 4.

An exemplary prediction of the output measurements is provided in Figure 5. Evidently, all data points lie within the support of the prediction with a tendency to regions with a higher possibility density. The obtained results can even be used to verify Proposition 3; the empirical probability of the reference measurements y_k^{exp} being contained in the α-cuts $\mathcal{C}^{\alpha}_{\widetilde{Y}_k}$ stays well above the guaranteed value of $1 - \alpha$, see Figure 6.

Similarly, the descriptions of process noise and measurement error can be employed to estimate the states x_1 and x_2 from the experimental data. The reference data are estimated accurately, i.e. they tend to lie in high-membership regions, and the order of magnitude of the area of the confidence sets is reasonably small in order to enable meaningful inference. Again, the relative frequency of

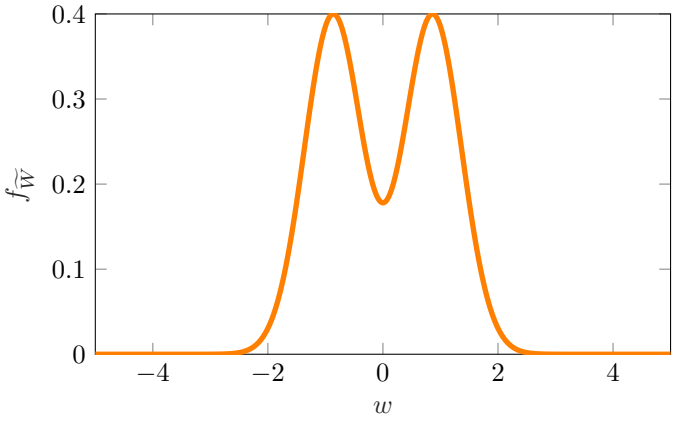


Figure 4. Probability density of $\widetilde{W}$ in the reference simulation.

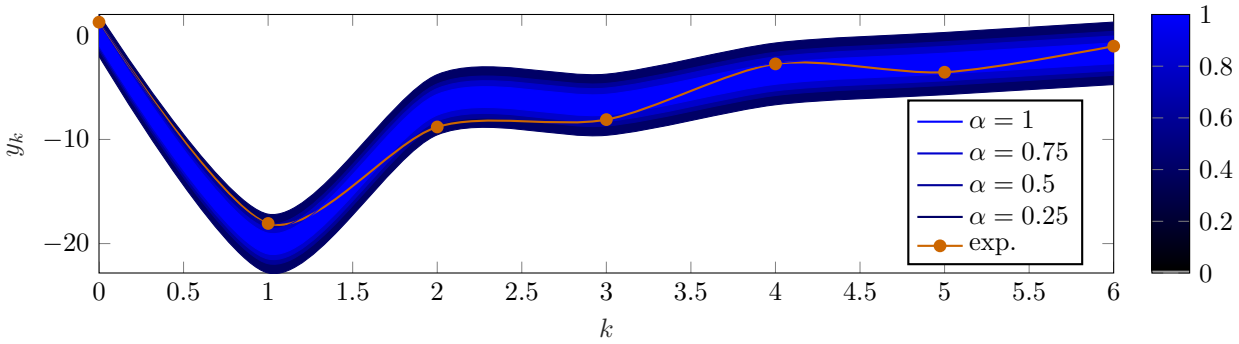


Figure 5. α-cuts of the predicted measurement and simulated reference.

the confidence sets containing the actual state samples stays well above the guaranteed level of confidence as illustrated in Figure 8, adding further verification to Proposition 4.

6. Conclusions

In this contribution, computationally exact solutions to a selection of forward and inverse problems in possibilistic LTI system analysis have been provided. The representation of convex fuzzy variables by the proposed quadruple description has been shown to be very well-suited for this purpose. It is conceptually easy to implement and allows for an intuitive design of a possibilistic batch filter.

The presented approach is evidently closely related to existing filtering formulations, such as moving horizon estimation (Rao et al., 2003). The novelty is that, while the presented approach relies on the propagation of convex sets, i.e. the α-cuts, this approach is able to also carry on statistical information in the form of the size of these sets for different values $\alpha \in [0, 1]$ without requiring a precise statistical modeling. Hence, it is possible to arrive at more expressive results than purely set-based techniques, but it does not force the practitioner to specify unique, yet

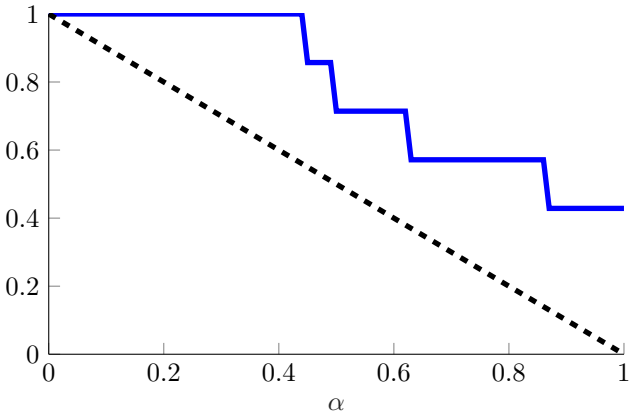


Figure 6. Empirical probabilities of the measurements y_k^{exp} being contained in the α-cuts $\mathcal{C}^{\alpha}_{\widetilde{Y}_k}$ compared to the guaranteed probability $1-\alpha$.

unwarranted, probability distributions thereon. It can, furthermore, serve as a reference solution for a more efficient (particle) filter, which has not yet been developed, with a recursive prediction and updating formulation similar to set-membership filtering approaches, refer e.g. to (Leong and Nair, 2016). However, this requires the derivation of a sensible procedure for the aggregation of confidence sets.

Other perspectives with respect to this contribution, which could prove to be of interest in the future, include the extension to to non-linear system models and the consideration of the inferred information in decision-making and the connection to results concerning robust optimization and control in a possibilistic context (Hose et al., 2019).

Acknowledgements

The authors gratefully acknowledge the support provided by the German Research Foundation (DFG) in the framework of the research project no. 319924547 (HA2798/9-1).

References

Augustin, T., Coolen, F. P., De Cooman, G., and Troffaes, M. C. (2014). *Introduction to Imprecise Probabilities*. John Wiley & Sons.

Baudrit, C. and Dubois, D. (2006). Practical representations of incomplete probabilistic knowledge. *Computational Statistics & Data Analysis*, 51(1):86–108.

Baudrit, C., Dubois, D., and Guyonnet, D. (2006). Joint propagation and exploitation of probabilistic and possibilistic information in risk assessment. *IEEE Transactions on Fuzzy Systems*, 14(5):593–608.

Couso, I., Montes, S., and Gil, P. (2001). The necessity of the strong α-cuts of a fuzzy set. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 9(02):249–262.

Destercke, S., Dubois, D., and Chojnacki, E. (2009). A consonant approximation of the product of independent consonant random sets. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 17(06):773–792.

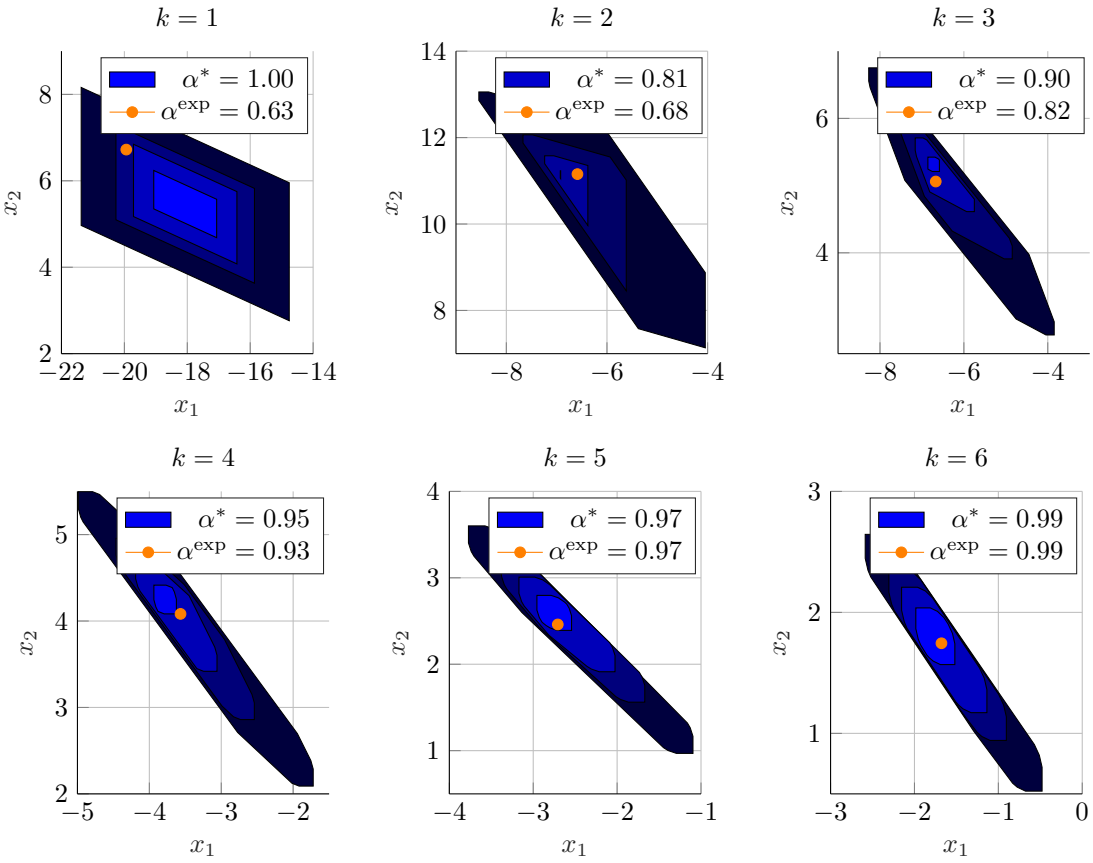


Figure 7. Confidence sets of the past system states (■) compared to the simulated reference (●). The coloring scheme represents the same α-levels as in Fig. 1.

Dubois, D., Foulloy, L., Mauris, G., and Prade, H. (2004). Probability-possibility transformations, triangular fuzzy sets, and probabilistic inequalities. *Reliable Computing*, 10(4):273–297.

Dubois, D. and Prade, H. (1988). *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Plenum Press.

Dubois, D. and Prade, H. (1992). When upper probabilities are possibility measures. *Fuzzy Sets and Systems*, 49(1):65–74.

Hanss, M. (2005). *Applied Fuzzy Arithmetic: An Introduction with Engineering Applications*. Springer.

Hose, D. and Hanss, M. (2019a). Consistent inverse probability and possibility propagation. In *Proceedings of the 2019 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology*. Atlantis Press.

Hose, D. and Hanss, M. (2019b). Possibilistic calculus as a conservative counterpart to probabilistic calculus. *Mechanical Systems and Signal Processing*, 133:106290.

Hose, D., Mäck, M., and Hanss, M. F. (2019). Robust optimization in possibility theory. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, 5(4).

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, 82(1):35–45.

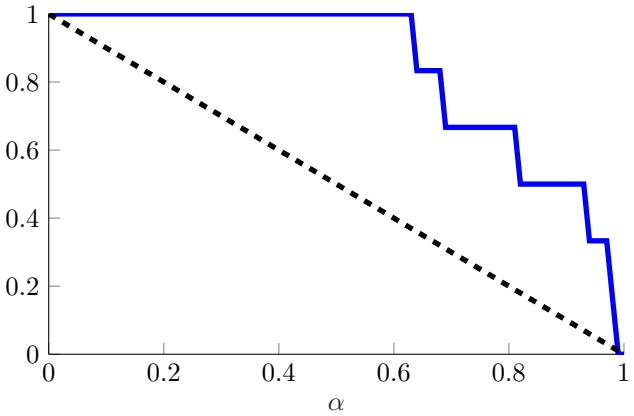


Figure 8. Relative frequency of the reference states $\mathbf{x}_k$ being contained in the α-cuts $\mathcal{C}^{\alpha}_{\widetilde{X}_k}$ compared to the guaranteed confidence level.

Langson, W., Chryssochoos, I., Raković, S., and Mayne, D. Q. (2004). Robust model predictive control using tubes. *Automatica*, 40(1):125–133.

Leong, P. H. and Nair, G. N. (2016). Set-membership filtering using random samples. In *2016 19th International Conference on Information Fusion (FUSION)*, pages 1087–1094. IEEE.

Rao, C. V., Rawlings, J. B., and Mayne, D. Q. (2003). Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations. *IEEE Transactions on Automatic Control*, 48(2):246–258.

Skogestad, S. and Postlethwaite, I. (2007). *Multivariable Feedback Control: Analysis and Design*, volume 2. Wiley New York.

Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning - I. *Information Sciences*, 8(3):199–249.

A Bimodal Distribution Function with Fuzzy Regression for Predicting Truck Load Population including Overloads

Bora Jang[1*] and Jamshid Mohammadi[2]
[1]*Sharma & Associates, Inc., USA*
[2]*Department of Civil, Architectural and Environmental Engineering, Illinois Institute of Technology, USA*
**Corresponding author: bjang5@hawk.iit.edu*

Abstract: This paper presents a statistical model for truckload frequencies including those from overloads. A Bimodal distribution function with statistical parameters obtained using fuzzy regression is presented for predicting truck loads including the contribution from overloads. Overload trucks often appear as a sizeable portion of truck populations on highways. In applications when damage estimation of transportation facilities such as pavements and bridges is desired, theoretical models providing a reasonable representation of truck load populations including overloads will be useful. Load populations mostly exhibit inconsistent patterns - often with two or more distinct peaks. This is because of a combination of loaded and empty trucks as well as overloads in the population. And as such, simple statistical distribution models fail to portray a realistic representation of truck load populations. In this paper, several truck load populations for 5-axle vehicles were acquired from the States of Illinois and Michigan in US and used to demonstrate the suitability of Bimodal models in representing the data. The goodness-of-fit tests using both the traditional Kolmogorov-Smirnov (K-S) and more modern Anderson-Darling (A-D) methods were used to demonstrate the suitability of a particular model to represent the data. The results show that a combination of beta and lognormal distribution can conveniently be used as a suitable model to represent the truck load populations that were analyzed in this study. Using the Bimodal model, theoretical distributions are developed to (1) represent the entire truck load population with WIM data; and (2) predict the population with limited data. If no truck load population is available, the Bimodal distribution model can still be used, with certain assumptions, based on fuzzy regression using traffic pattern information available for the roadway.

Keywords: Truck overloads, statistical models, mixed distribution function, Kolmogorov-Smirnov (K-S) test, Anderson-Darling (A-D) test, fuzzy regression

1. Introduction

1.1. Truckload Distributions and Overloads

Truck load data play an important role in condition assessment and life-cycle management of transportation systems such as pavements and bridges. Methods used in condition assessment of these systems require a realistic estimation of wear and tear and damage accumulation as a result of the repeated application of truck loads. For a given system, damage assessment can be done by using truck load data directly in a discrete format (Jang and Mohammadi, 2018). However, such a method will only be specific to the system for which the truck load data are available. Furthermore, since the data is used in a discrete format, the frequencies of load occurrences can only be described with specific ranges. Each range will have an upper and lower limit;

© 2021 *by Bora Jang and Jamshid Mohammadi.*

and any other load value within the limits must be approximated with either the upper or the lower limit. A more versatile and realistic approach in damage assessment can be achieved by using a theoretical model that best represents the truck load data. The advantage of using a theoretical truck load data model is that (1) it affords the formulation of the damage assessment independent of the type of structure and in a rather theoretical format; and (2) the model allows the continuity in the load data and eliminates the issues with the discrete data ranges as described.

For a given class of trucks, the intensity and frequency of load occurrences in the population vary to a great extent. The load population does not often show a consistent pattern; and as such, a single probability distribution function is generally unable to represent the entire population. Often the distribution exhibits two or more distinct peaks. This is attributed to the variety of loads in the population; and in most cases, the appearance of two peaks indicates a combination of loaded and empty trucks in the mix (Jang and Mohammadi, 2017). Load data also contain overload frequencies that may be significant in some cases resulting in multiple peaks in the population. The term overload, as used here, refers to truck weights in excess of the 356 kN (80 kips) limit.

Studies on the effect of truck load on highway systems have shown that the load population often contains occurrences of overloads that in certain cases may be frequent (Mohammadi and Shah, 1992; Snyder et al., 1985; and Timm et al., 2005). According to a recent report from the National Cooperative Highway Research Program, overload occurrences increase over time based on the trend in the Federal Highway Administration (FHWA) data (CPCS, 2016). Observing dual peaks appearing in the load data for 5-axle trucks, Mohammadi and Shah (1992) used a theoretical mixed distribution model to represent the occurrence of the dual peaks in truck load populations. Timm et al. (2005) used a distribution model as a mix of multiple probability functions to represent axle load spectra. They offered a combination of lognormal and normal distributions, which was supported by some statistical analyses.

Available related studies on truckloads in general and truck overloads in particular have primarily focused on the significance of load effects in highway bridges and presenting load models that can be used in bridge design and evaluations. Among related studies addressing truck live loads, those by Nowak (1994) and Nowak and Hong (1991) are reported here. Nowak (1994) used load data from truck survey for the purpose of developing live load models for bridges in Ontario, Canada. Although no specific truckload model is introduced, the live load models from load effects such as bending moments and shear forces are developed and recommended for use in bridge design. The study by Nowak and Hong (1991) is probability-based and provides a statistical representation of the bridge live load and related parameters (bridge span length and number of lanes). Again, no specific reference to a model that can represent load truck by axles and especially overloads is provided. Miao and Chan (2002) conduct a similar study using WIM data from Hong Kong and extreme bridge girder load effects (such as estimates of daily occurrences of extreme bending moments and shear forces) and developed statistical models for bridge live loads. Specific to this study is the introduction of statistical models for gross vehicle weight estimation using the inverse normal probability values and Type I extreme value distribution. They compared their load models with those in AASHTO and noted a relatively close agreement. Their study indicates that using such models, the gross vehicle weight for the bridge design life is 41.8 tons with a probability of 0.95 and 60 tons with a probability of 0.98. These values translate respectively to 409 kN (92 kips) and 587 kN (132 kips). Both these values correspond to overload situations in in US; and the load models presented indicate that there are 0.05 and 0.02 probability that the truck weight will actually exceed 409 kN (92 kips) and 587 kN (132 kips), respectively.

Fu and Hag-Elsafi (2000) used the New York State WIM data in developing load models for bridges including overloads. They present samples of gross vehicle data, which includes a significant presence of

overloads. Their model is primarily based on a lognormal distribution and focuses on providing equations applicable to load effects (such as bending moment) by treating the load source (i.e., gross vehicle weight) as a random variable. In presenting load distributions with overloads, they suggest combining probability values for normal load and overloads. Sample results presented for populations with significant overload presence clearly show the appearance of two peak values as expected and reported by others as well.

Tabatabai et al. (2017) use WIM data in Wisconsin for a period of one year to develop statistics for extreme truck loads for Class 9 category (which includes 5-axle single trailer trucks). For the data analyzed, the Class 9 constitutes about 62% of the entire data. The data indicates a 95% probability for a 467 kN (105 kips) with the maximum load recorded at 1,079 kN (242 kips). The study indicates that the load modeling for this class of trucks can be used to represent all other categories. Statistical analyses conducted considered the multi-modal effect of the data thus proposing to use mixed probability values. However, each probability function for a specific load range was determined around the peak value in that region. The model was developed by combining these different probability functions in which each function appears in the equation by a weight. The weights correspond to the number of sample data in different data ranges. The model indicates a significant amount of overload gross vehicle weight present in the load population. It was also compared with simulation data.

The main objective of this work is to use samples of truck load data and develop a mathematical model to represent a reasonable distribution for overload trucks. The use of such models is in applications where methods for transportation system damage assessment require information on the frequency and magnitude of loads, especially overloads. Specific transportation systems that can benefit from these models include bridges and pavements. In most part, real WIM data are used for damage assessment of these systems. However, in areas where WIM data is not available, models (such as those presented in this paper) will be helpful to estimate information on load frequency and magnitude.

1.2. STATISTICAL ANALYSIS

Specific to truck load data, available data is in discrete forms compiled primarily using a Weigh-in-Motion (WIM) system. As such data are compiled in only a limited number of stations across states, their use may become limited to the locations and states from which the data have been obtained. The use of mathematical models, that describe the truck load data through theoretical statistical distributions, may be an approach offering models that are more universally applicable to all locations. Since the discrete truck load distributions often show multiple peaks, conceivably, "mixed-type" distributions may be suitable models to represent them (Mohammadi and Shah, 1992). The overloads especially add more complexity to the shape of the load distribution, since they occur with less frequencies and constitute the area of the load population at the extreme upper tail of any theoretical distribution model.

A goodness-of-fit test is used to validate a theoretical distribution model when a particular probability distribution is specified to develop the model for a random set of data. Several studies have performed goodness-of-fit tests in addition to regression analyses to obtain suitable theoretical distributions to truckload data (Mohammadi and Shah, 1992; Timm et al., 2005). Such studies involve compiling actual data from the WIM systems and conduct statistical tests to determine the type of probability distribution models that would represent the load data. When mixed-type distribution models are involved, an extra parameter of the distribution model includes the common value that boundaries between any two functions.

Statistical test methods used in this paper are Kolmogorov-Smirnov (K-S) and Anderson-Darling (A-D) tests (Ang and Tang, 2007; Jang, 2018). The A-D test has advantages over the K-S test; and it may be suitable

for application to the WIM data, since the sample-size is very large. Since most overloads occur at the extreme end (tail) of the load population, the A-D test offers another advantage in the sense that it considers weights for these extreme values.

2. Weight-In-Motion Data used in Analysis

Nine different sets of WIM data are used in this study to find the characteristics of the truckload distribution and the truck overloads. The WIM data were collected from six different stations by the Illinois Department of Transportation (IDOT) and three different stations by Michigan Department of Transportation (MDOT) from June, 2014 to May, 2015 (Jang and Mohammadi, 2017). WIM data sets used in this study are calibrated by providers, IDOT and MDOT, to control any errors associated with the vehicle class, axle spacing and weight, speeds, etc.

A long enough period of time is needed to capture a realistic representation of truckload frequencies. In highway bridges, a two- or three-day period seems to be satisfactory (Hahin et al., 1993). Approximately, 4 weeks and a year of data represent 90% and 99% of the level of reliability, respectively. The data acquisition process may have to be conducted in several two- or three-day sessions at various times during the year to include any seasonal changes in the traffic pattern. Seasonal changes are usually caused by weather or other local factors. The effect of these changes is a periodic pattern in time that completes a cycle within a calendar year and is thus repeated on a yearly basis (Qu et al., 1997). The specific features of truck load population remain almost the same although the average yearly truck traffic changes from one highway to another (Mohammadi and Polepeddi, 2000). As such, a reasonable estimate of the population of overloads can be obtained based on the yearly truck traffic data available for the condition assessment a given system.

To better understand the distribution of truck weights in WIM data, it is necessary to analyze the data further by looking closely at the axle load and vehicle class data and specifically determining the presence of overloaded gross vehicle weight (GVW) in the data pool. The information on the frequency and distribution of the overloads in the data pool is especially important considering the fact that these loads, especially when occurred frequently, may accelerate the damage to pavements and bridges.

The load spectra from the WIM data are obtained in terms of the number of axles and the vehicle classifications in different data locations by developing a MATLAB program (MATLAB, 2016). Figure 1 shows the truck population of vehicle class 9 in the five-axle truck category for the combination of nine WIM data. In terms of the truck population according to the number of axles, the five-axle trucks constitute the bulk of the truck population except for the data from two of the nine stations.

Furthermore, it is noted that the truck population indicates that the vehicle class 9 is dominant in the five-axle category in the population except for the data from two of the nine stations. In addition to data from the five-axle and vehicle class 9 trucks; these two data sets appear to also include data from two-axle and the vehicle class 2 and 5 trucks (which does not include overloads). It is indicated that a significant portion of the overload data is in fact in the load data from the vehicle class 9 and the five-axle truck population. Furthermore, these types of trucks make up of a relatively large percentage of the overall truck populations on highways (as portrayed in WIM station data). Therefore, the WIM data for five-axle truck and vehicle class 9 were considered in this study, as the representative of a typical truck load data.

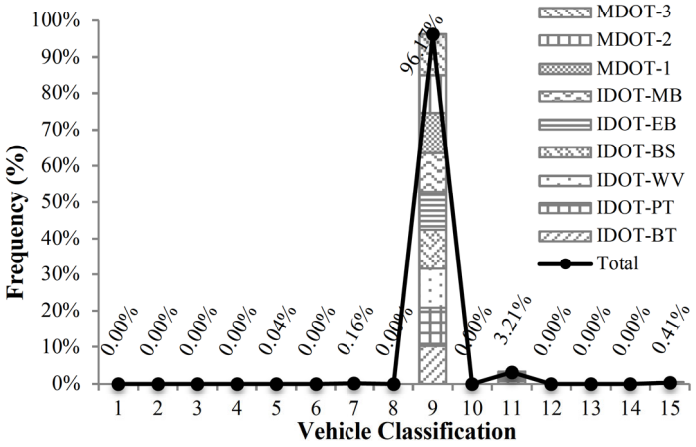


Figure 1. Truck population versus the vehicle class 9 in the five-axle truck category

3. A Bimodal Distribution Model for Overloads

The truckload distribution for the gross vehicle weight that exhibits two distinctive peaks is referred to a Bimodal distribution. A Bimodal distribution indicates a need for two different continuous probability distributions (f_1 and f_2) in which the overall truckload data is divided into two parts with a common load, S_L, representing the boundary between the two functions. As a simple method, the common load can be selected subjectively by visual examination of the data through trial and error. Alternatively, one may treat the common load as an extra parameter in the overall Bimodal distribution model, which will need to be estimated using statistical methods such as the method of maximum likelihood.

In developing the Bimodal distribution model, it was decided to use beta and lognormal distributions because of (1) the fact that they only accept positive values; and (2) others have also recommended them (e.g., Mohammadi and Shah 1992). Several combinations of beta and lognormal distributions were tried for functions, f_1 and f_2. The combinations considered included beta-beta, lognormal-lognormal, and beta-lognormal distributions (Jang, 2018). The beta distribution is among a few distributions that are appropriate for bounded random variables. The beta distribution, for a zero-lower bound, is (see for example, Ang and Tang, 2007)

$$f(s) = \frac{1}{B(q,r)} \frac{s^{q-1}(s_0 - s)^{r-1}}{{s_0}^{q+r-1}} \tag{1}$$

In which q and r are the shape parameters of the beta distribution, s_0 is the upper limit, and $B(q,r)$ is the beta function. Parameters, q and r, are positive and determine the shape of distribution. The beta function is written in the following form.

$$B(q,r) = \int_0^1 x^{q-1}(1-x)^{r-1}dx \tag{2}$$

The lognormal distribution may especially be suitable to represent overloads in the truckload frequency distribution with a one-sided tail effect. The function is written as follows (see for example, Ang and Tang, 2007):

$$f(s) = \frac{1}{\sqrt{2\pi}(\zeta s)} exp\left[-\frac{1}{2}\left(\frac{ln\, s - \lambda}{\zeta}\right)^2\right] \quad (3)$$

In which λ and ζ are the location and shape parameters of the lognormal distribution, respectively.

Figure 2 describes the concept of mixed density function with consideration for the common load.

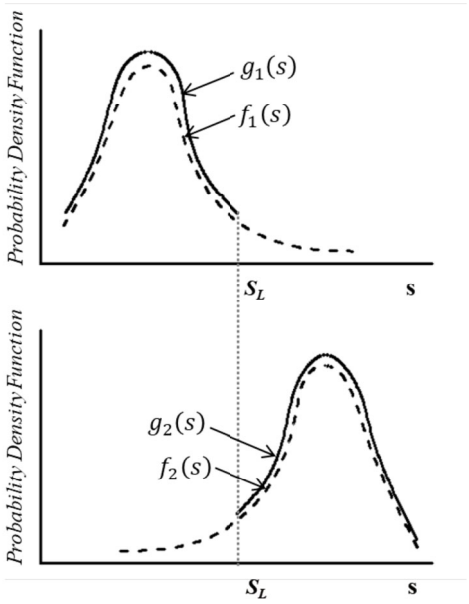


Figure 2. Schematic of mixed density function with consideration for the common load (S_L)

The probability density function for each part in the population will need to be adjusted since f_1 and f_2 are valid for $s < S_L$ and $s > S_L$, respectively (see Mohammadi and Shah, 1992). The adjusted density functions, g_1 and g_2, are therefore obtained from the proportion of each probability in the following forms:

$$g_1(s) = \frac{f_1(s)}{F_1(S_L)} \quad (4)$$

and

$$g_2(s) = \frac{f_2(s)}{1 - F_2(S_L)} \quad (5)$$

In which F_1 and F_2 are the probability distribution functions of f_1 and f_2. Using these adjusted probability density functions, the mixed density function is then written as

$$g(s) = \begin{cases} \alpha_1 g_1(s), & for\ s \le S_L \\ \alpha_2 g_2(s), & for\ s > S_L \end{cases} \quad (6)$$

where parameters α_1 and α_2 represent the proportions of entire load population that are below and above the common load, respectively. Note that $\alpha_1 + \alpha_2 = 1$.

4. Testing Goodness-of-fit of Load Distribution Models

As indicated earlier, to test the validity of a mix probability function to represent the data, statistical methods such as the Kolmogorov-Smirnov (K-S) test and the Anderson-Darling (A-D) test are useful (Ang and Tang, 2007; Anderson and Darling, 1954; Stephens, 1974). The WIM data were analyzed and tested against several types of statistical models when combined together. For each type, the aforementioned K-S and A-D tests were conducted in an effort to identify a Bimodal distribution model to represent the truck load data including the significance of the overloads that appear in the extreme right-hand tail of the model. The analysis shows that a combination of beta and lognormal distributions is especially useful in representing the data.

4.1. KOLMOGOROV-SMIRNOV (K-S) TEST

The details of the K-S test for similar engineering applications can be found for example in Khisty et al. (2012). The test is considered to be suitable for comparing the observed cumulative frequency with the cumulative distribution function of an assumed theoretical distribution. The maximum difference between the observed cumulative frequency function, $S_n(x)$, and the theoretical cumulative distribution function, $F_X(x)$, is the measure of discrepancy between the two. The maximum difference, D_n, is determined by

$$D_n = \max_x |F_X(x) - S_n(x)| \tag{7}$$

For a given significance level, the test passes if the value of the maximum difference is smaller than a critical value. We can consequently believe that the theoretical distribution model is appropriate to represent the observed data.

4.2. ANDERSON-DARLING (A-D) TEST

Since the application of this test is less known in engineering, the details of the test are presented herein. The test is sensitive to discrepancies between the observed and theoretical frequencies; and as such, it places more weight at the tails of the distribution rather than near the median. This is certainly an advantage over the more traditional K-S test. The A-D statistic can be calculated by the following equation (see for example, Ang and Tang, 2007):

$$AD^2 = -\sum_{i=1}^{n}\left[\frac{(2i-1)\{lnF_X(x_i) + ln[1 - F_X(x_{n+1-i})]\}}{n}\right] - n \tag{8}$$

In which, AD^2 is the Anderson-Darling statistic, $F_X(x)$ is the theoretical cumulative distribution function, and n is the sample size. A small sample size is not valid for the A-D test and hence the sample size should be larger than 7.

The A-D statistic would receive more contributions from the tails of distribution, since it is expressed in terms of the logarithm of the probabilities. However, the critical values for the A-D test are dependent on the specific distributions being tested. Tables of critical values are currently not available exclusively for the beta distribution and for any combinations of multiple distribution functions. Nevertheless, the values of AD^2 are valuable for use as a comparable measure when various combinations of different distribution models are considered. For example, if AD^2 is relatively large, the reason may be attributed to the application of Bimodal distribution in representing the data with the appearance of a cluster of data triggering the larger value for AD^2.

5. Results and Discussions of Statistical Assessment

Each of the nine load spectra of the WIM data is modeled using several possible combinations of beta and lognormal distributions. The statistical tests require rearranging data into discrete ranges of the loads. After several trials, load ranges in 22.25 kN (5 kips) intervals (or less) are found to be suitable for the analysis. For the common load values lower than 200 kN (45 kips) and greater than 311 kN (70 kips), the Bimodal distribution model could not be established with distinct multiple peaks (which are the characteristics of the truck load populations). Thus, a suitable common load value is anywhere between 200 to 311 kN (45 to 70 kips) based on the frequencies of various load values in the entire truck load population.

Based on the aforementioned conditions, the proposed mixed probability distribution models were developed for various common loads corresponding to three kinds of the combination of beta and lognormal distributions as described earlier (i.e., beta-beta, lognormal-lognormal, and beta-lognormal). The goodness-of-fit tests are conducted to verify the validity of the proposed mixed probability distribution models for each case by the K-S and A-D methods. It is noted that the K-S and A-D tests are conducted to also determine the appropriate upper bound of beta distribution for the density function f_2 (see Eqs. 4-6) in the model made up of a combination of two beta functions. Common load values between 200 to 311 kN (45 to 70 kips) are considered in performing the K-S and A-D tests. Figure 3 illustrates the results of K-S and A-D tests for different values of the upper bound of beta distribution and different WIM stations. The minimum values of the test statistic (i.e., the maximum difference between theoretical and observed values in the K-S and AD^2 in the A-D tests) in terms of different values for the common loads are computed and plotted as shown in Figure 3. A theoretical distribution model, in which the upper bound is 556 kN (125 kips), results in the lower values of both tests for all WIM stations. Thus, the results using this upper bound value will be used to compare the three possibilities for the combination of beta and lognormal distributions, as described earlier.

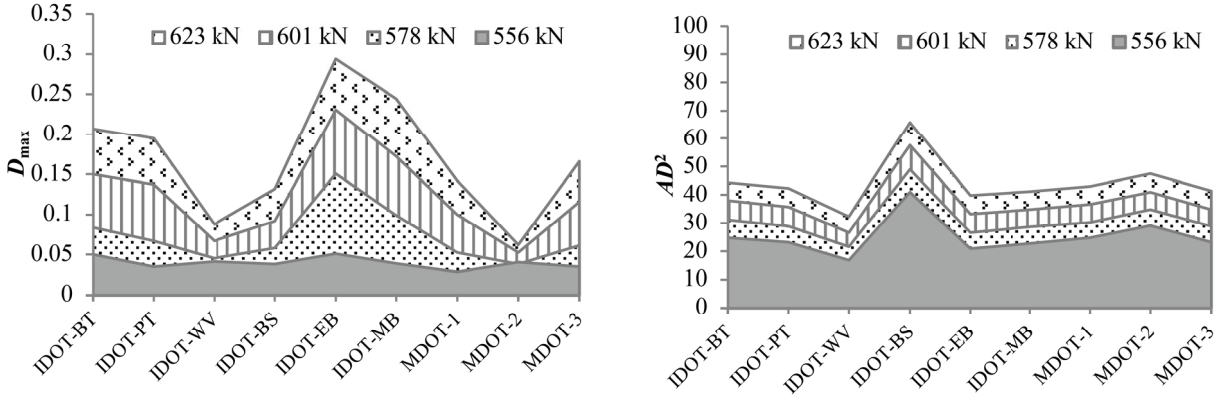


Figure 3. Test statistics (D_{max} in K-S and AD^2 in A-D tests) in terms of various values of the upper bound of beta distribution for nine WIM stations

Figure 4 represents the results of K-S and A-D tests for different values of the common load and combinations of the Bimodal distribution models for all WIM stations. As seen in the results, no test statistics value for the combination of two beta distributions is listed near the common load over 245 or 267 kN (55 or

60 kips). This is because the beta distribution is found to be unsuitable to present the distribution for load values larger than the common load.

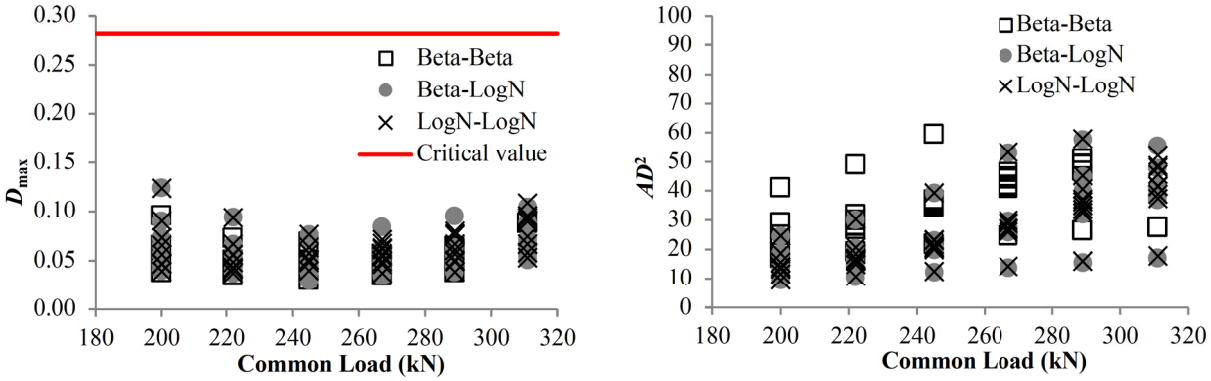


Figure 4. Results of K-S and A-D tests with various common loads and combinations of Bimodal distribution for the nine WIM stations

For the K-S test, all theoretical distributions for nine WIM stations are determined to be acceptable to model the load data at 5% significance level (critical value = 0.282). Considering various common load values, a combination of the two beta and beta-lognormal distributions with the common load at 245 kN (55 kips) is most suitable to represent the entire truck load population per results from the K-S test.

Since the critical value in the A-D test is not known, the method is used as a way of comparatively deciding the suitability of one model over the others. The A-D statistics mainly decrease as the common load decrease for most cases. The combination of two beta distributions results in greater values for the test statistics (AD^2 values) than the other kinds of Bimodal distribution for all common load cases. Furthermore, Bimodal distributions containing at least one lognormal distribution represent very similar values in the A-D statistics for all common load cases. This is attributed to the fact that the lognormal distribution tends to dominate the load population at load values larger than the common load. The overall results show that a combination of the beta and lognormal distribution yields a favorable model with the smallest A-D statistics. It is reasonable to believe that the lognormal distribution is more suitable in representing the data in the portion of the population where the second peak appears, which includes the overloads in the tail of the distribution model. The results indicate that the combination of beta-lognormal distribution model with the common load of 200 kN (45 kips) is perhaps well suited to represent the data.

The K-S test generally draws conclusions based on the middle portion of the data. And as such, the data at extremes (at tails) is not very well represented. On the other hand, the A-D test draws conclusions more on the extreme loads. Therefore, if only the overload distribution is of interest, the A-D test is perhaps more appropriate. However, in general because the entire population is needed for damage estimations, it makes sense to rely on both tests (i.e., K-S and A-D tests) in any given problem. For this reason, we suggest a range for the common load be specified (in this case, 200 to 245 kN (45 to 55 kips)), so that the significance of both the middle and the tail data effects is incorporated in damage estimation of an affected transportation system (pavement or bridge). Depending on the focus of analysis, however, one of three values (200, 222, or 245 kN (45, 50, or 55 kips)) may be used in practice. For analysis in this study, as a result of the two tests, and at least within the limited data analyzed for the Illinois and Michigan truckloads, it is reasonable to conclude that a

combination of the beta and lognormal distributions with the common load of 200 kN (45 kips) is a suitable theoretical distribution for truckload data (including the overloads).

It is noted that for most situations, and especially for 5-axle in vehicle class 9, the range of common loads specified would be the representative of the data in a Bimodal distribution model. In fact in situations where the overloads make up a significant portion of the overall load population, and when overloads are at unusually large values, the common load may become larger than the range specified in this study. However, such conditions are very rare, especially considering the fact that very large overloads happen under special conditions and very infrequently.

6. Proposed Distribution Model when WIM Data is available

In situations when the truck load population is available, the theoretical distribution proposed in this study may be used to describe the entire truck load population. Figure 5 shows a flowchart on how the load population may be quantified for the case if WIM data is available.

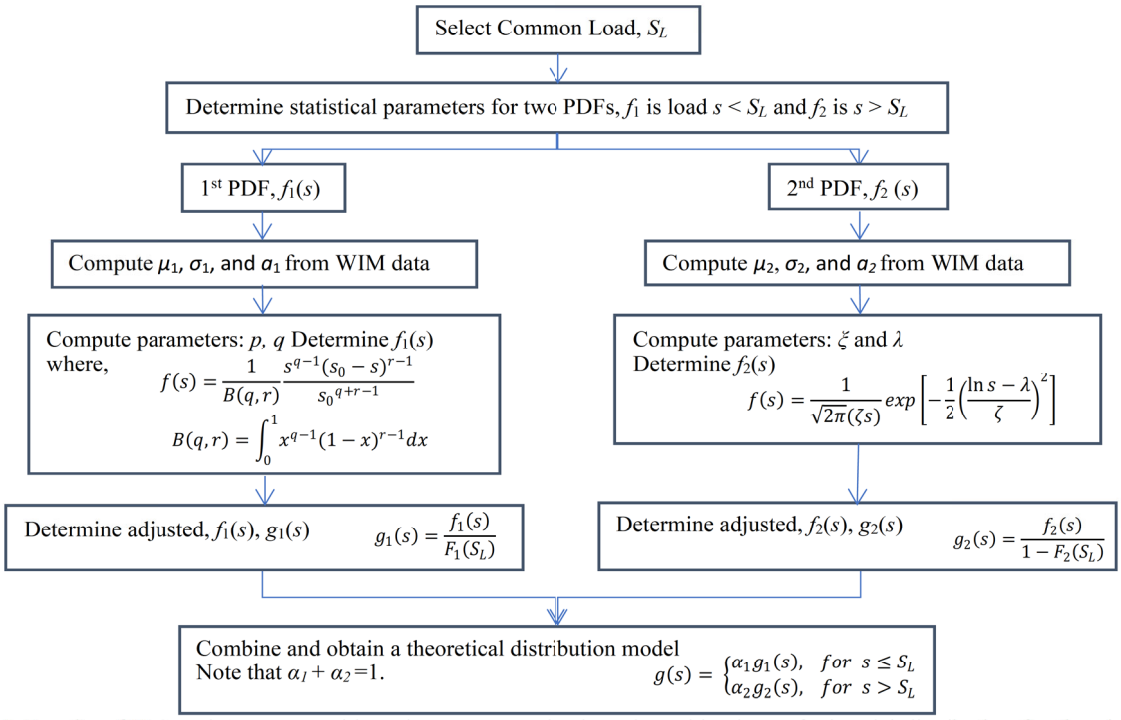
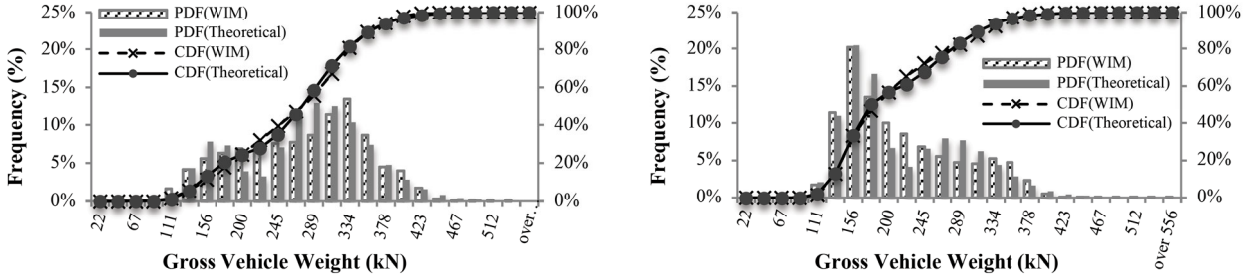


Figure 5. Results of K-S and A-D tests with various common loads and combinations of Bimodal distribution for the nine WIM stations

The parameters driving the Bimodal distribution modes are then obtained using the available WIM data. A common load, S_L, needs to be selected among possible values (i.e. 200, 222, or 245 kN (45, 50, or 55 kips)) for the truck load population. Then the mean and standard deviation (i.e. μ_1, σ_1, μ_2, and σ_2) for the two parts of the load population (i.e., the portion representing values of load lower than S_L and those representing values

of load larger than S_L) are computed. Furthermore, the percentage of the load frequencies in each portion, i.e. α_1 and α_2, are computed, in which α_1 is the ratio of the sample points in $s < S_L$ divided by the total sample points; while α_2 is the ratio of the sample points in $s > S_L$ divided by the total sample points. The theoretical distribution can then be determined from these parameters.

As an example, the frequency and cumulative probability distribution for the observed WIM data and the theoretical model using the Bimodal distributions for four WIM stations are illustrated for the common load of 200 kN (45 kips) in Figure 6.



(a) IDOT-WV (b) MDOT-2

Figure 6. The frequency and cumulative probability distribution for the observed WIM data and the theoretical model for four WIM stations

7. Proposed Distribution Model when WIM Data is not Available

If no truck load population is available, one may still use the theoretical models proposed in this study with certain assumptions and based on the roadway traffic and other information. The average daily truck traffic (ADTT) and the number of overload permits that an agency issues for bridges along the roadway may be used as basic information in developing a model for the load distribution. As a first step, using the available information, the percentage of overloads in the entire truck load population can be determined as:

$$OL^* = \frac{the\ number\ of\ permits\ per\ year}{ADTT \times 365} \tag{9}$$

In which, OL^* is the minimum percentage of overloads in the entire truck load population and $ADTT$ is the annual daily truck traffic. In the absence of WIM data, statistical parameters such as μ_1^*, σ_1^*, μ_2^*, and σ_2^* have to be estimated based on the available information using regression analysis – either through numerical or logical schemes.

7.1. FUZZY REGRESSION ANALYSIS

In general, a regression analysis, especially, a linear regression model, is commonly used to analyze relationship and correlation between dependent and independent variables and to predict the trend in the data set (Ang and Tang, 2007). Usually, a large number of collected data is required. If the data set is of small size, the relation between variables is imprecise. In such cases, rather than the statistical linear regression model, logical regression analysis may be used as described in this paper. Statistical parameters (i.e., μ_1^*, σ_1^*,

μ_2^*, σ_2^*, etc.) depend on volume of traffic, which can be used as the basic information in deriving these parameters.

A fuzzy regression model was developed by Tanaka et al. (1982), based on the fuzzy logic to treat the imprecise and uncertain phenomenon in a data set. The main approach in this model is to minimize the data fuzziness (and thus uncertainty) using a linear programming method. Over decades, many researchers have conducted studies to improve the outcome of the fuzzy regression model (Tanaka et al., 1982; Lee and Tanaka, 1999; Wang and Tsaur, 2000; Chen and You, 2014). In this paper, the classical regression analysis with crisp-input (i.e., the percentage of overload, *OL*) and fuzzy-output (i.e., statistical parameters required in the Bimodal distribution for predicting the truck load population) is presented in evaluating the relationship between variables. The model includes a confidence interval in treating the fuzzy nature of the relationship between variables.

The general form of fuzzy linear output by Tanaka et al. (1982) is as follows:

$$\widetilde{Y}_i = \widetilde{A_0} + \sum_{j=1}^{m} \widetilde{A}_j x_{ij} \tag{10}$$

In which, $\widetilde{Y}_i$ is the fuzzy linear output with $i = 1, 2, \ldots, n$, $\widetilde{A_0} = (a_0, c_0)$, $\widetilde{A}_j = (a_j, c_j)$ are fuzzy coefficient with $j = 1, 2, \ldots, m$, $x_{ij} = [x_{1j}, x_{2j}, \ldots, x_{nj}]$ and is a vector of input variables with respect to j^{th} fuzzy coefficient. In the fuzzy coefficient (a_j, c_j), a_j is a mode and c_j is the spread in the membership function (MF). A symmetric triangular model is considered in this study; and the expression for MF is written as follows:

$$MF(a_j) = \begin{cases} max\left(1 - \dfrac{|a_j - c_j|}{c_j}\right), & a_j - c_j \le a_j \le a_j + c_j \\ 0, & otherwise \end{cases} \tag{11}$$

The shape of membership function can be selected based on the distribution of data and can be in an asymmetric triangular or trapezoidal form.

To obtain the result of fuzzy linear regression analysis with crisp-input and fuzzy-output, the total degree of fuzziness (spread) is minimized as follows:

Minimize

$$\sum_{i}^{n}\left(c_0 + \sum_{j=1}^{m} c_j|x_{ij}|\right)$$

Minimize

$$\begin{aligned} &a_0 + \sum_{j=1}^{m} a_j x_{ij} + (1-h)\left[c_0 + \sum_{j=1}^{m} c_j|x_{ij}|\right] \ge y_i + (1-h)e_i \\ &a_0 + \sum_{j=1}^{m} a_j x_{ij} - (1-h)\left[c_0 + \sum_{j=1}^{m} c_j|x_{ij}|\right] \le y_i - (1-h)e_i \\ &c_j \ge 0, \quad 0 \le h \le 1, \quad i = 1,2,\ldots,n, \quad j = 1,2,\ldots,m \end{aligned} \tag{12}$$

In which, in addition to the aforementioned variables, y_i is the mode and e_i is the spread in a symmetric triangular membership function for the fuzzy output, respectively, and h is a factor to control the level of support (spread) in the membership function. The support of the fuzzy regression model prediction shows the range of possible values. All data points should be included in the interval created by the h index. Increasing the value of h the possible values of output which is not presented in the data set will be considered in the

interval. The value of h can be determined through an expert knowledge.

After estimating the fuzzy output number based on the fuzzy regression, the crisp value of output can be determined by the defuzzification procedure. Figure 7 presents a popular method, which involves the calculation of the centroid for the area under the membership function. It is suggested to select the defuzzification method based on the form of membership function used in the analysis.

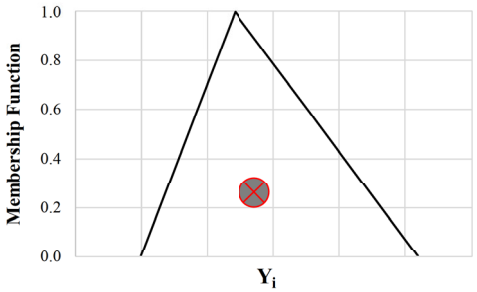


Figure 7. The fuzzy membership function (MF) and defuzzification: triangle shape and centroid calculation

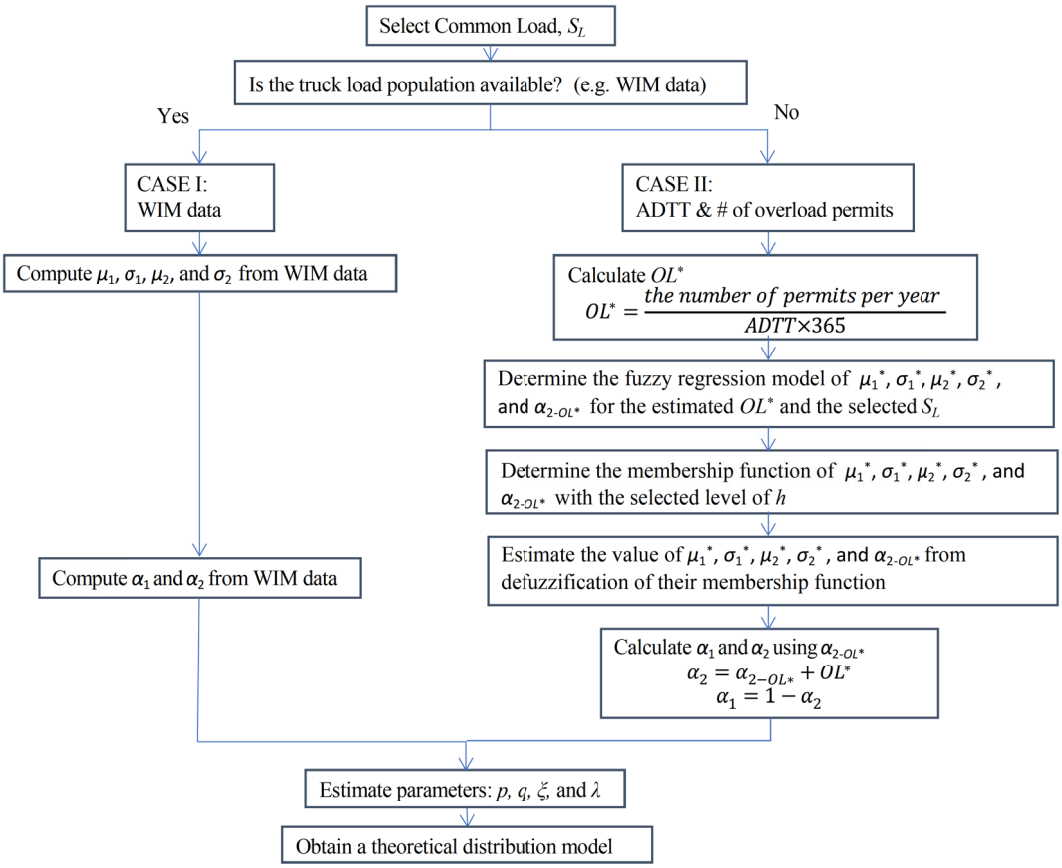


Figure 8. Flowchart for obtaining a suitable statistical truck load distribution model

Statistical parameters determined from the fuzzy regression model are then used to obtain the Bimodal distribution for the load population in the absence of any WIM data. Figure 8 shows a flowchart on how the load population may be estimated for the two cases of (1) WIM data is available; and (2) no WIM data is available.

7.2. Numerical Example and Discussions of Fuzzy Regression Model

Upon selecting the common load, i.e. 200, 222, or 245 kN (45, 50, or 55 kips), for the truck load population, approximate values for the mean and standard deviation of the two portions of the load population (i.e. μ_1^*, σ_1^*, μ_2^*, and σ_2^*) and the parameters α_1 and α_2 can be estimated. It is noted that we must first compute an auxiliary parameter (i.e. $\alpha_{2\text{-}OL}^*$) from OL^* in Eq. 9 and then computing α_1 and α_2 as shown in Eq. 13. The estimates for α_1 and α_2 are:

$$\alpha_2 = \alpha_{2-OL}(OL^*) + OL^* \text{ and then } \alpha_1 = 1 - \alpha_2 \tag{13}$$

As an example, considering a common load, S_L = 200 kN (45 kips), the membership function of fuzzy output from the regression analysis is developed and presented in Figs. 8 and 9.

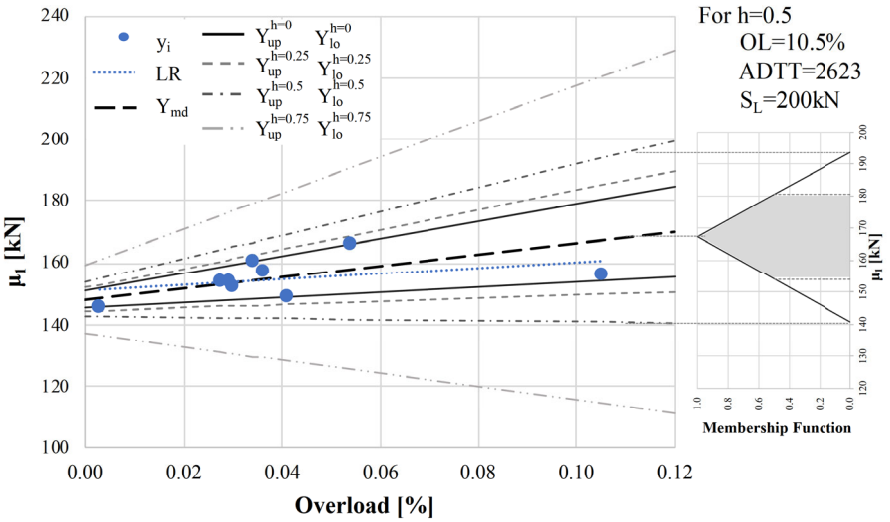


Figure 8. The fuzzy linear regression model for μ_1^* for S_L = 200 kN (45 kips) and a triangular MF

Using the fuzzy regression model given h, the values of the statistical parameters for μ_1^*, σ_1^*, μ_2^*, σ_2^*, and $\alpha_{2\text{-}OL}^*$ can be estimated. In Figure 8, the shaded area represents the observed fuzzy output without the consideration for h. The additional area (not shaded) expands the confidence interval and increase the probability to capture the data beyond the available set. It is further noted that $\alpha_{2\text{-}OL}^*$ is also bounded by 0 and 1 limits, i.e., $0 \leq \alpha_{2-OL}^* \leq 1$.

The fuzzy linear regression model results provided in Figures 8 and 9 are developed based on the data set collected. These results are reasonable predicted values of the parameters according to the fuzzy regression model with limited available data. When additional data becomes available, the shape of boundary and membership function may be calibrated for more robust estimates for the parameters of the truck load distribution.

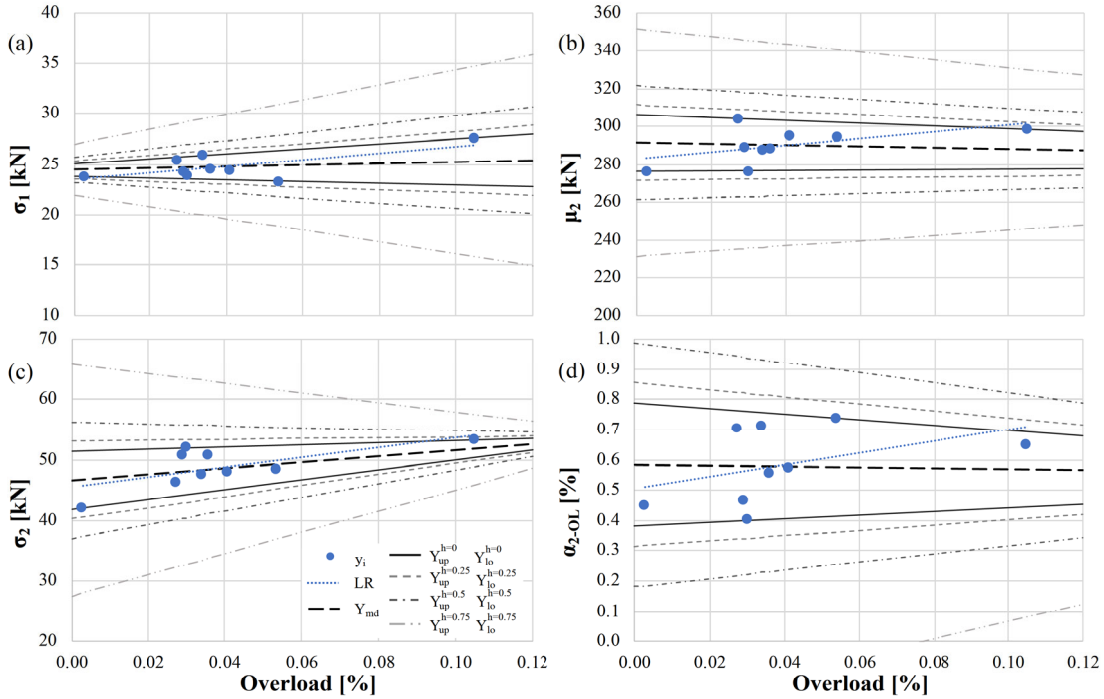


Figure 9. The fuzzy linear regression model for σ_1^*, μ_2^*, σ_2^*, and $\alpha_{2\text{-}OL}^*$ for S_L = 200 kN (45 kips)

8. Comparing Results from the Model with those from other Studies

Using the model presented in this paper, the gross vehicle weights were estimated for 95% and 98% probability values. Studies by Tabatabai et al. (2017) and Miao and Chan (2002) provide load levels corresponding to 95% and 98% values also. Fu and Hag-Elsafi (2000) do not specifically provide load estimates for these probability values, instead they present, in graphics form, the frequencies at different load levels including overloads. The graph can then be used to also read off the values corresponding to 95 and 98% probabilities. Table I provides a summary of the values from these references and those predicted using the models in this paper. Although the estimates from the current study show lower than those indicated by others, it is emphasized that there are marked differences in the data used in the three references listed in Table I.

The WIM data used in the three references listed in Table I all had a significantly higher GVW data as opposed to the data used in the current study. Furthermore, heavier truck weights were also present in their data. The overload content in the WIM data used in our study was at about 10% of the total loads with the maximum GVW at lower levels than those reported in the three references. Nevertheless, by knowing an estimate for the percentage of the overload and the largest load in the population, the model presented in this study provides a useful tool to predict load populations that can be used for damage assessment in payments and bridges, in the absence of having specific WIM data.

Table I. Gross vehicle weight estimates for 95 and 98% probability level in kN (kips)

Probability Level (%)	Model Values[1]	Tabatabai, Titi and Zhao (2017)[2]	Miao and Chan (2001)	Fu and Hag-Elsafi (2000)[3]
95	378.3 (85.0)	471.0 (104.9)	413.0 (92.0)	422.8 (95.0)
98	400.5 (90.0)	669.7 (150.5)	587.0 (132.0)	600.1 (135.0)

[1] Maximum values estimated are provided.
[2] Loads effects provided in this reference were used to arrive at GVW.
[3] Values were obtained from GVW graphs presented in this reference.

9. Limitations of Study

The model proposed in this study was intended for demonstrating how mathematical Bimodal distribution models can be used to represent truckload data. Available WIM data from two sources (in the US States of Illinois and Michigan) were used for this purpose. Having some type of mathematical model is useful in applications where truckload (including overloads) information is needed. The model presented in this paper can also be used in situations where WIM data is not readily available, as long as some preliminary information (for example the percentage of overloads in the truck load population) is known. There are several limitations to this study as explained below.

1. The model developed is exclusive to the sources from which WIM data was acquired. Since there are variations expected across truckload data from different states, the model is not intended to offer a universal method for estimating truckloads and overloads. WIM data from more states would be needed for development of a universal model, which also requires introducing additional parameters in it.

2. The study utilized the WIM data as received from the sources. State agencies in compiling WIM data use their own methods of quality assurance and control in maintaining the accuracy of their data. Thus, this study did not include any additional verification and/or validation on the original WIM data received.

3. A key parameter in the proposed Bimodal distribution model is the “common load.” The suggested values for this parameter are based on the WIM data analyzed; and as such, they represent the 5-axle vehicle class 9 category. Although this class of vehicles is reported to make up a major percentage of the total population of truckloads, in rare cases when frequent occurrences of very heavy trucks are involved in the population, the common load will need to be studied further and revised.

4. In the fuzzy regression analysis, various type of membership functions may be employed. It is noted that more complex membership function forms, especially asymmetric ones, require an additional set of data for their validation. A simple triangular linear function is approximate, yet provides a convenient means for using the logical regression analysis method explained in this paper.

10. Conclusions

The following presents the main conclusions from this study.

1. When the entire population of truck load in a given WIM station is analyzed, the distribution often appears with two distinct peaks indicating that a Bimodal distribution function is perhaps needed to represent the data.

2. Based on the WIM data compiled, and statistical tests, it is shown that a Bimodal distribution comprised of a beta distribution and a lognormal distribution would provide a reasonable model for the entire truck load data, including the overloads.

3. The study further enhances the statistical load analysis process by introducing using Anderson-Darling test in addition to the more traditional Kolmogorov-Smirnov test.

4. Specific to the Illinois and Michigan truckload data, a combination of the beta and lognormal distribution function with a common load value at 245 kN (55 kips) appears to suitably represent the WIM truck load data including the overloads.

5. When no WIM data is available and the truck load distribution need to be developed, this study proposes the model development based on some preliminary roadway information (such as ADTT and the number of annual overload permit issued).

6. Using the percentage of overloads in the truck load population estimated and the fuzzy regression model for the required statistical parameters, the proposed Bimodal distribution can be developed for cases when no or limited WIM data are available.

Acknowledgements

The writers wish to acknowledge the Illinois and Michigan Department of Transportation for their cooperation and support of the research in providing the WIM data.

References

Anderson, T. W. and Darling, D. A. A Test of Goodness of Fit. *Journal of the American Statistical Association*, 49(268):765-769, 1954.

Ang, A. H. and Tang, W. H. *Probability Concepts in Engineering – Emphasis on Applications to Civil and Environmental Engineering*. John Wiley & Sons, Inc., 2007.

Chen, S. P. and You, S. S. A Fuzzy Linear Regression Model Based on Dissemblance Index. *International Journal of Mathematical and Computational Sciences*, 8(9):1278-1283, 2014.

CPCS, Perkins Motor Transport, Inc., and Portscape, Inc. Multi-State, Multimodal, Oversize/Overweight Transportation. NCHRP Report 830, Transportation Research Board, Washington, D.C., 2016.

Fu, G. and Hag-Elsafi, O. Vehicular overloads: Load model, bridge safety, and permit checking. *Journal of Bridge Engineering*, ASCE, 5(1):49-57, 2000.

Hahin. C., South. J. M., Mohammadi. J., and Polepeddi, R. Accurate and rapid determination of fatigue damage in steel bridges. *Journal of Structural Engineering*, ASCE, 119(1):150-168, 1993.

Jang, B. *Bridge Rating with Considerations for Fatigue Damage Effect from Truck Overloads*. PhD Dissertation, Illinois Institute of Technology, Chicago, IL, 2018. Civil Engineering PhD.

Jang, B. and Mohammadi, J. Bridge rating modification to incorporate fatigue damage from truck overloads. *Bridge Structures*, 13(2-3):01-107, 2017.

Jang, B. and Mohammadi, J. Significance of fatigue damage from overload trucks on bridge load rating. *MATEC Web of Conferences*, 165, 10005, 2018.

Khisty, C. J., Mohammadi, J., and Amedkudzi, A. *Systems Engineering with economics probability and statistics*. Ross Publishing company, Fort Lauderdale, FL., 2012.

Lee, H. and Tanaka, H. Fuzzy Approximations with Non-Symmetric Fuzzy Parameters in Fuzzy Regression Analysis. *Journal of the Operations Research*, Society of Japan, 42(1):98-112, 1999.

MATLAB and Statistics Toolbox Release 2016b. The MathWorks, Inc., Natick, Massachusetts, United States, 2016.

Miao, T. J. and Chan, T. H. Bridge live load models from WIM data. *Engineering Structures*, 24(8):1071-1084, 2002.

Mohammadi, J. and Shah, N. Statistical Evaluation of Truck overloads. *Journal of Transportation Engineering*, ASCE. 118(5):651-665, 1992.

Mohammadi, J. and Polepeddi, R. Bridge rating with consideration for fatigue damage from overloads. *Journal of Bridge Engineering*, ASCE, 5(3):259-265, 2000.

Nowak, A. S. and Hong, Y. K. Bridge live-load models. *Journal of Structural Engineerin*g, ASCE, 117(9):2757-2767, 1991.

Nowak, A. S. Load model for bridge design code. *Canadian Journal of Civil Engineering*, 21(1):36-49, 1994.

Qu, T., Lee, C. E., and Huang, F. Traffic-load Forecasting using Weigh-In-Motion Data. TX-99/987-6, Texas Department of Transportation, Austin, TX, 1997.

Snyder, R. E., Likins, G. E., and Moses, F. Loading spectrum experienced by bridge structures in the United States. Report RD-851012, Federal Highway Administration, Washington, D.C., 1985.

Stephens, M. A. EDF Statistics for Goodness of Fit and Some Comparisons. *Journal of the American Statistical Association*, 69(347):730-737, 1974.

Tabatabai, H., Titi, H., and Zhao, J. WIM-based assessment of load effects on bridges due to various classes of heavy trucks. *Engineering Structures*, 140:189-198, 2017.

Tanaka, H., Uejima, S., and Asai, K. Linear Regression Analysis with Fuzzy Model. *IEEE Transactions on Systems, Man, And Cybernetics*, Smc-12(6):903-907, 1982.

Timm, D. H., Tisdale, S. M., and Turochy, R. E. Axle load spectra characterization by mixed distribution modeling. *Journal of Transportation Engineering*, 132(2):83-88, 2005.

Wang, H. F. and Tsaur, R. C. Bicriteria Variable Selection in a Fuzzy Regression Equation. *Computers and Mathematics with Applications*, 40:877-883, 2000.

Sensitivity of an hysteretic material model for random vibration analysis of base-isolated rigid-block historical artifacts

Salvatore Sessa, Davide Pellecchia, Nicoló Vaiana, Luciano Rosati
DIST, University of Naples Federico II, 80125, Napoli, Italia, salvatore.sessa2@unina.it

Abstract. We present the computation and the implementation of the sensitivity relevant to a recently formulated phenomenological nonlinear material. Such a constitutive model is capable of reproducing several typologies of hysteretic loops without requiring iterative algorithms to determine its response. Moreover, it is based on a set of parameters having a peculiar physical meaning and provides responses characterized by smooth derivative. Hence, the model is particularly feasible to be implemented within a finite-element based reliability code aiming to perform tail-equivalent linearization and random vibration analysis. Within this context, the implemented model has been tested by analyzing a base-isolated structural system inspired at the Riace Bronze A. Numerical tests confirm the effectiveness of the sensitivity formulation and of its capabilities for tail-equivalent linearization purposes.

Keywords: art objects; nonlinear dynamics; tail-equivalent linearization.

1. Introduction

Nonlinear response of rigid blocks has received a significant attention during the last decade due to the need of assessing the seismic risk of museum artifacts (Roussis et al., 2008). Such a response depends on a complex interaction between the block base and a shaking surface, often presenting nonlinear behavior, and the dynamics of a rigid block.

In such a context, seismic isolation is conveniently adopted in order to reduce the vulnerability of art objects. Among several isolation technologies currently available, elastomeric bearings usually present hysteretic behaviors due to the flexibility and energy dissipation capacity of the rubber, whereas the same effects in sliding bearings rely upon friction and contact between two flat or curved steel surfaces. Moreover, wire ropes devices are typically adopted when the isolated system needs to be centered after a seismic excitation and when the static loads are not significant.

The hysteresis of such bearing typologies strongly influences the maximum displacements and accelerations of the isolated rigid block subjected to a base excitation. Within a seismic context, it is essential to determine the first excursion probability and to assess the properties of the bearing devices in order to maintain it within a safety interval.

To this end, Tail-equivalent Linearization Method (TELM), developed by Fujimura and Der Kiureghian (2007), is an effective and appealing procedure to perform random vibration analysis of such structural typologies since a failure condition of the isolated rigid block can be expressed by means of a limited set of responses.

©2021 *by S. Sessa, D. Pellecchia, N. Vaiana, L. Rosati.*

In particular, the TELM defines the base excitation as a sequence of random pulses. Then, fixed a response of interest and a set of threshold values, the procedure determines a *performance point* as the random process for which the response is equal to the threshold and turns out to be associated with the maximum probability of occurrence. Finally, the performance point permits the definition of a linearized system as a collection of impulse response functions.

The determination of the performance point is carried out by the application of a First-Order Reliability Method (FORM) actually consisting in a constrained-optimization algorithm (Ditlevsen and Madsen, 1996). Specifically, the FORM aims to maximize the probability relevant to the considered random variables (for the TELM case, these are the values in time of the base acceleration random process) by fulfilling a nonlinear limit state function. The latter is defined as difference between the response of interest and a fixed threshold.

An important issue of TELM consists in the fact that all the available algorithms are of iterative nature requiring, at each iteration, the computation of the gradient of the limit state function with respect to the random variables (Broccardo et al., 2017). For what concerns the TELM technique, this consists in computing the derivative of the structural response with respect to a set of acceleration values in which the base excitation process is discretized.

In seismic applications, accurate TELM analyses require to adopt properly discretized random processes lasting at least the time required by the structural system to reach stationarity. For this reason, the base excitation, and the related response gradient, can be made of several hundreds of components. Thus, a finite-difference-based computation of the response gradient would be excessively demanding from the computational point of view.

An essential step of the TELM is represented by the Direct Differentiation Method (DDM) which is able to numerically compute the response gradient with affordable computational burden (Haukaas and Der Kiureghian, 2006). However, its implementation represents the main limitation to a widespread diffusion of TELM since DDM is not common in commercial codes and requires the evaluation of the sensitivity of the response relevant to each object which is part of the analysis procedure.

Specifically, within the context of finite elements, the OpenSees framework includes an effective implementation of the DDM and of the TELM algorithm although this has been developed for a limited set of constitutive models. In particular, the most versatile one is the Bouc-Wen model.

Besides of its capability of reproducing several typologies of smooth hysteresis loops, such a model is affected by two significant drawbacks. First, the model is defined by means of an implicit differential equation governing the derivative of a non-linear generalized strain. This implies the necessity of determining the response by iterative approaches such as the Newton's or the Backward Euler algorithms. As a second, and more important consideration, the Bouc Wen material is a purely numerical model defined by means of a set of coefficients which lacks of physical meaning and can be identified, by matching experimental data, with significant difficulty.

To overcome such drawbacks, a new class of phenomenological constitutive models has been recently proposed by Vaiana et al. (2018) including, in particular, the *algebraic material* formulated in (Vaiana et al., 2019).

The latter model was developed for reproducing the hysteresis loops of several typologies of seismic devices including elastomeric bearings. Its main benefit consists in the fact that the response is computed in closed form so that no iterative algorithms are required. Moreover, the model is

defined by a set of physically meaningful parameters directly identifiable by experimental data so that their calibration is straightforward (Sessa et al., 2020).

Within this context, the paper illustrates the development of the response sensitivity of the algebraic material proposed in (Vaiana et al., 2019) to be implemented within a DDM in order to perform the TELM analysis of a base-isolated rocking rigid block simulating the Riace Bronze A.

Specifically, a brief review of the TELM is reported in Section 2 while the computation of the response gradient relevant to the algebraic material is formulated in Section 3. A numerical application consisting in a vulnerability analysis of the Riace Bronze A is illustrated in Section 4 and, finally, conclusions are discussed in Section 5.

2. Review of the Tail-Equivalent Linearization Method

The Tail-Equivalent Linearization Method, developed by Fujimura and Der Kiureghian (2007), defines a linearized model of a nonlinear structural system by means of a collection of Impulse Response Functions (IRFs) associated with different threshold values of a response of interest.

Such a strategy adopts a discretized description of a zero-mean, second-order Gaussian base excitation which is expressed as an array of random pulses:

$$F(t) = \sum_{i=1}^{n} s_i(t)\, \upsilon_i = \mathbf{s}^T(t)\, \boldsymbol{\upsilon} \tag{1}$$

where n is the total number of pulses, $\boldsymbol{\upsilon}$ is a vector of standard normal variables and $\mathbf{s}(t)$ is a deterministic vector depending on the covariance of the base excitation process.

We also denote by $X_{NL}(t_n, \boldsymbol{\upsilon})$ a response of interest of the nonlinear structural system at time t_n associated with a realization of the input random pulses $\boldsymbol{\upsilon}$. Fixed a threshold x of such a response, a limit state function is defined as:

$$\Theta(x, t_n, \boldsymbol{\upsilon}) = x - X_{NL}(t_n, \boldsymbol{\upsilon}) \tag{2}$$

Denoting by *tail probability* the quantity $\Pr[\Theta(x, t_n, \boldsymbol{\upsilon}) \leq 0]$, i.e., the probability that the response of interest is greater than the fixed threshold x, its first-order approximation is computed by applying the FORM algorithm (Ditlevsen and Madsen, 1996) provided that the limit state function has continuous gradient.

The FORM algorithm determines a *performance point* $\boldsymbol{\upsilon}^\star(x, t_n)$, associated with threshold x at time t_n, as the solution of the following constrained optimization problem:

$$\boldsymbol{\upsilon}^\star(x, t_n) = \arg\min\{\|\boldsymbol{\upsilon}\| \;|\Theta(x, t_n, \boldsymbol{\upsilon}) = 0\} \tag{3}$$

In particular, the first-order approximation of the tail probability turns out to be:

$$\Pr[\Theta(x, t_n, \boldsymbol{\upsilon}) \leq 0] \propto \Phi[-\beta(x, t_n)] \tag{4}$$

where Φ denotes the standard-normal cumulative probability function and β is the reliability index:

$$\beta(x, t_n) = \boldsymbol{\alpha}(x, t_n)\, \boldsymbol{\upsilon}^\star(x, t_n) \tag{5}$$

where:

$$\boldsymbol{\alpha}\left(x,\, t_n\right) = -\frac{\nabla_{\boldsymbol{\upsilon}}\Theta\left(x,\, t_n,\, \boldsymbol{\upsilon}^{\star}\right)}{\left\|\nabla_{\boldsymbol{\upsilon}}\Theta\left(x,\, t_n,\, \boldsymbol{\upsilon}^{\star}\right)\right\|} \tag{6}$$

is the negative unit vector parallel to the gradient of the limit state function at the performance point $\boldsymbol{\upsilon}^{\star}$.

The *tail-equivalent linear oscillator* associated with time t_n and threshold x is therefore defined as a linear, single degree-of-freedom system having a tail-probability equal to the first-order tail probability of the nonlinear structure. In this sense, the linear response $X\left(t\right)$ can be expressed by means of its Impulse Response Function (IRF) $h\left(t\right)$ by a convolution integral:

$$X\left(t\right) = \int_0^t h\left(t-\tau\right) \sum_{i=1}^{n} s_i\left(\tau\right) \upsilon_i \, d\tau = \mathbf{a}\left(t\right)^T \boldsymbol{\upsilon} \tag{7}$$

where the deterministic vector $\mathbf{a}\left(t\right)$ is defined by the components:

$$a_i\left(t\right) = \int_0^t h\left(t-\tau\right) s_i\left(\tau\right) \, d\tau \tag{8}$$

Referring to the original paper (Fujimura and Der Kiureghian, 2007) for further details, we emphasize that enforcing equality between the response of the linear and of the nonlinear systems, and exploiting geometrical considerations, it is possible to relate such a deterministic vector to the performance point:

$$\mathbf{a}\left(t_n\right) = \frac{x}{\left\|\boldsymbol{\upsilon}^{\star}\left(x,\, t_n\right)\right\|} \frac{\boldsymbol{\upsilon}^{\star}\left(x,\, t_n\right)}{\left\|\boldsymbol{\upsilon}^{\star}\left(x,\, t_n\right)\right\|} \tag{9}$$

Hence, computed the performance point associated with time t_n and threshold x, it is possible to determine in turn the vector $\mathbf{a}\left(t_n\right)$ and an impulse response function $h^{\star}\left(x,\, t_n,\, t_i\right)$. Such a function, referred to the linearized system, is time discretized in terms of the set of instants t_i and associated with x and t_n.

The *Tail-Equivalent Linearized System* (TELS) it determined by adopting a time t_n sufficiently large to make the nonlinear system reach stationarity and an exhaustive set of thresholds x; in this sense, the TELS is defined by means of a collection of impulse response functions.

3. Sensitivity of the algebraic material

A key issue of the TELM concerns the determination of the performance point by the FORM algorithm applied to the optimization problem in Eq. (3) as well as the computation of the tail probability. Both tasks need the gradient $\nabla_{\boldsymbol{\upsilon}}\Theta\left(x,\, t_n,\, \boldsymbol{\upsilon}^{\star}\right)$ of the limit state function and, subsequently, of the response $X_{NL}\left(t_n,\, \boldsymbol{\upsilon}\right)$.

It is convenient to determine such a gradient by a DDM procedure in order to limit the computational burden. To this end, the constitutive models adopted in the analysis have to be equipped with their sensitivity with respect to all variables depending on the base excitation.

For this reason we present a brief review of the algebraic material formulated in (Vaiana et al., 2019) and report the detailed computation of its sensitivity required by the TELM procedure.

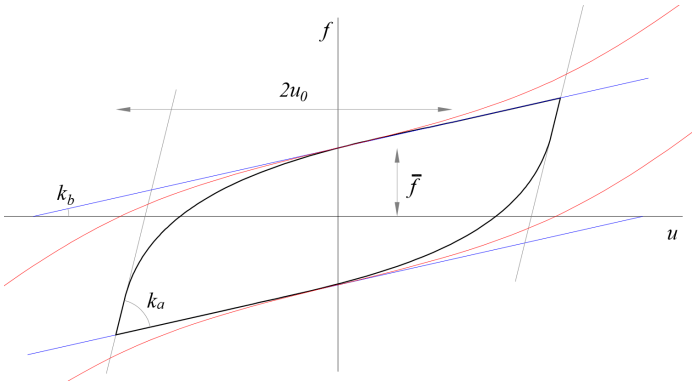


Figure 1. Riace Bronze A resting on a seismic-isolated basement.

The algebraic material is defined by means of five constitutive parameters. Namely, k_a is the initial stiffness, k_b is the asymptotic stiffness and α is a transition parameter between the initial elastic phase and the inelastic one. Parameters β_1 and β_2 influence the shape of the hysteretic loop by introducing local maxima and pinching.

Moreover, the model uses the following auxiliary parameters:

$$u_0 = \frac{1}{2}\left[\left(\frac{k_a - k_b}{\delta}\right)^{\frac{1}{\alpha}} - 1\right]; \quad \bar{f} = \frac{k_a - k_b}{2}\left[\frac{(1+2u_0)^{1-\alpha} - 1}{1-\alpha}\right] \tag{10}$$

where δ denotes a geometric tolerance which is usually set equal to 10^{-20}.

A physical interpretation of the parameters is schematized in Figure 1. Specifically, the hysteresis loop presents an initial stiffness k_a and asymptotically tends to a boundary line or curve with vertical intercept at $\bar{f}$. If β_1 and β_2 are both zero, the loop tends to the asymptotic lines with tangent k_b represented in blue in Figure 1. Parameters β_1 and β_2 introduce a curvature of the asymptotic boundary by means of a 5th degree polynomial function (see, e.g., the red curves in Figure 1).

The auxiliary variable u_0 is defined so that if displacement equal to $2u_0$ is applied to a load-displacement state belonging to one of the boundary curves, the distance between the new load-displacement state and the boundary curve is equal to the tolerance δ.

To determine the closed-form expression of the load, a history variable u_j is necessary:

$$u_j = u_c + s_t\,(1+2u_0) - s_t\left\{\frac{s_t\,(1-\alpha)}{k_a - k_b}\left[f_c - \beta_1 u_c^3 - \beta_2 u_c^5 - k_b u_c - s_t\bar{f} + (k_a - k_b)\,\frac{(1+2u_0)^{1-\alpha}}{s_t\,(1-\alpha)}\right]\right\}^{\frac{1}{1-\alpha}} \tag{11}$$

where u_c denotes the current displacement and s_t the algebraic sign of the velocity.

Denoting by f_c the force associated with the displacement u_c and by u_t a *trial displacement*, the trial force is given by:

$$f\,(u_t, u_c,\, s_t) = \beta_1 u_t^3 + \beta_2 u_t^5 + k_b u_t + (k_a - k_b)\left[\frac{(1 + s_t u_t - s_t u_j + 2u_0)^{1-\alpha}}{s_t\,(1-\alpha)} - \frac{(1+2u_0)^{1-\alpha} - 1}{1-\alpha}\right] + s_t\bar{f} \tag{12}$$



Figure 2. Riace Bronze A resting on a seismic-isolated basement.

Within the TELM framework, the sensitivity of the trial force must be computed with respect to quantities related to the sequence of acceleration pulses. In particular, the derivatives of the displacements u_c and u_t are computed by the DDM procedure disposing of the derivative of $f\left(u_t, u_c,\, s_t\right)$ with respect to such displacements as well as with respect to f_c.

Derivatives of the auxiliary parameters are zero while those of the history variable turn out to be:

$$\frac{\partial u_j}{\partial f_c} = 1 - \frac{s_t^2}{k_a - k_b}\left\{\frac{s_t\left(1-\alpha\right)}{k_a - k_b}\left[f_c - \beta_1 u_c^3 - \beta_2 u_c^5 - k_b u_c - s_t \bar{f} + \left(k_a - k_b\right)\frac{\left(1+2u_0\right)^{1-\alpha}}{s_t\left(1-\alpha\right)}\right]\right\}^{\frac{\alpha}{1-\alpha}} \tag{13}$$

$$\frac{\partial u_j}{\partial u_c} = \frac{\partial u_j}{\partial f_c}\left[-3\beta_1 u_c^2 - 5\beta_2 u_c^4 - k_b\right] \tag{14}$$

while the derivatives of the trial force are:

$$\frac{\partial f\left(u_t, u_c,\, s_t\right)}{\partial f_c} = -\frac{\partial u_j}{\partial f_c}\left(k_a - k_b\right)\left[1 + s_t u_t - s_t u_j + 2u_0\right]^{-\alpha} \tag{15}$$

$$\frac{\partial f\left(u_t, u_c,\, s_t\right)}{\partial u_c} = -\frac{\partial u_j}{\partial u_c}\left(k_a - k_b\right)\left[1 + s_t u_t - s_t u_j + 2u_0\right]^{-\alpha} \tag{16}$$

and

$$\frac{\partial f\left(u_t, u_c,\, s_t\right)}{\partial u_t} = 3\beta_1 u_t^2 + 5\beta_2 u_t^4 + k_b + \left(k_a - k_b\right)\left[1 + s_t u_t - s t u_j + 2u_0\right]^{-\alpha} \tag{17}$$

The algebraic material has been implemented in OpenSees, an object oriented and open-source framework for finite element analysis in which the DDM and the TELM algorithm have been implemented within a structural reliability environment (Haukaas and Der Kiureghian, 2006).

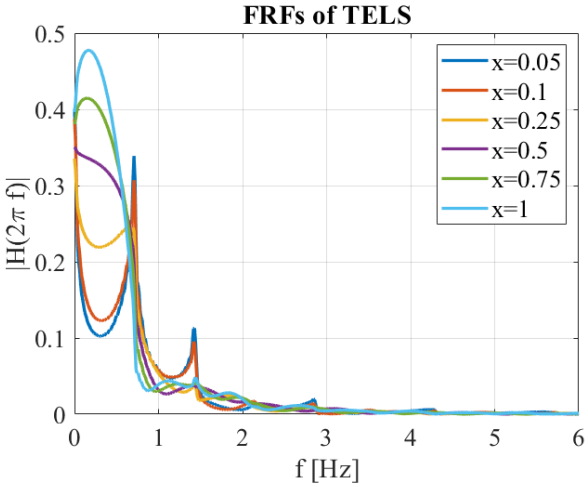


Figure 3. Frequency response functions of the TELS.

4. TELM analysis of the Riace Bronze A

One of the most appealing applications of TELM concerns seismic isolation since the safety of isolated structures can be easily analyzed by computing a single response of interest. Moreover, significant applications of base isolation have been recently developed for archaeological artifacts, e.g. slender statues, because of their vulnerability with respect to seismic actions (Calió and Marletta, 2004; Zuccaro et al., 2019).

In particular, base isolation has recently implemented on the Riace Bronzes (De Canio, 2012) in order to avoid vibrations induced by earthquakes. Bronze A, shown in Figure 2, has been chosen as benchmark for testing the implementation of the algebraic material in TELM.

The bronze, whose inertial properties are reported in (Pellecchia et al., 2020), is constrained to a marble basement equipped with elastomeric bearings. The basement-statue system has a mass of 4235 kg while the isolation device has been modeled by a truss element, having unit length and area of $0.1\,m^2$, made of the algebraic material with parameters reported in Table I.

Table I. Parameter set used in the numerical application.

k_a, $[kN/m]$	k_b, $[kN/m]$	α	β_1, $[kN/m^3]$	β_2, $[kN/m^5]$
1714.95	85.74	25	0	0

Base excitation has been modeled by an unfiltered Gaussian white noise lasting $t_n = 25\,s$ and discretized with time step of $0.01\,s$ corresponding to a cutoff frequency of 50 Hz. The horizontal displacement of the basement has been adopted as response of interest of the TELM procedure for which a set of 20 increasing values of the threshold have been adopted.

The Tail-Equivalent Linearized System (TELS) has been defined by means of a set of Impulse Response Functions. For convenience, it is possible to represent the TELS functions in the frequency

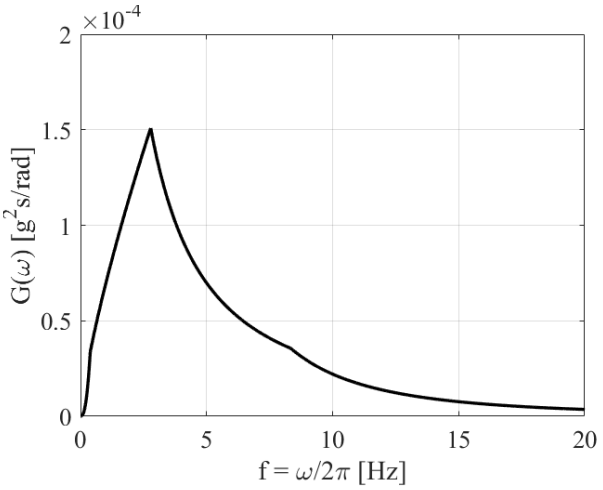


Figure 4. Single-sided Power Spectral Density adopted in the computations (Barone et al., 2015)

domain by computing their Fourier transform:

$$H^{\star}\left(x,\, t_n,\, \omega\right) = \int_0^{\infty} h^{\star}\left(x,\, t_n,\, \tau\right) \exp\left(-i\omega\tau\right)\, d\tau \tag{18}$$

Such Frequency Response Functions (FRFs) represent the amplitude of the steady-state response induced by an external excitation $f\left(t\right) = \exp\left(i\omega t\right)$. A limited set of FRFs belonging to the TELS is shown in Figure 3.

The main properties of the FRFs are consistent with what is usually expected by the application of TELM. Specifically, as the threshold increases, the peaks of the FRFs, representing the main frequencies of the system, tend to translate leftwards and the amplitude of each peak turns out to be inversely proportional to the threshold.

Such a behavior mainly depends on two aspects: the shifting to the lower frequencies depends on the smaller stiffness of the hysteretic material as long as the displacements increase; the smaller value of the FRF depends on the higher dissipation of elastic energy during the hysteresis as long as force and displacement become larger making the cycles wider.

To perform random vibration analysis, the base excitation power spectral density proposed by Barone et al. (2015) has been adopted. The main benefit of such a function consists in the fact that it is related to the design response spectrum functions provided by the Italian Structural Code so that its parameters can be calibrated consistently with the code hazard coefficients.

The single-sided power spectral density is defined by:

$$G_{FF}\left(\omega\right) = \begin{cases} G_0 \left(\frac{\omega_D}{\omega_C}\right)^{e_2} \left(\frac{\omega}{\omega_D}\right)^{e_1} & 0 \leq \omega \leq \omega_D \\ G_0 \left(\frac{\omega}{\omega_C}\right)^{e_2} & \omega_D \leq \omega \leq \omega_C \\ G_0 \left(\frac{\omega}{\omega_C}\right)^{e_3} & \omega_C \leq \omega \leq \omega_B \\ G_0 \left(\frac{\omega_B}{\omega_C}\right)^{e_3} \left(\frac{\omega}{\omega_B}\right)^{e_4} & \omega \geq \omega_B \end{cases} \tag{19}$$

where $\omega = 2\pi/T$ is the pulsation, G_0 is the intensity, ω_B, ω_C and ω_D are the pulsations associated with standard code periods T_B, T_C and T_D, respectively, and $e_1 \dots e_4$ are exponents calibrated with respect to the design response spectrum.

Coefficients of the power spectral density reported in Table II have been chosen according to (Barone et al., 2015) and are relevant to the Messina Strait area. The function is represended in Figure 4.

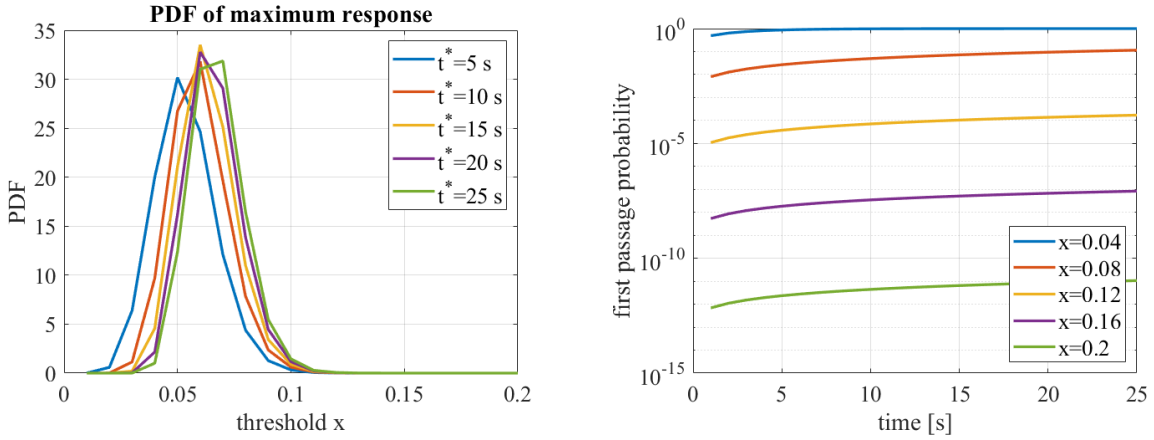


Figure 5. Probability density function of the maximum response and First Passage Probability of the TELS.

A series-system analysis (Der Kiureghian and Fujimura, 2009) permits to compute the statistics of the maximum response with respect to different time intervals $t^\star$ and to the First-Passage Probability of the structural system, both reported in Figure 5.

The probability density functions of the maximum response show that the procedure is capable of determining non-Gaussian distributions. Moreover, as larger time intervals are adopted, the modal value of the PDF increases as well as the corresponding maximum of the probability density while the standard deviation decreases.

In order to check the effectiveness of the isolation technique for seismic protection, it is useful to compute the statistics of the acceleration of the isolated block. In fact, we remind that the Riace Bronzes were isolated in order to reduce vibrations induced by seismic motions.

To this purpose, the peak ground acceleration of the base motion, equal to 0.248g, can be compared with the square-rooted mean square $\sigma_{\ddot{X}\ddot{X}}$ of the second derivative of the response of interest. Referring to (Lutes and Sarkani, 2004) for further details, we recall that the mean-square of the (acceleration) process representing the second derivative of the response can be computed

Table II. Parameters of the adopted Power Spectral Density.

G_0	T_B	T_C	T_D	e_1	e_2	e_3	e_4
$[g^2 s/rad]$	$[s]$	$[s]$	$[s]$	–	–	–	–
$1.51 \cdot 10^{-4}$	0.12	0.359	2.592	2.5094	0.7594	−1.3177	−2.6209

as:

$$\sigma^2_{\ddot{X}\ddot{X}}(\omega,\, x,\, t_n) = \int_0^\infty G_{FF}(\omega)\, |H^\star(x,\, t_n,\, \omega)|^2\, \omega^4\, d\omega \tag{20}$$

where dependency on x and t_n is due to the fact that such a parameter is relevant to the linearized system.

Figure 6 shows the values of such acceleration's root-mean-square as function of the threshold. Actually, because of the isolation system, the acceleration acting at the base of the statue is at least 30% smaller than the ground acceleration, this confirming the convenience in adopting such a technique. However, we remark that such a comparison is purely qualitative since the constitutive parameters of the seismic isolating devices have not been determined by a proper design procedure; actually the purpose of the present application is to check the effectiveness of the the algebraic material sensitivity within the TELM context.

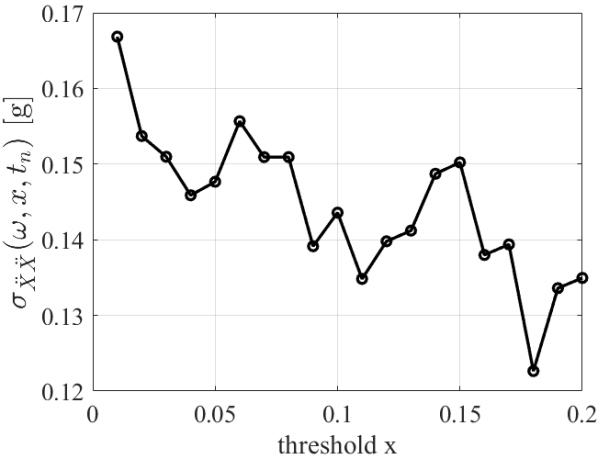


Figure 6. Square root of the mean-square of the acceleration of the isolated block as function of the threshold.

5. Conclusions and future work

The implementation of the sensitivity of a recently formulated phenomenological constitutive model has been presented. Such a constitutive model is capable of reproducing several typologies of hysteretic loops by means of a closed-form analytical expression and does not require iterative procedures to determine the response.

The sensitivity of the constitutive model, intended as the derivative of the response with respect to all quantities depending by the base acceleration, namely, the displacement and the initial force, has been analytically computed and implemented in OpenSees, an open source framework for finite element analysis. Within the context of a Direct Differentiation Method, the implemented material can be used in Tail Equivalent Linearization

Numerical application concerning the random vibration analysis of a base-isolated archaeological artifact shows the effectiveness of the implementation and of the whole procedure.

Acknowledgements

The present research was supported by the University of Naples Federico II and the *Compagnia di San Paolo*, which are gratefully acknowledged by the author, as part of the Research Project *MuRA – Multi-risk assessment and structural protection of archaeological vestiges in volcanic scenarios*, FRA grants, CUP E69C21000250005.

References

Barone, G., Iacono, F.L., Navarra, G., Palmeri, A. A novel analytical model of power spectral density function coherent with earthquake response spectra. In: *Proceedings of the 1st ECCOMAS Thematic Conference on Uncertainty Quantification in Computational Sciences and Engineering, UNCECOMP 2015*, 394-406, 2015.

Broccardo, M., Alibrandi, U., Wang, Z., Garré, L. The tail equivalent linearization method for nonlinear stochastic processes, genesis and developments. In: P. Gardoni editor *Risk and Reliability Analysis: Theory and Applications. Springer Series in Reliability Engineering*. Springer, Cham, 109-142, 2017.

Calió, I., Marletta, M. On the mitigation of the seismic risk of art objects: case-studies. In: *Proceedings of the 13th World Conference on Earthquake Engineering*, 2828, 2004.

De Canio, G. Marble devices for the Base isolation of the two Bronzes of Riace: a proposal for the David of Michelangelo. In: *Proceedings XV World Conference on Earthquake Engineering-WCEE*, 24-28, 2004.

Der Kiureghian, A., Fujimura, K. Nonlinear stochastic dynamic analysis for performance-based earthquake engineering. *Earthquake Engineering and Structural Dynamics*, 38(5): 719-738, 2009.

Ditlevsen, O., Madsen, H. O. Structural reliability methods. *Wiley, New York.*, 1996.

Fujimura, K., Der Kiureghian, A., Tail-equivalent linearization method for nonlinear random vibration. *Probabilistic Engineering Mechanics*, 22(1): 63-76, 2007.

Haukaas, T., Der Kiureghian, A., Strategies for finding the design point in non-linear finite element reliability analysis. *Probabilistic Engineering Mechanics*, 21(2): 133-147, 2006.

Lutes, L.D., Sarkani, S., Random Vibrations. Elsevier Butterworth-Heinemann, Burlington, MA, USA., 2004.

Pellecchia, D., Sessa, S., Vaiana, N., Rosati, L., Comparative Assessment on the Rocking Response of Seismically Base-Isolated Rigid Blocks. *Procedia Structural Integrity*, 29(C): 95-102, 2020.

Roussis, P.C., Pavlou, E.A., Pisiara, E.C., Base-isolation technology for earthquake protection of art objects. In: *Proceedings of the 14th World Conference on Earthquake Engineering: Innovation Practice Safety*, China, 2008.

Sessa, S., Vaiana, N., Paradiso, M. Rosati, L. An inverse identification strategy for the mechanical parameters of a phenomenological hysteretic constitutive model. *Mechanical Systems and Signal Processing*, 139: 106622, 2008.

Vaiana, N., Sessa, S., Marmo, F., Rosati, L. A class of uniaxial phenomenological models for simulating hysteretic phenomena in rate-independent mechanical systems and materials. *Nonlinear Dynamics*, 93: 1647-1669, 2018.

Vaiana, N., Sessa, S., Marmo, F., Rosati, L. An accurate and computationally efficient uniaxial phenomenological model for steel and fiber reinforced elastomeric bearings. *Composite Structures*, 211: 196-212, 2019.

Zuccaro, G., Dato, F., Cacace, F., De Gregorio, D., Sessa, S. Seismic collapse mechanisms analyses and masonry structures typologies: A possible correlation. *Ingegneria Sismica*, 34(4): 121-149, 2019.

An Evolutive Probability Transformation Method for the Dynamic Stochastic Analysis of Structures

Rossella Laudani and Giovanni Falsone
Department of Engineering, University of Messina, Italy. E-mail: rlaudani@unime.it

Abstract. The problem of the stochastic response determination of dynamic systems, once that the random properties of the actions are assigned, is studied. In the present work, an extension of the Probability Transformation Method (PTM) to the dynamic systems has been proposed. In particular, an evolutive PTM (EPTM) is introduced. Based on the properties of the mean square random calculus with the principle of conservation of probability, the EPTM applies the PTM contextually to the mean square Riemann integral of the system solution process. In this way, the EPTM gives the expression of the single-time evolutive characteristic function (CF) of the system's output. The application to a shear-type plane system subject to assigned stochastic process excitations and random initial conditions has revealed the efficiency of the stochastic procedure.

Keywords: Evolutive Probability Transformation Method, stochastic dynamic analysis, random process

1. Introduction

Over the last two decades, it has been recognized the uncertain character in the natural phenomena. Most physical behaviors exhibit appreciable randomness which can not be adequately represented by deterministic models. Stochastic differential equations are often used to model the stochastic dynamics of uncertain systems. Hence there is the need to carry out a stochastic analysis that describes the response in probabilistic terms. Efficient probabilistic characterization of the dynamical response of a system excited by random actions often requires a high computational effort. In the last fifty years, many significant results have been obtained in this field. Most of the approaches in the literature consider the solving input-output relations in terms of evolutive moments, or cumulants, or quasi-moments, or, for having a multiple times definition, in terms of correlation functions of various order (Lin, 1967; Roberts and Spanos, 1991; Wu and Lin, 1984; Lutes and Sarkani, 2004; Di Paola *et al.*, 1992; Di Paola and Falsone, 1994; Falsone, 1994; Di Paola and Falsone, 1997a,b; Falsone, 2005; Morikawa and Kameda, 1997; Makarios, 2012; Gioffrè and Gusella, 2002; Mazelsky, 1954; Bucher and Schueller, 1991). Unfortunately, all these quantities suffer the drawback of having high dimensions, above all for large systems.

In some problems, such as the structural reliability evaluation or the stochastic limit analysis, the accurate knowledge of the response PDF is essential, above all at the PDF tails, (Di Paola and Falsone, 1994; Falsone, 1994). However, only a few works perform dynamic stochastic analyses in terms of probability density function (PDF) (Conte and Peng, 1996; Adhikari, 2007; Hussein and Selim, 2015; Li, 2016; Calatayud *et al.*, 2018b,a; Liu and Liu, 2018; Kalogeris and Papadopoulos,

© 2021 *by Rossella Laudani and Giovanni Falsone.*

2018; Meimaris *et al.*, 2019; Mamis *et al.*, 2019). In particular, the present authors have extended to the dynamic analysis an approach, previously introduced for static analyses and called probability transformation method (PTM) (Falsone and Settineri, 2013a,b). This last one is essentially based on the rules of random variable transformations and on the principle of probability conservation. The above-cited extension to the dynamic analyses was made through a time discretization in the integral expression of the structural response (Falsone and Laudani, 2018, 2020). In this way, the input-output relationships are considered as algebraic equations and the PTM can be applied as in the static analyses.

In the present work, a study on the evaluation of the response of dynamic systems governed by first-order differential equations will be presented. Problems described by this type of equations arise in a wide variety of applied engineering areas. Within the framework of the mean square random calculus, a stochastic procedure that combines the properties of the PTM with an approach able to find numerically the relationship between a process and its time integral by working in terms of characteristic functions (CF) will be developed. Therefore, an evolutive PTM (EPTM) will be is introduced. In this way, while in the previous approach the PTM is applied after the numerical discretization (Falsone and Laudani, 2018, 2020), the EPTM applies the PTM contextually to the integral procedure. This makes the EPTM a very efficient approach for the dynamic stochastic analyses of systems described by first-order differential systems, as has been verified in a shear-type plane system that will be presented. In particular, this work performs the response PDF considering the following two stochastic actions of the system, (i) time-dependent actions represented as random processes, (ii) system initial conditions represented as random variables. For both these two possibilities, which could happen contemporaneously, the expression of the single-time varying CF output of a stochastic dynamic system has been obtained.

2. Preliminary concepts

In this section, some basic concepts, which will be useful in the following sections, are shown. In particular, the PTM is recalled.

The fundamental aspects of the PTM can be found in the theory of the space transformation of random vectors as well as in the *principle of probability conservation*. It is possible to state the principle of probability conservation as following: *the probability carried by a random event is conserved without introducing other stochastic factors. In other words, if the random factors involved in a stochastic system are preserved, i.e no new random factors arise nor existing factors vanish in a physical process, then the probability will be preserved in the evolution process of the system* (Li and Chen, 2009; Soong, 1973). In particular, the PTM allows working directly in terms of input and output probability density functions (PDFs) of two random vectors related to each other by the deterministic law corresponding to the assigned space transformation.

Let consider a n-dimensional random vector, $\mathbf{x}$, with joint probability density function (JPDF), $p_{\mathbf{x}}(\mathbf{x})$, and let $\mathbf{h}(\cdot)$ be a n-dimensional invertible application, with inverse $\mathbf{h}(\cdot)^{-1} = \mathbf{g}(\cdot)$. If $\mathbf{z} = \mathbf{h}(\mathbf{x})$ is a random vector whose JPDF, under the respect of the principle of probability conservation, can

be expressed as follows:

$$p_{\mathbf{z}}(\mathbf{z}) = \frac{1}{|\det\left[\mathbf{J}_{\mathbf{h}}(\mathbf{g}(\mathbf{z}))\right]|} p_{\mathbf{x}}(\mathbf{g}(\mathbf{z})) = |\det\left[\mathbf{J}_{\mathbf{g}}(\mathbf{z})\right]|\, p_{\mathbf{x}}(\mathbf{g}(\mathbf{z})) \tag{1}$$

where $\mathbf{J}_{\mathbf{h}}(\cdot)$ and $\mathbf{J}_{\mathbf{g}}(\cdot) = \mathbf{J}_{\mathbf{h}}^{-1}(\cdot)$ are the Jacobian matrices corresponding to the relation $\mathbf{z} = \mathbf{h}(\mathbf{x})$ and the inverse $\mathbf{x} = \mathbf{g}(\mathbf{z})$. The relationship in Eq. (1) holds when $\mathbf{x}$ and $\mathbf{z}$ have the same number of components. Nevertheless, this is not a restriction, in Falsone and Laudani (2019a) some procedures are given to apply it even when these numbers of components are different.
In Falsone and Settineri (2013a,b) it was introduced an efficient way to evaluate the marginal PDF of $\mathbf{z}$, $p_{z_j}(z_j)$. By using the properties of the multidimensional Dirac delta function, $\delta(\cdot)$ Eq. (1) is rewritten in the following form:

$$p_{\mathbf{z}}(\mathbf{z}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_{\mathbf{x}}(\mathbf{y}) \delta(\mathbf{z} - \mathbf{h}(\mathbf{y})) dy_1 \cdots dy_n. \tag{2}$$

The use of the latter expression is particularly useful if the marginal PDF evaluation of $\mathbf{z}$ is required. For example, if only the element $z_j = h_j(\mathbf{x})$ has to be studied, then Eq. (1) gives:

$$p_{z_j}(z_j) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_{\mathbf{x}}(\mathbf{y}) \delta(z_j - h_j(\mathbf{y})) dy_1 \cdots dy_n \tag{3}$$

while, if the JPDF of the elements $z_j = h_j(\mathbf{x})$ and $z_k = h_k(\mathbf{x})$ has to be evaluated, then the following expression is suitable:

$$p_{z_j,z_k}(z_j, z_k) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_{\mathbf{x}}(\mathbf{y}) \delta(z_j - h_j(\mathbf{y})) \delta(z_k - h_k(\mathbf{y})) dy_1 \cdots dy_n. \tag{4}$$

It is worth noting that Eqs. (2) to (4) do not require the evaluation of the inverse relation $\mathbf{g}(\cdot) = \mathbf{h}(\cdot)^{-1}$, that often can be a great problem. Moreover, the order of the vector $\mathbf{z}$ does not influence them. On the contrary, they have the drawback of requiring n integrations respect to the component of $\mathbf{x}$. This last problem can be easily solved in terms of characteristic function (CF) when $h_j(\mathbf{x})$ is given as a linear combination of the elements of $\mathbf{x}$. This means the relation $h_j(\mathbf{x}) = \mathbf{h}_j^T \mathbf{x}$, where $\mathbf{h}_j$ is the n-vector made by the combination constants. Indeed, the CF of z_j is given by:

$$\begin{aligned} M_{z_j}(\omega) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} p_{z_j}(z_j) \exp(-i\omega z_j) dz_j \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_{\mathbf{x}}(\mathbf{y}) \delta(z_j - \mathbf{h}_j^T \mathbf{y}) dy_1 \cdots dy_n \right] \exp(-i\omega z_j) dz_j. \end{aligned} \tag{5}$$

that, taking into account the properties of the Dirac delta function, gives:

$$M_{z_j}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_{\mathbf{x}}(\mathbf{y}) \exp(-i\omega \mathbf{h}_j^T \mathbf{y}) dy_1 \cdots dy_n = (2\pi)^{n-1} M_{\mathbf{x}}(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\omega \mathbf{h}_j}. \tag{6}$$

This expression evidences the fundamental result that the response CF, $M_{z_j}(w_j)$, is always obtainable in closed form, once that the multidimensional CF of the input is known, without the

necessity of any integration. If the joint CF (JCF) of the output random variables z_j and z_k has to be evaluated, then the following expression, easily obtained by generalizing the previous one can be used:

$$M_{z_j z_k}(\omega_j, \omega_k) = (2\pi)^{n-2} M_{\mathbf{x}}(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\omega_j \mathbf{h}_j + \omega_k \mathbf{h}_k} \tag{7}$$

The generalization to greater order JCFs is immediate. Once that the characteristic functions are evaluated the corresponding PDF can be obtained by Fourier anti-transform operations.

Therefore, from this section, it is worth highlighting the following essential concept: *in a generic stochastic system if no other stochastic factors are involved then the probability carried by a random event is conserved.*

3. Stochastic first-order differential equation systems

A large class of physical behaviors and natural phenomena can be described by ordinary differential equations. Taking into account the effects of the various inherent uncertainties in nature when the dynamic analysis of deterministic systems is performed, it may arise the necessity to perform a stochastic analysis. It is clear that although a generic system of equations involves a set of higher-order differential equations, it is always possible, by introducing more variables, to convert higher-order differential equations to a set of coupled first-order equations. In view of these considerations, in this section, stochastic first-order differential systems will be treated. In general, the uncertain behavior of the system is due to (i) time-dependent actions represented as random processes, (ii) system initial conditions represented as random variables. Both these possibilities, which could happen contemporaneously, will be investigated in the following subsections, respectively.

3.1. first-order differential systems excited by a given random process

Let consider a first-order differential system governed by the following equation:

$$\dot{\mathbf{X}}(t) = \mathbf{A}\mathbf{X}(t) + \mathbf{B}\mathbf{F}(t); \quad \mathbf{X}(t_0) = \mathbf{X}_0 \tag{8}$$

where $\mathbf{X}(t)$ is the n-vector of the response state variables, $\mathbf{A}$ is the $n \times n$ deterministic matrix which takes into account the physical-geometrical characteristics of the system, $\mathbf{B}$ is the $n \times m$ matrix defining the distribution of the external loads on the system, $\mathbf{F}(t)$ is the m-random vector whose probabilistic characterization is known and, at last, the vector $\mathbf{X}_0$ defines the initial conditions of the system that, here, are assumed to be deterministic. Assuming that the m.s. derivative $\dot{\mathbf{X}}(t)$ of $\mathbf{X}(t)$ exists, the solution of Eq. (8) is well known:

$$\mathbf{X}(t) = \boldsymbol{\Theta}(t - t_0)\mathbf{X}_0 + \int_{t_0}^{t} \boldsymbol{\Theta}(t - \tau)\mathbf{B}\mathbf{F}(\tau)d\tau \tag{9}$$

$\boldsymbol{\Theta}(t) = \exp(-\mathbf{A}t)$ being the fundamental matrix of the system. The first of the two quantities composing $\mathbf{X}(t)$ is deterministic, under the assumptions made here, while the second one defines a vector random process. Hence, it is useful to write:

$$\mathbf{X}(t) = \mathbf{Y}_0(t) + \mathbf{Y}(t); \quad \text{with} \quad \mathbf{Y}_0(t) = \boldsymbol{\Theta}(t - t_0)\mathbf{X}_0, \quad \mathbf{Y}(t) = \int_{t_0}^{t} \boldsymbol{\Theta}(t - \tau)\mathbf{B}\mathbf{F}(\tau)d\tau. \tag{10}$$

By assuming zero initial conditions, the attention is paid only on the vector random process $\mathbf{Y}(t)$ and on its relationship with the load random process $\mathbf{F}(t)$.
If $\mathbf{F}(t)$ is a Gaussian vector process, then it is simple to show that $\mathbf{Y}(t)$ is Gaussian, too. Hence, it is characterized by the knowledge of its mean vector and its cross-covariance matrix. When $\mathbf{F}(t)$ is not-Gaussian, the probabilistic characterization of the response becomes more complicated. At this scope, the following vector process is introduced:

$$\mathbf{Z}(\tau) = \mathbf{\Theta}(t-\tau)\mathbf{B}\mathbf{F}(\tau) \tag{11}$$

and Eq. (10) is rewritten as:

$$\mathbf{Y}(t) = \int_{t_0}^{t} \mathbf{\Theta}(t-\tau)\mathbf{B}\mathbf{F}(\tau)d\tau = \int_{t_0}^{t} \mathbf{Z}(\tau)d\tau. \tag{12}$$

Its JPDF $p_{\mathbf{Z}}(\mathbf{z},\tau)$ can be easily obtained through the use of the PTM rules to the linear transformation $\mathbf{F}(\tau) \to \mathbf{Z}(\tau)$. The last step to do is the characterization of the random vector process $\mathbf{Y}(t)$, once that the characterization of the vector process $\mathbf{Z}(t)$ has been found. Due to the time integral relationship between these two processes, an efficient numerical solution can be obtained by using the following theorem on the JCFs of $\mathbf{Y}(t)$ and $\mathbf{Z}(t)$:

THEOREM 1. *((Soong, 1973) If the m.s. integral* $\mathbf{X}(t)$*,* $t \in T$*, exists, then:*

$$M_{\mathbf{Y}}(\omega,t) = \lim_{m\to\infty \Delta m\to 0} M_{\mathbf{Z}}(\omega(\tau_1-\tau_0),\tau_1';\cdots;\omega(\tau_m-\tau_{m-1}),\tau_m') \quad \tau_j' \in (\tau_{j-1},\tau_j) \tag{13}$$

$\Delta_m = \max(\tau_{j-1}-\tau_j)$.

The result of this theorem is mainly based on the mean square calculus as well as on the principle of conservation of probability. In fact, employing some of the theory of random differential equations and assuming that $\mathbf{Y}(t)$ is a mean square Riemann integral over a generic interval $[a,b]$, it can be written:

$$\int_a^b \mathbf{Y}(t)dt \equiv \lim_{\Delta m\to 0} \sum_{j=1}^{m} \mathbf{Y}(t_j')\Delta t_j \tag{14}$$

where a collection of all finite partitions of the interval $[a,b]$ was considered as in Figure 1. Then, in accordance with the principle of conservation of probability, the formal relationship between the characteristic functions of $\mathbf{Y}(t)$ and that of $\mathbf{Z}(t)$ (Eq. (13)) can be obtained. Therefore, by truncating the value of m to a sufficiently high value, the expression of this theorem becomes an efficient way to evaluate numerically the response CF. Then, the response PDF is obtained by the Fourier anti-transform of the response CF. Moreover, if the characterization of the response random process is required at more instant, Eq. (13) must be generalized. This is always possible, even if the corresponding numerical evaluation becomes more and more heavy increasing the number of time instants.

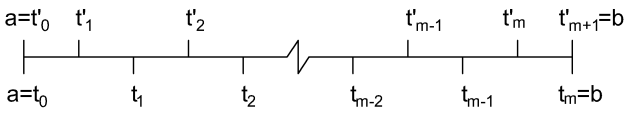


Figure 1.: Finite partitions of an interval $[a,b]$.

Overall, the application of the PTM with the fundamental results of the above Theorem 1 gives a stochastic procedure, here called EPTM, able to characterize the response of dynamical systems in terms of evolutive PDF.

3.2. FIRST-ORDER DIFFERENTIAL SYSTEMS WITH RANDOM LOAD AND RANDOM INITIAL CONDITIONS

Here, besides the loads acting on the system, even the initial conditions are assumed to be random variables. In particular, in this section, the EPTM approach in the assumption of independence between these two kinds of actions will be developed. In taking the expression of the solution $\mathbf{X}(t)$ given in Eq. (9), there is no need to assume the independence of $\mathbf{F}(\tau)$ and $\mathbf{X}_0$, but the initial condition is generally independent of the forcing term in practice. The expression of the response vector can be rewritten as follows:

$$\mathbf{X}(t) = \bar{\mathbf{X}}_0(t) + \bar{\mathbf{X}}_F(t) \tag{15}$$

$\mathbf{X}_0$ and $\bar{\mathbf{X}}_F(t)$ being two independent random processes given by:

$$\begin{aligned} \bar{\mathbf{X}}_0(t) &= \boldsymbol{\Theta}(t)\mathbf{X}_0; \\ \bar{\mathbf{X}}_F(t) &= \int_0^t \boldsymbol{\Theta}(t-\tau)\mathbf{B}\mathbf{F}(\tau)d\tau \end{aligned} \tag{16}$$

Each of these processes can be easily characterized through the evaluation of their CFs. Indeed, $\bar{\mathbf{X}}_0(t)$ is a linear combination of the elements of $\bar{\mathbf{X}}_F(t)$ and, thus, its CF can be evaluated in the form given in Eq. (6). The process $\bar{\mathbf{X}}_F(t)$ is the time integral of the process $\mathbf{Z}(\tau) = \boldsymbol{\Theta}(t-\tau)\mathbf{B}\mathbf{F}(\tau)$. In the previous section, a procedure giving numerically the CF has been given through Eq. (13). Lastly, the response CF $M_{\mathbf{X}(t)}(\omega, t)$ can be obtained by considering an important property of the CF of the sum of two independent random processes (Soong, 1973), such that:

$$M_{\mathbf{X}(t)}(\omega, t) = M_{\bar{\mathbf{X}}_0(t)}(\omega, t) M_{\bar{\mathbf{X}}_F(t)}(\omega, t) \tag{17}$$

Even in this case, the response PDF is obtained by the Fourier anti-transform of the above CF.

4. Numerical example

Consider an n-degree of freedom structure, subject to a time-dependent vector-force $\mathbf{F}(t)$. The equation of motion is:

$$\begin{aligned} \mathbf{M}\ddot{\mathbf{U}}(t) + \mathbf{C}\dot{\mathbf{U}}(t) + \mathbf{K}\mathbf{U}(t) &= \mathbf{F}(t) \\ \mathbf{U}(0) &= \mathbf{U}_0 \end{aligned} \tag{18}$$

where $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{K}$ are the mass, damping and stiffness matrices, respectively; $\mathbf{U}(t)$ is the n-vector of the response displacements, $\mathbf{U}_0$ defines the initial conditions of the response and $\mathbf{F}(t)$ represents the excitation vector.

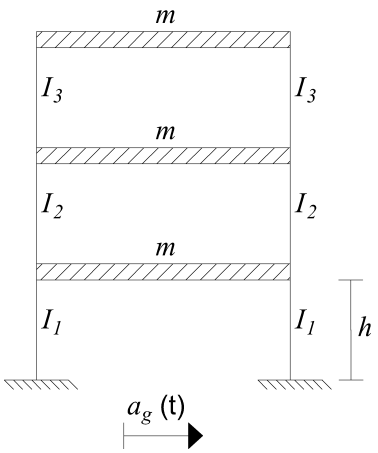


Figure 2.: Shear-type plane system.

By introducing the state variables vector $\mathbf{X}^T(t) = \left(\mathbf{U}^T(t) \;\; \dot{\mathbf{U}}^T(t)\right)$, Eq. (18) can be converted into the following first-order differential system:

$$\begin{aligned} \dot{\mathbf{X}}(t) &= \mathbf{D}\mathbf{X}(t) + \mathbf{v}\mathbf{F}(t); \\ \mathbf{X}(0) &= \mathbf{X}_0 \end{aligned} \tag{19}$$

where:

$$\mathbf{D} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{pmatrix}; \quad \mathbf{v} = \begin{pmatrix} \mathbf{0} \\ -\mathbf{M}^{-1} \end{pmatrix}; \quad \mathbf{X}_0 = \begin{pmatrix} \mathbf{U}(0) \\ \dot{\mathbf{U}}(0) \end{pmatrix} \tag{20}$$

The response $\mathbf{X}(t)$ is evaluated taking into account the Duhamel integral, that is:

$$\mathbf{X}(t) = \mathbf{\Theta}(t)\mathbf{X}_0 + \int_0^t \mathbf{\Theta}(t-\tau)\mathbf{v}\mathbf{F}(\tau)d\tau \tag{21}$$

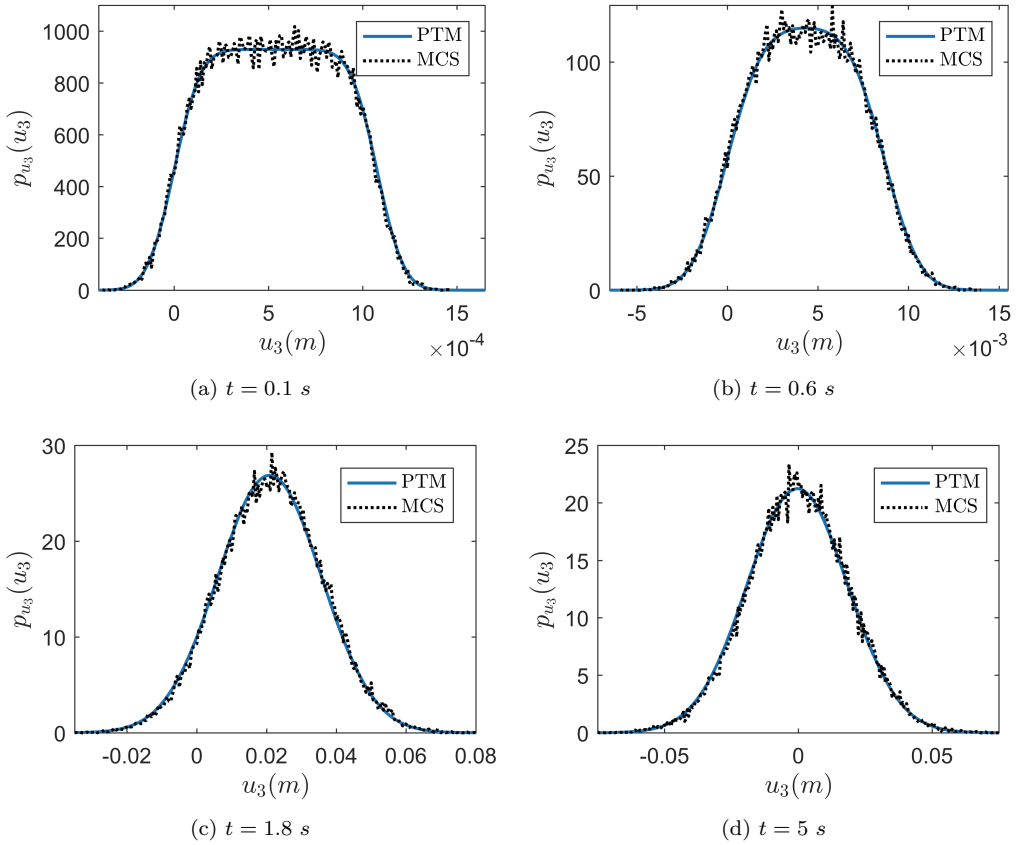
$\mathbf{\Theta}(t)$ being the transition (or fundamental) matrix corresponding to the dynamical system; it is expressed as:

$$\mathbf{\Theta}(t) = \exp(\mathbf{D}t) = \sum_{n=0}^{\infty} \frac{1}{n!}\mathbf{D}^n t^n. \tag{22}$$

Now, the shear-type plane system represented in Figure 2 is considered. The Young's modulus value is $E = 31 \times 10^9 \; N/m^2$; all the columns have the same length $h = 3.2 \; m$, while the moments of inertia of each column are $I_1 = I_2 = I_3 = 0.0054 \; m^4$; for each floor, a mass $m = 50,000/g \; kg$ is assumed. The system is forced by a the zero-mean Gaussian stationary ground acceleration $a_g(t)$ defined by its one-side power Clough–Penzien spectra density having the expression:

$$S_a(\omega) = \frac{\omega_r^4 + 4\xi_r^2\omega_r^2\omega^2}{\left(\omega_r^2 - \omega^2\right)^2 + 4\xi_r^2\omega_r^2\omega^2} \frac{\omega^4}{\left(\omega_p^2 - \omega^2\right)^2 + 4\xi_p^2\omega_p^2\omega^2} \frac{0.141\xi_r^2 a_{g0}^2}{\omega_r\sqrt{1 + r\xi_r^2}} \tag{23}$$

where the following filtering coefficients have been considered: $\omega_p = 2.0$, $\omega_r = 19$, $\xi_p = 0.6$ and $a_{g0} = 0.25g$. Therefore, the vector of the external excitations $\mathbf{F}(t)$ of Eq. $(19_{(1)})$ is expressed as $\mathbf{M}\boldsymbol{\tau}a_g(t)$, where $\boldsymbol{\tau}$ is the structural incidence vector. While, the vector of the initial displacement condition in Eq. $(19_{(2)})$, $\mathbf{X}_0$, is assumed as a random vector described by random variables uniformly distributed with $\sigma_{x_0} = 0.15$.



(a) $t = 0.1\ s$ (b) $t = 0.6\ s$

(c) $t = 1.8\ s$ (d) $t = 5\ s$

Figure 3.: Displacement PDF evaluated for four different instants. PTM (continuous line); MCS (dashed line)

The probabilistic definition of the displacement in correspondence of the j-th DOF at a fixed time t_k, that is $X^{(j)}(t_k)$, has been obtained by the application of EPTM. Figure 3, shows the PDF of the displacement u_3, in correspondence of 4 time instants. From the inspection of this figure, it is possible to appreciate that for the first instants, the random characteristics of the PDF outputs are significantly influenced by the random initial conditions, then the responses PDF are gradually filtered and depend only on the random characteristics of the excitation. These results are considered together with those coming out from an MCS characterized by $50,000$ samples. Overall, the analysis of the results evidences the goodness of the proposed stochastic procedure.

5. Conclusions

In this work, a stochastic procedure for the dynamic analyses of systems characterized by uncertainties in the external actions and in their initial conditions has been presented. Based on PTM, the EPTM combines the main properties of the mean square random calculus with the principle of conservation of probability. The proposed procedure allows to preserve the system's probability in time step by step and gives the random response directly in terms of CFs; then the response PDF of the system can be evaluated easily by the inverse Fourier transform. The application of the EPTM for the stochastic dynamic analysis of a shear-type system has been confirmed the goodness of the results in terms of evolutive PDF.

References

Adhikari, S. Uncertainty propagation in linear systems: an exact solution using random matrix theory. *48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, p. 1957, 2007

Bucher, C.G. and Schueller, G.I. Non-Gaussian response of linear system. *Structural Dynamics*, 103–127, 1991

Calatayud, J., Cortés, J.C., and Jornet, M. On the approximation of the probability density function of the randomized heat equation. *arXiv preprint arXiv:1802.04188*, 2018

Calatayud, J., Cortés, J.C., and Jornet, M. The damped pendulum random differential equation: A comprehensive stochastic analysis via the computation of the probability density function. *Physica A*, **512**, 261–279, 2018

Conte, J.P. and Peng, B.F. An explicit closed-form solution for linear systems subjected to nonstationary random excitation. *Probabilistic Engineering Mechanics*, **11**(1), 37–50, 1996

Di Paola, M. and Falsone, G. Non-Linear Oscillators Under Parametric and External Poisson Pulses. *Journal of Nonlinear Dynamics*, **5**(3), 337–352, 1994

Di Paola, M. and Falsone, G. Higher Order Statistics of the Response of MDOF Linear Systems Excited by Linearly Parametric White Noises and External Excitations. *Probabilistic Engineering Mechanics*, **12**(3), 179–188, 1997

Di Paola, M. and Falsone, G. Higher Order Statistics of the Response of MDOF Systems under Polynomials of Filtered Normal White Noises. *Probabilistic Engineering Mechanics*, **12**(3), 189–196, 1997

Di Paola, M., Falsone, G. and Pirrotta, A. Stochastic Response Analysis of Nonlinear Systems under Gaussian Inputs. *Probabilistic Engineering Mechanics*, **7**(1), 15–21, 1992

Falsone, G. Cumulants and Correlations for Linear Systems under Non-Stationary Delta-Correlated Processes. *Probabilistic Engineering Mechanics*, **9**(3), 157–165, 1994

Falsone, G. An extension of the Kazakov relationship for non-Gaussian random variables and its use in the non-linear stochastic dynamics. *Probabilistic Engineering Mechanics*, **20**(1), 45–56, 2005

Falsone, G. and Laudani, R. A probability transformation method (PTM) for the dynamic stochastic response of structures with non-Gaussian excitations. *Engineering Computations*, **35**(5), 1978–1997, 2018

Falsone, G. and Laudani, R. Matching the principal deformation mode method with the probability transformation method for the analysis of uncertain systems. *Engineering Computation*, 118(7): 395–410, 2019.

Falsone, G. and Laudani, R. Multi-time probability density functions of the dynamic non-Gaussian response of structures. *Structural Engineering and Mechanics*, **76**(5): 631, 2020

Falsone, G. and Settineri, D. Explicit solutions for the response probability density function of linear systems subjected to random static loads. *Probabilistic Engineering Mechanics*, 33: 86–94, 2013a.

Falsone, G. and Settineri, D. Explicit solutions for the response probability density function of nonlinear transformations of static random inputs. *Probabilistic Engineering Mechanics*, 33: 79–85, 2013b.

Gioffrè, M. and Gusella, V. Numerical analysis of structural systems subjected to non-Gaussian random fields. *Meccanica*, **37**(1-2), 115–128, 2012

Hussein, A. and Selim, M.M. Limit analysis of stochastic structures in the framework of the Probability Density Evolution Method. *The European Physical Journal Plus*, **30**(12), 249, 2015

Kalogeris, I. and Papadopoulos, V. Solution of the stochastic generalized shallow-water wave equation using RVT technique. *Engineering Structures*, **160**, 304–313, 2018

Li, J. (2016), Probability density evolution method: Background, significance and recent developments. *Probabilistic Engineering Mechanics*, **44**, 111–117, 2016

Li, J. and Chen, J. Stochastic dynamics of structures. *John Wiley & Sons*, 2009

Lin, Y.K. Probabilistic theory of structural dynamics. *McGraw Hill*, New York, NY, USA, 1967

Liu, Z. and Liu, Z. Random function representation of stationary stochastic vector processes for probability density evolution analysis of wind-induced structures. *Mechanical Systems and Signal Processing*, **106**, 511–525, 2018

Lutes, L.D. and Sarkani, S. Random vibration analysis of structural and mechanical system. *Butterworth-Heinemann*, Oxford, United Kingdom, 2004

Mamis, K.I., Athanassoulis G.A. and Kapelonis Z.G. A systematic path to non-Markovian dynamics: new response probability density function evolution equations under Gaussian coloured noise excitation. *Proceedings of the Royal Society A*, **475**, 2226, 2019

Makarios, T. K. Evaluating the effective spectral seismic amplification factor on a probabilistic basis. *Structural Engineering and Mechanics*, **42**(1), 121–129, 2012

Mazelsky, B. Extension of power spectral methods of generalized harmonic analysis to determine non-Gaussian probability functions of random input disturbances and output responses linear systems. *Aeronaut*, **21**(3), 145–153, 1954

Meimaris, A.T., Kougioumtzoglou, I.A. and Pantelous A.A. A closed form approximation and error quantification for the response transition probability density function of a class of stochastic differential equation. *Probabilistic Engineering Mechanics*, **54**, 87–94, 2019

Morikawa, H. and Kameda, H. Stochastic interpolation of earthquake ground motions under spectral uncertainties. *Structural Engineering and Mechanics*, **5**(56), 839–851, 1997

Roberts, J.B. and Spanos, P.D. Random vibration and statistical linearization. *Wiley*, Chichester, NY, USA, 1991

Soong, TT. Random Differential Equations in Science and Engineering. *Academic Press, New York*, 1973

Wu, W.F. and Lin, Y.K. Cumulant-neglect closure for non-linear oscillators under random parametric and external excitations. *International Journal of Non- Linear Mechanics*, **15**(5), 910–916, 1984

AI based bridge health assessment

G. Marasco[1)], B. Chiaia[1)], and G. Ventura[1)]
1)Department of Structural, Geotechnical and Building Engineering, Politecnico di Torino, Torino 10129, Italy, giulia.marasco@polito.it, bernardino.chiaia@polito.it, giulio.ventura@polito.it

Abstract: Starting from the data extracted from a long-term monitoring system installed on a steel bridge, it has been possible to outline the undamaged behaviour of the structure. The structure under monitoring is a steel suspended arch bridge of long span that has been instrumented with several types of sensors, e.g. triaxial accelerometers, load cells and environmental sensors. The records of the measurements during the first period of structural life and the lack of construction problems ensure the good respect of the structural nominal conditions.

The accelerometric data stored during this period have been used to extrapolate the dynamic characteristics of the bridge: natural frequencies, damping ratios and modal shapes. The use of a specific stochastic subspace technique (SSI-UPCX), allowed to obtain not only the modal parameters but also their uncertainty. In this way, the range of variation of modal parameters, e.g. affected by environmental factors, has been calculated and a minimum and maximum threshold for each parameter has been determined. Consequently, the assessment and control of structural health is updated and linked to these ranges of variation.

In addition, a promising modern approach to tackle the problem is the use of machine learning techniques within the broad field of AI. After the selection/reduction of the parameters that better represent the data, signal detection has been used and the obtained outcomes compared. In the light of both the above approaches, albeit in a different way, it is possible to create a model of the normal operating condition of the structure and consider the deviations from the pattern as an anomaly.
The work represents a first step and a benchmark for the wider damage and ageing identification problem to figure out which method is the most appropriate and effective for this specific case of structural assessment, in terms of effort and accuracy.

Keywords: Structural health monitoring; Modal parameters; Machine learning; Damage detection.

1. Introduction

One of the most crucial problems in the civil engineering field is the identification of structural damage. Accurate results, reduction of false and missing alarms, and detection time (DT) are fundamental aspects of this problem. They have a great weight both on economic and on safety point of view.
Similar to the human body, that during the individual life can be subjected to diseases, the facilities are complex systems that during their useful life are liable to degradation. In both cases, to tackle in a successful way the problem, an interdisciplinary approach is required.

In the infrastructural field, this issue is particular evident for large scale structure subjected to increased loads and different environmental conditions, as the bridges. The presence of some phenomena, e.g. ageing, fatigue and corrosion, implies a change of material properties, structural characteristics, boundary conditions,

© 2021 *by G. Marasco, B. Chiaia, and G. Ventura.*

and element connectivity. This entails the need of checking the structure and developing strategies to optimize the maintenance operations by allocating correctly the available resources. Indeed, early warnings of anomalies are essential for facilitating effective remedying actions. Strategies within this aim belong to structural health monitoring (*SHM*) field and foster a correct development of prognostics and health management (*PHM*).

SHM includes three fundamental approaches. The first consists in continuous monitoring, that allows for verifying alteration of the global structural response. The second includes non- destructive periodic testing (*NDT*), like acoustic emission, ultrasonic, magnetic fields, and radiography that enable a structural local analysis. The last is realized by spot tests. The common target of the three approaches is to highlight early-stage damage and to reduce its spread to avoid the achievement of a critical damage level leading to failure.
Operational modal analysis techniques are one of the most used techniques to address this problem. They belong to the vibration-based methods and allow the extraction of modal structural parameters (natural frequencies, damping ratios and mode shapes) in an inexpensive way and without interfering with the service structural operations. Changes in modal parameters are attributed to damage structural state. For example, a formation of a crack would imply a geometry change with consequent stiffness shifting that could be easily grasped by frequency change. Scour of a bridge pier and loosing of a bolted connection are other types of damage that can be captured by modal parameters alterations.

It is worth pointing out that the data processed to extract the features characterizing the structure, are often related to its normal behaviour, namely to the undamaged structural state. Indeed, the cases in which experimental data are available in damaged cases are very rare (Farrar and Worden, 2012; Diez at al., 2016). On the other hand, the data produced through models (e.g. F.E.M. model), due to high complexity of the systems, to intrinsic inaccuracy of a model, and to the difficulty in defining effective damage cases, are not always reliable for representing the effective structural status. For these reasons, damage diagnosis is often relied on a developed normal condition model.

Unsupervised learning techniques allow to distinguish the damaged from the undamaged structural status avoiding the major problem encountered by machine learning techniques in the field of the *SHM*, consisting in the need of damaged data as supervisor in learning process. These data, in general, are missing or difficult to gather. These techniques are often defined, in the machine learning area, with the term *novelty detection* and their success is linked to the *representativeness* of training dataset (Michau at al., 2019) corresponding to the normal structural conditions. In fact, a deficiency of representativeness will imply a missing distinction between faults and operating conditions (Michau at al., 2018) producing false alarm when new operating conditions, that has not been observed in training period, will classify as faults (Wang at al., 2019). Some approaches are developed to face the low representativeness of many training data sets, due for example to the short observation periods. They are based on the concept of transferring information between systems (units of a fleet) with similar characteristics but that have been under different conditions (Michau and Fink, 2019; Michau at al., 2018). On the other hand, promising approaches have been elaborated in learning the relevant features (Michau at al., 2020). This aspect is essential in the applications where the features *dimensionality*, i.e. the number of parameters monitored and used for damage detection, is very high.
Once used the reference dataset (healthy dataset) to train the chosen learning algorithm, without knowledge about the damaged system conditions (Michau at al., 2020), and designed the decision boundary, an *health indicator (HI)* is frequently used for the damage detection. It consists in the distance between the testing data and the reference data.

One of the criteria to differentiate the novelty detection techniques is the shape of the features distribution in the nominal condition (Farrar and Worden, 2012). For Gaussian distributed normal condition, techniques

that involve the discordance calculation between the test data and the normal condition with subsequent comparison between the calculated value and an alarm threshold are often used. In the multivariate case the discordance is calculated by means of Mahalanobis squared-distance and the threshold with the Monte Carlo method. For non-Gaussian normal condition, auto-associative neural network (AANN) can be a valid method to address the problem. Research in this field has produced a lot of improvement. Among these, (Hu at al., 2017) proposed a new approach defined by the terms “Auto-Associative Extreme Learning Machines”. This last looks promising due to its marked learning ability and its low computational cost.

Further approaches involve the use of clustering techniques as the k-means algorithm (Mehdinia at a., 2017; Bounzenad at al., 2019). They, by means of the creation of clusters based on the similarities between the features (Tryon, 1939), are able to distinguish the healthy from the damaged structural states. The distance between the feature vector of each data and the cluster centroid of nominal condition highlights the variation of the structural behaviour in case a damage event occurs. Larger is this distance, more the damage is extended.

This paper presents an AI application in which the k-means algorithm, within the proposed damage identification flowchart, is used for a steel arch highway bridge.

2. Structure under monitoring and sensor network

The structure under monitoring is a highway steel arch bridge in the northern of Italy. Its span is 250 meters. Vertical steel cables have variable number of strands and link the arch to the inferior way. The arch has a trapezoidal section, while the inferior way is composed by a chain hexagonal beam and transverse cantilevers, located along the bridge axis every 8 meters. A view of the bridge is reported in Figure 1 and more details are reported in (Chiaia at al., 2020a).



Figure 1. View of the suspended arch steel viaduct

The bridge is equipped with a customised monitoring network system that, for its capacity in real-time structural assessment, has been defined as “Active Monitoring System”. In a previous paper (Chiaia at al., 2020b) an in-depth illustration of the system has been presented.

Different types of sensors form the monitoring network. There are high-resolution servoinclinometers, steel surface temperature probes, air temperature and humidity sensors, triaxial accelerometers, differential wind pressure transducers, strain gauges at runway cantilevers, load cells at each suspension cable.

3. Feature extraction: Operational Modal Analysis

The signals must be first preprocessed in order to extract their characterizing features.
To this end, accelerometric signals, covering unevenly a period of about one and a half years, have been analysed. As frequently happens in civil structures, output-only techniques exploiting the environmental excitation have been used to avoid the interruption of the infrastructure service. By means the algorithm DD-SSI-UPCX (Data Driven Stochastic Subspace Identification Extended Unweighted Principal Component), present in the ARTeMIS software (Andersen, 2010), an operational modal analysis has been done to extract the features of interest: natural frequencies, damping ratios, and modal shapes. This particular algorithm utilizes a parametric model to fit the raw data. The fundamental assumptions of the SSI methods are infinite amount of data, linear system, and white noise excitation. The algorithm has good performance also in presence of not very large datasets. The validation of the identified structural modes is done by means of the Stabilisation diagram that through more or less restrictive stabilization criteria is able to avoid different types of errors.

The pick of this specific algorithm is supported by several reasons. Among them there are: its capacity to correctly analyse flexible structures (i.e. with low vibration frequencies), its robustness, its calculation speed, and its ability to provide also the uncertainties for modal parameters. This last characteristic allows to comprehend how reliable the obtained results are or not. As known from the literature (Rainieri and Fabbrocino, 2014), the uncertainties related to damping ratios estimation are greater than those related to the natural frequencies. Consequently, for identifying damage, the natural frequencies will be the utilized characteristics.

4. Detection of anomalous behavior

4.1 TRADITIONAL APPROACH: THRESHOLDS FOR SCATTERING RANGE

Having available the structural response in time in terms of modal parameters, it has been possible to define a range that encompasses the environmental variability. The first six natural frequencies have been reported in Figure 2. In Table I on page 5, the range of variability of each of them has been defined.
The values, maximum and minimum for each range, represent nominal conditions boundaries.
A departure from these values represents an anomalous behaviour.

The simplicity of this approach hides an important limit. In presence of very small damages, namely in the early stage of the development of an anomalous phenomenon, the extracted features could appear within the nominal defined range due to factors that are different by damage effects (Li at al., 2010). Among them, there are environmental factors (temperature, solar radiation, wind velocity, and humidity), operational factors (intensity of traffic flow and potential traffic jam), and errors due to poor data set and limits of processing techniques. Indeed, the probability density function of the damaged and undamaged state could be overlapped. For the investigated bridge, the variation due to EOVs (environmental and/or operational variations) turns out smaller than that shown for other highway bridges case studies (Ko and Ni, 2005) but

still not negligible. To improve the effectiveness of this approach, techniques of elimination/mitigation of EOVs, like regression modelling and machine learning approaches, should be implemented but it is demonstrated that successful results can be obtained only with a robust data normalization. This last is feasible merely in presence of a large volume (Magalhães at al. ,2012) and a high accuracy of experimental data, which are missing in this case study, so it is not possible to reach valid results following these procedures.
The followed strategy to attenuate the variability of the features has been to widen the observation window of the signals to extract the modal parameters. In detail, the observation window considered includes 31 signals. In this way, a more accurate estimate of modal parameters is possible, and many highly uncertain values located near the boundary of the range are deleted. Moreover, very long records already include a temperature variability that allows to smooth the structural behaviour in the examined temperature range.

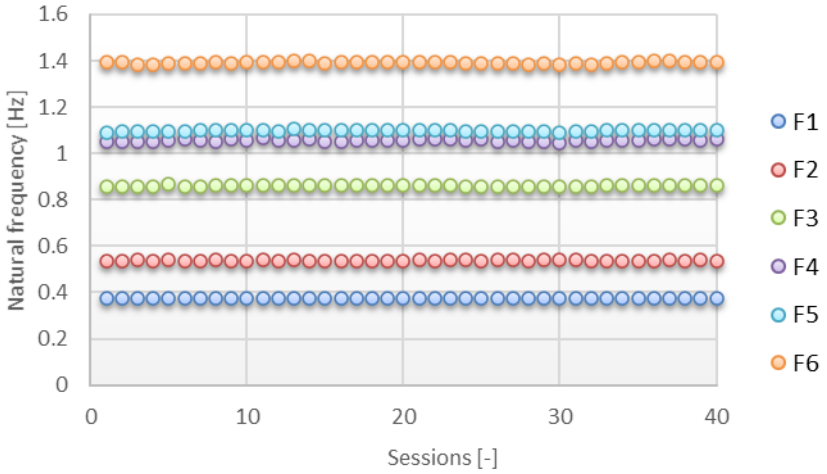


Figure 2. Natural frequency for definition of nominal condition

Table I. Nominal conditions: range of frequency scattering

VIBRATION MODE	**MODE1:** BENDING OUT OF PLANE	**MODE2:** DOUBLE VERTICAL INFLECTION	**MODE3:** SINGLE VERTICAL INFLECTION	**MODE4:** OUT OF PLANE DOUBLE INFLECTION	**MODE5:** TORSIONAL	**MODE6:** TRIPLE VERTICAL INFLECTION
FREQUENCY	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**
F_MIN [HZ]	0.373	0.532	0.856	1.044	1.091	1.383
F_MAX [HZ]	0.3772	0.540	0.866	1.065	1.105	1.400
VAR REL [%]	1.069	1.398	1.167	1.979	1.239	1.238

The percentage of signals not leading to an alarm, for very small damage (weak signals), can be extremely high. As will be shown in the following, more than half of the damaged signals will not exceed the nominal condition boundaries. This means that 50% of the damaged signals would not be considered as such and the process of damage identification would be significantly less sensitive. Analysing signals, belonging to undamaged and damaged structural states, the percentages of false and missing alarm has been calculated. The False (FA) and Missing Alarm (MA) have been expressed, respectively, as:

$$FA = \frac{N_d}{N_{test_ud}} \cdot 100$$

$$MA = \frac{N_{ud}}{N_{test_d}} \cdot 100$$

Where N_d and N_{ud} are the number of data identified as damaged and undamaged, respectively. On the other hand, N_{test_ud} and N_{test_d} are the number of tested undamaged and damaged data. Three potential damaged states have been simulated imposing variations in the acceleration signals.

Many of the most dangerous phenomena for the structural integrity like cracks, foundation settlements, malfunction of a bearing devices, and losses of connections imply change in structural stiffness. A variation in this structural characteristic, occurring locally or in a widespread way, causes a change in the signal frequency. In particular, a stiffness reduction occurs with an ensuing increase of the signal period. Thus, to simulate damaged signals, a delay in the structural response (stretching of time axis) has been imposed.
Analyses carried out on other real bridges (Dilena and Morassi, 2011; Chang, 2016; Magalhães at al., 2012), both with real and simulated damages, have been taken into account to establish the magnitude order of a realistic frequency shift produced by a potential structural damage. Notches, specifically, that can simulate slight damage stemmed from impacts of objects, corrosion or overload, lead to a natural frequency shift of the order on average of 0.1-0.5%.

Initially, the analysed response in natural frequency is that related to the Mode1, Mode2, Mode3 and Mode6, see Table II. These four modes have been considered the most relevant for implementing damage identification strategies.

Table II. Errors for each natural frequency.

DAMAGE LEVELS	**DEGREE OF SEVERITY**	**SIGNAL VARIATION**	**FREQUENCY**	**F1**	**F2**	**F3**	**F6**
			FA[%]	0.0	12.5	0.0	0.0
LD7	LOW	~0.33%	MA[%]	75.0	97.9	62.5	70.8
LD8	MEDIUM-LOW	~0.66%	MA[%]	41.6	75.0	27.1	39.6
LD9	MEDIUM	~1.0%	MA[%]	2.1	35.4	2.1	8.3

A combination criterion, based on the alarm trigger when one of the features is outside the limits of the ranges defined in Table I on page 5, has been used to improve the effectiveness in damage detection.
As can be noted from Table II, the contribution of the frequency of second mode is the worst both in term of false and missing alarm. The second mode gives a contribution to decrease the missing alarm if a combination of the four frequencies is considered, see Table III.

Table III. Errors for combination F1-F2-F3-F6.

Combination F1-F2-F3-F6	**MA [%]**	**FA [%]**	**TE [%]**
LD7	56.3	12.5	69.0
LD8	14.6	12.5	27.1
LD9	0.0	12.5	12.5

In Table IV, the results in terms of missing, false, and total errors have been reported for a combination of F1, F3, and F6. In this case the total errors decrease, due to the lack of false alarms. Nevertheless, the missing error, in particular way for LD8, increases of about thirty per cent.

Table IV. Errors for combination F1-F3-F6.			
Combination F1-F3-F6	**MA [%]**	**FA [%]**	**TE [%]**
LD7	58.3	0.0	58.3
LD8	18.8	0.0	18.8
LD9	0.0	0.0	0.0

For the smallest damage level (LD7), a large part of the monitored frequencies slips back into the nominal condition range. Of course, as can be deduced from the comparison between Table II, III, and IV, the missing alarm percentages decrease if the alarm is emitted when only one of the features is out of the limit values. The reduction of missing alarm percentages is more marked for the damage level LD8.
Concluding, for this structure and for the specific type of damage, the classical approach is able to identify discreetly anomalies produced by a damage level equal and greater than LD8, namely a medium-low damage.

4.2 UNSUPERVISED MACHINE LEARNING: CLUSTERING TECHNIQUE

The traditional approach results deficient for very low damage levels. Nevertheless, it is a complement tool to visual inspection and non-destructive test because it can assess the structural condition even when the damage location is not known a priori and the deteriorated part is not accessible. A specific clustering technique has been used to meet the need to improve the potential structural damages identification. A cluster analysis is applicable when the data have not assigned labels and it consists in researching groups of data, namely clusters, such that the data in a group will be similar (or related) to one another and different from (or unrelated to) the data in other groups. There are three fundamental types of clustering algorithms: K-means and its variants, hierarchical clustering, and density-based clustering. For the problem at hand, K-means algorithm has been chosen and exploited.

4.2.1. *Selection of the features*

Damage-sensitive features are essential to reach a good damage identification. Thus, it appears reasonable to exclude the features that are more sensible to the environmental/operational factors. The contribution of the second, the fourth, and the fifth natural frequency has been considered not relevant for damage identification purpose, as mentioned in subsection 4.1. due to different reasons: the fourth frequency, as underlined by the value in Table I on page 5, has a high relative variation already in the nominal conditions; the fourth and the fifth modes are much closed space modes and therefore are easily confused; finally, Table II on page 6 shows that the second frequency is the least damage-sensitive.

4.2.2. *K-means algorithm: theoretical aspects and SHM perspective*

The flowchart, displayed in Figure 3, describes the steps of the clustering process, that allows to attribute a structural “health/damaged” state to the signals, based on the k-means algorithm.

K-means can be classified as an unsupervised learning algorithm and is a partitional clustering approach. It allows data, represented by a vector of features, to be divided into non-overlapping subsets (clusters). This algorithm makes it possible to visualise similar data in clusters based on a specific metric. Its aim is to create groups of data in the light of the feature vector distances. High homogeneity among the data belonging to the same group and low homogeneity among those belonging to different groups are the criteria to build clusters. Once defined the number of clusters, k centers must be found so to minimize the intra-cluster distance. This algorithm requires considerable calculation effort (Bouzehand, 2019). Several iterative processes have been

proposed to converge rapidly to an optimum. One of them (Lloyd, 1982) is composed by three fundamental following steps:

1. Choice of k point centers, namely k centroids.

2. Assign each data/signal, characterize by its features vector, to the closest cluster centroid.

3. Calculate the new k point centers.

The last two steps are repeated as long as convergence is achieved. At each iteration, the k-means algorithm minimizes the intra-cluster distance (Bouzenad at al., 2019). To assess the performance of this algorithm, a matching matrix can be used. There are three factors that can negatively affect the performance. Indeed, K-means performs not well when clusters are of differing sizes, densities, and non-globular shapes. Often, to get over these limitations many clusters are used. Another drawback of this approach is that a different choice of initial centers can imply various created clusters. To overcome this problem, it is good practice to repeat the algorithm several times. Major details related to the k-means theory can be found in (MacKay and Ma Kay, 2003).

From *SHM* point of view, the classical k-means approach allows to analyse collected data/signals and to distinguish between two clusters. One will correspond to healthy and one to damaged state (in the hypothesis of a single damage source). This can provide a "photograph" of the structural state but prevents from tracking structural changing in time and possible evolution of damage from its early stages before damage reaches a critical size (Bouzenad at al., 2019). To intercept initial damage signs, the outlined process in the flowchart in Figure 3 can be considered. The starting point is the creation of an undamaged state, with the Td training data, represented by one cluster (k=1). The centroid (C) and the maximum distance (Dmax) from it are the benchmarks characterizing this nominal state. Then, new signals are analysed one at a time. For each new signal, the counter c is updated and a distance (d) between the new signal and the nominal centroid is calculated. If this quantity is greater than Dmax, the counter (count) is incremented, by revealing an anomaly. When this value (count) is equal to a persistence number (N), the number of cluster k is imposed equal to 2. The features of the undamaged pattern and the persistence number are used to reduce false alarms. Indeed, if the k-means algorithm with a number of cluster greater than 1 were used every time that a new signal was recorded or when a very small number of signals presented d > Dmax, the number of false alarms would be very high, and the data would most likely appear separated in two groups that would not linked to healthy/damaged state rather to warm/cool state. To track the evolution of the damage level, a check of the optimal number of clusters is done when critical time units are reached. The number of signals contained in a critical time unit is defined by Ccritical. The critical time unit is therefore a control unit that, through a discretized calculation of the kopt, allows a speeding up of the analysis process. On the other hand, it corresponds to a time in which initially at most one level damage can be developed. Indeed, in this process, the hypothesis made at the beginning is that in a first critical time unit, the level of potential damage is one and there is no check of the ability of the algorithm to build a model of the underlying structure in the data with a different value of k that results optimal for the problem at hand. The concept that maximum one level of damage can be present in the first critical unit and that then a verification is required (e.g. for c=2*Ccritical, c=3*Ccritical etc) is due to the fact that for the first critical units the appraisal of some factors like materials properties, environmental impact and load turns out simpler and the hypothesis made results realistic. In conclusion in the developed process, the value of k is equal to:

- 1 if count < N
- 2 if (count >= N and Kopt <2)
- Kopt calculated for n*Ccritical if (count >= N, n*Ccritical < c < n+1 Ccritical, and Kopt >2), with n natural number starting from 2

4.2.3. *Application aspects*

The algorithm implementation has been realised in Matlab. The choice of the k initial cluster centroids, corresponding to the point 1 of the list on page 8, is done by means of k-means++ algorithm, that optimizes the running time and the final solution (Arthur and Vassilvitskii, 2007). The points 2 and 3 of the just mentioned list correspond to so-called *batch updates phase* aims at minimizing the sum of point-to-centroid distances. This phase could not converge to the correct solution and it is sometimes followed by so-called *online updates phase*. For the problem at hand, the *batch updates phase* turns out sufficient and thus it is the only performed phase. Thereby, the process is very fast. Having the possibility to choose among several distance metric, it should be noted that Squared Euclidean distance has been used. Besides, as maximum number of iterations to reach the convergence has been used the default value in Matlab (100). As regard the empty clusters, they have been removed but the algorithm has been set to keep track of their presence. As underlined in the previous subsubsection, it is good practice to repeat the algorithm several times with different initial centers so to maximize the inter-cluster distance. In this case, the argument 'Replicates' has been set equal to 10. Another important aspect that must not be forgotten, is the data normalization (Mohamad and Usman, 2013). This is an important pre-processing task, that scaling data in a specific range is able to ensure the same weight to every attribute. For this specific problem, a Min-Max normalization has been applied. In addition to the aspects just analysed, k-means clustering is also very sensitive to outliers because their presence can skew the right positions of the k point centers (Alamdari at al., 2017). In the present case, the removal of the outliers does not lead to benefits.

4.2.4.1. *Learning of undamaged pattern.* The number of clusters is imposed equal to 1 in the nominal condition. The localization of the normalized centroid and the maximum normalized distance (Dmax) are reported in Table V.

Table V. Normalised centroid and maximum distance from it for nominal conditions

M1	M3	M6	Dmax
0.518	0.420	0.470	0.588

4.2.4.2. *Damaged pattern.* The number of cluster k is updated to the value of two when a number N of signals has a distance from the normal condition centroid higher than the defined threshold (Dmax).

Furthermore, some criteria may be considered to evaluate if there are structural changes, which in case of early-stage damage recognition can indicate grow of negative phenomena in time. They allow to evaluate the optimal number of clusters (Kopt), verifying thus the existence of a potential development of the damage. If it is growing, an increase in the damage level can be deduced. Among the criteria that can be found in the literature there are the Calinski-Harabasz clustering criterion (Calinski and Harabasz, 1974), the Silhouette index (Rousseeuw, 1987; Kaufman and Rousseeuw, 1990), the Davies-Bouldin criterion (Davies, 1979) and the gap statistic criterion (Tibshirani at al., 2001). The Calinski-Harabasz relies upon the sums of squared Euclidean distance between the feature vectors and the centroids of the predicted clusters and the optimal value of k corresponds to its maximum value. The Silhouette index evaluates the difference between inter-cluster distances and intra-clusters distance and the optimum value of k must maximize this index. The

Davies-Bouldin criterion utilizes the intra and the inter-clusters distance as well and, for this criterion, the optimal value of k corresponds to its minimum. Finally, the gap criterion calculates the logarithmic mean of the pairwise distance. In this case the optimum value of k is correlated to the maximization of this criterion. In this way, there will not be the need to choose thresholds that will define damage levels. The algorithm, on the basis of available data, will point out by means of calculated Kopt, a potential worsening of the situation.

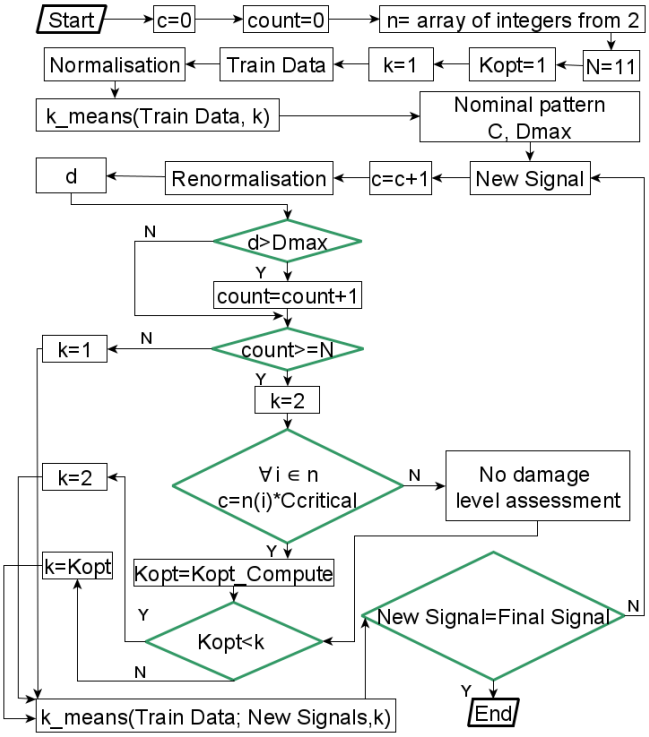


Figure 3. Proposed flowchart for damage identification based on k-means algorithm.

4.2.4.3. *Evaluation of the accuracy: the matching matrix.* The matching matrix has been used to assess the accuracy of the clustering process and to understand if there are improvements with respect to the traditional approach. The comparison between real and predicted clustered signals gives an idea about the capacity of the algorithm in correctly identifying the structural condition. The rows and the column of the matching matrix correspond to the real and to the predicted classes of the signal, respectively.
A combination of test undamaged dataset (Test_UD) and datasets corresponding to the different levels of damage, described in subsection 4.1, are used. The total error (TE) has been defined as:

$$TE\ [\%] = MA[\%] + FA[\%]$$

Where MA and FA correspond to missing and false alarms defined in the aforementioned subsection.

For all the three cases, the number of signals is less than the hypothesized Critical threshold, so the number of used clusters is equal to 2. The following figures 4, 5, and 6 display the division of the data (signals) for each analysed Dataset. Cluster 1 includes the signals belonging to the undamaged condition while Cluster 2 contains the ones of the damaged condition. Of course, both are associated with a centroid.

– Dataset 1: Test_UD and LD7

Table VI. Matching matrix Datasets 1			
Real / Predicted	Healthy	Damaged	TE
Healthy	68.8%	31.2% (FA)	60.3%
Damage	29.1% (MA)	70.1%	

As can be seen in Table VI, the total error is similar (variation of about 3%) to the one obtained by the traditional approach. The error induced by missed alarm is reduced by half at the expense of the error of false alarm. Thus, although the error level is still high, the clustering technique results on the safe side.

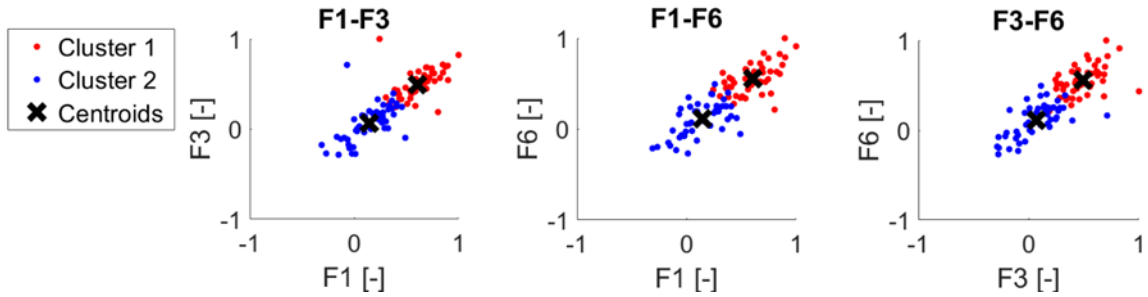


Figure 4. Clustered data (Dataset 1)

– Dataset 2: Test_UD and LD8

Table VII. Matching matrix Datasets 2			
Real / Predicted	Healthy	Damaged	TE
Healthy	77.1%	22.9% (FA)	27.0%
Damage	4.1% (MA)	95.9%	

From the comparison between Table VI and Table VII, it is clear that the total error is drastically decreased in the passage from Dataset 1 to Dataset 2. Compared to the traditional approach, the total error is increased (about 50%) but it is worth to stress that the error due to missed alarm has decreased more than four times. In this case, this approach results more expensive than the traditional approach, due to the fact that about 1/5 of the undamaged signals have been indicated as damaged. On the other hand, the drastic reduction of missing alarm makes this technique on the safe side.

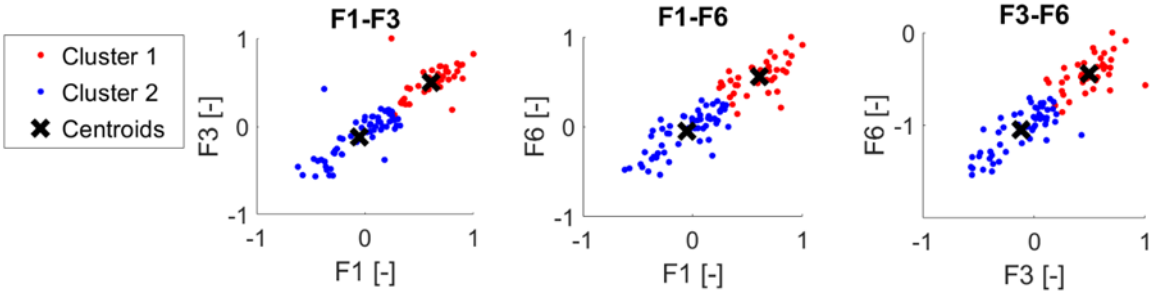


Figure 5 Clustered data (Dataset 2)

– Dataset 3: Test_UD and LD9

Table VIII. Matching matrix Datasets 3			
Real / Predicted	Healthy	Damaged	TE
Healthy	100.0%	0.0% (FA)	2.1%
Damage	2.1% (MA)	97.9%	

In this case, the cluster approach shows a very good performance but slightly lower than that of the traditional approach.

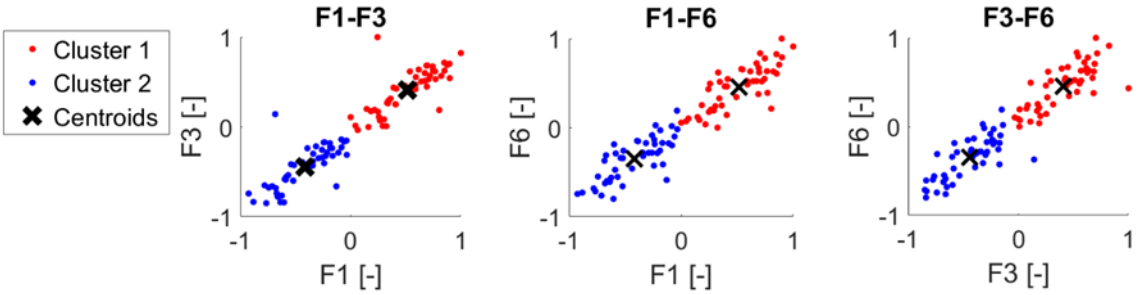


Figure 6 Clustered data (Dataset 3)

Better results in terms of missing alarms have been achieved for all the three damage levels if as number of clusters the maximum value, among the optimal value calculated by means of the three criteria, is used. The smallest damage level LD7 shows the most marked absolute improvement. Indeed, for LD7 the percentage of errors for missing alarms goes down to the value of 6.25%. There is a reduction of about four times compared to the previous case. The obtained total error is less than the previous, showing a value of 52.05%. Of course, the error for false alarms increases. However, this growth is much lower compared to the decrease of the missing alarms errors. From these considerations it can be deduced that, if the structure has been just built and so the first hypothesized damage level is very small (LD7 - LD8), the calculation and the use of kopt would be appropriate also if there are a number of signals less than 2*Ccritical. It could be greater than 2 dues to the presence of environmental effects.

Indeed, a very strong similarity could be present between the undamaged states measured at high temperatures and damaged states measured at low temperatures. On the other hand, in cases in which a greater level of damage is immediately present (LD9) the use of the maximum value of k implies worst performance of the algorithm due to the fact that a small gain in terms of missing alarms and a huge increase of false alarms occurs. Concluding, the k-means algorithm shows promising results in particular for very small damage levels. Its utilization implies a marked decrease of the error due to missing alarm. It produces acceptable outcomes from the safety point of view already for the damage level LD7. It avoids several missing alarms, paying the price of increasing the false alarms. For the early-stage damage levels it is preferable to the traditional approach as it ensures higher safety.

4.2.4.4. *Recognition of increasing damage.* When the critical unit of time is reached, the optimal value of k is calculated. A higher value of k, in the succession of control units, could point out an increase in the level of structural damage. By utilizing the three criteria for the calculation of the optimal value of k, in the analysis of the data covering all the three damage levels (LD7-LD8-LD9), it is possible to verify the capacity of recognition of an increasing damage. The results for the first three time units are displayed in Figure 7, 8 and 9, respectively. A change in the optimum value of k is displayed for all the three units by the Gap criterion and for the last two units by the Calinski-Harabasz criterion.

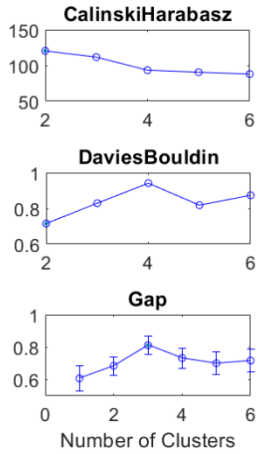


Figure 7. First unit

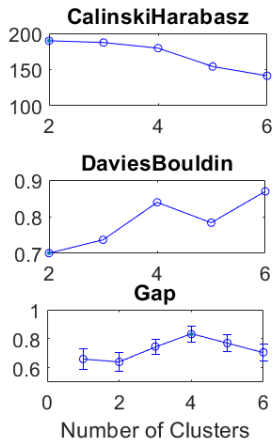


Figure 8. Second unit

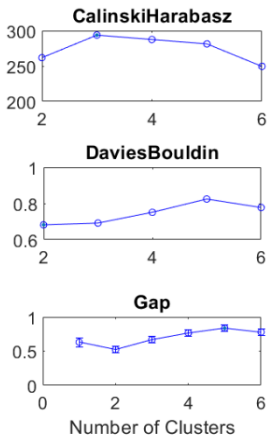


Figure 9. Third unit

5. Conclusions

This paper has illustrated two different approaches to address the bridge health assessment. Both the methods start from the elaboration of the acceleration data by means of the SSI-UPCX method that provides not only modal parameters but also their uncertainty. This additional information allows an increased awareness of the reliability of the modal parameters. The natural frequencies, due to their very low uncertainty, have been selected for the damage identification goal. The analysis aims to compare the performance of the two approaches using a test sample incorporating variability induced by environmental effects.

The main findings, observations, and conclusions stemmed from this survey are summarised as follows:

- The traditional approach shows relevant errors for very small damage levels. Such errors are solely due to the worst part of the total error, namely the missing alarms.
- For the small damage levels analysed, the clustering technique has a better performance. It greatly reduces the error due to missing alarms but increases the one due to false alarms. From the engineers' point of view, this swap turns out undoubtedly positive. It is safer to correct a false positive by means of inspections than not to grasp the presence of damage.
- For very small damage level, an increased computational effort in the calculation of the optimum value of numbers of clusters in the application of the k-means algorithm produces a further improvement as regards missing alarm errors. Thus, this clustering approach is preferable if a high level of safety must be ensured. Another positive aspect of this method is the possibility to observe the evolution of the optimal number of clusters for assessing the increase of damage. Indeed, it is possible to deduce a changing in the data and a potential growth of damage severity from a growing value of optimal k over time.
- A greater uniformity, completeness, and quality of data would imply an improvement of the already promising performance of the clustering algorithm. The role of the persistence number (N) and the optimal number of clusters (Kopt) is decisive in mitigating the expected increase in error, in terms of FA and MA, in critical scenarios for the k-means algorithm (clusters with sizes and densities markedly different).
- In the application of the clustering algorithm, the uncertainty in the choice of the optimum value of clusters due to the not-perfect correspondence among all the values indicated by various criterions is an aspect that must be improved. Certainly this, and the consequent not-perfect correspondence between each optimal value and the real number of data groups, suggest the research of more performant features as a perspective of this work. More effective features would imply a betterment for both the investigated approaches. In particular, with higher quality of the signals, the Spectral Moments (SMs) could be exploited as damage features. In (Alamdari at al., 2017) the high ability of these features, covering the whole frequency range, in detecting subtle differences between normal and distorted signals have been evidenced. Moreover, the use of indicators for modal shapes like MAC and COMAC, could be accounted due to the fact that they are the less sensitive to other factors. Then again, the combination of several parameters could lead to better damage-sensitive features.

References

Alamdari, M. M., T. Rakotoarivelo, and N. L. D. Khoa. A spectral-based clustering for structural health monitoring of the Sydney Harbour Bridge. Mechanical Systems and Signal Processing, 87: 384-400, 2017.

Andersen, P. ARTeMIS Extractor Online Help. Structural Vibration Solutions A/S, 2010.

Arthur, D. and S. Vassilvitskii. K-means++: The Advantages of Careful Seeding. In SODA '07: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, 2007.

Bouzenad, A. E., M. El Mountassir, S. Yaacoubi, F. Dahmene, M. Koabaz, L. Buchheit, and W. Ke. A Semi-Supervised Based K-Means Algorithm for Optimal Guided Waves Structural Health Monitoring: A Case Study. Inventions, 4(1): 17, 2019.

Caliński, T. and J. Harabasz. A dendrite method for cluster analysis. Communications in Statistics-theory and Methods, 3(1): 1-27, 1974.

Chang, K. C. and C.W Kim. Modal-parameter identification and vibration-based damage detection of a damaged steel truss bridge. Engineering Structures, 122: 156-173, 2016.

Chiaia, B., G. Marasco, G. Ventura, and C. Zannini Quirini. Customised active monitoring system for structural control and maintenance optimisation. J Civil Struct Health Monit, 10: 267- 282, 2020a.

Chiaia, B. G. Ventura, C. Zannini Quirini, and G. Marasco. Bridge active monitoring for maintenance and structural safety. In Arêde A, Costa C (eds) Proceedings of ARCH 2019. ARCH 2019. Structural integrity, vol 11. Springer, Cham, 866–873, 2020b.

Davies, D.L. and D.W. Bouldin. A cluster separation measure. IEEE transactions on pattern analysis and machine intelligence, 2: 224-227, 1979.

Diez, A., N. L. D. Khoa, M.M. Alamdari, Y. Wang, F. Chen, and P. Runcie. A clustering approach for structural health monitoring on bridges. Journal of Civil Structural Health Monitoring, 6(3): 429-445, 2016.

Dilena, M. and A. Morassi. Dynamic testing of a damaged bridge. Mechanical Systems and Signal Processing, 25(5): 1485-1507, 2011.

Farrar, C. R. and K. Worden. Structural health monitoring: a machine learning perspective. John Wiley & Sons, 2012.

Hu, Y., T. Palmé, and O. Fink. Fault detection based on signal reconstruction with Auto-Associative Extreme Learning Machines. Engineering applications of artificial intelligence, 57:105-117, 2017.

Kaufman, L. and P.J. Rousseeuw. Finding groups in data: an introduction to cluster analysis. John Wiley & Sons, 2009.

Ko, J. M. and Y. Q. Ni. Technology developments in structural health monitoring of large-scale bridges. Engineering structures, 27(12): 1715-1725, 2005.

Li, H., S.Li, J. Ou, and H. Li. Modal identification of bridges under varying environmental conditions: temperature and wind effects. Structural Control and Health Monitoring 17(5): 495-512, 2010.

Lloyd, S. Least squares quantization in PCM. IEEE Transactions on information theory, 28(2): 129-137,1982.

Mackay, D.J. and D.J. Mac Kay. Information theory, inference and learning algorithms. Cambridge university press, 2003.

Magalhães, F., A. Cunha, and E. Caetano. Vibration based structural health monitoring of an arch bridge: from automated OMA to damage detection. Mechanical Systems and Signal Processing, 28: 212-228, 2012.

Mehdinia, S., Y. Hamishebaharb, M. Bitarafc, A. Kayvan, and M.R. Ghorbani-Tanhad. Structural damage detection using K-means Clus-tering algorithm under Unknown Environmental and Operational Conditions.

Michau, G., T. Palmé, and O. Fink. Fleet PHM for critical systems: bi-level deep learning approach for fault detection. In Proceedings of the European Conference of the PHM Society, 2018.

Michau, G. and O. Fink. Unsupervised Fault Detection in Varying Operating Conditions. In 2019 IEEE International Conference on Prognostics and Health Management (ICPHM), 1-10, IEEE,2019.

Michau, G., Y. Hu, T. Palmé, and O. Fink. Feature learning for fault detection in high-dimensional condition monitoring signals. In Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 2020.

Mohamad, I.B. and D. Usman. Standardization and its effects on K-means clustering algorithm. Research Journal of Applied Sciences, Engineering and Technology, 6(17): 3299-3303, 2013.

Rainieri, C. and G. Fabbrocino, Influence of model order and number of block rows on accuracy and precision of modal parameter estimates in stochastic subspace identification. Int J Lifecycle Perform Eng 1(4):317-334, 2014.

Rousseeuw, P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics, 20: 53-65, 1987.

Tibshirani, R., G. WALTHER, and T. Hastie. Estimating the number of clusters in a data set via the gap statistic. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(2): 411-423, 2001.

Tryon, R.C. Cluster analysis: correlation profile and orthometric analysis for the isolation of unities in mind and personality. Ann Arbor: Edward Brothers, 1939.

Wang, Q., G. Michau, and O. Fink Domain adaptive transfer learning for fault diagnosis. In 2019 Prognostics and System Health Management Conference (PHM-Paris). IEEE, 279-285, 2019.

On constrained distribution-free p-boxes and their propagation

Marco Daub[1)], Stefano Marelli[1)], Bruno Sudret[1)]
[1)] *ETH Zürich, Stefano-Franscini-Platz 5, 8093 Zürich, Switzerland, daub@ibk.baug.ethz.ch, marelli@ibk.baug.ethz.ch, sudret@ethz.ch*

Abstract. Propagating uncertainties through computational models is a key ingredient in uncertainty quantification in engineering. In general, uncertainty can be characterized as aleatoric, caused by variability, and epistemic, caused by lack of knowledge. Imprecise probability theory offers a natural framework to deal with both through, through sets of probability distributions. Among the latter, probability-boxes (p-boxes), which specify upper and lower bounds on admissible cumulative distribution functions (CDFs), are well established in the literature. We hereby introduce a novel class of p-boxes, *constrained distribution-free p-boxes*, that is based on imposing constraints on the admissible distributions (e.g. bound moments, symmetry, derivatives, etc.) on otherwise distribution-free p-boxes. We demonstrate that this class maintains most of the flexibility of classical distribution-free p-boxes, while avoiding most of the non-physical configurations it can be associated with. We also show how constrained distribution-free p-boxes can influence uncertainty bounds in the model predictions, thus improving the quality of the resulting uncertainty estimation.

Keywords: uncertainty quantification, uncertainty propagation, imprecise probabilities, probability-boxes, Monte Carlo simulation

1. Introduction

Over the past decades, *uncertainty quantification (UQ)* for engineering systems has become of great interest, see (Soize, 2017). Popular disciplines of UQ concern, for instance, with the analysis or the selection of a design for reliability or robustness, (Moustapha and Sudret, 2019; Teixeira et al., 2021). Here, computational models that analyze the behavior of mechanical systems together with the increased availability of computational resources have made major contributions: they are used to propagate the uncertainty in input parameters to the output space by performing multiple model evaluations. Usually, the uncertain parameters in the input space are modeled by probabilistic models to account for their variation, i.e., intrinsic randomness. This type of uncertainty is also called *aleatoric uncertainty*, and Monte Carlo simulation (MCS) is usually applied for their propagation. However, there are often situations in engineering in which the actual probabilistic model is unknown. Hence, *epistemic uncertainty*, which is due to lack of knowledge and can be reduced in principle, adds to the aleatoric uncertainty, see (Der Kiureghian and Ditlevsen, 2009).

In general, aleatoric and epistemic uncertainties can be treated in a unified framework called *imprecise probability theory*, see (Augustin et al., 2014; Walley, 1991). For the modeling of uncertain parameters, imprecise probability models consider sets of probability distributions. A popular representative hereof are *(distribution-free) p-boxes*, which provide upper and lower bounds for

© 2021 *by M. Daub, S. Marelli, B. Sudret.*

the unknown cumulative distribution functions (CDF) to define such a set. Because of this simple construction and their intuitive visualization, p-box are especially useful in engineering, (Ferson et al., 2003). The more information about the unknown CDF of the uncertain parameter is available, the tighter the bounding CDFs can be chosen. This yields a unique CDF in the limit case. The bounding CDFs are usually constructed based on given information. In general, it is distinguished between construction methods that are based on incomplete distribution properties, or a given dataset with samples of the uncertain parameters (see (Ferson et al., 2003) for an overview). Unfortunately, the use of these methods may lead to wide bounds and CDFs enclosed by the p-box that do not match the given information, see (Beer et al., 2013). For example, the p-box covering all CDFs of random variables with specific mean and variance values, constructed by the Chebychev inequality, as done in (Oberguggenberger and Fellin, 2008), also covers CDFs of random variables with different mean and variance values. This motivates to constrain the feasible CDFs within the p-box bounds. A possible restriction would be to limit the considerations to a specific CDF family, which yields a so-called *parametric p-box*. However, reasons for this restriction might be lacking and thus, the epistemic uncertainty of the uncertain parameters is not represented well enough. Nevertheless, a remedy can be found by using either generalized distribution families for parametric p-boxes, see (Daub et al., 2021b), or by constraining the feasible CDFs of a distribution-free p-box. In this paper, the latter idea is considered, and *constrained distribution-free p-boxes* are proposed.

This paper is organized as follows: first, the basics of p-boxes are explained, and constrained distribution-free p-boxes are derived from distribution-free p-boxes. Then, both p-box types are discretized for numerics, and a pragmatic propagation method is presented. This is applied to the two-dimensional Rosenbrock function as a computational model and the results for various constrained distribution-free p-boxes are discussed.

2. P-boxes

2.1. From distribution-free to constrained distribution-free p-boxes

In order to describe variables facing aleatoric uncertainty, random variables are used in probability theory. Here, a *real-valued random variable* X is a measurable function which is defined on a probability space and maps from its sample space to $\mathbb{R}$. Conventionally, X is characterized by a *cumulative distribution function (CDF)* F_X, which states the probability that X takes on a value less than or equal to $x \in \mathbb{R}$, i.e.,

$$F_X : \mathbb{R} \to [0,1],\ F_X(x) = P(X \leq x). \tag{1}$$

A CDF has the property that it is non-decreasing, right-continuous, and converges to 0 for $x \to -\infty$ and 1 for $x \to \infty$. Its derivative $f_X = \frac{\mathrm{d}}{\mathrm{d}x} F_X$ is called *probability density function (PDF)*. Hence, the function value of F_X at $x \in \mathbb{R}$ can be also expressed as

$$F_X(x) = \int_{\mathbb{R}} I_x(x') f_X(x') \,\mathrm{d}x'. \tag{2}$$

where I_x is an indicator function with $I_x(x') = 1$ for $x' \leq x$ and $I_x(x') = 0$ for $x' > x$, $x' \in \mathbb{R}$.

In the context of imprecise probabilities, the actual CDF of X faces epistemic uncertainty, i.e., is unknown due to lack of knowledge, and only sets in which it is contained can be provided. One approach to define such a set of possible CDFs for X is the modeling as a *probability-box*, or short *p-box*. A p-box is defined as a pair of an upper CDF $\overline{F}_X$ and a lower CDF $\underline{F}_X$, see (Ferson et al., 2003). This results in the set

$$[\overline{F}_X, \underline{F}_X] = \{F_X \in \mathcal{F} \mid \underline{F}_X(x) \leq F_X(x) \leq \overline{F}_X(x),\ x \in \mathbb{R}\} \tag{3}$$

for the unknown CDF of X, where $\mathcal{F}$ is the set of all CDFs on $\mathbb{R}$. As $\mathcal{F}$, and thus $[\underline{F}_X, \overline{F}_X]$, allows for arbitrary distribution types, $[\overline{F}_X, \underline{F}_X]$ is also called a *distribution-free p-box*.

If the feasible CDFs can be limited to a specific CDF family $\mathcal{F}_\Theta$ with parameters θ_i, $i = 1, \ldots, n_\theta$, collected in $\boldsymbol{\theta} \in \mathbb{R}^{n_\theta}$, a *parametric p-box*, also called *distributional p-box* is obtained. Here, bounds are put on the parameters θ_i instead of directly on the CDFs $F_X(\cdot, \boldsymbol{\theta})$. This is usually done in the form of lower bounds $\underline{\theta}_i$ and upper bounds $\overline{\theta}_i$, yielding intervals $[\underline{\theta}_i, \overline{\theta}_i]$ for each parameter θ_i, $i = 1, \ldots, n_\theta$. The envelope of a parametric p-box is a distribution-free p-box defined by the bounding CDFs

$$\overline{F}_X(x) = \max\{F_X(x, \theta) \mid \underline{\theta}_i \leq \theta_i \leq \overline{\theta}_i,\ i = 1, \ldots, n_\theta\}, \tag{4}$$

$$\underline{F}_X(x) = \min\{F_X(x, \theta) \mid \underline{\theta}_i \leq \theta_i \leq \overline{\theta}_i,\ i = 1, \ldots, n_\theta\} \tag{5}$$

for $x \in \mathbb{R}$. Note that the parametric p-box is usually a proper subset of its envelope, i.e., they are not equal.

As opposed to focusing on a specific distribution family for using only selected CDFs of a distribution-free p-box, also the feasible CDFs of a distribution-free p-box can be constrained in a non-parametric way. This leads to a *constrained distribution-free p-box*, which is proposed in the following. The idea of this p-box type is to narrow the feasible CDFs by putting further constraints on the p-box than the fundamental p-box constraint in Equation (3). For example, upper and lower bounds can be put on the moments of X, like the mean $\mu = E(X)$ and the variance $\sigma^2 = \mathrm{Var}(X)$, or on the values of the PDF f_X and its further derivatives. In addition, the distributional shape can be constrained for, e.g., symmetry or unimodality. Thus, a set of feasible CDF $\mathcal{F}' \subseteq \mathcal{F}$ that only comprises CDFs fulfilling these constraints can be used to define a constrained distribution-free p-box for the unknown CDF of X via

$$[\overline{F}_X, \underline{F}_X]' = [\overline{F}_X, \underline{F}_X] \cap \mathcal{F}'. \tag{6}$$

Basically, this idea is already included in (Beer et al., 2013), in which a p-box is represented as a quintuple $\langle \overline{F}_X, \underline{F}_X, [\underline{\mu}, \overline{\mu}], [\underline{\sigma}, \overline{\sigma}], \mathcal{F}' \rangle$. Here, the bounds on the mean and the standard deviation, given by $[\underline{\mu}, \overline{\mu}]$ and $[\underline{\sigma}, \overline{\sigma}]$, are separated from further restrictions on the CDFs, given by $\mathcal{F}' \subseteq \mathcal{F}$. However, this remains rather a theoretical construct and appropriate techniques on how to handle such constraints for computations are lacking. To counteract this, a discretization scheme for CDFs that allows the integration of constraints is deployed in the following. First, distribution-free p-boxes with no constraints are considered again.

2.2. Numerics with distribution-free p-boxes

In this paper, the discretization is done with respect to the real space $\mathbb{R}$. Here, a simple approach is to put an equidistant grid on an interval $[\underline{x}, \overline{x}] \subseteq \mathbb{R}$. The lower bound $\underline{x}$ is chosen as the largest values such that $\overline{F}_X(\underline{x}) = 0$ holds and the upper bound $\overline{x}$ as the smallest value such that $\underline{F}_X(\overline{x}) = 1$ holds. In case of an unbounded support, these bounds are chosen application-specific so that the approximations $\overline{F}_X(\underline{x}) \approx 0$ and $\underline{F}_X(\overline{x}) \approx 1$ can be accepted. Let $N \in \mathbb{N}$ be the number of grid points inside the interval $[\underline{x}, \overline{x}]$, then the grid is defined as

$$[\underline{x}, \overline{x}]_h = \{x_h^j \in \mathbb{R} \mid x_h^j = \underline{x} + jh,\ j \in \{0, \dots, N+1\}\} \tag{7}$$

with $h = \frac{\overline{x}-\underline{x}}{N+1}$. Accordingly, the values of the CDF F_X at the grid points x_h^j are denoted by $F_{X,h}^j = F_X(x_h^j) \in [0,1]$ for $j = 0, \dots, N+1$. In order to guarantee that F_X is a proper CDF which is an element of the p-box $[\overline{F}_X, \underline{F}_X]$, the following inequalities must be fulfilled for $F_{X,h}^j$:

$$F_{X,h}^j \ \leq\ F_{X,h}^{j+1}, \tag{8}$$

$$F_{X,h}^j \ \leq\ \overline{F}_X(x_h^j), \tag{9}$$

$$F_{X,h}^j \ \geq\ \underline{F}_X(x_h^j), \tag{10}$$

$j = 1, \dots, N$. Moreover, $F_{X,h}^0 = F_X(\underline{x}) = 0$ and $F_{X,h}^{N+1} = F_X(\overline{x}) = 1$ hold or are least assumed for $i = 0, N+1$. The unknown $F_{X,h}^j$, $j = 1, \dots, N$ are collected in $\boldsymbol{F}_{X,h} = (F_{X,h}^1, \dots, F_{X,h}^N) \in [0,1]^N$. As the inequalities (8)-(10) are all linear, they can be stated as

$$\boldsymbol{A}\boldsymbol{F}_{X,h} \leq \boldsymbol{b}, \tag{11}$$

for $\boldsymbol{F}_{X,h} \in [0,1]^N$, where $\boldsymbol{A} \in \mathbb{R}^{(3N+1)\times N}$ and $\boldsymbol{b} \in \mathbb{R}^{3N+1}$. The set of all $\boldsymbol{F}_{X,h}$ which fulfill the system of linear inequalities (11) is denoted by $\mathcal{S}_{\boldsymbol{F}_{X,h}}$, i.e.,

$$\mathcal{S}_{\boldsymbol{F}_{X,h}} = \{\boldsymbol{F}_{X,h} \in [0,1]^N \mid \boldsymbol{A}\boldsymbol{F}_{X,h} \leq \boldsymbol{b}\}. \tag{12}$$

For numerical computations, often the values of F_X at non-grid points are also required. To account for this with a simple approach, linear interpolation is used in this paper, i.e.,

$$F_X(x) = F_{X,h}^j + \frac{F_{X,h}^{j+1} - F_{X,h}^j}{x_h^{j+1} - x_h^j}(x - x_h^j) \tag{13}$$

for $x \in (x_h^j, x_h^{j+1})$. It ensures that F_X is non-decreasing when it fulfills Equation (8). The set of all piecewise linear CDFs defined by Equation (13) for which $\boldsymbol{F}_{X,h} \in \mathcal{S}_{\boldsymbol{F}_{X,h}}$ holds is denoted by $[\overline{F}_X, \underline{F}_X]_h$. For $N \to \infty$, i.e., $h \to 0$, $F_X \in [\overline{F}_X, \underline{F}_X]_h$ is capable to approximate an arbitrary CDF of the p-box $[\overline{F}_X, \underline{F}_X]$ in $(\underline{x}, \overline{x})$. However, by using Equation (8), $F_X \in [\overline{F}_X, \underline{F}_X]_h$ can violate the p-box bounds if $\overline{F}_X$ is not concave or if $\underline{F}_X$ is not convex. This is accepted here, as the maximum error decreases with growing N. Moreover, $F_X \in [\overline{F}_X, \underline{F}_X]_h$ is not differentiable at the grid points. If differentiability is desired, a more profound interpolation approach is monotone (cubic) interpolation, see, e.g., (Fritsch and Carlson, 1980). Below, an example that illustrates the use of piecewise linear CDFs for distribution-free p-boxes is provided.

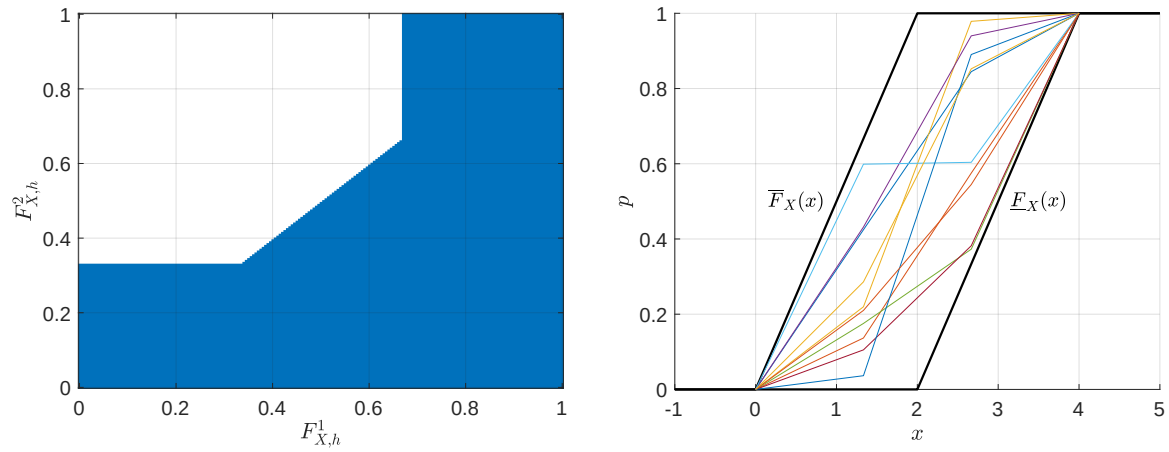


Figure 1. Distribution free p-box $[\overline{F}_X, \underline{F}_X]_h$ for Example 1 and $N = 2$. *Left:* Feasible set $\mathcal{S}_{\boldsymbol{F}_{X,h}}$ (white region) and unfeasible set (blue region) for $\boldsymbol{F}_{X,h}$. *Right:* P-box bounds of $[\overline{F}_X, \underline{F}_X]_h$ with examples of feasible CDFs.

EXAMPLE 1. *Let a distribution-free p-box be defined by the bounding CDFs*

$$\overline{F}_X(x) = \begin{cases} 0 & \text{for } x \leq 0, \\ 0.5x & \text{for } 0 < x \leq 2, \\ 1 & \text{for } x > 2, \end{cases} \qquad \underline{F}_X(x) = \begin{cases} 0 & \text{for } x \leq 2, \\ 0.5x - 1 & \text{for } 2 < x \leq 4, \\ 1 & \text{for } x > 4 \end{cases} \tag{14}$$

for $x \in \mathbb{R}$, which belong to the uniform distribution family. For the purpose of visualization, a discretization of $[\underline{x}, \overline{x}] = [0, 4]$ is done by only $N = 2$ inner grid points, first, and the feasible CDFs of $[\overline{F}_X, \underline{F}_X]$ are represented by piecewise linear CDFs $F_X \in [\overline{F}_X, \underline{F}_X]_h$ defined via $\boldsymbol{F}_{X,h} = (F^1_{X,h}, F^2_{X,h})$. CDF examples of $[\overline{F}_X, \underline{F}_X]_h$ are shown in Figure 1 together with their feasible set $\mathcal{S}_{\boldsymbol{F}_{X,h}}$. Then, the same is done for $N = 100$ to illustrate the full capability of the approach, and CDF examples of $[\overline{F}_X, \underline{F}_X]_h$ are shown in Figure 2.

2.3. Numerics with constrained distribution-free p-boxes

The approach for distribution-free p-boxes can be extended for constraint distribution-free p-boxes by introducing, in addition to the inequalities (8)-(10), inequalities accounting for specific constraints. Exemplarily, inequalities to yield upper and lower bounds on the PDF and the mean, and symmetry around $\frac{1}{2}(\underline{x} + \overline{x})$ for the piecewise linear CDFs as defined in Equation (13) are listed in the following:

– *Bounds on the PDF*: As a piecewise linear CDF with $\boldsymbol{F}_{X,h} \in \mathcal{S}_{\boldsymbol{F}_{X,h}}$ is not differentiable at the grid-points in a strong sense, only the derivative values in $x \in (x_h^j, x_h^{j+1})$, $j = 0, \ldots, N$, are considered. There are a lower bound $\underline{f} \in \mathbb{R}$ and an upper bound $\overline{f} \in \mathbb{R}$ on the PDF if the

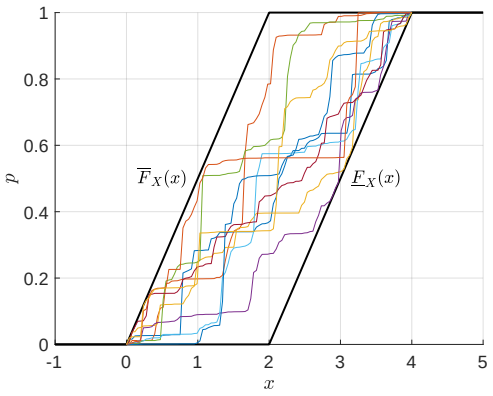


Figure 2. P-box bounds of the distribution free p-box $[\overline{F}_X, \underline{F}_X]_h$ with examples of feasible CDFs for Example 1 and $N = 100$.

inequalities

$$\underline{f} \leq \frac{F_{X,h}^{j+1} - F_{X,h}^{j}}{h} \leq \overline{f} \tag{15}$$

are fulfilled for $j = 0, \ldots, N$.

– *Bounds on the mean*: There are a lower bound $\underline{\mu} \in \mathbb{R}$ and an upper bound $\overline{\mu} \in \mathbb{R}$ on the mean for a piecewise linear CDF with $\boldsymbol{F}_{X,h} \in \mathcal{S}_{\boldsymbol{F}_{X,h}}$ if the inequalities

$$\underline{\mu} \leq \sum_{j=0}^{N} \left(F_{X,h}^{j+1} - F_{X,h}^{j}\right) \left(x_h^j + \frac{h}{2}\right) \leq \overline{\mu} \tag{16}$$

are fulfilled.

– *Symmetry around* $\frac{1}{2}(\underline{x} + \overline{x})$: A piecewise linear CDF with $\boldsymbol{F}_{X,h} \in \mathcal{S}_{\boldsymbol{F}_{X,h}}$ is symmetric around $\frac{1}{2}(\underline{x} + \overline{x})$ if the equations

$$F_{X,h}^{j} = 1 - F_{X,h}^{N+1-i} \tag{17}$$

are fulfilled for $i = 1, \ldots, \frac{N}{2}$ and even N, and for $i = 1, \ldots, \frac{N+1}{2}$ and odd N.

Similar to the inequalities (8)-(10), the inequalities (15) and (16) are linear. Hence, they can be integrated into the system of linear inequalities (11), yielding another system of linear inequalities

$$\boldsymbol{A}'\boldsymbol{F}_{X,h} \leq \boldsymbol{b}', \tag{18}$$

for $\boldsymbol{F}_{X,h} \in [0,1]^N$, where $\boldsymbol{A} \in \mathbb{R}^{N' \times N}$, $\boldsymbol{b} \in \mathbb{R}^{N'}$, and $N' \geq 3N + 1$. Here, the set of all $\boldsymbol{F}_{X,h}$ which fulfill the system of linear inequalities (18) is denoted by $\mathcal{S}'_{\boldsymbol{F}_{X,h}}$, i.e.,

$$\mathcal{S}'_{\boldsymbol{F}_{X,h}} = \{\boldsymbol{F}_{X,h} \in [0,1]^N \mid \boldsymbol{A}'\boldsymbol{F}_{X,h} \leq \boldsymbol{b}'\}, \tag{19}$$

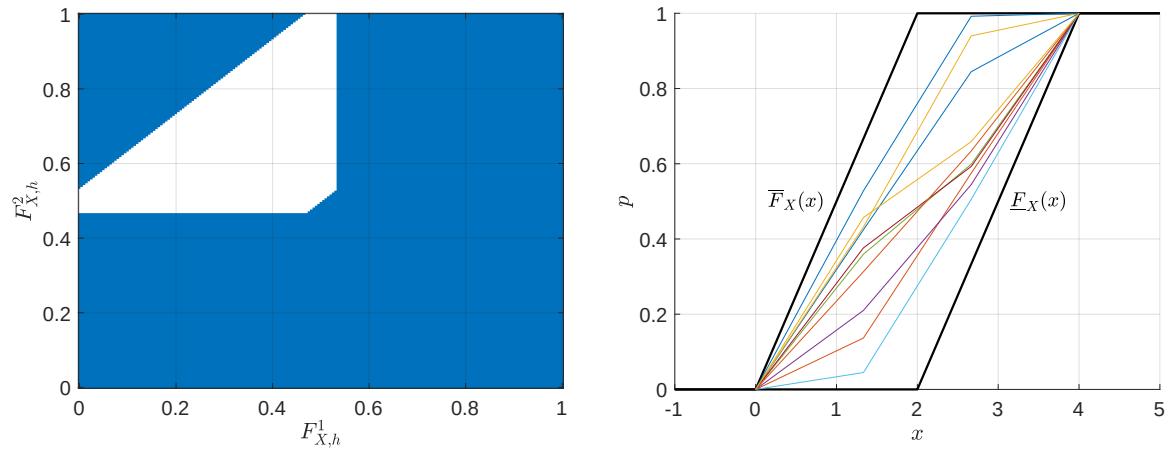


Figure 3. Constrained distribution-free p-box $[\underline{F}_X, \overline{F}_X]'_h$ for Example 2 and $N = 2$. *Left: Feasible set* $\mathcal{S}'_{\boldsymbol{F}_{X,h}}$ *(white region) and unfeasible set (blue region) for* $\boldsymbol{F}_{X,h}$. *Right: P-box bounds of* $[\underline{F}_X, \overline{F}_X]'_h$ *with examples of feasible CDFs.*

and it holds $\mathcal{S}'_{\boldsymbol{F}_{X,h}} \subseteq \mathcal{S}_{\boldsymbol{F}_{X,h}}$. Note that Equation (17) can be considered either in an additional system of linear equations, as each two inequalities in the system of linear inequalities (18), or by reducing the variables $F^j_{X,h}$, $j = 1, \ldots, N$.

Also note that a piecewise linear CDF with $\boldsymbol{F}_{X,h} \in \mathcal{S}_{\boldsymbol{F}_{X,h}}$ has an intrinsic upper bound on the PDF given by $\overline{f} = \frac{1}{h}$ because of the discretization. Accordingly, further inequalities for specific CDF properties can be defined. The set of all $F_X \in [\overline{F}_X, \underline{F}_X]_h$ with $\boldsymbol{F}_{X,h} \in \mathcal{S}'_{\boldsymbol{F}_{X,h}}$ is denoted by $[\overline{F}_X, \underline{F}_X]'_h \subseteq [\overline{F}_X, \underline{F}_X]_h$. Subsequently, the bounding CDFs of Example 1 are used to visualize constrained distribution-free p-boxes with piecewise linear CDFs.

EXAMPLE 2. *Given the bounding CDFs of Example 1, defined by Equation (14). In this example, a constrained distribution-free p-box* $[\underline{F}_X, \overline{F}_X]'$ *that accounts for an upper bound of the PDF with* $\overline{f} = 0.4$ *is considered. Again, the discretization of* $[\underline{x}, \overline{x}] = [0, 4]$ *is done by only* $N = 2$ *inner grid points, first, and the feasible CDFs of* $[\underline{F}_X, \overline{F}_X]'$ *are represented by linear CDFs* $F_X \in [\overline{F}_X, \underline{F}_X]'_h$ *defined via* $\boldsymbol{F}_{X,h} = (F^1_{X,h}, F^2_{X,h})$. *CDF examples of* $[\overline{F}_X, \underline{F}_X]'_h$ *are shown in Figure 3 together with the feasible set* $\mathcal{S}'_{\boldsymbol{F}_{X,h}}$. *Then, the same is done for* $N = 100$ *to illustrate the full capability of the approach, and CDF examples of* $[\overline{F}_X, \underline{F}_X]'_h$ *are shown in Figure 4. In comparison with Example 1, it can be seen that CDFs with large PDF values, i.e. derivative values, are excluded from* $[\underline{F}_X, \overline{F}_X]'_h$.

2.4. Propagating p-boxes

Next, it is considered how constrained distribution-free p-boxes can be propagated through a computational model $\mathcal{M}$ that maps from the n_x-dimensional input space $\mathbb{R}^{n_x}$ to the one-dimensional output space $\mathbb{R}$, i.e.,

$$\mathcal{M} : \mathbb{R}^{n_x} \to \mathbb{R}, \ (x_1, \ldots, x_{n_x}) = \boldsymbol{x} \mapsto y = \mathcal{M}(\boldsymbol{x}). \tag{20}$$

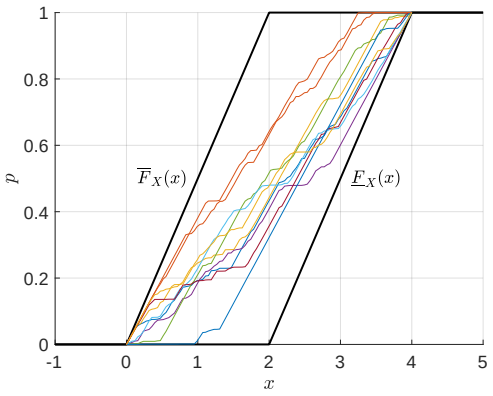


Figure 4. P-box bounds of the distribution free p-box $[\overline{F}_X, \underline{F}_X]'_h$ with examples of feasible CDFs for Example 1 and $N = 100$.

For uncertainty in the input variables x_i, represented by the random variables X_i with marginal CDFs F_{X_i}, $i = 1, \ldots, n_x$, there is uncertainty in the output variable y, too. Hence, it is $Y = \mathcal{M}(\boldsymbol{X})$, where $\boldsymbol{X}$ is the n_x-dimensional random vector comprising the real-valued random variables X_i, $i = 1, \ldots, n_x$, and Y is the real-valued random variable representing the uncertainty in the output variable. The CDF F_Y of Y depends on the joint CDF $F_{\boldsymbol{X}}$, or rather the joint PDF $f_{\boldsymbol{X}}$, of the random variables X_i, $j = 1, \ldots, n_x$, i.e.,

$$F_Y(y) = P(\mathcal{M}(X) \leq y) = \int_{\mathbb{R}} I_y(\mathcal{M}(\boldsymbol{x})) f_{\boldsymbol{X}}(\boldsymbol{x}) \, \mathrm{d}x, \tag{21}$$

where Equation (2) was used. As equation (21) is in general hard to solve, numerical techniques like *Monte Carlo simulation (MCS)* can be used, see (Kalos, 2008) for an overview. Here, realizations $\boldsymbol{x}^{(j)}$, $j = 1, \ldots, M$, of a independent random sample from the CDF $F_{\boldsymbol{X}}$ of length M are required. Then, $F_Y(y)$ can be estimated as

$$F_Y(y) \approx \frac{1}{M} \sum_{j=1}^{M} I_y(\mathcal{M}(\boldsymbol{x}^{(j)})) \tag{22}$$

for $y \in \mathbb{R}$. In general, the entries of $\boldsymbol{x}^{(j)}$, denoted by $x_i^{(j)}$, $i = 1, \ldots, n_x$, can be generated by *inverse transform sampling* for $j = 1, \ldots, M$, (Kalos, 2008). For reasons of simplicity, statistical independence between the random variables X_i, $i = 1, \ldots, n_x$ is assumed in this paper. Then, the joint PDF can be computed as the product of the marginal PDFs, and realizations $u_i^{(j)}$ of an independent random sample from the standard uniform distribution of length $n_x M$ can be used to obtain

$$x_i^{(j)} = F_{X_i}^{-1}(u_i^{(j)}), \tag{23}$$

$i = 1, \ldots, n_x$, $j = 1, \ldots, M$. Note that statistical dependence could be also considered using an *isoprobabilistic transform*, see, e.g., (Torre et al., 2019).

In the context of p-boxes with unknown marginal CDFs $F_{X_i} \in [\overline{F}_{X_i}, \underline{F}_{X_i}]$, $i = 1, \ldots, n_x$, the CDF F_Y is an element of a set of feasible CDFs for Y. This set can be grasped by its envelope $[\overline{F}_Y, \underline{F}_Y]$, which is a distribution-free p-box. Then, the challenge in propagating the p-boxes from the input to the output space becomes the computation of the bounding CDFs $\overline{F}_Y$ and $\underline{F}_Y$. In general, there are various methods to compute these bounds numerically if parametric or distribution-free p-boxes are considered, see (Faes et al., 2020) for an overview. However, because of the CDF constraints for constrained distribution-free p-boxes $[\overline{F}_{X_i}, \underline{F}_{X_i}]_h$, $i = 1, \ldots, n_x$, they are often not transferable. In the following, an intuitive approach, based on MCS and the presented discretization with linear interpolation, is proposed.

For constrained distribution-free p-boxes $[\overline{F}_{X_i}, \underline{F}_{X_i}]'_h$, $i = 1, \ldots, n_x$, the dependency of F_Y reduces to the function values of the marginal CDFs F_{X_i}, $i = 1, \ldots, n_x$, at the grid points. Without loss of generality, it is assumed that they are all represented by N-dimensional tuples $\boldsymbol{F}_{X_i,h} \in \mathcal{S}'_{\boldsymbol{F}_{X_i,h}}$, $i = 1, \ldots, n_x$. Thus, these N-dimensional tuples can be summarized in an $n_x N$-dimensional tuple $\boldsymbol{F}_{\boldsymbol{X},h} = (\boldsymbol{F}_{X_1,h}, \ldots, \boldsymbol{F}_{X_{nx},h}) \in \mathcal{S}'_{\boldsymbol{F}_{\boldsymbol{X},h}}$ with $\mathcal{S}'_{\boldsymbol{F}_{\boldsymbol{X},h}} = \mathcal{S}'_{\boldsymbol{F}_{X_1,h}} \times, \cdots \times \mathcal{S}'_{\boldsymbol{F}_{X_{nx},h}}$. Then, the dependency of F_Y on $\boldsymbol{F}_{\boldsymbol{X},h}$, can be expressed as $F_Y(\cdot, \boldsymbol{F}_{\boldsymbol{X},h})$ and the bounding CDFs $\overline{F}_Y$ and $\underline{F}_Y$ can be determined by

$$\overline{F}_Y(y) = \max_{\boldsymbol{F}_{\boldsymbol{X},h} \in \mathcal{S}'_{\boldsymbol{F}_{\boldsymbol{X},h}}} F_Y(y, \boldsymbol{F}_{\boldsymbol{X},h}), \tag{24}$$

$$\underline{F}_Y(y) = \min_{\boldsymbol{F}_{\boldsymbol{X},h} \in \mathcal{S}'_{\boldsymbol{F}_{\boldsymbol{X},h}}} F_Y(y, \boldsymbol{F}_{\boldsymbol{X},h}) \tag{25}$$

for $y \in \mathbb{R}$. Computationally, these optimization problems can be treated using MCS, i.e., Equation (22), which results the optimization problem

$$\begin{aligned} &\text{maximize}_{\boldsymbol{F}_{\boldsymbol{X},h}} \quad \frac{1}{M} \sum_{j=1}^{M} I_y \left(\mathcal{M} \left((F_{X_1}^{-1}(u_1^{(j)}, \boldsymbol{F}_{X_1,h}), \ldots, F_{X_{n_x}}^{-1}(u_{n_x}^{(j)}, \boldsymbol{F}_{X_{n_x},h})) \right) \right) \\ &\text{subject to} \qquad \boldsymbol{A}' \boldsymbol{F}_{\boldsymbol{X},h} \leq \boldsymbol{b}' \end{aligned} \tag{26}$$

for estimating $\overline{F}_Y(y)$, and the same optimization problem where "maximize" is replaced by "minimize" for estimating $\underline{F}_Y(y)$, $y \in \mathbb{R}$. Note that the quality of these estimations increases with both the number of grid points N and the number of random samples M.

Because of its non-smooth objective function and type of optimization constraints, problem (26) can be solved by generic population-based metaheuristic optimization, i.e., *evolutionary algorithms (EA)*, (De Jong, 2006). In this paper, the *genetic algorithm (GA)*, a type of EA is used, (Goldberg, 1989). Note, however, that the GA does not necessarily find a global optimum and can become computationally inefficient, in particular, when N is large. This motivates using advanced global optimization techniques for future research, like *efficient global optimization (EGO)*, (Jones et al., 1998). In the next section, the two-dimensional Rosenbrock function is considered as an example of a computational model $\mathcal{M}$. As an analytic function, it can be evaluated fast. For computational models for which this is not the case, e.g., where $\mathcal{M}$ is a black-box function, surrogate models like *Kriging* or *(sparse) polynomial chaos expansion (PCE)* can be used to approximate $\mathcal{M}$, see, e.g., (Lüthen et al., 2020; Schöbi, 2017), before solving problem (26).

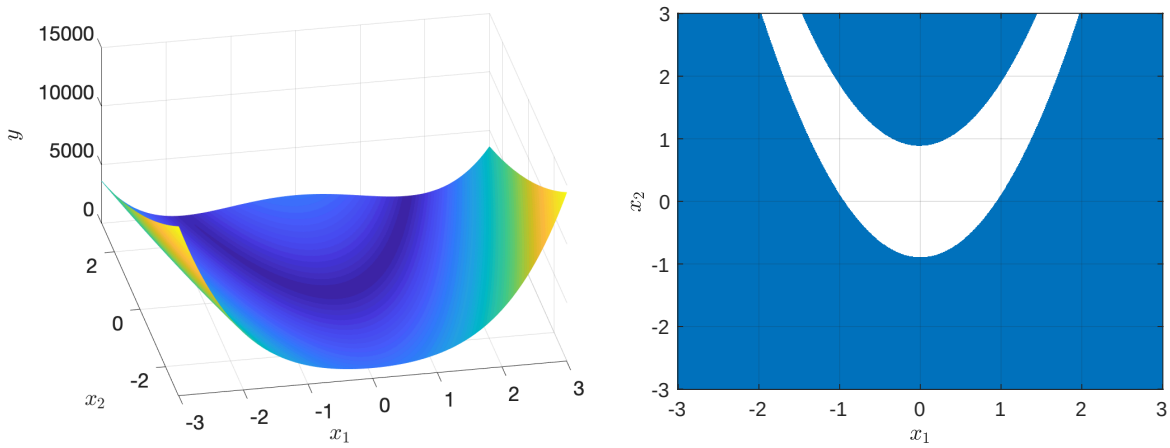


Figure 5. Two-dimensional Rosenbrock function as an example of a computational model $\mathcal{M}$. *Left:* Graph of $\mathcal{M}$ for $-3 \leq x_1, x_2 \leq 3$. *Right:* Set of $\boldsymbol{x}$ with $\mathcal{M}(\boldsymbol{x}) \leq 80$ (white region) and $\mathcal{M}(\boldsymbol{x}) > 80$ (blue region).

3. Application to the Rosenbrock function

3.1. Problem description

The two-dimensional Rosenbrock function, see (Rosenbrock, 1960), is used in (Schöbi, 2017; Schöbi and Sudret, 2017) as an example of a computational model for the propagation of distribution-free p-boxes. It reads

$$\mathcal{M}(\boldsymbol{x}) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \tag{27}$$

for $\boldsymbol{x} \in \mathbb{R}^2$ and is visualized in Figure 5. The two-dimensional Rosenbrock function has a global minimum at $\boldsymbol{x} = (1, 1)$ with $\mathcal{M}(\boldsymbol{x}) = 0$ and a codomain of $[0, \infty)$.

The variables x_1 and x_2 are modeled as random variables X_1 and X_2, whose feasible CDFs are considered for different cases of constraint distribution-free p-boxes. The CDFs forming the envelope of a parametric p-box from a normal distribution with mean value $\mu \in [-0.5, 0.5]$ and standard deviation value $\sigma \in [0.7, 1]$ are considered as bounding CDFs, i.e.,

$$\overline{F}_{X_i}(x_i) = \max_{\mu \in [-0.5,0.5], \sigma \in [0.7,1]} F_{\mathcal{N}}(x_i, \mu, \sigma) = \begin{cases} F_{\mathcal{N}}(x_i, -0.5, 1) & \text{for } x \leq -0.5, \\ F_{\mathcal{N}}(x_i, -0.5, 0.7) & \text{for } x > -0.5, \end{cases} \tag{28}$$

$$\underline{F}_{X_i}(x_i) = \min_{\mu \in [-0.5,0.5], \sigma \in [0.7,1]} F_{\mathcal{N}}(x_i, \mu, \sigma) = \begin{cases} F_{\mathcal{N}}(x_i, 0.5, 0.7) & \text{for } x \leq 0.5, \\ F_{\mathcal{N}}(x_i, 0.5, 1) & \text{for } x > 0.5, \end{cases} \tag{29}$$

$i = 1, 2$, where $F_{\mathcal{N}}(\cdot, \mu, \sigma)$ is the CDF of a normal distribution with mean μ and standard deviation σ, see Figures 6 and 7. Note that these bounding CDFs also correspond to the ones considered in (Schöbi, 2017; Schöbi and Sudret, 2017).

Here, the following cases of constraints are considered for both F_{X_1} and F_{X_2}:

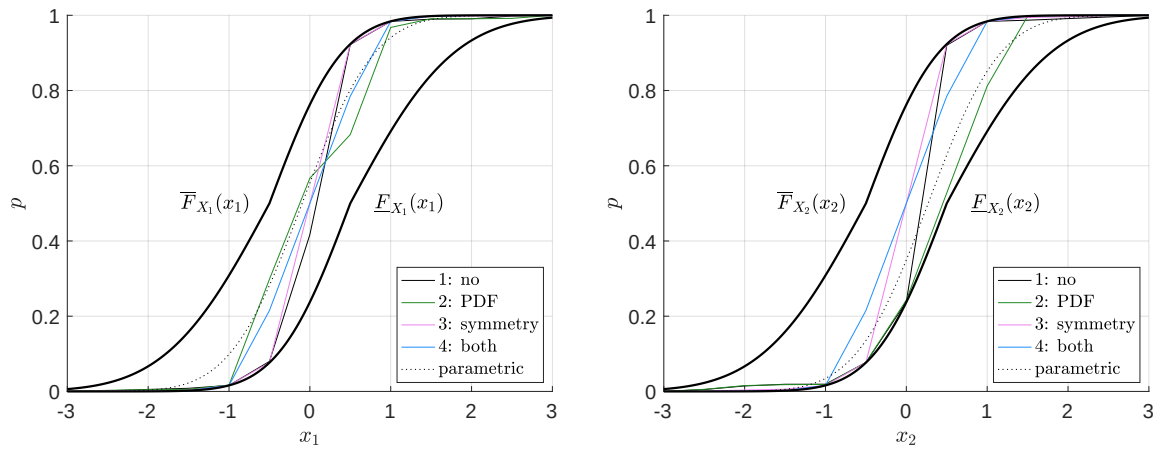


Figure 6. CDFs responsible for $\overline{F}_Y(80)$ of the constrained distribution-free p-boxes (cases 1-4) compared to the one of the parametric p-box for the Rosenbrock example. *Left:* F_{X_1}. *Right:* F_{X_2}.

1. *No additional constraints*: This case corresponds to the case of distribution-free p-box, where only the fundamental p-box constraints, i.e., inequalities (8)-(10), are considered.

2. *Bounds on the PDF*: An upper bound on the PDF is considered with

$$\overline{f} = \max_{x\in\mathbb{R},\mu\in[-0.5,0.5],\sigma\in[0.7,1]} f_{\mathcal{N}}(x,\mu,\sigma) \approx 0.57,$$

   i.e., inequality (15). Here, $f_{\mathcal{N}}(\cdot,\mu,\sigma)$ is the PDF of a normal distribution with mean μ and standard deviation σ. This type of constraint distribution-free p-box is a superset of the parametric p-box from the normal distribution with $\mu \in [-0.5, 0.5]$, $\sigma \in [0.7, 1]$.

3. *Symmetry around* 0: A symmetry constraint around $x_1 = x_2 = 0$ is considered, i.e., Equation (17). Hence, it follows $\mu_1 = \mu_2 = 0$, and this type of a constrained distribution-free p-box is no longer a superset of the parametric p-box with $\mu \in [-0.5, 0.5]$.

4. *Both bounds on the PDF and symmetry around* 0: The bounds on the PDF and the symmetry constraint around $x_1 = x_2 = 0$, as defined above, are considered.

Subsequently, the bounding CDFs $\overline{F}_Y$ and $\underline{F}_Y$ are computed for these cases of constrained distribution-free p-boxes using the proposed computational approach and compared to the ones of the parametric p-box from the normal distribution with $\mu \in [-0.5, 0.5]$, $\sigma \in [0.7, 1]$.

3.2. Results for $\overline{F}_Y$ and $\underline{F}_Y$

For discretizing the constraint distribution-free p-boxes, the grid $[-3, 3]_{h=0.5}$ with $N = 11$ inner grid points and $M = 10^4$ samples for the MCS are used in this paper. The results for $\overline{F}_Y(y)$ and $\underline{F}_Y(y)$, $y \in [0, 1{,}000]$, are visualized in Figure 8. Here, the case $y = 80$ is investigated in more detail.

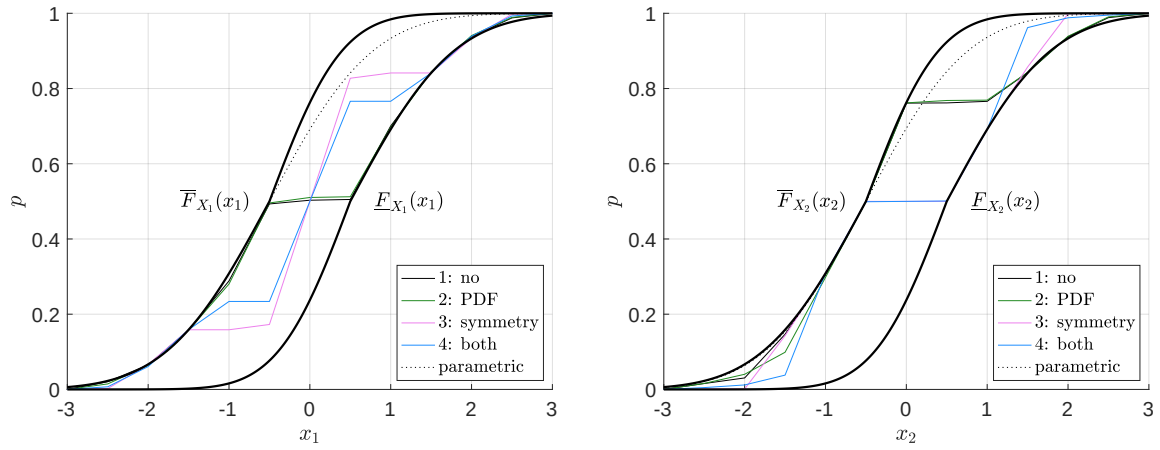


Figure 7. CCDFs responsible for $\underline{F}_Y(80)$ of the constrained distribution-free p-boxes (cases 1-4) compared to the one of the parametric p-box for the Rosenbrock example. *Left:* F_{X_1}. *Right:* F_{X_2}.

Table I. Values of $\overline{F}_Y(80)$ and $\underline{F}_Y(80)$ for the constrained distribution-free p-boxes (cases 1-4) compared to the ones of the parametric p-box for the Rosenbrock example.

Case	1 (no)	2 (PDF)	3 (symmetry)	4 (both)	parametric
$\overline{F}_Y(80)$	0.90	0.76	0.88	0.76	0.69
$\underline{F}_Y(80)$	0.16	0.16	0.21	0.22	0.33

The CDFs responsible for $\overline{F}_Y(80)$ are depicted in Figure 6 and the CDFs responsible for $\underline{F}_Y(80)$ in Figure 7. Furthermore, the numerical values of $\overline{F}_Y(80)$ and $\underline{F}_Y(80)$ are shown in Table I.

In Figure 8, it can be seen that using constrained distribution-free p-boxes (cases 2-4) compared to the distribution-free p-box with no constraints (case 1) in the input space can lead to tighter p-box bounds in the output space. For case 1, $\overline{F}_Y(y)$, $y \in [0, 200]$, raises fast, from 0 to almost 1. In contrast, $\underline{F}_Y(y)$ raises much slower and is below 0.7 for $y \in [0, 1{,}000]$. The results for cases 2-4 follow this overall trend, although in a different manner: While the upper bounding CDF $\overline{F}_Y(y)$, $y \in [0, 1{,}000]$, for case 3 follows the one for case 1 closely, the ones for cases 2 and 4 raise, with a similar rate, slower until $y \in [0, 140]$. For $y > 140$, the upper bounding CDF for case 2 tends faster to 1 than the one for case 4. The lower bounding CDFs for cases 1-4 behave similarly to their upper bounding CDFs for $y \in [0, 60]$. Then for $y > 60$, this effect reverses and $\underline{F}_Y(y)$ for case 2 follows the one for case 1 closer, than the ones for case 3 and 4. Moreover, the bounding CDFs $\overline{F}_Y(y)$ and $\underline{F}_Y(y)$ for the parametric p-box are enclosed between the bounding CDFs for all cases, $y \in [0, 1{,}000]$.

All observations for constrained distribution-free p-boxes can be explained by examining the constraints of cases 2-4 in comparison to case 1 for a specific $y \in [0, 1{,}000]$. Exemplary, this is done

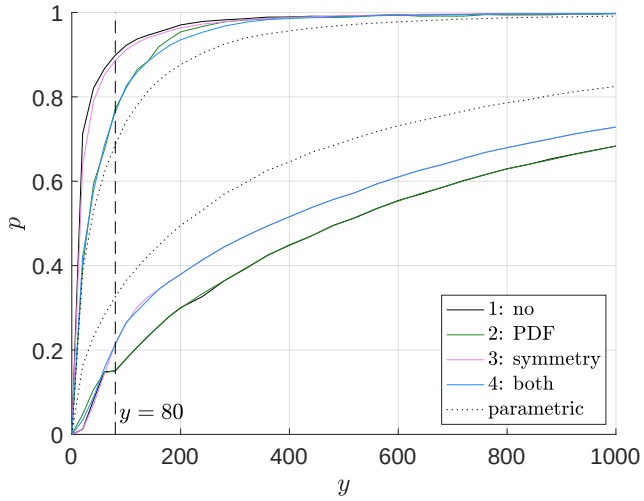


Figure 8. Graph of $\overline{F}_Y$ and $\underline{F}_Y$ of the constrained distribution-free p-boxes (cases 1-4) compared to the one of the parametric p-box for the Rosenbrock example.

for $y = 80$. Here, the CDFs of case 1 are very steep at $x_1 = 0$ and $x_2 = 0$ for $\overline{F}_Y(80)$. This cannot be achieved by cases 2 and 4 due to the bounds on the PDF. For $\underline{F}_Y(80)$, the CDFs of case 1 are moderately steep at both $x_1 \in \{-1, 1\}$ and $x_2 \in \{-1, 1\}$ with a larger focus on -1 for x_2. This cannot be achieved by cases 3 and 4 due to the symmetry around $x_2 = 0$.

Besides the errors in the estimation of $\overline{F}_Y(y)$ and $\underline{F}_Y(y)$, $y \in [0, 1{,}000]$, due to N and M, a further error can occur when the GA does not find the global optimum. In order to increase the computational efficiency for determining $\overline{F}_Y$ and $\underline{F}_Y$, this can be combined with regression methods, compare (Schöbi, 2017; Schöbi and Sudret, 2017). Furthermore, the optimization results of $\overline{F}_Y(y)$ and $\underline{F}_Y(y)$ can be improved by using advanced global optimization strategies, as discussed above. An approach for selecting the grid point non-equidistantly could also be beneficial to further improve the optimization results.

4. Conclusion

In this paper, the propagation of uncertainty modeled as probability-boxes (p-boxes) through a computational model is considered. Traditionally, p-boxes, which are defined as upper and lower bounds on cumulative distribution functions (CDFs), consider either only a specific distribution family or all CDFs between the bounds as feasible. The former are referred to as parametric p-boxes and the latter as distribution-free p-boxes. Because parametric p-boxes might lead to bounds in the

output space that are too tight, and distribution-free p-boxes might lead to bounds that are too wide, constrained distribution-free p-boxes are proposed here. The idea of this new p-box type is based on distribution-free p-boxes plus a narrowing of the feasible CDFs. This is achieved by putting constraints on the CDFs, e.g., bounds on their derivatives, bounds their moments, or restrictions on their distributional shape. In order to treat constrained distribution-free p-boxes numerically, an approach, which is based on a discretization of the input space and on linear interpolation for CDFs defined at the resulting grid points, is presented. Then, the discretized p-boxes are propagated to p-boxes in the output space using Monte Carlo simulation and optimization. The effects of considering constraint distribution-free p-boxes are demonstrated using a two-dimensional analytical example. They influence the bounds of the p-box in the output space depending on the type of constraints, resulting in tighter bounds than the ones from distribution-free p-boxes but also wider than the ones from parametric p-boxes if the constraints allow. Hence, constrained distribution-free p-boxes can enhance uncertainty quantification for practical applications whenever more information than the information of bounds on the CDFs is available, but a limitation to a specific distribution is too restrictive.

For the application of constrained distribution-free p-boxes in more complex problems, it could be helpful to improve the numerical propagation by adjustment concerning, e.g., the selection of grid points, the interpolation scheme, or the optimization algorithm. Pointers on how this could be achieved are given throughout this paper.

Acknowledgements

Marco Daub acknowledges the support of the German Academic Exchange Service (DAAD) with a postdoc fellowship.

References

Augustin, T., F. P. A. Coolen, G. de Cooman, and M. C. M. Troffaes. *Introduction to imprecise probabilities*. Wiley Series in Probability and Statistics. Wiley, 2014.

Beer, M., S. Ferson, and V. Kreinovich. Imprecise probabilities in engineering analyses. *Mechanical Systems and Signal Processing*, 37(1-2):4–29, 2013.

Daub, M., S. Marelli, and B. Sudret. Uncertainty propagation with constrained distribution-free vs. generalized parametric p-boxes for engineering applications. Technical report, ETH Zürich, Zurich, Switzerland, 2021.

De Jong, K. A. *Evolutionary computation: A unified approach*. MIT Press, Cambridge, MA, United States, 2006.

Der Kiureghian, A. and O. Ditlevsen. Aleatory or epistemic? Does it matter? *Structural Safety*, 31(2):105–112, 2009.

Faes M., M. Daub, and M. Beer. Engineering analysis with imprecise probabilities: a state-of-the-art review on p-boxes. *Proceedings of the 7th Asian-Pacific Symposium on Structural Reliability and its Applications (APSSRA2020)*, Tokyo, Japan, 2020.

Ferson, S., V. Kreinovich, L. Ginzburg, D. S. Myers, and K. Sentz. Constructing probability boxes and Dempster-Shafer structures. Technical report, Sandia National Laboratories, Albuquerque and Livermore, United States, 2003.

Fritsch, F. N. and R. E. Carlson. Monotone piecewise cubic interpolation. *SIAM Journal on Numerical Analysis*, 17(2):238–246, 1980.

Goldberg, D. E. . *Genetic algorithms in search, optimization and machine learning.* Addison-Wesley Longman Publishing Co., Inc., Boston, MA, United States, 1989.

Jones, D. R., M. Schonlau and W. J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13:455—492, 1998.

Kalos, M. H. and P. A. Whitlock. *Monte Carlo methods*. Wiley-VCH, Berlin, 2008.

Lüthen, N., S. Marelli, and B. Sudret. Sparse polynomial chaos expansions: Literature survey and benchmark. *SIAM/ASA Journal on Uncertainty Quantification* (accepted), arXiv:2002.01290.

Moustapha, M., and B. Sudret. Reliability bounds through random sets: Non-parametric methods and geotechnical applications. *Structural and Multidisciplinary Optimization*, 60:2157–2176, 2019.

Oberguggenberger, M. and W. Fellin. Reliability bounds through random sets: Non-parametric methods and geotechnical applications. *Computers & Structures*, 86(10):1093–1101, 2008.

Rosenbrock, H. H. An automatic method for finding the greatest or least value of a function. *The Computer Journal*, 3(3):175–184, 1960.

Schöbi, R. *Surrogate models for uncertainty quantification in the context of imprecise probability modeling.* PhD thesis, ETH Zürich, Zurich, Switzerland, 2017.

Schöbi, R. and B. Sudret. Uncertainty propagation of p-boxes using sparse polynomial chaos expansions. *Journal of Computational Physics*, 339:307–327, 2017.

Soize, C. *Uncertainty Quantification*. Springer International Publishing, Cham, Switzerland, 2017.

Teixeira, R., M. Nogal, and A. O'Connor. Adaptive approaches in metamodel-based reliability analysis: A review *Structural Safety*, 102019, 2021.

Torre, E., S. Marelli, P. Embrechts, and B. Sudret. A general framework for data-driven uncertainty quantification under complex input dependencies using vine copulas *Probabilistic Engineering Mechanics*, 55:1–16, 2019.

Walley, P. *Statistical Reasoning with Imprecise Probabilities*. Chapman & Hall, London, UK, 1991.

Distribution-free P-box processes: definition and simulation

Matthias G.R. Faes
KU Leuven, Department of Mechanical Engineering, LMSD Division, matthias.faes[at]kuleuven.be
Institute for Risk and Reliability, Leibniz Universität Hannover

Matteo Broggi
Institute for Risk and Reliability, Leibniz Universität Hannover, broggi[at]irz.uni-hannover.de

Kok-Kwang Phoon
National University of Singapore, Department of Civil & Environmental Engineering,kkphoon[at]nus.edu.sg

Michael Beer
Institute for Risk and Reliability, Leibniz Universität Hannover, beer[at]irz.uni-hannover.de
Institute for Risk and Uncertainty and School of Engineering, University of Liverpool, Peach Street, Liverpool L69 7ZF, UK
International Joint Research Center for Engineering Reliability and Stochastic Mechanics, Tongji University, 1239 Siping Road, Shanghai 200092, P.R. China

Abstract. Typically, non-deterministic models of spatial or time dependent uncertainty are modelled using the well-established random field framework. However, while tailored for exactly these types of time and spatial variations, stochastic processes and random fields currently have only limited success in industrial engineering practice. This is mainly caused by its computational burden, which renders the analysis of industrially sized problems very challenging, even when resorting to highly efficient random field analysis methods such as EOLE. Apart from that, also the methodological complexity, high information demand and rather indirect control of the spatial (or time) variation has limited its cost-benefit potential for potential end-users. This data requirement was recently relaxed by the first author with the introduction of imprecise random fields, but so far the method is only applicable to parametric p-box valued stochastic processes and random fields. This paper extends these concepts by expanding the framework towards distribution-free p-boxes. The main challenges addressed in this contribution are related to both the non-Gaussianity of realisations of the imprecise random field in between the p-box bounds, as well as maintaining the imposed auto-correlation structure while sampling from the p-box. A case study involving a dynamical model of a car suspension is included to illustrate the presented concepts.

Keywords: uncertainty quantification, stochastic processes, imprecise probabilities

© 2021 *by M.G.R. Faes, M. Broggi, K-K. Phoon, and M. Beer.*

1. Introduction

1.1. General rationale

Stochastic processes are widely used in domains such as engineering (Stefanou, 2009) and financial economics to represent stochastic quantities that vary over time- and/or space (Vanmarcke, 1983). However, due to the theoretical and computational difficulties, usually these processes are assumed to be Gaussian, which might not always be a truthful representation of reality (Grigoriu, 2010). Furthermore, the definition of a stochastic process requires the rigorous description of the governing distribution function, which includes selecting the appropriate distribution family as well as the governing hyper-parameters. In practice, this might not always be possible due to limitations on the available data (quantity of the data, corrupted or missing data, etc.), but also conflicting sources of information (e.g., expert opinions). Recently introduced approaches based on Bayesian compressed sensing alleviate this problem (see e.g., (Wang et al., 2018; Montoya-Noguera et al., 2018) by making the estimated stochastic process robust to missing data, even when these processes are non-Gaussian and have an unknown non-stationary auto-correlation function (Wang et al., 2019). Furthermore, also detrending is not required in these cases (Wang et al., 2019). In (Ching and Phoon, 2020), these methods were extended to also account for the general case of multivariate, uncertain, unique, sparse, incomplete spatial data (denoted MUSIC-X by the authors).

As a possible alternative pathway to the issue of low data availability is to resort to the more general framework of imprecise probabilities (Beer, Ferson and Kreinovich , 2013). According to this framework, epistemic uncertainty resulting from potential data deficiencies are taken explicitly into account to allow for rigorous analysis. In the context of processes, parametric p-box valued stochastic processes have been introduced in e.g., (Dannert et al., 2018), (Faes and Moens, 2019) and (Fina et al., 2020). However, such approaches still require the definition of a distribution family, which is not always possible. Furthermore, these approaches are based on Gaussian distributed processes. This paper aims to go further than the available methods for simulating from imprecise stochastic processes by introducing a type of distribution-free p-box stochastic process. Such process is obtained by passing a standard normal Gaussian stochastic process through an interval-valued translation map, effectively providing a distribution-free p-box stochastic process. An example is included to illustrate the definition and propagation of these structures. The paper is constructed as follows: the remainder of this section recalls some important concepts concerning the definition and simulation of non-Gaussian stochastic processes; Section 2 introduces the approach for defining distribution-free p-box processes; Section 3 briefly discusses a double-loop approach to propagate these structures; Section 4 provides a case study as illustration of the approach; Section 5 lists the conclusions of the work.

1.2. Stochastic processes

A finite-dimensional stochastic process $x(t,\omega)$ describes a set of correlated random variables $x(\omega)$, which are assigned to a number of locations $t \in \Omega_d$ in the model domain $\Omega_d \subset \mathbb{R}^d$ with dimension $d \in \mathbb{N}$. Note that Ω_d may comprise both space and/or time dimensions. Each random variable $x(\omega) : (\Omega, \varsigma, P) \mapsto \mathbb{R}$ as such maps from a complete probability space to the real domain, with $\omega \in \Omega$ a coordinate in sample space Ω and ς the sigma-algebra. In this paper, we require that $x(t,\Omega) \in$

$\mathcal{L}^2(\Omega, P)$, with $\mathcal{L}^2(\Omega, P)$ the Hilbert space of second-order random variables (i.e., finite variance). For a given event $\omega_i \in \Omega$, the corresponding $x(t, \omega_i)$ is a realization of the stochastic process. A stochastic process is considered Gaussian if the distribution of $(x(t_1, \omega), x(t_2, \omega), \ldots, x(t_{n_t}, \omega))$, with $n_t \in \mathbb{N}$, is jointly Gaussian $\forall t \in \Omega_d$. In this case, $x(t, \omega)$ is completely described by its mean function $\mu_x(t) : \Omega_d \mapsto \mathbb{R} = E_\omega[x(t, \omega)]$ and its auto-covariance function $\mathbf{C}_{xx}(t, t') : \Omega_d \times \Omega_d \mapsto \mathbb{R}$, given by $\mathbf{C}_{xx}(t, t') = E_\omega[x(t, \omega), x(t', \omega)]$ (Vanmarcke, 1983). In the remainder of the paper, univariate stochastic processes (i.e., $\Omega_d \in \mathbb{R}^1$) are considered for the sake of conciseness of notation. Note that the concepts explained in this paper scale straightforwardly to multivariate or multi-index stochastic processes (e.g., random fields) as well.

Generally, when applying stochastic fields in an engineering context, for instance to represent a spatially uncertain input quantity of a finite element model, the field has to be discretized over Ω_d. In this context, the Karhunen-Loève expansion is a very powerful tool to represent stochastic fields (Spanos, 1989). Specifically, following the Karhunen-Loève (KL) series expansion, a stochastic process $x(t, \omega)$ is represented as:

$$x(t, \omega) = \mu_x(t) + \sigma_x \sum_{i=1}^{\infty} \sqrt{\lambda_i} \psi_i(t) \xi_i(\omega), \tag{1}$$

with σ_x the standard deviation of the random field and where the quantities $\lambda_i \in (0, \infty)$ and $\psi_i(t) : \Omega_d \mapsto \mathbb{R}$ are respectively the eigenvalues and eigenfunctions of the continuous, bounded, symmetric and positive (semi-)definite stationary auto-correlation function $\rho_{xx}(\tau) : \Omega_d \times \Omega_d \mapsto [0, 1]$, where Ω_d is a bounded and closed interval, in accordance with Mercer's theorem:

$$\rho_{xx}(\tau) = \sum_{i=1}^{\infty} \lambda_i \psi_i(t) \psi_i(t'), \tag{2}$$

These quantities are in practice obtained by solving the homogeneous Fredholm integral equation of the second kind:

$$\int_{\Omega_d} \rho_{xx}(\tau) \psi_i(t') dt' = \lambda_i \psi_i(t), \tag{3}$$

where $t' = t + \tau$ for which many efficient discretization schemes exist (Betz et al., 2014). Since $\rho_{xx}(\tau)$ is bounded, symmetric and positive semi-definite, and furthermore in most practical cases can be assumed positive definite, these eigenvalues λ_i are non-negative and the eigenfunctions $\psi_i(t)$ satisfy the following orthogonality condition:

$$\langle \psi_i(t), \psi_j(t) \rangle = \int_{\Omega_d} \psi_i(t) \psi_j(t) dt = \delta_{ij} \tag{4}$$

with δ_{ij} the Kronecker delta and $\langle \cdot, \cdot \rangle : \Omega_d \times \Omega_d \mapsto \mathbb{R}$ an inner product on the space $\mathcal{L}^2(\Omega, dt)$. In this case, the series expansion in Eq. 2 can be shown to be optimally convergent (Spanos, 1989).

The variables $\xi_i(\omega), i = 1, \ldots, \infty$, introduced in Eq. 1, are uncorrelated random variables, which are determined according to:

$$\xi_i(\omega) = \frac{1}{\sqrt{\lambda_i}} \int_{\Omega_d} [x(t, \omega) - \mu_x(t)] \psi_i(t) dt, \tag{5}$$

which can be shown to be independent standard normally distributed in the case of a Gaussian random field. For practical reasons, the infinite series expansion in Eq. 1 is usually truncated after a finite number of terms $n_{KL} \in \mathbb{N}$:

$$x(t,\omega) = \mu_x(t) + \sigma_x \sum_{i=1}^{n_{KL}} \sqrt{\lambda_i}\psi_i(t)\xi_i(\omega), \tag{6}$$

where n_{KL} should be selected such that a well-chosen variance error metric is minimized.

For non-Gaussian fields, the $\xi_i(\omega), i = 1, \ldots, n_{KL}$ represented in Eq. 5 are non-Gaussian too and their distribution needs to be solved for explicitly. Furthermore, in this case, the corresponding random variables $\xi_i(\omega)$ may be uncorrelated but dependent. These dependencies may be difficult to quantify (Grigoriu, 2010). Finally, Eq. 5 reveals that the distribution of $\xi_i(\omega)$ depends on sample path realisations $x(t,\omega)$ of the stochastic process, and hence, iterative methods such as presented by (Huang, Quek and Phoon, 2001) need to be applied. Alternatively, also translation theory as introduced by Grigoriu (Grigoriu, 1998) provides a viable approach towards simulating from (strongly) non-Gaussian stochastic processes.

1.3. Translation stochastic fields

Translation process theory, as introduced by Grigoriu (Grigoriu, 1998), provides a different pathway for the simulation of non-Gaussian stochastic fields. Specifically, a stationary Gaussian stochastic process $\eta(t,\omega)$ with autocorrelation function $\rho_{xx}(\tau)$ is transformed using a nonlinear transformation into a non-Gaussian field $x(t,\omega)$, which is formally expressed as:

$$x(t,\omega) = F_X^{-1} \circ \Phi(\eta(t,\omega)) = g(\eta(t,\omega)) \tag{7}$$

with $g := F_X^{-1} \circ \Phi$ the so-called translation mapping, F_X^{-1} the inverse of the target non-Gaussian cumulative distribution function (CDF) that represents the distribution of the non-Gaussian stochastic process and Φ the marginal standard normal CDF, i.e., $\Phi(\eta) = P(\eta(t,\omega) < \eta)$.

It can be shown (see (Grigoriu, 1998)) that the mean μ_x, variance σ_x^2 and correlation function $r_{xx}(\tau)$ of $x(t,\omega)$ have closed-form solutions, that are given respectively by:

$$\mu_x = E_\eta[x(t,\omega)] = E_\eta[g(\eta(t,\omega))] = \int_{-\infty}^{\infty} g(\eta)\phi(\eta)d\eta \tag{8}$$

$$\sigma_x^2 = E_\eta[x(t,\omega) - \mu_x]^2 = \int_{-\infty}^{\infty} (g(\eta) - \mu_x)^2\phi(\eta)d\eta \tag{9}$$

$$r_{xx}(\tau) = E[x(t,\omega) - \mu_x]E[x(t+\tau,\omega) - \mu_x] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (g(\eta_1) - \mu_x)(g(\eta_2) - \mu_x)\phi(\eta_1, \eta_2, \rho_{xx}(\tau))dydz \tag{10}$$

with $\phi(\eta_1, \eta_2, \rho_{xx}(\tau))$ the density of a bivariate standard Gaussian vector with correlation coefficient $\rho_{xx}(\tau)$, given by:

$$\phi(\eta_1, \eta_2, \rho_{xx}(\tau)) = \frac{1}{\sqrt{2\pi(1-\rho_{xx}(\tau)^2)}} \exp -\frac{\eta_1^2 + \eta_2^2 - 2\eta_1\eta_2\rho_{xx}(\tau)}{2(1-\rho_{xx}(\tau))^2} \tag{11}$$

These equations can be readily solved by standard numerical quadrature schemes. Furthermore, the forward transformation of a Gaussian process to a non-Gaussian process is always possible (Grigoriu, 1998).

On the other hand, given a non-Gaussian $r_{xx}(\tau)$, it is not always possible to determine the corresponding Gaussian autocorrelation $\rho_{xx}(\tau)$ that, when transformed, yields $r_{xx}(\tau)$ (Grigoriu, 1998). This happens when either the inverse of Eq. 10 yields an autocorrelation function that is not positive semi-definite, or when the normalized autocorrelation $\xi(\tau)$ has values that lie outside of the admissible range $[\xi(\tau^{min}), \xi(\tau^{max})]$ which can be found by setting $\rho_{xx}(\tau)$ to respectively 1 and -1 in Eq. 10. In literature, approximative methods have been introduced to find a representative $r_{xx}(\tau)$ that is an admissible autocorrelation function (see e.g., (Kim and Shields, 2015)).

2. P-box stochastic fields

In practical engineering cases, it is not always possible to define a crisp distribution function $F_X(x)$ to construct the non-Gaussian stochastic process. In this context, p-boxes can provide a valuable tool to represent the uncertainty an analysist has on the specification of the appropriate distribution function. This section introduces p-box-valued stochastic processes.

A scalar distribution-free p-box is usually described by a lower CDF $\underline{F}_X \in \mathbb{F}$ and an upper CDF $\overline{F}_X \in \mathbb{F}$, where $\mathbb{F}$ expresses the the set of all CDFs on $\mathbb{R}$. They are collected as a pair $\left[\underline{F}_X, \overline{F}_X\right]$ which yields a set of possible CDFs via $\underline{F}_X(x) \leq F_X(x) \leq \overline{F}_X(x),\ x \in \mathbb{R}$. A distribution-free p-box as such corresponds to defining a lower probability $\underline{P}$ and upper probability $\overline{P}$ on events $\{X \leq x\} = (-\infty, x]$, i.e., $\underline{P}(X \leq x) = \underline{F}_X(x)$ and $\overline{P}(X \leq x) = \overline{F}_X(x)$ for $x \in \mathbb{R}$, which define a credal set of probability measures. In case additional information on the uncertainty is available, constraints on the p-box can be enforced. For instance, if the (class of) distribution functions $\mathcal{F}$ is known, the set of possible CDFs $\{F_X(\cdot, \theta) \in \mathcal{F} \mid \theta \in D_\theta\}$ can be defined conditional on a vector of hyper-parameters θ. Since this is a special case of the distribution-free p-box, the following discussion on distribution-free p-box stochastic fields is equally applicable.

As explained in Section 1, a (precise) stochastic process $x(t, \omega)$ can be considered as a collection of n_t correlated random variables distributed throughout the model domain Ω_d. Imprecise stochastic processes and fields $\hat{x}(t, \omega)$ can be regarded as a natural extension of this idea, where for each discrete location $t_i \in \Omega_d$, a scalar p-box is defined. However, since all $x(t_i, \omega)$ are correlated according to $\rho_x(\tau)$, these scalar p-boxes also are correlated to each other. This observation complicates the analysis of distribution-free imprecise stochastic processes drastically, since the direct simulation from a set of correlated distribution-free p-boxes is far from trivial from both a theoretical as a numerical point of view. Furthermore, since distribution-free p-boxes are considered, also non-Gaussian processes are inherently included in the imprecise stochastic description.

A potential solution to this issue is to start from a precise standard normal Gaussian stochastic process $\eta(t, \omega)$ with predefined correlation function $\rho_{xx}(\tau)$ and pass this representation through an imprecisely defined translation map which is defined as:

$$\hat{x}(t, \omega) = \left[\underline{F}_X^{-1}, \overline{F}_X^{-1}\right] \circ \Phi(\eta(t, \omega)), \tag{12}$$

which can be further expanded as:

$$\hat{x}(t,\omega) = \left[\underline{F}_X^{-1}, \overline{F}_X^{-1}\right] \circ \Phi\left(\sum_{i=1}^{n_{KL}} \sqrt{\lambda_i}\psi_i(t)\xi_i(\omega)\right), \tag{13}$$

where $\Phi\left(\sum_{i=1}^{n_{KL}} \sqrt{\lambda_i}\psi_i(t)\xi_i(\omega)\right)$ represents a re-scaling of $\eta(t,\omega)$ to the interval $[0,1]$. Since a CDF is by definition a monotonic function, i.e., $F_X(x_1) \leq F_X(x_2) \iff x_1 < x_2$, the bounds of $\hat{x}(t,\omega)$ are determined by $\left[\underline{F}_X^{-1}, \overline{F}_X^{-1}\right]$. Note that in case F_X is not strictly monotonic (i.e.,$F_X(x_1) < F_X(x_2) \iff x_1 < x_2$, pseudo-inverses should be used to calculate the inverse of F_X. Furthermore, since $\left[\underline{F}_X^{-1}, \overline{F}_X^{-1}\right] \circ \Phi$ represents an interval-valued mapping, each realisation of the stochastic process $\eta(t,\omega_j)$, corresponding to the event ω_j is translated towards an interval field that is consistent with the bounds on the CDF (due to the monotonicity of the CDF), and which is given as:

$$x^I(t,\omega_j) = \left[\underline{F}_X^{-1}, \overline{F}_X^{-1}\right] \circ \Phi\left(\sum_{i=1}^{n_{KL}} \sqrt{\lambda_i}\psi_i(t)\xi_i(\omega_j)\right), \tag{14}$$

with the lower bound given as:

$$\underline{x}(t,\omega_j) = \overline{F}_X^{-1} \circ \Phi\left(\sum_{i=1}^{n_{KL}} \sqrt{\lambda_i}\psi_i(t)\xi_i(\omega_j)\right), \tag{15}$$

and the upper bound defined as:

$$\overline{x}(t,\omega_j) = \underline{F}_X^{-1} \circ \Phi\left(\sum_{i=1}^{n_{KL}} \sqrt{\lambda_i}\psi_i(t)\xi_i(\omega_j)\right). \tag{16}$$

It should be noted that this is not an explicit interval field, i.e., an interval field that is represented as a series expansion with interval-valued coefficients as described in (Faes and Moens, 2020a), since the interval-valued nature in this field stems from the mapping that is performed on a single realisation of the crisp Gaussian field, rather than from a series expansion with interval-valued weights. As such, typically applied interval propagation methods, as described in (Faes and Moens, 2020b), cannot be applied straightforwardly to propagate this interval field. Furthermore, the auto-dependence function of realisations within this interval field becomes interval-valued too. Note that in case of intervals, auto-dependence is used to describe the dependency throughout Ω_d, rather than autocorrelation or autocovariance. Intermediate realisations of $x^k(t,\omega_j) \in x^I(t,\omega_j)$ can be generated by drawing admissible CDFs $F_X^{-1,k} \in \left[\underline{F}_X^{-1}, \overline{F}_X^{-1}\right]$. For reasons of clarity, the explanation of a possible procedure to do so is deferred to a later section.

Conversely, when collecting all $x^k(t,\omega)$, that correspond to a certain realisation $F_X^{-1,k} \in [\underline{F}_X^{-1}, \overline{F}_X^{-1}]$, this becomes again a crisp random field, the properties of which can be computed by virtue of translation field theory as:

$$\mu_{x^k(t,\omega)} = \int_{-\infty}^{\infty} F_X^{-1,k} \circ \Phi(\eta)\phi(\eta)d\eta \tag{17}$$

$$\sigma^2_{x^k(t,\omega)} = \int_{-\infty}^{\infty} (F_X^{-1,k} \circ \Phi(\eta) - \mu_{x^k(t,\omega)})^2 \phi(\eta) d\eta \tag{18}$$

$$r_{x^k x^k}(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (F_X^{-1,k} \circ \Phi(\eta_1) - \mu_{x^k(t,\omega)})(F_X^{-1,k} \circ \Phi(\eta_2) - \mu_{x^k(t',\omega)}) \phi(\eta_1, \eta_2, \rho_{xx}(\tau)) d\eta_1 d\eta_2 \tag{19}$$

with $\phi(\eta_1, \eta_2, \rho_{xx}(\tau))$ as defined in Eq. (11). Note that this random field is stationary in case the underlying Gaussian field is stationary since it is mapped through a crisp $F_X^{-1,k}$. Evidently every $F_X^{-1,k}$ will yield a random field with generally different central moments for every $F_X^{-1,k} \in [\underline{F}_X^{-1}, \overline{F}_X^{-1}]$.

Based on the preceding discussion, it can be seen that an imprecise stochastic process can jointly be regarded as a stochastic collection of interval fields as well as as a credal set of stochastic fields. Since each stochastic realisation of the imprecise stochastic process is an interval field, and vice versa, each realisation within the p-box corresponds to a stochastic process, also the mean of the corresponding imprecise field is interval-valued. Due to the monotonicity of the translation map, this mean is formally given as:

$$\mu_x^I = [\underline{\mu}_x, \overline{\mu}_x] = \left[E_\omega\left[\underline{x}(t,\omega)\right], E_\omega\left[\overline{x}(t,\omega)\right]\right]. \tag{20}$$

These bounds can be computed in a straightforward manner by invoking Eq. 17 twice: once on the ensemble of lower bounds and once on the ensemble of upper bounds. As an illustration of these concepts, consider the example in Figure 1. This figure shows two stochastic realisations of the crisp zero-mean Gaussian stochastic process, as well as their transformation into two realisations of the distribution-free p-box random field, which manifest themselves in the shape of two interval fields. From this figure, it can also be seen that the collection of all upper bounds of these interval fields, denoted $\overline{x}(t,\omega)$, tends towards a stochastic process that has $\underline{F}_X(x)$ as a distribution when the number of terms in the KL expansion tends to infinity. The same obviously holds for the lower bound, as well as any intermediate realisation $F_X^{-1,k} \in \left[\underline{F}_X^{-1}, \overline{F}_X^{-1}\right]$.

As a final comment, it should be noted that the auto-correlation function of $x^I(t,\omega)$, being $r_{xx}(\tau)$ also has become an interval field due to the interval-valued translation map. This might have important implications for structural dynamical problems, where the match of a dominant frequency of the loading process might interfere with a natural frequency of the structure. The detailed treatment of this issue and a potential solution hereto however fall outside the scope of this paper.

3. Propagation of distribution-free imprecise stochastic fields

Usually, an analyst who is confronted with imprecise probabilistic model quantities is concerned with finding the bounds on some probabilistic measure $\mathcal{P}$ of the model's responses of interest. In case a crisp density function $f_{\mathbf{X}}$ is known, the n^{th} central moment of the model's response $E_{\mathbf{y}}[Y^n]$ or the probability of failure p_F is determined by evaluating an integral equation of the following form:

$$\mathcal{P} = \int_{\mathbb{R}^{n_{KL}}} \mathcal{H}(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) \, \mathrm{d}\mathbf{x}, \tag{21}$$

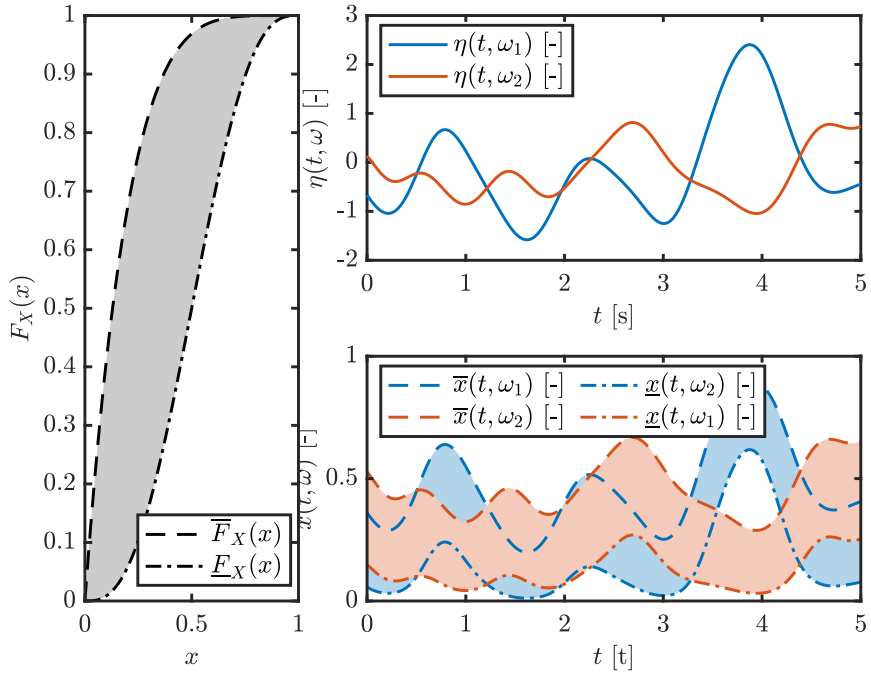


Figure 1. Illustration of the transformation of two stochastic realisations of the crisp zero-mean Gaussian stochastic process into two realisations of the distribution-free p-box random field, which manifest themselves in the shape of two interval fields.

where $\mathcal{P}$ denotes, depending on the context, the n^{th} central moment of the model's response $E_{\mathbf{y}}[Y^n]$ or the probability of failure p_F. In this equation, the random variable $\mathbf{X} = [\xi_1, \xi_2, \ldots, \xi_{n_{KL}}]$ represents a vector of i.i.d. standard normal random variables. These stem from the finite-dimensional approximation of the random field $\eta(t,\omega)$, as given in Eq. (6). Hence, $f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^{n_{KL}} \phi(x_i)$. In case $\mathcal{P} \equiv E_{\mathbf{y}}[Y^n]$ is considered, $\mathcal{H} \equiv r^n(\mathbf{x})$, where $r(\mathbf{x})$ represents the so-called performance function of the model. On the other hand, in case the calculation is aimed at computing p_F, $\mathcal{H} \equiv I_r(\mathbf{x})$, where I_r is the indicator function which is 1 in case $r(\mathbf{x}) \leq 0,\ \mathbf{x} \in \mathbb{R}^{n_x}$, and 0 otherwise. In this context, $\mathbf{x}$ is used to denote

In the specific case of distribution-free imprecise random fields, the probabilistic measure $\mathcal{P}$ becomes conditional on the realisation F_X^k of the p-box that represents the translation map CDFs, i.e., $\mathcal{P}(F_X^k)$. As explained in Section 2, the distribution of a random field realisation $x^k(t,\omega)$ of the imprecise random field $\hat{x}(t,\omega)$, as described in Eq. (13) tends towards F_X^k as $n_{KL} \to \infty$. In this case, Eq. (22) reads:

$$\mathcal{P}(F_X^k) = \int_{\mathbb{R}^{n_{KL}}} \mathcal{H}\left(F^{-1,k} \circ \Phi\left(\boldsymbol{\Psi}\sqrt{\boldsymbol{\Lambda}}\mathbf{x}\right)\right) f_{\mathbf{X}}^k(\mathbf{x})\,\mathrm{d}\mathbf{x}, \tag{22}$$

where f_X^k represents the i.i.d. standard normal random variables that have been passed through the translation map and $\boldsymbol{\Psi}$ and $\boldsymbol{\Lambda}$ representing respectively a matrix and vector collecting the

eigenfunctions and -values of the autocorrelation function of the underlying Gaussian field. The main point in calculating this time or space-dependent probabilistic measure lies in the definition of the performance function. For instance, the performance function can be defined such that it converts the time-dependent problem into a series event by considering a first passage probability.

To infer the bounds on $\mathcal{P}$, two optimization problems need to be solved to actively search the parameter space spanned by $[\underline{F}_X(x), \overline{F}_X(x)]$. Specifically, the lower bound is obtained as:

$$\underline{\mathcal{P}} = \min \mathcal{P}(F_X^k), \tag{23}$$

where the minimum is taken over all distribution functions F_X^k such that $[F_X^k \in \underline{F}_X(x), \overline{F}_X(x)]\ \forall x$. Sinilary, the upper bound is determined as:

$$\overline{\mathcal{P}} = \max \mathcal{P}(F_X^k) \tag{24}$$

Note that each realisation F_X^k drawn from this interval represents a non-Gaussian random field with auto-correlation structure as described in Eq. (10). As such, this is effectively a double-loop approach, which might entail a non-negligible computational cost to solve. The main difficulty associated with solving these optimization problems lies in the fact that the optimisation has to be performed over the infinite-dimensional space of bounded, strictly monotonically increasing functions over the support of x. Such calculation is intractable, even for the most simple cases. In this paper, it is therefore proposed to approximate these optimization problems as discrete problems. Specifically, it is aimed at solving following problems:

$$\begin{aligned} \underline{\mathcal{P}} &= \min_{\mathbf{F}_X \in \mathbf{F}_X^I} \mathcal{P}(F_X^k) \\ \overline{\mathcal{P}} &= \max_{\mathbf{F}_X \in \mathbf{F}_X^I} \mathcal{P}(F_X^k) \end{aligned} \tag{25}$$

subject to:

$$\mathbf{A}\mathbf{F}_X \leq 0 \tag{26}$$

where $\mathbf{F}_X = F_X(\mathbf{x}_s)$, with $\mathbf{x}_s \in \mathbb{R}^{n_s}$ representing n_s equally spaced sample points throughout the support of $[\underline{F}_X(x), \overline{F}_X(x)]$. Similarly, $\mathbf{F}_X^I = [\underline{F}_X(\mathbf{x}_s), \overline{F}_X(\mathbf{x}_s)] \in \mathbb{IR}^{n_s}$ represents an n_s dimensional interval vector collecting the bounds of the p-box for each sample point in the support. The inequality shown in Eq. (26) enforces the realisations drawn from the interval vector $\mathbf{F}_X^I$ to be strictly monotonic to ensure that they represent admissible CDFs, where $\mathbf{A} \in \mathbb{R}^{n_s-1 \times n_s}$ represents an upper-triangular band matrix with $A_{1,:} = [1 \ -1 \ 0 \ \dots \ 0]$. As such, the infinite-dimensional optimisation problem is converted to a linear-inequality-constrained optimization problem over n_s variables. Finally, it can be noted that the translation mapping explained in section 2 requires the calculation of the inverse of the CDF (see e.g., Eq. 17). This is for instance required to generate the required sample paths of $x^k(t,\omega)$ to estimate the performance function $r(\mathbf{x}^k(t,\omega))$. Hereto, a piece-wise cubic Hermite polynomial interpolation is performed using the Fritsch-Carlson algorithm to estimate a functional relationship between $\mathbf{x}_s$ and $\mathbf{F}_X$. This approach is selected as this allows for generating a strictly monotonic, C^1 continuous interpolation of the inverse of $\mathbf{F}_X(\mathbf{x}_s)$ (Fritsch and Carlson, 1980).

4. Case study

The case study represents a quarter-car model, which is a 2-DOF idealisation of the realistic dynamics of the suspension of a car. Specifically, this case study is concerned with assessing the bounds on several comfort metrics of a vehicle suspension, given a p-box process-valued base excitation. The quarter-car dynamics can be represented as a set of two ordinary differential equations:

$$m_s\ddot{x}_s + c_s(\dot{x}_s - \dot{x}_{us}) + k_s(x_s - x_{us}) = 0 \tag{27}$$

$$m_{us}\ddot{x}_{us} - c_s(\dot{x}_s - \dot{x}_{us}) - k_s(x_s - x_{us}) + c_t(\dot{x}_{us} - \dot{x}_0) + k_t(x_{us} - x_0) = 0 \tag{28}$$

with $\dot{\bullet}$ the time derivative of $\bullet$, x_{us} the displacement of the unsprung mass (i.e., the suspension components, wheel and other components directly connected to them), x_s the displacement of the sprung mass (i.e., all components resting on the suspension), m_{us} and m_s the unsprung and sprung mass of a quarter of the car, c_s and c_t respectively the damping coefficients of the suspension and tire, k_s and k_t respectively the stiffness coefficients of the suspension and tire. Finally, x_0 and $\dot{x}_0$ are the displacement and velocity in vertical direction that excite the bottom of the wheel (i.e., the road profile). The complete road profile is denoted $x_0(t)$. The dynamics of the car are simulated over a distance of 50 (m), when the car is travelling at a speed of 10 (m/s). The one-dimensional spatial domain is discretized into 1000 equidistant points and the time domain is discretized into time intervals of 0.005 (s). A schematic representation of the model is given in figure 2.

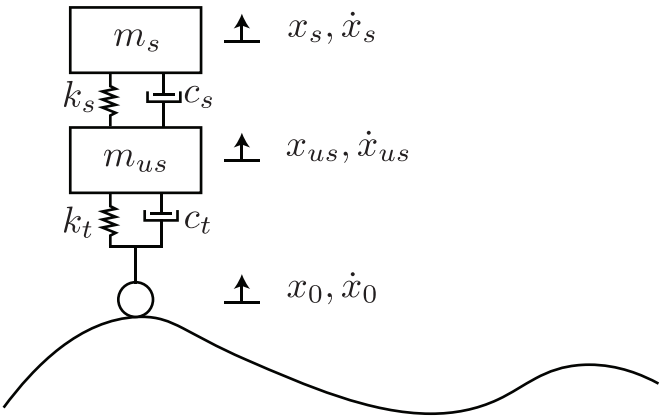


Figure 2. Schematic illustration of the quarter-car model

For the solution of this coupled system of ODEs, a state-space model is employed:

$$\frac{d}{dt}\begin{bmatrix} x_{us} - x_0 \\ \dot{x}_{us} \\ x_s - x_{us} \\ \dot{x}_s \end{bmatrix} = A\begin{bmatrix} x_{us} - x_0 \\ \dot{x}_{us} \\ x_s - x_{us} \\ \dot{x}_s \end{bmatrix} + \begin{bmatrix} -1 \\ \frac{c_t}{m_{us}} \\ 0 \\ 0 \end{bmatrix}\dot{x}_0 \tag{29}$$

with the matrix A equal to:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{-k_t}{m_{us}} & \frac{-(c_s+c_t)}{m_{us}} & \frac{k_s}{m_{us}} & \frac{c_s}{m_{us}} \\ 0 & -1 & 0 & 1 \\ 0 & \frac{c_s}{m_s} & \frac{-k_s}{m_s} & \frac{-c_s}{m_s} \end{bmatrix} \tag{30}$$

Four state variables are considered, being respectively the tire deflection $(x_{us} - x_0)$; the unsprung mass velocity $\dot{x}_{us}$; the suspension stroke $x_s - x_{us}$, and sprung mass velocity $\dot{x}_s$. Typically, in the context of assessing the dynamical comfort of a car, two parameters are of interest: the suspension stroke (i.e., the relative displacement of the car body with respect to the unsprung mass) and the acceleration of the sprung mass. In the proceeding study, the damping effect of the tire, c_t is considered negligible.

Table I. Considered case studies for the distribution free p-box field

Case	$\underline{F}_X(x)$	$\overline{F}_X(x)$	b_t
1	$\min[\mathcal{B}(1,3), \mathcal{B}(5,5)])$	$\max[\mathcal{B}(1,3), \mathcal{B}(5,5)])$	0.5
2	$\min[\mathcal{B}(3,1), \mathcal{B}(5,5)])$	$\max[\mathcal{B}(3,1), \mathcal{B}(5,5)])$	0.5
3	$\min[\mathcal{B}(1,1), \mathcal{B}(2,5)])$	$\max[\mathcal{B}(1,1), \mathcal{B}(2,5)])$	0.5
4	$\min[\mathcal{B}(1,0.2), \mathcal{B}(5,5)])$	$\max[\mathcal{B}(1,.02), \mathcal{B}(5,5)])$	0.5
5	$\min[\mathcal{N}(0,0.75), \mathcal{B}(1,0.2)])$	$\max[\mathcal{N}(0,0.75), \mathcal{B}(1,0.2)])$	2

The complete road profile $x_0(t)$ is modelled as a p-box valued stochastic process. The auto-correlation of the underlying Gaussian process $\eta(t,\omega)$ is governed by a squared exponential auto-correlation function with a correlation length of 0.5 m. Sample path realisations of $\eta(t,\omega)$ are generated using the Karhunen–Loève series expansion, while retaining 32 terms. The stochastic content of the imprecise stochastic process is represented using a distribution-free p-box. For illustrative reasons, the bounds of the p-box, $[\underline{F}_X(x), \overline{F}_X(x)]$ are generated by taking the extremes of a set of distributions. Hereto, 5 different case studies are considered, which are summarized in Table I. The corresponding p-boxes area also visualized in figure 3. Note that these cases do not necessarily represent a physical phenomenon. Rather, they are selected for illustrative reasons. Each random field corresponding to a realisations of these p-boxes is given as:

$$x_0^k(t,\omega) = (F_X^k)^{-1} \circ \Phi\left(\sum_{i=1}^{n_{KL}} \sqrt{\lambda_i}\psi_i(t)\xi_i(\omega)\right) \tag{31}$$

which are generated by the optimization algorithms introduced in Section 4. Applying the double-loop optimization algorithm introduced in Section 4, the bounds on the probability of failure of the structure are computed. In this context, the performance function $r(\mathbf{x})$ of the car model is given as:

$$r(x_0) = 1 - \max_{i=1,\ldots,m}\left(\frac{|x_s(\mathbf{x},t_i) - x_{us}(\mathbf{x},t_i)|}{b_t}\right) \tag{32}$$

where the threshold value b_t is also given in Table I. This corresponds to a first passage probability. Since the process is non-Gaussian, highly efficient and dedicated sampling methods such as Directional Importance Sampling (Misraji et al., 2020), as also applied in the context of imprecise probabilities in (Faes et al., 2020a) or (Faes et al., 2020b), are not applicable. Therefore, the integral equation in the inner loop of the optimization is solved using Subset-∞, as presented in (Au & Patelli, 2016), with an initial sample size of 5000 and a proposal standard deviation of 0.1. The discretisation of $F_X(\mathbf{x}_s)$, as described in Section 4, is performed using $n_s = 40$ slices, yielding a 40-dimensional optimization problem, which is solved using a gradient-free pattern search optimization algorithm. Pattern search is specifically selected to avoid the need to calculate gradients of p_F.

The results of performing the double-loop optimization problem are shown in Table II. In this table, $\underline{p}_F^*$ and $\overline{p}_F^*$ indicate the bounds on p_F obtained by means of optimization, whereas $\underline{p}_F$ and $\overline{p}_F$ are the failure probabilities corresponding to the bounds of the p-box. As is clear, the bounds obtained by just propagating $\underline{F}_X(x)$ and $\overline{F}_X(x)$ are not conservative. This is a direct result from the fact that the car model acts as a filter on the excitation towards the responses of interest.

Table II. Bounds on the probability of failure based on propagating the bounds, as well as performing optimization.

Case	$\underline{p}_F$	$\overline{p}_F$	$\underline{p}_F^*$	$\overline{p}_F^*$
1	$3.25 \cdot 10^{-4}$	0.065	$2.77 \cdot 10^{-4}$	0.385
2	$3.25 \cdot 10^{-4}$	0.061	$9.95 \cdot 10^{-6}$	0.378
3	0.0045	0.400	0.0028	0.624
4	0.0038	0.469	$3.01 \cdot 10^{-4}$	0.571
5	0.0062	0.015	$9.43 \cdot 10^{-6}$	0.423

A further explanation of these results can be given based on Figure 3. This figure shows clearly that the CDF corresponding to the highest probability of failure pushes the probability mass as much as possible towards the bounds of the p-box. This makes sense from a physical standpoint since the performance function contains an absolute value operation, and hence, positive and negative responses contribute both equally to the failure. Furthermore, the considered system is a 2-degree-of-freedom oscillator where the quantity of interest is the relative displacement between the two masses. This as such constitutes a perfect symmetric system. The CDF that minimizes p_F on the other hand aims at getting as much of the probability mass as possible towards the centre of the support.

5. Conclusions

This paper discusses the concept of distribution-free p-box stochastic processes and fields. To generate realisations of such a process, it is proposed to pass realisations of a standard Gaussian process through an imprecisely defined translation map such that the auto-correlation of the original

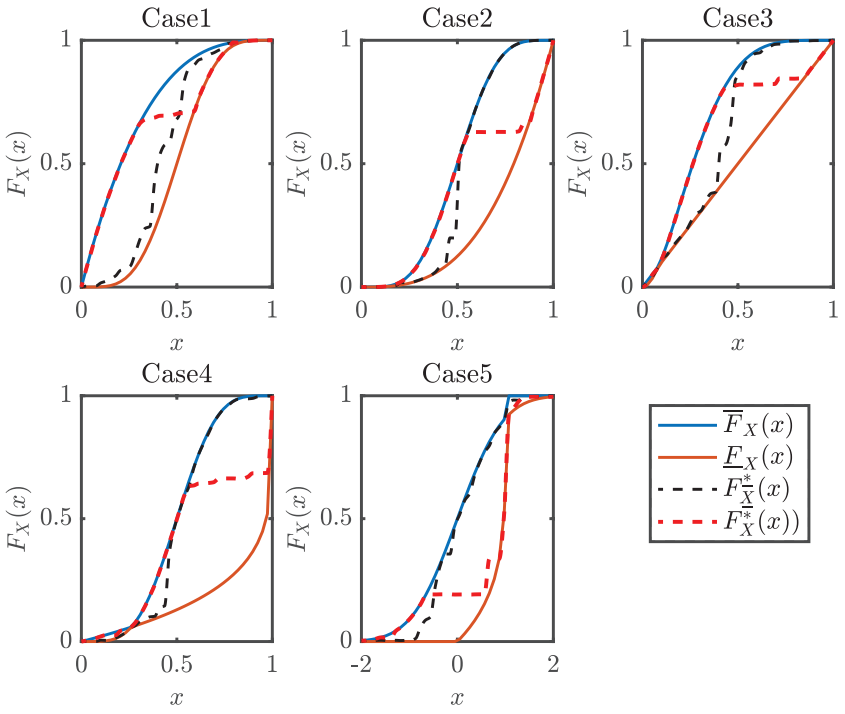


Figure 3. P-boxes corresponding to the 5 cases, as well as the realisations $F^*_{\underline{X}}$ and $F^*_{\overline{X}}$ that yield respectively $\underline{p}^*_F$ and $\overline{p}^*_F$.

process is largely maintained. Furthermore, an optimization approach is introduced to actively look for those realisations inside the p-box that yield a stochastic process that yields an extreme in a probabilistic measure of a response of interest. A case study on a quarter car model illustrated that the bounds of the P-box in fact do not necessarily coincide with the bounds on the probability of failure, which motivates the application of optimization algorithms.

Future work will focus on propagating the p-box field such that a target auto-correlation of the p-box field can be predefined by looking for an appropriate pre-mapped auto-correlation function for each realisation F^k_x in the p-box. Furthermore, the application of more advanced approaches such as e.g., based on sparse polynomial chaos expansions as discussed in (Schöbi and Sudret, 2017) to propagate the p-box will be investigated.

Acknowledgements

Matthias Faes acknowledges the financial support of the Research Foundation Flanders (FWO) under grant number 12P3519N as well as the Alexander von Humboldt foundation under the Humboldt Research Fellowship framework. The authors would further like to thank the anonymous

reviewer for their elaborate and helpful comments on the article. We greatly appreciate the time and effort the reviewer spent in providing those comments.

References

G. Stefanou. The stochastic finite element method: Past, present and future. Computer Methods in Applied Mechanics and Engineering, 198(9–12) (2009) 1031–1051.

E. Vanmarcke, Random Fields: Analysis and Synthesis, MIT Press, Cambridge, 1983.

Y. Wang, T. Zhao, K.-K. Phoon. Direct simulation of random field samples from sparsely measured geotechnical data with consideration of uncertainty in interpretation. Canadian Geotechnical Journal, 55(6) (2018), 862–880.

S. Montoya-Noguera, T. Zhao, Y. Hu, Y. Wang, K.-K. Phoon. Simulation of non-stationary non-Gaussian random fields from sparse measurements using Bayesian compressive sampling and Karhunen-Loève expansion. Structural Safety, 79 (2019), 66–79.

Y. Wang, T. Zhao, Y. Hu, K.-K. Phoon. Simulation of Random Fields with Trend from Sparse Measurements without Detrending. Journal of Engineering Mechanics, 145(2) (2019), 04018130.

J. Ching, K.-K. Phoon. Constructing a Site-Specific Multivariate Probability Distribution Using Sparse, Incomplete, and Spatially Variable (MUSIC-X) Data. Journal of Engineering Mechanics, 146(7) (2020), 1–18.

M. Beer, S. Ferson, V. Kreinovich. Imprecise probabilities in engineering analyses. Mechanical Systems and Signal Processing, 37(1–2) (2013), 4–29.

P. Spanos, R. Ghanem, Stochastic finite element expansion for random media, Journal of Engineering Mechanics 115 (5) (1989) 1035–1053.

W. Betz, I. Papaioannou, D. Straub, Numerical methods for the discretization of random fields by means of the Karhunen-Loève expansion, Computer Methods in Applied Mechanics and Engineering 271 (2014) 109–129.

M. Grigoriu, Probabilistic models for stochastic elliptic partial differential equations, Journal of Computational Physics 229 (22) (2010) 8406 – 8429.

S. P. Huang, S. T. Quek, K. K. Phoon, Convergence study of the truncated karhunen–loeve expansion for simulation of stochastic processes, International Journal for Numerical Methods in Engineering 52 (9) (2001) 1029–1043.

H. Kim, M. D. Shields, Modeling strongly non-Gaussian non-stationary stochastic processes using the Iterative Translation Approximation Method and Karhunen-Loève expansion, Computers and Structures 161 (2015) 31–42.

M.M. Dannert, A. Fau, R.M.N Fleury, M. Broggi, U. Nackenhorst, M. Beer. A probability-box approach on uncertain correlation lengths by stochastic finite element method. PAMM (Proceedings in Applied Mathematics and Mechanics), 18(1) (2018) e201800114.

M. Faes, D. Moens. Imprecise random field analysis with parametrized kernel functions. Mechanical Systems and Signal Processing, 134 (2019), 106334.

M. Fina, P. Weber, W. Wagner. Polymorphic uncertainty modeling for the simulation of geometric imperfections in probabilistic design of cylindrical shells. Structural Safety, 82 (2020), 101894.

M. Grigoriu, Simulation of stationary non-Gaussian translation processes, Journal of Engineering Mechanics 124 (2) (1998) 121–126.

M. Faes, D. Moens, On auto- and cross-interdependence in interval field finite element analysis, International Journal for Numerical Methods in Engineering 121 (9) (2020) 2033–2050.

M. Faes, D. Moens, Recent Trends in the Modeling and Quantification of Non-probabilistic Uncertainty, Archives of Computational Methods in Engineering 27 (3) (2020), 633–671.

F.N. Fritsch, R.E. Carlson. Monotone Piecewise Cubic Interpolation. SIAM Journal on Numerical Analysis 17(2) (1980), 238–246.

M.A. Misraji, M.A. Valdebenito, H.A. Jensen, C.F. Mayorga, Application of directional importance sampling for estimation of first excursion probabilities of linear structural systems subject to stochastic Gaussian loading. Mechanical Systems and Signal Processing, 139 (2020) 106621.

M.G.R. Faes, M.A. Valdebenito, D. Moens, M. Beer. Bounding the first excursion probability of linear structures subjected to imprecise stochastic loading. Computers & Structures, 239 (2020) 106320.

M.G.R. Faes, M.A. Valdebenito, X. Yuan, P. Wei, M. Beer. Augmented Reliability Analysis for Estimating Imprecise First Excursion Probabilities in Stochastic Linear Dynamics (2020). Preprint Submitted to Elsevier.

S.-K. Au, E. Patelli. Rare event simulation in finite-infinite dimensional space. Reliability Engineering & System Safety, 148 (2016) 67–77.

R. Schöbi, B. Sudret. Uncertainty propagation of p-boxes using sparse polynomial chaos expansions. Journal of Computational Physics, 339 (2017) 307–327.

Distribution-free uncertainty propagation

Scott Ferson[1] and Ander Gray[2]
Institute for Risk and Uncertainty, University of Liverpool, UK
[1]*Scott.Ferson@liverpool.ac.uk,* [2]*Ander.Gray@liverpool.ac.uk*

Abstract. Elementary formulas for propagating information about means and variances through mathematical expressions have long been used by analysts. Yet the precise implications of such information are rarely articulated. This paper explores distribution-free techniques of uncertainty propagation that do not require simulation, sampling or approximation of any kind. We describe best-possible bounds on exceedance risks (probabilities of extreme events) that can be inferred given only information about the range, mean and variance of a random variable. These bounds generalize the classical Chebyshev inequality in an obvious way, yet apparently have not been described elsewhere. We also collect in convenient tables several formulas for propagating range and moment information through calculations involving 7 binary convolutions (addition, subtraction, multiplication, division, powers, minimum, and maximum) and 9 unary transformations (scalar multiplication, scalar translation, exponentiation, natural and common logarithms, reciprocal, square, square root and absolute value) commonly encountered in uncertain expressions. These formulas are rigorous rather than approximate, and in most cases are either exact or mathematically best possible. The formulas can be used effectively even when only interval estimates of the moments are available. Although most discussions of moment propagation assume stochastic independence among variables, this paper shows the assumption to be unnecessary and generalizes formulas for the case when no assumptions are made about dependence. These formulas can be viewed as a distribution-free risk analysis.

Keywords: uncertainty propagation, moment propagation, distribution-free risk analysis, imprecise probabilities

1. Introduction

Many authors have suggested propagating means and variances of variables through mathematical expressions as a crude form of risk analysis. This approach is sometimes called first-order error analysis, and it is a widely used approach for making risk estimates. In traditional probability theory, these calculations are called moment propagation and are considered a fundamental part of mathematical statistics (see, for example, Wilks 1962). Despite this wide use, there has always been a disconnect between moment propagation and what these calculations would imply about risks of extreme values of the variable. For instance, after reviewing some moment propagation formulas, Cullen and Frey (1999, page 184) gave a rather pessimistic conclusion:

> Although the results of [the formulas] are useful in some cases for propagating the mean and variance trough a simple linear model, they do not imply anything about the shape of the

© 2021 *by Ferson and Gray.*

model output distribution. Thus, if we were interested in making predictions regarding the 95th percentile of the model output for a linear function of independent random variables, we would not have sufficient information based solely on the properties of the mean and variance to do this.

Their pessimistic view is based on the fact that current moment propagation methodologies:

- Require stochastic independence
- Require moments to be perfectly known (point values)
- Give no information about output distributions without assumptions (e.g. normality)

In this paper, we suggest that one can combine the methods of moment propagation with elementary interval analysis to obtain results that are better than can be obtained from either analysis separately. Rowe (1988) considered the problem of computing moments of certain kinds of transformations such as exp, log, sqrt, etc. from sparse structural information such as first moments and ranges of the operands. We extend this approach to the context of convolutions between poorly characterized random variables, and provide formulae for moment propagation which require no assumptions about stochastic dependence. Rowe's methods, together with the present extension, creates what may be characterized as a *distribution-free* risk analysis that lets analysts compute bounds on uncertain expressions without making assumptions about the precise distributions of the underlying variables. We also show that information about moments actually does enable us to make rigorous conclusions about the shape and, indeed, the percentiles of the output distributions that will be useful in many real-world risk assessments (contra Cullen and Frey 1999, page 184).

2. Means and variances always 'exist'

Mathematically, the distribution of a random variable may fail to have a mean or variance. For instance, Student's t distribution with two degrees of freedom theoretically has no variance because its formula does not converge to a finite value. Similarly, the quotient of independent unit normals, which follows a Cauchy distribution, has neither a variance nor mean. Wiwatandate and Claycamp (2000) suggested that a risk calculation based on simple formulas for means and variances can only be applied in situations where the moments all exist.

As a practical matter, however, we do not consider the nonexistence of moments to be of any real significance for risk analysts. Infinite means and variances are merely mathematical bêtes noires that need not concern the practically minded. All random variables relevant to real-world risk analyses come from bounded distributions. As an example, consider human body weight. There are no infinitely massive body weights (despite recent trends in western dietary health). The largest recorded human body weight was 635 kg. Although a person could probably exceed this weight, perhaps even substantially, there are clearly bounds that human body mass cannot exceed. Therefore, as a practical matter, even a very comprehensive risk analysis need never include a mathematically infinite distribution for body weight. Similar arguments apply to other variables. Analysts concerned with infinite tails of distributions are addressing mathematical problems, not

risk analysis problems. All the moments of any bounded distribution are finite and therefore 'exist' in the mathematical sense.

On the other hand, just because the moments are finite, does not imply they are determinate. In fact, it may usually be the case that only an indeterminate estimate of a mean or variance is available. In such situations, we can use intervals to represent the value, whatever it is, in some range. We can then use interval arithmetic (Moore 1966) to manipulate the estimate and propagate it through calculations even though we cannot specify its value precisely.

3. Propagating range and moment information

In this section, we review formulas for bounds on the range and first two moments (mean and variance) for imprecisely specified random variables. Bounds are considered "rigorous" or "true" bounds if they are certain to contain the value (given the assumptions). All of the formulas in the tables in this paper are rigorous, so the true moments are guaranteed to be inside the given bounds so long as the inputs are within their respective bounds. This means that none of the table entries is merely approximate. Bounds are considered "best possible" if they cannot be any tighter. If a formula in the table is exact or best possible, it is displayed in boldface. Most of the other formulas yield fairly narrow results and are still quite good for practical purposes even though they may not be mathematically best possible.

Table I summarizes formulas that can be used to estimate the least and greatest possible value of a distribution arising from a transformation or convolution. In this and the following tables, X and Y are two random numbers and k is an arbitrary constant. $\underline{X}$ and $\overline{X}$ denote respectively the least and greatest possible value of X. EX denotes the expectation or mean of X, and VX denotes its variance. Following Rowe (1988), we define the variance with a denominator of n instead of $n-1$, and emphasize that the quantities under consideration are moments of finite data populations, which are not necessarily samples of anything. In other respects, the random variables are arbitrary except for restrictions implied by the mathematical operations. For instance, the entries in the square root rows assume X cannot take on negative values, and the rows for division assume that the random variable Y does not straddle zero.

The formulas in Table I are essentially a synopsis of standard interval arithmetic (Moore 1966) and, apart from the row for subtraction perhaps, are probably not very surprising. Monotone increasing transformations are especially easy, because the endpoint of the transformation is just the transformation of the endpoint. For instance, the least possible value of the square root of some variable is simply the square root of the least possible value of the variable. The relevant endpoints are reversed for monotone decreasing transformations. For instance, the greatest possible value of the reciprocal of some variable is the reciprocal of its least possible value. Non-monotone functions, such as absolute value, are more troublesome to account for because values inside the range of the variable can play a role in determining the endpoints of the transformation of the variable. For instance, the least possible value of the absolute value of some variable that ranges between $+2$ and -2 is zero (which is neither endpoint).

The formulas in Table II review the basic arithmetic operations on moments without dependence assumptions. These formulas generally yield intervals rather than precise values. In part, the results

Table I. Rigorous formulas for least and greatest possible values of 9 transformations and 7 convolutions of random variables (all the formulations in this table are mathematically best-possible).

	Least possible value	*Greatest possible value*
$k + X$ (shifting)	$\mathbf{k + \underline{X}}$	$\mathbf{k + \overline{X}}$
kX (rescaling)	$\begin{cases} \mathbf{k\underline{X}}, & \text{if } 0 \le k \\ \mathbf{k\overline{X}}, & \text{if } k < 0 \end{cases}$	$\begin{cases} \mathbf{k\overline{X}}, & \text{if } 0 \le k \\ \mathbf{k\underline{X}}, & \text{if } k < 0 \end{cases}$
e^X	$\mathbf{e^{\underline{X}}}$	$\mathbf{e^{\overline{X}}}$
$\ln(X)$ for $0 < X$	$\mathbf{ln(\underline{X})}$	$\mathbf{ln(\overline{X})}$
$\log_{10}(X)$ for $0 < X$	$\mathbf{log_{10}(\underline{X})}$	$\mathbf{log_{10}(\overline{X})}$
$\frac{1}{X}$ for $0 \notin X$	$\mathbf{1/\overline{X}}$	$\mathbf{1/\underline{X}}$
X^2	$\begin{cases} \mathbf{0}, & \text{if } 0 \in X \\ \mathbf{min(\underline{X}^2, \overline{X}^2)}, & \text{otherwise} \end{cases}$	$\mathbf{max(\underline{X}^2, \overline{X}^2)}$
$\lvert X\rvert$ (absolute value)	$\begin{cases} \mathbf{0}, & \text{if } 0 \in X \\ \mathbf{min}(\lvert\underline{\mathbf{X}}\rvert, \lvert\overline{\mathbf{X}}\rvert), & \text{otherwise} \end{cases}$	$\mathbf{max}(\lvert\underline{\mathbf{X}}\rvert, \lvert\overline{\mathbf{X}}\rvert)$
$\sqrt{X}$ for $0 \le X$	$\mathbf{\sqrt{\underline{X}}}$	$\mathbf{\sqrt{\overline{X}}}$
$X + Y$	$\mathbf{\underline{X} + \underline{Y}}$	$\mathbf{\overline{X} + \overline{Y}}$
$X - Y$	$\mathbf{\underline{X} - \overline{Y}}$	$\mathbf{\overline{X} - \underline{Y}}$
$X \times Y$	$\mathbf{min(\underline{XY}, \underline{X}\overline{Y}, \overline{X}\underline{Y}, \overline{XY})}$	$\mathbf{max(\underline{XY}, \underline{X}\overline{Y}, \overline{X}\underline{Y}, \overline{XY})}$
$\frac{X}{Y}$ for $0 \notin Y$	$\mathbf{min(\underline{X}/\underline{Y}, \underline{X}/\overline{Y}, \overline{X}/\underline{Y}, \overline{X}/\overline{Y})}$	$\mathbf{max(\underline{X}/\underline{Y}, \underline{X}/\overline{Y}, \overline{X}/\underline{Y}, \overline{X}/\overline{Y})}$
X^Y for $0 < X$ or $0 < Y$	$\mathbf{min(\underline{X}^{\underline{Y}}, \underline{X}^{\overline{Y}}, \overline{X}^{\underline{Y}}, \overline{X}^{\overline{Y}})}$	$\mathbf{max(\underline{X}^{\underline{Y}}, \underline{X}^{\overline{Y}}, \overline{X}^{\underline{Y}}, \overline{X}^{\overline{Y}})}$
$\min(X, Y)$	$\mathbf{min(\underline{X}, \underline{Y})}$	$\mathbf{min(\overline{X}, \overline{Y})}$
$\max(X, Y)$	$\mathbf{max(\underline{X}, \underline{Y})}$	$\mathbf{max(\overline{X}, \overline{Y})}$

are indeterminate because we are not specifying the stochastic dependence between the random variables X and Y (this reason is reflected in the occasional appearance of the $\pm$ operator in the table). This indeterminism would be present even if the estimates of means and variances used as inputs were precise. But, of course, these inputs may well start out as intervals, perhaps because they were previously computed using the tabled formulas or because they were imprecisely estimated from statistical data or by subjective judgment.

Some of these formulas, such as those in the first two rows, are elementary and can be found in any textbook on mathematical statistics (e.g. Wilks 1962). (Rowe, 1988) describes several bounds on transformations of random variables that have constant-sign derivatives, including exponentiation, logarithms, reciprocal, square and square root. Rowe showed how to make use of information about the minimum and maximum values to obtain surprisingly tight bounds on the mean and variance with simple closed-form expressions. These expressions do not require approximation and are extremely fast when implemented on a computer. In the table, we use rowe (Rowe's mean estimate) and rowevar (Rowe's variance estimate) to denote his functional templates:

Table II. Rigorous formulas for the mean and variance for 9 transformations and 7 convolutions of random variables (best-possible formulations in boldface).

	Mean	*Variance*
$k + X$ (shifting)	$\mathbf{k + EX}$	$\mathbf{VX}$
kX (rescaling)	$\mathbf{kEX}$	$\mathbf{k^2VX}$
e^X	rowe(exp)	rowevar(exp)
$\ln(X)$ for $0 < X$	rowe(ln)	rowevar(ln)
$\log_{10}(X)$ for $0 < X$	rowe($\log_{10}$)	rowevar($\log_{10}$)
$\frac{1}{X}$ for $0 \notin X$	rowe(reciprocal)	rowevar(reciprocal)
X^2	$\mathbf{(EX)^2 + VX}$	rowevar(square)
$\lvert X\rvert$ (absolute value)	$\begin{cases} \mathbf{EX}, & \text{if } 0 \le \underline{X} \\ -\mathbf{EX}, & \text{if } \overline{X} \le 0 \\ [\lvert EX\rvert, \lvert EX\rvert + \sqrt{VX}(\pi - \operatorname{atan}(\frac{\lvert EX\rvert}{\sqrt{VX}}))], & \text{if } 0 \in X \end{cases}$	$\max(0, EX^2 + VX - E(\lvert X\rvert)^2)$
$\sqrt{X}$ for $0 \le X$	rowe($\sqrt{}$)	rowevar($\sqrt{}$)
$X + Y$	$\mathbf{EX + EY}$	$(\sqrt{\mathbf{VX}} \pm \sqrt{\mathbf{VY}})^\mathbf{2}$
$X - Y$	$\mathbf{EX - EY}$	$(\sqrt{\mathbf{VX}} \pm \sqrt{\mathbf{VY}})^\mathbf{2}$
$X \times Y$	$\mathbf{EXEY} \pm \sqrt{\mathbf{VXVY}}$	"Goodman"
$\frac{X}{Y}$ for $0 \notin Y$	$E(X \times (1/Y))$	$V(X \times (1/Y))$
X^Y for $0 < X$ or $0 < Y$	$E(\exp(\ln(X) \times Y))$	$V(\exp(\ln(X) \times Y))$
$\max(X, Y)$	"Bertsimas max"	$\operatorname{env}(\max(VX, VY), 0)$
$\min(X, Y)$	"Bertsimas min"	$\operatorname{env}(\max(VX, VY), 0)$

$$\text{rowe}(t) = \text{env}(Pt(\underline{X}) + (1-P)t(EX + \frac{VX}{EX - \underline{X}}), Qt(\overline{X}) + (1-Q)t(EX + \frac{VX}{EX - \overline{X}})) \tag{1}$$

$$\text{rowevar}(t) = \text{env}(\frac{t(\underline{\nu}) - t(\underline{X})}{(\underline{\nu} - \underline{X})^2}(VX + (\underline{\nu} - EX)^2), \frac{t(\overline{\nu}) - t(\overline{X})}{(\overline{\nu} - \overline{X})^2}(VX + (\overline{\nu} - EX)^2)) \tag{2}$$

where t denotes one of the transformations exp, ln, $\log_{10}$, square root or reciprocal $(1/X)$, and where *env* denotes the interval envelope:

$$\text{env}(a, b) = [\min(a, b), \max(a, b)] \tag{3}$$

P and Q in equation 1 are:

$$P = 1/(1 + (EX - \underline{X})^2/VX) \tag{4}$$

$$Q = 1/(1 + (EX - \overline{X})^2/VX) \tag{5}$$

and ν in equation 2 is the anti-transformation of the Rowe mean estimate (which generally gives an interval result):

$$\nu = t^{-1}(\text{rowe}(t)) \tag{6}$$

For example, the mean of $\ln(X)$ would be estimated by:

$$\text{env}(P\ln(\underline{X}) + (1-P)\ln(EX + \frac{VX}{EX - \underline{X}}), Q\ln(\overline{X}) + (1-Q)\ln(EX + \frac{VX}{EX - \overline{X}})) \tag{7}$$

and the variance would be estimated by:

$$\text{env}(\frac{\ln(\underline{\nu}) - \ln(\underline{X})}{(\underline{\nu} - \underline{X})^2}(VX + (\underline{\nu} - EX)^2), \frac{\ln(\overline{\nu}) - \ln(\overline{X})}{(\overline{\nu} - \overline{X})^2}(VX + (\overline{\nu} - EX)^2)) \tag{8}$$

where ν is the exp (antilog) of the mean estimate. Thus, if X ranges over $[10, 30]$ and has a mean of 15 and a variance of 3, then the mean of $\ln(X)$ is sure to be within the interval $[2.699, 2.704]$, and a variance sure to be in $[0.006437, 0.02002]$, and has a range of $[2.3025, 3.4012]$. Although these templates are a bit complicated for manual calculation, they are very amenable to implementation on a computer and require only two dozen elementary floating-point operations and four evaluations of the transformation function. Rowe's approach works for all transformations that have constant-sign second derivatives.

Some of the formulae for moment propagation under any dependence are too lengthy to be placed in Table II. We therefore expand them here. The Goodman formula (Goodman, 1960) for the variance of product is:

$$V(XY) = (EX)^2VY + (EY)^2VX + 2EXEYE_{11} + 2EXE_{12} + 2EYE_{21} + E_{22} - E_{11}^2 \tag{9}$$

where E_{ij} are the higher bivariate moments: $E_{ij} = E[(X - EX)^i(Y - EY)^j]$ (e.g. E_{11} is covariance). These are generally not tracked by the method, however they may be expressed in terms of the marginal moments and the other formulae described here:

$$\begin{aligned} E_{11} &= E[(X - EX)(Y - EY)] \\ &= E[XY - XEY - YEX + EXEY] \\ &= E[XY] - EXEY \end{aligned}$$

$$\begin{aligned} E_{21} &= E[(X - EX)^2(Y - EY)] \\ &= E[X^2Y + E[X]^2Y - X^2E[Y] + 2E[X]E[Y]X - 2E[X]XY - E[X]^2E[Y]] \\ &= E[X^2Y] + E[X]^2E[Y] - E[X^2]E[Y] + 2E[X]^2E[Y] - 2E[X]E[XY] - E[X]^2E[Y] \\ &= E[X^2Y] - E[X^2]E[Y] + 2E[X]^2E[Y] - 2E[X]E[XY] \end{aligned}$$

$$\begin{aligned} E_{12} &= E[(X-EX)(Y-EY)^2] \\ &= E[XY^2 + XE[Y]^2 - E[X]Y^2 + 2E[X]E[Y]Y - 2E[Y]XY - E[X]E[Y]^2] \\ &= E[XY^2] + E[X]E[Y]^2 - E[X]E[Y^2] + 2E[X]E[Y]^2 - 2E[Y]E[XY] - E[X]E[Y]^2 \\ &= E[XY^2] - E[X]E[Y^2] + 2E[X]E[Y]^2 - 2E[Y]E[XY] \end{aligned}$$

$$\begin{aligned} E_{22} &= E[(X-EX)^2(Y-EY)^2] \\ &= E[E[X]^2E[Y]^2 - 2E[X]E[Y]^2X + E[Y]^2X^2 - 2E[X]^2E[Y]Y + 4E[X]E[Y]XY \\ &\quad -2E[Y]X^2Y + E[X]^2Y^2 - 2E[X]XY^2 + X^2Y^2] \\ &= E[X]^2E[Y]^2 - 2E[X]^2E[Y]^2 + E[Y]^2E[X^2] - 2E[X]^2E[Y]^2 + 4E[X]E[Y]E[XY] \\ &\quad -2E[Y]E[X^2Y] + E[X]^2E[Y^2] - 2E[X]E[XY^2] + E[X^2Y^2] \\ &= -3E[X]^2E[Y]^2 + E[X^2]E[Y]^2 + E[X]^2E[Y^2] + 4E[X]E[Y]E[XY] \\ &\quad -2E[Y]E[X^2Y] - 2E[X]E[XY^2] + E[X^2Y^2] \end{aligned}$$

The Bertsimas (Bertsimas et al, 2006) formulae for the expectation of max is:

$$E[\max(X,Y)] = \text{env}(\max(EX,EY), \max(\overline{X},\overline{Y})) \cap (EX + EY - \text{env}(\min(EX,EY), \min(\underline{X},\underline{Y}))) \tag{10}$$

and min being:

$$E[\min(X,Y)] = \text{env}(\min(EX,EY), \min(\underline{X},\underline{Y})) \cap (EX + EY - \text{env}(\max(EX,EY), \max(\overline{X},\overline{Y}))) \tag{11}$$

3.1. Independence need not be assumed (but can be)

Unlike the formulations usually given for moments of the sums, products, quotients, etc. of random variables (e.g., Wiwatandate and Claycamp 2000), the formulas in Table II do *not* assume that X and Y are stochastically independent. Our formulas are guaranteed to give correct results whenever their inputs enclose the respective extremes, means and variances. However, if an analyst is willing to assume independence, then the formulas in Table II can be improved substantially. Table III gives the preferred formulas for such cases. We hasten to point out that an independence assumption is extremely strong, and it is very widely abused in risk analysis. Some uses of the assumption border on the ridiculous, such as the assumption that body weight and skin surface area are independent, or the assumption, echoed even in the paper of Wiwatandate and Claycamp (2000), that body mass and height are independent.

Analysts should take care to use assumptions of independence and the formulas of Table III only when justified by theoretical argument or comprehensive empirical information. In contrast,

Table III. Improved formulas for the mean and variance for convolutions of random variables under an assumption of stochastic independence (best-possible formulations in boldface).

	Mean	*Variance*
$X + Y$	$\mathbf{EX + EY}$	$\mathbf{VX + VY}$
$X - Y$	$\mathbf{EX - EY}$	$\mathbf{VX + VY}$
$X \times Y$	$\mathbf{EXEY}$	$\mathbf{(EX)^2VY + (EY)^2VX + VXVY}$
$\frac{X}{Y}$ for $0 \notin Y$	$E(X \times (1/Y))$	$V(X \times (1/Y))$
X^Y for $0 < X$ or $0 < Y$	$E(\exp(\ln(X) \times Y))$	$V(\exp(\ln(X) \times Y))$
$\max(X, Y)$	$\begin{cases} \mathbf{EX}, & \text{if } Y < X \\ \mathbf{EY}, & \text{if } X < Y \\ \text{“Bertsimas max”}, & \text{otherwise} \end{cases}$	$\begin{cases} \mathbf{VX}, & \text{if } Y < X \\ \mathbf{VY}, & \text{if } X < Y \\ \text{env}(\max(VX, VY), 0), & \text{otherwise} \end{cases}$
$\min(X, Y)$	$\begin{cases} \mathbf{EX}, & \text{if } X < Y \\ \mathbf{EY}, & \text{if } Y < X \\ \text{“Bertsimas min”}, & \text{otherwise} \end{cases}$	$\begin{cases} \mathbf{VX}, & \text{if } X < Y \\ \mathbf{VY}, & \text{if } Y < X \\ \text{env}(\max(VX, VY), 0), & \text{otherwise} \end{cases}$

the formulas of Tables I and II are appropriate for all situations and need not be justified by special argument or evidence.

3.2. Using the formulas with interval inputs

Even if one starts out with point estimates for means and variances, applying the formulas in the tables generally yields interval results. Thus, if one must propagate uncertainty through multiple arithmetic operations, one needs to be able to handle interval estimates for the moments. The above formula can be readily evaluated with intervals for EX and VX and will surely bound the transformed mean and variances; however the tightness of the result depends on the number times a variable appears in the expression. If the variables appears just once, then the result is tightest possible. But if variables appear multiple times in an expression (such as in the variance of the product in Table III), then the interval result will be artificially inflated. This is the well known *repeated variables problem*, and has several numerical solutions such as significance arithmetic (Hyman, 1982), affine arithmetic (Rump and Kashiwagi, 2015), Taylor models (Makino, 1998) and more recently zone arithmetic (Gray, et al 2021). Where possible, expressions can be rearranged in such a way that the variables appear only once, for example realising that $a^2 + a = (a + \frac{1}{2})^2 - \frac{1}{4}$. This process may be automated by an uncertainty compiler, as suggested by Gray et al (2019).

A simple-to-implement solution (although more computationally expensive than the above suggestions) is *sub-intervalisation*, where the interval is split into n (usually linearly spaced) sub-intervals, and the expression is evaluated n times with each sub-interval. The resulting range is then the union of the propagated sub-intervals. Usually the main drawback from this method is that it suffers from the curse of dimensionality, that is if a function has m inputs, then n^m interval

calculations are required. However, since the expressions proposed in this paper only require 2 variables (EX and VX) to be sub-intervalised, this is an appropriate technique here.

4. What do the range and moments say about risks?

What does knowing something about the mean and variance of a random number tell us about the probability distribution of that variable? Generally, people expect that it is unlikely for a random value to be many standard deviations away from the mean. But what exactly is the chance of being, say, 5 standard deviations (or more) larger than the mean? If we assume the underlying distribution is standard normal, the risk is roughly 1 in 3.5 million. Such a value seems very small and might be considered an acceptable risk by planners and decision makers.

But what can one say about such risks *without assuming normality*? What inferences can be drawn about the risks of exceedance that are free of assumptions about the particular shape of the distribution? This question was posed by Chebyshev (1874) and answered by Markov (1887) for the case when only the mean and variance are known. The answer we need for risk analysis is embodied in a version of the classical Chebyshev inequality (Feller 1968, page 152; Allen 1990, page 79). The upper bound on the probability that the variable X will exceed a value as large as x is:

$$Prob(x \leq X) \leq \begin{cases} 1/(1+(x-EX)^2/VX), & \text{if } EX < x \\ 1, & \text{if } x \leq EX \end{cases} \tag{12}$$

where EX and VX are the mean and variance of X. The lower bound on the same probability is:

$$Prob(x \leq X) \geq \begin{cases} 1/(1+VX/(x-EX)^2), & \text{if } x < EX \\ 1, & \text{if } EX \leq x \end{cases} \tag{13}$$

If we use the Chebyshev inequality to ask how large the chance might be without any assumption about the shape of the underlying distribution (with mean 0 and variance 1 at 5 standard deviations), we find it is somewhere between zero and $1/(1+(5-0)^2/1) = 0.03846$, or 1 in 26. Omitting the normality assumption causes the risk to go from 0.000000286 to almost $[0, 0.04]$, which represents a potential risk increase of over five orders of magnitude. What engineer designing a safety system for a nuclear power plant, or for that matter, the razor burn guard on an electric shaver, would be happy with a potential risk of 1 in 26?

The Chebyshev bounds can be tightened substantially in some cases by the addition of knowledge about one endpoint of the range, i.e., either the minimum or the maximum of the underlying distribution. This improvement is expressed in the classical Cantelli inequalities, which give rigorous and best possible bounds on the distribution function for a nonnegative random variable X having mean EX and variance VX. The Cantelli inequalities are a combination of the Markov and Chebyshev inequalities. The upper bound on the probability that the variable X will be no larger than a value x is:

$$Prob(x \le X) \le \begin{cases} 0, & \text{if } x \le 0 \\ 1/(1+(x-EX)^2/VX), & \text{if } 0 \le x \le EX \\ 1, & \text{if } EX < x \end{cases} \quad (14)$$

This function forms the left side of a p-box for X. The right side is the lower bound on the same probability, which is:

$$Prob(x \le X) \ge \begin{cases} 0, & \text{if } x \le EX \\ 1 - EX/x, & \text{if } EX \le x \le EX + VX/EX \\ 1/(1+VX/(x-EX)^2), & \text{if } EX + VX/EX < x \end{cases} \quad (15)$$

If the minimum value of X is not zero, we can encode the information in a new variable Y whose minimum value is zero with the transformations:

$$\begin{aligned} Y &= X - \underline{X}, \\ EY &= EX - \underline{X}, \\ VY &= VX, \end{aligned}$$

then apply the inequalities to obtain the p-box for Y, and finally back-transform this p-box to get the bounds in terms of the original variable X by adding LX to it. If it is the maximum, rather than the minimum that is known, we can use the encoding:

$$\begin{aligned} Z &= -X \\ EZ &= -EX, \\ VZ &= VX, \end{aligned}$$

then apply the inequalities (possibly also encoding to make the new minimum zero), and finally negate the resulting Z p-box to reexpress it in terms of the original variable.

Using the above formulation, it is possible to construct a p-box using a minimum, mean and variance and one using the maximum, mean and variance. A p-box using both endpoints, mean and variance can thus be found by intersecting these two p-boxes. Figure 1. shows this for range $= [0, 6]$, $EX = 3$ and $VX = 5$.

5. Conclusions

The limitations of linearity and independence mentioned by Cullen and Frey are real and serious, but they can be relaxed. In this paper we bring the following extensions to moment propagation:

– Independence between variables need not be assumed.

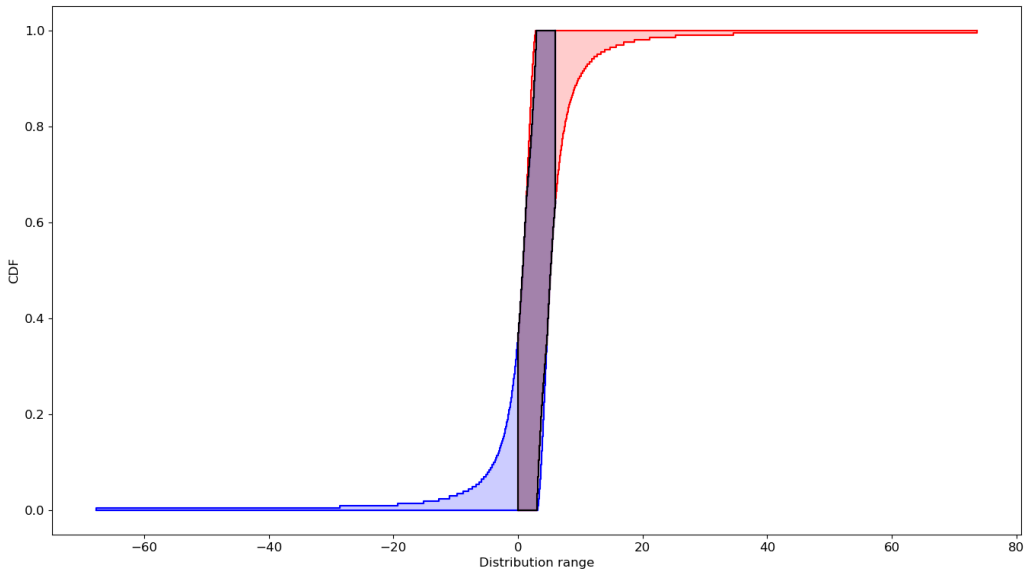


Figure 1. Shows a p-box a constructed using minimum, maximum, mean and variance information (black). Constructed by intersecting the p-box using Cantelli's inequalities with minimum, mean and variance (red) and maximum, mean and variance(blue)

- Moment propagation formulae may be evaluated with intervals.
- Assumptions about input distributions is no longer necessary.
- Distributional information may be obtained from moment and range information in models other than linear.

One important application of the methods to be developed in this paper is to the area of risk analysis. In this discipline, predictions are made about the magnitudes or probabilities of structural failures or other adverse extreme events such as patients receiving toxic doses of therapeutic drugs or endangered species going extinct. These forecasts are often computed from limited empirical information. In traditional "worst case" analyses, the elementary methods of interval analysis are applied to risk formulations estimating, for instance, the difference between a structure's strength and some stress acting on it, or the delivered dose of a drug, or the population size of the endangered species, etc. In this paper, we provide convenient tables for moment propagation formulae for the independence case as well as the case with no knowledge about dependence, and we suggest that one can combine the methods of moment propagation with elementary interval analysis to obtain results that are better than can be obtained from either analysis separately. We provide a method for bounding distributional information solely from moments and ranges, allowing for

risks of extreme values to be calculated in models other than linear and without assumptions about input distributions, as is required by standard moment propagation practices.

Acknowledgements

The authors would like to thank the gracious support from the EPSRC iCase studentship award 15220067. We also gratefully acknowledge funding from UKRI via the EPSRC and ESRC Centre for Doctoral Training in Risk and Uncertainty Quantification and Management in Complex Systems. This research was supported by the EPSRC through grant EP/R006768/1, which is acknowledged for its funding and support. This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

References

Allen, A.O. Probability, Statistics and Queueing Theory with Computer Science Applications, second edition., Academic Press, Boston (1990)

Bertsimas, D., Natarajan, K., Teo, C.P., Tight bounds on expected order statistics, Probability in the Engineering and Informational Sciences 20 (4) (2006) 667.

Bogomolov, S., Forets, M., Frehse, G., Potomkin, K., Schilling, C., Juliareach: a toolbox for set-based reachability, in: Proceedings of the 22nd ACM International Conference on Hybrid Systems: Computation and Control, 2019, pp. 39–44.

Chebyshev, P. L., Sur les valeurs limites des intégrales, Imprimerie de Gauthier-Villars, 1874.

Cullen, A. C., Frey, H. C., Probabilistic techniques in exposure assessment: a handbook for dealing with variability and uncertainty in models and inputs, Springer Science & Business Media, 1999.

Feller, W. An Introduction to Probability Theory and Its Applications. Volume 2, Second edition., John Wiley Sons, New York (1968)

Goodman, L. A., On the exact variance of products, Journal of the American statistical association 55 (292) (1960) 708–713.

Gray, A. and De Angelis, M., Ferson, S., and Patelli, E., What's Z X, when $Z = X + Y$? Dependency tracking in interval arithmetic with bivariate sets, Reliable Engineering Computing (2021).

Gray, N. and De Angelis, M. and Ferson, S., *Computing with Uncertainty: introducing Puffin the Automatic Uncertainty Compiler*. UNCECOMP, 2019.

Hyman, J. M., Forsig: an extension of fortran with significance arithmetic, Tech. rep., Los Alamos National Lab., NM (USA) (1982).

Makino, K., Rigorous analysis of nonlinear motion in particle accelerators. PhD thesis, Michigan State University, 1998.

Markoff, A., et al., Sur une question de maximum et de minimum: Proposée par m. tchebycheff, Acta Mathematica 9 (1887) 57–70.

Moore, R. E., Interval analysis, Vol. 4, Prentice-Hall Englewood Cliffs, 1966.

Rowe, N. C., Absolute bounds on the mean and standard deviation of transformed data for constant-sign-derivative transformations, SIAM journal on scientific and statistical computing 9 (6) (1988) 1098–1113.

Rump, S. M., Kashiwagi, M., Implementation and improvements of affine arithmetic, Nonlinear Theory and Its Applications, IEICE 6 (3) (2015) 341–359.

Wilks, S. S., Mathematical statistics. new york: Johnwiley & sons', Inc. Wilks Mathematical Statistics (1962).

Wiwatanadate, P., Claycamp, H. G., Exact propagation of uncertainties in multiplicative models, Human and Ecological Risk Assessment 6 (2) (2000) 355–368.

Fusion of Probabilistic Knowledge as Foundation for Sliced-Normal Approach

Michael Beer[1], Olga Kosheleva[2], and Vladik Kreinovich[3]
[1]*Institute for Risk and Reliability, Leibniz University Hannover*
30167 Hannover, Germany, beer@irz.uni-hannover.de
Departments of [2]*Teacher Education and* [3]*Computer Science*
University of Texas at El Paso
El Paso, TX 79968, USA, {olgak,vladik}@utep.edu

Abstract. In many practical applications, it turns out to be efficient to use Sliced-Normal multi-D distributions, i.e., distributions for which the logarithm of the probability density function (pdf) is a polynomial – – to be more precise, it is a sum of squares of several polynomials. This class is a natural extension of normal distributions, i.e., distributions for which the logarithm of the pdf is a quadratic polynomial.

In this paper, we provide a possible theoretical explanation for this empirical success. Specifically, we show that if a family of distributions is closed under information fusion, is scale- and shift-invariant, and has the property that every piece of knowledge can be obtained by fusing small pieces of information, then all distributions from this family are Sliced-Normal multi-D distributions.

Keywords: sliced-normal distribution, fusion of probabilistic knowledge, scale-invariance, shift-invariance

1. Formulation of the Problem

Sliced-normal distributions are efficient. In many practical applications, it turns out to be efficient to use Sliced-Normal multi-D distributions, i.e., distributions for which the logarithm of the probability density function (pdf) $f(x_1, \ldots, x_n)$ is a polynomial (to be more precise, it is a sum of squares of several polynomials); see, e.g., (Colbert, Crespo, and Peet, 2019; Crespo, 2019; Crespo, Colbert, Kenny, and Giesy, 2019):

$$\ln(f(x_1, \ldots, x_n)) = P(x_1, \ldots, x_n) \tag{1}$$

for some polynomial $P(x_1, \ldots, x_n)$, so

$$f(x_1, \ldots, x_n) = \exp(P(x_1, \ldots, x_n)). \tag{2}$$

This class is a natural extension of normal distributions, i.e., distributions for which the logarithm of the pdf is a quadratic polynomial; see, e.g., (Sheskin, 2011).

But why? This is what we try to explain. While the sliced-normal distributions have been empirically successful, there seems to be no convincing theoretical explanation for their empirical success. The main goal of this paper is to provide such an explanation.

© 2021 *by M. Beer, O. Kosheleva, and V. Kreinovich.*

2. Let Us Formulate This Problem in Precise Terms

Need for a finite-parametric family. In principle, we can have many different probability density functions. The class of all the functions is infinite-dimensional – which means that, to select a single probability density function out of all possible such functions, we need to know the values of infinitely many parameters (e.g., values of the pdf at points with rational coordinates).

In practice, however, at any given moment of time, we only have finitely many observations. Based on these observations, we can determine only finitely many parameters. Thus, it makes sense to looks for families F of probability density functions that depend on finitely many parameters $c_1, \ldots, c_m$, i.e., on families of the type $F = \{f(x_1, \ldots, x_n, c_1, \ldots, c_m)\}_{c_1,\ldots,c_m}$.

The dependence should be continuous. All our information about the physical world comes from measurements and from expert estimates. Measurements are never 100% accurate (see, e.g., (Rabinovich, 2005)), expert estimates are even less accurate. So, we can only determine the values x_i and c_j with some accuracy. Based on these approximate values of x_i and c_j, we should make estimates of the corresponding values $f(x_1, \ldots, x_n, c_1, \ldots, c_m)$ of the probability density – and the more accurately we perform measurements, the more accurate should be our estimates.

In mathematical terms, this means that the dependence of the function $f(x_1, \ldots, x_n, c_1, \ldots, c_m)$ on all its $n + m$ inputs x_i and c_j should be continuous.

Moreover, small inaccuracy in x_i and c_j should lead to proportionally small inaccuracy in the resulting value of $f(x_1, \ldots, x_n, c_1, \ldots, c_m)$. Thus, the function $f(x_1, \ldots, x_n, c_1, \ldots, c_m)$ should be *differentiable*. Thus, we arrive at the following definition.

Definition 1. *Let n and m be positive integers. By an m-*parametric family of probability density functions on $\mathbb{R}^n$ *(or simply a* family*, for short), we mean a differentiable function $f(x_1, \ldots, x_n, c_1, \ldots, c_m)$ of $n+m$ variables for which, for each tuple $(c_1, \ldots, c_m)$, the corresponding function $x_1, \ldots, x_n \to f(x_1, \ldots, x_n, c_1, \ldots, c_m)$ is a probability density function, i.e.:*

- *we have $f(x_1, \ldots, x_n, c_1, \ldots, c_m) \geq 0$ for all x_i and c_j, and*
- *we have $\int f(x_1, \ldots, x_n, c_1, \ldots, c_m)\, dx_1 \ldots dx_n = 1$ for all tuples $(c_1, \ldots, c_m)$.*

3. The Class of Distributions Should Be Closed Under Fusion

Need for fusion. The very fact that we only know the *probability* of different tuples $x = (x_1, \ldots, x_n)$ means that we do not know which of the tuples describe the corresponding real-life situation. In other words, the fact that we have a probabilistic knowledge means that our knowledge is incomplete. It is therefore desirable to gain additional knowledge about the situation – either by performing additional measurements, or by requesting additional expert estimates.

This additional knowledge usually comes in the form of a probability distribution. Once we have this probability distribution, we need to fuse it with the distribution describing our original knowledge.

The class of distributions should be closed under fusion. The main objective in selecting a finite-parametric family of distributions is to come up with a reasonable family, a family that

describes reasonably well all possible states of our knowledge. From this viewpoint, it is reasonable to require that:

- if both fused pieces of knowledge are described by distributions from our family,
- then the result of fusing these two pieces of knowledge should also be described by distributions from our family.

In mathematical terms, this means that the desired family of probability distributions should be *closed* under fusion.

To describe this requirement in precise terms, let us describe fusion in precise terms.

How to describe fusion: a natural idea. In probability theory, if we have two independent events with probabilities p_1 and p_2, then the probability that both events will happen is equal to the product of these probabilities. Similarly, if we have two independent sources of information, so that:

- based on the information from the first source, we assign, to each of N alternatives $a_1, \ldots, a_N$, the probabilities $p_{11}, \ldots, p_{1N}$,
- based on the information from the second source, we assign, to each of N alternatives $b_1, \ldots, b_N$, the probabilities $p_{21}, \ldots, p_{2N}$,

then the probability that we have alternative a_i in the first case and alternative b_j in the second case is equal to the product of the corresponding probabilities $p_{1i} \cdot p_{2j}$.

If it turns out that in both cases, we have the exact same set of alternatives, then we need to consider *conditional* probabilities, namely probabilities under the condition that $i = j$. In general, the conditional probability $P(A \,|\, B)$ of an event A under the condition B can be obtained by dividing the probability $P(A \,\&\, B)$ of $A \,\&\, B$ by the probability $P(B)$ that the condition B is satisfied. In our case, this means that after the fusion, the probability of the i-th alternative is equal to $p_i = C \cdot p_{1i} \cdot p_{2i}$, where the coefficient $C \stackrel{\text{def}}{=} \dfrac{1}{P(B)}$ can be obtained from the requirement that the resulting probabilities add up to 1, i.e., that $\sum\limits_{i=1}^{N} p_i = C \cdot \sum\limits_{i=1}^{N} p_{1i} \cdot p_{2i} = 1$, so that $C = \dfrac{1}{\sum\limits_{i=1}^{N} p_{1i} \cdot p_{2i}}$.

Similar formulas can be obtained for continuous distributions: if we have two independent sources of information that lead to distributions $f_1(x)$ and $f_2(x)$, then the fusion of these two pieces of information is a probability distribution $f(x) = C \cdot f_1(x) \cdot f_2(x)$, where C is a normalization constant selected so as to guarantee that $\int f(x)\,dx$, i.e., $C = \dfrac{1}{\int f_1(x) \cdot f_2(x)\,dx}$.

Similarly, we can define the result of fusing several probability distributions.

Definition 2. *Let $f_1(x), \ldots, f_k(x)$ be probability density functions (pdfs) on $\mathbb{R}^n$. By the* result of fusing *these pfds, we mean a pdf $f(x) = C \cdot f_1(x) \cdot \ldots \cdot f_k(x)$, where $C = \dfrac{1}{\int f_1(x) \cdot \ldots \cdot f_k(x)\,dx}$.*

Definition 3. *We say that the family $f(x,c)$ is* closed under fusion *if for every k pdfs $f\left(x,c^{(1)}\right),\ldots,f\left(x,c^{(k)}\right)$ from this family, the result of fusing these pdfs also belongs to the same family, i.e., has the form $f(x,c)$ for some tuple c.*

4. Every Piece of Knowledge Can Be Obtained by Fusing Several "Smaller" Pieces of Information

Main idea. Sometimes, knowledge comes in one big step. However, more typically, to gain the knowledge, we must acquire it piece by piece, sometimes in two steps, sometimes in three steps, sometimes in four steps, etc. So, it is natural to come up with the following definition.

Definition 4. *We say that in a family $f(x,c)$,* every piece of knowledge can be obtained by fusing small pieces of information *if for every pdf $f(x,c)$ from this family and for every integer $M\geq 2$, there exists another pdf $f(x,c')$ from this family so that fusing M copies of $f(x,c')$ leads to $f(x,c)$.*

5. The Family of Distributions Should Not Depend on the Choices of Starting Points and Measuring Units for x_i

Possibility to change measuring unit and a starting point. We want to deal with physical quantities, but in reality, we deal with their numerical values. These numerical values depend on what measuring unit we use for measuring the quantity, and what starting point we select for this measurement. When we change the measuring unit and/or the starting point, the numerical values change.

For example, if we change the measuring unit from meters to centimeters, all the numerical values are multiplied by 100, so that, 2 m becomes 200 cm. In general, if we change from the original measuring unit to a new one which is a times smaller, then all the numerical values are multiplied by a: $x\to a\cdot x$. This transformation is known as *scaling*.

Similarly, if we change the starting point to the one which is b units before – as we can do for time, temperature, and many other quantities – then b is added to all the numerical values $x\to x+b$. This transformation is known as *shift*. A shift can also be viewed as a kind of re-scaling.

If we change both the measuring unit and the starting point, then numerical value change as $x\cdot a\cdot x+b$. These transformations change the pdf: if we had a pdf $f(x_1,\ldots,x_n)$, and we apply such transformation $x_i\to x_i'\stackrel{\text{def}}{=}a_i\cdot x_i+b_i$ to each of inputs, then in terms of the new numerical values $x_1',\ldots,x_n'$, the corresponding pdf takes a different form.

Definition 5. *Let $f(x_1,\ldots,x_n)$ be a pdf, let $a=(a_1,\ldots,a_n)$ be a tuple of positive numbers, and let $b=(b_1,\ldots,b_n)$ be a tuple of real numbers. By a (a,b)-*re-scaling *of the pdf f, we mean a pdf*

$$f'(x_1',\ldots,x_n')=\frac{1}{\prod\limits_{i=1}^{n}a_i}\cdot f\left(\frac{x_1'-b_1}{a_1},\ldots,\frac{x_n'-b_n}{a_n}\right). \tag{3}$$

A natural invariance requirement. We want to come up with a universal family of probability distributions, a family that would be applicable no matter what measuring units and what starting points we select for all the inputs. Thus, it is reasonable to require that our family is invariant with respect to the corresponding transformations.

Definition 6. *We say that a family F is* scale- and shift-invariant *if every pdf $f(x, c)$ from this family and for every two tuples a and b, the (a, b)-re-scaling of the pdf $f(x, c)$ also belongs to the family F.*

Now, we are ready for formulate and prove our main result.

6. Main Result

Proposition. *For every family F:*

- *which is closed under fusion,*
- *for which every piece of knowledge can be obtained by fusing small pieces of information, and*
- *which is scale- and shift-invariant,*

there exists an integer $d \le m + 1$ such that every probability density function from this family has the form $f(x, c) = \exp(P(x_1, \ldots, x_n))$ for some polynomial $P(x_1, \ldots, x_n)$ of degree $\le d$ with respect to each of its variables.

Comment. This result explains the empirical success of sliced-normal distributions.

Proof.

1°. Let F be the family that satisfies all the conditions described in the formulation of the Proposition. By a *log-function*, we will mean a function of the type $L(x, c, s) = \ln(f(x, c)) + s$ for some tuple c and some real number s. Let us denote the class of all log-functions by $\mathcal{L}$.

2°. Let us prove that the class of all log-functions is closed under addition, i.e., that for every two log-functions $L(x, c;, s')$ and $L(x, c'', s'')$, their sum is also a log-function.

Indeed, by definition, $L(x, c', s') = \ln(f(x, c')) + s'$ and $L(x, c'', s'') = \ln(f(x, c'')) + s''$. Since the family F is closed under fusion, the result of fusing the corresponding pdfs is also a pdf from the same family, i.e., $C \cdot f(x, c') \cdot f(x, c'') = f(x, c)$ for some tuple c. By taking logarithms of both sides of this equality, we conclude that

$$\ln(C) + \ln(f(x, c')) + \ln(f(x, c'')) = \ln(f(x, c)). \tag{4}$$

If we add $s' + s'' - \ln(C)$ to both sides of the resulting equality, we conclude that

$$(\ln(f(x, c')) + s') + (\ln(f(x, c'')) + s'') = \ln(f(x, c)) + (s' + s'' - \ln(C)), \tag{5}$$

i.e., that the sum of the two given log-functions is indeed a log-function:

$$L(x, c', s') + L(x, c'', s'') = L(x, c, s' + s'' - \ln(C)). \tag{6}$$

The statement is proven.

3°. Let us now prove that for each log-function $L(x, c, s)$ and for every integer $M \geq 2$, the function $M^{-1} \cdot L(x, c, s)$ is also a log-function.

By definition, $L(x, c, s) = \ln(f(x, c)) + s$. Since for the family F, every piece of knowledge can be obtained by fusing small pieces of information, we conclude that the pdf $f(x, c)$ can be obtained by fusing M instances of some other pdf $f(x, c')$, i.e., that $f(x, c) = C \cdot (f(x, c'))^M$. By taking logarithms of both sides of this equality, we get $\ln(f(x, c)) = M \cdot \ln(f(x, c')) + \ln(C)$, thus

$$M^{-1} \cdot \ln(f(x, c)) = \ln(f(x, c')) + M^{-1} \cdot \ln(C). \tag{7}$$

By adding $M^{-1} \cdot s$ to both sides, we get

$$M^{-1} \cdot (\ln(f(x, c)) + s) = \ln(f(x, c')) + M^{-1} \cdot (\ln(C) + s). \tag{8}$$

The left-hand side of this formula is exactly $M^{-1} \cdot L(x, c, s)$, and the right-hand side is a log-function. So, the statement is proven.

4°. Let us now consider the closure $\mathcal{C}$ of the set $\mathcal{L}$ of all log-functions – the closure in the usual topological sense, i.e., the set of all limit functions with respect to some natural topology on the class of all differentiable functions. Since the set $\mathcal{L}$ is closed under addition, its closure $\mathcal{C}$ is also closed under addition.

Let us prove that this closure is closed under multiplication by positive numbers. In other words, let us prove that for each function $f(x) \in \mathcal{C}$ and for every positive real number $r > 0$, the function $r \cdot f(x)$ also belongs to $\mathcal{C}$. Since $\mathcal{C}$ is the closure of the set of all log-functions, it is sufficient to prove that for each log-function $L(x, c, s)$ and for every positive real number $r > 0$, the function $r \cdot L(x, c, s)$ is a limit of log-functions.

Indeed, for every possible accuracy $\varepsilon > 0$, we can approximate, with this accuracy, the real number r by a rational number $\dfrac{N}{M}$. By Part 3 of this proof, the function $M^{-1} \cdot L(x, c, s)$ is also a log-function. Now, by Part 2 of this proof, the function $\dfrac{N}{M} \cdot L(x, c, s)$ is also a log-function – as the sum of N log-functions $M^{-1} \cdot L(x, c, s)$. When $\dfrac{N}{M}$ tends to r, the corresponding function $\dfrac{N}{M} \cdot L(x, c, s)$ tends to $r \cdot L(x, c, s)$. Thus, the function $r \cdot L(x, c, s)$ is indeed a limit of log-functions. The statement is proven.

5°. By combining Parts 2 and 4, we conclude that for every finite set of functions $C_1(x), \ldots, C_k(x)$ from the set $\mathcal{C}$, and for every tuple of positive numbers $r_1, \ldots, r_k$, the linear combination

$$r_1 \cdot C_1(x) + \ldots + r_k \cdot C_k(x) \tag{9}$$

also belongs to $\mathcal{L}$.

6°. Let us now prove that the set $\mathcal{C}$ cannot contain more than $m+1$ linearly independent functions.

Indeed, if this was the case, and we would have more than $m+1$ linearly independent functions, then we would have at least $m+2$ of them $C_1(x), \ldots, C_{m+2}(x)$ in the class $\mathcal{C}$. Then, due to Part 5 of this proof, the class $\mathcal{C}$ will contain a $(m+2)$-parametric family of functions

$$r_1 \cdot C_1(x) + \ldots + r_{m+2} \cdot C_{m+2}(x) \tag{10}$$

of different functions. However, the class $\mathcal{C}$ is the closure of the class $\mathcal{L}$ of functions of the type $\ln(f(x,c)) + s$ that depend on $m+1$ parameters:

- we have m parameters $c_1, \ldots, c_m$ and
- we have an additional parameter s.

So, the closure of this set is also of dimension $m+1$ (or less) – and thus, cannot contain more-dimensional subfamilies. The statement is proven.

7°. Let us denote by $\mathcal{S}$ the class of all linear combinations of functions from the class $\mathcal{C}$. Clearly, $\mathcal{C} \subseteq \mathcal{S}$, and, due to Part 6, the dimension d of the linear space $\mathcal{S}$ cannot exceed $m+1$. So, if we pick any basis $e_1(x), \ldots, e_d(x)$ in this class, then each function $f(x)$ from the class $\mathcal{S}$ can be represented as a linear combination of functions from this basis: $f(x) = C_1 \cdot e_1(x) + \ldots + C_d \cdot e_d(x)$, for some values $C_1, \ldots, C_d$.

We can pick the basis from the set $\mathcal{C}$. Moreover, since the closure does not change the dimension, we can pick it from the original class $\mathcal{L}$ of log-functions. All the pdf functions from the family F are, by definition of a family, differentiable. Thus, every log-function is also differentiable. Hence, we can choose the basis of differentiable functions.

8°. Let us prove that the class $\mathcal{L}$ is closed under arbitrary re-scalings, i.e., if a function $f(x_1, \ldots, x_n)$ is in this class, then for each tuple $a = (a_1, \ldots, a_n)$ of positive numbers and for each tuple $b = (b_1, \ldots, b_n)$ of real numbers, the function $f(a \cdot x_1 + b_1, \ldots, a_n \cdot x_n + b_n)$ also belongs to the class $\mathcal{L}$.

This follows from the requirement that the family F is scale- and shift-invariant, if we take logarithms of both sides and add appropriate constants to both sides.

9°. From Part 8, we can conclude that the closure class $\mathcal{C}$ is also invariant with respect to arbitrary re-scalings. Thus, the class $\mathcal{S}$ of all linear combinations of functions from $\mathcal{C}$ is also thus invariant.

10°. Let us first study the consequences of shift-invariance of the class $\mathcal{S}$ with respect to the first variable. This shift-invariance implies, in particular, that for each basis function $e_i(x_1, x_2, \ldots, x_n)$, the result of its shift $e_i(x_1 + b_1, x_2, \ldots, x_n)$ is also a function from the class $\mathcal{S}$, i.e., that

$$e_i(x_1 + b_1, x_2, \ldots, x_n) = \sum_{j=1}^{d} C_{ij}(b_1) \cdot e_j(x_1, x_2, \ldots, x_n), \tag{11}$$

for some coefficients C_{ij} that, in general, depend on b_1.

For a while, let us fix the values $x_2, \ldots, x_n$ and only consider the dependence on x_1. In other words, let us consider auxiliary functions $E_i(x_1) \stackrel{\text{def}}{=} e_i(x_1, x_2, \ldots, x_n)$. For these auxiliary functions, the above formula takes the form

$$E_1(x_1 + b_1) = C_{11}(b_1) \cdot E_1(x_1) + \ldots + C_{1d}(b_1) \cdot E_d(x_1);$$
$$\ldots \tag{12}$$
$$E_d(x_1 + b_1) = C_{d1}(b_1) \cdot E_1(x_1) + \ldots + C_{dd}(b_1) \cdot E_d(x_1).$$

Here, all the functions $E_1(x_1) \ldots, E_d(x_1)$ are differentiable – since they come by fixing some values from the basis functions $e_i(x_1, \ldots, x_n)$, and the basis functions are differentiable.

Let us prove that the dependencies $C_{ij}(b_1)$ are also differentiable. Indeed, for each i, let us pick d different values $x_{11}, \ldots, x_{1d}$ of x_1, then we get the following d linear equations for d unknowns $C_{i1}(b_1), \ldots, C_{in}(b_1)$:

$$E_i(x_{11} + b_1) = C_{i1}(b_1) \cdot E_1(x_{11}) + \ldots + C_{id}(b_1) \cdot E_d(x_{11});$$
$$\ldots \tag{13}$$
$$E_i(x_{1d} + b_1) = C_{i1}(b_1) \cdot E_1(x_{1d}) + \ldots + C_{id}(b_1) \cdot E_d(x_{1d}).$$

Each element $C_{ij}(b_1)$ solution to a system of linear equations can be described, by using the Cramer rule, as the ratio of two determinants, i.e., as a smooth function of all the coefficients. Since the coefficients $E_i(x_{1k} + b_1)$ smoothly depend on b_1, we conclude that the solutions $C_{ij}(b_1)$ are also differentiable functions of b_1.

Since all the functions $E_i(x_1)$ and $C_{ij}(b_1)$ are differentiable, we can differentiate both sides of all equalities describing $E_i(x_1 + b_1)$ with respect to b_1, and take $b_1 = 0$. Then, we get the following system of equations:

$$E_1'(x_1) = c_{11} \cdot E_1(x_1) + \ldots + c_{1d} \cdot E_d(x_1);$$
$$\ldots \tag{14}$$
$$E_d'(x_1) = c_{d1} \cdot E_1(x_1) + \ldots + c_{dd} \cdot E_d(x_1),$$

where $E_i'(x_1)$ denotes the derivative, and $c_{ij} \stackrel{\text{def}}{=} C_{ij}'(0)$.

In other words, for the functions $E_1(x), \ldots, E_d(x)$, we get a system of linear differential equations with constant coefficients. It is known that a general solution to such system of equations is a linear combination of functions of the type $x_1^k \cdot \exp((p + \mathrm{i} \cdot q) \cdot x_1)$, i.e., functions of the type $x_1^k \cdot \exp(p \cdot x_1) \cdot \cos(q \cdot x_1)$ and $x_1^k \cdot \exp(p \cdot x_1) \cdot \sin(q \cdot x_1)$, where $p + \mathrm{i} \cdot q$ are eigenvalues of the matrix c_{ij}, and k is a non-negative integer corresponding to duplicate eigenvalues.

For a $d \times d$ matrix, the multiplicity of an eigenvalue cannot exceed d, so $k \le d$.

11°. Let us now study the consequences of *scale*-invariance of the class $\mathcal{S}$ with respect to the first variable. This scale-invariance implies, in particular, that for each basis function $e_i(x_1, x_2, \ldots, x_n)$, the result of its re-scaling $e_i(a_1 \cdot x_1, x_2, \ldots, x_n)$ is also a function from the class $\mathcal{S}$, i.e., that

$$e_i(a_1 \cdot x_1, x_2, \ldots, x_n) = \sum_{j=1}^{d} D_{ij}(a_1) \cdot e_j(x_1, x_2, \ldots, x_n), \tag{15}$$

for some coefficients D_{ij} that, in general, depend on a_1. Thus,

$$E_1(a_1 \cdot x_1) = D_{11}(a_1) \cdot E_1(x_1) + \ldots + D_{1d}(a_1) \cdot E_d(x_1);$$
$$\ldots \tag{16}$$
$$E_d(a_1 \cdot x_1) = D_{d1}(a_1) \cdot E_1(x_1) + \ldots + D_{dd}(a_1) \cdot E_d(x_1).$$

Similarly to Part 10 of this proof, we can prove that the dependencies $D_{ij}(a_1)$ are also differentiable. By differentiating both sides of the above equations with respect to a_1 and taking $a_1 = 1$, we conclude that

$$x_1 \cdot E_1'(x_1) = d_{11} \cdot E_1(x_1) + \ldots + d_{1d} \cdot E_d(x_1);$$
$$\ldots \tag{17}$$
$$x_1 \cdot E_d'(x_1) = d_{d1} \cdot E_1(x_1) + \ldots + d_{dd} \cdot E_d(x_1),$$

where $d_{ij} \stackrel{\text{def}}{=} D_{ij}'(1)$.

In each equation, the left-hand side $x_1 \cdot \dfrac{dE_i}{dx_1}$ can be reformulated as $\dfrac{dE_i}{dx_1/x_1} = \dfrac{dE_i}{d(\ln(x_1))}$. Thus, for the new variable $X_1 \stackrel{\text{def}}{=} \ln(x_1)$, we get the system of linear differential equations with constant coefficients:

$$\frac{dE_1}{dX_1} = d_{11} \cdot E_1(X_1) + \ldots + d_{1d} \cdot E_d(X_1);$$
$$\ldots \tag{18}$$
$$\frac{dE_d}{dX_1} = d_{d1} \cdot E_1(X_1) + \ldots + d_{dd} \cdot E_d(X_1).$$

We already know that a general solution to this equation is a linear combination of functions $X_1^k \cdot \exp(p \cdot X_1) \cdot \cos(q \cdot X_1)$ and $x_1^k \cdot \exp(p \cdot X_1) \cdot \sin(q \cdot X_1)$. Substituting $X_1 = \ln(x_1)$ into these formulas and taking into account that $\exp(p \cdot \ln(x_1)) = (\exp(\ln(x_1))^p = x_1^p$, we conclude that a general solution is a linear combination of functions of the type $(\ln(x_1))^k \cdot x_1^p \cdot \cos(q \cdot \ln(x_1))$ and $(\ln(x_1))^k \cdot x_1^p \cdot \sin(q \cdot \ln(x_1))$.

12°. From Parts 10 and 11 of this proof, we get two different expressions for the functions $E_i(x_1)$. By comparing these expressions, one can easily see that the only functions that can be described in both forms are functions of the form x^k for some non-negative integer $k \le d$ – or their linear combinations. So, each function $E_i(x_1)$ is a linear combination of such functions – i.e., a polynomial.

13°. We have shown that for each combination of values of $x_2, \ldots, x_n$, the dependence of each function $e_i(x_1, x_2, \ldots, x_n)$ on x_1 can be described by a polynomial of degree $\le d$. Similarly, we can prove that for each combination of values $x_1, x_3, \ldots, x_n$), the dependence on x_2 is also described by a polynomial. Let us combine these two conclusions and prove that each i, and for all possible values of $x_3, \ldots, x_n$, the dependence of $e_i(x_1, x_2, x_2, \ldots, x_n)$ on x_1 *and* on x_2 can be described by a polynomial of two variables.

Indeed, let us denote $T(x_1, x_2) \stackrel{\text{def}}{=} e_i(x_1, x_2, x_3, \ldots, x_n)$. We know that:

– for each x_2, this expression is a polynomial is x_1, and
– for each x_1, this expression is a polynomial is x_2.

Let us prove that $T(x_1, x_2)$ is a polynomial of two variables.

Indeed, the fact that the dependence of e_i on x_1 can be described by a polynomial of order $\le d$ can be rewritten, in terms of the function $T(x_1, x_2)$, as

$$T(x_1, x_2) = a_0(x_2) + a_1(x_2) \cdot x_1 + \ldots + a_d(x_2) \cdot x_1^d. \tag{19}$$

In writing this expression, we took into account that, in general, for different values of x_2, the coefficients $a_0, \ldots, a_d$ of this polynomial may be different.

Let us substitute d_1 different values $x_{10}, \ldots, x_{1d}$ of x_1 into this formula. As a result, we have $d+1$ linear equations for $d+1$ unknowns $a_0(x_2), \ldots, a_d(x_1)$, with constant coefficients:

$$\begin{aligned} T(x_{10}, x_2) &= a_0(x_2) + a_1(x_2) \cdot x_{10} + \ldots + a_d(x_2) \cdot x_{10}^d; \\ &\ldots \\ T(x_{1d}, x_2) &= a_0(x_2) + a_1(x_2) \cdot x_{1d} + \ldots + a_d(x_2) \cdot x_{1d}^d. \end{aligned} \tag{20}$$

In general, each component in a solution to a system of linear equations is a linear combination of the right-hand sides. The right-hand sides $T(x_{1i}, x_2)$ are polynomials of x_2. Thus, each coefficient $a_i(x_2)$ is a linear combination of polynomials – thus, a polynomial itself. Since all the expressions $a_i(x_2)$ are polynomials, the whole above expression for $T(x_1, x_2)$ becomes a polynomial in two variables x_1 and x_2.

By adding variables one by one, we can prove that the dependence on x_1, x_2, and x_3 is a polynomial, etc. – all the way to proving that the dependence of each of the basis function $e_i(x_1, \ldots, x_n)$ on all n variables $x_1, \ldots, x_n$ is a polynomial. Thus, each element of the class $\mathcal{S}$ – which is a linear combination of the basis functions – is also a polynomial.

For each tuple of parameters c, the function $\ln(f(x, c))$ belongs to the class $\mathcal{L} \subseteq \mathcal{S}$ and is, thus, also a polynomial. So, indeed, each pdf $f(x, c)$ from the family F has the form $\exp(P(x_1, \ldots, x_n))$ for some polynomial $P(x_1, \ldots, x_n)$. The proposition is proven.

Acknowledgements

This work was supported in part by the US National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science) and HRD-1242122 (Cyber-ShARE Center of Excellence). It was also supported by the program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478.

The authors are thankful to the anonymous referees for valuable suggestions.

References

Colbert, B. K., L. G. Crespo, and M. M. Peet, A sum of squares optimization approach to uncertainty quantification, In: *Proceedings of the 2019 American Control Conference ACC*, Philadelphia, Pennsylvania, USA, July 2019, pp. 5378–5384.

Crespo, L., An introduction to sliced-normal distributions, In M. Beer and E. Zio, editors, *Proceedings of the 29th European Safety and Reliability Conference ESREL'2019*, Hannover, Germany, September 2019. Research Publishing, Singapore, p. 28.

Crespo, L. G., B. K. Colbert, S. P. Kenny, and D. P. Giesy, On the quantification of aleatory and epistemic uncertainty using Sliced-Normal distributions, *Systems and Control Letters*, 134, Paper 104560, 2019.

Rabinovich, S. G., *Measurement Errors and Uncertainties: Theory and Practice*, Springer, New York, 2005.

Sheskin, D. J., *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC, Boca Raton, Florida, 2011.

Interval Matrix Multiplication Using Fast Low-precision Arithmetic on GPU

Katsuhisa Ozaki[1], Takeshi Ogita[2], Daichi Mukunoki[3]

[1]*Department of Mathematical Sciences, Shibaura Institute of Technology, ozaki@sic.shibaura-it.ac.jp*

[2]*Division of Mathematical Sciences, Tokyo Woman's Christian University, ogita@lab.twcu.ac.jp*

[3]*RIKEN Center for Computational Science, daichi.mukunoki@riken.jp*

Abstract. This paper discusses interval arithmetic, in particular, interval matrix multiplication using Graphics Processing Units (GPUs). Interval matrix multiplication plays an important role in verification methods for numerical linear algebra. For some GPUs, single-precision arithmetic is performed much faster than double-precision arithmetic. In this paper, we propose fast methods for performing interval matrix multiplication based on an error-free transformation of matrix multiplication that uses fast single-precision arithmetic and Tensor Cores on a GPU. Numerical examples also illustrate the efficiency of the proposed methods in terms of tightness of intervals and computational speed.

Keywords: interval arithmetic, matrix multiplication, GPU computing, mixed-precision computations

1. Introduction

The concept of an interval and its arithmetic, e.g., (Sunaga, 1958; Moore, 1966) has been widely applied to scientific problems. In this paper, we focused on interval matrix multiplication using a Graphics Processing Unit (GPU). Interval matrix multiplication plays crucial role in verification methods for numerical linear algebra, for example, for matrix determinants (Rump, 2005), matrix square roots (Frommer and Hashemi, 2010), eigenproblems (Oishi, 2001; Miyajima et al., 2010; Miyajima, 2012; Hoshi et al., 2020) and Sylvester matrix equations (Frommer and Hashemi, 2012; Miyajima, 2013). Several fast methods (Rump, 1999; Ogita and Oishi, 2005; Rump, 2012; Ozaki et al., 2012; Ozaki et al., 2015) for performing interval matrix multiplication exist, though the tightness of computed intervals and computing times continue to pose problems.

In this study, we propose fast computation methods for utilizing single-precision arithmetic (binary32) in IEEE 754 (IEEE, 2008) in interval matrix multiplication that is originally computed with double-precision arithmetic (binary64) in IEEE 754. This method is based on the error-free transformations published by (Ozaki et al., 2012), and instead of binary32 arithmetic, it also allows one to use Tensor Cores, which is a binary16/binary32 mixed-precision matrix engine available on recent NVIDIA GPUs. In high performance computing, there has been research on utilizing low

©2021 *by K. Ozaki, T. Ogita, D. Mukunoki.*

precision arithmetic units, which are available in the market at low cost due to the huge demand, for scientific and engineering computations. Examples include the use of binary32 units originally developed for image and media processing (e.g., Cell/B.E. and GPU with CUDA, both appeared in 2006), and the use of fast low-precision arithmetic units for artificial intelligence (e.g., Tensor Cores, Google TPUs), which has been developing remarkably in recent years. Following this trend, this study aims to make effective use of low-cost computing resources by utilizing low-precision arithmetic, and to expand the potential of low-precision arithmetic itself.

2. Error-Free Transformation Using Low-Precision Arithmetic

Let $\mathbb{F}_q$ ($\subset \mathbb{R}$) be a set of binary floating-point numbers with q bit in IEEE 754, for example, $\mathbb{F}_{64}$ is a set of binary64 floating-point numbers. Notation $\mathrm{fl}_q(\cdot)$ indicates that all operations in the parenthesis are evaluated using q bit floating-point arithmetic. The rounding mode of $\mathrm{fl}_q(\cdot)$ is rounding to nearest mode. For simplicity, we assumed that neither overflow nor underflow occurs in $\mathrm{fl}_q(\cdot)$. Let $\mathbf{u}_q$ be the roundoff unit for q bit floating-point numbers, i.e., $\mathbf{u}_{32} = 2^{-24}$ and $\mathbf{u}_{64} = 2^{-53}$ in IEEE 754. For example for $\mathbb{F}_{64}$, the maximum floating-point number is $(1 - \mathbf{u}_{64})2^{1024}$ and the minimum number in normalized positive floating-point numbers is 2^{-1022}. Hence, overflow or underflow is rare to occur. Once overflow occurs for a computations of dot product, a result is one of `Inf`, $-$`Inf`, or `NaN`. Therefore, occurrence of overflow is easy to find from a computed result. It is also possible to treat the case of underflow by extending rounding error analysis. A function $\mathtt{RN}_r(a)$ for $a \in \mathbb{R}$ produces a nearest floating-point number in $\mathbb{F}_r$ to a real number a. Note that for $a \in \mathbb{R}$

$$\mathtt{RN}_r(a) = a + \delta, \quad |\delta| \le \mathbf{u}_r|a| \tag{1}$$

holds true except in the case that $|a|$ is smaller than the smallest positive number in the normalized floating-point number. For $x \in \mathbb{R}^n$, $|x|$ indicates $(|x_1|, |x_2|, \ldots, |x_n|)^T$. The manner $|\cdot|$ is straightforwardly extended to a matrix. A function $\mathtt{ufp}(a)$ for $a \in \mathbb{R}$, the unit in the first place of the binary representation of a, is defined as

$$\mathtt{ufp}(a) := \begin{cases} 0 & a = 0 \\ 2^{\lfloor \log_2 |a| \rfloor} & \text{otherwise} \end{cases}.$$

We briefly review an error-free transformation of matrix multiplication (Ozaki et al., 2012). Applying Algorithm 3 in (Ozaki et al., 2012) for two given matrices $A \in \mathbb{F}_q^{m\times n}$ and $B \in \mathbb{F}_q^{n\times p}$ produces $A^{(s)}$ and $B^{(t)}$, such that

$$A = \sum_{s=1}^{n_A} A^{(s)}, \quad n_A \in \mathbb{N}, \quad A^{(s)} \in \mathbb{F}_q^{m\times n}, \ 1 \le s \le n_A \tag{2}$$

$$B = \sum_{t=1}^{n_B} B^{(t)}, \quad n_B \in \mathbb{N}, \quad B^{(t)} \in \mathbb{F}_q^{n\times p}, \ 1 \le t \le n_B \tag{3}$$

in order to satisfy

$$A^{(i)}B^{(j)} = \mathrm{fl}_q\left(A^{(i)}B^{(j)}\right), \quad 1 \le i \le n_A, \quad 1 \le j \le n_B.$$

We assume that a divide and conquer method, such as Strassen's method (Strassen, 1969) or Winograd's method (Winograd, 1971), were not applied to matrix multiplication. The constants n_A and n_B depend on the inner dimension n and the difference in the magnitude of elements in the rows of A and columns of B. Since no rounding error exists for the floating-point evaluation of $A^{(i)}B^{(j)}$ for all (i,j) pairs, AB can be transformed into an unevaluated sum of n_An_B floating-point matrices after computing $\mathrm{fl}_q\left(A^{(i)}B^{(j)}\right)$ for all (i,j) pairs.

In this paper, we propose an error-free transformation of AB using r bit ($r < q$) floating-point arithmetic. A typical example is $r = 32$ (binary32) and $q = 64$ (binary64) in IEEE 754. We assume that when $A \in \mathbb{F}_q^{m\times n}$ and $B \in \mathbb{F}_q^{n\times p}$ are rounded to $A' \in \mathbb{F}_r^{m\times n}$ and $B' \in \mathbb{F}_r^{n\times p}$, respectively, neither overflow nor underflow occurs in the rounding. Let $\underline{A}^{(1)} := A$ and $\underline{B}^{(1)} := B$. We define a constant β as

$$\beta := \left\lceil -\log_2 \mathbf{u}_q - \frac{-\log_2 \mathbf{u}_r - \log_2 n}{2} \right\rceil. \tag{4}$$

We set two vectors $v^{(k)} \in \mathbb{F}_q^m$ and $w^{(k)} \in \mathbb{F}_q^p$ for $k = 1$. If $\max\limits_{1\le j\le n} \left|\underline{a}_{ij}^{(k)}\right| = 0$, then we define $v_i^{(k)} := 0$. Similarly, $w_j^{(k)} := 0$ in the case of $\max\limits_{1\le i\le n} \left|\underline{b}_{ij}^{(k)}\right| = 0$. Otherwise, we set

$$v_i^{(k)} := \left\lceil \log_2 \max_{1\le j\le n} \left|\underline{a}_{ij}^{(k)}\right| \right\rceil, \quad w_j^{(k)} := \left\lceil \log_2 \max_{1\le i\le n} \left|\underline{b}_{ij}^{(k)}\right| \right\rceil. \tag{5}$$

Next, vectors $\sigma^{(k)} \in \mathbb{F}_q^m$ and $\tau^{(k)} \in \mathbb{F}_q^p$ are defined as

$$\sigma_i^{(k)} := \mathrm{fl}_q\left(0.75 \cdot 2^{\beta} \cdot 2^{v_i^{(k)}}\right), \quad \tau_j^{(k)} := \mathrm{fl}_q\left(0.75 \cdot 2^{\beta} \cdot 2^{w_j^{(k)}}\right) \tag{6}$$

using (4) and (5). Then, $A^{(k)}$, $\underline{A}^{(k+1)}$, $B^{(k)}$, $\underline{B}^{(k+1)}$ are obtained by

$$\begin{aligned} a_{ij}^{(k)} &:= \mathrm{fl}_q\left(\left(\sigma_i^{(k)} + \underline{a}_{ij}^{(k)}\right) - \sigma_i^{(k)}\right), \quad \underline{a}_{ij}^{(k+1)} := \mathrm{fl}_q\left(\underline{a}_{ij}^{(k)} - a_{ij}^{(k)}\right), \\ b_{ij}^{(k)} &:= \mathrm{fl}_q\left(\left(\tau_j^{(k)} + \underline{b}_{ij}^{(k)}\right) - \tau_j^{(k)}\right), \quad \underline{b}_{ij}^{(k+1)} := \mathrm{fl}_q\left(\underline{b}_{ij}^{(k)} - b_{ij}^{(k)}\right). \end{aligned} \tag{7}$$

If we compute (5), (6) and (7) from $k = 1, 2, \ldots$, then there exist $n'_A, n'_B \in \mathbb{N}$ such that

$$A = \sum_{s=1}^{n'_A} A^{(s)}, \quad B = \sum_{t=1}^{n'_B} B^{(t)}, \quad \underline{A}^{(n'_A+1)} = \mathbf{O}_{mn}, \quad \underline{B}^{(n'_B+1)} = \mathbf{O}_{np}, \tag{8}$$

where $\mathbf{O}_{mn}$ is the m-by-n zero matrix. For the matrices in (8),

$$A^{(i)} \in \mathbb{F}_r^{m\times n}, \quad B^{(j)} \in \mathbb{F}_r^{n\times p}, \quad A^{(i)}B^{(j)} = \mathrm{fl}_r\left(A^{(i)}B^{(j)}\right), \quad 1 \le i \le n'_A, \quad 1 \le j \le n'_B \tag{9}$$

are satisfied. This means that

- $\mathtt{RN}_r\left(A^{(i)}\right) = A^{(i)}$ and $\mathtt{RN}_r\left(B^{(j)}\right) = B^{(j)}$, and

– if we use r bit floating-point arithmetic for $A^{(i)}B^{(j)}$, a rounding error never occurs in the evaluation.

These are proved in the subsequent section. Hence, AB is transformed into an unevaluated $n'_A n'_B$ sums of floating-point matrices such that

$$AB = \sum_{r=1}^{n'_A} \sum_{s=1}^{n'_B} A^{(r)}B^{(s)} = \sum_{r=1}^{n'_A} \sum_{s=1}^{n'_B} \mathrm{fl}_r \left(A^{(r)}B^{(s)} \right). \tag{10}$$

If we obtain an accurate approximation of AB using the error-free transformation, we perform three steps as follows:

Step 1 Split matrices: $A = A^{(1)} + \ldots + A^{(n'_A)}$ and $B = B^{(1)} + \ldots + B^{(n'_B)}$ using (5), (6) and (7).

Step 2 Compute matrix multiplications $A^{(i)}B^{(j)},\ 1 \le i \le n'_A,\ 1 \le j \le n'_B$.

Step 3 Compute the summation in (10).

For Step 3, a user can use accurate summation algorithms, e.g., (Rump et al., 2008; Ogita et al., 2005; Yamanaka et al., 2008; Rump, 2009; Demmel and Hida, 2004; Zhu and Hayes, 2009; Zhu and Hayes, 2010; Rump et al., 2008). Or, pure floating-point arithmetic with proper order of summation can produce an accurate numerical result in many cases (Kazal et al.,). In these steps, the cost of Step 2 is the most expensive in many cases. Compared with (2) and (3), $n'_A n'_B$ is larger than $n_A n_B$. However, if r bit arithmetic is performed very quickly compared with q bit arithmetic, the computation of error-free transformation using low-precision arithmetic is finished quickly. In fact, we observed that matrix multiplication with binary32 worked more than 15 times faster than that with binary64 on the Quadro RTX 5000 by NVIDIA. This benchmark will be shown in Section 5.

3. Proof for the Error-Free Transformation

In this section, the proof of (9) is derived. We first introduce a basic lemma for floating-point arithmetic, which is derived from the definition of the floating-point number defined by IEEE 754.

Lemma 1. *For $a \in \mathbb{R}$, assume that $|a|$ is a multiple of the minimum positive floating-point number, and less than the maximum floating-point number. If we find $\mathbb{N} \cup \{0\} \ni k \le \mathbf{u}_q^{-1}$ and s (a power of two) such that $s\mathbb{Z} \ni |a| \le ks$, then $a \in \mathbb{F}_q$.*

The next lemma was derived by Minamihata et al. (Minamihata et al., 2016).

Lemma 2. *For $x \in \mathbb{F}_q$ and $M \in \mathbb{N}$, let*

$$\sigma := \begin{cases} 0.75 \cdot 2^M \cdot 2^{\lceil \log_2 |x| \rceil} & x \neq 0 \\ 0 & x = 0 \end{cases}.$$

If x_1 and x_2 are obtained as follows:

$$x_1 = \mathrm{fl}_q \left((\sigma + x) - \sigma \right), \quad x_2 = \mathrm{fl}_q \left(x - x_1 \right),$$

then

$$x = x_1 + x_2, \quad |x_1| \le \mathtt{ufp}(\sigma)2^{-M}, \quad x_1 \in \mathbf{u}_q\mathtt{ufp}(\sigma)\mathbb{Z}$$

are satisfied.

Applying Lemma 2 into the computations (4), (5), (6) and (7) yields

$$\left|a_{ij}^{(k)}\right| \le 2^{-\beta}\mathtt{ufp}\left(\sigma_i^{(k)}\right), \quad \left|b_{ij}^{(\ell)}\right| \le 2^{-\beta}\mathtt{ufp}\left(\tau_j^{(\ell)}\right), \quad a_{ij}^{(k)} \in \mathbf{u}_q\mathtt{ufp}\left(\sigma_i^{(k)}\right)\mathbb{Z}, \quad b_{ij}^{(\ell)} \in \mathbf{u}_q\mathtt{ufp}\left(\tau_j^{(\ell)}\right)\mathbb{Z}. \tag{11}$$

We use these properties for the following theorems.

Theorem 1. *For the matrices $A^{(i)}$ and $B^{(j)}$ in (9), that are generated by (4), (5), (6) and (7),*

$$A^{(i)} \in \mathbb{F}_r^{m\times n}, \quad B^{(j)} \in \mathbb{F}_r^{n\times p} \tag{12}$$

are satisfied for $1 \le i \le n'_A$ and $1 \le j \le n'_B$.

Proof. From the definition of β in (4), we have $2^{-\beta} < \mathbf{u}_r^{-1}\mathbf{u}_q$. This and (11) yield

$$\begin{aligned}
\mathbf{u}_q \cdot \mathtt{ufp}\left(\sigma_i^{(k)}\right)\mathbb{Z} \ \ni\ |a_{ij}^{(k)}| \le 2^{-\beta}\mathtt{ufp}\left(\sigma_i^{(k)}\right) < \mathbf{u}_r^{-1}\mathbf{u}_q\mathtt{ufp}\left(\sigma_i^{(k)}\right),\\
\mathbf{u}_q \cdot \mathtt{ufp}\left(\tau_j^{(k)}\right)\mathbb{Z} \ \ni\ |b_{ij}^{(k)}| \le 2^{-\beta}\mathtt{ufp}\left(\tau_j^{(k)}\right) < \mathbf{u}_r^{-1}\mathbf{u}_q\mathtt{ufp}\left(\tau_j^{(k)}\right).
\end{aligned}$$

Therefore, (12) is satisfied from Lemma 1. □

Theorem 2. *For the matrices $A^{(k)}$ and $B^{(\ell)}$ in (9),*

$$A^{(k)}B^{(\ell)} = \mathrm{fl}_r\left(A^{(k)}B^{(\ell)}\right) \tag{13}$$

are satisfied for $1 \le k \le n'_A,\ 1 \le \ell \le n'_B$.

Proof. From the definition of β in (4),

$$2^{-2\beta} = \frac{1}{2^{2\beta}} = \frac{1}{2^{2\left\lceil -\log_2 \mathbf{u}_q - \frac{-\log_2 \mathbf{u}_r - \log_2 n}{2}\right\rceil}} \le \frac{1}{2^{-2\log_2 \mathbf{u}_q + \log_2 \mathbf{u}_r + \log_2 n}} = \frac{1}{n}\mathbf{u}_r^{-1}\mathbf{u}_q^2 \tag{14}$$

holds true. Applying (11) yields

$$a_{iz}^{(k)}b_{zj}^{(\ell)} \in \mathbf{u}_q^2 \cdot \mathtt{ufp}\left(\sigma_i^{(k)}\right)\mathtt{ufp}\left(\tau_j^{(\ell)}\right)\mathbb{Z}$$

for all i, z, j, k, and ℓ. Using this, (11) and (14) gives

$$\begin{aligned}
\mathbf{u}_q^2\mathtt{ufp}\left(\sigma_i^{(k)}\right)\mathtt{ufp}\left(\tau_j^{(\ell)}\right)\mathbb{Z} \ &\ni\ \sum_{z=1}^{n} a_{iz}^{(k)}b_{zj}^{(\ell)} \le \sum_{z=1}^{n} |a_{iz}^{(k)}||b_{zj}^{(\ell)}|\\
&\le\ n2^{-2\beta}\mathtt{ufp}\left(\sigma_i^{(k)}\right)\mathtt{ufp}\left(\tau_j^{(\ell)}\right) \le \mathbf{u}_r^{-1}\mathbf{u}_q^2\mathtt{ufp}\left(\sigma_i^{(k)}\right)\mathtt{ufp}\left(\tau_j^{(\ell)}\right).
\end{aligned}$$

Therefore, (13) is proved from Lemma 1. □

A computed result of the dot product depends on a computational order, for example, recursive, block-wise, and pair-wise orders, due to rounding errors. Note that Theorem 2 is valid for such orders of computation. This can be proved using an induction on binary trees as in (Jeannerod and Rump, 2013).

4. Application to Interval Matrix Multiplication

4.1. Basics of Methods for Interval Matrix Multiplication

Let $\mathbb{IR}$ be a set of real intervals. We define an inf-sup form $[a, b]$ with real numbers $a, b \in \mathbb{R}$ $(a \leq b)$ and a mid-rad form $\langle c, r\rangle$ with real numbers $c, r \in \mathbb{R}$ $(r \geq 0)$:

$$[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}, \quad \langle c, r\rangle = \{x \in \mathbb{R} \mid c - r \leq x \leq c + r\}.$$

These intervals are straightforwardly extended to vectors and matrices. We introduced floating-point arithmetic with directed rounding. Let

- $\mathrm{fl}_{\triangledown,q}(\cdot)$: a computed result using rounding downward mode (roundTowardNegative) using q bit floating-point arithmetic
- $\mathrm{fl}_{\triangle,q}(\cdot)$: a computed result using rounding upward mode (roundTowardPositive) using q bit floating-point arithmetic

For $A_m, A_r \in \mathbb{F}_q^{m\times n}$ and $B_m, B_r \in \mathbb{F}_q^{n\times p}$ ($A_r = |A_r|$ and $B_r = |B_r|$), an interval matrix multiplication $\langle A_m,\ A_r\rangle \cdot \langle B_m,\ B_r\rangle$ is examined. It is well known, e,g, (Rump, 1999), that

$$\langle A_m,\ A_r\rangle \cdot \langle B_m,\ B_r\rangle \subseteq \langle A_m B_m,\ T_1\rangle, \quad T_1 = |A_m|B_r + A_r(|B_m| + B_r). \tag{15}$$

Based on (15), Rump proposed the following method (Rump, 1999)

$$\langle A_m,\ A_r\rangle \cdot \langle B_m,\ B_r\rangle \subseteq [\mathrm{fl}_{\triangledown,q}(A_m B_m - T_1'),\ \mathrm{fl}_{\triangle,q}(A_m B_m + T_1')],$$

where

$$T_1' = \mathrm{fl}_{\triangle,q}(|A_m|B_r + A_r(|B_m| + B_r)).$$

In total, four matrix multiplications are necessary to compute the enclosure of interval matrix multiplication. Ozaki et al. (Ozaki et al., 2012) and Rump (Rump, 2012) have proposed a method with three matrix multiplications. For a computed result of matrix multiplication,

$$|A_m B_m - \mathrm{fl}_q(A_m B_m)| \leq n\mathbf{u}_q |A_m||B_m| \tag{16}$$

is satisfied barring overflow and underflow (Jeannerod and Rump, 2013). Then, we have

$$A_m B_m \in \langle \mathrm{fl}_q(A_m B_m),\ n\mathbf{u}_q |A||B|\rangle. \tag{17}$$

From (17) and (15),

$$\langle A_m,\ A_r\rangle \cdot \langle B_m,\ B_r\rangle \subset \langle \mathrm{fl}_q(A_m B_m),\ T_2\rangle \tag{18}$$

is satisfied, where

$$T_2 = |A_m|(n\mathbf{u}_q|B_m| + B_r) + A_r(|B_m| + B_r). \tag{19}$$

Using direct rounding, $\langle A_m,\ A_r\rangle \cdot \langle B_m,\ B_r\rangle$ is enclosed by

$$\langle A_m,\ A_r\rangle \cdot \langle B_m,\ B_r\rangle \subset \langle \mathrm{fl}_q(A_m B_m),\ T_2'\rangle, \tag{20}$$

where

$$T_2' = \mathrm{fl}_{\triangle,q}(|A_m|(n\mathbf{u}_q|B_m| + B_r) + A_r(|B_m| + B_r)). \tag{21}$$

The paper (Ozaki et al., 2015) have shown that if low-precision arithmetic is used to obtain the upper bounds of T_1 and T_2, the overestimation of the radius is almost negligible in many cases. Let $A_m', A_r' \in \mathbb{F}_r^{m\times n}$ and $B_m', B_r' \in \mathbb{F}_r^{n\times p}$ such that

$$|A_m| \le A_m' \in \mathbb{F}_r^{m\times n}, \quad A_r \le A_r' \in \mathbb{F}_r^{m\times n}, \quad |B_m| \le B_m' \in \mathbb{F}_r^{n\times p}, \quad B_r \le B_r' \in \mathbb{F}_r^{n\times p}.$$

Because

$$\begin{aligned} T_1 &= |A_m|B_r + A_r(|B_m| + B_r) \le A_m'B_r' + A_r'(B_m' + B_r') \le \mathrm{fl}_{\triangle,r}(A_m'B_r' + A_r'(B_m' + B_r')) := T_1'', \\ T_2 &= |A_m|(n\mathbf{u}_q|B_m| + B_r) + A_r(|B_m| + B_r) \le \mathrm{fl}_{\triangle,r}(A_m'(n\mathbf{u}_q B_m' + B_r') + A_r'(B_m' + B_r')) := T_2'', \end{aligned}$$

the interval matrix multiplication $\langle A_m,\ A_r\rangle \cdot \langle B_m,\ B_r\rangle$ is enclosed by

$$\langle A_m,\ A_r\rangle \cdot \langle B_m,\ B_r\rangle \ \subseteq\ [\mathrm{fl}_{\triangledown,q}(A_m B_m - T_1''),\ \mathrm{fl}_{\triangle,q}(A_m B_m + T_1'')], \tag{22}$$

$$\langle A_m,\ A_r\rangle \cdot \langle B_m,\ B_r\rangle \ \subseteq\ \langle \mathrm{fl}_q(A_m B_m),\ T_2''\rangle. \tag{23}$$

Obtaining the interval in (22) involves two matrix multiplications with q bit arithmetic and two matrix multiplications with r bit arithmetic. For (23), the cost is one matrix multiplication with q bit arithmetic and two matrix multiplications with r bit arithmetic.

Another possibility is to use Ogita and Oishi's technique. For $X = |X| \in \mathbb{F}_q^{m\times n}$ and $Y = |Y| \in \mathbb{F}_q^{n\times p}$, set four vectors $g \in \mathbb{R}^n$, $h \in \mathbb{R}^n$, $e \in \mathbb{F}^m$, and $f \in \mathbb{F}^p$ as follows

$$g_j = \max_i |X_{ij}|, \quad h_i = \max_j |Y_{ij}|, \quad e = (1, 1, \ldots, 1)^T \in \mathbb{F}^m, \quad f = (1, 1, \ldots, 1)^T \in \mathbb{F}^p. \tag{24}$$

Then, XY is bounded by

$$XY \le \min\left(e(gY), (Xh)f^T\right) \le \min\left(\mathrm{fl}_{\triangle,q}\left(e(gY)\right), \mathrm{fl}_{\triangle,q}\left((Xh)f^T\right)\right) =: F, \tag{25}$$

where $\min(C, D)$ returns a matrix with the smallest elements taken from $C \in \mathbb{F}^{n\times n}$ or $D \in \mathbb{F}^{n\times n}$. The result of a product of two positive matrices can be bounded without computing full matrix multiplication. We denoted a function $F = \mathtt{fn}(X, Y, q)$ for computation (25). Using this function, we have

$$|A_m|B_r + A_r(|B_m| + B_r) \le \mathrm{fl}_{\triangle,q}\left(\mathtt{fn}(|A_m|, B_r, q) + \mathtt{fn}(A_r, \mathrm{fl}_{\triangle,q}(|B_m| + B_r), q)\right) := T_1''',$$

and

$$\begin{aligned} & |A_m|(n\mathbf{u}_q|B_m| + B_r) + A_r(|B_m| + B_r) \\ \le\ & \mathtt{fn}(|A_m|, \mathrm{fl}_{\triangle,q}\left(n\mathbf{u}_q|B_m| + B_r\right), q) + \mathtt{fn}(A_r, \mathrm{fl}_{\triangle,q}\left(|B_m| + B_r\right), q) \\ \le\ & \mathrm{fl}_{\triangle,q}\left(\mathtt{fn}(|A_m|, \mathrm{fl}_{\triangle,q}\left(n\mathbf{u}_q|B_m| + B_r\right), q) + \mathtt{fn}(A_r, \mathrm{fl}_{\triangle,q}\left(|B_m| + B_r\right), q)\right) =: T_2'''. \end{aligned}$$

Then,

$$\langle A_m,\ A_r\rangle \cdot \langle B_m,\ B_r\rangle \ \subseteq\ [\mathrm{fl}_{\bigtriangledown,q}(A_mB_m - T_1'''),\ \mathrm{fl}_{\bigtriangleup,q}(A_mB_m + T_1''')], \tag{26}$$

$$\langle A_m,\ A_r\rangle \cdot \langle B_m,\ B_r\rangle \ \subseteq\ \langle \mathrm{fl}_q(A_mB_m),\ T_2'''\rangle. \tag{27}$$

Method (26) with two matrix multiplications was proposed by Ogita and Oishi (Ogita and Oishi, 2005), and method (27) with a matrix multiplication was proposed by Ozaki et al. (Ozaki et al., 2012).

4.2. Proposed Methods

We propose a method based on (15). Because we cannot switch rounding modes on GPUs for already compiled BLAS routines, for example, $\mathrm{fl}_{\bigtriangledown,q}(A_mB_m - T_1''')$, $\mathrm{fl}_{\bigtriangleup,q}(A_mB_m + T_1''')$ in (26) could not be computed using functions in cuBLAS (NVIDIA, 2007). Therefore, we obtained the enclosure of interval matrix multiplication by only rounding to nearest mode. For $0 \le a, b \in \mathbb{F}_q$,

$$a + b \le \mathrm{fl}_q((\theta_q(a + b)), \quad ab \le \mathrm{fl}_q(\theta_q(ab)), \quad \theta_q = 1 + 2\mathbf{u}_q \in \mathbb{F}_q \tag{28}$$

is satisfied from the standard floating-point arithmetic (Kearfott et al., 1994), and

$$1 + \frac{n\mathbf{u}_q}{1 - n\mathbf{u}_q} = \frac{1}{1 - n\mathbf{u}_q} \le \kappa_q, \quad \kappa_q := \mathrm{fl}_q\left(\frac{1}{1 - (n + 1)\mathbf{u}_q}\right), \quad (n + 1)\mathbf{u}_q < 1 \tag{29}$$

is obtained from (Ogita et al., 2005). For $0 \le a \in \mathbb{R}$, we have

$$a \le \mathrm{fl}_r(\theta_r \mathtt{RN}_r(a)). \tag{30}$$

For matrices $X = |X| \in \mathbb{F}_q^{m\times n}$ and $Y = |Y| \in \mathbb{F}_q^{n\times p}$, and the vectors in (24), XY is bounded by

$$XY \le \min\left(e(gY), (Xh)f^T\right) \le \mathrm{fl}_q\left(\theta_q \min\left(\mathrm{fl}_q\left(e(gY)\right), \mathrm{fl}_q\left((Xh)f^T\right)\right)\right) =: z \tag{31}$$

using only $\mathrm{fl}_q(\cdot)$ from (28). We define the function $\mathtt{fn2}(X, Y, q)$ producing z in (31).

For matrices $A \in \mathbb{F}_q^{m\times n}$ and $B \in \mathbb{F}_q^{n\times p}$, the paper (Rump, 2015) introduced

$$|A||B| \le \left(1 + \frac{n\mathbf{u}_q}{1 - n\mathbf{u}_q}\right) \mathrm{fl}_q(|A||B|). \tag{32}$$

To apply the error-free transformation for matrix multiplication to A_mB_m, we compute (5), (6) and (7) for A_m and B_m up to $k - 1$. Then, A_m and B_m are divided into k matrices, such that

$$\begin{aligned} A_m &= A_m^{(1)} + A_m^{(2)} + \ldots + \underline{A}_m^{(k)},\ A_m^{(i)} \in \mathbb{F}_r^{m\times n},\ 1 \le i \le k - 1, \quad A_m^{(k)} \in \mathbb{F}_q^{m\times n}, \\ B_m &= B_m^{(1)} + B_m^{(2)} + \ldots + \underline{B}_m^{(k)},\ B_m^{(j)} \in \mathbb{F}_r^{n\times p},\ 1 \le j \le k - 1, \quad B_m^{(k)} \in \mathbb{F}_q^{n\times p}. \end{aligned}$$

Let $\Delta_A \in \mathbb{F}_q^{m\times n}$ be $\Delta_A = \underline{A}_m^{(k)} - \mathtt{RN}_r\left(\underline{A}_m^{(k)}\right)$. Then, A_m is transformed into

$$A_m = A_m^{(1)} + A_m^{(2)} + \ldots + A_m^{(k-1)} + \mathtt{RN}_r\left(\underline{A}_m^{(k)}\right) + \Delta_A.$$

Now, we have

$$\langle A_m, A_r\rangle \subseteq \langle C_A,\ A_r + |\Delta_A|\rangle, \quad C_A \ = \ A_m - \Delta_A = A_m^{(1)} + A_m^{(2)} + \ldots + A_m^{(k-1)} + \mathtt{RN}_r\left(\underline{A}_m^{(k)}\right).$$

Let $A'_r = \mathrm{fl}_r(\theta_r\mathtt{RN}_r(\mathrm{fl}_q(\theta_q(A_r + |\Delta_A|))))$ and $B'_r = \mathrm{fl}_r(\theta_r\mathtt{RN}_r(B_r))$. Then, from (30) and (28), $A_r + |\Delta_A| \le A'_r$ and $B_r \le B'_r$ are satisfied, and we have $\langle A_m, A_r\rangle \subseteq \langle C_A,\ A'_r\rangle$ and $\langle B_m, B_r\rangle \subseteq \langle B_m,\ B'_r\rangle$. Therefore,

$$\langle A_m, A_r\rangle\langle B_m, B_r\rangle \subseteq \langle C_A, A'_r\rangle\langle B_m, B'_r\rangle \subseteq \langle C_AB_m,\ |C_A|B'_r + A'_r(|B_m| + B'_r)\rangle. \tag{33}$$

We focused on how to obtain enclosure of the center C_AB_m and the upper bound of the radius $|C_A|B'_r + A'_r(|B_m| + B'_r)$ in (33) by using only floating-point arithmetic with rounding to nearest mode.

First, we considered an upper bound of the radius in (33). Let $C'_A = \mathrm{fl}_r(\theta_r\mathtt{RN}_r(|C_A|))$ and $B'_m = \mathrm{fl}_r(\theta_r\mathtt{RN}_r(|B_m|))$. Using (28), (29) and (32) gives

$$|C_A|B'_r \le C'_AB'_r \le \left(1 + \frac{n\mathbf{u}_r}{1 - n\mathbf{u}_r}\right)\mathrm{fl}_r(C'_AB'_r) \le \kappa_r\mathrm{fl}_r(C'_AB'_r) \le \mathrm{fl}_r(\theta_r(\kappa(C'_AB'_r))) =: X_1,$$

and

$$\begin{aligned} A'_r(|B_m| + B'_r) \ &\le \ A'_r(B'_m + B'_r) \le \left(1 + \frac{n\mathbf{u}_r}{1 - n\mathbf{u}_r}\right)\mathrm{fl}_r(A'_rW) \\ &\le \ \kappa_r\mathrm{fl}_r(A'_rW) \le \mathrm{fl}_r(\theta_r(\kappa_r(A'_rW))) =: X_2, \end{aligned}$$

where $W = \mathrm{fl}_r(\theta_r(B'_m + B'_r))$. Finally, the radius in (33) is bounded by

$$|C_A|B'_r + A'_r(|B_m| + B'_r) \le \mathrm{fl}_q(\theta_q(X_1 + X_2)) =: Z.$$

Now, $\langle A_m,\ A_r\rangle \cdot \langle B_m,\ B_r\rangle \subseteq \langle C_AB_m,\ Z\rangle$. Note that it is possible to use the function `fn2` for the upper bound of $|C_A|B'_r + A'_r(|B_m| + B'_r)$.

Next, we treat the enclosure of C_AB_m in (33) that can be transformed into

$$C_AB_m = \sum_{i+j\le k} A_m^{(i)}B_m^{(j)} + \sum_{i=1}^{k-1} A_m^{(i)}\underline{B}_m^{(k-i+1)} + \mathtt{RN}_r(\underline{A}_m^{(k)})B_m. \tag{34}$$

There are $\dfrac{1}{2}k(k-1)$ matrix multiplications in (34).

We cannot apply low-precision arithmetic directly to (34), because

$$\underline{B}_m^{(k-i+1)} \notin \mathbb{F}_r^{m\times n}, \quad B_m \notin \mathbb{F}_r^{m\times n}$$

can be satisfied. Let

$$\underline{B}_m^{(k-i+1)} \ = \ \dot{B}_m^{(k-i+1)} + \Delta_{Bm}^{(k-i+1)}, \quad \dot{B}_m^{(k-i+1)} = \mathtt{RN}_r(B_m^{(k-i+1)}), \quad \Delta_{Bm}^{(k-i+1)} \in \mathbb{F}_q^{n\times p}, \tag{35}$$

$$B_m \ = \ \dot{B}_m + \Delta_{Bm}, \quad \dot{B}_m = \mathtt{RN}_r(B_m), \quad \Delta_{Bm} \in \mathbb{F}_q^{n\times p}. \tag{36}$$

From (1), we have

$$|\Delta_{Bm}^{(k-i+1)}| \le \mathbf{u}_r |\dot{B}_m^{(k-i+1)}|, \quad |\Delta_{Bm}| \le \mathbf{u}_r |\dot{B}_m|. \tag{37}$$

Note that

$$\mathrm{fl}_r\left(A_m^{(i)} B_m^{(j)}\right) = A_m^{(i)} B_m^{(j)}, \quad i+j \le k,$$

and it is possible that

$$\mathrm{fl}_r\left(A_m^{(i)} \dot{B}_m^{(k-i+1)}\right) \ne A_m^{(i)} \dot{B}_m^{(k-i+1)}, \quad 1 \le i \le k-1, \quad \mathrm{fl}_r\left(\mathtt{RN}_r(\underline{A}_m^{(k)})\dot{B}_m\right) \ne \mathtt{RN}_r(\underline{A}_m^{(k)})\dot{B}_m.$$

From (16), we have an enclosure of $A_m^{(i)} \dot{B}_m^{(k-i+1)}$ as

$$A_m^{(i)} \dot{B}_m^{(k-i+1)} \in \left\langle \mathrm{fl}_r\left(A_m^{(i)} \dot{B}_m^{(k-i+1)}\right), n\mathbf{u}_r |A_m^{(i)}||\dot{B}_m^{(k-i+1)}| \right\rangle.$$

Using (35), (16), (31), and (28), for $1 \le i \le k-1$ gives

$$\begin{aligned}
A_m^{(i)} \underline{B}_m^{(k-i+1)} &= A_m^{(i)}\left(\dot{B}_m^{(k-i+1)} + \Delta_{Bm}^{(k-i+1)}\right) \subseteq \left\langle A_m^{(i)} \dot{B}_m^{(k-i+1)},\ |A_m^{(i)}||\Delta_{Bm}^{(k-i+1)}| \right\rangle \\
&\subseteq \left\langle \mathrm{fl}_r\left(A_m^{(i)} \dot{B}_m^{(k-i+1)}\right),\ n\mathbf{u}_r |A_m^{(i)}||\dot{B}_m^{(k-i+1)}| + \mathbf{u}_r |A_m^{(i)}||\dot{B}_m^{(k-i+1)}| \right\rangle \\
&= \left\langle \mathrm{fl}_r\left(A_m^{(i)} \dot{B}_m^{(k-i+1)}\right),\ |A_m^{(i)}|\left((n+1)\mathbf{u}_r |\dot{B}_m^{(k-i+1)}|\right) \right\rangle \\
&\subseteq \left\langle \mathrm{fl}_r\left(A_m^{(i)} \dot{B}_m^{(k-i+1)}\right),\ |A_m^{(i)}|\left(\mathrm{fl}_q(\theta_q((n+1)\mathbf{u}_r |\dot{B}_m^{(k-i+1)}|)\right) \right\rangle \\
&\subseteq \left\langle C_m^{(i)},\ C_r^{(i)} \right\rangle
\end{aligned}$$

where

$$C_m^{(i)} = \mathrm{fl}_r\left(A_m^{(i)} \dot{B}_m^{(k-i+1)}\right), \quad C_r^{(i)} = \mathtt{fn2}\left(|A_m^{(i)}|,\ \left(\mathrm{fl}_q(\theta_q((n+1)\mathbf{u}_r |\dot{B}_m^{(k-i+1)}|)\right),\ q\right)$$

for $1 \le i \le k-1$. Similarly, from (36), (16), (31), and (28), we obtain

$$\begin{aligned}
\mathtt{RN}_r(\underline{A}_m^{(k)}) B_m &= \mathtt{RN}_r(\underline{A}_m^{(k)})(\dot{B}_m + \Delta_{Bm}) \subseteq \left\langle \mathtt{RN}_r(\underline{A}_m^{(k)})\dot{B}_m,\ |\mathtt{RN}_r(\underline{A}_m^{(k)})||\Delta_{Bm}| \right\rangle \\
&\subseteq \left\langle \mathrm{fl}_r\left(\mathtt{RN}_r(\underline{A}_m^{(k)})\dot{B}_m\right),\ n\mathbf{u}_r |\mathtt{RN}_r(\underline{A}_m^{(k)})||\dot{B}_m| + \mathbf{u}_r |\dot{A}_m^{(k)}||\dot{B}_m| \right\rangle \\
&= \left\langle \mathrm{fl}_r\left(\mathtt{RN}_r(\underline{A}_m^{(k)})\dot{B}_m\right),\ |\mathtt{RN}_r(\underline{A}_m^{(k)})|((n+1)\mathbf{u}_r |\dot{B}_m|) \right\rangle \\
&\subseteq \left\langle \mathrm{fl}_r\left(\mathtt{RN}_r(\underline{A}_m^{(k)})\dot{B}_m\right),\ |\mathtt{RN}_r(\underline{A}_m^{(k)})|\mathrm{fl}_q(\theta_q((n+1)\mathbf{u}_r |\dot{B}_m|)) \right\rangle \\
&= \langle C_m^{(k)},\ C_r^{(k)} \rangle
\end{aligned}$$

where

$$C_m^{(k)} = \mathrm{fl}_r\left(\mathtt{RN}_r(\underline{A}_m^{(k)})\dot{B}_m\right), \quad C_r^{(k)} = \mathtt{fn2}\left(|\mathtt{RN}_r(\underline{A}_m^{(k)})|,\ \mathrm{fl}_q(\theta_q((n+1)\mathbf{u}_r |\dot{B}_m|)),\ q\right)$$

Until now, we have

$$C_A C_B \in \left\langle \sum_{i+j\le k} \mathrm{fl}_r\left(A_m^{(i)} B_m^{(j)}\right),\ \mathbf{O} \right\rangle + \sum_{i=1}^{k} \langle C_m^{(i)},\ C_r^{(i)} \rangle$$
$$= \left\langle \sum_{i=1}^{\frac{1}{2}k(k-1)} D^{(i)} + \sum_{i=1}^{k} C_m^{(i)},\ \sum_{i=1}^{k} C_r^{(i)} \right\rangle,$$

where $D^{(1)} := \mathrm{fl}_r(A^{(1)}B^{(1)})$, $D^{(2)} := \mathrm{fl}_r(A^{(1)}B^{(2)})$, $D^{(3)} := \mathrm{fl}_r(A^{(2)}B^{(1)}), \ldots, D^{(\frac{1}{2}k(k-1))} := \mathrm{fl}_r(A^{(k-1)}B^{(1)})$. Let $F^{(1)} := D^{(1)}$, $G^{(1)} := \mathbf{O}$ and $H^{(1)} := C_r^{(1)}$. Compute

$$F^{(i+1)} := \mathrm{fl}_q\left(F^{(1)} + D^{(i+1)}\right), \quad G^{(i+1)} := \mathrm{fl}_q\left(\theta_q\left(G^{(i)} + \mathbf{u}_q|F_{(i+1)}|\right)\right),$$

for $i = 2, 3, \ldots, \frac{1}{2}(k^2 - k) - 1$ and

$$H^{(j+1)} := \mathrm{fl}_q\left(\theta_q\left(C_r^{(j+1)} + H^{(j)}\right)\right)$$

for $j = 1, \ldots, k-1$. $C_A B_m$ is enclosed by

$$C_A B_m \in \left\langle F^{(\frac{1}{2}(k^2-k))}, Q \right\rangle, \quad Q := \mathrm{fl}_q(\theta_q(G^{(\frac{1}{2}(k^2-k))} + H^{(k)})),$$

and we finally obtain

$$\langle A_m,\ A_r\rangle\langle B_m,\ B_r\rangle \in \langle F^{(\frac{1}{2}k(k-1))},\ \mathrm{fl}_q(\theta_q(Q + Z))\rangle. \tag{38}$$

Setting $q = 64$ and $r = 32$ derives

$$PA^{(i)} \in \mathbb{F}_{16}^{m\times n}, \quad B^{(j)}Q \in \mathbb{F}_{16}^{n\times p}$$

are satisfied where P, Q are diagonal matrices for proper scaling. Because Tensor Cores accepts binary16 format and compute it as binary32, no rounding error occurs in the evaluation of $A^{(i)}B^{(j)}$, $i + j \le k$ using Tensor Cores (Mukunoki et al., 2020). Performance is accelerated when the size of the matrices is sufficiently large. This will be shown in the next section.

5. Numerical Examples

In this section, we present numerical examples to illustrate the efficiency of the proposed methods. We used the following components: CPU: Core i7-8665U, GPU: Quadro RTX 5000, OS: Windows 10, Software: MATLAB R2020a with Parallel Computing Toolbox and CUDA version 10.0.130. First, we determined the difference in performance between binary32 arithmetic and binary64 arithmetic for both the CPU and GPU. Tables I and II show the computing time of a product of square matrices using MATLAB. Each item shows an average of 20 examples. Note that computing time

in this section did not include the transfer time of data between CPU and GPU. The number in parentheses is represented in the form: precision (bit), e.g., CPU (32) means binary32 on the CPU. GPU (16-32) means computation using Tensor Cores with cublasGemmEx in cuBLAS. The inputs in cublasGemmEx were binary16, and the output was binary32 in the numerical examples. Note that scaling for matrices was necessary for converting binary32 to binary16 in many cases: otherwise, overflow or underflow tended to occur due to a narrow range of the exponent of binary16. Hence, GPU (16-32) included the time for diagonal scaling. The following is the detail of the benchmark using MATLAB.

- CPU (32) and CPU (64): $A * B$ for randomly generated matrices A and B
- GPU (32) and GPU (64): $A * B$ (after $A = \mathrm{gpuArray}(A)$; and $B = \mathrm{gpuArray}(B)$;, the time for gpuArray is not included)
- GPU (16-32): transforming the matrices with binary32 into matrices with binary16 and calling cublasGemmEx in cuBLAS exploiting MATLAB executable. The code is compiled using mexcuda.

If the matrices' dimensions were over 2,000, the ratio GPU (64) / GPU (32) was more than 15. There is an advantage to using binary32 on the GPU. For example, among NVIDIA GPUs after the Maxwell architecture (released in 2014), GeForce, Tegra, and some low-cost Tesla products have binary64 performance less than 1/32 of binary32, and thus the proposed method will be effective.

However, because the ratio CPU (64) / CPU (32) was approximately 2, there was no merit in using binary32 on the CPU for this problem.

Table I. Computing times (sec) for matrix multiplication (CPU).

Dimension	CPU (32)	CPU (64)	ratio (64/32)
2,000	0.086	0.1592	1.85
4,000	0.5090	1.1211	2.20
6,000	1.8083	3.9009	2.15
8,000	4.3411	9.3109	2.14
10,000	8.1339	17.7706	2.18
12,000	13.3200	31.8232	2.38

Interval matrices $\langle A_m, A_r \rangle$ and $\langle B_m, B_r \rangle$ were generated using MATLAB2020a as follows

$$\begin{aligned} A_m &= \mathtt{randn}(n); \quad A_r = c_1 * \mathtt{rand}(n). * \mathtt{abs}(A_m); \\ B_m &= \mathtt{randn}(n); \quad B_r = c_2 * \mathtt{rand}(n). * \mathtt{abs}(B_m); \end{aligned}$$

Table II. Computing times (sec) for matrix multiplication (GPU).

Dimension	GPU (16-32)	GPU (32)	GPU (64)	ratio (64/(16-32))	ratio (64/32)
2,000	0.0049	0.0032	0.0524	10.7937	16.4890
4,000	0.0143	0.0225	0.3987	27.7866	17.7414
6,000	0.0485	0.0749	1.3573	27.9910	18.1275
8,000	0.0672	0.1932	3.2473	48.3451	16.8092
10,000	0.1581	0.3890	6.4355	40.7028	16.5422
12,000	0.3577	0.6031	11.0075	30.7695	18.2517

where c_1 and c_2 are positive constants and the width of the interval can be controlled by the constants c_1 and c_2.

Table III shows the average radii of the computed intervals. The notation, MD, M0, and M1 to M7, in the tables in this section indicates

- MD: The method (Rump, 1999) using CPU, as reference.
- M0: The method based on (20) using GPU (binary64 and binary32)
- Mk: The proposed method based on (38) using GPU, $k = 2, 3, ..., 7$

The method MD involves two matrix multiplications with double-precision and two matrix multiplications with single-precision:

$$\mathrm{fl}_{\nabla,64}(A_m B_m), \quad \mathrm{fl}_{\triangle,64}(A_m B_m), \quad \mathrm{fl}_{\triangle,32}(A'_m B'_r), \quad \mathrm{fl}_{\triangle,32}(A'_r S),$$

where $S := \mathtt{fl}((1 + 2u_{32})(B'_m + B'_r))$. The method M0 involves one matrix multiplication with double-precision and two matrix multiplications with single-precision:

$$\mathrm{fl}_{64}(A_m B_m), \quad \mathrm{fl}_{32}(A'_m P), \quad \mathrm{fl}_{32}(A'_r Q).$$

where $P, Q \in \mathbb{F}^{n\times p}$ are the upper bounds of $n\mathbf{u}_q B'_m + B'_r$ and $B'_m + B_r$, respectively. The method Mk has no matrix multiplication with double-precision and $\frac{1}{2}k(k-1) + 2$ matrix multiplications with single-precision.

Computing times for each method with and without Tensor Cores are shown in Table V. The time is an average of 10 examples. We underline comparable methods to M0. Notation TC indicates that cublasGemmEx in cuBLAS was used for $A^{(i)}B^{(j)}$, $i + j \leq k$. The ratio of computing time, the time without Tensor Cores is normalized to be one, is shown as AC in Table V. If the matrix dimension was 1,000, then using Tensor Cores slowed down performance because the cost of the scaling is not negligible. If the dimension was more than 3,000, we observed an acceleration of performance using Tensor Cores.

A summary of the numerical example from the tables is as follows:

Table III. Average of the radii ($n = 7,000,\ c_1 = 10^{-15}$).

c_2	MD	M0	M2	M3	M4	M5	M6	M7
10^{-15}	8.68e-12	3.47e-09	5.82e-01	2.19e-02	5.44e-04	1.16e-05	2.31e-07	4.38e-09
10^{-12}	2.23e-09	5.69e-09	5.82e-01	2.19e-02	5.43e-04	1.16e-05	2.32e-07	6.59e-09
10^{-9}	2.22e-06	2.23e-06	5.82e-01	2.18e-02	5.43e-04	1.38e-05	2.45e-06	2.23e-06
10^{-6}	2.22e-03	2.22e-03	5.84e-01	2.41e-02	2.77e-03	2.24e-03	2.22e-03	2.22e-03

Table IV. Average of the radii ($n = 7,000,\ c_1 = c_2$).

$c_1,\ c_2$	MD	M0	M2	M3	M4	M5	M6	M7
10^{-15}	8.68e-12	3.47e-09	5.82e-01	2.19e-02	5.44e-04	1.16e-05	2.31e-07	4.38e-09
10^{-12}	4.45e-09	7.92e-09	5.82e-01	2.19e-02	5.43e-04	1.16e-05	2.35e-07	8.83e-09
10^{-9}	4.45e-06	4.46e-06	5.82e-01	2.19e-02	5.47e-04	1.60e-05	4.68e-06	4.46e-06
10^{-6}	4.45e-03	4.45e-03	5.86e-01	2.64e-02	5.00e-03	4.46e-03	4.45e-03	4.45e-03

$c_2 = 1e-15$ and $c_2 = 1e-12$: The tightness of the interval by M0 and M7 was comparable. M7 was slightly slower than M0.

$c_2 = 1e-09$: The tightness of the interval by M0 and M6 was comparable. M6 was slightly faster than M0.

$c_2 = 1e-06$: The tightness of the interval by M0 and M5 was comparable. M5 was much faster than M0.

If the radius was sufficiently small compared with the magnitude of the center, the original method M0 was better than the proposed methods. Otherwise, the proposed methods were faster than M0 whereas tightness of the produced intervals was comparable.

Conclusion

We proposed methods for interval matrix multiplication using low-precision arithmetic based on error-free transformation. Tensor Cores efficiently accelerated the performance of interval matrix multiplication when matrix dimension was more than 3,000. The numerical examples illustrated that the proposed methods have an advantage of higher performance compared with those from previous works. Exploiting batched BLAS will be our future work for small matrices.

Table V. Computing time (sec) with and without Tensor Cores.

Dimension		MD	M0	M2	M3	M4	M5	M6	M7
		0.053	0.028	0.018	0.022	0.026	0.030	0.034	0.039
1000	TC	-	-	0.018	0.025	0.032	0.042	0.053	0.065
	AC	-	-	0.984	1.121	1.240	1.391	1.557	1.6523
		0.785	0.296	0.151	0.194	0.253	0.318	0.395	0.479
3000	TC	-	-	0.142	0.182	0.226	0.275	0.340	0.422
	AC	-	-	0.946	0.934	0.894	0.865	0.861	0.880
		3.580	1.190	0.542	0.698	0.906	1.152	1.476	1.802
5000	TC	-	-	0.495	0.627	0.787	0.944	1.178	1.452
	AC	-	-	0.913	0.899	0.868	0.819	0.798	0.805
		9.417	3.011	1.224	1.662	2.171	2.855	3.561	4.426
7000	TC	-	-	1.127	1.533	1.906	2.333	2.947	3.694
	AC	-	-	0.921	0.922	0.877	0.816	0.827	0.834
		20.139	6.097	2.466	3.316	4.479	5.879	7.570	9.509
9000	TC	-	-	2.219	2.779	3.481	4.147	4.983	6.211
	AC	-	-	0.899	0.838	0.777	0.705	0.658	0.653

Acknowledgements

This work was supported by JST CREST Grant Number JPMJCR14D4, Japan.

References

ANSI/IEEE 754-2008: IEEE Standard for Floating-Point Arithmetic, New York, 2008.

Basic Linear Algebra on NVIDIA GPUs, https://developer.nvidia.com/cublas.

Demmel J. and Hida Y. Accurate and efficient floating point summation. *SIAM Journal on Scientific Computing*, 25(4) , 1214–1248, 2004.

Frommer, A. and Hashemi, B. Verified computation of square roots of a matrix. *SIAM Journal on Matrix Analysis and Applications*, SIAM, 31(3):1279–1302, 2010.

Frommer A. and Hashemi B. Verified error bounds for solutions of Sylvester matrix equations. *Linear Algebra and its Applications*, 436(2):405–420, 2012.

Hoshi T., Ogita T., Ozaki K. and Terao T. An a posteriori verification method for generalized real-symmetric eigenvalue problems in large-scale electronic state calculations. *Journal of Computational and Applied Mathematics*, 376:112830, 2020.

Jeannerod, C.-P. and Rump S. M. Improved error bounds for inner products in floating-point arithmetic. *SIAM Journal on Matrix Analysis and Applications*, 34(2):338–334, 2013.

Kazal N.Y., Mukhlash I.,Arry S.B., Hidayat N. and Ozaki K. Extended use of error-free transformation for real matrix multiplication to complex matrix multiplication. *Journal of Physics: Conference Series*, to appear.

Kearfott, R. B., et al. Algorithm 737: INTLIB: a portable Fortran 77 interval standard-function library. ACM Transactions on Mathematical Software (TOMS), 20(4):447–459, 1994.

Minamihata A., Ozaki K., Ogita T. and Oishi S. Improved Extraction Scheme for Accurate Floating-point Summation. The 35th JSST Annual Conference, International Conference on Simulation Technology (JSST2016), Kyoto University (2016/10/27-29).

Miyajima S., Ogita T., Rump S.M. and Oishi S. Fast verification for all eigenpairs in symmetric positive definite generalized eigenvalue problem. *Reliable Computing*, 14:24–45, 2010.

Miyajima S. Numerical enclosure for each eigenvalue in generalized eigenvalue problem. *Journal of Computational and Applied Mathematics*, 236(9):2545–2552, 2012.

Miyajima S. Fast enclosure for solutions of Sylvester equations. *Linear Algebra and its Applications*, 439(4):856–878, 2013.

Moore, R. E. *Interval analysis*. Prentice-Hall, Englewood Cliffs, NJ, 1966.

Mukunoki D., Ozaki K., Ogita T. and Imamura T. DGEMM Using Tensor Cores, and Its Accurate and Reproducible Versions. In: Sadayappan P., Chamberlain B., Juckeland G., Ltaief H. (eds) High Performance Computing. ISC High Performance 2020. Lecture Notes in Computer Science, 12151:230–248, 2020.

Ogita, T. and Oishi, S. Fast Inclusion of Interval Matrix Multiplication. *Reliable Computing*, 11(3):191–205, 2005.

Ogita T., Rump S.M. and Oishi S. Verified solution of linear systems without directed rounding. Technical Report 2005-04, Advanced Research Institute for Science and Engineering, Waseda University, 2005.

Ogita T., Rump S.M. and Oishi S. Accurate sum and dot product. *SIAM Journal on Scientific Computing*, 26(6), 1955–1988, 2005.

Oishi S. Fast enclosure of matrix eigenvalues and singular values via rounding mode controlled computation. *Linear Algebra and its Applications*, 324(1–3): 133–146, 2001.

Ozaki, K., Ogita, T., Bünger. F. and Oishi, S. Accelerating interval matrix multiplication by mixed precision arithmetic. *Nonlinear Theory and its Applications*, 6(3):364–376, 2015.

Ozaki, K., Ogita, T. and Oishi, S. Error-free Transformation of Matrix Multiplication with a Posteriori Validation. *Numerical Linear Algebra with Applications*, 23(5):931–946, 2016.

Ozaki, K., Ogita, T., Oishi, S. and Rump S. M. Fast Algorithms for Floating-point Interval Matrix Multiplication. *Journal of Computational and Applied Mathematics*, 236:1795–1814, 2012.

Ozaki, K., Ogita, T., Oishi, S. and Rump S. M. Error-Free Transformation of Matrix Multiplication by Using Fast Routines of Matrix Multiplication and its Applications. *Numerical Algorithms*, 59(1):95–118, 2012.

Rump S. M. Fast and parallel interval arithmetic. *BIT Numerical Mathematics*, 39(3):539–560, 1999.

Rump S.M. Computer-assisted proofs and Self-Validating Methods. In B. Einarsson, editor, Handbook on Accuracy and Reliability in Scientific Computation, SIAM, 195-240, 2005.

Rump S.M. Ultimately fast accurate summation. *SIAM Journal on Scientific Computing*, 31(5), 3466–3502, 2009.

Rump S. M. Fast Interval Matrix Multiplication. *Numerical Algorithms*, 61(1):1–34, 2012.

Rump S. M. Computable backward error bounds for basic algorithms in linear algebra. *Nonlinear Theory and its Applications*, 6(3):360–363, 2015.

Rump, S. M., Ogita, T. and Oishi S. Accurate floating-point summation part I: Faithful rounding. *SIAM Journal on Scientific Computing*, 31(1):189–224, 2008.

Rump, S.M., Ogita T. and Oishi S. Accurate floating-point summation part II: sign, K-kold faithful and rounding to nearest. *SIAM Journal on Scientific Computing*, 31 (2), 1269–1302, 2008.

Strassen V. Gaussian elimination is not optimal. *Numerische Mathematik*, 13:354–356, 1969.

Sunaga, T. Theory of interval algebra and its application to numerical analysis. *RAAG memoirs*, 2:29–46, 1958.

Winograd S. On multiplication of 2 x 2 matrices. *Linear Algebra and its Applications*, 4:381–388, 1971.

Yamanaka N., Ogita T., Rump S.M. and Oishi S. A parallel algorithm of accurate dot product. *Parallel Computing*, 34 (6–8), 392–410, 2008.

Zhu Y.-K. and Hayes W.B. Correct rounding and a hybrid approach to exact floating-point summation. *SIAM Journal on Scientific Computing*, 31 (4), 2981–3001, 2009.

Zhu, Y.-K. and Hayes W.B. Algorithm 908: online exact summation of floating-point streams. *ACM Transactions on Mathematical Software*, 37 (3), 37:1–37:13, 2010.

Guaranteed Minimization of the Bit Error Ratio for Correlated MIMO Systems

Ekaterina Auer and Andreas Ahrens
Department of Electrical Engineering, University of Applied Sciences Wismar, 23966 Wismar, Germany, {ekaterina.auer, andreas.ahrens}@hs-wismar.de

Abstract. The multiple-input multiple-output (MIMO) mechanism is a method to increase the capacity of communication links by employing multiple transmitting and receiving antennas. An important quality criterion for such systems is the bit error probability characterizing the bit error ratio (BER). If simple MIMO transmission channel and data source models are assumed, BER can be computed analytically and optimized with respect to major parameters it depends on. In this contribution, we consider BER minimization for MIMO links with uncertainty in their parameters under good and poor scattering conditions, illustrated by close-to-life simulation examples. The obtained results show that the achievable performance depends strongly on the number of active MIMO layers, the assigned transmit power per layer and the bits per number of transmitted symbols.

Keywords: methods with automatic result verification, BER, power allocation, bit allocation

1. Introduction

The strategy of placing multiple antennas at the transmitter and receiver sides, well-known as multiple-input multiple-output (MIMO) method, improves both the capacity and the integrity of wireless systems through the use of the spatial characteristics of the underlying channel (Foschini, 1996; Telatar, 1999). A simple linear stochastic model for a frequency flat MIMO link consisting of n_{T} transmitting and n_{R} receiving antennas is

$$\vec{y} = H \cdot \vec{a} + \vec{n}, \quad \vec{y}, \vec{n} \in \mathbb{C}^{n_{\mathrm{R}}},\ \vec{a} \in \mathbb{C}^{n_{\mathrm{T}}},\ H \in \mathbb{C}^{n_{\mathrm{R}} \times n_{\mathrm{T}}}\ , \tag{1}$$

where $\vec{y}$ is the received data vector, $\vec{a}$ is the transmitted signal vector, $\vec{n}$ is the vector of the additive white Gaussian noise at the receiver side with the zero mean and the variance σ^2 in both real and imaginary parts, and H is the channel matrix. The channel matrix organizes the individual descriptions of the paths from every transmit antenna to every receive antenna. In wireless communications, those paths are found to be appropriately simulated by the Rayleigh distribution. That is, for a frequency flat MIMO link, the paths' descriptions (the coefficients of the $(n_{\mathrm{R}} \times n_{\mathrm{T}})$ channel matrix H) are assumed to be independently and identically distributed Rayleigh fading channels (Sklar, 1997) with equal variance δ^2. This model can be easily extended to reflect frequency selective channel conditions (Raleigh and Cioffi, 1998). In this paper, however, we focus on the frequency flat case.

©2021 *by E. Auer and A. Ahrens.*

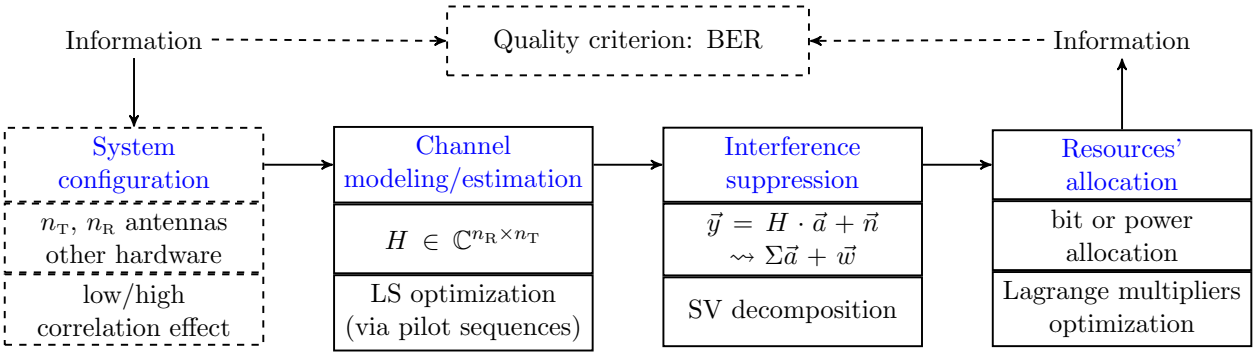


Figure 1. Modeling and simulation of a MIMO digital channel.

An important quality criterion for MIMO links is the bit error ratio (BER), the expectation value of which is characterized by the bit error probability. Among various factors influencing the BER are transmission channel noise (reflected in the linear stochastic model given above) and the interference between the different antenna data streams. Having established the configuration and having estimated the corresponding channel matrix H (cf. Figure 1), we can transform the MIMO link into a number of independent, weighted, frequency flat single-input single-output (SISO) layers using, for instance, singular value decomposition (SVD) to suppress the interference. As a rule, this can be achieved only for an *ideal* H and its SVD (perfect channel state information). The imperfect estimation of the channel coefficients can, in the worst case, lead to a loss of orthogonality in the SVD. It certainly causes a deviation of the ideal weights (or singular values) of the SISO links from the estimated ones and represents a source of uncertainty influencing the BER.

The quality of a MIMO transmission system (with such fixed parameters as the number of transmit and receive antennas, modulation level, and data throughput) can be improved by estimating the channel state more accurately, for example, by using more symbols in training (pilot) sequences (Weikert and Zölzer, 2007). However, the larger the length of such sequences the less data can be put through the channel. For this reason, it is necessary to study the influence of inaccuracy and uncertainty on the achievable transmission quality. To guarantee reliability of a communication with high data rates, the transmission parameters need to be adapted to current channel conditions, for which purpose reliable channel state information is crucial.

The channel matrix and therefore its singular values are influenced by spacial scattering conditions in the MIMO system configuration. The so-called antennas' *correlation effect* appears because of the proximity of the multiple antennas available at the transmitter and receiver sides. In consequence, transmit-to-receive antenna paths become too similar, which decreases the channel capacity and increases the BER. The antennas correlation effect has been extensively studied and analyzed wrt. space diversity possibilities (Lee, 1973; Wang et al., 2009).

For simple MIMO transmission channel and data source models, BER can be computed analytically. It should be as small as possible for the whole link. For quadrature amplitude modulated (QAM) signals, the bit error probability in the general form for a transmission (SISO) layer l is given as

$$P_b^{(l)} = f(M_l) \cdot \operatorname{erfc}\left(g(M_l, \lambda_l, \sigma^2, P_s^{(l)})\right) \quad . \tag{2}$$

Here, f, g are positive-valued functions depending on the constellation size M_l (and the number of bits per symbol $\log_2 M_l$), the noise variance σ^2, the available transmit power per layer $P_s^{(l)}$ and the singular value λ_l corresponding to the considered layer; $\mathrm{erfc}(\cdot)$ is the complementary error function. The weights λ_l are not necessarily equal for each SISO channel, which can be countered by bit and power allocation (BA, PA, resp.). That is, the analytical BER representation can be optimized wrt. the mentioned parameters[1] using, for example, the Lagrange multipliers approach. In poor scattering conditions with high antenna correlation, where the weighting of the SISO channels might turn strongly unequal, such optimization, or the process of bit and power allocation, becomes challenging. A unique indicator of the unequal weighting of the MIMO layers is the ratio ϑ between the smallest and the largest singular value which also characterizes the correlation effect. BER of the correlated system can become significantly higher than that of the uncorrelated one.

After the stage of channel estimation, the process in Figure 1 is influenced by uncertainty in λ_l and σ^2, whereas M_l and $P_s^{(l)}$ can be known exactly or allocated appropriately. For example, (numerical or/and measurement) errors might affect the singular values, which can be assumed to be bounded by the interval $[\ \lambda_l - \varepsilon_l, \lambda_l + \varepsilon_l\ ]$ if the errors are are bounded. From lab simulations, ε_l can be assessed as being up to one order of magnitude smaller than λ_l itself for larger singular values and of the same order of magnitude for small ones. That is, in the linear model in Eq. (1), both stochastic and bounded uncertainty can be present. In (Auer et al., 2018b; Auer and Ahrens, 2020), we employed mixed analytical-numerical strategies relying on methods with automatic result verification to deal with BER minimization under uncertainty in the parameters given above. From the point of view of application, such strategies are of interest, for example, for delay-critical real time interactive voice or video communication systems, where the BER has to be minimized at a fixed data rate under uncertain conditions.

Interval analysis (Moore et al., 2009) and other so-called *methods with result verification* can take into account and propagate bounded uncertainty in parameters directly through an implementation of a mathematical model (provided that the appropriate implementation exists). Moreover, these methods address the question of system reliability by proving formally that the outcome of a simulation implemented on a computer is correct (assuming that the underlying implementation is correct). The results are usually sets of floating point numbers which with certainty contain the exact solution to the model. As a common drawback, the possibility of too wide bounds for the solution sets should be mentioned (e.g., between $-\infty$ and $+\infty$). This drawback is caused by the dependency problem or the wrapping effect (Lohner, 2001). A literature overview and a suggestion for a result verification scheme and dealing with uncertainty using interval analysis in the overall process from Figure 1 are given in (Auer et al., 2018a).

In this paper, we take a closer look at the difference between correlated and uncorrelated systems under uncertainty using the criterion of BER. In general, the performance of MIMO systems is significantly affected by the choice of the number of bits per symbol $\log_2 M_l$ and the appropriate allocation of the transmit power $P_s^{(l)}$ per SISO layer l. A further important factor is the number of activated layers L. To compare the behaviour of correlated and uncorrelated MIMO links, we study the influence of $P_s^{(l)}$, $\log_2 M_l$ and L on them for a fixed throughput. As data, we consider

[1] Note that the noise variance σ^2 is usually considered to be fixed.

two sets for a correlated and an uncorrelated channel with 5000 realizations each (obtained in a lab simulation).

The paper is structured as follows. First, we describe in detail how the data for the comparison we intend to perform are generated, with the focus on channel matrices exhibiting high correlation and on interference suppression, in Section 2. Techniques for channel estimation are to be found, for example, in (Weikert and Zölzer, 2007), and are outside the scope of this paper. In Section 3, we describe approaches to bit and power allocation under bounded uncertainty in parameters, which rely on methods with result verification. A comparison of BA and PA results for the above mentioned channels with low and high correlation is in Section 4. Conclusions and an outlook on future research are in the last section.

2. Correlated Channel Simulation and Interference Suppression

In this section, we describe classical techniques for channel simulation taking into account the correlation effect as well as for interference suppression. They were used to produce the data sets necessary for the comparison in Section 4. Whereas we generated independent and Rayleigh distributed coefficients with equal variance to obtain realizations of the uncorrelated channel matrix, the approach from Section 2.1 was necessary in the case of the correlated MIMO channel. After that, the techniques from Section 2.2 were used for interference suppression. In both cases, MATLAB was employed for implementation.

2.1. MIMO Channel Correlation

While estimating the channel matrix H, it is quite common to assume that its coefficients are independent and Rayleigh distributed with equal variance. However, correlations between the transmit and receive antennas, respectively, cannot be ignored in many cases. The way to include the antenna signal correlation into the MIMO channel model for Rayleigh flat-fading channels is given by (Oestges, 2006) and results in

$$\mathrm{vec}(H) = R_{\mathrm{HH}}^{1/2} \cdot \mathrm{vec}(G) \tag{3}$$

where G is a $(n_{\mathrm{R}} \times n_{\mathrm{T}})$ uncorrelated channel matrix with independent, identically Rayleigh distributed complex elements and $\mathrm{vec}(\cdot)$ being the operator stacking the matrix G into a vector columnwise. The matrix R_{HH} describing the correlation within the channel coefficients $h_{\nu,\mu}$ (with $\nu = 1, \ldots, n_{\mathrm{R}}$ and $\mu = 1, \ldots, n_{\mathrm{T}}$) is defined as

$$R_{\mathrm{HH}} = \mathrm{E}\left\{\mathrm{vec}(H) \cdot \mathrm{vec}(H)^{*\,\mathrm{T}}\right\} \tag{4}$$

with $\mathrm{vec}(H)$ resulting for the example of a system with two transmit and two receive antennas in

$$\mathrm{vec}(H) = \begin{pmatrix} h_{1,1} \\ h_{2,1} \\ h_{1,2} \\ h_{2,2} \end{pmatrix} \quad . \tag{5}$$

Assuming that the correlation introduced by the antenna elements at the transmitter side is independent from the correlation introduced by the antenna elements at the receiver side, the correlation matrix can be defined using the transmitter side correlation matrix R_{TX} and the receiver side correlation matrix R_{RX} as

$$R_{\mathrm{HH}} = R_{\mathrm{TX}} \otimes R_{\mathrm{RX}} \ , \tag{6}$$

where $\otimes$ represents the Kronecker product. For a (2×2) MIMO system, the receiver side correlation matrix R_{RX} is given by

$$R_{\mathrm{RX}}^{(2\times2)} = \begin{pmatrix} \rho_{1,1}^{(\mathrm{RX})} & \rho_{1,2}^{(\mathrm{RX})} \\ \rho_{2,1}^{(\mathrm{RX})} & \rho_{2,2}^{(\mathrm{RX})} \end{pmatrix} = \begin{pmatrix} 1 & \rho^{(\mathrm{RX})} \\ \rho^{*\,(\mathrm{RX})} & 1 \end{pmatrix} \ . \tag{7}$$

The receiver side correlation coefficient can be calculated as

$$\rho_{m,n}^{(\mathrm{RX})} = E\{h_{m,k} \cdot h_{n,k}^{*}\} \ . \tag{8}$$

It describes the correlation between the receive antennas m and n, independent from the transmit antenna k. It should reflect the fact that the value of the correlation coefficient depends on the reference antenna, that is,

$$\rho_{m,n}^{(\mathrm{RX})} = E\{h_{n,k} \cdot h_{m,k}^{*}\} = \rho_{n,m}^{*\,(\mathrm{RX})} \ . \tag{9}$$

Hence, the elements of the correlation matrix symmetric wrt. the main diagonal are complex conjugate. This relationship is due to the sign change in the distance difference between antennas depending on the antenna reference. Similarly, the transmitter-side correlation matrix R_{TX} results in

$$R_{\mathrm{TX}}^{(2\times2)} = \begin{pmatrix} \rho_{1,1}^{(\mathrm{TX})} & \rho_{1,2}^{(\mathrm{TX})} \\ \rho_{2,1}^{(\mathrm{TX})} & \rho_{2,2}^{(\mathrm{TX})} \end{pmatrix} = \begin{pmatrix} 1 & \rho^{(\mathrm{TX})} \\ \rho^{*\,(\mathrm{TX})} & 1 \end{pmatrix} \ . \tag{10}$$

The transmitter side correlation coefficient is obtained as

$$\rho_{k,\ell}^{(\mathrm{TX})} = E\{h_{m,k} \cdot h_{m,\ell}^{*}\} \ . \tag{11}$$

Analogously to the receiver side, it describes the correlation between the transmit antennas k and ℓ, independent from the receive antenna m. Finally, the overall correlation matrix R_{HH} in the example (2×2) system is given by the elements

$$R_{\mathrm{HH}}^{(2\times2)} = \begin{pmatrix} \rho_{1,1,1,1} & \rho_{1,1,1,2} & \rho_{1,2,1,1} & \rho_{1,2,1,2} \\ \rho_{1,1,2,1} & \rho_{1,1,2,2} & \rho_{1,2,2,1} & \rho_{1,2,2,2} \\ \rho_{2,1,1,1} & \rho_{2,1,1,2} & \rho_{2,2,1,1} & \rho_{2,2,1,2} \\ \rho_{2,1,2,1} & \rho_{2,1,2,2} & \rho_{2,2,2,1} & \rho_{2,2,2,2} \end{pmatrix} \tag{12}$$

where $\rho_{k,\ell,m,n}$ are computed as

$$\rho_{k,\ell,m,n} = E\{h_{m,k} \cdot h_{n,\ell}^{*}\} = \rho_{k,\ell}^{(\mathrm{TX})} \cdot \rho_{m,n}^{(\mathrm{RX})} \ . \tag{13}$$

For the model in Equation (3), the square root of the matrix R_{HH} can be computed using Cholesky decomposition as $R_{\mathrm{HH}}^{1/2} = \left(R_{\mathrm{TX}}^{1/2} \otimes R_{\mathrm{RX}}^{1/2}\right)$.

2.2. Interference Suppression

Having generated the channel matrix H, we decompose it as $H = U \cdot \Sigma \cdot V^\dagger$, where U and V are unitary matrices, Σ is the diagonal matrix with real elements and $V^\dagger$ denotes the Hermitian adjoint of V. The matrix Σ contains the positive square roots of the eigenvalues ξ_l of $H^\dagger H$ in descending order on the main diagonal (singular values denoted by $\lambda_l = \sqrt{\xi_l}$ throughout the paper). If a pre-processed data vector $\vec{x} := V \cdot \vec{a}$ is considered and the corresponding receive signal $\vec{z} := H\vec{x} + \vec{n}$ is post-processed by $U^\dagger$, then the new receive signal is

$$\vec{u} := U^\dagger \vec{z} = U^\dagger \left(U \Sigma V^\dagger \right) V\vec{a} + U^\dagger \vec{n} = \Sigma \vec{a} + \vec{w} \ , \tag{14}$$

where the vector $\vec{a}$ is the transmitted signal vector and $\vec{n}$ is the Gaussian noise vector as explained in the Introduction. In this way, the MIMO link is transformed (ideally) into $L = \min\{n_\mathrm{T}, n_\mathrm{R}\}$ independent, non-interfering SISO layers u_l having (unequal) weights λ_l:

$$u_l = \lambda_l a_l + w_l \quad \text{for} \quad l = 1 \ldots L \ . \tag{15}$$

There are a number of difficulties during this stage due to numerics or uncertainty. For example, $U^\dagger \cdot U$ or $V^\dagger \cdot V$ are not exactly identity matrices because of numerical errors, which might create interference between the SISO links. Moreover, the elements of the matrix H can never be obtained perfectly in a real life situation. That is, H includes bounded uncertainty in its elements in addition to the stochastic one modeled by the Rayleigh distribution, making the weights λ_l also uncertain.

The possibilities to apply methods with result verification (or other set-based techniques) to SVD under the assumption of bounded uncertainty in H are limited. It is possible to obtain (verified) bounds on (real) singular values of H (Deif, 1991; Li and Chang, 2003; Su et al., 2011; Hladík et al., 2010; Rump, 2011). However, the interval equivalents of matrices U and V do not convey the same meaning as their floating-point counterparts as explained in (Kearfott, 1996). This prohibits any kind of set-based decomposition in the manner given by Eq. (14) and (15). (Note that as long as the matrices themselves are not needed, the bounds for λ_l can be used in the same way as in Eq. (15).) However, a MIMO link can be also decomposed into several SISO links based on the geometric mean decomposition (Ahrens et al., 2016), the study of which from the point of view of verified techniques is a subject of our future work. In this paper, we take lowest and highest values for each λ_l from the lab simulation data. Note that the proposed approach would work according to the same principles with the analytical bounds. Additionally, since the number of simulations is moderately large (5000) and the model in Eq. (1) takes into account the Gaussian noise, bounds for λ_l based on these data should not be too far from the analytical or real life values.

2.3. Comparison of Correlated and Uncorrelated Channel Realizations

Due to the proximity of the antennas at the transmitter and the receiver side, the unequal weighting of the MIMO layers can become stronger (antennas correlation effect as described in the Introduction). The ratio ϑ between the smallest and the largest singular values can be considered as the unique indicator of the unequal weighting of the MIMO layers. In Figure 2, the frequencies of occurrence for the values of the weighting factor ϑ are shown for the example of an uncorrelated

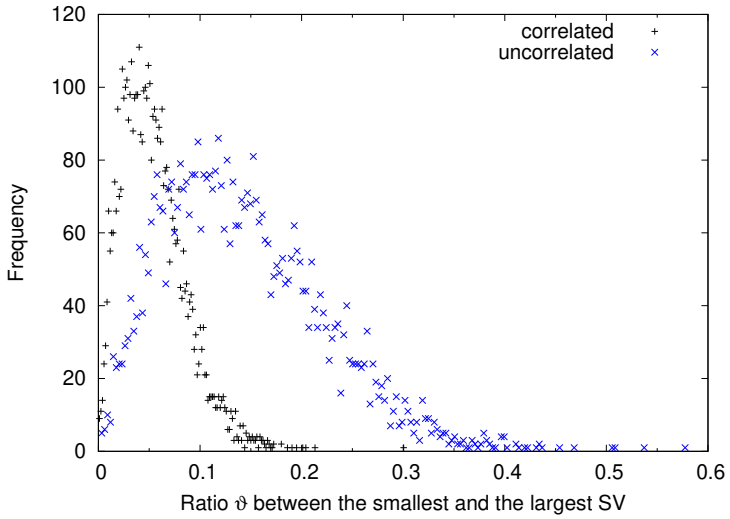


Figure 2. Frequencies of values for the ratio ϑ for uncorrelated ($\times$) and correlated ($+$) frequency flat (4×4) MIMO channel.

and a correlated frequency flat (4×4) MIMO system. (The corresponding frequencies for the amplitudes of the values are shown in Figure 3.) Figure 3 demonstrates how the difference between the largest and the smallest singular values (i. e., the unequal weighting) increases for the correlated system. This means that the ratio between the smallest and the largest singular value decreases as the correlation increases, which can be observed from Figure 2. That is, the probability of having predominant layers also increases. As a result, the use of resource allocation techniques seems an appropriate solution to optimize the layer behavior. It seems that no power should be allocated to the MIMO layer having the smallest singular value because the overall performance would be deteriorated. This claim is substantiated and extended in Section 4.

3. Guaranteed BER Minimization Through Power and Bit Allocation

After the SISO weights λ_l are computed for a given MIMO link, it is necessary to think about how to minimize its overall BER. In this Section, we describe possibilities for power and bit allocation under uncertainty in in the weights. In (Auer and Ahrens, 2020), more information is given on this overall subject and, in particular, on possibilities to minimize wrt. other parameters. Additionally, the number of activated layers $L \leq \min\{n_\mathrm{T}, n_\mathrm{R}\}$ might play an important role. Not every layer $l \in \{1, \ldots, L\}$ of the MIMO link should necessarily be active to achieve the best performance. In Section 4, we consider possible choices for active layers in detail. Here, we assume that L is fixed within a certain suitable range.

3.1. Enclosing the BER for Uncertain Parameters

We work with the following formula to compute the the overall BER (for derivation, see, for example, (Auer et al., 2018b)):

$$P_b = \frac{2}{\sum_{l=1}^{L} \log_2 M_l} \sum_{l=1}^{L} \left(1 - \frac{1}{\sqrt{M_l}}\right) \cdot \operatorname{erfc}\left(\frac{\lambda_l}{2\sigma}\sqrt{\frac{3 \cdot P_s}{L(M_l - 1)}}\right) \quad , \tag{16}$$

with P_s being the overall available transmit power. The coefficient $\dfrac{2}{\sum_{l=1}^{L} \log_2 M_l}$ is constant for a fixed throughput and is independent of the number of effectively activated layers.

From Eq. (16), it is evident that the major charasteristics influencing the BER are the singular values (layer weights), the noise variance and the number of bits per symbol. Initially, the transmit power per layer $P_s^{(l)} = P_s/L$ is supposed to be equal for each layer. As shown in (Auer and Ahrens, 2020), if $\lambda_l \in [\ \underline{\lambda}_l, \overline{\lambda}_l\ ]$, where $\underline{\lambda}_l$, $\overline{\lambda}_l$ are known lower and upper bounds, respectively, and the standard deviation $\sigma \in [\ \underline{\sigma}, \overline{\sigma}\ ]$, then a conservative upper bound on the BER can be obtained as

$$P_b(\sigma, \lambda_1 \dots \lambda_L) \leq \frac{2}{\sum_{l=1}^{L} \log_2 M_l} \sum_{l=1}^{L} \left(1 - \frac{1}{\sqrt{M_l}}\right) \cdot \operatorname{erfc}\left(\frac{\underline{\lambda}_l}{2\overline{\sigma}}\sqrt{\frac{3P_s}{L(M_l - 1)}}\right) \quad . \tag{17}$$

That means that, due to monotonicity of the involved functions, it is not necessary to work with actual ranges but with their bounds only, which makes verified optimization easier.

3.2. Power Allocation

If the number of transmitted bits per layer is fixed, a possible approach to BER minimization is the power allocation technique assigning more power to the layers with small weights. This becomes necessary because small λ_l lead to large values of bit error probability $P_b^{(l)}$ per MIMO layer l since $\operatorname{erfc}(\cdot)$ is monotonically decreasing with λ_l. If the Lagrange multipliers method is used, then the following cost function needs to be minimized:

$$\begin{aligned} J(\pi_1 \dots \pi_L, \mu) \ = \ & \frac{2}{\sum_{l=1}^{L} \log_2 M_l} \sum_{l=1}^{L} \left(1 - \frac{1}{\sqrt{M_l}}\right) \\ & \cdot \ \operatorname{erfc}\left(\frac{\pi_l \lambda_l}{2\sigma}\sqrt{\frac{3 \cdot P_s}{L(M_l - 1)}}\right) + \mu \left(\sum_{l=1}^{L} \pi_l^2 - L\right) \longrightarrow \min, \end{aligned} \tag{18}$$

where $\pi_1 > 0, \dots, \pi_L > 0$ are the power allocation parameters with which we modify the weights λ_l from Eq. (15) in order to improve P_b from Eq. (16) and μ is the Lagrange multiplier assigned to

the constraint $\sum_{l=1}^{L} \pi_l^2 = L$. With the notations

$$k_l = k_l(M_1 \dots M_L) \;:=\; \frac{2}{\sum_{l=1}^{L} \log_2 M_l} \cdot \left(1 - \frac{1}{\sqrt{M_l}}\right), \tag{19}$$

$$c_l = c_l(M_l, \sigma, P_s, L) \;:=\; \frac{1}{2\sigma}\sqrt{\frac{3 \cdot P_s}{L(M_l - 1)}},\; l = 1...L\;, \tag{20}$$

the Lagrange multipliers approach produces the nonlinear system of equations (21) for stationary points of the cost function (18) :

$$\frac{\partial J(\pi_1 \dots \pi_L, \mu)}{\partial \pi_l} = -\frac{2k_l}{\sqrt{\pi_l}}\left(c_l \lambda_l e^{-c_l^2 \lambda_l^2 \pi_l^2}\right) + 2\mu\pi_l = 0, \quad \sum_{l=1}^{L} \pi_l^2 - L = 0\;, \tag{21}$$

where $\pi_l > 0, l \in \{1, \dots, L\}$. Moreover, it is clear from the first L equations that μ must be positive.

The second derivative $\dfrac{\partial^2 J}{\partial \pi_l \partial \pi_m} = 0$ for $l \neq m$ and is positive for $l = m$. The bordered Hessian is symmetric and has the form

$$\begin{pmatrix} 0 & 2\pi_1 & \cdots & 2\pi_L \\ 2\pi_1 & 2\mu + \dfrac{4k_1 c_1^2 \lambda_1^2}{\sqrt{\pi_1}} \pi_1 e^{-c_1^2 \lambda_1^2 \pi_1^2} & \cdots & 0 \\ 2\pi_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 2\pi_L & 0 & \cdots & 2\mu + \dfrac{4k_L c_L^2 \lambda_L^2}{\sqrt{\pi_L}} \pi_L e^{-c_L^2 \lambda_L^2 \pi_L^2} \end{pmatrix}$$

The determinants of all relevant $L - 1$ leading principal minors have the structure

$$(l+1) \times (l+1): \qquad \begin{vmatrix} 0 & a_1 & \cdots & a_l \\ a_1 & d_1 & \cdots & 0 \\ \vdots & & \ddots & \\ a_l & 0 & \cdots & d_l \end{vmatrix} = -\sum_{i=1}^{l} \left(a_i^2 \prod_{k=1, k\neq i}^{l} d_k \right)$$

for $l = 2 \dots L$. The elements d_k are positive for $\pi_k > 0$ $(k = 1 \dots l)$, $\mu > 0$, that is, the determinants are always negative. For one constraint, this means that a stationary point is a local minimum. Using a solver for systems of non-linear equations based on methods with result verificaiton (e.g., C-XSC Toolbox (Hofschuster et al., 2008)), we can compute a guaranteed enclosure for it from the system in Eq. (21). Working in a verified way has the advantages of taking care of numerical errors and of the possibility to prove the uniqueness of the solution (leading to the proof that the minimum is global) on a computer.

3.3. Influence of Bits per Symbol

The number of bits per transmitted symbol and layer influences the overall BER. For example, we studied the case of $M_1 = M_2 = 16$ and $M_1 = 64$, $M_2 = 4$ for the (4×4) MIMO system with a fixed throughput of 8 bit/s/Hz and two active layers in (Auer et al., 2018b). The latter case had a better BER for the same signal-to-noise ratio (SNR) and MIMO links with both high and low correlation coefficients. Therefore, it is necessary to determine optimal constellations (possibly, under additional consideration of power allocation).

The optimization with respect to bits per symbol is complicated by the fact that M_l should be positive integer powers of 2 fitting the desired throughput, which leads to a non-linear mixed-integer programming problem (Lee and Leyffer, 2011). The BER to minimize is

$$P_b(M_1 \dots M_L) = \frac{2}{T} \sum_{l=1}^{L} \left(1 - \frac{1}{\sqrt{M_l}}\right) \cdot \operatorname{erfc}\left(\frac{\hat{c}_l}{\sqrt{M_l - 1}}\right) \longrightarrow \min \quad \text{s.t.} \quad \sum_{l=1}^{L} \log_2 M_l = T, \tag{22}$$

where $\hat{c}_l = \hat{c}_l(\lambda_l, \pi_l, \sigma, P_s, L) := \dfrac{\lambda_l \pi_l}{2\sigma} \sqrt{\dfrac{3P_s}{L}}$ and T the desired throughput. The problem can be solved for small T by a brute force (trial and error) approach using every admissible positive integer solution to the problem in Eq. (22). This is the case for a fixed throughput of 8 bit/s/Hz we consider in the next section. In general, it is necessary to augment the approach by a branch and bound algorithm.

4. Numerical Results: BER Optimization for a Correlated and Uncorrelated Wireless MIMO Link with Four Antennas

In this section, we illustrate the ideas using a practical problem. We consider a correlated and uncorrelated wireless frequency flat MIMO link with $n_{\mathrm{T}} = n_{\mathrm{R}} = 4$ antennas. The desired throughput is 8 bit/s/Hz and the available transmit power $P_s = 1\mathrm{W}$ throughout the simulations.

4.1. Simulation Settings

Two data sets with 5000 elements (channel realizations) each were obtained for both variants of the wireless channel in a non-verified simulation with $\delta^2 = \frac{1}{2}$ as described in Sections 2.1, 2.2 (data available on demand via email). The correlation coefficients at the transmitter and receiver sides were chosen as $\rho^{(\mathrm{RX})} = \rho^{(\mathrm{TX})} = 0.2375$. From Figure 2, it can be seen that the ratio ϑ for the correlated channel is significantly lower in the produced data sets. The frequencies of occurrence for the amplitudes of all four singular values are shown in Figure 3.

We provide numerical results for the SNR of $\left(\frac{E_s}{N_0}\right)_{\mathrm{dB}} = 10$ dB, where E_s denotes the symbol energy and N_0 the noise power spectral density with the corresponding value for σ computed as $\sigma = \sqrt{\dfrac{P_s}{2 \cdot \exp\left(\frac{\ln 10}{10} \cdot \left(\frac{E_s}{N_0}\right)_{\mathrm{dB}}\right)}} = \sqrt{\frac{1}{20}} \approx 0.2236$. All the following simulations are performed with result verification using C-XSC Toolbox (Hofschuster et al., 2008). The shown bounds are rounded

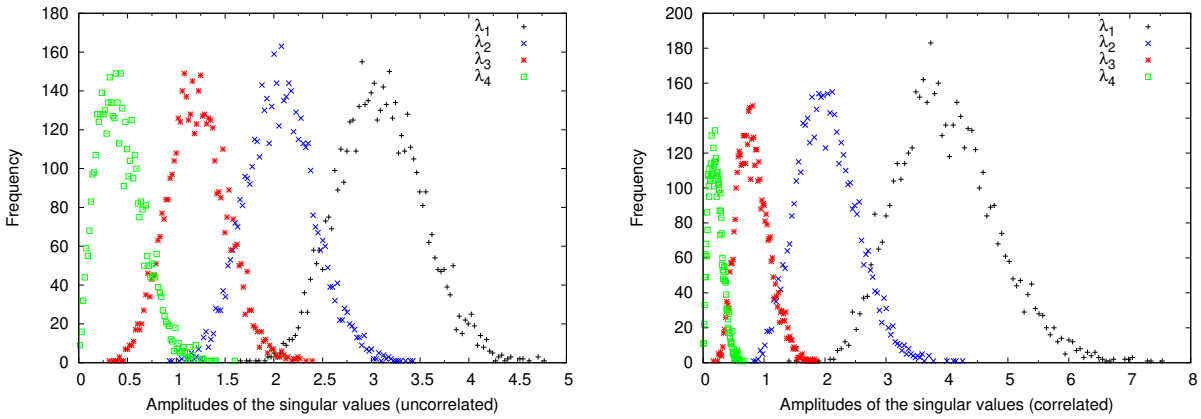


Figure 3. Amplitudes of the singular values in the considered data sets for the uncorrelated (left) and the correlated (right) channels.

outwards to the fourth digit after the decimal point; the actual interval widths are somewhat smaller. In practice, it is accurate enough to store the data up to the fourth digit after the decimal point since the changes in digits in places after that do not have much influence on the overall system. If the widths of enclosures are small, we reproduce only one (rounded) value.

4.2. Power and Bit Allocation

The lower bounds for λ_i for the data set in the uncorrelated case are $\underline{\lambda}_1 = 2.4208$, $\underline{\lambda}_2 = 0.9180$, $\underline{\lambda}_3 = 0.3022$, $\underline{\lambda}_4 = 0.0049$. The correlation coefficient for these bounds is $\vartheta \approx 20 \times 10^{-4}$, with the worst possible one from the data set equal to approx. 18×10^{-4}. That is, the upper bound for the BER at 10 dB for this system with all four layers active and $M_1 = 4$, $M_2 = 4$, $M_3 = 4$ and $M_4 = 4$ is 0.2217 (the outer bound). The power allocation reduces it to the value of 0.1974. The lower singular value bounds in the correlated case are $\underline{\lambda}_1 = 1.3791$, $\underline{\lambda}_2 = 0.5526$, $\underline{\lambda}_3 = 0.1609$, $\underline{\lambda}_4 = 0.0013$ ($\vartheta \approx 9.4 \times 10^{-4}$, the worst from the data set 2.5×10^{-4}). Here, the upper bound on the BER under the same conditions is 0.2761, reduced to 0.2553 by PA, that is, significantly higher than in the uncorrelated case. Note that the actual worst case BER for the simulated data are 0.1909 and 0.2180, respectively. That is, the bounds obtained with interval arithmetic are conservative. In Figure 4, the BER with and without power allocation is shown for each crisp set of 5000 λ_l values in both the uncorrelated (left) and correlated (right) case for $L = 4$. For better presentation, only every 100th data point is plotted. It can be seen that the BER is worse for the correlated case.

In general, the BER is the worst in both the correlated and uncorrelated case if all four MIMO layers are active as can be seen from Table I. Especially, the choice of $M_1 = 4$, $M_2 = 4$, $M_3 = 4$ and $M_4 = 4$ (number 14 in the Table) is unfavorable for both systems' BER. In the Table, the results of optimization wrt. the transmission power, the number of bits per symbol and the number of active layers at 10 dB are shown for the uncorrelated and correlated case as intervals with the best and worst possible bound from the respective data sets. The best constellations are highlighted in bold face, with the second-best shown in italics. Since the $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$, putting more

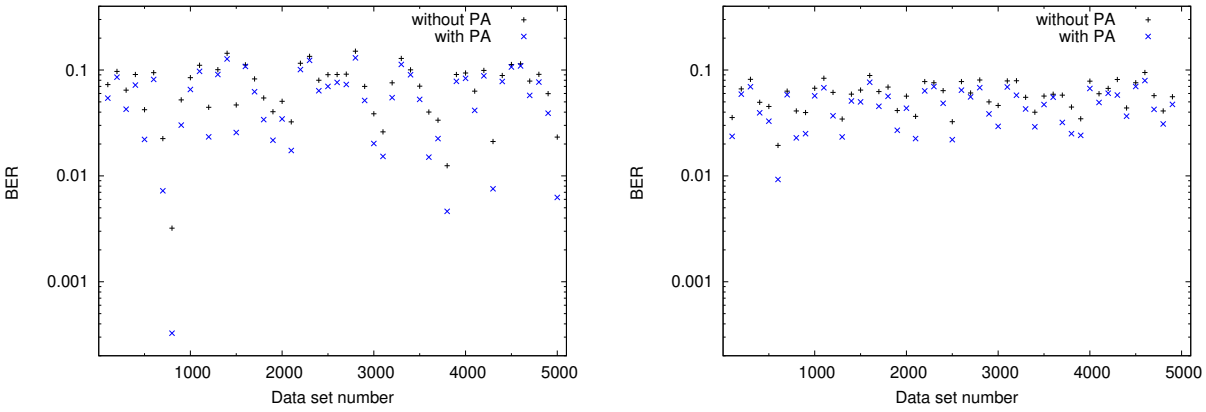


Figure 4. BER with and without power allocation for 5000 possibilities for λ_l (at 10 dB) in the uncorrelated (left) and correlated (right) case with active four layers and $M_1 = 4$, $M_2 = 4$, $M_3 = 4$, $M_4 = 4$.

symbols through the weaker layers does not improve the BER so that only the situations where $M_1 \geq M_2 \geq M_3 \geq M_4$ need to be considered.

The Table is consistent with the knowledge that correlated MIMO systems have a higher BER. Moreover, we can notice that the power allocation technique does not reduce the BER as much in the correlated case as in the uncorrelated one. Additionally, we can conclude that the number of active layers for these (4×4) MIMO systems needs to be either two or three, since they result in the best BER in both cases. Almost every choice of symbols per layer with $L = 2$ and $L = 3$ results in a good BER in the uncorrelated case aside from the marginal cases $M_1 = 128$, $M_2 = 2$ and $M_1 = 64$, $M_2 = 2$, $M_3 = 2$ (with $M_1 = 32$, $M_2 = 8$; $M_1 = 16$, $M_2 = 8$, $M_3 = 2$ or $M_1 = 32$, $M_2 = 4$, $M_3 = 2$ being the best constellations). In the correlated case, the resource allocation plays a more important role since only two constellations lead to a good BER, namely, $M_1 = 32$, $M_2 = 8$ and $M_1 = 32$, $M_2 = 4$, $M_3 = 2$.

5. Conclusions

In this paper, we presented a comparison of correlated and uncorrelated MIMO systems from the point of view of resource allocation. For this purpose, we relied on examples of corresponding (4×4) links with 5000 channel realisations each obtained from MATLAB simulations. Note that the resource allocation was performed with result verification, that is, the achieved minima are proved to be unique. We studied both MIMO systems wrt. power and bit allocation as well as wrt. the number of active layers. The important conclusions are:

— All four layers should never be activated at the same time, that is, the weakest layer should be switched off;

— For correlated systems, the resource allocation plays an especially important role; as a rule, working with two active layers results in the best performance for the BER.

Table I. Analysis of BER for 5000 channel realizations at 10 dB with throughput of 8 bit/s/Hz (high and low correlation, with and without power allocation).

	Layer M_1,M_2,M_3,M_4	BER (correlated)	BER-PA (correlated)	BER (uncorrelated)	BER-PA (uncorrelated)
			One active layer		
1	256, 0, 0, 0	[0.0023,0.1492]	the same	[0.0240, 0.13423]	the same
			Two active layers		
2	128, 2, 0, 0	[0.0022,0.1449]	[0.0001,0.1220]	[0.0232, 0.1304]	[0.0059, 0.1036]
3	64, 4, 0, 0	[$55{\cdot}10^{-6}$,0.1103]	[$4{\cdot}10^{-6}$,0.0959]	[0.0044, 0.0928]	[0.0007, 0.0749]
4	32, 8, 0, 0	**[$10{\cdot}10^{-7}$,0.0808]**	**[$2{\cdot}10^{-7}$,0.0773]**	**[0.0002, 0.0599]**	*[0.0001, 0.0556]*
5	16, 16, 0, 0	[$40{\cdot}10^{-7}$,0.1092]	[$3{\cdot}10^{-7}$,0.0981]	[$1.4{\cdot}10^{-4}$, 0.06710]	[$1.1{\cdot}10^{-4}$, 0.0589]
			Three active layers		
6	64, 2, 2, 0	[$8{\cdot}10^{-4}$,0.1279]	[$8{\cdot}10^{-6}$,0.1002]	[0.0127, 0.1121]	[0.0009, 0.0771]
7	32, 4, 2, 0	*[$12{\cdot}10^{-6}$,0.0926]*	*[$2{\cdot}10^{-6}$,0.0775]*	[0.0015, 0.0739]	**[$6{\cdot}10^{-5}$, 0.0533]**
8	16, 8, 2, 0	[$11{\cdot}10^{-6}$,0.1006]	[$5{\cdot}10^{-6}$,0.0936]	*[0.0001, 0.06417]*	[$2{\cdot}10^{-5}$, 0.0584]
9	16, 4, 4, 0	[$11{\cdot}10^{-5}$,0.1015]	[$1{\cdot}10^{-5}$,0.0972]	[$9{\cdot}10^{-5}$, 0.0850]	[$1{\cdot}10^{-5}$, 0.0785]
10	8, 8, 4, 0	[0.0001,0.1429]	[$7{\cdot}10^{-5}$,0.1282]	[$2{\cdot}10^{-5}$, 0.1048]	[$1{\cdot}10^{-5}$, 0.0916]
			Four active layers		
11	32, 2, 2, 2	[0.0106,0.1532]	[0.0032,0.1255]	[0.0073, 0.1426]	[0.0005, 0.1129]
12	16, 4, 2, 2	[0.0071,0.1252]	[0.0023,0.1181]	[0.0006, 0.1099]	[$7{\cdot}10^{-5}$, 0.1010]
13	8, 4, 4, 2	[0.0109,0.1665]	[0.0038,0.1529]	[$7{\cdot}10^{-5}$, 0.1419]	[$4{\cdot}10^{-5}$, 0.1344]
14	4, 4, 4, 4	[0.0414,0.2180]	[0.0228,0.2028]	[0.0014, 0.1909]	[0.0002, 0.1785]

However, we did not consider explicitly the influence of the noise variance σ^2 on the results. All simulations were performed for the SNR of 10 dB. To study this influence is a subject of our future work.

References

Ahrens, A., Benavente-Peces, C., and Cano-Broncano, F. (2016). Power Allocation in SVD- and GMD-assisted MIMO Systems. *Optimization and Engineering*.

Auer, E. and Ahrens, A. (2020). Guaranteed Minimization of the Bit Error Ratio for MIMO Systems: A Mathematical Viewpoint. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems Part B: Mechanical Engineering*. accepted.

Auer, E., Ahrens, A., Sandman, A., and Benavente-Peces, C. (2018a). Influence of uncertainty and numerical errors in the context of MIMO systems. In *8th International Workshop on Reliable Engineering Computing*, pages 21–30.

Auer, E., Benavente-Peces, C., and Ahrens, A. (2018b). Solving the Power Allocation Problem Using Methods with Result Verification. *Int. J. Reliability and Safety*, 12(1/2):86–102.

Deif, A. (1991). Singular values of an interval matrix. *Linear Algebra and its Applications*, 151:125 – 133.

Foschini, G. (1996). Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Labs Technical Journal*, 1(2):41 – 59.

Hladík, M., Daney, D., and Tsigaridas, E. P. (2010). Bounds on real eigenvalues and singular values of interval matrices. *SIAM Journal of Matrix Analysis and Applications*, 31(4):2116–2129.

Hofschuster, W., Krämer, W., and Neher, M. (2008). C-XSC and Closely Related Software Packages. In Cuyt, A. M., Krämer, W., Luther, W., and Markstein, P. W., editors, *Numerical Validation in Current Hardware Architectures LNCS 5492*, pages 68–102.

Kearfott, R. B. (1996). *Rigorous Global Search: Continuous Problems*. Kluwer.

Lee, J. and Leyffer, S. (2011). *Mixed Integer Nonlinear Programming*. Springer Publishing Company, Incorporated.

Lee, W.-Y. (1973). Effects on Correlation between two Mobile Radio Base-Station Antennas. *IEEE Transactions on Vehicular Technology*, 22(4):130–140.

Li, W. and Chang, Q. (2003). Inclusion intervals of singular values and applications. *Computers & Mathematics with Applications*, 45(10):1637–1646.

Lohner, R. (2001). On the Ubiquity of the Wrapping Effect in the Computation of Error Bounds. In Kulisch, U., Lohner, R., and Facius, A., editors, *Perspectives on Enclosure Methods*, pages 201–218. Springer-Verlag.

Moore, R. E., Kearfott, R. B., and Cloud, M. J. (2009). *Introduction to Interval Analysis*. Society for Industrial and Applied Mathematics, Philadelphia.

Oestges, C. (2006). Validity of the Kronocker Model for MIMO Correlated Channels. In *Vehicular Technology Conference*, volume 6, pages 2818–2822, Melbourne.

Raleigh, G. G. and Cioffi, J. M. (1998). Spatio-Temporal Coding for Wireless Communication. *IEEE Transactions on Communications*, 46(3):357–366.

Rump, S. M. (2011). Verified bounds for singular values, in particular for the spectral norm of a matrix and its inverse. *BIT Numerical Mathematics*, 51(2):367–384.

Sklar, B. (1997). Rayleigh fading channels in mobile digital communication systems. I. Characterization. *IEEE Communications Magazine*, 35(7):90–100.

Su, F., Xu, Z., Dong, Q., and Li, H. (2011). Bounds analysis of singular values for real symmetric interval matrices. *International Journal of Research and Reviews in Applied Sciences*, 6(4):391 – 391.

Telatar, I. E. (1999). Capacity of multi-antenna Gaussian channels. *EUROPEAN TRANSACTIONS ON TELECOMMUNICATIONS*, 10:585–595.

Wang, H., Wang, P., Ping, L., and Lin, X. (2009). How Does Correlation Affect the Capacity of MIMO Systems with Rate Constraints? In *IEEE Global Telecommunications Conference (GLOBECOM)*, pages 1–5.

Weikert, O. and Zölzer, U. (2007). Efficient MIMO Channel Estimation With Optimal Training Sequences. In *Proceedings of 1st Workshop on Commercial MIMO Components and Systems*, Duisburg, Germany.

Quantification of Data and Production Uncertainties for Tire Design Parameters

Maria Böttcher, Wolfgang Graf, Michael Kaliske
Institute for Structural Analysis, TU Dresden, 01062 Dresden, Germany,
{*maria.boettcher; wolfgang.graf; michael.kaliske*}*@tu-dresden.de*

Abstract. The design process for products such as tires constitutes a multi-objective optimization to enhance the product's performance for a specific usage. However, the selected geometry and material parameters as well as the boundary conditions, which are the basis of the performance prediction resulting from a numerical simulation, cannot be guaranteed to coincide with the real parameters of a final product in use. Most parameters are subjected to variation in production and service, which can lead to significant variation in performance.
This contribution has the objective of describing how to model uncertainties of tire design parameters with regard to production variation as well as to non-precise data (Viertl, 1996), to enable uncertainty analysis for tire performance. Thus, the choice of the product design parameters will not be solely based on the predicted value, but also on the reliability of the performance. Uncertainties of design parameters, divided into the two types aleatory and epistemic with respect to their causation, are quantified for several geometry and material parameters based on the situation of available data (Leichsenring et al., 2018) as well as their sensitivity (Pannier and Graf, 2014; Götz et al., 2018). Using statistical analysis methods for provided data from production plants, different approaches for probabilistic, possibilistic and polymorphic uncertainty modeling, combining different types of uncertainties, are selected, including random variables, fuzzy variables, probability-boxes (p-boxes) and other fuzzy stochastic approaches (Möller et al., 2007).

Keywords: polymorphic uncertainty, non-precise data, tire design, production variation

1. Introduction

The objective of any product design is finding the optimal parameter configuration for its intended usage. This can include a variety of different objectives: style, size, durability, weight, costs, sustainability and more. In engineering tasks, the focus often lies on durability and performance for the intended load exposure. Specifically in tire design, the selection of parameters is based on its intended driving situation, e.g. an increased tread pattern depth for better grip and safety when snow and ice are present on the roads, whereas sleeker, stiffer tires are used for racing scenarios. Relevant tire performances include, but are not limited to, dry and wet braking, noise, rolling resistance, wear and cornering stiffness. A small difference in these performances can have large impact on gas consumption, longevity, driving comfort and safety over the

© 2021 *by M. Böttcher, W. Graf, M. Kaliske.*

lifetime of the tire. Therefore, a lot of effort is spent on optimizing these performances. However, viewing the design process as a purely deterministic problem, the performance deviation based on production and service (e.g. tire pressure) fluctuations is neglected. Yet, given the severe impact of minor performance deviations, a consideration of the parameter uncertainty can improve the process of the tire design, by allowing for not only a performance prediction, but for an analysis of the robustness of that performance. The robustness consideration has grown more important in recent years, made possible by surrogate models replacing computationally expensive simulations and, thus, allowing for an uncertainty analysis. The modeling of parametric uncertainty has been applied in numerous examples. (Serafinska et al., 2016) describes a fail-safe design optimization process by computing crack propagation with interval and fuzzy geometry parameters. In (Leichsenring et al., 2018), the fluctuating material parameters of wood are modeled using fuzzy-probability based random variables. A comparison of a probabilistic with an interval modeling approach is given in (Faes et al., 2018) for the variability in the forming process of plate sheets.

This paper describes strategies of modeling parametric uncertainty in tire design, with the focus lying on the performance of the tire's cornering stiffness. A distinction between the two types of uncertainty - aleatory and epistemic - is necessary, as different modeling approaches are commonly used to account for them. Based on both existing data and expert knowledge, the selection of models to quantify the parameters' uncertainties is discussed. Another step in the process of uncertainty modeling is the sensitivity analysis of each design parameter on the cornering stiffness, made possible by a computationally inexpensive performance prediction tool, which allows for the generation of a large data base. The level of detail of the uncertainty quantification can then be selected based on a parameters sensitivity to the performance, i.e. a high level of detail for highly sensitive parameters, whereas non-sensitive parameters can be considered as deterministic.

The structure of the paper is as follows. Section 2 provides a compact overview of uncertainty modeling and gives a brief definition of several existing modeling strategies. In the subsequent Section 3, sensitivity analysis approaches are sketched, both sampling and non-sampling based, with application recommendations and limitations. Section 4 follows with a description of the data situation for different tire design parameters, from the scope of geometry and material, as well as the main results of an exploratory data analysis and the selected uncertainty modeling strategies. A summary and an outlook conclude this contribution in Section 5.

2. Models for Parametric Uncertainty Quantification

For the selection of a model to quantify parameter uncertainties, it is essential to first consider the causes of these uncertainties. There is consensus in the scientific community, that two types can be distinguished: aleatory and epistemic (Möller and Beer, 2008; Der Kiureghian, 2009). The group of aleatory uncertainties considers the inherent variability of parameters, e.g. resulting from production fluctuation or load variation on structures. They can be measured but not reduced. As a modeling strategy, a stochastic approach is commonly used. As for epistemic uncertainties, they originate from a lack of knowledge or information, including incomplete data, e.g. caused by a small sample size, as well

as imprecise data, i.e. samples with a certain measurement error. Given the necessary resources to accumulate additional data as well as to improve the measurement accuracy, the epistemic uncertainty is reducible, yet not completely avoidable. An adequate modeling strategy here is subject of discussion as seen in (Ferson et al., 2004) and (Mullins et al., 2016), with different approaches including evidence theory, possibility theory and interval analysis.
Since the two types of uncertainty are present simultaneously, a joint model is necessary to combine both of them. Also, it is preferable to allow for a distinction of the uncertainty causes, i.e. aleatory or epistemic, as a result of the parameter uncertainty propagation in the uncertainty analysis. Given the results, it is possible to decide if spending additional resources is expedient to reducing the output uncertainty based on the influence of the epistemic uncertainty. Combined modeling strategies include the Bayesian approach (O'Hagan and Oakley, 2004) and imprecise probability, e.g. evidence theory, fuzzy randomness, probability-boxes and fuzzy-probability based randomness. However, since Bayesian probability is a purely stochastic method, it results in a stochastic output variable and, thus, does not allow for an inference of the uncertainty causation, the focus here lies on imprecise probabilities. These modeling strategies can be considered as an extension of a probabilistic model by including possibility theory, in order to add different aspects of epistemic uncertainty.
The subject of model uncertainty is to be viewed separately from that of parametric uncertainty, as discussed in, e.g., (Ferson et al., 2003; Der Kiureghian, 2009). However, it lies outside the scope of this contribution and will not be considered further.

2.1. BASIC UNCERTAINTY CONCEPTS

As a probabilistic model, the random variable is used. It is based on the probability space (Ω, Σ, P) with Ω as the set of all elementary events, Σ as a σ-algebra as a set containing all subsets of elements in Ω, and P as a measure of probability, assigning to each event a probability $P : \Sigma \mapsto [0, 1]$. The random variable is a function mapping an event to a real value $X : \Omega \mapsto \mathbb{R}$.
To describe the function of a random variable, the cumulative distribution function (CDF) is defined as

$$F_X = P(X \leq x) = \int_{-\infty}^{x} f_X(t)dt, \tag{1}$$

with f_X denoting the probability density function (PDF), for which $\int_{-\infty}^{\infty} f_X(t)dt = 1$. The PDF can be considered as the relative frequency $h(x)$ resulting from an infinite number of experiments.
In contrast to a random variable, a fuzzy variable follows possibility theory and is defined as a set A of pairs, assigning a possibility to each value, by assigning to each value x a possibility of occurrence using a membership function μ

$$A = \{(x,\, \mu(x)) \,|\, x \in \mathbb{R},\, 0 \leq \mu(x) \leq 1\};\; \exists\, x \,|\, \mu(x) = 1, \tag{2}$$

$$\mu(x) : \mathbb{R} \to [0,\, 1]. \tag{3}$$

The possibility of a value is to be interpreted linguistically only and does not assign a value of probability.
An alternative way of describing a fuzzy variable is through a family $A = (A^\alpha)_{\alpha\in]0,1]}$ of α-level intervals

$$A^\alpha = \{x \in \mathbb{R} \,|\, \mu(x) \geq \alpha\},\ \alpha \in]0, 1], \tag{4}$$
$$A^\alpha = [\underline{x}_\alpha, \overline{x}_\alpha],\ A^\alpha \subseteq \mathbb{R}. \tag{5}$$

Often, fuzzy variables are described by a simple shape, e.g. as a triangle or a trapezoid. Their definition can be simplified by using the sets $x = \langle x_{0,l}, x_1, x_{0,u}\rangle$ and $x = \langle x_{0,l}, x_{1,l}, x_{1,u}, x_{0,u}\rangle$, respectively. Also, an interval variable can be considered as a special case of a fuzzy variable, where the membership function can merely assume the values $\{0, 1\}$, with $\mu(x) = 1\ \forall\, x \in I$. The interval can be described by a lower and an upper bound $I = [\underline{x}, \overline{x}]$, given a fuzzy variable described by the set $x = \langle x_l, x_u\rangle$. A visualization of these fuzzy variables can be found in Figure 1.

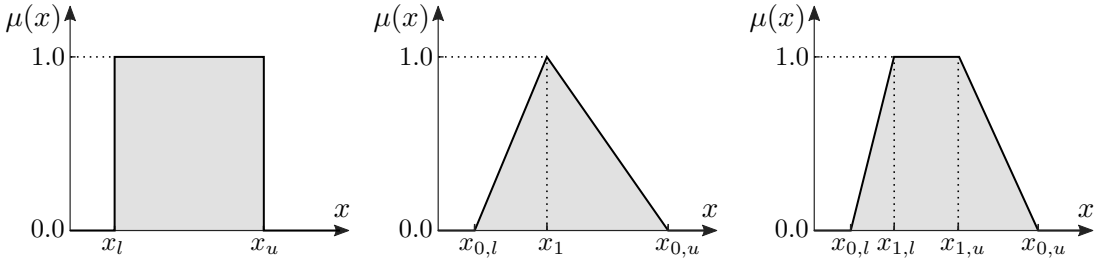


Figure 1. Illustration of an interval (left), a triangular (middle) and a trapezoid (right) fuzzy variable.

2.2. Consideration of parameter uncertainty interactions

As there often exist several parameters for, e.g., one part of the product, the production variation of these can be assumed to show a correlation to some extend. Considering as an example the width of the lower belt in a tire as x_1 and the difference of the lower and upper belt as x_2, then it stands to reason that an increased value in x_1 leads also to an increase in x_2. Accounting for these dependencies is an important part of the uncertainty quantification, since a neglect leads to an overestimation of the predicted performance uncertainty.
If design parameters are modeled as random variables, then interacting parameters can be considered a multivariate random variable, constituted of their individual distribution and joined through a covariance matrix. As for fuzzy variables, (Schietzold et al., 2018) uses a spatial plane dividing a multivariate fuzzy variable in a permissible and a non-permissible region.

2.3. Polymorphic uncertainty

The concept of polymorphic uncertainty, which was introduced in (Graf et al., 2013) and extended in (Graf et al., 2015; Götz et al., 2015; Götz, 2017), describes different approaches combining probabilistic and possibilistic approaches, including probability-boxes, fuzzy randomness and fuzzy-probability based randomness. A guide for selecting

an appropriate uncertainty model for different data situations is given as follows:

- few, imprecise, non-assessed data : interval
- few, imprecise, assessed data : fuzzyness
- very many, deterministic data : randomness

} basic uncertainty models

- many, imprecise, non-assessed data : probability-box
- many, imprecise, assessed data : fuzzy random variable
- some deterministic data : fuzzy-probability based randomness

} polymorphic uncertainty models

In the following, the three mentioned polymorphic uncertainty models will be shortly introduced.
In (Ferson and Ginzburg, 1996) and (Ferson et al., 2003), the uncertainty model probability-box (p-box) is proposed. It deviates from the traditional random variable in the definition of its cumulative density function, which is now not described as a singe function F_X, but by an lower and upper bound $\underline{F}_X$ and $\overline{F}_X$, and, therefore, giving an interval for possible values of F_X

$$\underline{F}_X \leq F_X \leq \overline{F}_X \tag{6}$$

for a random variable X. Thereby, the restriction based on selecting and fitting of the data to a predefined type of function is somewhat reduced. For the lower and upper bounds of the CDF, both predefined and empirical distributions can be used. In the context of a p-box, each realization of the random variable can be considered an interval $X = [\underline{X}, \overline{X}]$, i.e. mapping

$$X : \Omega \mapsto \mathcal{I}(\mathbb{R}), \tag{7}$$

with $\underline{X} : \underline{F}_X(x)$ and $\overline{X} : \overline{F}_X(x)$.
A p-box is described by the tuple $\langle \underline{F}_X, \overline{F}_X, m, v, \mathbf{F}_X \rangle$ in (Götz, 2017), where $\mathbf{F}_X$ includes all possible distribution functions F_X and parameters m and v are the intervals of mean value and variance, for which

$$\int_{-\infty}^{\infty} x \frac{dF_X}{dx} dx \in m, \tag{8}$$

$$\left(\int_{-\infty}^{\infty} x \frac{dF_X}{dx} dx \right) - \left(\int_{-\infty}^{\infty} x \frac{dF_X}{dx} dx \right)^2 \in v. \tag{9}$$

A further extension of the random variable is the model of fuzzy randomness (fr). While for a p-box each realization of the random variable is defined as an interval, here a realization is described by a fuzzy variable

$$X : \Omega \mapsto \mathcal{F}(\mathbb{R}), \tag{10}$$

giving

$$X = ([\underline{X}_\alpha, \overline{X}_\alpha])_{\alpha \in]0,1]}. \tag{11}$$

The definition of the fuzzy cumulative probability function is, similarly to that of a p-box, for each α-level

$$\underline{F}_{X,\alpha} \leq F_{X,\alpha} \leq \overline{F}_{X,\alpha}. \tag{12}$$

Another model consisting of both fuzzy and random variables is fuzzy-probability based randomness (fp-r). Contrary to fuzzy randomness though is the mapping of the random variable to a real value

$$X : \Omega \mapsto \mathbb{R}. \tag{13}$$

To combine epistemic uncertainty quantification with the stochastic model, the parameters θ of the distribution function, from hereon called hyperparameters, are fuzzified

$$F_X = (F_\theta \,|\, \theta \in \theta^\alpha)_{\alpha \in]0,1]}. \tag{14}$$

Therefore, each realization of X has, for each α-level, an interval of probability

$$F_X = ([\underline{F}_{X,\alpha}, \overline{F}_{X,\alpha}])_{\alpha \in]0,1]}. \tag{15}$$

A visualization of these three modeling approaches can be found in Figure 2.

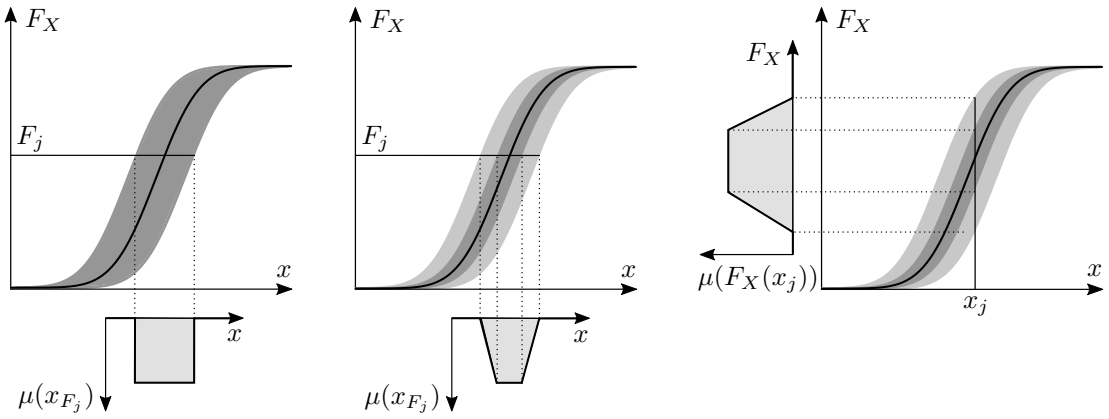


Figure 2. Illustration of the polymorphic uncertainty models p-box (left), fr (middle) and fp-r (right).

3. Sensitivity Analysis

Another important step in quantification of design parameter uncertainties for a specific performance is the investigation of the importance of each parameter, i.e. to what extend the resulting performance is influenced by them. To select adequate analysis methods, the types of the design parameters need to be regarded, given that certain analysis procedures are limited to, e.g., only continuous parameters, and cannot be applied to discrete, categorical or functional ones.

Dependent on the task at hand, the importance of each design parameter can be assigned based on expert knowledge as well as from information of white-box models, if available. Otherwise, or in addition, a sensitivity analysis is performed on data samples $\mathcal{D}$ mapping

input to output $\mathbf{x} \mapsto \mathbf{y}$. Here, a variety of methods can be utilized, including, but not limited to, visualization, measures of correlation, regression coefficients and variance-based indices. A methodology overview with recommendations for selecting a procedure for different tasks can be found in (Pianosi et al., 2014).
When performing a sensitivity analysis, the main question is if the analysis is to be performed as non-sampling, i.e. is based on a set of existing data samples $\mathcal{X}$, or if a simulation model $f(\mathbf{x})$ is available for sampling after a specific design plan (Design of Experiments, DoE) and how computationally expensive a single model run is.

3.1. Non-sampling-based approaches

For a non-sampling sensitivity analysis, the method selection as well as the accuracy of the results are somewhat limited. However, they can allow for a good approximation and validation of expert knowledge. A commonly used approach is data visualization, including scatter- and swarmplots, parallel coordinate plots, Andrew plots and self-organizing maps (SOM). A variety of visualization methods can be found in (Schnell, 1994). However, visual analysis is subjective and difficult to quantify. Also, for high-dimensional input and/or output domains, as is the case in most engineering tasks, the disregard of higher order sensitivities can lead to false interpretation of the results. As a quantification of spatial dependencies, coefficients of correlation can be employed. For linear correlation, the Pearson coefficient

$$\rho_P(x,y) = \frac{cov(x,y)}{\sigma_x \cdot \sigma_y} \in [-1,1] \tag{16}$$

is commonly used. Alternatively, a rank correlation coefficient, e.g. Spearman

$$\rho_S(x,y) = \frac{\sum_i (r(x_i) - \overline{r}_x)(r(y_i) - \overline{r}_y)}{\sqrt{\sum_i (r(x_i) - \overline{r}_x)^2} \sqrt{\sum_i (r(y_i) - \overline{r}_y)^2}} \in [-1,1] \tag{17}$$

with $r(y_i) = j \in \mathbb{N}$ yielding the rank, i.e. the position of the sample y_i within the sample set $\mathcal{Y}$, sorted in ascending order, also shows non-linear dependencies, while reacting less sensitive to outliers. The benefit of a correlation coefficient is the information about positive and negative correlation, which leads to an improved insight of the output dependencies.
With a standardized regression coefficient (SRC), a regression model, e.g. $\hat{y} = b_0 + \sum_{j=1}^{n_x} b_j x_j$ as a linear model, is fitted. To compute the constants b_j, an optimization is performed to minimize the squared error $\sum_{i=1}^{N} (\hat{y}_i - y_i)^2$. The sensitivity of a parameter x_j is proportional to its associated factor b_j. In addition, using the coefficient of determination (COD)

$$R^2 = \frac{SS_{reg}}{SS_{tot}} = \frac{\sum_{i=1}^{N} (\hat{y}_i - \bar{y})}{\sum_{i=1}^{N} (y_i - \bar{y})} \in [0,1], \tag{18}$$

the fitness of the regression model can be evaluated. If $R^2 \approx 1$, the selected model gives a good approximation of the data samples, whereas a low value for R^2 shows that the

model and, therefore, the sensitivity measure is insufficient.
The analysis of variance (ANOVA) is based on the variance decomposition of the output into

$$y = f_0 + \sum_{i=1}^{k} f_i(\mathbf{x}_i) + \sum_{i<1}^{k} f_{i,j}(\mathbf{x}_i, \mathbf{x}_j) + \ldots + f_{1,2,\ldots,k}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k) \tag{19}$$

with

$$\begin{aligned} f_0 &= E(y), \\ f_i(\mathbf{x}_i) &= E(y|\mathbf{x}_i) - f_0, \\ f_{i,j}(\mathbf{x}_i, \mathbf{x}_j) &= E(y|\mathbf{x}_i, \mathbf{x}_j) - f_0 - f_i - f_j, \\ &\ldots. \end{aligned} \tag{20}$$

Based on this decomposition, the Sobol indices S_i were introduced by their namesake as a parameter x_i sensitivity by defining S_i as the variance of the conditional expected value of the output, normalized by the unconditional variance of the output

$$S_{I,i} = \frac{V[E(y|x_i)]}{V(y)}, \quad \sum_i S_{I,i} \leq 1. \tag{21}$$

While Eq. (21) describes the first order Sobol index, i.e. the main effect of a parameter x_i on the output y, thereby neglecting any interaction effects with one or more other parameters x_j, the total Sobol index can be computed as

$$S_{tot,i} = 1 - \frac{V[E(y|\mathbf{x}_{\sim i})]}{V(y)}, \quad \sum_i S_{tot,i} \geq 1, \tag{22}$$

with $E(y|\mathbf{x}_{\sim i})$ as the expected value of the output by varying all parameters except x_i. A sum of first order indices $\sum_i S_{I,i}$ close to 1 indicates that almost only main effects of the parameters exist, while a lower value shows that interaction effects are not to be ignored.
For a approximation of the first order Sobol indices for existing data samples (KVest), (Götz, 2017) proposes a separation of the input domain $\mathcal{X}_i$ in k equidistant intervals and a computation of $\hat{S}_{I,i}$ using

$$\hat{S}_{I,i} = \frac{V(\mathbf{E})}{V(\mathbf{y})}, \quad E_j = E(\mathbf{y}|\mathbf{x_i})_j \; \forall \; j \in k \tag{23}$$

given a sufficient and somewhat equally distributed number of samples.
Another variance-based measure is achieved by performing the fast Fourier sensitivity test (FAST) method (Cukier et al., 1978), which, while also being perform-able as an approximation for existing data samples using the effective algorithm for computing sensitivity indices (EASI), is limited to continuous input and output parameters.

3.2. Sampling-based approaches

Is a simulation model for the generation of (additional) data samples available, a DoE can be constructed. For a visualization or a regression-based analysis, it is most desirable to create samples that cover the whole design space. To that purpose, possible experiments can follow full-factorial or fractional-factorial designs for categorical or discrete variables with a small number of possible realizations (mostly ≤ 3). Otherwise, random sampling (RS) or latin hypercube sampling (LHS) is utilized, whereas LHS shows faster convergence (Aistleitner et al., 2013) than RS and is often preferred. With the goal of approximating first order and total Sobol indices, (Saltelli et al., 2010) suggests specific sampling designs (radial sampling, winding stairs sampling) for a more accurate indices approximation.
In addition to the described data based computation of sensitivity measures, some surrogate models allow for an analytical result. In (Sobol and Kucherenko, 2009) a derivative-based approach is described. (Montano and Palmer, 2003) establishes a method for weight-based sensitivity measures based on feedforward neural network models.

3.3. Design parameter sensitivities for cornering stiffness

As for the tire design performance cornering stiffness, an inexpensive performance prediction tool is available, allowing for generation of data samples for sensitivity analysis. The structural model is a black-box and, therefore, does not give any information about parameter sensitivities based on the underlying functions/equations. The $n_x = 31$ input parameters consist of 24 continuous and 4 discrete, as well as 3 categorical variables. Since the objective of this sensitivity analysis is to contribute to the uncertainty quantification of the examined input parameters, categorical parameters are not of interest, since they are considered as deterministic. Table I gives an overview of the continuous and discrete input parameters, separated into geometry, compound and service. Each parameter index provides information about its corresponding tire part (grouped into A-F), or indicates an association to the categories of general tire dimensions (e.g. nominal width, rim diameter) or of service during usage. The model output consists of two continuous cornering stiffness parameter $\mathbf{y} = \{C_{\alpha,1}, C_{\alpha,2}\}$, for two different loads F_z.
A visual analysis proves to be only somewhat useful, since the input dimension is rather large, and will not be further discussed for the sake of brevity. The results of coefficient of correlation after Spearman (Eq. (17)) are depicted in Figure 3. The results for the

Table I. Tire design parameter for cornering stiffness prediction, sorted by tire part, with (g) for geometry, (c) for compound and (m) for miscellaneous parameter.

	general tire dimensions	tire part A	tire part B	tire part C	tire part D	tire part E	tire part F	service
g	$\{x^g_{0,i}, i \in 5\}$	$\{x^g_{a,i}, i \in 4\}$	$\{x^g_{b,i}, i \in 2\}$	$\{x^g_{c,i}, i \in 3\}$	$\{x^g_{d,i}, i \in 2\}$	$\{x^g_{e,i}, i \in 2\}$		
c		$\{x^c_{a,i}, i \in 2\}$	$\{x^c_{b,i}, i \in 2\}$	$\{x^c_{c,i}, i \in 1\}$	$\{x^c_{d,i}, i \in 1\}$	$\{x^c_{e,i}, i \in 2\}$	$\{x^c_{f,i}, i \in 1\}$	
m								$\{x_s\}$

PEARSON correlation coefficient (Eq. (16)) have been found to coincide, with a maximum deviation of $\Delta(\rho_S, \rho_P) = 0.02$. The same result was acquired as the result of a regression analysis when using a linear model. However, the coefficients ot determination $R^2 = 0.71$ for $C_{\alpha,1}$ and $R^2 = 0.52$ for $C_{\alpha,2}$ indicate an insufficient representation of the target function by a linear model. An improvement was achieved by separating the general tire dimension parameter nominal tire width, resulting in $R^2 > 0.8$ for each category and still leading to similar results for the standardized regression coefficients. In addition, the radial sampling scheme described in (Saltelli et al., 2010) has been performed in order to approximate the total SOBOL indices $S_{tot,i}$. Here, each nominal width was investigated separately, and the maximum of the resulting indices was selected for each parameter. For that reason, the sum of all total SOBOL indices is much larger than 1, without necessarily indicating the presence of higher order interaction effects. The resulting sensitivities are depicted in Figure 3.

With the focus on the total SOBOL indices, the design parameters are separated based on their sensitivity as shown in Table II. The total SOBOL index was selected instead of the first order one, because the total effect of a parameter is of more relevance for the description of parameter importance in the context of uncertainty quantification. Also, a faster convergence during the numerical approximation of the total indices compared to the first order was observed.

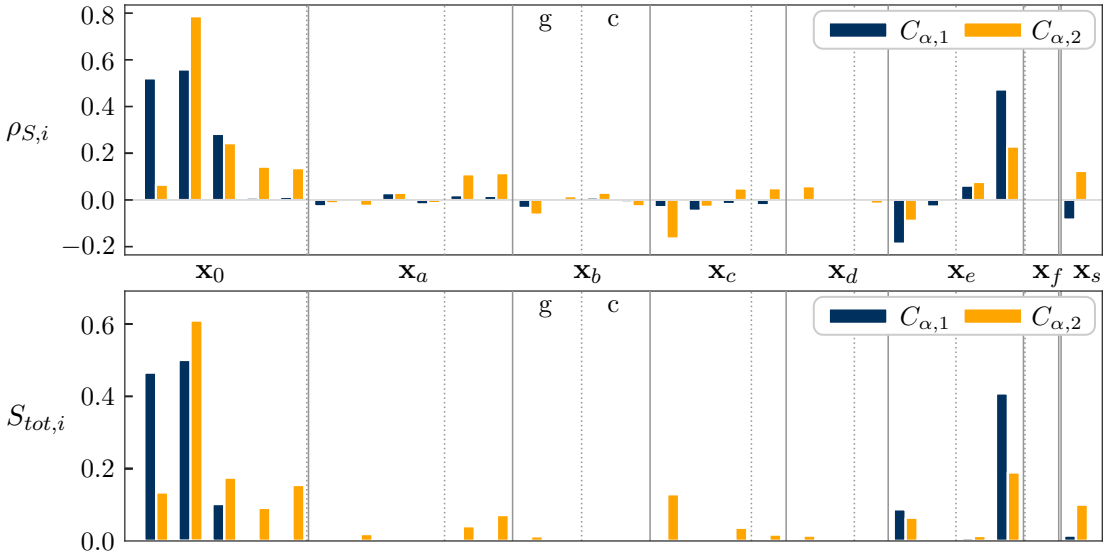


Figure 3. Coefficient of Correlation (SPEARMAN) $\rho_{S,i}$ (top) and total SOBOL indices $S_{tot,i}$ (bottom) for tire design parameter, with distinction of tire parts, grouped into (g) geometry and (c) compound (if applicable).

Table II. Tire design parameter for cornering stiffness prediction, separated by sensitivity.

parameters with medium to high sensitivity	parameters with low to no sensitivity
$\{x^g_{0,i}, i \in 5\}$	
$\{x^c_{a,i}, i \in 2\}$	$\{x^g_{a,i}, i \in 4\}$
	$\{x^g_{b,i}, i \in 2\}$, $\{x^c_{b,i}, i \in 2\}$
$\{x^g_{c,i}, i \in 2\}$	$\{x^g_{c,3}\}$ $\{x^c_{c,i}, i \in 1\}$
	$\{x^g_{d,i}, i \in 2\}$, $\{x^c_{d,i}, i \in 1\}$
$\{x^g_{e,1}, x^c_{e,2}\}$	$\{x^g_{e,2}, x^c_{e,1}\}$
	$\{x^c_{f,i}, i \in 1\}$
$\{x_s\}$	

4. Uncertainty quantification for tire design parameters

4.1. Data situation

The performance and, therefore, the design of a tire is dependent on a variety of parameters, including different types of rubber compounds and layers of textile and steel, the geometry and the molding process. For $n^g = 7$ geometry parameters $x^g_i, i \in n^g$ and $n^c = 47$ compound parameters $x^c_i, i \in n^c$, measurements taken in the process of tire production were available.
Each data sample for a geometry parameter includes information about the target value, the actual value and the general tire geometry as well as randomized information about the product specification, the production plant and date

$$x^g_i = \langle \Delta x, x_{Target}, Prod, Plant, Date \rangle. \tag{24}$$

As for the compound data, each sample consists of three measurements of the materials storage shear moduli G' for the strains $\varepsilon = \{1, 10, 100\}\%$, with values statistically normalized to $\overline{x}_i = \frac{x_i - \mu_X}{\sigma_X}$, as well as the date the test was performed. The description of the compound is randomized, yet allows for a distinction between body (B) and tread (T) compounds

$$x^c_i = \langle G'_1, G'_{10}, G'_{100}, Date \rangle. \tag{25}$$

No performance measure is associated with these data samples, so no sensitivity analysis based on these experimental values is possible.
In context of the data preparation, extreme outliers are discarded, i.e. data samples with $\Delta x \notin [Q_1 - 3(Q_3 - Q_1), Q_3 + 3(Q_3 - Q_1)]$ with Q_1 and Q_3 being the 25% and 75% quartiles of all samples Δx_j.

4.2. Uncertainties present in data samples

The data samples for the geometric and material parameters show, that a variability is present for each one of the parameters. This type of uncertainty, denoted as the aleatory type, is commonly modeled using a stochastic model. Figure 4 depicts the deviation from

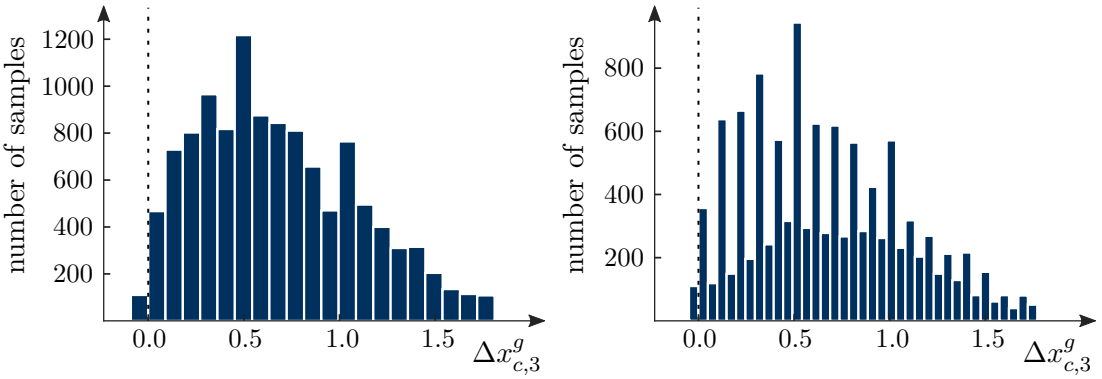


Figure 4. Histogram deviation of the tire geometry parameter $\Delta x^g_{c,3}$ with different number of bins.

the target value of the geometry parameter $x^g_{c,3}$ in form of a histogram with a positively skewed distribution. In addition, by adjusting the number of bins, an alternation of bin heights, i.e. in the number of contained samples, becomes apparent, indicating the limitations of measurement accuracy for the parameter. This lack of accuracy is classified as epistemic uncertainty and can not be disregarded when establishing a detailed uncertainty model. Given that a number of 11 706 measurements (after discarding of outliers) for the parameter $x^g_{c,3}$ can be considered a large sample size, epistemic uncertainty caused by incompleteness will be disregarded.
Between the geometry parameters of the same tire part, a correlation of the variation can be noted. Also, given the information about the target value in addition to the variation, some parameters display a dependency of these values. However, since these parameters show low to no sensitivity and will therefore be set as deterministic, these dependencies will be neglected. Otherwise, separate models for each target level would be advisable.

4.3. Selected uncertainty modeling strategies

As described before, the uncertainty quantification for the tire design parameters is based on their sensitivity for the performance cornering stiffness as well as the available data and expert knowledge. Table III gives an overview of the selected uncertainty modeling strategies.
For sensitive parameters with no production variation data, a fuzzy model is employed, to avoid false assumptions by assigning a specific type of probability distribution. Low sensitivity parameters are considered deterministic, for the sake of reducing the dimensionality of the later following fuzzy analysis in context of an uncertainty analysis and thereby minimizing the numerical effort. For parameters with considerable sensitivity as well as a data base, the measurement accuracy bounds constitute each data sample as an interval and are therefore used to construct p-boxes. Figure 5 shows the p-box modeling on the example of parameter $x^g_{c,3}$. A stochastic modeling of the body compound uncertainty can be seen in Figure 6 for two different strains ε. The fitting of different probability distributions was performed using the maximum-likelihood-estimate, and evaluated with the χ^2-test. As the p-value was higher than the conservatively selected α value of 0.10, the distribution can be considered an accurate fit in both cases. As

the sensitivity of body compounds is comparatively low, a more detailed uncertainty quantification has been refrained from.

Table III. Uncertainty quantification models for different parameters relevant for the performance cornering stiffness in a tire design process

	parameter	sensitivity	data base	modeling strategy
geometry	$x^g_{0,i}, i \in 5$	high	–	individual trapezoid fuzzy variables, based on expert knowledge
	$x^g_{c,3}$	medium	+	p-box with measurement accuracy as bounds
	$x^g_{c,1}, x^g_{c,3}$ $x^g_{e,1}$		–	individual fuzzy interval variables, based on expert knowledge
	$x^g_{a,i}, i \in 4$	low/none	–	deterministic
	$x^g_{b,i}, i \in 2$		–	
	$x^g_{c,2}$		–	
	$x^g_{d,i}, i \in 2$		+/–	
	$x^g_{e,2}$		+	
compound	tread $x^c_{T,i}$	high	+	p-box with measurement accuracy as bounds
	body x^c_B	low/none	+	multivariate (for different strains ε) random variable $X^c_B \sim t$ with covariance matrix
service	x_s	medium	–	individual trapezoid fuzzy variables, based on expert knowledge

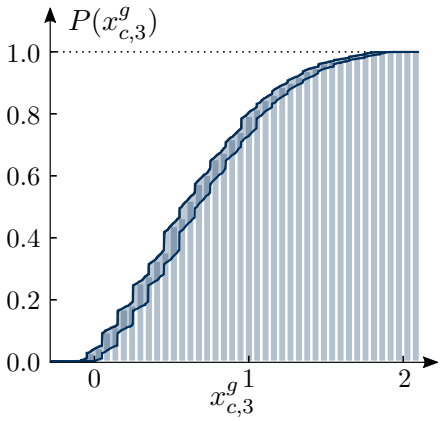


Figure 5.: p-box of parameter $x^g_{c,3}$.

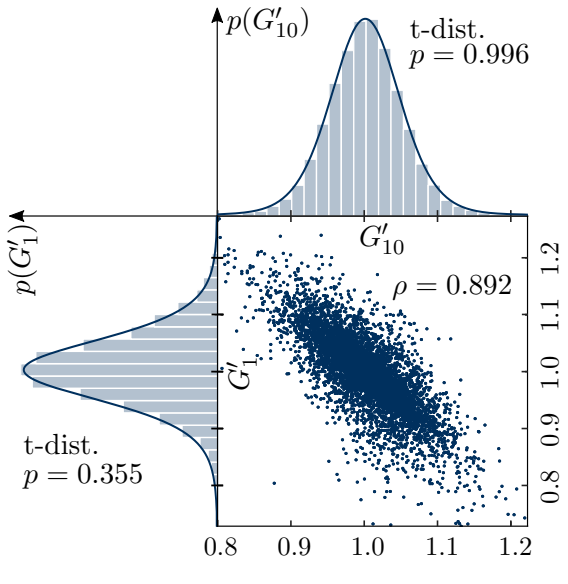


Figure 6.: Histogram & Scatterplot for measurements of body compound G' at $\varepsilon = \{1, 10\}\%$.

5. Conclusion and outlook

The quantification of parametric uncertainty is an important step in many design tasks, as it allows for a robustness analysis of the model output in addition to a deterministic prediction. As has been shown by the data collected from several production plants over a considerable amount of time, the fluctuation of different design parameters is considerable, as there are tolerance limits during any production process. Therefore, an intensive study of modeling these uncertainties is necessary.
The focus of this paper lies on the uncertainty quantification of tire design parameters for the performance cornering stiffness. An overview of basic as well as polymorphic uncertainty modeling approaches has been described. Since a specific product performance was investigated, a sensitivity analysis was performed to distinguish those design parameters which are most important, i.e. who's deviation has the biggest impact on the performances variance, and fixing not important ones as deterministic. For some of the tire design parameters, a data base of the production variation was available and analyzed as a basis for an uncertainty quantification. The resulting selected modeling concepts have been illustrated by examples based on the data situation, the parameters' sensitivities and expert knowledge.
It has been shown that an extensive data base is necessary for a detailed description of the aleatory parameter uncertainty, and, therefore, if not available, an assumption like, e.g., a normal distribution model, is not sufficient, as it may falsify results and under- or overestimates the robustness. In such a case, a quantification of epistemic uncertainty is necessary as well as the aleatory one, therefore a combined modeling is required.
Future work will include an uncertainty analysis for the tire performance run on an inexpensive simulation model, as well as a comparison of different modeling approaches and their influence on the performance variation. A model for robustness evaluation will be developed as a guideline for selecting a tire design which is considered as a multi-criteria optimization for both, the performance value and its robustness.

Acknowledgements

The provision of tire production data and expert knowledge on the subject of tire design as well as the financial support of Continental Reifen Deutschland GmbH is highly appreciated. Special thanks goes to the departments of Mechanics and Simulation Development under R. Nojek and Virtual Development and Computational Mechanics under J. Dehnert as well as the project team members C. Gerstenberger, C. Marschler and C. von Zobeltitz for the continuous and helpful cooperation in all tire-related questions.

References

Aistleitner, C., Hofer, M. and Tichy, R A central limit theorem for Latin hypercube sampling with dependence and application to exotic basket option pricing. *Journal of Theoretical and Applied Finance*, 15:1–20, 2013.

Cukier, R. I., Levine, H. B. and Shuler, K. E. Nonlinear sensitivity analysis of multiparameter model systems. *Journal of Computational Physics*, 26:1–42, 1978.

Der Kiureghian, A. Aleatory or epistemic? Does it matter?. *Structural Safety*, 31:105–112, 2009.

Faes, M., Van Doninck, B., Imholz, M. and Moens, D. Product reliability optimization under plate sheet forming process variability. *8th International Workshop on Reliable Engineering Computing, "Computing with Confidence"*, Liverpool, 2018.

Ferson, S. and Ginzburg, L. R. Different methods are needed to propagate ignorance and variability. *Reliability Engineering & System Safety*, 54:133–144, 1996.

Ferson, S., Kreinovich, V., Ginzburg, L. R. and Myers, D. S. *Constructing Probability Boxes and Dempster-Shafer Structures.* Sandia National Laboratories, Livermore, 2003.

Ferson, S. and Hajagos, J. G. Arithmetic with uncertain numbers: rigorous and (often) best possible answers. *Reliability Engineering & System Safety*, 85:135–152, 2004.

Ferson, S., Joslyn, C. A. Helton, J. C., Oberkampf, W. L. and Sentz, K. A Summary from the epistemic uncertainty workshop: consensus amid diversity. *Reliability Engineering & System Safety*, 85: 355–369, 2004.

Graf, W., Götz, M. and Kaliske, M. Analysis of dynamical processes under consideration of polymorphic uncertainty. In: G. Deodatis, B. R. Ellingwood, and D. M. Frangopol (Eds.), *Safety, Reliability, Risk and Life-Cycle Performance of Structures & Infrastructures: Proceedings of the 11th International Conference on Structural Safety and Reliability (ICOSSAR 2013)*, CRC Press, Taylor and Francis Group, pp. 453–460, 2013.

Graf, W., Götz, M. and Kaliske, M. Analysis of dynamical processes under consideration of polymorphic uncertainty. *Structural Safety*, 52(B):194–201, 2015.

Götz, M., Graf, W. and Kaliske, M. Structural design with polymorphic uncertainty models. *International Journal of Reliability and Safety*, 9:112–131, 2015.

Götz, M., Leichsenring, F., Kropp, M., Müller, P., Falk, T., Graf, W., Kaliske, M. and Drossel, W. Data mining and machine learning methods applied to a numerical clinching model. *Computer Modeling in Engineering & Sciences*, 117(3):387–423, 2018.

Götz, M. *Numerische Entwurfsmethoden unter Berücksichtigung polymorpher Unschärfe*. Dissertation, Institute for Structural Analysis, Technische Universität Dresden 2017. Vol. 32.

Jansen, M. J. W. Analysis of variance designs for model output. *Computer Physics Communications*, 117:35–43, 1999.

Leichsenring, F., Jenkel, C., Graf, W. and Kaliske, M. Numerical simulation of wooden structures with polymorphic uncertainty in material properties. *International Journal of Reliability and Safety*, 12:24–45, 2018.

Möller, B. and Beer, M. Engineering computation under uncertainty – Capabilities of non-traditional models. *Computers & Structures*, 86:1024–1041, 2008.

Möller, B., Graf, W., Sickert, J. U. and Reuter, U. Numerical simulation based on fuzzy stochastic analysis. *Mathematical and Computer Modelling of Dynamical Systems*, 13:349–364, 2007.

Montaño, J. J. and Palmer, A. Numeric sensitivity analysis applied to feedforward neural networks. *Neural Computation & Application*, 12:119–125, 2003.

Mullins, J., Ling, Y., Mahadevan, S., Sun, L. and Strachan, A. Separation of aleatory and epistemic uncertainty in probabilistic model validation. *Reliability Engineering & System Safety*, 147:49–59, 2016.

Pannier, S. and Graf, W. Sectional global sensitivity measures.. *Reliability Engineering & System Safety*, 134:110–117, 2014.

Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B. and Wagener, T. Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software*, 79:214–232, 2014.

Plischke, E. An effective algorithm for computing global sensitivity indices (EASI). *Reliability Engineering & System Safety*, 95:354–360, 2010.

O'Hagan, A. and Oakley, J. E. Probability is perfect, but we can't elicit it perfectly. *Reliability Engineering & System Safety*, 85:239–248, 2004.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*, 181:259–270, 2010.

Saltelli, A., Tarantola, S. and Chan, K. P.-S. A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output. *Technometrics*, 41:39–56, 1999.

Schietzold, F. N., Graf, W. and Kaliska, M. Polymorphic uncertainty modeling for optimization of timber structures. *8th International Workshop on Reliable Engineering Computing, "Computing with Confidence"*, Liverpool, 2018.

Schnell, R. *Grafisch gestützte Datenanalyse*. Verlag Oldenbourg, München, 1994.

Serafinska, A., Özenc, K., Kaliske, M. and Graf, W. A Fail-Safe Design Approach Based on the Fracture Mechanical Analysis and Epistemic Uncertainty Quantification. *Proceedings of the 7th International Workshop on Reliable Engineering Computing (REC) 2016, Bochum*, 143–158, 2016.

Sobol, I. M. Global sensitivity indices for the investigation of nonlinear mathematical models. *Matematicheskoe Modelirovanie*, 19:23–24, 2007.

Sobol, I. M. and Kucherenko, S. Derivative based global sensitivity measures and their link with global sensitivity indices. *Mathematics and Computers in Simulation*, 79:3009–3017, 2009.

Viertl, R.. *Statistical Methods for Non-Precise Data*. CRC Press, Boca Raton, 1996.

Serviceability assessment of footbridges via interval analysis

Roberta Santoro[1)], Alba Sofi[2)], and Federica Tubino[3)]
[1)] *Department of Engineering, University of Messina, Messina, 98166, Italy, roberta.santoro@unime.it*
[2)] *Department of Architecture and Territory, University "Mediterranea" of Reggio Calabria, Reggio Calabria, 89124, Italy, alba.sofi@unirc.it*
[3)] *Department of Civil, Chemical and Environmental Engineering, University of Genoa, Genova, 16145, Italy, federica.tubino@unige.it.*

Abstract: This paper studies serviceability assessment of footbridges through a non-deterministic approach. Both the parameters defining pedestrian-induced loading and the structural dynamic properties are uncertain quantities, and they can be characterized through possible ranges of variation. Thus, a suitable tool for taking uncertainties into account is interval analysis. In this paper, the improved interval analysis is applied together with classical optimization that allows obtaining the bounds of the spectral moments of the response, of the cumulative distribution function of the maximum footbridge acceleration, and of the expected value of the maximum acceleration. Based on this approach, an interval of probability of reaching a suitable comfort level can be estimated.

Keywords: Footbridges; Human-induced vibration; Interval Analysis; Serviceability.

1. Introduction

The increasing strength of new materials and aesthetic requirements lead to the design of lightweight, slender and, consequently, vibration-sensitive structures. As regards footbridges, this trend is further stimulated by the relatively small service loads. For these structures, the vibration serviceability under human-induced loading has become a key design criterion.

In order to estimate human-induced response of footbridges, pedestrian traffic should be modelled probabilistically, considering several sources of randomness among which pedestrian arrivals, step frequencies and velocities, force amplitudes and pedestrian weights. The method generally used to statistically characterize the maximum dynamic response of the footbridge accounting for the probabilistic nature of pedestrian excitation is Monte Carlo simulation. However, due to the computational burden of Monte Carlo simulations, which makes them not practical in view of engineering applications, current guidelines (e.g., SETRA, 2006; HIVOSS, 2009) provide simplified procedures to assess footbridges' serviceability. An interesting alternative to simplified procedures is the introduction of spectral models of pedestrian excitation (e.g., Brownjohn et al., 2004; Piccardo and Tubino, 2012), that allow dealing with the problem of serviceability assessment in the frequency domain through the methods of random dynamics. The Equivalent Spectral Model (ESM) introduced by Piccardo and Tubino (2012) has been derived analytically, based on suitable probabilistic models of all the parameters involved. It allows deriving closed-form expressions for the spectral moments of the dynamic response, as functions of the loading and structural parameters.

© 2021 *by R. Santoro, A. Sofi, F. Tubino.*

Serviceability assessment of footbridges is commonly carried out, based on current guidelines, by comparing a deterministic value of the expected maximum acceleration and suitable thresholds defining different comfort classes. Actually, both pedestrian-induced loading parameters (i.e. dynamic loading factor, pedestrian weight and step frequency) and the structural dynamic properties (i.e. modal damping ratios, natural frequencies and modal masses) are affected by remarkable uncertainties. Thus, serviceability assessment should be tackled through a non-deterministic approach. A first attempt accounting only for the errors connected with the estimate of structural parameters and based on simplified loading models has been introduced through a fuzzy analysis (Lievens et al., 2016). In principle, serviceability assessment could be carried out based on Monte Carlo simulations of pedestrian excitation, taking into account the uncertainties of the involved parameters, resulting in very time consuming analyses. However, the ESM constitutes a very efficient tool to perform uncertainty propagation analyses, thanks to the availability of closed-form solutions for the spectral moments of the response as explicit functions of the uncertain parameters.

Tubino et al. (2020) carried out a literature analysis in order to characterize probabilistically the uncertainties in the loading and structural parameters. The mean values and coefficient of variation of the involved uncertain parameters have been estimated, and the propagation of uncertainties on the serviceability assessment of footbridges has been studied adopting the Taylor series expansion technique. Furthermore, Monte Carlo simulations have been performed assuming probability distributions of the uncertain parameters, and the probability distribution of the mean value of the maximum acceleration has been estimated numerically (Tubino et al., 2020).

Actually, probability distributions of the loading and structural parameters are not available, but the literature analysis in Tubino et al. (2020) provides possible ranges of variation of these parameters. Thus, a suitable tool for taking into account uncertainties in serviceability assessment of footbridges is interval analysis (Moore et al., 2009; Muscolino and Sofi, 2012). Indeed, the key idea of the interval model of uncertainty is to describe the uncertain parameters as interval variables with given lower bound and upper bound without requiring a full probabilistic characterization. Since the mid-1990s, several studies have been devoted to the development of efficient procedures for the propagation of interval uncertainties in engineering problems (see e.g., Faes and Moens, 2020). The main challenge is the evaluation of tight bounds of system's response, which is described by an interval rather than assuming a single value. Both static (see e.g., Sofi et al. 2019) and dynamic (see e.g., Liu et al., 2013) behaviors of structures with interval parameters have been thoroughly investigated in literature. Over the last decade, the analysis of randomly excited systems in the presence of interval uncertainties has also been addressed (see e.g., Muscolino and Sofi, 2012; Muscolino et al., 2016).

In this paper, starting from the analytical expression for the spectral moments of the structural response (Piccardo and Tubino, 2012), the *improved interval analysis* (Muscolino and Sofi, 2012) is applied together with classical optimization that allows obtaining the bounds of the standard deviation, of the cumulative distribution function (Muscolino et al., 2016) and of the expected value of the maximum footbridge acceleration. Based on this approach, serviceability assessment is carried out within a non-probabilistic framework, and an interval of probability of reaching a suitable comfort level can be estimated.

2. Analytical formulation

2.1. Equation of Motion

Let us consider a footbridge, modelled as a linear mono-dimensional classically damped dynamical system. Its equation of motion can be written as follows:

$$m_s(x)\frac{\partial^2 q(x,t)}{\partial^2 t}+\mathcal{C}\left[\frac{\partial q(x,t)}{\partial t}\right]+\mathcal{L}\left[q(x,t)\right]=f(x,t) \tag{1}$$

where q is the vertical displacement of the footbridge; x is the abscissa along the structure; t is the time; $m_s(x)$ is the structural mass; $\mathcal{C}$ is the damping operator; $\mathcal{L}$ is the stiffness operator; $f(x,t)$ is the external force. Focusing attention only on the first walking harmonic for each pedestrian, the force induced by N_p pedestrians can be expressed as:

$$f(x,t)=\sum_{i=1}^{N_p}\alpha_i G_i \sin\left(\Omega_i(t-\tau_i)+\Psi_i\right)\delta\left[x-c_i(t-\tau_i)\right]\left[H(t-\tau_i)-H\left(t-\tau_i-\frac{L}{c_i}\right)\right] \tag{2}$$

where $\delta(\bullet)$ and $H(\bullet)$ are the Dirac function and the Heaviside function, respectively; L is the span length of the structure; furthermore, F_i ($=\alpha_i G_i$), ω_i, Ψ_i, c_i and τ_i are, respectively, the force amplitude, the step circular frequency, the phase-angle, the walking velocity and the arrival time of the i-th pedestrian, while α_i and G_i are the Dynamic Loading Factor (DLF) and the weight of the i-th pedestrian. All these quantities can be considered as random variables (Racic et al. 2009; Tubino and Piccardo, 2016).

Under the hypothesis of classical damping, Eq. (1) is usually solved applying the principal transformation. Considering the contribution of one vibration mode (namely the j-th), the structural displacement can be expressed as:

$$q(x,t)=\varphi_j(x)\,p_j(t) \tag{3}$$

where $\varphi_j(x)$ is the j-th mode of vibration, $p_j(t)$ is the corresponding principal coordinate.

The equation of motion of the j-th principal coordinate is expressed as:

$$\ddot{p}_j(t)+2\xi_j\omega_j\dot{p}_j(t)+\omega_j^2 p_j(t)=\frac{1}{M_j}F_j(t) \tag{4}$$

where ξ_j, ω_j and M_j are, respectively, the j-th modal damping ratio, natural circular frequency and modal mass; $F_j(t)$ is the j-th modal force:

$$F_j(t)=\int_0^L f(x,t)\varphi_j(x)\,\mathrm{d}x \tag{5}$$

2.2. Equivalent Spectral Model of the Loading

According to the ESM, multi-pedestrian loading can be modelled as an equivalent Gaussian stationary random process. Considering the first modal shape of a simply-supported beam $\varphi_j(x) = \sin(\pi x/L)$, the power spectral density function (PSDF) of the modal force F_j (Eq. (5)) is given by:

$$S_{F_j}(\omega) = \frac{N_p}{4}(\alpha G)^2 \, p_\Omega(\omega) \tag{6}$$

where N_p is the number of pedestrians; α is the mean value of the DLF; G is the mean value of the weight of pedestrians; $p_\Omega(\omega)$ is the probability density function (PDF) of the circular step frequency of pedestrians. The circular step frequency of pedestrians can be assumed as normally distributed, and its PDF is given by:

$$p_\Omega(\omega) = \frac{1}{\sigma_{\omega p}\sqrt{2\pi}} \exp\left[-\frac{(\omega - \mu_{\omega p})^2}{2\sigma_{\omega p}^2} \right] \tag{7}$$

being $\mu_{\omega p}$ and $\sigma_{\omega p}$ the mean-value and standard deviation of the circular step frequency of pedestrians, respectively.

The PSDF of the acceleration of the j-th principal coordinate is then given by:

$$S_{\ddot{p}_j}(\omega) = \omega^4 \left| H_j(\omega) \right|^2 S_{F_j}(\omega) \tag{8}$$

where the PSDF of the modal force $S_{Fj}(\omega)$ is provided by Eq. (6), and $H_j(\omega)$ is the frequency response function of the j-th principal coordinate:

$$H_j(\omega) = \frac{1}{M_j} \frac{1}{\omega_j^2 - \omega^2 + 2\mathrm{i}\xi_j \omega_j \omega} \tag{9}$$

where $\mathrm{i} = \sqrt{-1}$ is the imaginary unit. Taking into account that the PSDF of the modal force $S_{Fj}(\omega)$ is nearly constant around the j-th natural circular frequency ω_j, the spectral moments of zero- and second-order of the acceleration of the j-th principal coordinate can be approximated as follows:

$$\lambda_{0,\ddot{p}_j} = \sigma_{\ddot{p}_j}^2 = \int_0^{+\infty} S_{\ddot{p}_j}(\omega)\mathrm{d}\omega = \int_0^{+\infty} \omega^4 \left| H_j(\omega) \right|^2 S_{F_j}(\omega)\mathrm{d}\omega \approx \frac{\pi \omega_j S_{F_j}(\omega_j)}{4\xi_j M_j^2} \tag{10}$$

$$\lambda_{2,\ddot{p}_j} = \int_0^{+\infty} \omega^2 S_{\ddot{p}_j}(\omega)\mathrm{d}\omega = \int_0^{+\infty} \omega^6 \left| H_j(\omega) \right|^2 S_{F_j}(\omega)\mathrm{d}\omega \approx \frac{\pi \omega_j^3 S_{F_j}(\omega_j)}{4\xi_j M_j^2} \tag{11}$$

2.3. EXTREME VALUE OF THE ACCELERATION

Modelling the loading through the ESM, the footbridge dynamic response is a stationary Gaussian random process. The maximum acceleration of the j-th principal coordinate in the time interval T is then defined as the following random process:

$$\ddot{p}_{j,\max}(T) = \max_{0 \le t \le T} \left| \ddot{p}_j(T) \right| \tag{12}$$

where the time interval T can be conventionally defined as a multiple of the average crossing time of pedestrians, $T=NL/c_m$, being $c_m=0.9\mu_{\omega p}/2\pi$ (m/s) the mean walking velocity and assuming N=10 (Piccardo and Tubino, 2012).

Under the Poisson's assumption of independent up-crossings of a prescribed threshold, the probability that the extreme value acceleration process is less than or equal to the critical level b>0 within the time interval [0 T] is defined by the following cumulative distribution function (CDF):

$$L_{\ddot{p}_{j,\max}}(b,T) = \Pr\left[\ddot{p}_{j,\max}(T) \le b \right] \approx \exp\left[-2T \, \nu_{\ddot{p}_j} \exp\left(-\frac{b^2}{2\lambda_{0,\ddot{p}_j}} \right) \right] \tag{13}$$

where $\nu_{\ddot{p}_j}$ is the expected frequency of the acceleration of the *j*-th principal coordinate, given by:

$$\nu_{\ddot{p}_j} = \frac{1}{2\pi}\sqrt{\frac{\lambda_{2,\ddot{p}_j}}{\lambda_{0,\ddot{p}_j}}} \approx \frac{\omega_j}{2\pi}. \tag{14}$$

Considering Eq. (3), and the first modal shape of a simply-supported beam $\varphi_j(x) = \sin(\pi x/L)$, the maximum mid-span acceleration is coincident with the maximum acceleration of the *j*-th principal coordinate. According to Davenport's formulation (Davenport, 1964), it is given by:

$$\ddot{q}_{\max} = \mathrm{E}\left[\ddot{p}_{j_{\max}}\right] = g_{\ddot{p}_j}\sigma_{\ddot{p}_j} \tag{15}$$

being the peak factor $g_{\ddot{p}_j}$ defined as:

$$g_{\ddot{p}_j} = \sqrt{2\ln(2\nu_{\ddot{p}_j}T)} + \frac{0.5772}{\sqrt{2\ln(2\nu_{\ddot{p}_j}T)}} \tag{16}$$

2.4. COMFORT REQUIREMENTS

Serviceability assessment of footbridges is commonly carried out, based on current guidelines, by comparing a deterministic value of the expected maximum acceleration and suitable thresholds defining different comfort classes. Table I reports the threshold values as defined by SETRA (2006).

Table I. Comfort classes (SETRA)

Comfort Class		*Min*	*Max*
Maximum	[m/s^2]	0	0.5
Medium	[m/s^2]	0.5	1
Minimum	[m/s^2]	1	2.5
Unacceptable	[m/s^2]	2.5	

Assuming deterministic values for all the parameters involved, the serviceability assessment can be carried out estimating the mean-value of the maximum acceleration through Eq. (15) and checking the comfort class through Table I. As an alternative, the probability of falling within each comfort class can be estimated through Eq. (13).

3. Interval serviceability assessment

Interval serviceability assessment is based on the assumption of uncertain loading and structural parameters, represented as interval variables according to the interval model of uncertainty:

$$\mathbf{x}^I = \left\{\mu_{\omega p}^I \quad \sigma_{\omega p}^I \quad \alpha^I \quad G^I \quad \omega_j^I \quad \xi_j^I \quad M_j^I\right\}^{\mathrm{T}}. \tag{17}$$

Each component of the interval vector $\mathbf{x}^I$, i.e. the *k*-th interval variable x_k^I, is defined as:

$$x_k^I = x_{0,k}\left(1 + \Delta x_k \hat{e}_k^I\right) \tag{18}$$

where $\hat{e}_k^I = [-1, 1]$ is the *extra unitary interval* (Muscolino and Sofi, 2012), $x_{0,k}$ and Δx_k are the midpoint and normalized deviation amplitude, defined as follows:

$$x_{0,k} = \frac{\overline{x}_k + \underline{x}_k}{2}; \tag{19}$$

$$\Delta x_k = \frac{\overline{x}_k - \underline{x}_k}{2x_{0,k}}. \tag{20}$$

In Eqs. (19) and (20), $\underline{x}_k$ and $\overline{x}_k$ represent, respectively, the lower bound (LB) and upper bound (UB) of the interval variable x_k^I.

Based on the interval representation of the involved variables, all the quantities in Eqs. (6)-(16) are interval functions/ variables. The impact of each uncertain parameter on serviceability assessment of the footbridge can be predicted performing sensitivity analysis. Sensitivities of the standard deviation of the acceleration, of the CDF of the extreme value acceleration and of the expected value of the maximum acceleration can be evaluated analytically by direct differentiation of Eqs. (10), (13) and (15) with respect to the parameters $x_k \in x_k^I$.

The LB and UB of the standard deviation of the modal acceleration are defined as:

$$\underline{\sigma}_{\ddot{p}_j} = \min_{\mathbf{x}\in\mathbf{x}^I}\left\{\sqrt{\frac{\pi\omega_j S_{F_j}(\omega_j)}{4\xi_j M_j^2}}\right\}; \qquad \overline{\sigma}_{\ddot{p}_j} = \max_{\mathbf{x}\in\mathbf{x}^I}\left\{\sqrt{\frac{\pi\omega_j S_{F_j}(\omega_j)}{4\xi_j M_j^2}}\right\}. \tag{21}$$

Analogously, the LB and UB of the interval CDF of the extreme value acceleration process are defined as:

$$\begin{aligned}
\underline{L}_{\ddot{p}_j,\max}(\mathbf{x};b,T) &= \min_{\mathbf{x}\in\mathbf{x}^I}\left\{\exp\left[-2T\,\nu_{\ddot{p}_j}\exp\left(-\frac{b^2}{2\lambda_{0,\ddot{p}_j}}\right)\right]\right\} \\
\overline{L}_{\ddot{p}_j,\max}(\mathbf{x};b,T) &= \max_{\mathbf{x}\in\mathbf{x}^I}\left\{\exp\left[-2T\,\nu_{\ddot{p}_j}\exp\left(-\frac{b^2}{2\lambda_{0,\ddot{p}_j}}\right)\right]\right\}
\end{aligned} \tag{22}$$

Finally, the LB and UB of the maximum mid-span acceleration are given by:

$$\begin{aligned}
\underline{\ddot{q}}_{\max} &= \min_{\mathbf{x}\in\mathbf{x}^I}\left\{\left[\sqrt{2\ln(2\nu_{\ddot{p}_j}T)} + \frac{0.5772}{\sqrt{2\ln(2\nu_{\ddot{p}_j}T)}}\right]\sigma_{\ddot{p}_j}\right\} \\
\overline{\ddot{q}}_{\max} &= \max_{\mathbf{x}\in\mathbf{x}^I}\left\{\left[\sqrt{2\ln(2\nu_{\ddot{p}_j}T)} + \frac{0.5772}{\sqrt{2\ln(2\nu_{\ddot{p}_j}T)}}\right]\sigma_{\ddot{p}_j}\right\}
\end{aligned} \tag{23}$$

Since analytical expressions of the functions of interest are available, the global minimum and maximum under the constraint that the optimization parameters range within prescribed intervals can be evaluated using the built-in functions “Minimize” and “Maximize” of Wolfram Mathematica (Wolfram Research 2020).

4. Numerical application

In this Section, serviceability assessment of an ideal footbridge is carried out. Results of a conventional evaluation are compared with those deriving from interval analysis. In particular, at first interval analysis is

carried out assuming fixed normalized deviation amplitudes of the parameters. Then, an increasing degree of uncertainty is considered according to the following definition of the interval variables:

$$x_k^I(\beta) = x_{0,k}\left(1+\beta\Delta x_k \hat{e}_k^I\right), \qquad 0 \le \beta \le 1. \tag{24}$$

Furthermore, three different analyses are carried out: considering all the parameters as uncertain, i.e. full uncertainty analysis (FU), taking into account only the uncertainties in the structural parameters (SU), and taking into account only uncertainties in the loading parameters (LU).

An ideal, steel, simply-supported beam with span length L=50 m, a deck width b=2 m is analyzed (Tubino et al. 2020). Considering load category III, sparse crowd (SETRA, 2006), a pedestrian density ρ=0.5 pedestrians/m^2 is assumed, i.e. N_p=50, and the predicted maximum acceleration applying SETRA procedure is $\ddot{q}_{\max}$ =2.15 m/s^2, falling into a minimum comfort class according to Table I.

Table II reports the midpoint and the normalized deviation amplitude of the uncertain parameters assumed in numerical simulations, derived from the literature analysis in Tubino et al. (2020). It should be remarked that the degree of uncertainty of the damping is very large. However, the assumed range of variation is in accordance with what proposed by HIVOSS Guidelines for steel footbridges, suggesting a minimum damping 0.2% and a mean value of 0.4%.

Table II. Midpoint and normalized deviation amplitude

Parameter		$x_{0,k}$	Δx_k
$\mu_{\omega p}$	[rad/s]	2 π 1.88	0.03
$\sigma_{\omega p}$	[rad/s]	2 π 0.17	0.17
α	[ ]	0.35	0.20
G	[N]	700	0.10
ω_j	[rad/s]	2 π 1.88	0.10
ξ_j	[%]	0.42	0.560
M_j	[kg]	50000	0.10

The nominal solutions for the standard deviation and the mean-value of the maximum acceleration are $\sigma_{\ddot{p}_j} = 0.50$ m/s^2 (Eq. (10)), $\ddot{q}_{\max} = 1.94$ m/s^2 (Eq. (15)). Table III reports the LB and UB of $\sigma_{\ddot{p}_j}^I$ (see Eq. (21)) and $\ddot{q}_{\max}^I$ (see Eq. (23)) obtained in the three analyses. The significant deviation amplitudes of the uncertain parameters (Table II) lead to very large interval of variation of the response, especially if all the parameters (structural and loading) are considered as uncertain (FU). However, the uncertainty in the dynamic response is mainly governed by the uncertainties in the structural parameters. When only loading parameters are assumed as uncertain (LU), the interval of variation of the response is significantly reduced. It should be remarked that the maximum acceleration obtained applying SETRA procedure is higher than the nominal value and it is close to the UB provided by interval analysis if only loading uncertainties are accounted for (LU). In all the other cases (FU and SU), SETRA prediction is well below the UB provided by interval analysis, and it can be non-conservative.

	Table III. LB and UB of the standard deviation of the response and of the maximum acceleration			
		FU	*SU*	*LU*
$\underline{\sigma}_{\ddot{p}_j}$	[m/s^2]	0.13	0.25	0.32
$\overline{\sigma}_{\ddot{p}_j}$	[m/s^2]	1.29	0.88	0.72
$\underline{\ddot{q}}_{\max}$	[m/s^2]	0.49	0.97	1.26
$\overline{\ddot{q}}_{\max}$	[m/s^2]	5.02	3.42	2.81

Figure 1 provides the UB and LB of $\sigma^I_{\ddot{p}_j}$ versus the coefficient β which measures the degree of uncertainty according to Eq. (24): results obtained for the three different analyses (FU, SU, LU) are reported. It can be deduced that, while the LB of the standard deviation tends to decrease almost linearly with the degree of uncertainty β, the increase of the UB is not linear with β. As a consequence, the LB and UB are not symmetric with respect to the nominal value and the amplitude of the interval of variation is not increasing linearly with β. The comparison among the different analyses (FU, SU, LU) confirms that the uncertainty in the standard deviation is mainly governed by the structural uncertainties. In order to have a narrow interval of variation of the response, the degree of uncertainty should be reduced (i.e. β should be significantly smaller than 1).

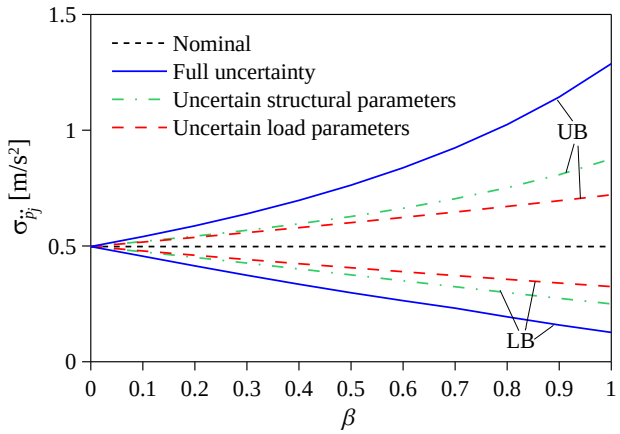


Figure 1. UB and LB of the standard deviation for increasing degree of uncertainty.

A probabilistic assessment of the footbridge serviceability can be obtained through the analysis of the CDF of the extreme value of the acceleration. Figure 2 provides the LB and UB of the CDF of the maximum acceleration for two degrees of uncertainty (β=0.5, β=1), for the full uncertainty analysis (FU, Fig. 2(a)), taking into account only the uncertainties in the structural parameters (SU, Fig. 2(b)), and taking into account only uncertainties in the loading parameters (LU, Fig. 2(c)). From Fig. 2, the LB and UB of the probability of falling within each comfort class can be estimated. It can be deduced that if the nominal values of the parameters are assumed, the footbridge will surely fall within a minimum comfort class. When all the parameters are considered as uncertain (Fig. 2(a)) with the maximum degree of uncertainty, the UB and LB of the CDF are extremely far from each other: considering the LB of the CDF, the maximum footbridge acceleration is surely higher than 4 m/s^2, and thus the comfort level is unacceptable, while

considering the UB of the CDF, the maximum acceleration is almost surely lower than 0.5 m/s^2 and the maximum comfort level is assured. When only the structural or the loading parameters are considered as uncertain (Fig. 2(b) and 2(c)), the interval of variation of the CDF becomes narrower. For example, if only loading parameters are considered as uncertain (Fig. 2(c)) and β=0.5, both LB and UB of the CDF provide a minimum comfort class.

As an alternative to the analysis of the CDF, an interval-based serviceability assessment of the footbridge can be carried out focusing on the mean value of the maximum acceleration $\ddot{q}^I_{\max}$. Figure 3 provides the UB and LB of $\ddot{q}^I_{\max}$, compared with the comfort limits provided by SETRA, versus the coefficient β which measures the degree of uncertainty according to Eq. (24). Results obtained for the three different analyses (FU, SU, LU) are reported. It can be deduced that, if all the parameters are considered as uncertain and the maximum degree of uncertainty is assumed, then the interval of variation of the expected value of the maximum response is so wide that the footbridge could fall in any comfort class (from unacceptable to maximum comfort). If only the uncertainties of the loading are considered, the footbridge is almost surely in a minimum comfort class, as predicted by SETRA guideline. As regards structural uncertainties, if the degree of uncertainty is small then interval analysis provides a minimum comfort class, while for large degree of uncertainty a comfort level ranging from minimum to unacceptable can be obtained.

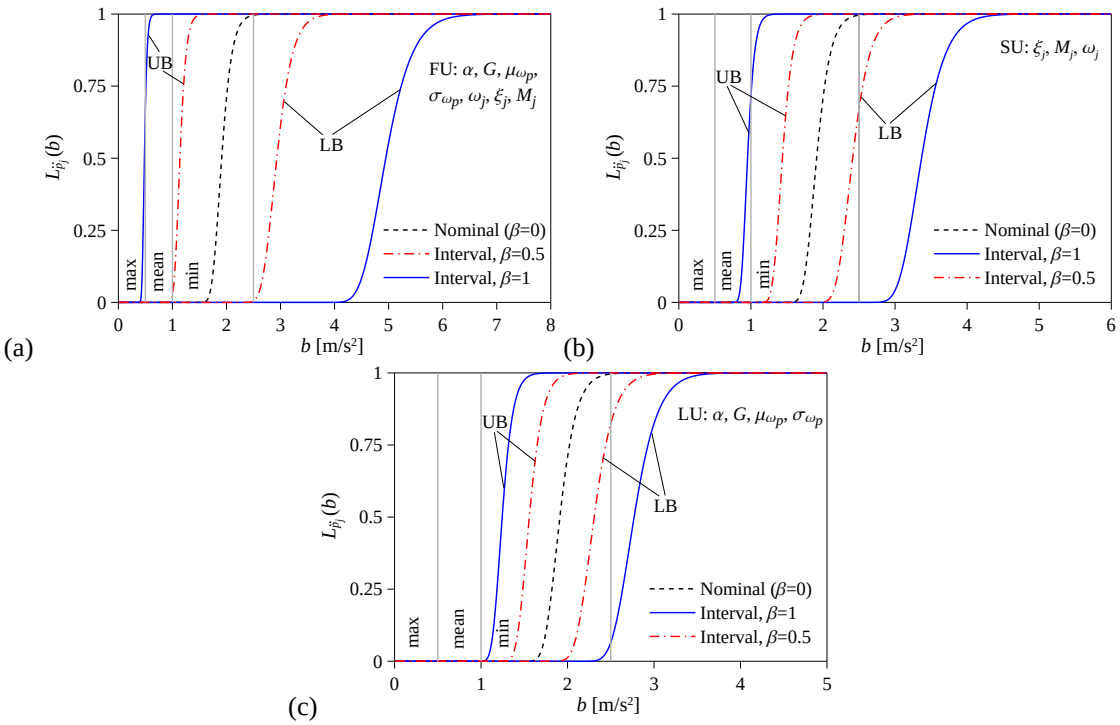


Figure 2. UB and LB of the CDF: full uncertainty (a), structural uncertainty (b), loading uncertainty(c).

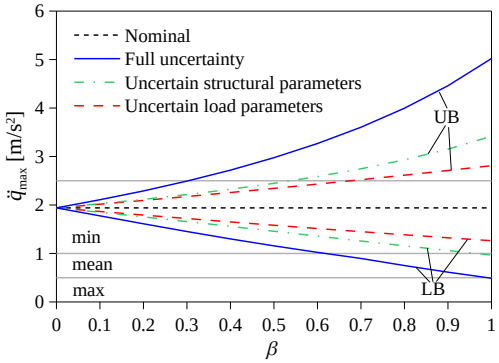


Figure 3. UB and LB of the mean value of the maximum acceleration for increasing degree of uncertainty.

5. Conclusions

Serviceability assessment of footbridges is studied through a non-deterministic approach, modelling structural and loading parameters as interval variables. Based on this approach, the intervals of variation of the standard deviation of the footbridge acceleration, of the cumulative distribution function of the maximum dynamic response and of the expected value of the maximum acceleration are defined. The analysis of the cumulative distribution function of the extreme value of the acceleration allows carrying out a probabilistic assessment of the footbridge serviceability.

The application to an ideal case study has shown that the high degree of uncertainty in the loading and structural parameters leads to very large intervals of variation of the response. The separate analysis of the role of the structural and loading parameters has shown that uncertainty in the structural parameters provides a determinant contribution to the variation of the response. In order to obtain narrow intervals of variation of the response, a reliable characterization of the loading and of the structural properties is mandatory. If all the parameters are considered as uncertain and a relatively high degree of uncertainty is assumed, then the interval of variation of the response is so wide that the footbridge could fall in any comfort class (from unacceptable to maximum comfort).

References

Brownjohn, J.M.W., Pavic, A. and P. Omenzetter. A spectral density approach for modelling continuous vertical forces on pedestrian structures due to walking. *Canadian Journal of Civil Engineering*, 31:65–77, 2004.

Davenport, A.G. Note on the distribution of the largest value of a random function with application to the gust loading. *Proc Inst Civ Eng* 28(2): 187–196, 1964.

HIVOSS. *Design of footbridges guideline. Human induced vibrations of steel structures*, http://www.stb.rwth-aachen.de/projekte/2007/HIVOSS/download.php RFS2-CT-2007-00033, 2009.

Faes, M. and D. Moens. Recent trends in the modeling and quantification of non-probabilistic uncertainty. *Archives of Computational Methods in Engineering*, 27: 633–671, 2020.

Lievens, K., Lombaert, G., De Roeck, G. and P. Van den Broeck. Robust design of a TMD for vibration serviceability of a footbridge. *Engineering Structures*, 123: 408-414, 2016.

Moore, R.E., Kearfott, R.B. and M.J. Cloud. *Introduction to Interval Analysis*, SIAM, Philadelphia, 2009.

Muscolino, G. and A. Sofi. Stochastic analysis of structures with uncertain-but-bounded parameters via improved interval analysis. *Probabilistic Engineering Mechanics*, 28: 152-163, 2012.

Muscolino, G., Santoro, R. and A. Sofi. Reliability assessment of structural systems with interval uncertainties under spectrum-compatible seismic excitations. *Probabilistic Engineering Mechanics*, 44: 138-149, 2016.

Piccardo, G. and F. Tubino. Equivalent spectral model and maximum dynamic response for the serviceability analysis of footbridges. *Engineering Structures*, 40: 445-456, 2012.

Racic, V., Pavic, A. and J.M.W. Brownjohn. Experimental identification and analytical modeling of human walking forces: Literature review. *Journal of Sound and Vibration*, 326: 1–49, 2009.

SETRA. *Footbridges – Assessment of vibrational behaviour of footbridges under pedestrian loading*. Paris: Technical Department for Transport, Roads and Bridges Engineering and Road Safety. Ministry of Transport and Infrastructure, 2006.

Sofi, A., Romeo, E., Barrera, O. and A. Cocks. An interval finite element method for the analysis of structures with spatially varying uncertainties. *Advances in Engineering Software*, 128: 1-19, 2019.

Tubino, F., Pagnini, L. and G. Piccardo. Uncertainty propagation in the serviceability assessment of footbridges. *Structure and Infrastructure Engineering*, 16(1): 123-137, 2020.

Tubino F. and G. Piccardo. Serviceability assessment of footbridges in unrestricted pedestrian traffic condition. *Structure and Infrastructure Engineering*, 12(12): 1652-1662, 2016.

Wolfram Research, Inc., Mathematica, Version 12.1, Champaign, IL, 2020.

Virtual Corrosion Initiation Monitoring for Infrastructures Considering Environmental Uncertainty

Yuguo Yu
Centre for Infrastructure Engineering and Safety, School of Civil and Environmental Engineering, The University of New South Wales, Sydney, NSW 2052, Australia Email: yuguo.yu@unsw.edu.au

Wei Gao
Centre for Infrastructure Engineering and Safety, School of Civil and Environmental Engineering, The University of New South Wales, Sydney, NSW 2052, Australia Email: w.gao@unsw.edu.au

Qihan Wang
Centre for Infrastructure Engineering and Safety, School of Civil and Environmental Engineering, The University of New South Wales, Sydney, NSW 2052, Australia Email: qihan.wang@unsw.edu.au

Bin Dong
Centre for Infrastructure Engineering and Safety, School of Civil and Environmental Engineering, The University of New South Wales, Sydney, NSW 2052, Australia Email: bin.dong@unsw.edu.au

Abstract: Material decay and corrosion of reinforced concrete infrastructures have long been recognized as major damage indicators in modern structural health monitoring. In order to further transform from reactive monitoring to proactive prediction and risk prevention, the key resides in the developments of high-fidelity physics-based modelling method to generate robust prior estimates considering uncertainty. To this end, this paper is concerned with a novel metamodeling on the chloride-induced corrosion initiation of reinforced concrete infrastructures. The physics-based modelling of uncertain material decay represents a chemophysical process, involving various coupling effects and environmental uncertainties, leading to a complex stochastic system that requires computational-heavy approaches to generate useful statistical analyses. In this regard, the computational intensiveness is addressed by the recently reported extended support vector regression (X-SVR) method. The advanced performance of the proposed approach is explored by modelling a real-life reinforced concrete parking infrastructure.

Keywords: Chemophysical modelling; Degradation; Machine learning; Virtual corrosion monitoring

1. Introduction

Structural health is concerned with the probabilistic responses of materials and structures in widespread areas. In civil engineering, one major durability concern is the reinforcement corrosion (Panesar and Ching 2018), which is found in concrete infrastructures subjected to environmental decays. The mechanism of corrosion initiation is known to be caused by the penetration and the accumulation of chloride (Muthulingam and Rao 2014, Hornbostel et al. 2016, Jin et al. 2018). However, many uncertain factors could substantially complicate the detailed assessment (Ann and Song 2007).

© 2021 *by Yuguo Yu, Wei Gao, Qihan Wang, Bin Dong.*

This paper deals with the virtual corrosion monitoring on the probabilistic corrosion initiation of a real-life parking infrastructure under the seasonal application of de-icing salt. Most of the reported studies on this issue were deterministic thus far (Samson and Marchand 2007), and simplifications on the degradation models were heavily used in the reported probabilistic studies (Hackl and Kohler 2016, Li and Ye 2018, Zhang 2018). In durability assessment, most of the models for chloride ingress were based on the Fick's law and prescribed binding isotherms, where many coupling effects were ignored. As a result, the low fidelity of the reported methods in the real-life scenarios is suspected to generate only poor estimates.

The technical difficulties hindering the comprehensive durability analysis under uncertainty are twofold. First, modelling the material decay represents a chemophysical problem (Yu et al. 2015, Yu and Zhang 2017, Yu and Zhang 2018, Yu and Zhang 2018, Yu et al. 2019), where the analysis under uncertainty involves solving stochastic partial differential equations (SPDEs) of multi-species transportation coupled with algebraic equations for chemical reactions. In order to achieve a informed statistical analysis, solving such a complex system often requires exhausting Monte Carlo simulations (MCS) (Yu et al. 2018). The computational intensiveness thus becomes a major concern, which represents the second difficulty. To address these issues, the metamodeling method combining machine learning with the stochastic chemophysical modelling of corrosion initiation is exploited in this paper.

In the following context, the developments of the novel machine learning algorithm and the stochastic chemophysical finite element analysis are presented. Focusing on the virtual corrosion monitoring of reinforced concrete infrastructure subjected to seasonal applications of de-icing salt, the details of the stochastic chemophysical models are presented in Section 2. The recently reported machine learning algorithm, i.e. the extended support vector regression (X-SVR), is then presented in Section 3. The effectiveness and efficiency of performing the machine learning aided corrosion initiation analysis is explored in Section 4, with some concluding remarks presented in Section 5.

2. Modelling chloride-induced corrosion initiation

To conduct high-fidelity corrosion initiation monitoring, a variety of chemical reactions associated with the hydrated cement system must by innovatively incorporated. Firstly, to reduce computational cost, the operator splitting approach (OSA) is implemented, and chemical reactions are evaluated by the thermodynamic modelling method (Yu et al. 2015). Herein, the general format of chemical equilibrium model following thermodynamic modelling is given below, where the details regarding the complete set of chemical reactions as considered in the modelled were reported in (Yu et al. 2019).

$$K_m = \prod_i [u_i]^{\nu_{i,m}} \gamma_i^{\nu_{i,m}} \Big/ \prod_j [u_j]^{\nu_{j,m}} \gamma_j^{\nu_{j,m}},\ m \in [1, M] \tag{1}$$

where K_m is equilibrium constant (-) for the *m*th reaction, $[u_i], \gamma_i, \nu_i$ and $[u_j], \gamma_j, \nu_j$ are concentration (M), activity coefficient (-) and valence number (-) of the *i*th product and the *j*th reactant respectively.

Furthermore, in addition to the general chemical equilibrium of the hydrated cement system, the most crucial impact to be considered in virtual corrosion monitoring is the chloride binding effect during exposure. Herein, the chloride binding is modelled in detail by differentiating the binding mechanism, i.e.,

physical and chemical bindings. In specific, physical binding describes the ionic exchange between the free ionic chloride and the hydroxyl group at the electrical double layer of the C-S-H gel. The free chloride is physically attached to the binding sites without creating new compounds, hence the physical binding.

$$\underbrace{\left[\text{C-S-H}-\text{OH}^-\right]}_{\text{C-S-H gel}}+\text{Na}^++\text{Cl}^-\rightleftarrows\underbrace{\left[\text{C-S-H}-\text{Cl}^-\right]}_{\text{C-S-H gel}}+\text{Na}^++\text{OH}^- \tag{2}$$

Following the double layer theory (Pan et al. 2010), the physical chloride binding is calculated as:

$$c_b^{\text{phy}}=\frac{\text{A}_\text{s}\text{N}_\text{s}K_{\text{phy}}\left[\text{Cl}^-\right]}{\left[\text{OH}^-\right]+K_{\text{phy}}\left[\text{Cl}^-\right]}N_{\text{C-S-H}} \tag{3}$$

where N_s is total binding capacity (mol/m^2) on the C-S-H surface, which is reported to be $1.793\times10^{-7}\,\text{mol}/\text{m}^2$ (Pan et al. 2010), K_{phy} is physical binding coefficient (-), reported to be 0.0698 (Baroghel-Bouny et al. 2012), A_s is surface area of the C-S-H gel ($250916\,\text{m}^2/\text{mol}$ of the C-S-H gel), $N_{\text{C-S-H}}$ is remaining solid content of the C-S-H gel ($\text{mol}/\text{m}^3=\text{mmol}/\text{L}$ of the material), and c_b^{phy} is amount of (mmol/L of material) physically-bound chloride content.

On the other hand, chemical binding indicates the reaction between free ionic chloride and monosulfate, leading to the formations of new compounds. The chemical binding is generally attributed to the formation of the Friedel's salt due to its dominant amount (Baroghel-Bouny et al. 2012).

$$\underbrace{\left[\text{AFm}-\text{SO}_4^{2-}\right]}_{\text{Monosulfate}}+2\text{Cl}^-\rightleftarrows\underbrace{\left[\text{AFm}-2\text{Cl}^-\right]}_{\text{Friedel's salt}}+\text{SO}_4^{2-} \tag{4}$$

Unlike the dissolution of Friedel's salt modelled as a part of the AFm solid solution, the formation of Friedel's salt follows another ionic exchange (Lito et al. 2012), and is calculated as:

$$c_b^{\text{che}}=\frac{2\times K_{\text{che}}\left\{\text{Cl}^-\right\}^2}{\left\{\text{SO}_4^{2\text{-}}\right\}+K_{\text{che}}\left\{\text{Cl}^-\right\}^2}N_{\text{AFm}} \tag{5}$$

where {...} indicates activity of diffusive reactant (mol/L of solution), K_{che} is equilibrium coefficient for ionic exchange, assumed to be 0.1 at 23°C (Bothe and Brown 2004), N_{AFm} (mmol/L of material) represents the remaining amount of AFm solid solution that is occupied by Friedel's salt and monosulfate, and c_b^{che} is amount (mmol/L of material) of the chemically-bound chloride content.

The ionic exchange of chloride binding would result in retardation of chloride ingress, which should be included in the reactant transportation. To do so, the derivatives of ionic exchange are calculated:

$$\frac{\partial c_b^{\text{phy}}}{\partial c_{\text{f}}}=\frac{\partial c_b^{\text{phy}}}{\partial\left[\text{Cl}^-\right]}=\frac{\text{A}_{\text{s}}\text{N}_{\text{s}}K_{\text{phy}}\left[\text{OH}^-\right]}{\left(\left[\text{OH}^-\right]+K_{\text{phy}}\left[\text{Cl}^-\right]\right)^2}N_{\text{C-S-H}}\ \text{and}\ \frac{\partial c_b^{\text{che}}}{\partial c_{\text{f}}}=\frac{\partial c_b^{\text{che}}}{\partial\left[\text{Cl}^-\right]}=\frac{4\times K_{\text{che}}\gamma_{\text{Cl}}^2\left[\text{Cl}^-\right]\left\{\text{SO}_4^{2-}\right\}}{\left(\left\{\text{SO}_4^{2-}\right\}+K_{\text{che}}\left\{\text{Cl}^-\right\}^2\right)^2}N_{\text{AFm}} \quad (6)$$

By doing so, the stochastic Nernst-Planck equation for chloride penetration is formulated as follow:

$$\left(\frac{\partial c_b^{\text{phy}}}{\partial c_{\text{f}}}+\frac{\partial c_b^{\text{che}}}{\partial c_{\text{f}}}\right)\frac{\partial c_{\text{f}}}{\partial t}+\frac{\partial w_{\text{L}}c_{\text{f}}}{\partial t}=\text{div}\left[\begin{array}{l}\boldsymbol{D}_{\text{f}}w_{\text{L}}\text{grad}(c_{\text{f}})+\dfrac{\boldsymbol{D}_{\text{f}}\mathbf{z}\text{F}}{\text{R}T}c_{\text{f}}\text{grad}(\psi)+\boldsymbol{D}_{\text{f}}w_{\text{L}}c_{\text{f}}\text{grad}(\ln\gamma)\\+\dfrac{\boldsymbol{D}_{\text{f}}c_{\text{f}}\ln(\gamma c_{\text{f}})}{T}\text{grad}(T)+c_{\text{f}}\boldsymbol{D}_{\text{L}}\text{grad}(RH)\end{array}\right]+Q_{\hat{U}}\frac{\partial\boldsymbol{W}_{\hat{U}}}{\partial t} \quad (7)$$

where w_{L} is volumetric liquid content (m^3/m^3), $\mathbf{z}$ is valence number of the diffusive reactant (-), F is Faraday constant (96488.46C/mol), R is ideal gas constant (8.3143J/mol/K), T is thermodynamic temperature (K), ψ is electrical potential (V), γ is chemical activity coefficient (-), $\boldsymbol{D}$ is diffusivity vector (m^2/s), RH is relative humidity (%). In addition, $Q_{\hat{U}}\partial\boldsymbol{W}_{\hat{U}}$ are the noise terms, and $Q_{\hat{U}}$ is the covariance relying on the solutions to diffusive reactants, chemical reactions, and other coupling effects.

3. Extended support vector regression (X-SVR)

To achieve stochastic analyses with acceptable accuracy, the scale of the MCS often ranges from 10000 to a couple of millions. Thus, machine learning based methods, as a surrogate model (Ben Abdessalem and El-Hami 2015), become a popular solution to mitigate computational intensiveness. Herein, a recently developed support vector regression (SVR) algorithm is introduced, which was extended from the doubly regularised SVM (DrSVM) for binary classification to regression (Dunbar et al. 2010, Li et al. 2016), hence the extended support vector regression (X-SVR).

In X-SVR, instead of using the linear ε-insensitive loss function as in the conventional ε-SVR (Smola and Schölkopf 2004), a quadratic ε-insensitive loss function (l_2^{ε}) is adopted.

$$l_2^{\varepsilon}(y_{train}^i-\hat{f}(\mathbf{x}_i))=\left|y_{train}^i-\hat{f}(\mathbf{x}_i)\right|^2,i=1,...m \quad (8)$$

where $\mathbf{x}$ and $\mathbf{y}_{train}\in\Re^m$ are the input vector and response from the available knowledge, i.e. the training dataset, $\hat{f}(\mathbf{x}_i)$ indicates surrogate model prediction.

Although applying the quadratic ε-insensitive loss function enhances the numerical stability in X-SVR, the algorithm may still struggle to model problems of high nonlinearity. In the present case of predicting the stochastic corrosion initiation, the multi-physical problem and the associated uncertainty would lead to a highly nonlinear regression. Under this circumstance, a surrogate model may not be found directly based on

the input features, and further processing on $\mathbf{x}_{train}$ is required. To do so, it is achieved by mapping raw input data from the low-dimensional space $\Re^n$ into a higher-dimensional feature space F with an empirical mapping function $\mathbf{\Phi}(\mathbf{x}_i)$ (Kung 2014). Such an approach leads to the development of a kernelized X-SVR.

$$\mathbf{x}_i = [x_{i,1}, x_{i,2}, ..., x_{i,n}]^T \mapsto \hat{\mathbf{k}}(\mathbf{x}_i) = \begin{bmatrix} \mathbf{\Phi}(\mathbf{x}_1)^T \mathbf{\Phi}(\mathbf{x}_i) \\ \mathbf{\Phi}(\mathbf{x}_2)^T \mathbf{\Phi}(\mathbf{x}_i) \\ \vdots \\ \mathbf{\Phi}(\mathbf{x}_m)^T \mathbf{\Phi}(\mathbf{x}_i) \end{bmatrix} = \begin{bmatrix} K(\mathbf{x}_1, \mathbf{x}_i) \\ K(\mathbf{x}_2, \mathbf{x}_i) \\ \vdots \\ K(\mathbf{x}_m, \mathbf{x}_i) \end{bmatrix} \tag{9}$$

Through kernel mapping, the empirical feature vector $\hat{\mathbf{k}}(\mathbf{x}_i)$ is regarded as training samples, and the surrogate model generated from kernelized X-SVR model is derived by solving the constrained optimization problem as follow:

$$\min_{\mathbf{z}_k, \gamma} : \frac{1}{2}(\mathbf{z}_k^T \hat{\mathbf{C}}_k \mathbf{z}_k + \gamma^2) + \lambda_2 \mathbf{b}_k^T \mathbf{z}_k \quad s.t. \quad (\hat{\mathbf{A}}_k + \mathbf{I}_{4m\times 4m})\mathbf{z}_k + (\varepsilon \mathbf{I}_{4m\times 4m} + \gamma \hat{\mathbf{G}}_k)\hat{\mathbf{e}}_k + \hat{\mathbf{d}}_k \geq \mathbf{0}_{4m} \tag{10}$$

where $\mathbf{I}_{4m\times 4m} \in \Re^{4m\times 4m}$ and $\mathbf{0}_{4m} \in \Re^{4m}$ denote the identity matrix and the zero vector. The matrices $\hat{\mathbf{C}}_k$, $\hat{\mathbf{G}}_k$ and $\hat{\mathbf{A}}_k$, as well as the vectors $\mathbf{b}_k$, $\hat{\mathbf{e}}_k$, $\hat{\mathbf{d}}_k$ and $\mathbf{z}_k$ were carefully defined in (Yu et al. 2019). More importantly, the proof that the X-SVR would lead to global optimum solution have also been discussed intensively by (Yu et al. 2019).

4. Results and analyses

The performance of the proposed method is examined here by studying the real-life case as found in a 20-year-old concrete parking structure (Samson and Marchand 2007). The structure was in Canada, where the harsh winter led to seasonal adoption of de-icing salt.

The concrete material of the parking structure is reported to be made of the CSA Type-10 ordinary Portland cement with a *w*/*b* ratio of 0.45. The detailed material properties, such as hydrate contents, solution composition, microstructure and thermal conductivity, are referred to the reported study (Samson and Marchand 2007). The physical monitoring was conducted by measuring the chloride profile in the core samples extracted from the slab after 20-year exposure. A schematic diagram describing the field corrosion initiation monitoring is presented in the Figure 1.

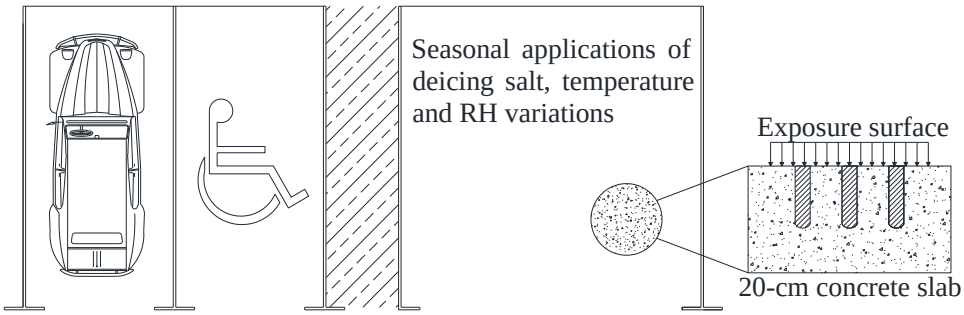


Figure 1. Schematic diagram of physical corrosion monitoring on a parking infrastructure

Since field structure is bound to interact with environment, the representations of boundary conditions are the decisive factor for the accurate virtual monitoring. According to (Samson and Marchand 2007), the data as collected from the local weather station reflecting the annual variations of temperature, RH and surface concentration of de-icing salt are adopted to represent the in-situ boundary conditions, see Figure 2.

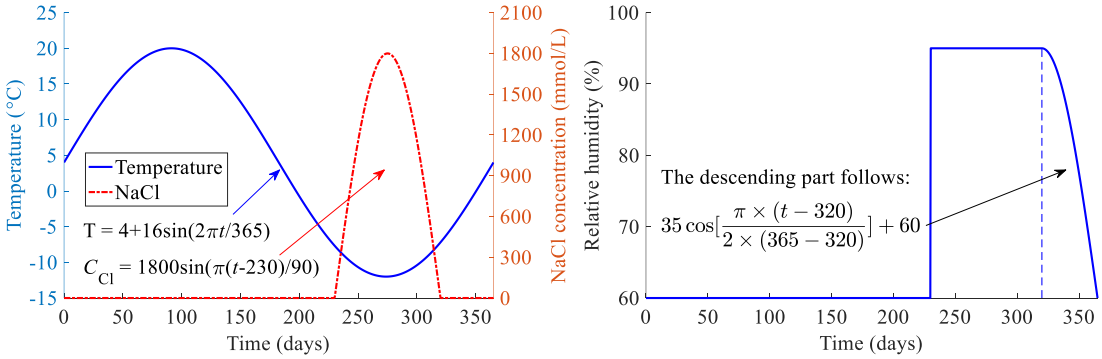


Figure 2. Annual variations in the boundary conditions: (a) temperature and NaCl; (b) RH

Considering the variations in environmental conditions, the associated uncertainties are thus assigned to the peak dosage of de-icing salt ($C^R_{\mathrm{Cl}_{\max}}$), annual average temperature (T^R_{ave}) and relative humidity (RH^R_{ave}). The assumed statistical representations for environmental uncertainties are summarised in the Table I.

Table I. Random parameters for the durability assessment

Variables	Mean	COV	Distribution type
RH^R_{ave}	60%	0.03	Lognormal
T^R_{ave}	$4\,^{\circ}C$	0.05	Normal
$C^R_{\mathrm{Cl}_{\max}}$	$1800\,\mathrm{mmol/L}$	0.056	Uniform

The full-scale MCS (1000,000 function evaluations) using the stochastic chemophysical model is conducted first. The modelling of the long-term chloride ingress involves of a 20-year seasonal exposure of concrete to the de-icing salt. The MCS results are obtained for the chloride profiles at Year 5, 10, 15 and 20

to demonstrate the chloride ingress process and the uncertainty propagation. In addition, the field measurements after the 20-year exposure are also presented for inspection.

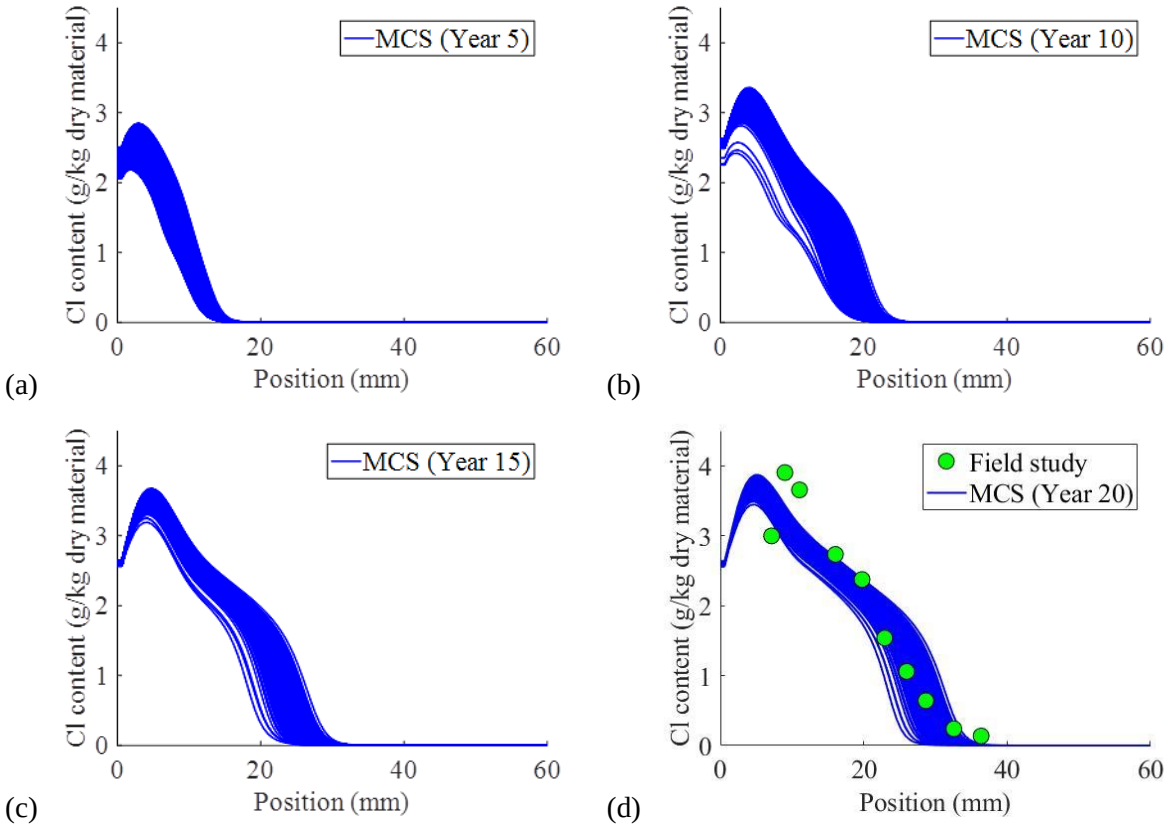


Figure 3. Comparisons between the full-scale MCS and the field measurement

Examining *Figure 3*, near the exposure surface, the chloride profile experiences a localized peak due to the strong advection effect. However, this peak of chloride content is not critical for assessing corrosion initiation. Instead, the chloride content at the reinforcement surface are the main focuses. Herein, both ε-SVR and X-SVR are implemented with a Gaussian kernel. The number of training samples used for deriving the surrogate model is set to be 50, and the obtained results are compared to the full-scale MCS. Specifically, by assuming the cover depth of 20mm and 25mm for a typical slab structure as adopted in public parking infrastructure, the chloride content at the reinforcement surface at Year 10 and 20 are examined, see the Figures 4-7.

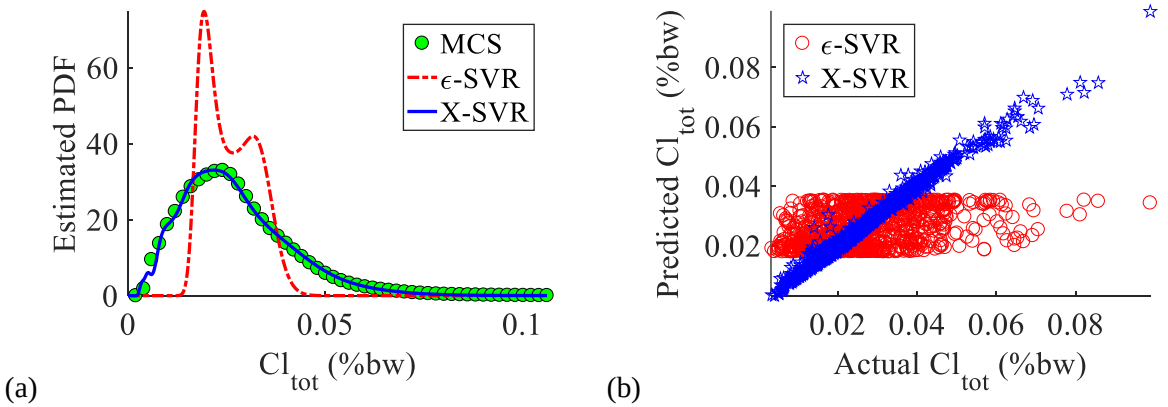


Figure 4. Total chloride content at the presumed concrete cover of 20mm after 10-year exposure

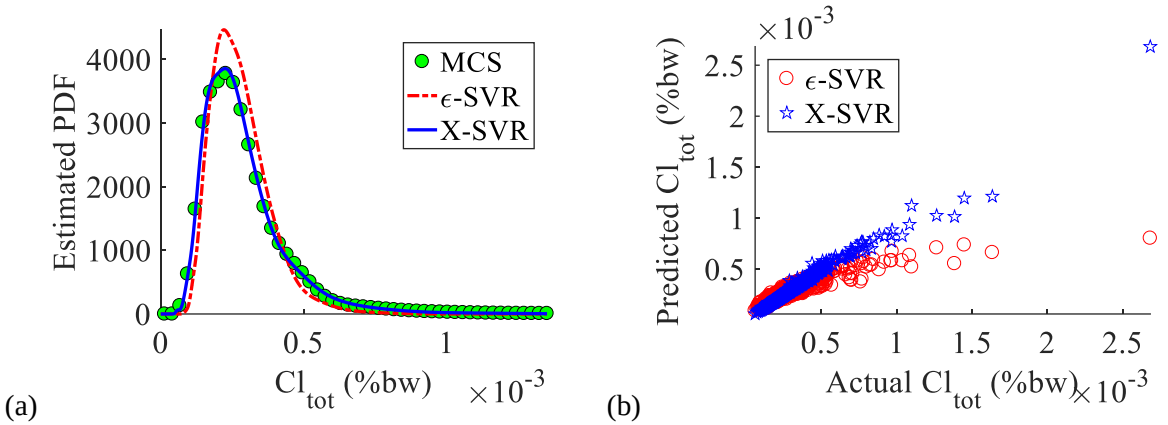


Figure 5. Total chloride content at the presumed concrete cover of 25mm after 10-year exposure

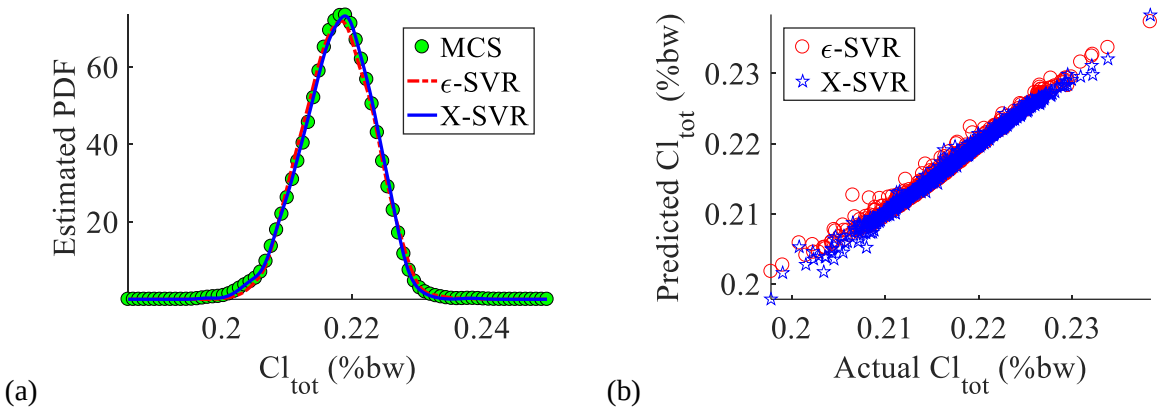


Figure 6. Total chloride content at the presumed concrete cover of 20mm after 20-year exposure

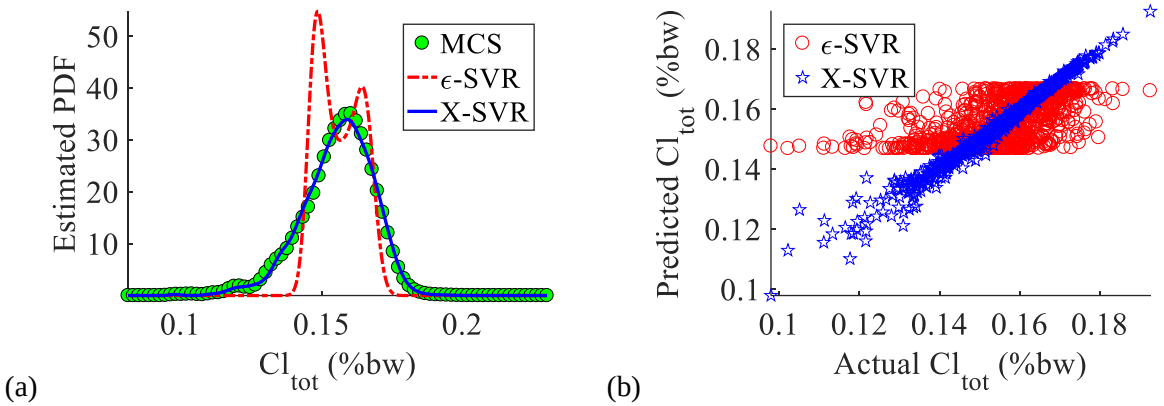


Figure 7. Total chloride content at the presumed concrete cover of 25mm after 20-year exposure

For surrogate modelling the chloride content at the cover depth, the conventional ε-SVR experiences more difficulty in precisely predicting the full-scale MCS. This is mainly due to the complex and nonlinear chemophysical mechanisms regarding the chloride transportation in a real-life case. As discussed in the previous sections, many interrelated factors would affect the overall transportation process. For instance, the changing boundaries would affect diffusivity, reactivity and moisture movement, leading to different paces in terms of the progressive degradation. The uncertain material decay process leads to more detailed differences as found in hydrate compositions, hence different chloride binding capacity and chloride penetration degrees in general. Furthermore, the distinct pace in the microstructural evolutions also brings contrasts to the material diffusivity and permeability. All these variations would cause interrelated influences within the chemophysical modelling. As a result, obtaining accurate correlation between chloride accumulation at the reinforcement surface and the uncertain environmental input indicates a complex regression problem for surrogate modelling, which is governed by the complex chemophysical process of high nonlinearity. As shown in the above figures, by using a low number of actual function evaluations, the X-SVR generates good predictions on the total chloride content regardless of cover depth and in-service duration. In comparison, the ε-SVR fails to generate meaningful predictions in multiple occasions.

5. Summary

Excellent performances of the recently developed X-SVR are demonstrated to be an effective and efficient surrogate modelling approach for the physics-based virtual monitoring of corrosion initiation considering environmental uncertainties. The X-SVR well predicts the total chloride accumulation at the reinforcement surface with low function evaluations required. i.e. training samples. In comparison, the performance of ε-SVR is inconsistent, which often fails to generate meaningful estimations on the stochastic responses.

Acknowledgements

This research work presented has been supported by Australian Research Council project IH150100006. The presented numerical computations are undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI), which is supported by the Australian Government.

References

Ann, K. Y. and H.-W. Song. Chloride threshold level for corrosion of steel in concrete. Corrosion Science, 49(11): 4113-4133, 2007.

Baroghel-Bouny, V., X. Wang, M. Thiery, M. Saillio and F. Barberon. Prediction of chloride binding isotherms of cementitious materials by analytical model or numerical inverse analysis. Cement and Concrete Research, 42(9): 1207-1224, 2012.

Ben Abdessalem, A. and A. El-Hami. A probabilistic approach for optimising hydroformed structures using local surrogate models to control failures. International Journal of Mechanical Sciences, 96-97(143-162, 2015.

Bothe, J. V. and P. W. Brown. PhreeqC modeling of Friedel's salt equilibria at 23±1 °C. Cement and Concrete Research, 34(6): 1057-1063, 2004.

Dunbar, M., J. M. Murray, L. A. Cysique, B. J. Brew and V. Jeyakumar. Simultaneous classification and feature selection via convex quadratic programming with application to HIV-associated neurocognitive disorder assessment. European Journal of Operational Research, 206(2): 470-478, 2010.

Hackl, J. and J. Kohler. Reliability assessment of deteriorating reinforced concrete structures by representing the coupled effect of corrosion initiation and progression by Bayesian networks. Structural Safety, 62(12-23, 2016.

Hornbostel, K., U. M. Angst, B. Elsener, C. K. Larsen and M. R. Geiker. Influence of mortar resistivity on the rate-limiting step of chloride-induced macro-cell corrosion of reinforcing steel. Corrosion Science, 110(46-56, 2016.

Jin, M., S. Gao, L. Jiang, H. Chu, M. Lu and F. F. Zhi. Degradation of concrete with addition of mineral admixture due to free chloride ion penetration under the effect of carbonation. Corrosion Science, 138(42-53, 2018.

Kung, S. Y. (2014). Kernel methods and machine learning. United Kingdom, Cambridge University Press.

Li, J., Y. Wang, Y. Cao and C. Xu. Weighted doubly regularized support vector machine and its application to microarray classification with noise. Neurocomputing, 173(595-605, 2016.

Li, Q. and X. Ye. Surface deterioration analysis for probabilistic durability design of RC structures in marine environment. Structural Safety, 75(13-23, 2018.

Lito, P. F., S. P. Cardoso, J. M. Loureiro and C. M. Silva (2012). Ion Exchange Equilibria and Kinetics. Ion Exchange Technology I: Theory and Materials. I. Dr and M. Luqman. Dordrecht, Springer Netherlands: 51-120.

Muthulingam, S. and B. N. Rao. Non-uniform time-to-corrosion initiation in steel reinforced concrete under chloride environment. Corrosion Science, 82(304-315, 2014.

Pan, T., K. Xia and L. Wang. Chloride binding to calcium silicate hydrates (C-S-H) in cement paste: a molecular dynamics analysis. International Journal of Pavement Engineering, 11(5): 367-379, 2010.

Panesar, D. K. and G.-H. Ching. Implications of coupled degradation mechanisms of cement based materials exposed to cold climates. International Journal of Mechanical Sciences, 144(865-876, 2018.

Samson, E. and J. Marchand. Modeling the effect of temperature on ionic transport in cementitious materials. Cement and Concrete Research, 37(3): 455-468, 2007.

Smola, A. J. and B. Schölkopf. A tutorial on support vector regression. Statistics and Computing, 14(3): 199-222, 2004.

Yu, Y., X. Chen, W. Gao, Q. Li, D. Wu and M. Liu. Stochastic leaching analysis on cementitious materials considering the influence of material uncertainty. Construction and Building Materials, 184(186-202, 2018.

Yu, Y., X. Chen, W. Gao, D. Wu and A. Castel. Impact of atmospheric marine environment on cementitious materials. Corrosion Science, 148(366-378, 2019.

Yu, Y., D. Wu, Q. Wang, X. Chen and W. Gao. Machine learning aided durability and safety analyses on cementitious composites and structures. International Journal of Mechanical Sciences, 160(165-181, 2019.

Yu, Y. and Y. X. Zhang. Coupling of chemical kinetics and thermodynamics for simulations of leaching of cement paste in ammonium nitrate solution. Cement and Concrete Research, 95(95-107, 2017.

Yu, Y. and Y. X. Zhang. Effect of capillary connectivity and crack density on the diffusivity of cementitious materials. International Journal of Mechanical Sciences, 144(849-857, 2018.

Yu, Y. and Y. X. Zhang. Numerical modelling of mechanical deterioration of cement mortar under external sulfate attack. Construction and Building Materials, 158(490-502, 2018.

Yu, Y., Y. X. Zhang and A. Khennane. Numerical modelling of degradation of cement-based materials under leaching and external sulfate attack. Computers & Structures, 158(1-14, 2015.

Zhang, H. Durability reliability analysis for corroding concrete structures under uncertainty. Mechanical Systems and Signal Processing, 101(26-37, 2018.

Free vibration of functionally graded graphene-platelets reinforced composite (FG-GPLRC) plate with interval parameters

Youqin Huang and Donghui Luo
Research Centre for Wind Engineering and Engineering Vibration, Guangzhou University, Guangzhou 510006, China, yqhuang@gzhu.edu.cn

Abstract: Functionally graded graphene-platelets (GPLs) reinforced composite (FG-GPLRC) is a recently proposed and promising nanocomposite, whose mechanical behaviors have been extensively studied in the field of composite structures. Studies on the uncertainty effect of material properties on performances of FG-GPLRC are very few, so an interval analysis is conducted in this work to investigate the influence of uncertain Young's modulus on natural frequencies of an FG-GPLRC plate. The study indicates that the position of the layer with interval parameter play a vital role on the uncertainty degree of first order natural frequency of the plate when only a single layer has an interval parameter. The weight fraction of GPLs in the whole plate should be decreased for controlling the deviation of natural frequency if all layers have interval parameters.

Key words: functionally graded graphene-platelets reinforced composite (FG-GPLRC), free vibration, interval analysis, generalized eigenvalue problem

1. Introduction

The material composition and mechanical properties of functionally graded materials (FGMs)can be tailored to continuously vary over one or more dimension(s) for simultaneously meeting different mechanical requirements (Suresh and Mortensen, 1998). Considering the extraordinary mechanical, thermal and electrical properties of graphene, Yang and his coworkers firstly proposed the functionally graded graphene-platelets (GPLs) reinforced composite (FG-GPLRC) (Yang et al. 2017; Song et al. 2017; Feng et al., 2017). Mechanical performances of FG-GPLRC composite have quickly attracted worldwide attentions in the research field of composite materials (Zhao et al. 2020).

The free vibration of FG-GPLRC polymer nanocomposite plate has been studied extensively. Zhang et al. (2020) investigated the free vibration of FG-GPLRC annular plate resting on elastic substrate and subjected to nonlinear temperature gradient. Quang et al. (2020) presented a three-variable high order shear deformation plate theory (THSDT) for free vibration analysis of FG-GPLRC porous plates. Thai and Phung-Van (2020) carried out free vibration analysis of FG-GPLRC plates with complicated shapes by a meshfree approach using a naturally stabilized nodal integration (NSNI). Chiker et al. (2020) compared the free vibrational behavior of FG-GPLRC plates and multilayer functionally graded carbon nanotube reinforced composite (FG-CNTRC) plates. Noroozi et al. (2020) developed a meshfree radial point interpolation method (RPIM) to investigate the free vibration of FG-GPLRC perforated plates. Majidi-Mozafari et al. (2020) provided an analytical solution for free vibration of an FG-GPLRC sandwich plate enclosed by piezoelectric layers according to Maxwell's equation.

© 2021 *by Youqin Huang and Donghui Luo.*

It should be mentioned that all the existing studies on FG-GPLRC composites have not considered the uncertainty of material parameters such as Yong's modulus. However, the presence of uncertainty in material performance is inevitable owing to that the fabrication of FG materials is still immature. Therefore, uncertainty free vibration analysis on an FG-GPLRC plate is conducted in this work by assuming the uncertain material parameter of Young's modulus as an interval variable.

2. Methods based on interval analysis

An FG-GPLRC plate with length a, width b and thickness h is shown in Figure 1. Due to the constraint of current manufacturing technology, the FG-GPLRC plate is simulated by a multilayer structure consisting of sufficient number of layers where GPL weight fraction (w.f. = weight of GPLs / weight of composite) follows a layer-wise change. The plate is composed of N_L layers and each layer has the same thickness $\Delta h = h / N_L$. Neighboring layers are firmly bonded, and there exists no slide among layers. The effective material parameters of the k-th layer including Young's modulus $E_C^{(k)}$ can be determined by the Voigt-Reuss mode (Guzmán and Miravete, 2007).

Four types of FG patterns are considered as demonstrated in Figure 2. The w.f. of GPLs varies linearly among layers. Pattern 1 is a special case with uniform GPLs in all layers. For Pattern 2, w.f. of GPLs decreases from midplane to surface of plate, and Pattern 3 is the opposite. In Pattern 4, w.f. of GPLs decreases from one side to the other side of the plate.

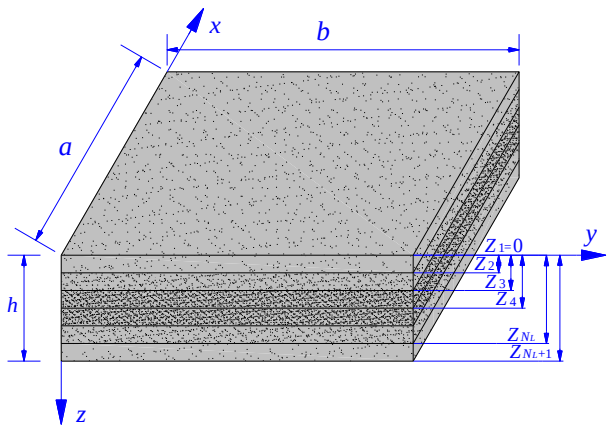
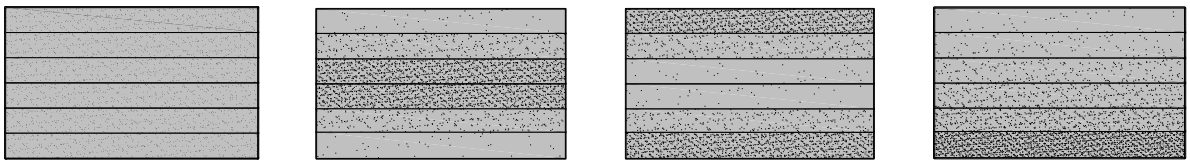


Figure 1. Multi-layer FG-GPLRC plate.



(a) Pattern 1 (b) Pattern 2 (c) Pattern 3 (d) Pattern 4

Figure 2. Different GPL distribution patterns.

Free vibration of FG-GPLRC plate with interval parameters

According to the work of Song et al. (2017), free vibration of an FG-GPLRC plate corresponds to a generalized eigenvalue problem:

$$\mathbf{K}\boldsymbol{\varphi} = \lambda\mathbf{M}\boldsymbol{\varphi} \tag{1}$$

where **M** and **K** are matrices of mass and stiffness; λ and $\boldsymbol{\varphi}$ are eigenvalue and eigenvector respectively, and λ is square of natural frequency, i.e., $\lambda = \omega^2$.

Expressions of elements in **M** and **K** can be found in Song et al. (2017), indicating that ω is related with Young's modulus $E_C^{(k)}$. Considering that $E_C^{(k)}$ could be uncertain owing to difficult manufacturing processes of FG materials and probabilistic information of such an uncertain parameter is not easy to be obtained, *Ec* is to be modeled as non-probabilistic interval variable in this study, and the interval analysis is carried out to evaluate the effect of uncertain Young's modulus of each layer on the natural frequencies of the FG-GPLRC plate.

An interval variable α^I denotes a variable closed by a lower bound $\underline{\alpha}$ and an upper bound $\overline{\alpha}$, i.e.

$$\alpha^I = [\underline{\alpha}, \overline{\alpha}] = \{\alpha \in R : \underline{\alpha} \le \alpha \le \overline{\alpha}\} \tag{2}$$

In engineering interval analysis, interval variable is generally represented by its midpoint and uncertainty degree:

$$\alpha^C = \frac{\overline{\alpha} + \underline{\alpha}}{2} \tag{3}$$

$$\Delta\alpha^I = \frac{\overline{\alpha} - \underline{\alpha}}{\overline{\alpha} + \underline{\alpha}} \tag{4}$$

For a system with r interval variables, the interval variables can be integrated as an interval vector, $\boldsymbol{\alpha}^I = \{\alpha_1^I, \alpha_2^I, \cdots, \alpha_r^I\}$, and free vibration of the system is governed by:

$$\mathbf{K}(\boldsymbol{\alpha})\boldsymbol{\varphi} = \lambda\mathbf{M}(\boldsymbol{\alpha})\boldsymbol{\varphi}, \ \boldsymbol{\alpha} \in \boldsymbol{\alpha}^I = [\underline{\boldsymbol{\alpha}}, \overline{\boldsymbol{\alpha}}] \tag{5}$$

For the studied FG-GPLRC plate, N_L interval parameters are considered, i.e.

$$\boldsymbol{\alpha}^I = \{E_C^{(1)I}, \cdots E_C^{(k)I}, \cdots, E_C^{(N_L)I}\} \tag{6}$$

The lower and upper bounds of natural frequency of any order can be obtained by solving two deterministic eigenvalue problem (Sofi et al., 2015):

$$\mathbf{K}(\boldsymbol{\alpha}^{(\mathrm{LB})})\boldsymbol{\varphi} = \underline{\lambda}\mathbf{M}(\boldsymbol{\alpha}^{(\mathrm{LB})})\boldsymbol{\varphi} \tag{7}$$

$$\mathbf{K}(\boldsymbol{\alpha}^{(\mathrm{UB})})\boldsymbol{\varphi} = \overline{\lambda}\mathbf{M}(\boldsymbol{\alpha}^{(\mathrm{UB})})\boldsymbol{\varphi} \tag{8}$$

The elements of $\boldsymbol{\alpha}^{(\mathrm{LB})}$ and $\boldsymbol{\alpha}^{(\mathrm{UB})}$ are determined by:

$$\alpha_i^{(\mathrm{UB})} = \overline{\alpha}_i, \alpha_i^{(\mathrm{LB})} = \underline{\alpha}_i \ \text{when} \ S_{\alpha_i} > 0 \tag{9}$$

$$\alpha_i^{(\mathrm{LB})} = \overline{\alpha}_i, \alpha_i^{(\mathrm{UB})} = \underline{\alpha}_i \ \text{when} \ S_{\alpha_i} < 0 \tag{10}$$

with the sensitivity function:

$$S_{\alpha_i} = \left.\frac{\partial\lambda(\boldsymbol{\alpha})}{\partial\alpha_i}\right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^C} = \boldsymbol{\phi}_0^{\mathrm{T}} \left.\frac{\partial\mathbf{K}(\boldsymbol{\alpha})}{\partial\alpha_i}\right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^C} \boldsymbol{\phi}_0 - \lambda_0\boldsymbol{\phi}_0^{\mathrm{T}} \left.\frac{\partial\mathbf{M}(\boldsymbol{\alpha})}{\partial\alpha_i}\right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^C} \boldsymbol{\phi}_0 \tag{11}$$

Hence, after all sensitivity functions are computed for all interval parameters, the bounds of natural frequency of any order of the FG-GPLRC plate can be obtained by solving Eqs. (7) and (8).

3. Results and discussions

An FG-GPLRC plate simply supported at all edges with $a \times b \times h$ = 0.45m×0.45m×0.045m is studied, with the total number of layers as N_L = 10. The w.f. of GPLs in the whole plate is g_{GPL} = 1%. The average length, width and thickness of GPLs is $l_{GPL} \times w_{GPL} \times h_{GPL}$ = 2.5μm×1.5μm×1.5nm, and the mass density, Young's modulus and Poisson's ratio of GPL or polymer matrix are ρ_{GPL} = 1.06g / cm3, ρ_M = 1.2g / cm3, E_{GPL} = 1.01TPa, E_M = 3GPa, v_{GPL} = 0.186 and v_M = 0.34, respectively, based on which the midpoints of Young's modulus of each layer $E_C^{(k)C}$ can be obtained for any pattern of GPLs according to modified Halpin-Tsai model. The uncertainty degree of Young's modulus is set as $\Delta\alpha_E = 0.1$ for any layer.

In order to verify the results of eigenvalue analysis in this work, the code written based on the above theory is used to compute the first order dimensionless natural frequency ($\varpi = \omega h\sqrt{\rho_M / E_M}$) of the plate with deterministic material parameters and compare the results with those from Song et al. (2017), as demonstrated in Table I. It can be seen that this study produces identical results with the literature.

Table I. First order dimensionless natural frequency

	Pattern 1	Pattern 2	Pattern 3	Pattern 4
This work	0.1216	0.1020	0.1378	0.1118
Song et al. (2017)	0.1216	0.1020	0.1378	0.1118

The signs of sensitivities of first five orders of eigenvalues with respect to the interval Young's modulus of each layer are shown in Table II. It is indicated that the sensitivities of eigenvalue to Young's moduli of all layers are positive, so the eigenvalue becomes larger when Yong's modulus of any layer increases.

Table II. Sign of sensitivities of eigenvalues

λ_j	$E_C^{(1)}$	$E_C^{(2)}$	$E_C^{(3)}$	$E_C^{(4)}$	$E_C^{(5)}$	$E_C^{(6)}$	$E_C^{(7)}$	$E_C^{(8)}$	$E_C^{(9)}$	$E_C^{(10)}$
j = 1	+	+	+	+	+	+	+	+	+	+
j = 2	+	+	+	+	+	+	+	+	+	+
j = 3	+	+	+	+	+	+	+	+	+	+
j = 4	+	+	+	+	+	+	+	+	+	+
j = 5	+	+	+	+	+	+	+	+	+	+

Based on the sensitivities, bounds of natural frequencies of any order can be calculated according to Eqs. (7)-(10). The obtained bounds of first five orders of natural frequencies and the results from Monte Carlo Simulation (MCS) for Pattern 3 are given in Table III. It is shown that the proposed method generates very close results to MCS for all orders of eigenvalue.

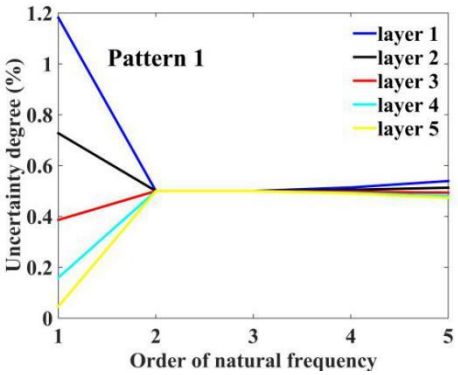
Free vibration of FG-GPLRC plate with interval parameters

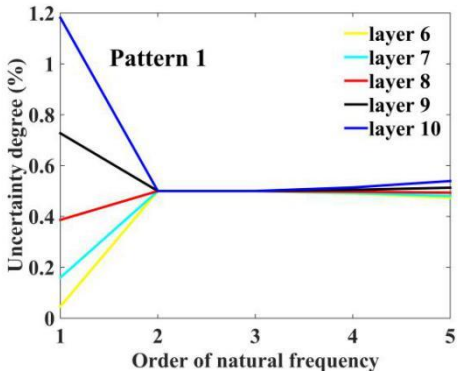
Table III. Bounds of natural frequencies

Mode	This study		MCS	
	Upper bound	Lower bound	Upper bound	Lower bound
1	0.1431	0.1324	0.1431	0.1328
2	0.5848	0.5465	0.5846	0.5469
3	1.0164	0.9498	1.0162	0.9498
4	4.2581	3.9114	4.2580	3.9116
5	4.3966	4.0377	4.3966	4.0377

Under various distribution patterns, the effects of uncertain Young's modulus at any layer on natural frequencies are shown in Figure 3. It is clear that the number of layer (i.e. the position) where uncertain Young's modulus is located obviously affects the uncertainty degree of natural frequencies.

When the weight fraction (w.f.) of GPLs is uniform among layers (Pattern 1), the position mainly affects the first order of natural frequency while generates few effects on natural frequencies of higher orders. However, when GPLs are distributed ununiformly among layers (Patterns 2-4), the position also influences higher orders of natural frequencies. Therefore, for higher orders of natural frequencies, the position of interval Young's modulus is not important but the amount of w.f. of GPLs in the uncertain layer is important. This can be proved by comparing the uncertainty degree of natural frequencies of higher orders of Pattern 2 and Pattern 3, which indicates that the uncertainty degree of natural frequency is similar when the layer with highest w.f. of GPLs is uncertain no matter which layer it is located at.





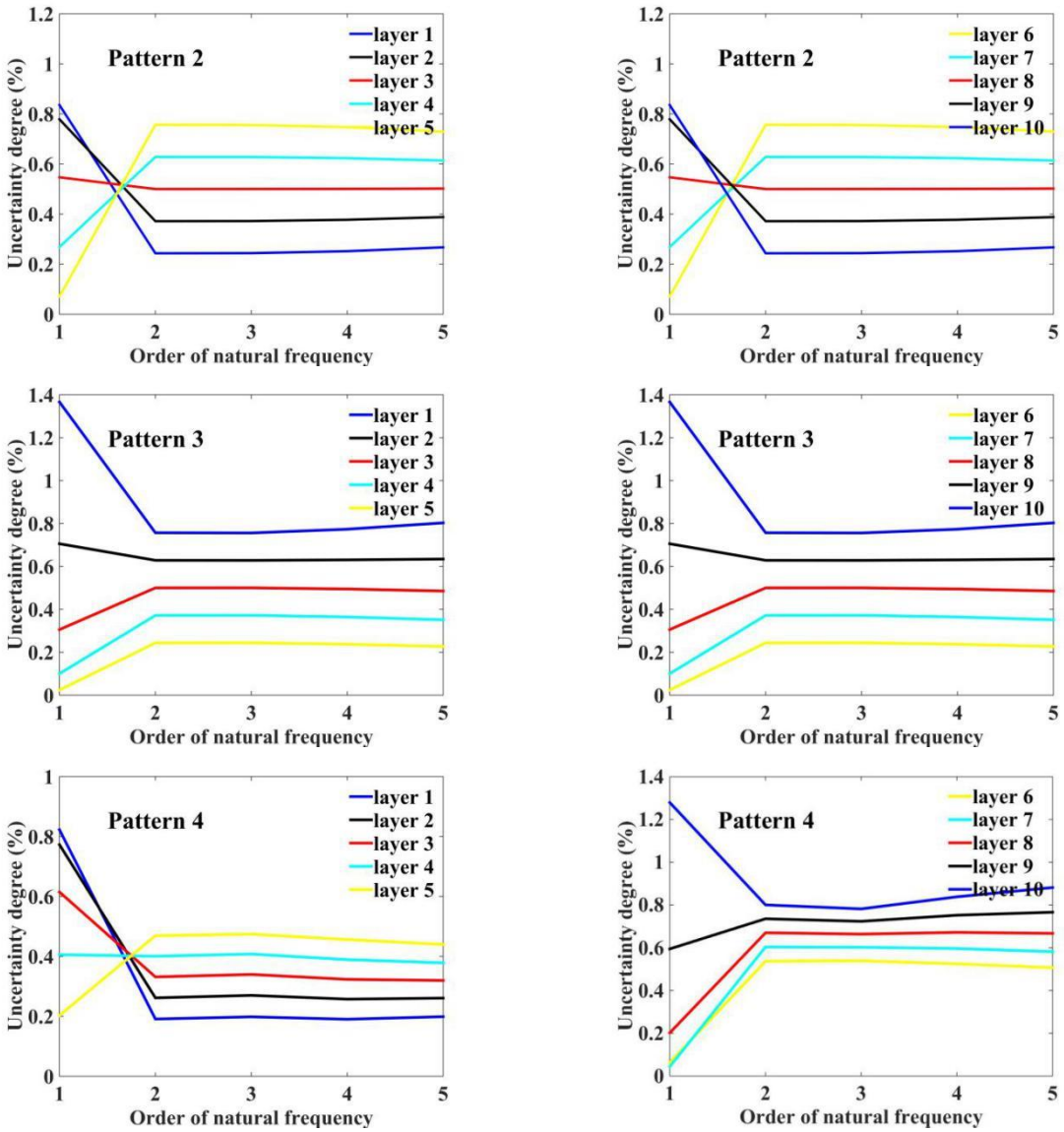


Figure 3. Uncertainty degree of natural frequency for a single uncertain layer.

Assuming that two layers of symmetrical positions in the plate have interval parameters, the uncertainty degree of natural frequency is shown in Figure 4. It can be observed that the corresponding uncertainty degree doubles for all orders of natural frequencies comparing with the case with single interval layer, so the effect of w.f. of GPLs on the uncertainty degree of natural frequency can be linearly superimposed for the layer with the same weight fraction of GPLs. For Pattern 4, although the curves seem different from those with single uncertain layer, they are also the superimposition of results of individual layers.

When more layers are uncertain simultaneously, the curve becomes flat gradually owing to the superimposition effect (Figure 5 and Figure 6). That is, the effects of material uncertainty on natural

frequency uncertainty accumulate even though the uncertain layers have different w.f. of GPLs. Finally, all orders of natural frequencies have the same uncertainty degree. Therefore, for controlling the uncertainty degree of natural frequencies, the w.f. of GPLs in the whole plate should be decreased.

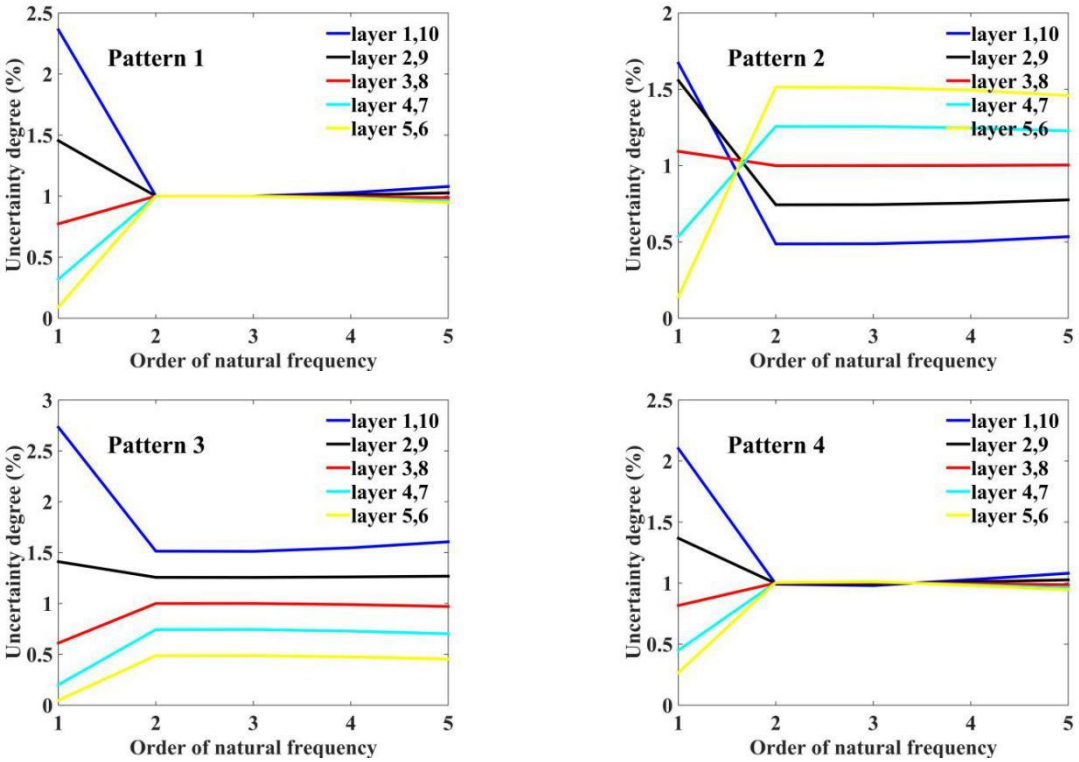
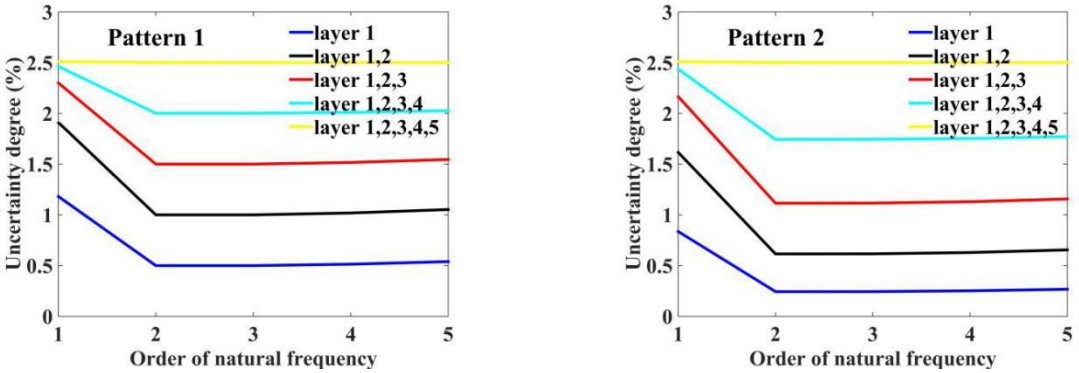


Figure 4. Uncertainty degree of natural frequency under two symmetrical interval layers.



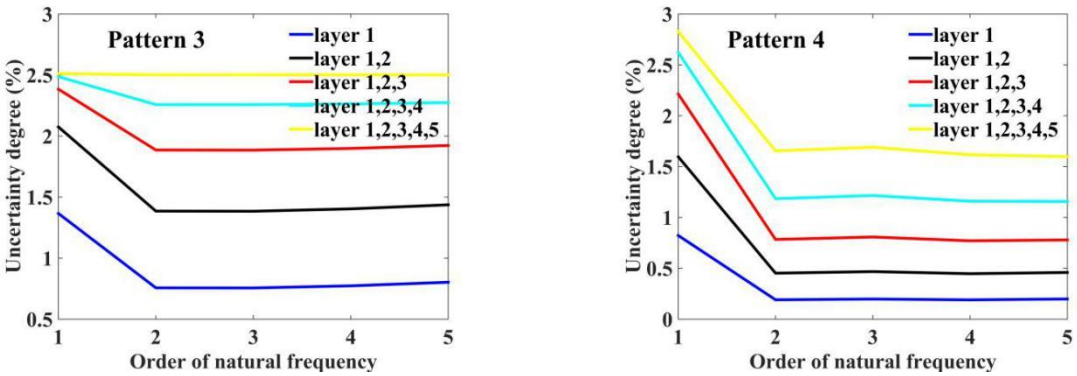


Figure 5. Uncertainty degree of natural frequency under increasing number of interval layers.

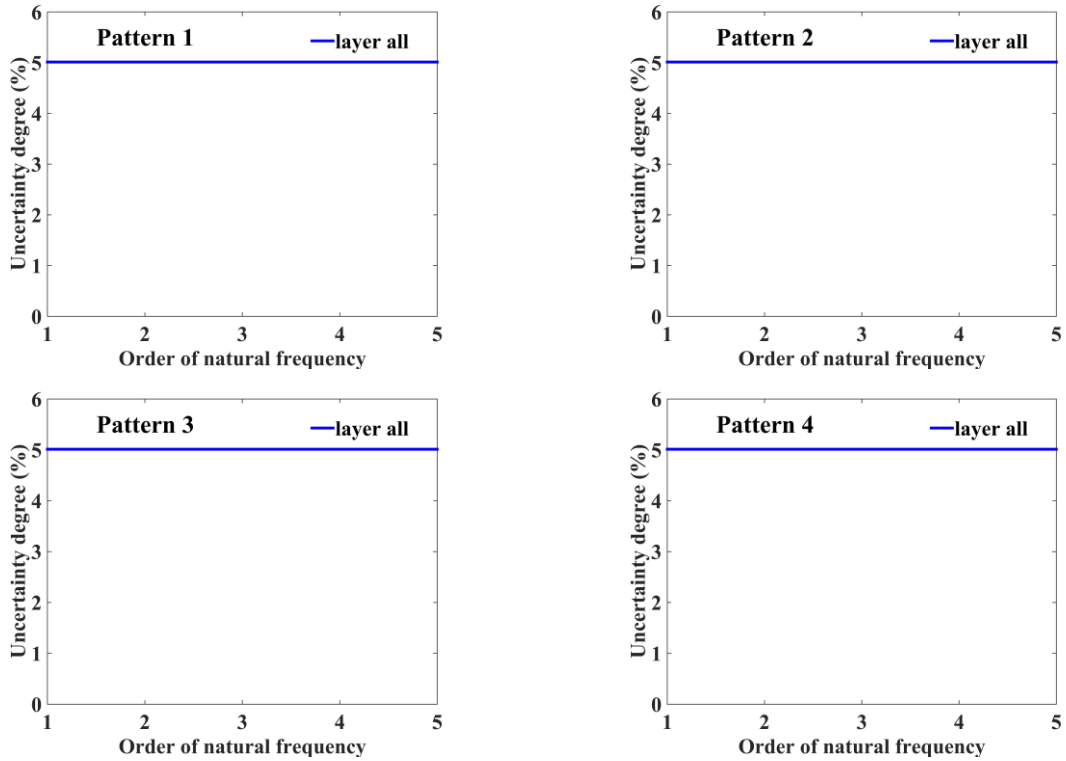


Figure 6. Uncertainty degree of natural frequency when all layers are uncertain.

4. Conclusions

Owing to the difficult fabrication process of functionally graded (FG) materials, uncertainty of material properties exists inevitably. This study deals with the free vibration of an functionally graded graphene-platelets (GPLs) reinforced composite (FG-GPLRC) plate with interval uncertain Young's modulus at each

layer of the FG plate. The investigations show that the position of uncertain layer apparently affects the uncertainty degree of natural frequencies. When the weight fraction (w.f.) of GPLs is uniform among layers, the position mainly affects the first order of natural frequency while has few effects on natural frequencies of higher orders. When GPLs are distributed ununiformly among layers, the position also influences higher orders of natural frequencies. The effect of w.f. of GPLs on the uncertainty degree of natural frequency can be superimposed. From the respect of reducing the deviation of natural frequencies, the w.f. of GPLs in the whole plate should be decreased.

Acknowledgements

The research reported was financially supported by the Technology Planning Project of Guangzhou City in China (No. 201904010307).

References

Chiker, Y., M. Bachene, M. Guemana, B. Attaf, and S. Rechak. Free vibration analysis of multilayer functionally graded polymer nanocomposite plates reinforced with nonlinearly distributed carbon-based nanofillers using a layer-wise formulation model. *Aerospace Science and Technology*, 104: 105913, 2020.

Feng, C., S. Kitipornchai, and J. Yang. Nonlinear bending of polymer nanocomposite beams reinforced with non-uniformly distributed graphene platelets (GPLs). *Composites Part B-Engineering*, 110: 132–140, 2017.

Guzmán de Villoria, R. and A. Miravete. Mechanical model to evaluate the effect of the dispersion in nanocomposites. *Acta Materialia*, 55(9): 3025–3031, 2007.

Majidi-Mozafari, K., R. Bahaadini, and A. R. Saidi. An analytical solution for vibration analysis of sandwich plates reinforced with graphene nanoplatelets. *Engineering with Computers*, DOI: 10.1007/s00366-020-01183-9, 2020.

Nguyen, Q. H., L. B. Nguyen, H. B. Nguyen, and H. Nguyen-Xuan. A three-variable high order shear deformation theory for isogeometric free vibration, buckling and instability analysis of FG porous plates reinforced by graphene platelets. *Composite Structures*, 245: 112321, 2020.

Noroozi, A. R., P. Malekzadeh, R. Dimitri, and F. Tornabene. Meshfree radial point interpolation method for the vibration and buckling analysis of FG-GPLRC perforated plates under an in-plane loading. *Engineering Structures*, 221: 111000, 2020.

Sofi, A., G. Muscolino, and I. Elishakoff. Natural frequencies of structures with interval parameters. *Journal of Sound and Vibration*, 347: 79–95, 2015.

Song, M., S. Kitipornchai, and J. Yang. Free and forced vibrations of functionally graded polymer composite plates reinforced with graphene nanoplatelets. *Composite Structures*, 59: 579–588, 2017.

Suresh, S. and A. Mortensen. Fundamentals of functionally graded materials. Institute of Materials Communications Ltd., London, 1998.

Thai, C. H. and P. Phung-Van. A meshfree approach using naturally stabilized nodal integration for multilayer FG GPLRC complicated plate structures. *Engineering Analysis with Boundary Elements*, 117: 346–358, 2020.

Yang, J., H. Wu, and S. Kitipornchai. Buckling and postbuckling of functionally graded multilayer graphene platelet-reinforced composite beams. *Composite Structures*, 161: 111–118, 2017.

Zhang, G., C. Xiao, A. Rahimi, and M. Safarpour. Thermal and mechanical buckling and vibration analysis of FG-GPLRC annular plate using higher order shear deformation theory and generalized differential quadrature method. *International Journal of Applied Mechanics*, 12(2): 2050019, 2020.

Zhao, S., Z. Zhao, Z. Yang, L. Ke, and J. Yang. Functionally graded graphene reinforced composite structures: a review. *Engineering Structures*, 210: 110339, 2020.

Hysteretic Tuned Mass Dampers for Seismic Protection

A.Boccamazzo, B. Carboni, G. Quaranta, W. Lacarbonara
Department of Structural and Geotechnical Engineering,
Sapienza University of Rome, Rome 00184, Italy,
antonio.boccamazzo@uniroma1.it, biagio.carboni@uniroma1.it,
giuseppe.quaranta@uniroma1.it, walter.lacarbonara@uniroma1.it

February 7, 2021

Abstract. The main shortcoming of linear Tuned Mass Dampers employed as anti-seismic device, is related to the loss of tuning, also called *detuning*, due to changes in the structural stiffness that may occur after the onset of damage. To take into account this detrimental effect, in this study the main system is represented through a hysteretic nonlinear constitutive model. An identification procedure to fast identify the dynamic behavior of a multistory steel-scale structure is presented. Moreover, in the present work an absorber that exhibits a hysteretic restoring force with pinching is considered. The goal is to tune the nonlinear behavior of the TMD with the nonlinear behavior of the structure to control to make the TMD effective for seismic applications overcoming the detuning effect.

Keywords: Hysteretic TMD, Nonlinear identification, TMD optimization

1. Introduction

Tuned Mass Dampers (TMDs) are recognized as effective device to mitigate wind-induced vibrations in civil structures whereas their seismic protection capability still represents an open issue (Elias and Matsagar, 2017). The main shortcoming is related to the loss of tuning, or *detuning*, due to changes in the structural stiffness that may occur after the onset of damage (Wong and Harris, 2012). To take into account this detrimental effect, the main system is represented through a hysteretic constitutive model (Boccamazzo et al., 2018).

Moreover, since real structures exhibit nonlinear behavior even at relatively low amplitudes and linear TMDs can suffer the drawback of detuning, in the present work an absorber that exhibits a hysteretic restoring force with pinching is considered (Carboni et al., 2015; Carboni and Lacarbonara, 2016). In order to make the TMD effective for seismic applications, the objective of the present study is to tune the nonlinear behavior of the TMD with the nonlinear behavior of the structure.

With the goal of considering realistic structures, the main system parameters are evaluated considering the experimental data previously collected for a five-story scale building model (Carboni and Lacarbonara, 2016). Since the structure shows a nonlinear dynamic behavior (specifically, softening backbone), it is well described by the Bouc-Wen model. After an extensive discussion about the dynamic properties of this hysteretic model, a new analytic identification procedure to fast identify the dynamic behavior of the structure is presented. This identification is compared with

© 2021 *by A. Boccamazzo, B. Carboni, G. Quaranta, W. Lacarbonara.*

the results obtained by a second numerical identification that exploits the Differential Evolutionary strategy (Boccamazzo et al., 2020).

Later, a new optimization procedure for the TMD constitutive parameters search, proposed in (Boccamazzo et al., 2020), is implemented. Finally, the vibration mitigation performances of the optimized hysteretic TMDs mounted on the nonlinear main system are carefully evaluated.

2. The Bouc-Wen model of hysteresis

The Bouc–Wen (BW) is one of the most popular model used to describe non-linear hysteretic systems. It was introduced by Bouc (Bouc, 1967) and later extended by Wen (Wen, 1976) who showed its versatility in producing a variety of hysteretic patterns.

The overall restoring force of the BW model is:

$$f(x) = k_e\, x + z(x). \tag{1}$$

It is the sum of a linear elastic force (denoted by $k_e x$, where x is the displacement and k_e the elastic stiffness) and a hysteretic force (denoted by z) whose evolution in time is given by the following first-order differential equation:

$$\dot{z} = \{k_d - [\gamma + \beta\,\mathrm{sgn}(\dot{x}z)]|z|^n\}\dot{x}, \tag{2}$$

where the overdot denotes differentiation with respect to time, the parameters γ and β are related to the hysteresis shape, n rules the transition between the elastic and post-elastic phase. The stiffness k_d is a fraction of the initial stiffness at the origin $k_0 = k_e + k_d$. The tangent stiffness (denoted by z_x) of the hysteretic force is:

$$\frac{dz}{dx} = z_x = [k_d - (\gamma + \beta\,\mathrm{sgn}(\dot{x}z)]|z|^n. \tag{3}$$

It is possible to devise three different nonlinear behaviors during loading ($\dot{x} > 0$): softening hysteresis (stiffness degrading) if $\gamma + \beta > 0$, quasilinear hysteresis (constant stiffness) if $\gamma + \beta = 0$, and hardening hysteresis (stiffness increasing) if $\gamma + \beta < 0$. A collection of these types of hysteresis loops are shown in Fig. 1.

Depending on the signs of z ad $\dot{x}$, the hysteretic cycles can be divided into four branches; the different branches are shown in Fig. 2, where the characteristic displacements, namely the transition points that divide the four branches, are also indicated. In particular $x_0 = -x_2$ are the maximum displacements reached, and $x_1 = -x_3$ are the points where the hysteretic force vanishes. Each of them admits the general solution reported in Tab. I. The terms b_1, b_2, b_3, b_4 are constants of integration that can be found imposing the initial conditions.

The hysteretic force reaches the maximum when the z_x tends to zero. Consequently, imposing the stationarity of z, the upper and the lower bounds are given by

$$z_{max} = -z_{min} = \frac{k_d}{\gamma + \beta}. \tag{4}$$

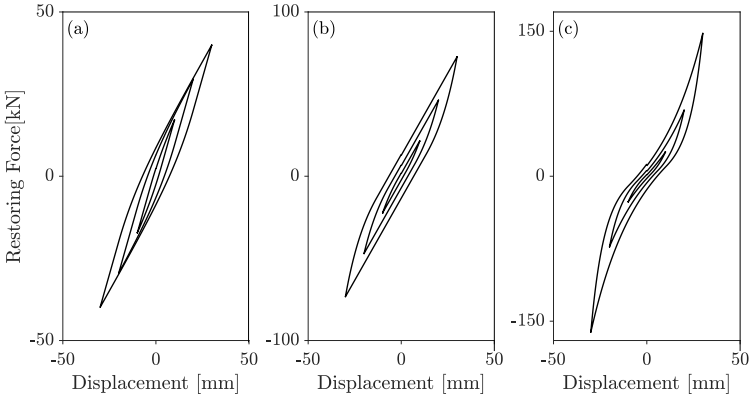


Figure 1. Loops of the total restoring force for (a) softening hysteresis ($\gamma = \beta = 0.05$ mm^{-1}), (b) quasilinear hysteresis ($-\gamma = \beta = 0.05$ mm^{-1}), (c) hardening hysteresis ($\gamma = -0.08$ mm^{-1} and $\beta = 0.05$ mm^{-1}), described by the BW model when $k_e = k_d = 1$ kN/mm and $n = 1$.

Table I. BW model solution for each branch.

Branch	z	$\dot{x}$	z_x	z
1	> 0	< 0	$k_d + (\gamma - \beta)z$	$\frac{k_d}{\gamma-\beta} + b_1 e^{x(\beta-\gamma)}$
2	< 0	< 0	$k_d - (\gamma + \beta)z$	$-\frac{k_d}{\gamma+\beta} + b_2 e^{x(\beta+\gamma)}$
3	> 0	< 0	$k_d - (\gamma - \beta)z$	$\frac{k_d}{-\gamma+\beta} + b_3 e^{x(-\beta+\gamma)}$
4	> 0	> 0	$k_d + (\gamma + \beta)z$	$\frac{k_d}{\gamma+\beta} + b_4 e^{x(-\beta-\gamma)}$

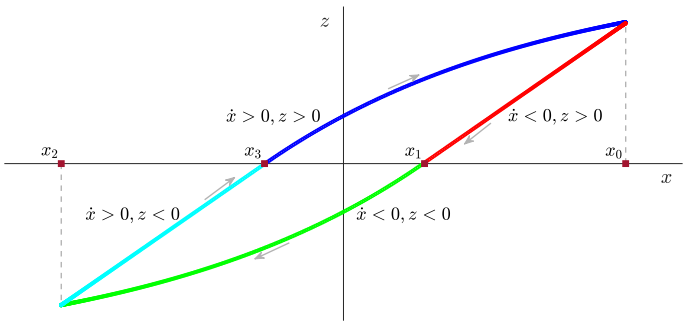


Figure 2. Typical force-displacement hysteretic cycles. The four branches are highlighted by the red, green, cyan and blue line and the transition points are described by the characteristic displacement x_0, x_1, x_2,and x_3,

Consider the closed form solution of the loading branch of the BW model (branch 4 in Tab. I). Imposing the initial condition $z(0) = 0$ is possible to find the constant $b_4 = -\frac{k_d}{\gamma+\beta}$ and, hence, the

particular solution of the first loading (z_{fl}) branch reads

$$z_{fl}(x) = \frac{(1 - e^{-x(\gamma+\beta)})k_d}{\gamma + \beta}. \tag{5}$$

This solution is a monotonically increasing exponential function and the value of z_{max} represents an horizontal asymptote such that $\lim_{x\to\infty} z_{fl}(x) = z_{max}$. For this reason it makes sense to define a saturation displacement for which the hysteretic force is close to the asymptotic value. This characteristic displacement, defined as x_{95}, is found imposing that the $z(x_{95}) = 0.95\, z_{max}$. Its meaning is shown in the Fig. 3. After few simple steps it can be stated that

$$x_{95} = \frac{-ln(0.05)}{\gamma + \beta} \simeq \frac{3}{\gamma + \beta}. \tag{6}$$

Therefore, the stiffness corresponding to this point is:

$$k_{d95} = \frac{0.95\, z_{max}}{x_{95}} = \frac{0.95\, k_d}{-ln(0.05)} \simeq 0.317\, k_d. \tag{7}$$

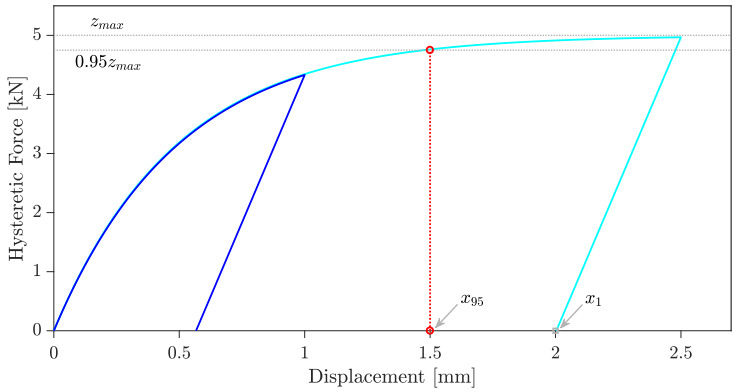


Figure 3. The meaning of x_{95} for the BW model, when $k_d = 10$ kN/mm, $\gamma = \beta = 1$ mm^{-1} and $n = 1$.

Consider a single-degree-of-freedom (SDOF) system composed by a mass m connected to the ground through a nonlinear spring that exerts the restoring hysteretic force $f(x)$ (function of the displacement x) described by the BW model. The equation of motion of the SDOF system reads:

$$m\ddot{x} + f(x) = -m\ddot{x}_g, \tag{8}$$

$$f(x) = k_e x + z(x), \tag{9}$$

$$\dot{z} = \{k_d - [\gamma + \beta\, \mathrm{sgn}(\dot{x}z)]|z|^n\}\dot{x}, \tag{10}$$

where $\ddot{x}_g$ is the base acceleration. In Fig. 4 the Frequency Response Curves (FRCs) for a hysteretic oscillator with $m = 1$ kg, $k_e = 10$ N/m, $k_d = 30$ N/m, $\gamma = \beta = 0.5$ mm^{-1} , and $n = 1$ are shown.

Note that the system is characterized by an initial (f_{in}) and an ultimate (f_{ult}) resonance frequencies for small and large oscillation amplitude, respectively, determined as follows

$$f_{ul} = \frac{1}{2\pi}\sqrt{\frac{k_e}{m}} = 1.00\,\text{Hz} \tag{11}$$

$$f_{in} = \frac{1}{2\pi}\sqrt{\frac{k_e+k_d}{m}} = 0.50\,\text{Hz} \tag{12}$$

The relationship between natural frequency and amplitude is commonly known as Backbone Curve and it is represented by the red dashed line in Fig. 4. It should be noted that the backbone passes through the point highlighted in blue (x_{95}, f_{95}), where

$$f_{95} = \frac{1}{2\pi}\sqrt{\frac{k_e+0.317k_d}{m}} = 0.70\,\text{Hz} \tag{13}$$

$$x_{95} = \frac{3}{\gamma+\beta} = 3.00\,\text{mm} \tag{14}$$

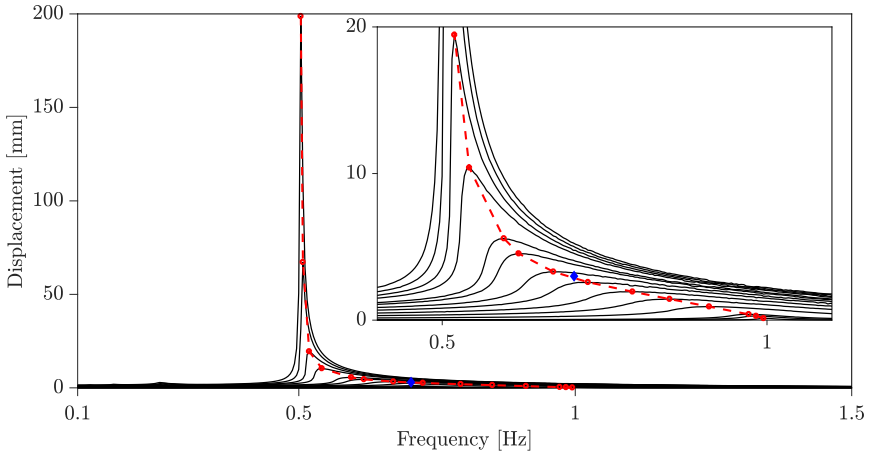


Figure 4. FRCs (in black), Backbone Curves (in red) and the point (f_{95}, x_{95}) (in blue) for the Hysteretic Oscillator, charcterized by the constitutive parameters: $m = 1$ kg, $k_e = 10$, N/m, $k_d = 30$ N/m, $\gamma = \beta = 0.5$ mm^{-1} , and $n = 1$.

The damping ratio is computed by exploiting the equivalence with the dissipated energy of the linear viscoelastic oscillator whose restoring force is $f_l(x) = c\dot{x} + k_e x$ where c is the linear damping coefficient. The damping ratio is defined as

$$\xi_0 = \frac{c}{c_{cr}}. \tag{15}$$

Substituting in Eq. (15) the critical damping coefficient $c_{cr} = 2\omega_0 m$, where $\omega_0 = \sqrt{k/m}$ is natural pulsation (in radians), it holds

$$\xi_0 = \frac{c\,\omega_0}{2\,k_e} \tag{16}$$

The elastic energy is

$$\mathcal{E}_E = \frac{1}{2}\,k_e\,x_0^2 \tag{17}$$

while the dissipated energy reads

$$\mathcal{E}_D = \pi\, c\, \overline{\omega}\, x_0^2 \tag{18}$$

where $\overline{\omega}$ is excitation pulsation. Making explicit the terms k and c from Eqs. (17) and (18), respectively, the damping ratio can be written as:

$$\xi_0 = \frac{\mathcal{E}_D}{4\,\pi\,\mathcal{E}_E}\frac{\omega_0}{\overline{\omega}} \tag{19}$$

The pulsation ratio, namely the ratio between the natural pulsation of the system and that of the external excitation, is usually set to 1, so that the damping ratio is computed as

$$\xi_0 = \frac{\mathcal{E}_D}{4\,\pi\,\mathcal{E}_E}. \tag{20}$$

The damping ratio and the dissipated end elastic energies are shown in Figs. 5-(a) and 5-(b), respectively.

The dissipated energy of the Hysteretic Oscillator is calculated as the area enclosed by the force-displacement curve during a sinusoidal displacement cycle (with maximum displacement x_0), so that

$$\mathcal{E}_D = \int_{-x_0}^{x_0} z\, dx. \tag{21}$$

For very large displacement, as shown in Fig. 5-(a), this dissipated area can be approximated as

$$\mathcal{E}_D \simeq 4\, z_{max}\, x_0. \tag{22}$$

Substituting Eqs. (4) and (22), the damping ratio ξ_0 can be expressed as funtion of the displacement x

$$\overline{\xi}_0(x) = \frac{2\,k_d}{k_e \pi(\gamma+\beta)\,x}. \tag{23}$$

In a novel variant of the BW model proposed in (Carboni et al., 2015) a function $H(x)$ is introduced in order to provide a pinching along the force-displacement cycle. This is described by the following exponential function:

$$H(x) = 1 - \xi_H e^{-x^2/x_H} \tag{24}$$

where $\xi_H \in [0,1)$ and $x_H > 0$. Specifically, the function $H(x)$ modulates the tangent stiffness k_d of the BW model in the neighborhood of the origin. Therefore, the evolution in time of the hysteretic force is described as follows:

$$\dot{z} = \{k_d H(x) - [\gamma + \beta\,\mathrm{sgn}(\dot{x}z)]|z|^n\}\dot{x}. \tag{25}$$

The pinching function $h(x)$ depends on two parameters, namely, ξ_h and x_H. The first one governs the intensity of the pinching effect, while the latter regulates the width of the pinched region of the hysteresis loop. Specifically, low values of x_H entail a localized pinching around the origin of the force-displacement cycles, whereas high values determine the pinching presence over a larger

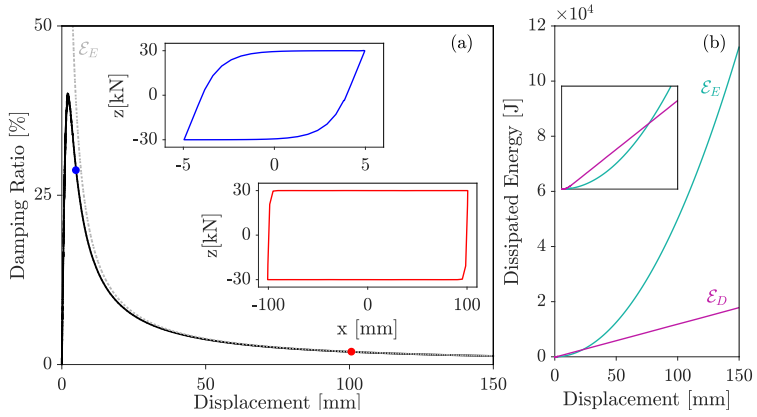


Figure 5. Damping ratio (black solid line), Eq. (23) (dashed grey line), and force-displacement cycles (blue and red solid lines) in (a). Elastic and Dissipated Energies in (b).

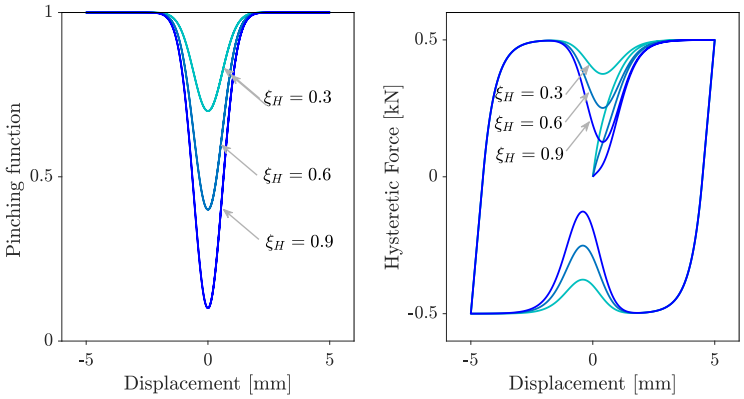


Figure 6. Effect of the parameter ξ_H on the pinching function $H(x)$ (left) and on the hysteretic force $z(x)$ (right). In all plots $k_d = 1$ kN/mm, $\gamma = \beta = 1$ mm^{-1}, $x_H = 0.7$ mm^2, $n = 1$ while ξ_H is set to 0.3, 0.6 and 0.9.

displacement range. Obviously, the modified BW model can be consider a generalization of the BW model, since the modified model returns the classical model when ξ_H is null. The effects of the parameter ξ_H and x_H on the hysteretic force $z(x)$ and pinching function $H(x)$ are shown in Figs. 6 and 7.

3. Mechanical model and equations of motion

The main structure and the TMD are modeled as a SDOF system: specifically, the first one by employing the BW model, while the latter, by means of the pinched variant previously described. Indeed, the the overall elastoplastic behavior of the steel structure is well described by means of the classical BW model of hysteresis. The pinched BW model is able to simulate a large variety of nonlinear behaviors within the context of novel absorbers, instead. Particularly, it intends to

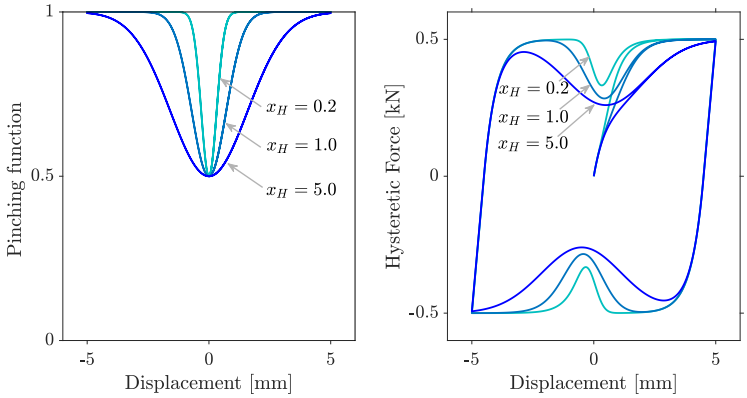


Figure 7. Effect of the parameter x_H on the pinching function $H(x)$ (left) and on the hysteretic force $z(x)$ (right). In all plots $k_d = 1$ kN/mm, $\gamma = \beta = 1$ mm^{-1}, $\xi_H = 0.5$, $n = 1$ while x_H is set to 0.2, 1 and 5 mm^2.

simulate here a nonlinear TMD, consisting of a single mass that dissipates hysteretic energy through suitable assemblies of NiTiNOL and steel wire ropes, that has been characterized experimentally and numerically in the context of previous studies (Carboni et al., 2015; Carboni and Lacarbonara, 2016), in which were demonstrated that the pinching behavior is due to the coupling of frictional dissipation and phase transformations in SMA material. As a consequence, the protected structure turns out to be a nonlinear two-degree-of-freedom system (2DOF).

Its schematic is given in Fig. 8. The structure displacement is indicated as x_1; x_2 is the dispacement relative to the main structure, x_g is the ground displacement: so that, $p_1 = x_g + x_1$ and $p_2 = x_g + x_1 + x_2$ are the absolute dispacement of structure and TMD, respectively. The terms $f_1(x_1) = k_{e1}x_1 + z_1(x_1)$ and $f_2(x_2) = k_{e2}x_2 + z_2(x_2) + k_{c2}x_2^3$ are the overall nonlinear restoring forces of structure and TMD, respectively. Note that the cubic elastic term $k_{c2}x_2^3$ is also introduced to take into account the effect of a hardening behavior in the absorber restoring force.

The equations of the motion of the 2DOF system under seismic base acceleration are given hereafter, where the time dependence will be tacitly omitted for the ease of notation.

$$m_1\ddot{x}_1 + k_{e_1}x_1 + c_1\dot{x}_1 + z_1 - k_{e_2}x_2 - k_{c2}x_2^3 - z_2 = -m_1\ddot{x}_g, \tag{26}$$

$$m_2\ddot{x}_2 + m_2\ddot{x}_1 + k_{e_2}x_2 + k_{c2}x_2^3 + z_2 = -m_2\ddot{x}_g, \tag{27}$$

$$\dot{z}_1 = \{k_{d1} - [\gamma_1 + \beta_1 \operatorname{sgn}(\dot{x}_1 z_1)]|z_1|^{n_1}\}\dot{x}_1, \tag{28}$$

$$\dot{z}_2 = \{k_{d2}H(x_2) - [\gamma_2 + \beta_2 \operatorname{sgn}(\dot{x}_2 z_2)]|z_2|^{n_2}\}\dot{x}_2. \tag{29}$$

Hereafter, the subscript 1 indicates the quantities that refer to the main structure, while the subscript 2 indicates the quantities that refer to the absorber.

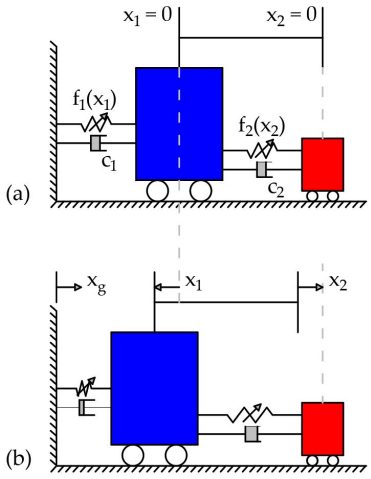


Figure 8. Schematic of two-DOF model; undeformed (a) and deformed (b) configuration.

4. Analytical and numerical nonlinear structural identifications

The steel multistory structure analysed in (Carboni and Lacarbonara, 2016) is here considered. The experimental data (the blue line in Fig. 9) show a softening behavior characterized by a decrease of the main resonance frequency and simulated through the BW oscillator.

The parameters to be identified are collected in the vector $\boldsymbol{\theta} = \{k_{e_1}, k_{d_1}, \gamma_1, \beta_1\}$ while n_1 is set to 1. The weight of the main structural mass is estimated to be 565.82 kg. Two strategies of identification are proposed.

The first strategy is analytic and exploits the relationship previously introduced. The parameters $\overline{k}_{e1}$, $\overline{k}_{d1}$, $\overline{\gamma}_1$ and $\overline{\beta}_1$ are determined as follows.

$$\overline{f}_{ul} = \frac{1}{2}\sqrt{\frac{\overline{k}_{e1}}{m_1}} = 4.23\,\text{Hz} \rightarrow \overline{k}_{e1} = 0.3610\,\text{kN/mm} \tag{30}$$

$$\overline{f}_{in} = \frac{1}{2}\sqrt{\frac{\overline{k}_{e1}+\overline{k}_{d1}}{m_1}} = 4.02\,\text{Hz} \rightarrow \overline{k}_{d1} = 0.0387\,\text{kN/mm} \tag{31}$$

$$\overline{f}_{95} = \frac{1}{2}\sqrt{\frac{\overline{k}_{e1}+0.317\,\overline{k}_{ed}}{m_1}} = 4.0875\,\text{Hz} \rightarrow \overline{x}_{95} \cong 10\,\text{mm} \tag{32}$$

$$\overline{x}_{95} = \frac{3}{\overline{\gamma}_1+\overline{\beta}_1} \rightarrow \overline{\gamma}_1 + \overline{\beta}_1 = 0.30\,\text{mm}^{-1} \tag{33}$$

and setting $\overline{\gamma}_1$ equal to $\overline{\beta}_1$ it results $\overline{\gamma}_1 = 0.15$ mm $^{-1}$.

The second strategy is numerical and exploits the Differential Evolution (DE) algorithm (Storn and Price, 1997). The DE algorithm is implemented to find the minimum of the following objective function:

$$\Phi_{id}(\boldsymbol{\theta}) = \sum_{i=1}^{M} w_f \frac{(f_{ex,i} - f_{m,i}(\boldsymbol{\theta}))}{f_{ex,i}} + w_\xi \frac{(\xi_{ex,i} - \xi_{m,i}(\boldsymbol{\theta}))}{\xi_{ex,i}}. \tag{34}$$

For the ith base excitation amplitude, this objective function measures the difference between the experimental frequencies and damping ratios ($f_{ex,i}$ and $\xi_{ex,i}$, respectively) and those numerically estimated ($f_{m,i}$ and $\xi_{m,i}$, respectively). The weight functions w_f and w_ξ can promote the agreement in frequency or damping identification. The BW damping is estimated by means of Eq. (20).

In order to save time, a new identification procedure is, here, proposed. Indeed, the standard identification strategy needs the evaluation of the entire FRC, for different oscillation amplitude, to get its peak so as to build the backbone. Obviously, the FRC evaluation requires the stationary response of the SDOF in several points. In the proposed method a harmonic displacement is imposed and the overall restoring force is evaluated. In this way is possible to build the force-displacement cycles. The resonance frequency is estimated from the average stiffness along the hysteresis cycles as follows:

$$k(x) = \frac{\int_{\mathcal{C}} k_e + z_x(x)dx}{\int_{\mathcal{C}} dx}, \tag{35}$$

where $\mathcal{C}$ is the displacement path along the loading and unloading branches. The proposed method is very fast because it requires only one stationary response of the SDOF for each excitation level. It is worth noting another great advantage of proposed method: since the displacement is imposed, the oscillation amplitude at the resonance frequency is known.

The identified structural parameters, according to the two different approaches, are listed in table II. Several identifications are carried out by considering different weight functions.

Table II. Identified structural parameters by exploiting the analytical and numerical approaches for different weight functions.

	w_f [–]	w_d [–]	k_{e1} $\left[\frac{\text{kN}}{\text{mm}}\right]$	k_{d1} $\left[\frac{\text{kN}}{\text{mm}}\right]$	γ_1 $\left[\frac{\text{kN}^{1-n}}{\text{mm}}\right]$	β_1 $\left[\frac{\text{kN}^{1-n}}{\text{mm}}\right]$	n_1 [–]
Analytical	–	–	0.3610	0.0387	0.150	0.150	1
DE	0.7	0.3	0.356	0.058	0.312	0.076	1
DE	0.5	0.5	0.352	0.057	0.262	0.057	1
DE	0.1	0.9	0.354	0.032	0.001	0.126	1

Fig. 9 and Fig. 10 show the frequency and damping identified, respectively. In order to promote the agreement in frequency rather than that in damping, and since the experimental damping ratio is low and it does not affect much the structural behavior, from now on, let us to consider the weight coefficients w_e and w_d equal to 0.7 and 0.3, respectively.

5. TMD optimization and seismic performances

The TMD optimization strategy is formulated in the time domain and it is suitable for generally nonstationary base excitations. The DE algorithm is employed to look for the maximum of the objective function:

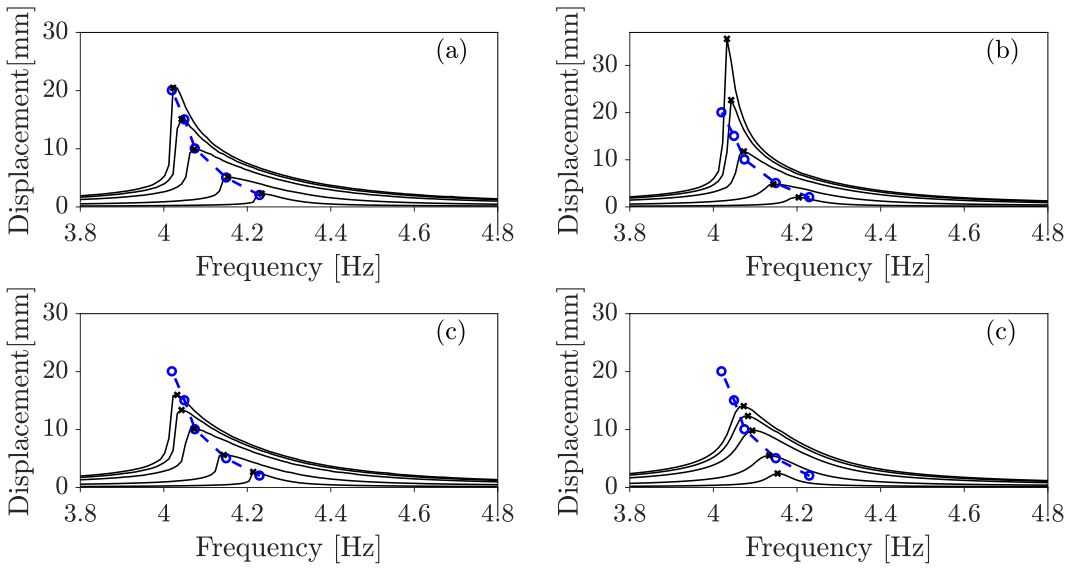


Figure 9. Displacement Frequency Response Function of the identified BW model (black line) and experimental data (blue line). Case a) $w_f = 0.7$, $w_d = 0.3$, b) analytic, c) $w_f = 0.5$, $w_d = 0.5$, d) $w_f = 0.1$, $w_d = 0.9$.

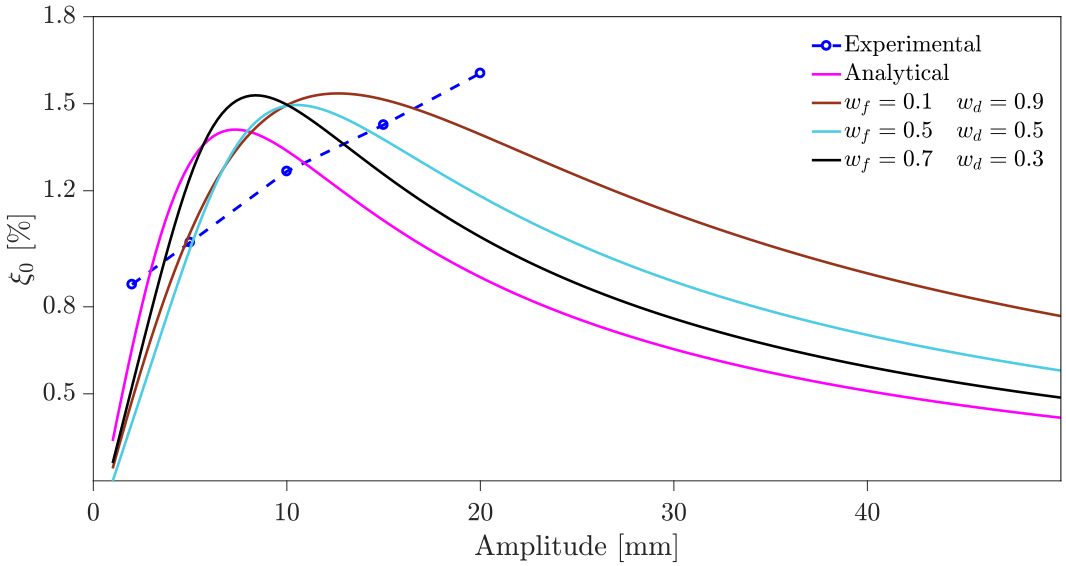


Figure 10. Damping of the identified BW model (solid line) for different weight functions and experimental data (dashed-blue line).

$$R_t(\mathbf{d}) = 1 - \frac{\sum_{r=1}^{S} \text{rms}[x_r^{con}(\mathbf{d})]}{\sum_{r=1}^{S} \text{rms}[x_r^{unc}]}. \tag{36}$$

Herein, $x_r^{con}(\mathbf{d})$ and x_r^{unc} denote the displacement of the controlled and uncontrolled system (Fig. 11) under the rth seismic acceleration time history, with $r = 1, \ldots, S$. This function represents the ratio between the sum of the controlled and uncontrolled root mean squares (rms) of the main system displacement under a set of earthquakes.

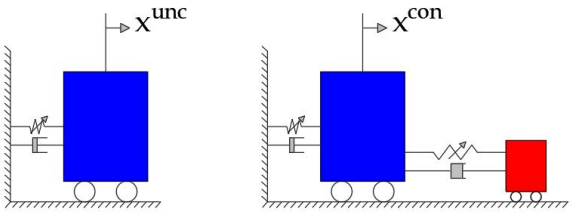


Figure 11. Uncontrolled (left) and controlled (right) scenarios.

The TMD parameters are collected in the vector $\mathbf{d}$ and their optimal are found by exploiting again the DE strategy. The design space of the TMD parameters is given in Tab. III, where the bounds are expressed as function of the structural parameters. The parameter n_2 is set equal to 1, while the upper limit for ξ_H is posed to 0.985 to avoid numerical problem.

Table III. Design space of the TMD model parameters.

	k_{e_2} $\left[\frac{\text{kN}}{\text{mm}}\right]$	k_{d_2} $\left[\frac{\text{kN}}{\text{mm}}\right]$	k_{c2} $\left[\frac{\text{kN}}{\text{mm}^3}\right]$	γ_2 $\left[\frac{\text{kN}^{1-n_1}}{\text{mm}}\right]$	β_2 $\left[\frac{\text{kN}^{1-n_1}}{\text{mm}}\right]$	ξ_H [–]	x_H $[\text{mm}^2]$
Lower bound	$\frac{k_{e_1}}{1000}$	$\frac{k_{e_1}}{1000}$	0	$\frac{\gamma_1}{1000}$	$\frac{\beta_1}{1000}$	0	20
Upper bound	$10k_{e_1}$	$30k_{e_1}$	k_{e1}	$100\gamma_1$	$100\beta_1$	0.985	40000

A set of 8 spectrum-compatible seismic acceleration time histories are generated: one half of this set is defined so that their spectra match the design spectrum for the Immediate Occupancy Limit State (SLD), while the second half the Life Safety Limit State (SLV). The design spectra are defined in in agreement with the Italian Building Code (NTC, 2018) (SLV and SLD have a probability of occurrence equal to 10% and 63% in 50 years, respectively) and the reference site is the city of Messina (Italy). The characteristic parameters for the elastic response spectra (namely the design ground acceleration, the spectral amplification factor, and the upper limit of the period of the spectral plateau) are the following: $a_g = 0.082\,g$, $F_0 = 2.318$, $T_c^* = 0.294$ s for SLD, and $a_g = 0.247\,g$, $F_0 = 2.411$, $T_c^* = 0.359$ s for SLV. Three different values for the mass ratio $\mu = m_2/m_1$ are considered, i.e. $\mu = (0.01, 0.02, 0.05)$. The convergence of the objective function

R_t, defined in Eq. (36), is shown in Fig. 12 for different mass ratios. Hysteretic TMDs endowed with mass ratios equal to 0.01, 0.02, and 0.05 attain meaningful reductions of the RMS of the main system displacement, namely, 40%, 50%, 60%, respectively.

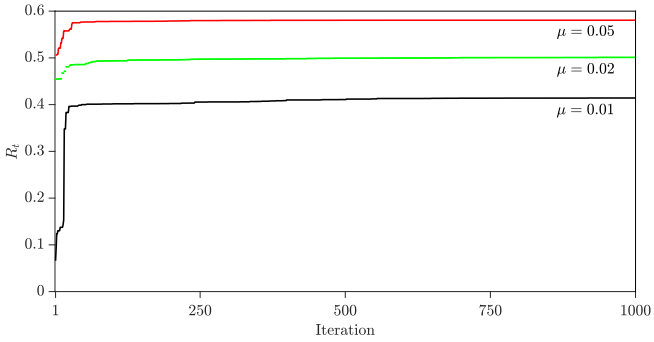


Figure 12. Convergence of the objective function R_t shown in Fig. 4. RMS displacement reductions for the controlled structure. The mass ratios are set to 0.01 (black line), 0.02 (green line), and 0.05 (red line).

After just 50 iterations, the optimizer sets to zero the k_{c2} for all the population vectors as shown in Fig. 13. If all the vector converge to a unique value, that value get fixed and it can not longer be turned in the following iterations. Obviously, this means that the optimized solution does not need the addition of the cubic elastic stiffness. Therefore, these results confirm the above conclusion: since the structure has a softening behavior, the optimized TMD has to exhibit a similar behavior characterized by softening. In this way the TMD can be tuned (remaining effective) with the system for a larger range of oscillation amplitude. The optimal TMD parameters are reported in Tab. IV.

Table IV. Optimum parameters of the TMD for the steel structure (set TMD-NS) from time domain-based optimization under nonstationary excitation.

μ	Unit	0.01	0.02	0.05
k_{e_2}	[kN/mm]	$3.31 \cdot 10^{-3}$	$6.41 \cdot 10^{-3}$	$1.41 \cdot 10^{-2}$
k_{d_2}	[kN/mm]	$9.19 \cdot 10^{-2}$	$2.35 \cdot 10^{-2}$	$6.19 \cdot 10^{-2}$
γ_2	[kN$^{1-n_2}$/mm]	$9.92 \cdot 10^{-2}$	$1.00 \cdot 10^{-3}$	$3.91 \cdot 10^{-4}$
β_2	[kN$^{1-n_2}$/mm]	$6.77 \cdot 10^{-2}$	$1.65 \cdot 10^{-1}$	$4.49 \cdot 10^{-1}$
ξ_H	[–]	0.985	0.898	0.870
x_H	[mm^2]	$4.00 \cdot 10^4$	$2.54 \cdot 10^3$	$2.00 \cdot 10^2$

Finally, the seismic effectiveness of the hysteretic TMD in protecting the identified structure is discussed by considering 3 new sets of 7 seismic acceleration time histories. The assortments are compatible with the SLD, SLV and SLC (Collapse Prevention Limit State, $a_g = 0.336\,g$, $F_0 = 2.446$,

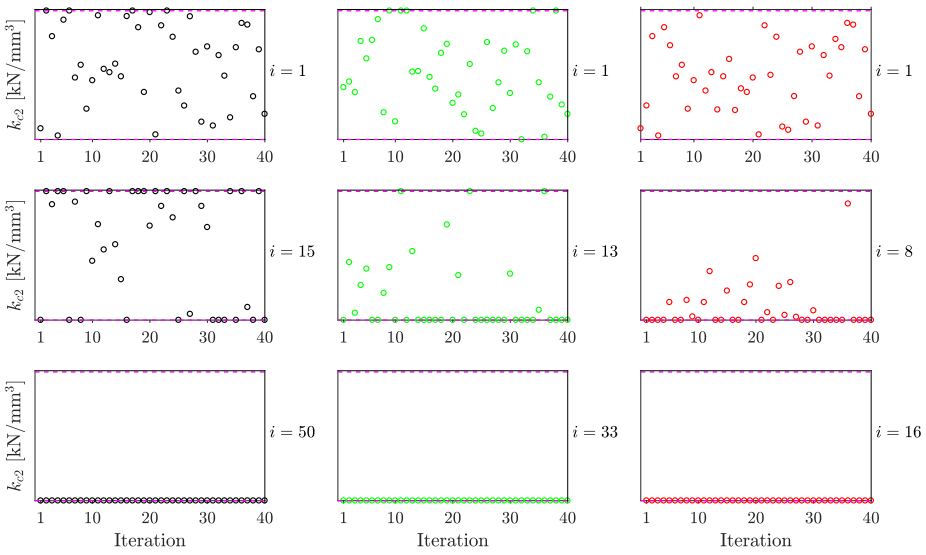


Figure 13. Cubic elastic stiffness k_{c2} for the population vectors at ith iteration. The dashed lines represent the lower and upper bounds of the search space. The mass ratios are set to 0.01 (black line), 0.02 (green line), and 0.05 (red line).

$T_c^* = 0.384$ s), respectively. The results are reported in Tab. V and clearly show that a nonlinear TMD reduces the loss of mitigation capability caused by the detuning effect.

Table V. Objective function R_t for each limit state.

μ	SLD	SLV	SLC
0.01	0.42	0.38	0.35
0.02	0.49	0.45	0.41
0.05	0.58	0.51	0.46

6. Conclusions

The present work is focused on the identification of a steel structure exhibiting a nonlinear dynamic behavior (specifically, softening backbone) and on the design of a passive hysteretic TMD for seismic protection purposes. Two identification strategies are proposed: the first strategy is analytic and exploits the properties of the BW model, while the latter is numerical and is based on the use of

DE. Numerical results have proven that the optimized hysteretic TMD has the ability to reduce the main structure vibrations under strong seismic base accelerations by exploiting its softening behavior similar to that exhibited by the main structure. Indeed, reduction of the RMS displacement close to 40% (for the SLV) were obtained with the hysteretic TMD endowed with a mass ratio equal to 1%. It has also been shown that adopting a TMD endowed with a nonlinear cubic stiffness does not improve mitigation performances.

References

Boccamazzo, A., B. Carboni, G. Quaranta, and W. Lacarbonara. Optimized Hysteretic TMD for Seismic Control of a Nonlinear Steel Structure. In *ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers Digital Collection, 2018.

Boccamazzo, A., B. Carboni, G. Quaranta, and W. Lacarbonara. Optimization Strategies of Hysteretic Tuned Mass Dampers for Seismic Control. In *Nonlinear Dynamics and Control*. Springer, Cham, 2020. 99 – 106.

Boccamazzo, A., B. Carboni, G. Quaranta, and W. Lacarbonara. Seismic effectiveness of hysteretic tuned mass dampers for inelastic structures.. In *Engineering Structures*, 216, 110591.

Bouc, R. Forced vibrations of mechanical systems with hysteresis. In *Proceedings of the Fourth Conference on Nonlinear Oscillations*, Prague, 1967.

Carboni, B., W. Lacarbonara, and F. Auricchio. Hysteresis of multiconfiguration assemblies of nitinol and steel strands: experiments and phenomenological identification. *Journal of Engineering Mechanics*, 141.3: 04014135, 2015.

Carboni, B. and W. Lacarbonara. Nonlinear vibration absorber with pinched hysteresis: theory and experiments. *Journal of Engineering Mechanics*, 142.5: 04016023, 2016.

Carboni, B. and W. Lacarbonara. Nonlinear dynamic characterization of a new hysteretic device: experiments and computations.. *Nonlinear Dynamics*, 83.1-2: 23 – 39, 2016.

Elias, S. and V. Matsagar. Research developments in vibration control of structures using passive tuned mass dampers.. *Annual Reviews in Control*, 44: 129 – 156, 2017.

NTC2018. Aggiornamento delle Norme Tecniche per le Costruzioni, Decree of the Minister of the Infrastructures, 14, 2018.

Storn, R. and K. Price. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, 11.4: 341 – 359, 1997.

Wen, Y. K. Method for random vibration of hysteretic systems. *Journal of the engineering mechanics division*, 102.2: 249 – 263, 1976.

Wong, K. K. and J.L. Harris. Seismic damage and fragility analysis of structures with tuned mass dampers based on plastic energy.. *The Structural Design of Tall and Special Buildings*, 21.4: 296 – 310, 2012.

Behavior of structures equipped with variable friction dissipative systems

Liborio Cavaleri
Department of Engineering, University of Palermo, Palermo, 90128, Italy, liborio.cavaleri@unipa.it

Mario Di Paola
Department of Engineering, University of Palermo, Palermo, 90128, Italy, mario.dipaola@unipa.it

Marco Filippo Ferrotto
Department of Engineering, University of Palermo, Palermo, 90128, Italy, marcofilippo.ferrotto@unipa.it

Vincenzo Sucato
Department of Engineering, University of Palermo, Palermo, 90128, Italy, vincenzo.sucato03@unipa.it

Abstract: Usually, in order to mitigate the stresses in framed structures, different strategies are used. Among them, the base isolation, the viscous dampers and tuned mass dumpers have been widely investigated.
In this paper a strategy of energy dissipation by friction is proposed, that is a braking system borrowed from the mechanical engineering. The constitutive law of such system is linearly dependent on the displacement of the device - therefore on the interstorey displacement of structures - and depends also on the signum function of the interstorey velocity.
The feasibility of the braking system, the amount of energy that can be dissipated and the overall performances of structures equipped with these type of systems are studied in detail in view of provide adequate levels of energy dissipation, being effective and not dangerous for structures that request limited interstorey drifts.

Keywords: Dissipative systems; variable friction dissipative device; Structural dynamics

1. Introduction

Structural vibration control involves the use of dissipation devices to reduce excessive vibration during winds or earthquakes to limit damage in the structural members. These devices can be used in the case of new constructions as well as for the retrofitting of existing buildings exploiting the effect of adding energy dissipation capacity.

The purpose of using these kind of strategies is that the structure must remain in the elastic range so that the bearing frames can be designed considering reduced seismic actions, while the task of facing the majority of the seismic loads is demanded to bearing isolators or to dissipative systems.

Bearing isolators are interposed between the foundation and the superstructure for the decoupling of the frequencies coming from the ground motion and the superstructure to achieve a longer period of vibration that makes the seismic excitation to be much lower (Ferrotto et al. 2019, 2020). Typical bearing isolators

© 2021 *by Liborio Cavaleri, Mario Di Paola, Marco Filippo Ferrotto, Vincenzo Sucato.*

are elastomeric and sliding devices. Very often these two systems are used together to improve their efficiency (Braga and Laterza, 2004). Base isolation is mainly used for new buildings since for the case of existing structures this approach is very expensive and involves significant structural modifications.

Dissipative systems are used to provide to the structures supplementary energy dissipation by means of devices that can be of different types. Among the most known there are fluid viscous dampers, yielding dissipaters, hysteretic, viscoelastic and friction dampers. In the case of existing buildings, the installation of these devices may not require excessive modifications.

Hysteretic yielding dampers are generally made of steel. They can be designed to reach the yielding strength in bending, shear or axially (Teruna et al., 2015). The yielding of the hysteretic dampers is designed to be reached before the yielding of the structural members. These systems provide also an increasing of stiffness and strength to the structure.

Visco-elastic dampers are generally formed by two layers of polymer bonded between a central driving plate and two outer plates. The damping force is linearly velocity dependent and produce an elliptical force-displacement function where the inclination of the major axis corresponds to the stiffness of the device.

Fluid viscous dampers are very efficient because of their velocity-dependent behavior. Dampers force are out of phase with the relative displacement and therefore the structural restoring forces; for this reason, the peak forces of a viscous damper does not add to the peak restoring forces; they can be designed with characteristics that do not change the stiffness of the structure (Alotta et al., 2016).

Friction dampers are special dissipative devices that act as fuses, having force-displacement behavior independent by the velocity and the frequency. The task of these devices is to limit the force on the members. Each damper has a typical Coulomb-friction law and it is manufactured to slip at a specified force, within a specified tolerance over the specified travel. The slip force is controlled by the tension in bolts holding together the sliding elements (Bhaskararao and Jangid, 2006).

In addition to the traditional supplementary dissipative systems, in the last period, a number of innovative energy dissipating devices have been studied (Cancellara et al., 2019). Dal Lago et al. (2018) presented a study of precast concrete structures with energy dissipating cladding panel connection systems with friction-based or plasticity-based devices interposed in between adjacent cladding panels or between the panels. Cook et al. (2018) proposed a dissipative system (named by the authors as Grip ‘n’ Grab GnG device) composed by a tension-only device with a ratcheting function to improve the behavior with respect to the buckling restrained braces (BRBs). In fact, the system proposed by the authors eliminates residual compressive forces and removes the need for buckling restraint. The device is designed to be used in conjunction with a dissipater, and energy dissipation in the GnG-dissipater system can be provided by various mechanisms such as yielding or friction. Bagheri et al. (2019) studied a cable bracing system comprising a pre-stressed cable and a drum interacting via frictional contact for lateral resistance of structures.

In this paper a braking system borrowed from the mechanical engineering is studied and designed as innovative method to be applied for the energy dissipation on moment resisting frame structures. The constitutive law of such system is linearly dependent on the displacement of the device and therefore on the interstorey displacement and depends on the signum function of the interstorey velocity.

The feasibility of the braking system, the amount of energy that can be dissipated and the overall performances of structures equipped with these type of systems are studied in details in view of provide adequate levels of energy dissipation, being effective and not dangerous for structures that request limited interstorey drifts.

2. Model for variable friction dissipative (VFD) device

The variable friction damper proposed in this study provides dissipative capacity as a function of two components: the first one, governed by pure friction, is constant and is a function of the signum of the velocity (typical Coulomb-friction law), while the second component depends on the relative displacement (in absolute value) and by the signum of the velocity. The latter is governed by a linear law that increases the dissipative force as the displacement increases and decreases as the displacement decreases until it reaches zero for relative zero displacements. In other words, the load-displacement response of the system is linear due to the existence of frictional contact as well as geometrical effects and it consists of two phases: a linear phase before gross slipping with a relatively specific force followed by a linear phase with gradually increasing stiffness (i.e. hardening). The mechanical model of the variable friction damper (VFD) device is shown in Fig. 1.

The device is composed by different components. In detail, two steel plates are fixed to the bottom of the beam at the upper floor, while an intermediate steel device is connected to the lower steel braces. The central device slides along the internal surfaces of the external plates that are bolted with a given prestressing induced by tightening (according to the classic friction damper devices).

The peculiarity of the novel device is that the internal surfaces of the two plates are inclined by a certain angle with respect to a horizontal sliding surface. This causes that the central device is forced to a convex funnel-shaped surface. This geometrical configuration generates additional restoring forces with the increasing of the relative outwards displacements of the central device. The tightening level of the bolts acts for the constant friction component, while the convex funnel-shaped surface provides an additional restoring force depending on the level of displacement.

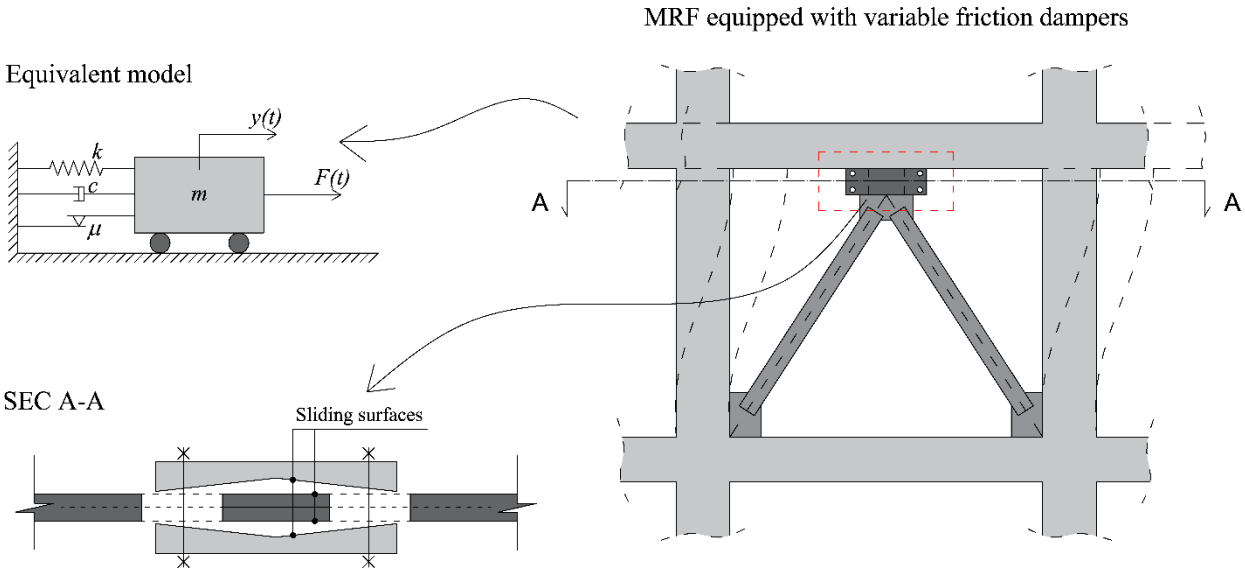


Figure 1. Mechanical model of VFD.

The governing equation of the variable friction damper (VFD) can be expressed as follow:

$$F = c_0 \cdot sign(\dot{x}) + c_1 \cdot |x| \cdot sign(\dot{x}) \tag{1}$$

where F is the dissipative force, $sign(\dot{x})$ is the signum function of the velocity, x is the relative displacement and c_0 and c_1 are constants. The first constant (c_0) depends by the pure friction like a constant brake system, while the second constant (c_1) depends on the shape and on the type of material used so it can act as a constant brake system but influenced by the relative displacement raised by the system during the increasing of the displacement.
In Eq. (1) the first dissipative component represents the constant friction damper (CFD) constitutive law. In Fig. 2 a comparison is shown in terms of load-displacement (Fig. 2a) and cycle-displacement (Fig. 2b) for CFD and VFD with and without the CFD component respectively. In detail, it can be seen that, while the force of the CFD is constant after the activation of the sliding, the force of the VFD is characterized also by initial normal stress at the interface equal to the slipping force (or zero if c_0=0), and the maximum dissipative force is attained at maximum displacement.

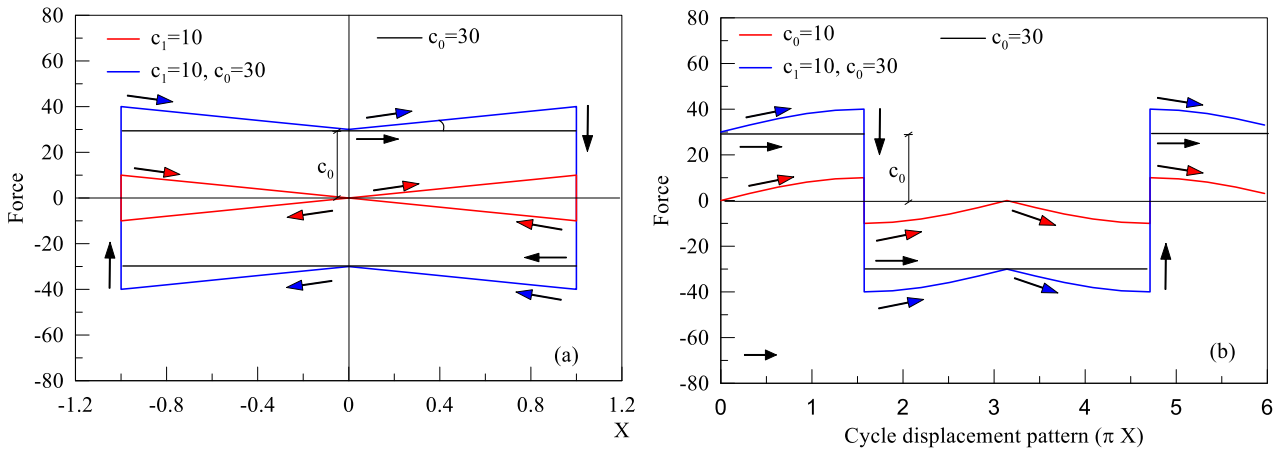


Figure 2. Comparison between VFD and FVD.

3. Implementing on a SDOF model

In this paper, a comparison is proposed on systems that can be modelled as single degree of freedom (SDOF) equipped alternatively with CFD and VFD. The seismic input is previously assigned. The responses are compared for a given level of additional maximum dissipation force to study the feasibility of applying the novel device.
The following equation of motion represents the behavior of the investigated systems:

$$\ddot{x} + 2\varsigma\omega\dot{x} + \frac{F_d}{m} + \omega^2 x = -\ddot{x}_g \tag{2}$$

where F_d is the dissipative force, m is mass, x is the displacement (the upper dot means time derivative), ς is the damping factor and ω^2 is the square own frequency obtained by dividing the stiffness by the mass. Moreover, $\ddot{X}_g$ is the external input in terms of ground accelerations and is modelled as a zero mean Gaussian white noise process characterized by the correlation function $R(\tau)$ given by:

$$R(\tau) = E\left[\ddot{X}_g(t)\,\ddot{X}_g(t+\tau)\right] = 2\pi S_0 \delta(\tau) \tag{3}$$

In Eq. (3) $E[\cdot]$ is the average operator, t means time, τ is a time delay, $\delta(\tau)$ is the Dirac's delta and S_0 is the Power Spectral Density of $\ddot{X}_g$.

4. Numerical application

To investigate on the mechanical and structural performances of the device, a series of time-history dynamic analyses are performed, starting from the response of the SDOF model with no dissipater in terms of displacements, velocities and the restoring forces. Then, the effect of using the VFD device in the reduction of the displacements was studied for different levels of additional dissipating forces designed as a ratio of the elastic restoring force. In detail the design procedure has been established as follows:

- Defining of the maximum dissipative force for the constant friction damper (CFD) (fixing c_0 in Eq. 1) as a ratio of the maximum elastic restoring force of the damped system;
- Performing the analysis to calculate the dynamic response (displacement, velocity, restoring force);
- Defining the linear variable part of the friction dissipater (VFD) (Eq. (1) fixing c_0 as for the CFD and c_1 to reach the 50% and the 100% of the pure friction maximum force respectively (Fig. 3);
- Performing the analysis to observe the differences in the dynamic response compared with the case of constant friction.

In addition to what above described, the efficiency of the VFD device has been investigated by keeping the same mechanical characteristics of the dissipater and varying the seismic input, using an accelerogram having higher peak ground acceleration (PGA). In doing so, an important consideration need to be pointed out that highlight the importance of the proposed novel device: differently from the constant friction damper (CFD) that provides always the same damping force regardless the intensity of the seismic input, the variable friction damper (VFD) take advantage in giving high dissipating force as much as the structure experience high displacements (that is the case of higher seismic inputs).

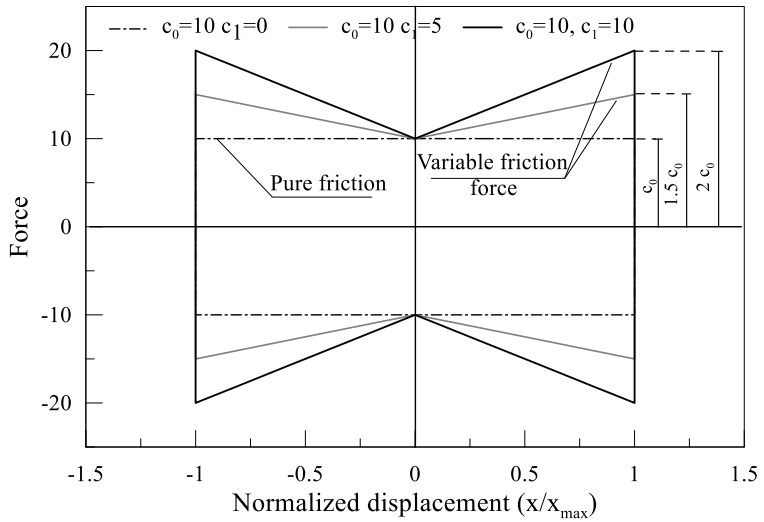


Figure 3. Case study for the VFD.

4.1. CHARACTERISTICS OF THE MODEL AND SEISMIC INPUT

The SDOF model was obtained starting from a single storey 3D reinforced concrete moment resisting frame structure of 4 meters' height, having columns with section dimension of 0.3 m x 0.5 m with the major inertia axis oriented along the direction of the seismic input. The modulus of elasticity of the reinforced concrete was assumed 30 GPa and a uniform distributed mass of 36 tons has been assigned to the floor. The system is shown in Fig. 4. For a given natural damping coefficient of 0.05, the dynamic characteristics of the SDOF model can be summarized as follows:

$$k = 0.5 \cdot 4\frac{12EI}{H^3} = 35156250\,\text{N/m}; \qquad \omega = \sqrt{\frac{k}{m}} = 31.25\,\text{sec}^{-1}; \qquad T = \frac{2\pi}{\omega} = 0.2\,\text{sec} \tag{4}$$

being k, ω and T the lateral stiffness, the own frequency and the vibration period respectively.

The seismic input was defined by two stationary accelerograms of 25 seconds (having the probabilistic characteristics expressed by Eq. 3), whose power spectral density S_0 have been fixed to obtain a peak ground acceleration (PGA) of 4.83 m/s^2 and 11.28 m/s^2 (corresponding to 0.5 g and 1.15g respectively). The dynamic response obtained from the first seismic input of 0.5g was used to design the dissipaters; then, the second seismic input of 1.15 g was used to investigate on the efficiency for seismic inputs higher than that used for the design. The accelerograms and the response spectrum curves are shown in Fig. 5a-b respectively.

Three set of analyses were performed beside the case of system without friction damper. In each set, the pure friction was maintained constant to exhibit a force approximately equal to 5%, 10% and 20% of the maximum restoring force of the system without friction damper respectively. The variable friction was varied to have the maximum variable component of the friction force about 15% of the maximum restoring

force of the system not with no damper. The analysis cases are summarized in Table 1 (the meaning of symbols and letters are explained in the next section).

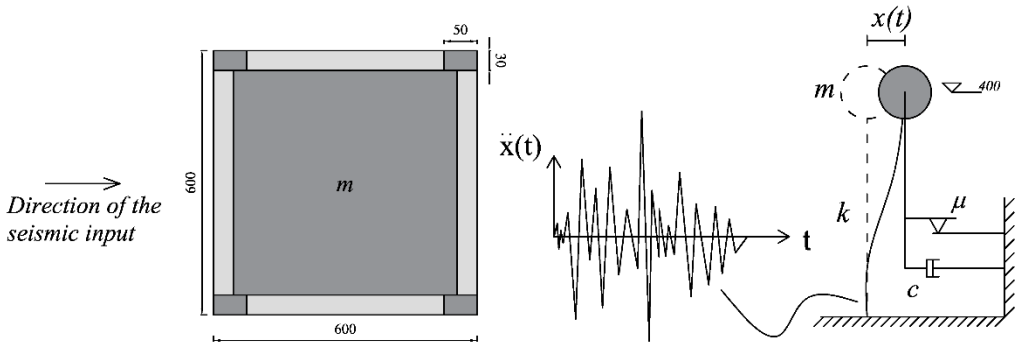


Figure 4. SDOF model.

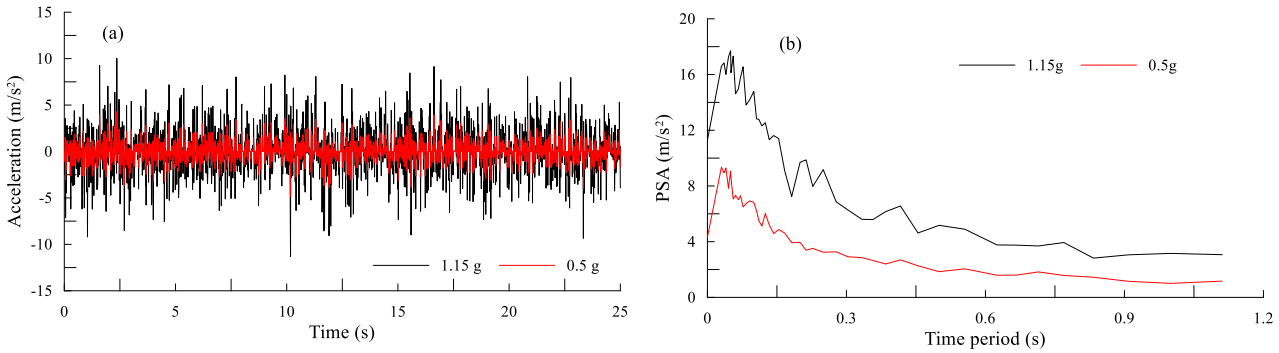


Figure 5. Accelerograms (a); Response spectra (b).

4.2. ANALYSIS RESULTS

Time-history analyses were performed calculating the dynamic responses. The analysis cases are labeled with the letter A and B for the accelerograms with PGA of 0.5g and 1.15g respectively. The results for the case A are summarized in Table 1 where number that follows A refers to ratio (in percentage) between the pure friction force and the maximum restoring force while the roman numbers I and II correspond respectively to a maximum variable friction force half and equal to the pure friction. For example, A-5 refers to a system exhibiting a constant friction force equal to 5% of the maximum restoring force exhibited by the system without dissipater (A) and without the contribution of the variable friction component of the force (c_1 = 0), A-5-I refers to a system exhibiting a constant friction of 5% and maximum variable component of friction equal to 50% of the constant friction, A-10-II refers to a system exhibiting a constant

friction of 10% and maximum variable component of friction equal to 100% of the constant friction and so on. The restoring and the damping forces are normalized with respect to the mass of the system.

Table I. Analysis scheme and results for the SDOF model equipped with VFD.

	Restoring force RF		Friction coefficients		Friction force FF		Dissipated energy	
CASE ANALYSIS	Maximum normalized (m/s^2)	Ratio "without VFD/ withVFD"	c_0 (m/s^2)	c_1 ($1/s^2$)	Maximum normalized (m/s^2)	FF/RF (%)	E	%
A	4.230	1.000	0	0	0	0	0.272	1.000
A-5	2.830	0.669	0.145	0	0.145	0.05	0.341	1.253
A-5-I	2.470	0.584	0.145	50	0.266	0.11	0.373	1.371
A-5-II	2.180	0.515	0.145	100	0.361	0.17	0.401	1.475
A-10	2.194	0.519	0.225	0	0.225	0.10	0.362	1.332
A-10-I	2.070	0.489	0.225	50	0.324	0.16	0.385	1.417
A-10-II	1.899	0.449	0.225	100	0.42	0.22	0.408	1.499
A-20	1.850	0.437	0.370	0	0.37	0.20	0.405	1.490
A-20-I	1.740	0.411	0.370	50	0.463	0.27	0.418	1.536
A-20-II	1.720	0.407	0.370	100	0.552	0.32	0.439	1.613

Figure 6 and Fig. 7 show the results in terms of displacements, velocities and restoring forces for the cases A-5, A-5-I, A-5-II and A-20, A-20-I, A-20- II on a 25 sec time window. The effectiveness of the dissipation device is evident and shows a considerable reduction in the dynamic response. Starting from the system with no VFD and analyzing the case with additional dissipation of 5%, a maximum displacement of 0.0043 m, decreasing to 0.0029 m, 0.0025 m and 0.0022 m for A-5, A-5-I, A-5- II is observed, showing a gradually marginal reduction. In the case of additional dissipation of 20%, a maximum displacements of 0.0021 m, 0.00185 m and 0.00181 m for cases A-20, A-20-I, A-20-II is obtained.

The relative reduction of the displacement (and therefore of the elastic restoring force) is higher in the case of additional dissipation of 5%, while the relative reduction is lesser for higher values of dissipation forces. These considerations give a preliminary idea about the efficiency of the target dissipation levels to be reached.

In Fig. 8, the comparisons between the restoring forces and the force experienced by the dissipaters are shown for a time interval between 10 and 11 seconds for the analysis cases A-5 and A-20 (for CFD and VFD). Moreover, Fig. 9 shows for a wide time interval (from 10 to 13 seconds) the dynamic response for the analysis cases A-10 and A-10-II to highlight the variation of the dynamic response depending on the type of friction device used (constant or linear variable).

Based on the results obtained by the time-history analyses, some considerations can be drawn. With a low level of constant additional friction dissipative force (as for the case A-5 in which the friction force is 5% of the elastic restoring force) given to the constant part, by using the novel device, higher dissipative forces can be obtained. In detail, the range varies from 5% to 10% to 17% depending on weather a device with capacity of 50% or 100% of the constant friction is used. If the case A-10 is analyzed, the novel device provides, starting from the 10%, 16% and 22% for the cases A-10-I and A-10-II respectively. Finally, for the case of 20% of dissipative forces, the 27% and the 32% are obtained for the cases A-20-I and A-20-II

respectively. However, as for the displacements, the increasing of the ratio between the force experienced by the friction dissipater and the elastic restoring force is marginal showing a non-linear trend. This confirms what assessed about a definition of an optimal threshold value to be defined for the additional dissipation. The results of what above described are shown in Tab. 1 and Fig. 10a. Fig. 10b shows the normal probability distribution of the displacements obtained by the analyzes. It can be noted that the dispersion in the case of system with no additional dissipation is much higher than in the case with VFD. The dispersion reduces with the increasing of the dissipation, giving information on the reduction of the range of displacements.

In Fig. 11, a comparison is made in terms of force-displacement cycles obtained in the case of CFD (Fig. 11a), in the case of VFD (Fig. 11b) with a capacity of 0.42. The case analyses are A-10 and A10-II respectively.

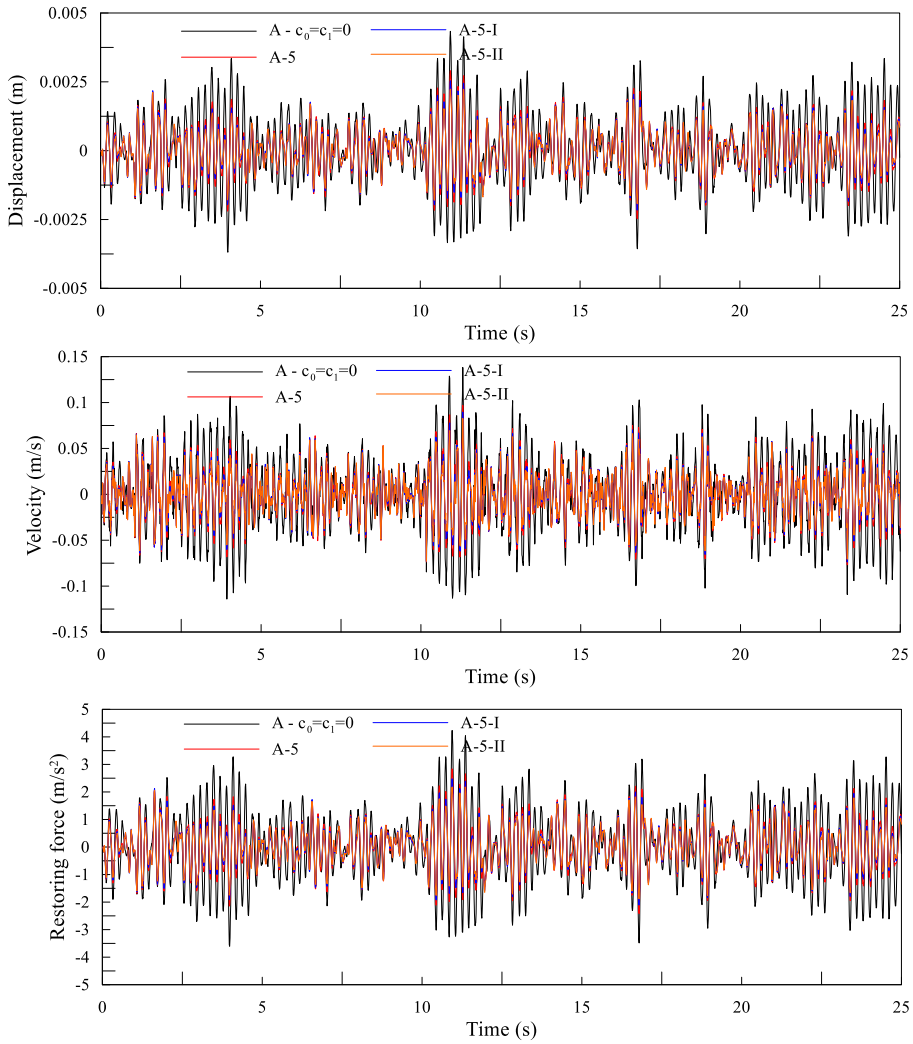


Figure 6. Time-history results for the case of constant friction dissipater (CFD) and variable friction dissipater (VFD) for the case analysis A, A-5, A-5-I and A-5-II.

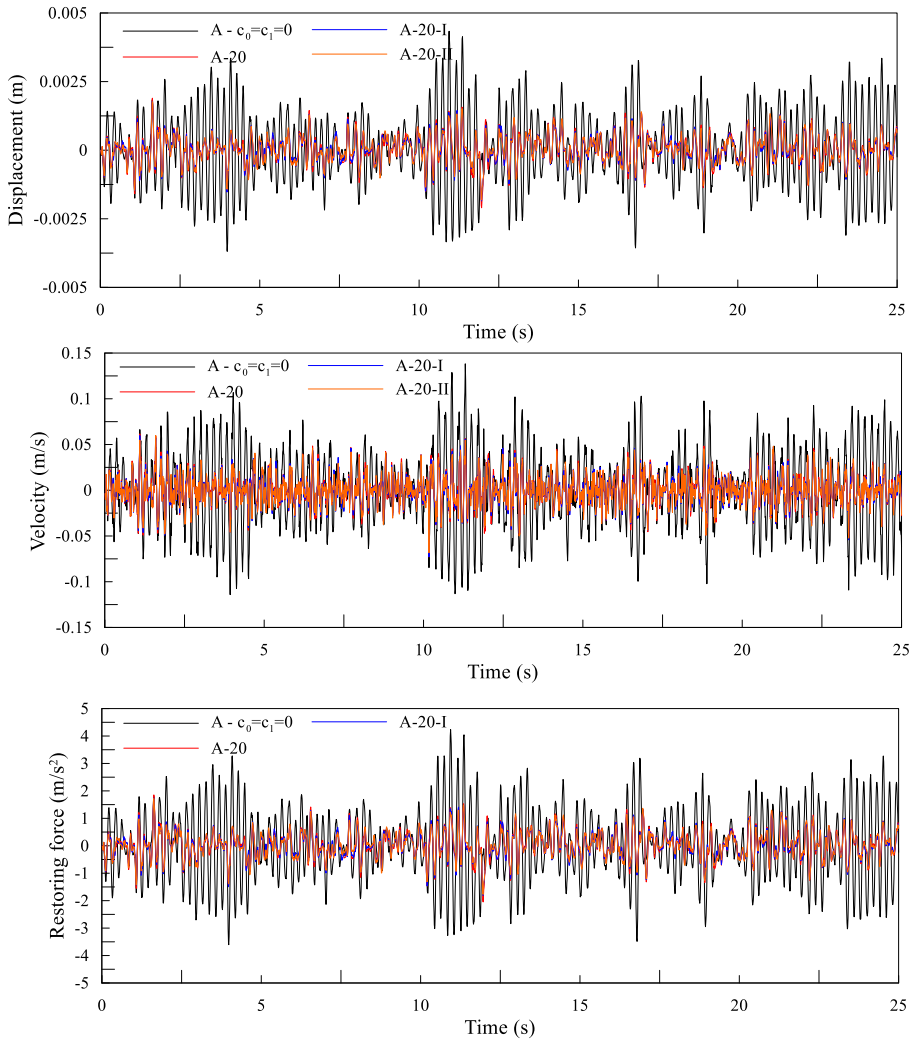


Figure 7. Time-history results for the case of constant friction dissipater (CFD) and variable friction dissipater (VFD) for the case analysis A, A-20, A-20-I and A-20-II.

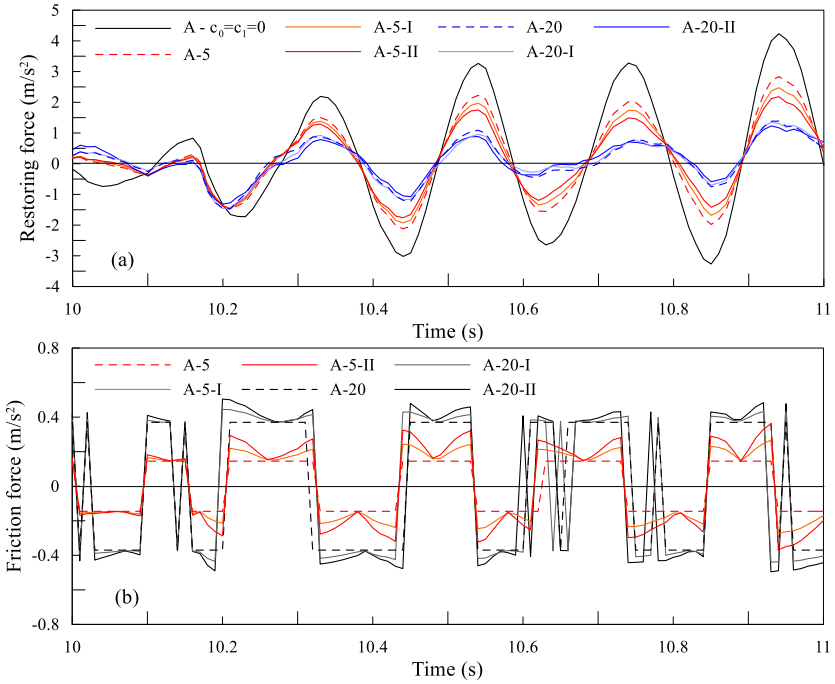


Figure 8. Restoring force and friction forces for time interval of 10-11 seconds for the various analysis cases.

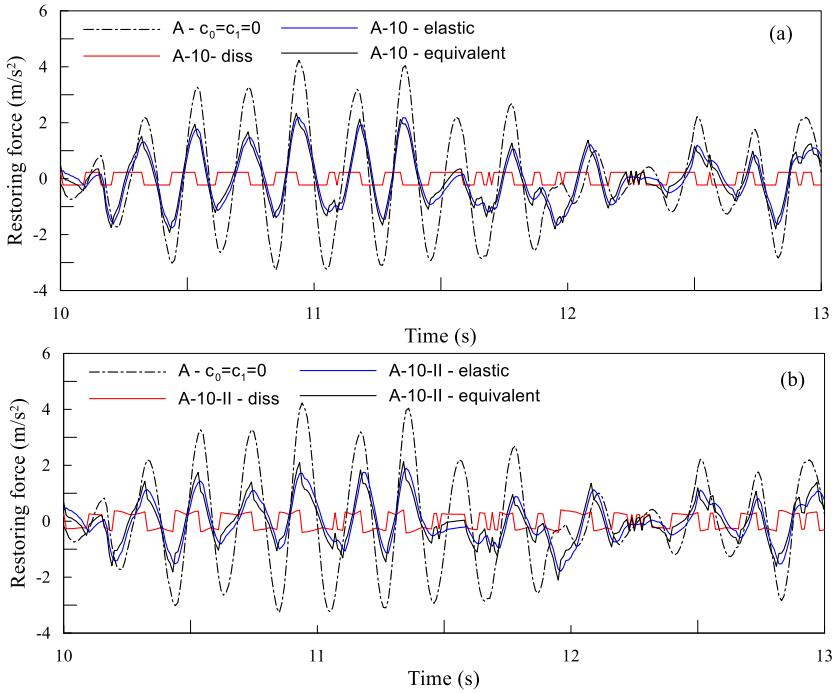


Figure 9. Response of the system equipped with CFD (a) and VFD (b).

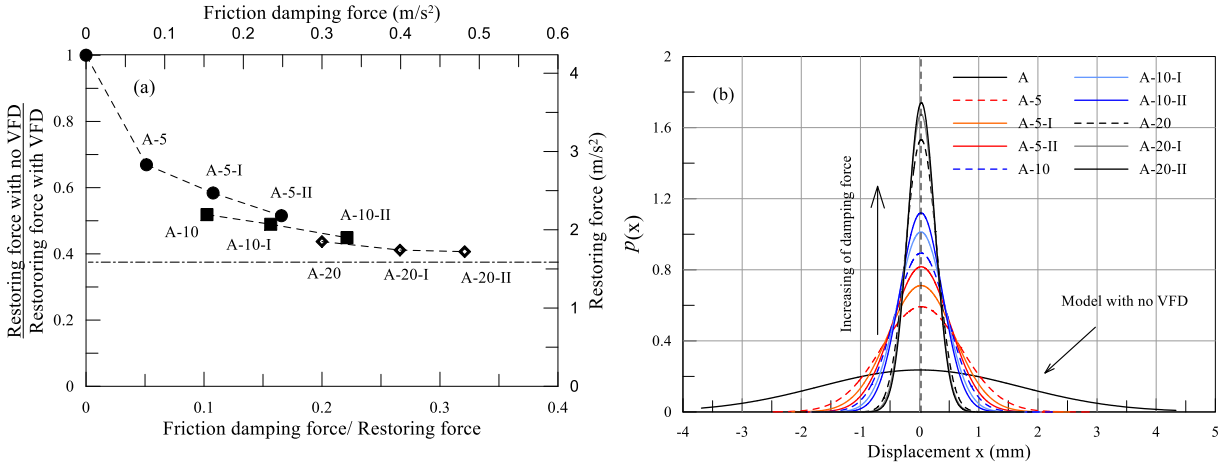


Figure 10. Reduction of the restoring force with the increasing of the damping force (a); Normal distribution of the displacements for the various analyses cases (b).

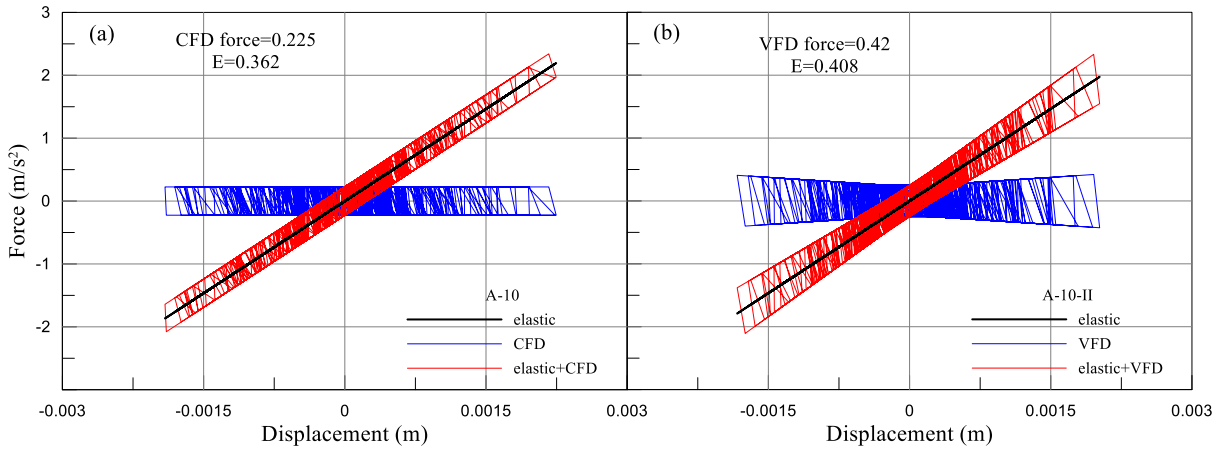


Figure 11. Force-displacement cycles: CFD (a); VFD (b).

An important consideration that can be drawn is that by using the novel device it is not necessary to define high friction forces to reach high levels of dissipation but it is possible to design a device that exploits the geometrical effect to increase the dissipative force with the increasing of the displacements. This effect is clearly highlighted in Fig. 12 a, where the maximum forces of the dissipaters are shown for the two seismic input A and B respectively with PGA of 0.5g and 1.15g. It is clearly shown that when the seismic input increases and, consequently, the displacements are higher, the maximum force of the VFD are higher, confirming the high efficiency of the VFD devices compared with the CFD ones. In other word, the constant friction damper (CFD) provides always the same damping force regardless the intensity of the seismic input, while, by using variable friction dampers (VFD) the advantage is obtained in giving to the

structure high dissipating force as much as the structure experience high displacements. Therefore, the higher is the displacement that the system tends to reach, the higher is the force exhibited by the dissipater. Obviously, the relative travel of the dissipater after the slipping force has to be accurately defined during the design.

The same considerations about the displacement reductions (Fig. 12b) can be obtained also in the case of the seismic input B (1.15g). The maximum displacement drops from 0.01 m for the system with no dampers to 0.006 m and 0.005 m for the case analysis B5-II and B20-II respectively, indicating a marginal reduction of the displacement for higher values of damping forces.

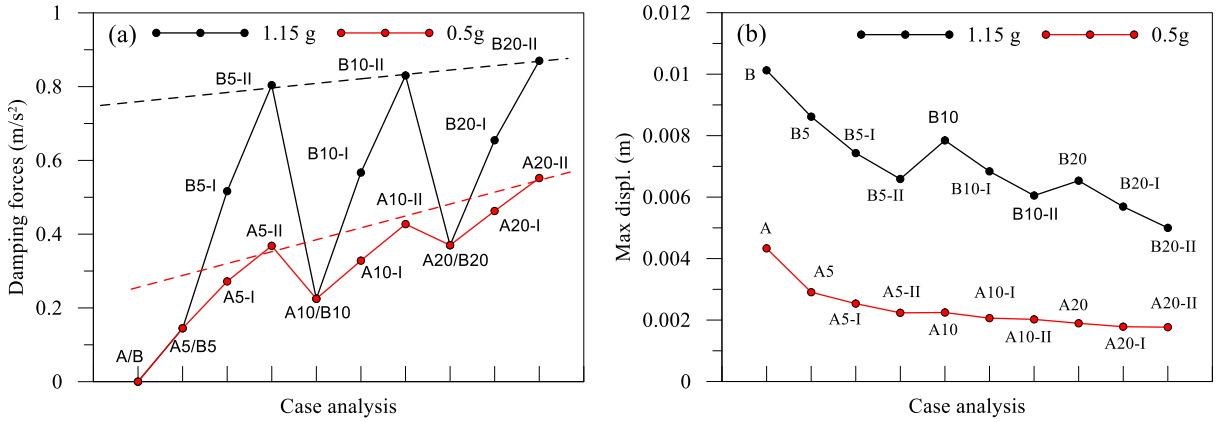


Figure 12. Maximum (a) friction forces and (b) displacements for seismic input A (0.5g) and B (1.15g)

5. Conclusions

In this paper a braking system borrowed from the mechanical engineering was presented as innovative method to be applied for the energy dissipation on moment resisting frame structures. The constitutive law of such a system is linearly dependent on the displacement of the device and therefore on the interstorey displacement and depends on the signum function of the interstorey velocity.

The feasibility of the braking system, the amount of energy that can be dissipated and the overall performance of structures equipped by these type of systems was studied in details for different intensities of the seismic input. The novel device resulted to be very efficient in terms of the reduction of the dynamic response, showing a capability of providing higher amount of dissipated energy and a significant reduction of the interstorey displacements. The system is certainly worthy of attention as it allows to reach greater dissipation forces than the classic friction dissipaters, exploiting the geometric effects of the device which causes the modification of the constitutive law by assigning a linear component dependent on the relative displacement.

References

Alotta, G., Cavaleri, L., Di Paola, M., Ferrotto, M.F. Solution for the Design and Increasing of Efficiency of Viscous Dampers. *The Open Construction and Building Technology Journal*, 10, (Suppl 1: M6):106-121, 2016.

Bagheri, S., Shishvan, S.S., Barghian, M. and Baniahmad, B. A new energy dissipative cable bracing system. *Advances in Structural Engineering*, 22(14):3134–3146, 2019.

Bhaskararao, A.V., Jangid, R.S. Seismic analysis of structures connected with friction dampers. *Engineering Structures*, 28 (2006) 690–703

Braga, F., Laterza, M. Field testing of low-rise base isolated building. *Engineering Structures* 26:1599-1610, 2004.

Cancellara, D., De Cicco, S., De Angelis, F. Assessment and vulnerability reduction of under-designed existing structures: Traditional vs innovative strategy. *Computers and Structures* 221:44–64, 2019.

Cavaleri, L., Di Trapani, F., Ferrotto, M.F. Experimental determination of viscous dampers parameters in low velocity ranges. *Ingegneria Sismica*, 34(2):64-74, 2017.

Cook, J., Rodgers, G.W. and MacRae, G.A. Design and Testing of Ratcheting, Tension-Only Devices for Seismic Energy Dissipation Systems. *Journal of Earthquake Engineering*, 2018. DOI: 10.1080/13632469.2018.1441765.

Dal Lago, B., Biondini, F., Toniolo, G. Seismic performance of precast concrete structures with energy dissipating cladding panel connection systems. *Structural Concrete*, 19: 1908–1926, 2018. https://doi.org/10.1002/suco.201700233.

Ferrotto, M.F., Asteris, P.G., Cavaleri, L. Strategies of Identification of a Base-Isolated Hospital Building by Coupled Quasi-Static and Snap-Back Tests, *Journal of Earthquake Engineering*, 2020, DOI: 10.1080/13632469.2020.1824877;

Ferrotto, M.F., Cavaleri, L., Di Trapani, F., & Castaldo, P. Full scale tests of the base-isolation system for an emergency hospital. In COMPDYN 2019 7th International Conference on Computational Methods in Structural Dynamics and Earthquake Engineering 2019 - vol.1 (pp. 2012-2025). M. Papadrakakis, M. Fragiadakis (eds.). Crete, Greece, 24–26 June 2019.

Teruna, D.R., Majid, T.A. Budiono, B. Experimental Study of Hysteretic Steel Damper for Energy Dissipation Capacity. *Advances in Civil Engineering*, 631726, 12 pages. 2015. http://dx.doi.org/10.1155/2015/631726.

Using a tuned fluid inerter for improving the seismic performance of structures isolated with friction pendulum system

Dario De Domenico[1], Giuseppe Ricciardi[2]
[1] *Department of Engineering, University of Messina, Italy, dario.dedomenico@unime.it*
[2] *Department of Engineering, University of Messina, Italy, gricciardi@unime.it*

Abstract: The fluid inerter is the hydraulic realization of the inerter. It consists of a piston-cylinder device that conveys a fluid through an external helical channel, thereby generating rotational inertia of a fluid mass. Unlike other variants of the inerter (e.g. ball screw, rack-and-pinion inerter, and inerter with clutch), the assumption of linear behavior for the fluid inerter is largely inaccurate, because this device is characterized by a marked nonlinear damping effect. While identification and modeling of the fluid inerter were discussed in the relevant literature, the effect of its nonlinear damping contribution in relationship to structural vibration control remains unclear. Aim of this contribution is to present a feasibility study in which the nonlinear damping contribution of the fluid inerter is examined numerically. Some experimental findings relevant to a small-scale prototype of fluid inerter justify the assumption of nonlinear power law damping, in parallel with linear inertance. The fluid inerter is here employed in combination with low-damping rubber isolators as linear restoring terms, thus realizing a novel control scheme called Tuned Fluid Inerter (TFI). In this paper, the TFI is adopted to improve the seismic performance of isolated buildings. Optimal parameters of the fluid inerter are determined using random vibration theory, by modeling the base acceleration as a zero-mean Kanai-Tajimi stationary random process, and resorting to statistical linearization to handle the nonlinear terms. The seismic performance of structures equipped with friction pendulum system (FPS) is comparatively analyzed considering a classical Tuned Mass Damper (TMD), the Tuned Mass Damper Inerter (TMDI) with mechanical inerter, and the novel TFI. Based on a benchmark six-story building, the seismic performance of the various structural control systems is analyzed in terms of isolators' displacement demand, interstory drift and acceleration response. It is shown that the effect of the inherent nonlinear damping of the fluid inerter is crucial for reducing the peak response under pulse-like ground motions that may occur in near-field earthquake events.

Keywords: Tuned Mass Damper, Fluid Inerter, Inerter, Nonlinear damping, Passive vibration control, Optimal design, Seismic base isolation, Friction pendulum isolators.

1. Introduction

The inerter is a relatively new mechanical device capable to develop a resisting force proportional to the relative acceleration of its two terminals. The constant of proportionality, called inertance, has dimensions of mass and is called inertance. The inerter device completes the force-current analogy between mechanical and electrical networks, besides the spring and the dashpot elements. Since its introduction in 2002 (Smith, 2002), the inerter has been studied and applied to different fields. Of particular relevance to the present contribution, the inerter has totally revolutionized the technology of vibration control because of its inherent capability to adjust the inertial properties of a system without allocating actual physical mass (Ikago et al., 2012). Just to quote a few prominent examples, inerter-based systems were used to improve the

© 2021 *by Dario De Domenico and Giuseppe Ricciardi.*

performance of tuned mass dampers (TMDs), by partly or totally replacing the tuned mass thus realizing a lightweight and cost-effective device called Tuned Mass Damper Inerter (TMDI) (Giaralis and Taflanidis, 2018; Pietrosanti et al. 2017) and Tuned Inerter Damper (TID) (Lazar et al., 2014; Gonzalez-Buelga et al., 2017), respectively. Inerters were also studied to realize nonlinear energy sinks (Javidialesaadi and Wierschem, 2019), for vibration mitigation of cables (Lazar et al., 2016), for wind-vibration control of high-rise buildings (Giaralis and Petrini, 2017; Wang et al., 2019), for seismic protection of buildings (Takewaki et al., 2012; Zhao et al., 2019) and storage tanks (Luo et al., 2016; Jiang et al., 2020), or in combination with traditional base isolation systems (De Domenico and Ricciardi 2018b, 2018c; Hashimoto et al., 2015; Di Matteo et al. 2019). Focusing on the latter field of application, in seismic base isolation the dis-placement demand is mostly concentrated at the level of the isolation devices, which should be flexible enough to shift the fundamental frequency of the structure to a low-frequency (high-period) region, where the energy transferred by the ground-motion base acceleration is typically low. While the isolators guarantee low base shear and structural accelerations in the superstructure, flexible isolation devices imply large displacement demands as a counter effect. These displacements can be of the order of 40-70 cm depending on the earthquake intensity and on the selected flexibility characteristics of the isolators. Moreover, base-isolated structures may suffer from undesired resonance effects under near-field long-period ground motions (Takewaki et al., 2011). Hence, various strategies have been proposed to reduce isolators' dis-placements, such as providing supplemental damping, using adaptive energy dissipation systems like gap dampers combined with isolators (Zargar et al., 2013; De Domenico et al. 2020a), or through hybrid control strategies combining base isolation with a TMD located above or below the isolation level (Taniguchi et al., 2008; De Domenico and Ricciardi 2018a). In this regard, the effectiveness and robustness of the TMD in mitigating isolators' displacement is closely related to the mass actually allocated, which may be hampered by feasibility considerations and architectural constraints. This is why the use of inerter-based vibration absorbers turns out to be particularly useful to reduce the displacements of isolators without employing large mass ratios.

Along this research line, an inerter-based vibration absorber is proposed in this paper to improve the seismic performance of base-isolated structures equipped with Friction Pendulum System (FPS). Unlike most of the literature studies that studied the performance of the inerter in its various mechanical configurations, e.g. rack-and-pinion inerter, ball screw inerter or inerter with clutch (Málaga-Chuquitaype et al., 2019), the present contribution addresses the hydraulic realization of the device, called *fluid inerter* (Swift et al., 2013). A peculiarity of the fluid inerter is related to its marked nonlinear damping effect, whose effect in the vibration control of structures is still unclear (Smith and Wagg, 2016). Therefore, the proposed study aims at investigating a novel Tuned Fluid Inerter (TFI) system for applications to base-isolated buildings, in particular, to buildings equipped with FPS. Optimal design of the fluid inerter parameters is conducted using random vibration theory, with the base acceleration being modeled as a stationary Kanai-Tajimi zero-mean Gaussian random process. The nonlinearities of FPS and fluid inerter are handled through statistical linearization. Seismic performance of the proposed system is finally analyzed considering a series of ground motion records having different frequency characteristics. The effect of the inherent nonlinear damping contribution of the fluid inerter is also analyzed and discussed.

2. Hysteretic properties of the helical fluid inerter

A sketch of the helical fluid inerter is shown in Figure 1. As can be seen, this device consists of a piston-cylinder block that, due to its linear movement, conveys a fluid through an external helical channel, which

generates rotational inertial of the fluid mass. Due to the rotational movement of the fluid, a main component of the resisting force is the so-called inerter contribution that is proportional to the relative acceleration of the two device terminals, and the proportionality coefficient (inertance b with dimensions of mass) is related to the volume of fluid and to the squared ratio of the cylinder area to helical channel area A_1 / A_2 (Swift et al., 2013)

$$F_{\text{inerter}} = b\ddot{x} \quad \text{with} \quad b = \frac{m_{hel}}{1+\left(h/2\pi r_4\right)^2}\left(\frac{A_1}{A_2}\right)^2 \tag{1}$$

where x denotes the linear movement of the piston (relative displacement of the two terminals of the device). In Eq. (1) $m_{hel} \approx \rho_f \ell A_2$ is the mass of the fluid in the helical channel, with ρ_f denoting the mass density of the fluid, while ℓ and $A_2 = \pi r_3^2$ represents the length and cross-sectional area of the helical channel, respectively. Based on Eq. (1), a desired value of inertance can be obtained by simply adjusting the geometrical parameters of the device for given physical properties of the fluid.

Besides the inerter force in Eq. (1), other contribution to the total resisting force are (Swift et al., 2013): 1) the damping effects due to the pressure drops caused by the viscosity of the internal fluid; 2) inlet and outlet forces at the transition from the cylinder to the helical tube and vice-versa; 3) shear friction between walls of the piston and the surrounding cylinder; 4) friction effects at the two seals at either end of the piston rod moving inside the cylinder. These four contributions have the following expressions

$$\begin{aligned}
&1)\ F_{d-\text{helical}} = c_p^{NL}\dot{x}^{1.75} \quad \text{with} \quad c_p^{NL} = \frac{0.0664\mu_f^{0.25}\rho_f^{0.75}\ell A_1}{r_3^{1.25}}\left(\frac{A_1}{A_2}\right)^{1.75} \\
&2)\ F_{d-\text{inlet}} = c_1\dot{x}^2 \ \text{ with } \ c_1 = 0.25A_1\rho_f\left(\frac{A_1}{A_2}\right)^2; \quad F_{d-\text{outlet}} = c_2\dot{x}^2 \ \text{ with } c_2 = 0.5A_1\rho_f\left(\frac{A_1}{A_2}\right)^2 \\
&3)\ F_{\text{shear}} = \frac{\mu_f 2\pi r_2 L}{\Delta r}\dot{x} \\
&4)\ F_{\text{friction}} = f_0\,\text{sgn}(\dot{x})
\end{aligned} \tag{2}$$

where μ_f is the dynamic viscosity of the fluid, Δr is the clearance between piston head and cylinder wall (typically in the order of $\Delta r = 0.1$mm (Swift et al., 2013)), and f_0 is a friction coefficient that can be calibrated from pseudo-static experimental test. Following Eqns. (1) and (2), the total resisting force of the fluid inerter is the sum of an inerter force F_{inerter} (proportional to the relative acceleration) and a total damping force (dependent on the relative velocity), the latter being the sum of all the contributions appearing in Eq. (2):

$$F = F_{\text{inerter}} + F_{\text{damping,tot}} \tag{3}$$

The influence of the various terms contributing to the total damping force is analyzed in an attempt to develop a simplified model that can be used for practical earthquake engineering purposes. In Figure 2 (left) the damping terms expressed in Eq. (2) are plotted in a semi-logarithmic scale with increasing piston velocities from 0.2 m/s to 1 m/s. The geometrical parameters of the fluid inerter are those of the patent

document (Glover et al., 2009, while the internal fluid is assumed to be water ($\rho_f = 1000\,\mathrm{kg/m^3}$, $\mu_f = 10^{-3}\,\mathrm{Pa\,s}$). It is clearly seen that, among the four damping terms, the most important contribution is the helical tube damping force, which almost coincides with the total damping force. To confirm this result, in Figure 2 (right) the force-displacement loops under a sinusoidal imposed motion having frequency 1.5 Hz e amplitude 50 mm (maximum velocity of 470 mm/s) for the complete model and for a simplified model including only the helical tube damping term are compared. As can be seen, the simplified model (whose dynamic sketch is reported in the bottom-right part of Figure 1) is accurate enough to describe the hysteretic characteristics of the fluid inerter device. As a result, for preliminary design purposes, the inlet and the outlet terms, as well as the shear forces can be practically neglected. The fluid inerter can be modeled as a linear inerter in parallel with a nonlinear dashpot (power law damping factor of 1.75).

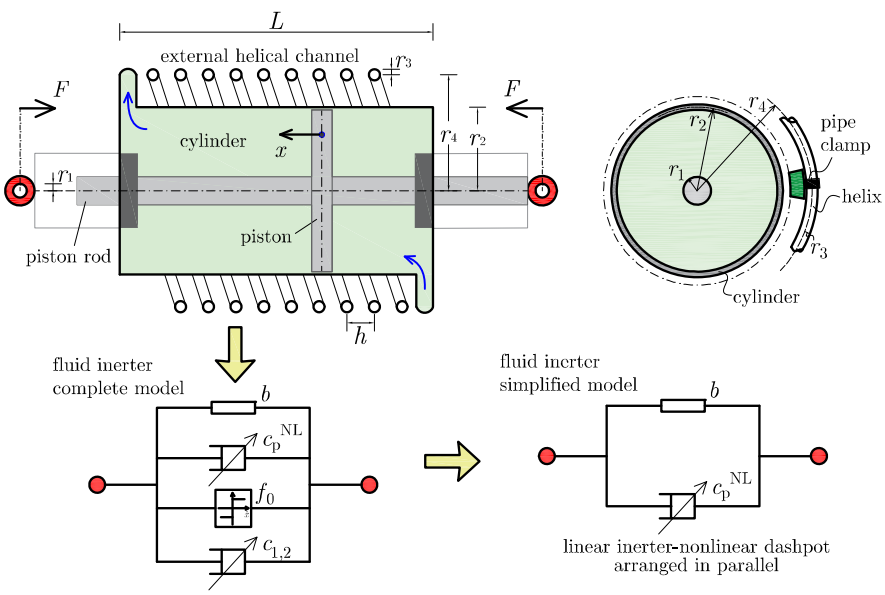


Figure 1. Helical fluid inerter (top) and related simplified mechanical models.

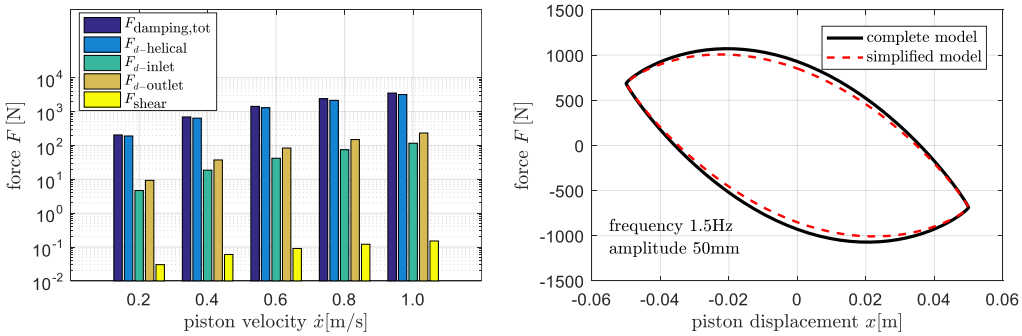


Figure 2. Contributions to the total damping force of the fluid inerter (left) and corresponding hysteretic curve un-der an imposed sinusoidal displacement.

The simplified model of a linear inerter in parallel with a nonlinear dashpot is validated against experimental results from the literature. In particular, a small-scale fluid inerter with 7 coils and filled with a silicone oil ($\rho_f = 802\,\text{kg/m}^3$ and $\mu_f = 0.00168\,\text{Pa s}$) was tested by Smith and Wagg (2016), and the corresponding experimental results (for a sinusoidal displacement amplitude of 17.5 mm and frequency 3 Hz) are shown in Figure 3, along with pertinent numerical simulations obtained by the complete and simplified model. It can be observed that the numerical models are accurate enough to capture the actual resisting force variation obtained by the experiments. Interestingly, the complete and simplified models are almost superimposed to each other, thus confirming the validity of the simplified model composed of a linear inerter in parallel with a nonlinear dashpot, which will be used afterwards in this paper.

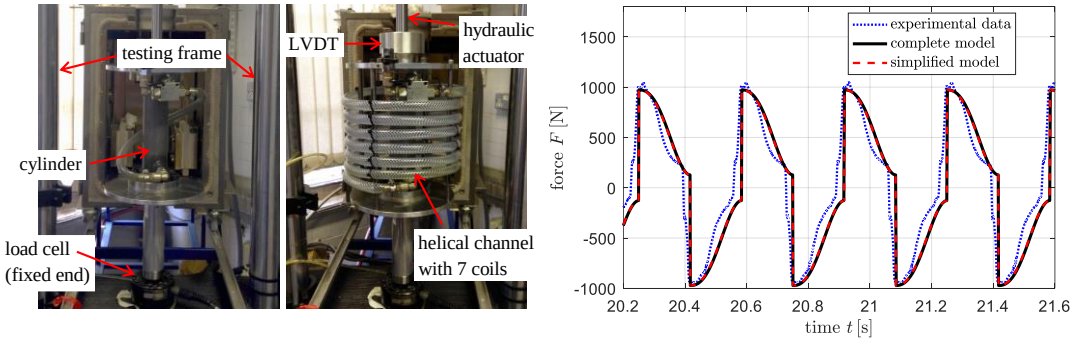


Figure 3. Laboratory tests on a small-scale fluid inerter (Smith and Wagg, 2016) and corresponding validation of the numerical model against experimental force results.

3. Building isolated with FPS equipped with tuned fluid inerter (TFI)

A simple model of *n*-story framed building isolated with FPS equipped with a tuned fluid inerter (TFI) is reported in Figure 4. The overall structure is thus composed of two systems, the primary system, i.e. the base-isolated building (with FPS), and the secondary system i.e. the TFI. In particular, the TFI is a grounded fluid inerter device that is connected to the base isolation floor through some linear restoring elements (providing the secondary system with the necessary re-centering function). Such linear restoring elements, idealized as a spring and a dashpot element arranged in parallel, can be practically realized via low-damping rubber bearings, as sketched in Figure 4. The FPS is modeled through a simplified Coulomb frictional model (rigid-plastic behavior), thus neglecting the complicated variability of the friction coefficient with axial load, sliding velocity and heating phenomena for simplicity. The equations of motion of the combined structure subjected to a horizontal earthquake-induced acceleration $\ddot{u}_\text{g}$ are expressed as follows

$$\begin{aligned} &\mathbf{M}_\text{s}\ddot{\mathbf{u}}_\text{sr} + \mathbf{C}_\text{s}\dot{\mathbf{u}}_\text{sr} + \mathbf{K}_\text{s}\mathbf{u}_\text{sr} = -\mathbf{M}_\text{s}\boldsymbol{\tau}_\text{s}\left(\ddot{u}_\text{g} + \ddot{u}_\text{b}\right) \\ &m_\text{b}\ddot{u}_\text{b} + F_{FPS} - F_{\text{s1,b}} - F_\text{rubber} = -m_\text{b}\ddot{u}_\text{g} \\ &m_\text{t}\ddot{u}_\text{t} + F_\text{rubber1} + F_\text{FI1} = -m_\text{t}\ddot{u}_\text{g} \end{aligned} \tag{4}$$

where $\mathbf{M}_\text{s}, \mathbf{C}_\text{s}, \mathbf{K}_\text{s}$ denote the mass, damping and stiffness matrices of the superstructure, respectively, $\mathbf{u}_\text{sr} = \mathbf{u}_\text{s} - \boldsymbol{\tau}_\text{s} u_\text{b}$ is the vector of superstructure displacements relative to the base-isolation floor, where,

$\mathbf{u}_s = [u_{s1},\ldots,u_{sn}]^T$ are the displacements relative to the ground, $\boldsymbol{\tau}_s$ is an n-dimensional vector of ones and u_b is the displacement of the FPS relative to the ground. Moreover, m_b indicates the mass of the base-isolation floor, $F_{FPS} = k_b u_b + f_b W_{tot}\,\mathrm{sgn}(\dot{u}_b)$ is the resisting force of the FPS, where $k_b = W_{tot}/R$ is the pendulum stiffness of the FPS (W_{tot} is the total weight of the overall structure) and f_b is the friction coefficient at the sliding interface, $\mathrm{sgn}(\cdot)$ is the signum function. Finally, $F_{s1,b} = c_1\dot{u}_{sr1} + k_1 u_{sr1}$ is the force transferred from the first floor of the superstructure to the FPS, $F_{rubber} = c_t\dot{u}_{tb} + k_t u_{tb}$ is the resisting force of the low-damping rubber bearings, dependent on relative displacements and velocities between TFI and FPS ($u_{tb} = u_t - u_b$ and $\dot{u}_{tb} = \dot{u}_t - \dot{u}_b$), and $F_{FI} = b\ddot{u}_t + c_p^{NL}\left|\dot{u}_t\right|^{1.75}\mathrm{sgn}(\dot{u}_t)$ is the nonlinear resisting force of the fluid inerter. The system of equations (4) is nonlinear due to the terms F_{FPS} and F_{FI}.

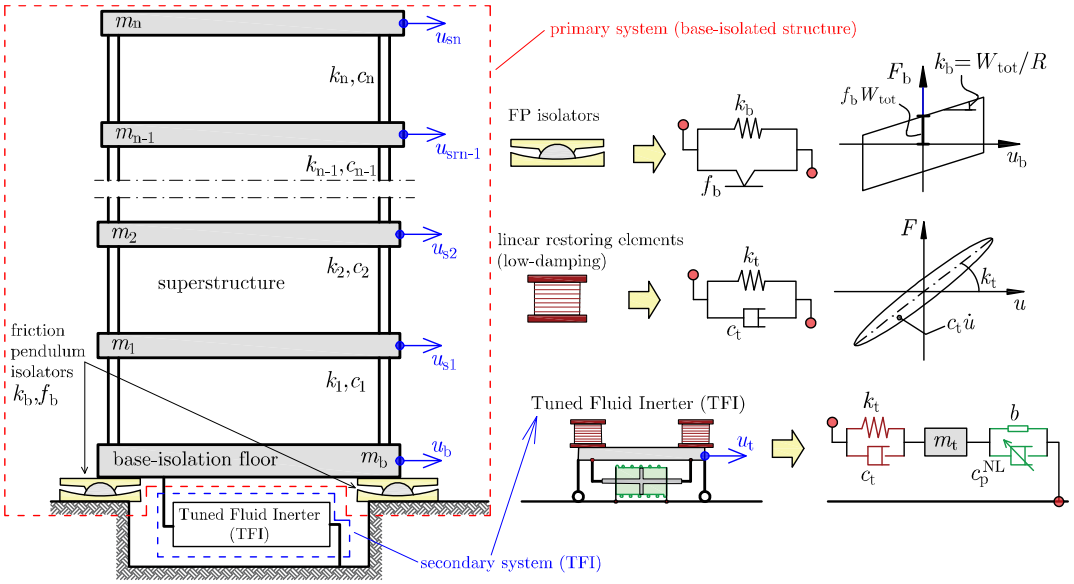


Figure 4. Model of a framed building with FPS equipped with tuned fluid inerter.

The base acceleration $\ddot{u}_g$ is modeled as a stationary zero-mean filtered Kanai-Tajimi random process, with filter parameters calibrated to capture different frequency characteristics of the earthquake excitation (De Domenico et al., 2020b)0. Moreover, the nonlinear terms F_{FPS} and F_{FI} can be linearized by resorting to the stochastic linearization technique, by converting the nonlinear contributions into linear viscous damping terms as follows

$$\begin{aligned} F_{FPS}^{SL} &= k_b u_b + c_{b,eq}\dot{u}_b \qquad \text{FPS} \\ F_{FI}^{SL} &= b\ddot{u}_t + c_{p,eq}\dot{u}_t \qquad \text{fluid inerter} \end{aligned} \tag{5}$$

where the linearization coefficients $c_{b,eq}$ and $c_{p,eq}$ can be obtained under the simplifying assumption of Gaussian response, which gives (De Domenico et al., 2019a; De Domenico et al., 2020b)

$$c_{\mathrm{b,eq}} = f_b W_{tot} \sqrt{\frac{2}{\pi}} \frac{1}{\sigma_{\dot{u}_\mathrm{b}}} \qquad \text{FPS}$$
$$c_{p,\mathrm{eq}} = 1.3952\, c_p^{NL} \sigma_{\dot{u}_\mathrm{t}}^{0.75} \qquad \text{fluid inerter} \tag{6}$$

in which $\sigma_{\dot{x}}$ denotes the standard deviation of the response process $\dot{x}$ (with $x = u_\mathrm{b}, u_\mathrm{t}$).

4. Optimal design of the fluid inerter parameters

The optimization of the fluid inerter parameters is carried out in a probabilistic framework, by assuming as objective function specific stochastic performance indices, i.e., combinations of the terms of the covariance matrices of the system response. In particular, once the linearization coefficients in Eq. (6) are defined, the linearized equations of motion can be solved in the frequency domain, so as to determine the response transfer functions and, by integration, the covariance matrices of the system response (De Domenico et al., 2020b).

In particular, the goal of the optimization procedure is to find the best inertance ratio $\beta = b / M_\mathrm{tot}$, and the best nonlinear damping ratio of the fluid inerter $\zeta_p^{NL} = c_p^{NL} / 2\sqrt{k_\mathrm{b} M_\mathrm{tot}}$ (where M_tot is the total mass of the structure) that minimize a specific objective function. The superstructure is set as a six-storey reinforced concrete building with a fundamental period $T_\mathrm{s} = 0.5\mathrm{s}$ and an inherent damping ratio of $\zeta_\mathrm{s} = 0.02$. The mass of each storey is assumed constant and equal to the mass of the base-isolation floor, so that the mass ratio $\mu_\mathrm{s} = M_{\mathrm{s,tot}} / M_\mathrm{tot} = 6/7$. A broad optimization study was discussed in a recent paper by the authors (De Domenico et al., 2020b). For the sake of brevity, in this contribution only a limited set of results are discussed. In particular, in Figure 5 we report the trend of the optimal fluid inerter parameters $\beta_\mathrm{opt}, \zeta_{p,\mathrm{opt}}^{NL}$, the corresponding linearization coefficients of both FPS and fluid inerter $\zeta_{\mathrm{b,eq}}, \zeta_{p,\mathrm{eq}}$, as well as the seismic performance of the system. The seismic performance is described in terms of four different response indicators, namely the variance of the top-storey displacement $\bar{\sigma}_\mathrm{us}^2$ (which is the objective function that is minimized in this case), the variance of the top-storey absolute acceleration $\bar{\sigma}_{\ddot{u}_\mathrm{As}}^2$, an energy-based indicator called filtered energy index (FEI) and the variance of the stroke of the TFI $\bar{\sigma}_\mathrm{utb}^2$. The FEI is a useful energy indicator quantifying the portion of input energy that is not dissipated by the passive vibration control system and, consequently, that filters into the structure. The FEI has been found to be a robust objective function for the optimization of a variety of passive vibration control systems, including hysteretic dampers, large-mass ratio TMDs (Reggio and De Angelis, 2015), TMDI (De Domenico and Ricciardi, 2018c) and viscous dampers (De Domenico et al., 2019b).

The above-mentioned response indicators are all shown in a normalized format (divided by the corresponding response indicator in the system without TFI), so that values below one indicate reduction. The results are plotted for four different earthquake frequency contents (firm, medium, soft soil and white-noise excitation as an idealization of an extremely broadband random process) in terms of the stiffness ratio $\kappa = k_\mathrm{t} / k_\mathrm{b}$. The other parameters are set as follows: $\mu_\mathrm{t} = m_\mathrm{t} / M_\mathrm{tot} = 0.005$; $f_\mathrm{b} = 0.04$; $\omega_\mathrm{b} = \sqrt{k_b / M_{tot}} = \sqrt{g / R} = \pi$; $\zeta_\mathrm{t} = c_\mathrm{t} / 2\sqrt{k_\mathrm{t} m_\mathrm{t}} = 0.1$.

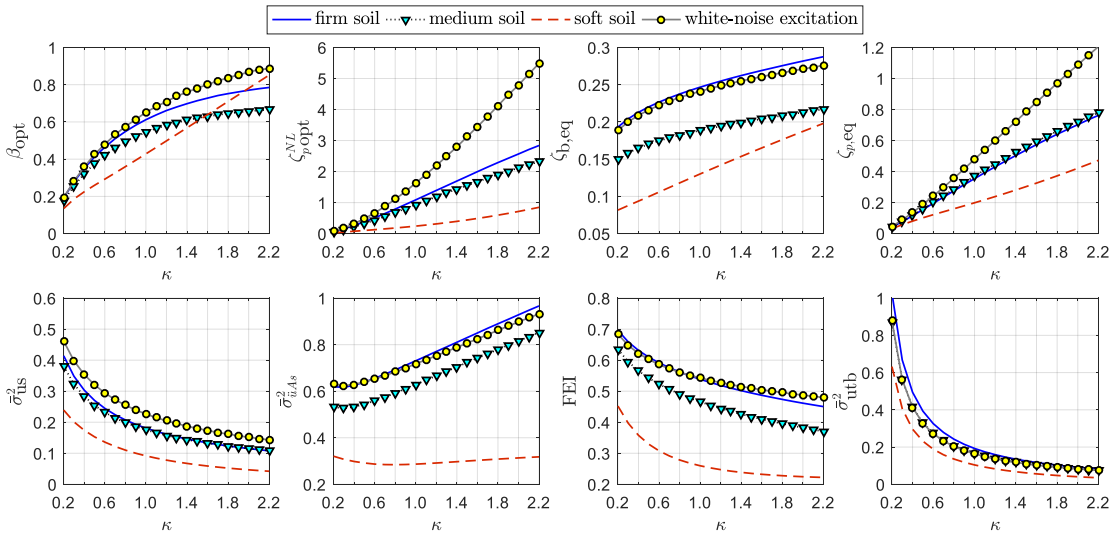


Figure 5. Optimal design and seismic performance of the FPS-TFI system for different soil conditions.

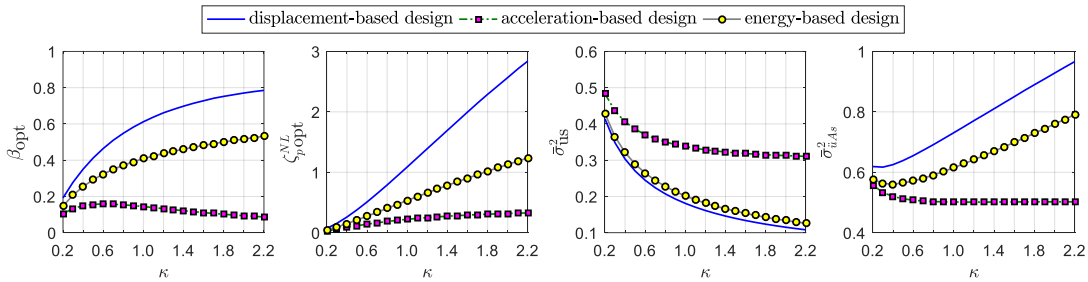


Figure 6. Optimal design and seismic performance of the FPS-TFI system for different design strategies.

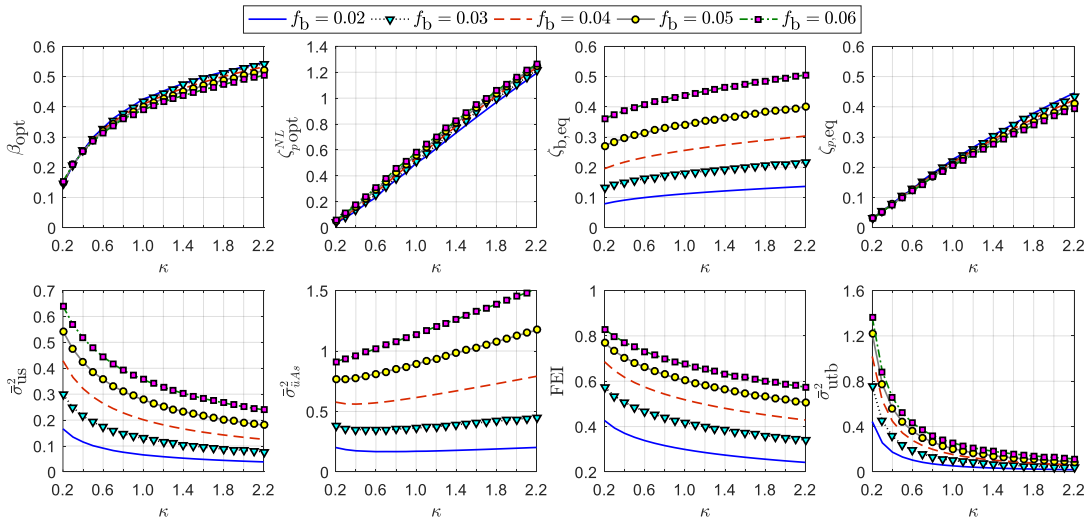


Figure 7. Optimal design and seismic performance of the FPS-TFI system for friction coefficient of the friction isolators.

Based on Figure 5, the values of β_{opt} and $\zeta_{p,\text{opt}}^{NL}$ increase with increasing κ. For the assumed parameters, the increase of the stiffness ratio κ produces a better performance in terms of displacement response, stroke of the TFI and energy dissipation, but is causes a slight increase of the absolute acceleration response. Based on the observed plots, a reasonable value for κ could be 0.8 to compromise between displacement and acceleration response. The soil conditions affect the trend of the optimal parameters: in particular, for soft soil conditions, the optimal fluid inerter parameters (both β_{opt} and $\zeta_{p,\text{opt}}^{NL}$) are lower than in the other soil conditions, at least for $\kappa < 1.4$.

The optimal fluid inerter parameters are obviously affected also by the design strategy, i.e., by the selection of the objective function to minimize. To demonstrate this, in Figure 6 the trend of the optimal design parameters along with the seismic performance in terms of displacement and absolute acceleration response are shown. Three curves are shown, each related to a specific objective function to minimize (displacement, acceleration or filtered energy index). In line with previous studies dealing with linear inerters (Pietrosanti et al., 2017; De Domenico and Ricciardi, 2018c), the energy-based design procedure (FEI) is a good compromise to obtain a reasonable reduction of both displacement and absolute acceleration. Moreover, the optimal fluid inerter parameters related to the FEI are bounded from below and from above by those obtained by the displacement- and acceleration-oriented optimal design procedures, respectively. Finally, the effect of the friction coefficient of the FPS is investigated in Figure 7, ranging from 2% to 6%. It is noted that the friction coefficient f_b has a modest influence on the optimal fluid inerter parameters, but significantly affects the control performance. In particular, higher f_b are associated with worse control performance, especially in terms of absolute acceleration response.

5. Seismic performance of FPS-TFI combined isolation system

The seismic performance of the proposed FPS-TFI combined isolation system is assessed via time history analyses considering a set of natural ground motion records. In particular, the FEMA P695 far-field (FF) and near-field (NF) record sets, comprising 44 and 56 records, respectively, are considered. Details of these records can be found in (De Domenico et al., 2020b). The records were all preliminarily scaled to have a common peak ground acceleration (PGA) of 0.3 g, so as to eliminate the randomness related to the input intensity level.

The superstructure is assumed to be a regular idealized six-storey framed building, as shown in Figure 8, with floor mass $m_f = 100000\text{kg}$ and constant lateral stiffness $k_f = 270\text{MN/m}$ equal at all storeys, which results in a fundamental period $T_{s1} \approx 0.50\text{s}$. The base-isolation system comprises 16 friction pendulum (FP) isolators. The FPS characteristics are assumed as $\omega_b = \pi$ and $f_b = 0.04$. The base-isolation floor mass is assumed equal to the floor mass, so as to obtain a total mass of $M_{tot} = 700000\,\text{kg}$. Based on the previous optimization results, a stiffness ratio $\kappa = 0.8$ is assumed, along with a damping ratio of $\zeta_t = 0.1$. The practical implementation of the TFI is realized with 4 additional low-damping rubber bearings, 4 fluid inerters and a set of bi-directional rollers, as sketched in Figure 8. The TFI is attached to the isolation floor through a rigid steel frame, located underneath the low-damping rubber bearings and able to transfer the shear force from such auxiliary isolators to the foundation through the fluid inerter. The mass of the overall rigid steel frame is $m_t = 3500\text{ kg}$, based on a reasonable set of HEB100 profiles plus transversal stiffeners.

It is worth noting that this mass is just 0.5 % of the total mass of the base-isolated structure, which corresponds to a very low mass ratio $\mu_t = 0.005$. The advantage of the mass-amplification effect induced by the fluid inerter is indeed related to the possibility of realizing lightweight vibration absorbers, thus avoiding architectural issues due to the need of allocating large masses. Based on the optimal design procedure explained in Section 4, the minimization of the objective function FEI leads to inertance ratio $\beta_{\text{opt}} = 0.341$ and damping ratio $\zeta_{p,\text{opt}}^{NL} = 0.362$.

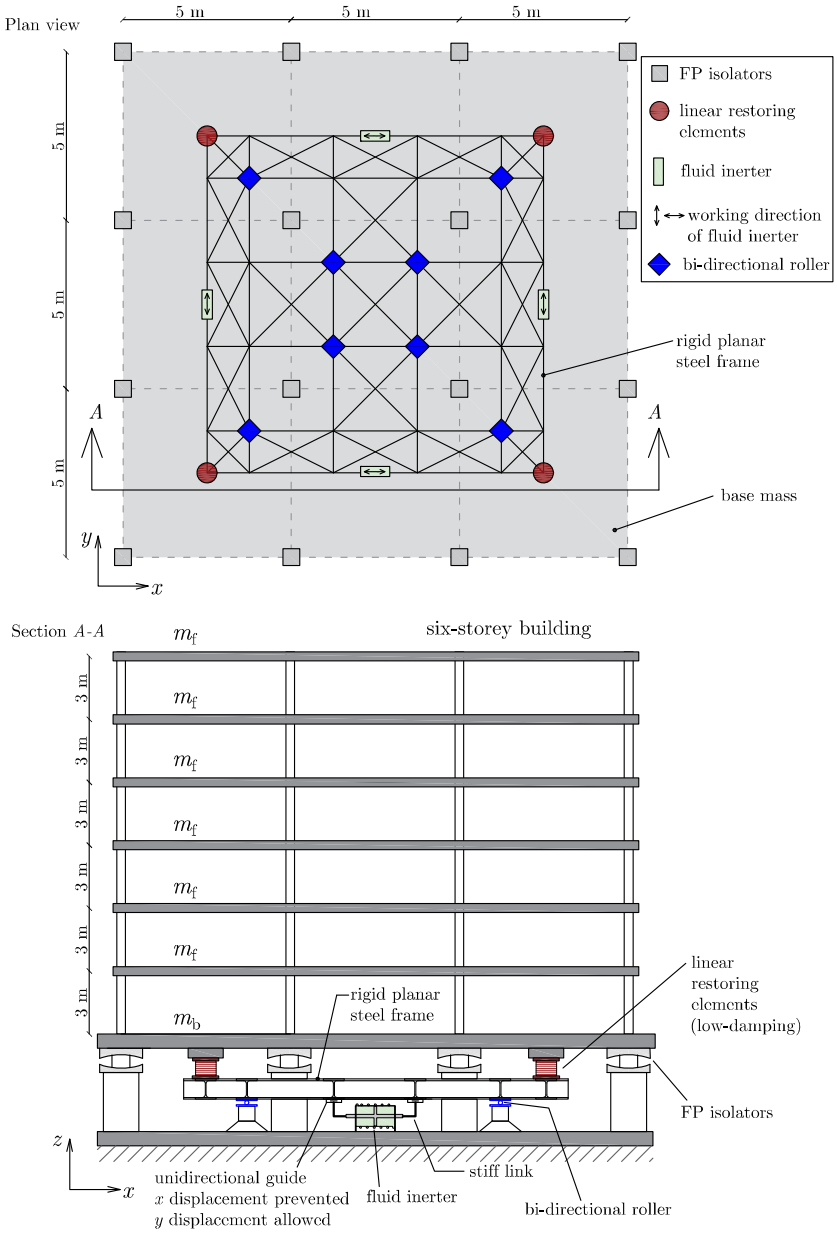


Figure 8. Six-storey framed building with proposed FPS-TFI system, plan view (top) and front view (bottom).

The seismic performance of the proposed FPS-TFI combined isolation system is compared to the classical base isolation system (FPS alone) in terms of displacements relative to the ground (normalized with respect to the top-floor displacement of the structure with FPS alone), interstory drift ratios (IDRs), assuming an interstorey height of 3.0 m, and absolute floor accelerations. To highlight the additional benefits of the combined inertance and damping properties of the fluid inerter in comparison to the classical mechanical inerter (TMDI scheme), the seismic performance of a FPS-TMDI coupled system (previously analyzed in 0) is also reported. Finally, a large-mass ratio TMD (with $\mu_t = 0.10$ in place of $\mu_t = 0.005$) coupled with the FPS is also reported in this comparative study.

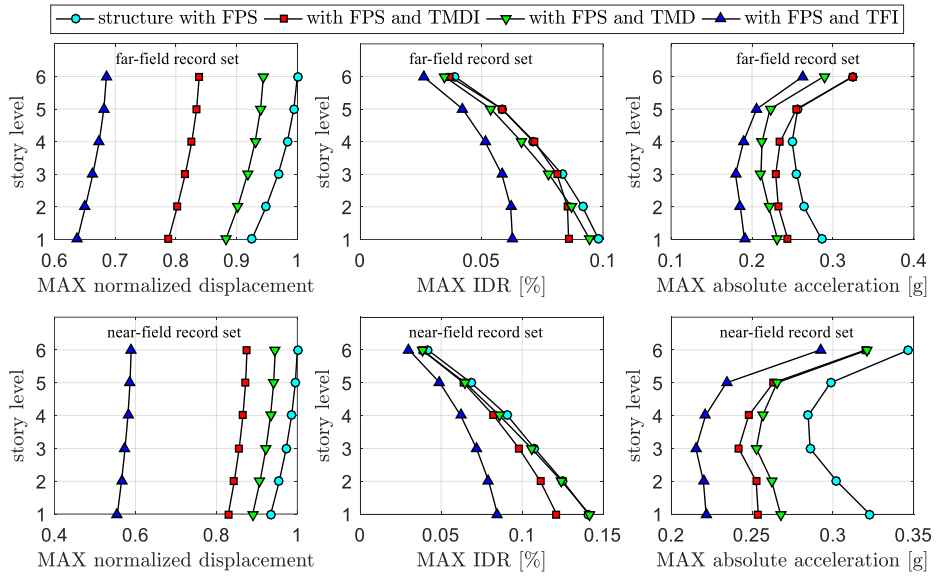


Figure 9. Average MAX displacements, interstorey drifts and absolute accelerations of benchmark structure with different structural control systems for both far-field and near-field FEMA P695 record sets.

Average maximum results (selectively considering the 44 records of the far-field set or the 56 records of the near-fault set of the FEMA P695 database) are illustrated in Figure 9. The FPS-TFI combined isolation system leads to a marked reduction of the response in comparison to the structure with FPS alone, not only of the isolators' displacements (of around 30% and 40% for FF and NF record sets, respectively), but also of the IDRs (of more than 30% for both FF and NF record sets) and of the absolute accelerations (of around 25% for both FF and NF record sets). Moreover, the FPS-TFI combined system is better able to mitigate the seismic response than the classical TMDI scheme, as well as than a large mass ratio TMD.

What is interesting to highlight is the significantly better performance of the FPS-TFI combined system under near-fault pulse-like earthquake events compared to both structure with FPS alone and structure with FPS-TMDI combined isolation system. To demonstrate this aspect, in Figure 10 the time histories of the last-floor displacement and of the last-floor absolute acceleration response under two near-fault pulse-like records, namely 1979 Imperial Valley RSN 182 and 1980 Irpinia RSN 292, are illustrated. As expected, it is noted that the strong motion phase is limited to within a very short time interval, which is typical of a pulse-like response. The inherent nonlinear damping features of the TFI are essential to mitigate the peak values

of the response in the strong motion phase, something that is not effectively achieved by the other two alternative structural control systems. In particular, the FPS-TFI system reduces the maximum displacement and the maximum absolute acceleration of around 50% and 30%, respectively, compared to the structure with FPS alone. The reduction percentages achieved by the classical TMDI scheme are, instead, relatively modest for this specific case of near-fault pulse-like events.

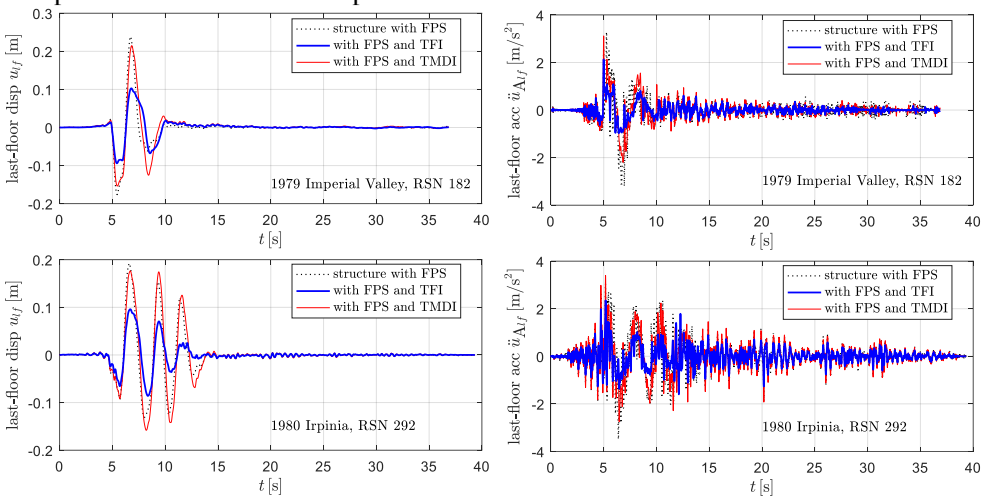


Figure 10. Last-floor displacement and absolute acceleration time histories under two near-fault pulse-like events.

6. Conclusions and critical remarks

This contribution has presented a feasibility study of the fluid inerter for applications to earthquake engineering. The fluid inerter mechanical principle is based on a piston-cylinder device that generates a movement of a fluid through an external helical channel, thus producing rotational inertia of a fluid mass. In this contribution, a so-called Tuned Fluid Inerter (TFI) scheme has been elaborated, comprising a grounded fluid inerter and some linear restoring elements that can be realized, for instance, with low-damping rubber bearings. This novel TFI scheme has been combined with a classical isolation system made of friction pendulum (FP) isolators. It has been pointed out that, unlike mechanical variants of the inerter, the assumption of linear inerter is not accurate for describing the hysteretic features of the fluid inerter. Instead, a nonlinear damping contribution, mainly related to the pressure drops occurring in the helical channels, should be introduced. Therefore, a simplified scheme of the fluid inerter proposed in this paper comprises a linear inerter in parallel with a nonlinear dashpot element. The effect of such inherent nonlinear damping contribution of the fluid inerter in relationship to earthquake protection purposes has been analyzed and discussed in this paper.

An optimal design procedure based on random vibration theory has been presented, aimed at minimizing a specifically selected objective function under the assumption of base acceleration modeled as a stationary Kanai-Tajimi zero-mean Gaussian random process. The nonlinear terms of the FP isolators and of the fluid inerter are incorporated in the optimal design procedure through statistical linearization technique. The influence of the soil conditions (firm, medium and soft soil), of the design strategy (displacement-, acceleration-, or energy-oriented design procedure), and of the stiffness characteristics of the low-damping rubber bearings has been analyzed. Then, the seismic performance of the proposed FPS-TFI combined isolation system has been assessed in the time domain considering an ensemble of 100

ground-motion records belonging to the FEMA P695 database. This database comprises both far-field and near-fault earthquake events for the sake of generality. It has been found that the proposed FPS-TFI combined isolation system significantly reduces the isolators' displacement demand, as well as the interstory drift and absolute acceleration response of the superstructure. Moreover, the fluid inerter is better able to mitigate the seismic response under near-fault pulse-like events than the classical mechanical inerter, due to the inherent damping contribution.

Based on the encouraging results of this paper, further investigation should be directed towards the actual experimental behavior of large-scale fluid inerter devices to assess the suitability of the proposed TFI for practical earthquake protection purposes. To the authors' best knowledge, only small-scale prototypes of the fluid inerter have been tested so far, whereas more reasonable dimensions and more realistic load scenarios relevant to earthquake engineering applications should be considered in laboratory tests, which is the object of an ongoing research.

References

De Domenico, D., Deastra, P., Ricciardi, G., Sims, N.D., and D.J. Wagg. Novel fluid inerter based tuned mass dampers for optimised structural control of base-isolated buildings. *Journal of the Franklin Institute*, 356(14), 7626-7649, 2019a.

De Domenico, D., Gandelli, E., and V. Quaglini. Effective base isolation combining low-friction curved surface sliders and hysteretic gap dampers, *Soil Dyn Earthq Eng*, 130, 105989, 2020a.

De Domenico, D., and G. Ricciardi. Earthquake-resilient design of base isolated buildings with TMD at basement: Application to a case study, *Soil Dyn Earth Eng*, 113, 503-521, 2018a.

De Domenico, D., and G. Ricciardi. Optimal design and seismic performance of tuned mass damper inerter (TMDI) for structures with nonlinear base isolation systems, *Earth Eng Struct Dyn*, 47(12), 2539-2560, 2018b.

De Domenico, D., and G. Ricciardi., An enhanced base isolation system equipped with optimal Tuned Mass Damper Inerter (TMDI). *Earthq Eng Struct Dyn*, 47, 1169-1192, 2018c.

De Domenico, D., Ricciardi, G., and I. Takewaki. Design strategies of viscous dampers for seismic protection of building structures: a review. *Soil Dynamics and Earthquake Engineering*, 118, 144-165, 2019b.

De Domenico, D., Ricciardi, G., and R. Zhang. Optimal design and seismic performance of tuned fluid inerter applied to structures with friction pendulum isolators, *Soil Dyn Earthq Eng*, DOI: 10.1016/j.soildyn.2020.106099, 2020b.

Di Matteo, A., Masnata, C., and A. Pirrotta. Simplified analytical solution for the optimal design of Tuned Mass Damper Inerter for base isolated structures. *Mechanical Systems and Signal Processing*, 134, 106337, 2019.

Giaralis, A., and F. Petrini. Wind-induced vibration mitigation in tall buildings using the tuned mass-damper-inerter. *J Struct Eng*, 143(9), 04017127, 2017.

Giaralis, A., and A.A. Taflanidis. Optimal tuned mass-damper-inerter (TMDI) design for seismically excited MDOF structures with model uncertainties based on reliability criteria. *Structural Control and Health Monitoring*, 25(2), e2082, 2018.

Glover, A.R., Smith, M.C., Houghton, N.E., and P.J.G. Long. Force-controlling hydraulic device (International Patent Application No: PCT/GB2010/001491, 2009).

Gonzalez-Buelga, A., Lazar, I.F., Jiang, J.Z., Neild, S.A., and D.J. Inman. Assessing the effect of nonlinearities on the performance of a tuned inerter damper. *Struct Control Health Monit*, 24(3), e1879, 2017.

Hashimoto, T., Fujita, K., Tsuji, M., and I. Takewaki. Innovative base-isolated building with large mass-ratio TMD at basement for greater earthquake resilience, *Future Cities and Environment*, 1:9, 1-19, 2015.

Ikago, K. , Saito, K., and N. Inoue. Seismic control of single-degree-of-freedom structure using tuned viscous mass damper, *Earthq Eng Struct Dyn*, 41(3), 453-474, 2012.

Javidialesaadi, A., and N.E. Wierschem. An inerter-enhanced nonlinear energy sink. *Mechanical Systems and Signal Processing*, 129, 449-454, 2019.

Jiang, Y., Zhao, Z., Zhang, R., De Domenico, D., and C. Pan. Optimal design based on analytical solution for storage tank with inerter isolation system. *Soil Dyn Earthq Eng*, 129, 105924, 2020.

Lazar, I.F., Neild, S.A., and D.J. Wagg. Using an inerter-based device for structural vibration suppression. *Earthq Eng Struct Dyn*, 43, 1129-1147, 2014.

Lazar, I.F., Neild, S.A., and D.J. Wagg. Vibration suppression of cables using tuned inerter dampers. *Eng Struct*, 122: 62-71, 2016.

Luo, H., Zhang, R., Weng, D. Mitigation of liquid sloshing in storage tanks by using a hybrid control method. *Soil Dyn Earth Eng*, 90, 183-195, 2016.

Málaga-Chuquitaype, C., Menendez-Vicente, C., and R. Thiers-Moggia. Experimental and numerical assessment of the seismic response of steel structures with clutched inerters. *Soil Dyn Earth Eng*, 121, 200-211, 2019.

Pietrosanti, D., De Angelis, M., and M. Basili. Optimal design and performance evaluation of systems with Tuned Mass Damper Inerter (TMDI). *Earthq Eng Struct Dyn*, 46(8), 1367-1388, 2017.

Reggio, A., and M. De Angelis. Optimal energy-based seismic design of non-conventional Tuned Mass Damper (TMD) implemented via inter-story isolation. *Earth Eng Struct Dyn*, 44(10), 1623-1642, 2015.

Smith, M.C. Synthesis of mechanical networks: the inerter. *IEEE Transactions on Automatic Control*, 47(10), 1648-1662, 2002.

Smith, N., and D. Wagg. A fluid inerter with variable inertance properties, In: *6th European Conference on Structural Control (EACS2016)*, Sheffield, UK, 11-13 July 2016, paper no. 199.

Swift, S.J., Smith, M.C., Glover, A.R., Papageorgiou, C., Gartner, B., and N.E. Houghton. Design and modelling of a fluid inerter. *Int J Control* 86(11), 2035-2051, 2013.

Takewaki, I., Murakami, S., Fujita, K., Yoshitomi, S., and M. Tsuji. The 2011 off the Pacific coast of Tohoku earthquake and response of high-rise buildings under long-period ground motions, *Soil Dyn Earthq Eng*, 31(11), 1511–1528, 2011.

Takewaki, I., Murakami, S., Yoshitomi, S., and M. Tsuji. Fundamental mechanism of earthquake response reduction in building structures with inertial dampers. *Struct Control Health Monit*, 19, 590-608, 2012.

Taniguchi, T., Der Kiureghian, A., and M. Melkumyan. Effect of tuned mass damper on displacement demand of base-isolated structures, *Eng Struct*, 30, 3478-3488, 2008.

Wang, Q., Qiao, H. , De Domenico, D., Zhu, Z., and Z. Xie. Wind-induced response control of high-rise buildings using inerter-based vibration absorbers. *Applied Sciences*, 9(23), 5045, 2019.

Zargar, H., Ryan, K.L., and J.D. Marshall. Feasibility study of a gap damper to control seismic isolator displacements in extreme earthquakes. *Struct Control Health Monit*, 20(8), 1159–1175, 2013.

Zhao, Z., Zhang, R., Jiang, Y., and C. Pan. Seismic response mitigation of structures with a friction pendulum inerter system. *Engineering Structures*, 193, 110-120, 2019.

Stochastic Response of Beams Equipped with Tuned Mass Dampers, considering Spring Inertial Effects, Subjected to Poissonian Loads

I. P. Dunn[1], A. Di Matteo[1], G. Failla[2], and A. Pirrotta[1]
1) Dipartimento di Ingegneria, Università degli Studi di Palermo, Palermo, 90128, Italy; IainPeter.Dunn@unipa.it, Alberto.dimatteo@unipa.it, Antonina.Pirrotta@unipa.it
2) Dipartimento di Ingegneria Civile, dell'Energia, dell'Ambiente e dei Materiali Università degli Studi Mediterranea di Reggio Calabria, Reggio Calabria, 89124, Italy; giuseppe.failla@unirc.it

Abstract: This contribution considers the dynamic response of Euler-Bernoulli beams equipped with multiple tuned mass dampers, subjected to random moving loads following a Poissonian. The method proposed for the mathematical solution utilises the theory of generalised functions to find the response variables at the exact locations of the discontinuities, in this case, tuned mass dampers in which the distributed mass of the spring is also considered. This involves deriving exact, complex eigenvalues and eigenfunctions from a characteristic equation built as the determinant of a (4 x 4) matrix as oppose to the classical method requiring an (n+1 x n+1) for n number of attachments. This is always the case, in the proposed method, regardless of the number or type of attachments. Orthogonality conditions are then built for the eigenfunctions and then, the stochastic response of the beam under Poissonian loading is evaluated. To show the applicability and accuracy of the proposed procedure, a numerical application is presented in which a beam with multiple tuned mass dampers, acted upon by random moving loads, in the form of a filtered Poissonian process. The means and standard deviations at the midpoint are presented for various beam models fitted with different numbers of tuned mass damper attachments, these are found by employing a Monte- Carlo simulation.

Keywords: Euler-Bernoulli Beam; Poissonian Loading; Tuned Mass Damper; Modal Analysis.

1. Introduction

The Euler-Bernoulli model of slender beams is commonly used to calculate the dynamic response of rail and road bridges to any type of loading and it has been shown to provide extremely accurate results when validated by the finite element method (FEM). This paper will present an extension to the novel mathematical method used to find the response of Euler-Bernoulli beams under Poisson type loading. Dynamic analysis is a field which has been understood since the early part of the 20th century, however it was not until relatively recently that the dynamics of structures were given the importance that they deserve in the design and construction of structures. This change in tactics was necessitated by a number of engineering disasters which forced design standards to be updated and made more robust and it was facilitated thanks to advances in computer technology allowing increasingly complex problems to be modelled and solved. Significant advances in the transport industry, such as automation and the increasing interest in high speed rail projects, have again highlighted the importance of dynamic analysis in structures which can be typically modelled as a discontinuous Euler-Bernoulli beam.

Poissonian loading of the type presented in this paper considers various moving loads traversing the full span of the beam, by analysing the effects of moving loads on this mathematical bridge model, a greater understanding of the dynamic response which we can expect to see in real world applications, can be obtained.

© 2021 *by I. P. Dunn, A. Di Matteo, G. Failla, and A. Pirrotta.*

This type of forcing action is used because beams subjected to moving loads have greater maximum deflections and maximum bending moments than beams which are subjected to static loads, the results should therefore, provide a more accurate picture of the actual stresses under which a bridge will suffer.

As bridges are which are subjected to higher dynamic forces from increasingly fast locomotives, this analysis takes on a greater importance due to the danger of the locomotive reaching the beam's critical velocity; at this velocity, serious damage to the beam can be sustained.

These high-speed rail projects, which are seen as significantly important to reducing carbon emissions by functioning as a viable alternative to short haul intercity flights, combined with the aging infrastructure which is present all over Europe posing a barrier to the implementation of these projects. Thorough analysis of current infrastructure combined with the addition of adequate damping to mitigate the effects of the increased forces, could reduce costs of these projects by removing the need for renewing and rebuilding existing infrastructure.

While most current studies focus on deterministic solutions to known loading, a number of studies consider the effect that random loading and a series of random moving loads have on the beams subjected to them. In many studies concerning the stochastic response of Euler-Bernoulli beams, the random loading applied follows a number of standard loading patterns, a Gaussian distribution (Di Paola et al., 1992), typical white noise (Pirrotta, 2005) a Poissonian distribution (Ricciardi, 1994), “constituted by a train of impulses of random amplitudes occurring at random time instants” (Di Matteo et al., 2016) giving a stochastic output due to these random elements (Bilello et al., 2002). Traffic loading (particularly considering road traffic, although this is also applicable to rail traffic) can be described as a Poissonian process as the magnitude of the forces is random as are their arrival times (Bucher and Di Paola, 2015).

Ricciardi (1994) proposed a method of modelling the forcing action as a filtered Poisson process, this is obtained by finding "the response of a linear undamped oscillator excited by a Poisson white noise process"
The dynamic response of a Euler-Bernoulli beam can only be accurately obtained by a very small number of methods, namely: computer models such as the finite element method (FEM), or the classical numerical method. These methods however each have significant drawbacks; the FEM is accurate at lower modes but as the number of modes under consideration increases so too does the percentage error in the eigenvalues and eigenfunctions obtained. This is also true as the number of spans increases, a beam with a high number of spans cannot be as effectively modelled by the FEM. The classical numerical method (CM) does not have this limitation, it provides exact eigenvalues and eigenfunctions regardless of the number of spans or the number of modes considered; the drawback with this method however is the large amount of computational time required to calculate the solution, as the number of spans increases, the complexity of the matrices contained in the solution of the CM grows and the required computing power increases exponentially.

This paper will expand on the work conducted by Adam et al. (2017) in which a novel numerical method was developed to find the response to a series of moving loads acting on a Euler-Bernoulli beam equipped with tuned mass dampers, this method can consider, as the CM does, complex eigenfunctions caused by damping elements, in this case from the tuned mass dampers, where the damping is localised at a specific point rather than distributed proportionally throughout the entire length of the beam as is often assumed. The method is then extended to consider a series of moving loads following a Poissonian distribution where they have random magnitudes and random arrival times, the results obtained should, therefore, more closely reflect the response of a road bridge subjected to normal traffic loading.
Further to this, in recent years, work has been undertaken to refine the mathematical model in the aim to obtain more accurate results which are closer to real-world systems. This refinement has taken the form of removing assumptions which are present in the mathematical models.

A TMD is most commonly modelled as a subsystem comprising of a lumped mass attached to the primary structural system by a spring and, in general, a viscous dashpot connected in parallel. In this standard model of a TMD, only three characteristics are generally considered, the spring stiffness, the damping coefficient and the magnitude of the lumped mass (Failla et al., 2019). It is generally assumed that the mass of the dashpot and spring are small enough in comparison to the total lumped mass to be negligible and as such, the inertial effects along the length of the spring, are neglected. This allows a major simplification of the problem by assuming a massless spring and only considering the inertial effects of the lumped mass.

In real-world applications however, this is not necessarily true as a lumped mass comprising of the total system weight will undoubtedly act differently from a distributed mass in dynamic analysis. As such, a number of studies were undertaken which aimed to better define the analytical model by including the effects from the spring's internal inertia. The model presented by Failla et al. (2019) is one such study which aimed to remove this assumption by considering the spring element's distributed mass and its effect on the beam-TMD system dynamics.

2. The Multi-Span Euler-Bernoulli Beam Problem

One of the simplest cases of a multi-span Euler-Bernoulli beam is a two-span beam fitted with one spring-mass attachment at the midpoint, this is a simplified model of a tuned mass damper (TMD) in which no dashpot is present to provide viscous damping. This allows the description of the proposed method without the introduction of too many terms thus, allowing the proposed model to retain a sufficient level of clarity. In this case however, the mathematical model of the simplified TMD is slightly more complex than that of the standard simplified TMD as the effect of the stiffness element's mass will also be considered. Typically, the spring is considered to be massless (Dunn et al., 2019), it should be clear, however, that the effects of the distributed mass will affect the dynamics of both the TMD and, by extension, the continuous beam itself. It should be emphasised however, that using the method proposed in this paper, any number of TMDs could be fitted at any point in the beam's span without changing the steps required to find the equation of motion (EoM), this will become clear.

Using the proposed formulation, supports, lumped masses, TMDs, and, indeed, any other attachment which does not cause localised rotation can be modelled as a shear discontinuity acting on a specific point. Shear discontinuities could otherwise be described as point forces located at the point of attachment of the TMD, this allows the EoM to take the form (Di Matteo et al., 2020)

$$EI\frac{\tilde{\partial}^4 w(x,t)}{\partial x^4}+\overline{m}\frac{\partial^2 w(x,t)}{\partial t^2}+R(x,t)=f(x,t) \tag{1}$$

where: the "tilde" represents the generalised derivative, EI is the flexural rigidity, $\overline{m}$ is the mass per unit length, $w(x,t)$ is the dynamic response of the beam in terms of the transversal displacement in the space and time domains, $R(x,t)$ is a generalised function (Biondi and Caddemi, 2007) used to account for the discontinuities present, in this case one single TMD and $f(x,t)$ is the forcing action. As the generalised function represents a single TMD, $R(x,t)$ takes the form Falsone (2002):

$$R(x,t) = -\sum_{j=1}^{N} P_j(t)\ \delta(x - x_j) \tag{2}$$

where: $P_j(t)$ is the shear reactionary force, a point load, and $\delta(x - x_j)$ is a Dirac's delta function positioning the reactionary force at x_j, the location of the TMD.

1.1. FREE VIBRATION

In free vibration, the left-hand side of the EoM is equal to zero and the EoM takes the form:

$$EI\frac{\tilde{\partial}^4 w(x,t)}{\partial x^4} + \bar{m}\frac{\partial^2 w(x,t)}{\partial t^2} + R(x,t) = 0 \tag{3}$$

At this point the separable variables approach is applied, wherein the time and space domains are considered independently:

$$w(x,t) = \psi(x)\, g(t) \tag{4}$$

where $g(t)$ can also be defined as: $\boldsymbol{g(t) = e^{i\omega t}}$

The space domain term $\psi(x)$ and the time domain term $g(t)$ are split; allowing for the solution of the beam's eigenvalues and eigenfunctions, however, the reactionary forces cannot be handled in this manner as the point force caused by the discontinuity is in the time domain only. $R(x,t)$ could easily be transferred into just the time and space domains by splitting $P_j(t)$ from $\delta(x - x_j)$, to retain the forcing term, $P_j(t)$ must be Fourier transformed (Failla, 2016) leading to the frequency dependent term:

$$P(x,\omega) = -\sum_{j=1}^{N} \varphi_j(\omega)\,\delta(x - x_j) \tag{5}$$

where $\delta(x - x_j)$ is a Dirac's delta function specifying the application point of the force as before and (Di Lorenzo et al., 2017):

$$\varphi_j(\omega) = -K_{TMD\,j}(\omega)\ \psi(x_j) \tag{6}$$

where $\psi(x_j)$ is the eigenfunction of deflection at the point x_j, and $K_{TMD\,j}(\omega)$ is the frequency dependent stiffness of the TMD given by the Fourier transform of the term in the time domain (Failla et al., 2019):

$$K_{TMD\,j}(\omega) = \sqrt{EA_{spring\ j}}\sqrt{\rho_{spring\ j}}\ \omega\left(\frac{\omega\ \mathrm{M}_{TMD\ j}\,\mathrm{Cos}\left(\frac{h_j\omega\sqrt{\rho_{spring\ j}}}{\sqrt{EA_{spring\ j}}}\right) + \sqrt{EA_{spring\ j}}\sqrt{\rho_{spring\ j}}\,\mathrm{Sin}\left(\frac{h_j\omega\sqrt{\rho_{spring\ j}}}{\sqrt{EA_{spring\ j}}}\right)}{-\sqrt{EA_{spring\ j}}\sqrt{\rho_{spring\ j}}\,\mathrm{Cos}\left(\frac{h_j\omega\sqrt{\rho_{spring\ j}}}{\sqrt{EA_{spring\ j}}}\right) + \omega\ \mathrm{M}_{TMD\ j}\,\mathrm{Sin}\left(\frac{h_j\omega\sqrt{\rho_{spring\ j}}}{\sqrt{EA_{spring\ j}}}\right)}\right) \tag{7}$$

where: $M_{TMD\,j}$ is the magnitude of the lumped mass of the j[th] TMD, $\sqrt{EA_{spring\ j}}$ the spring stiffness of the j[th] TMD's spring, $\sqrt{\rho_{spring\ j}}$ is the distributed mass of the spring which forms part of the j[th] TMD, h_j is the spring length, and ω is the frequency.

3. Eigensolution

The exact modes of vibration can be found by applying the separable variables method, equation 4, allowing the transversal displacement $w(x,t)$, rotation $\theta(x,t)$, bending moment $m(x,\tau)$, and shear $Q(x,t)$, to be expressed in dimensionless form as:

$$w(x,t)=\psi(x)e^{i\omega t}\ ;\ \theta(x,t)=\vartheta(x)e^{i\omega t};\quad m(x,\tau)=\mu(x)e^{i\omega\tau}\ ;\ Q(x,t)=\chi(x)e^{i\omega t} \tag{8}$$

This gives the four eigenfunctions of the response variables $\psi(x)$ deflection, $\vartheta(x)$ rotation, $\mu(x)$ bending moment, and $\chi(x)$ shear. These eigenfunctions are related through a derivative method where:

$$\begin{aligned}
&\vartheta(x)=\frac{\tilde{d}\psi(x)}{dx};\\
&\mu(x)=-\frac{\tilde{d}\vartheta(x)}{dx};\\
&\chi(x)=\frac{\tilde{d}\mu(x)}{dx};\\
&\frac{\tilde{d}\chi(x)}{dx}+\sum_{j=1}^{N}\varphi_j(\omega)\ \delta(x-x_j)+\sigma^2\psi(x)=0.
\end{aligned} \tag{9 a-d}$$

From these relations, the free vibration of the beam in the space domain can then be expressed in terms of the first eigenfunction:

$$\frac{\tilde{d}^4\psi(x)}{dx^4}+P(x,\omega)-\sigma^2\psi(x)=0 \tag{10}$$

where $\sigma^2=(\omega^2\overline{m}L^4)/EI$.

Eq. (7) shows that the reactionary force of the TMD attached at point x_j depends on the frequency dependent term concerning the spring stiffness, the damping coefficient and the attached mass.

In Eq. (5) the term $\varphi_j(\omega)$ representing the frequency dependent stiffness based on the TMD's parameters and the deflection at x_j is an unknown due to the presence of the eigenfunction of deflection. Therefore, the matrix approach must be applied:

$\mathbf{Y}(x)$ is a vector built from the response variables of the eigenfunctions:

$$\mathbf{Y}(x) = \begin{bmatrix} \psi(x) & \vartheta(x) & \mu(x) & \chi(x) \end{bmatrix}^T \tag{11}$$

Failla (2014) proposed the following method to find the unknown $\varphi_j(\omega)$ in a closed form. Through the solution of a linear function of the 4×1 vector $\boldsymbol{c}$ which is constructed from the four integration constants which are found from the solution of the homogeneous equation a closed form expression can be obtained. This leads to the following closed analytical expression of $\mathbf{Y}(x)$:

$$\mathbf{Y}(x) = \tilde{\mathbf{Y}}(x)\boldsymbol{c} \tag{12}$$

where $\tilde{\mathbf{Y}}(x)$ is a 4×4 matrix given by the standard solution of a fourth order, homogeneous, Euler-Bernoulli beam equation:

$$\tilde{\mathbf{Y}}(x) = \boldsymbol{\Omega}(x) + \sum_{j=1}^{N} \mathbf{J}(x, x_j)\,\varphi_j(\omega) \tag{13}$$

where and each row represents an individual eigenfunction from the homogeneous solution: deflection, rotation, bending moment, and shear respectively for rows 1 through 4:

$$\boldsymbol{\Omega}(x) = \begin{bmatrix} \Omega_{\psi_1} & \Omega_{\psi_2} & \Omega_{\psi_3} & \Omega_{\psi_4} \\ \Omega_{\vartheta_1} & \Omega_{\vartheta_2} & \Omega_{\vartheta_3} & \Omega_{\vartheta_4} \\ \Omega_{\mu_1} & \Omega_{\mu_2} & \Omega_{\mu_3} & \Omega_{\mu_4} \\ \Omega_{\chi_1} & \Omega_{\chi_2} & \Omega_{\chi_3} & \Omega_{\chi_4} \end{bmatrix} \tag{14}$$

In the interest of clarity, the matrix $\boldsymbol{\Omega}(x)$ is expanded below to show the terms contained in the general solution of the homogeneous equation and the derivative method used to relate the eigenfunctions throughout the subsequent rows:

$$\begin{array}{llll} \Omega_{\psi_1}(x) = e^{-\sigma x} & \Omega_{\psi_2}(x) = e^{\sigma x} & \Omega_{\psi_3}(x) = \cos(\sigma x) & \Omega_{\psi_4}(x) = \sin(\sigma x) \\ \Omega_{\vartheta_1}(x) = -\sigma e^{-\sigma x} & \Omega_{\vartheta_2}(x) = \sigma e^{\sigma x} & \Omega_{\vartheta_3}(x) = -\sigma \sin(\sigma x) & \Omega_{\vartheta_4}(x) = \sigma \cos(\sigma x) \\ \Omega_{\mu_1}(x) = -\sigma^2 e^{-\sigma x} & \Omega_{\mu_2}(x) = -\sigma^2 e^{\sigma x} & \Omega_{\mu_3}(x) = \sigma^2 \cos(\sigma x) & \Omega_{\mu_4}(x) = \sigma^2 \sin(\sigma x) \\ \Omega_{\chi_1}(x) = \sigma^3 e^{-\sigma x} & \Omega_{\chi_2}(x) = -\sigma^3 e^{\sigma x} & \Omega_{\chi_3}(x) = -\sigma^3 \sin(\sigma x) & \Omega_{\chi_4}(x) = \sigma^3 \cos(\sigma x) \end{array}. \tag{15}$$

The discontinuities are considered (as shown in equation 13) through the vector $\mathbf{J}(x, x_j)$ which can be expanded as (Failla et al., 2019):

$$\mathbf{J}(x, x_j) = \begin{bmatrix} J_{\psi}^{(p)} & J_{\vartheta}^{(p)} & J_{\mu}^{(p)} & J_{\chi}^{(p)} \end{bmatrix}^T \tag{156}$$

At this point, the boundary conditions at the extremes of the beam are enforced:

$$\mathbf{B}\,\mathbf{c} = \mathbf{0} \tag{17}$$

Where **B** is a 4 x 4matrix constructed from enforcing the boundary conditions on the matrix $\tilde{\mathbf{Y}}(x)$ and as before **c** is a 4 x 1 vector of the unknown constants.

From here, the characteristic equation can then be built as the determinant of the 4×4 matrix $\mathbf{B}$:

$$\det(\mathbf{B}) = 0 \tag{18}$$

Then, the non-trivial solutions of $\mathbf{c}$ are found and exact closed form expressions can be built for the beam's eigenfunctions. Due to the presence of a dashpot in the TMD model, localised damping will be present and therefore, the eigenfunctions will be complex as will the mode shapes.

4. Orthogonality Conditions

Following the procedure presented in Oliveto et al. (1997), the orthogonality conditions are built to derive the particular impulse response function of this beam.

Firstly the EoM in free vibration in the form shown below is considered:

$$\frac{\tilde{d}^4\psi_m(x)}{dx^4} - \sigma_m{}^2\psi_m(x) + \sum_{j=1}^{N} K_{TMD\ j}(\omega_m)\psi_m(x_j) = 0 \tag{169}$$

Considering modes m and n, multiplying the EoM at mode m by $\psi_n(x)$ and at mode n by $\psi_m(x)$ and then integrating between 0 and L with respect to x:

$$\int_0^L \frac{\tilde{d}^2\psi_m(x)}{dx^2}\frac{\tilde{d}^2\psi_n(x)}{dx^2}dx - \sigma_m{}^2\int_0^L \psi_{mn}(x)dx + \sum_{j=1}^{N} K_{TMD\,j}(\omega_m)\psi_{mn}(x_j) = 0 \tag{20}$$

where: $\psi_{mn}(x) = \psi_m(x)\psi_n(x)$

$$\int_0^L \frac{\tilde{d}^2\psi_n(x)}{dx^2}\frac{\tilde{d}^2\psi_m(x)}{dx^2}dx - \sigma_n{}^2\int_0^L \psi_{mn}(x)dx + \sum_{j=1}^{N} K_{TMD\,j}(\omega_n)\psi_{mn}(x_j) = 0 \tag{21}$$

Integrating by parts and subtracting Eq. (21) from Eq. (20) then yields the first orthogonality condition:

$$(\sigma_m{}^2 - \sigma_n{}^2)\int_0^L \psi_{nm}(x)dx + \sum_{j=1}^{N}[K_{TMD\ j}(\omega_n) - K_{TMD\ j}(\omega_m)]\psi_{nm}(x_j) = 0 \tag{22}$$

The second orthogonality condition is then found by multiplying Eq. (20) by σ_n and Eq. (21) by σ_m and then subtracting Eq. (21) from Eq. (20):

$$\sigma_{m-n}\int_0^L \frac{\partial^2\psi_{mn}(x)}{\partial x^2}dx + \sigma_m\sigma_n(\sigma_{m-n})\int_0^L \psi_{nm}(x)dx + [\sigma_m K_{TMD\ j}(\omega_n) - \sigma_n K_{TMD\ j}(\omega_m)]\psi_{nm}(x_j) = 0 \tag{23}$$

where $\frac{\partial^2\psi_{mn}(x)}{\partial x^2} = \frac{\partial^2\psi_m(x)}{\partial x^2}\frac{\partial^2\psi_n(x)}{\partial x^2}$ and $\sigma_{m-n} = (\sigma_m - \sigma_n)$

5. Forced Vibrations

These orthogonality conditions are then used to derive the beam's response to arbitrary loading. This is accomplished by using the complex modal superposition principle as defined by (Di Lorenzo et al., 2017) where the complex modal impulse response function is used. This leads to (Adam et al., 2017):

$$w(x,t) = \sum_{k=1}^{\infty} \psi_k(x) \frac{1}{i\,\Xi_k\,\omega_k} \int_0^t f(\tau)\, e^{i\omega_k(\tau - t)} d\tau \tag{24}$$

where: $f(\tau)$ is the moving load dependent term and Ξ_k is the effect that the beam's mass and the attached TMD have on the beam's response:

$$\Xi_k = 2\int_0^L \bar{m}(x)\left[\psi_k(x)\right]^2 dx + \sum_{j=1}^{N} TMD_j \left[\psi_k(x_j)\right]^2 \tag{25}$$

where

$$\mathrm{TMD_j} = \left(\frac{2\,\beta\,\mathrm{M}_{TMD\,\mathrm{j}}\,\omega_k \mathrm{Cos}(2\alpha) + \sqrt{\beta}\,\mathrm{Sin}(2\alpha)(-\beta + \omega_k{}^2\mathrm{M}_{TMD\,\mathrm{j}}{}^2) - 2\,\rho_{spring\,\mathrm{j}}\,\omega_k(EA_{spring\,\mathrm{j}}\,\mathrm{M}_{TMD\,\mathrm{j}} + h_j(\beta + \omega_k{}^2\mathrm{M}_{TMD\,\mathrm{j}}))}{2\,\omega_k\left(\sqrt{\beta}\,\mathrm{Cos}(\alpha) - \omega_k\,\mathrm{Sin}(\alpha)\mathrm{M}_{TMD\,\mathrm{j}}\right)^2} \right) \tag{26}$$

where $\alpha = \left(\frac{h_j \omega_k \sqrt{\rho_{spring\,\mathrm{j}}}}{\sqrt{EA_{spring\,\mathrm{j}}}} \right)$ and $\beta = EA_{spring\,\mathrm{j}}\,\rho_{spring\,\mathrm{j}}$.

For a moving load, the response equation takes the following form:

$$w(x,t) = \sum_{k=1}^{\infty} \left(\psi_k(x) \int_0^t e^{i\omega_k(t-\tau)} d\tau \right) \frac{\int_0^L \psi_k(x)\,\delta(x - V_0\,\tau)\,dx}{i\,\Xi_k\,\omega_k} \tag{27}$$

where: $\delta(\cdot)$ is a Dirac's delta function and V_0 is the velocity of the load.

Further, when considering multiple loads traversing the beam, the effects of preceding loads must also be accounted for (Dunn et al., 2019):

$$w(x,t) = \sum_{k=1}^{\infty} \psi_k(x) \frac{\int_0^L \psi_k(x)\,\delta(x - V_0\,\tau)\,dx}{i\,\Xi_k\,\omega_k} \left(\int_{\tau_L^0}^t e^{i\omega_k(t-\tau)} d\tau - \int_{\tau_L^E}^t e^{i\omega_k(t-\tau)} d\tau \right) \tag{28}$$

where: τ_L^0 and τ_L^E denote the start and end times of the L$^{\text{th}}$ load.

Due to the presence of complex conjugate pairs, Eq. (28) can revert to the following real form (Failla et al., 2019):

$$w(x,t) = 2\,\mathrm{Re}\left[\sum_{k=1}^{\infty} \psi_k(x) \frac{\int_0^L \psi_k(x)\,\delta(x - V_0\,\tau)\,dx}{i\,\Xi_k\,\omega_k} \left(\int_{\tau_L^0}^t e^{i\omega_k(t-\tau)} d\tau - \int_{\tau_L^E}^t e^{i\omega_k(t-\tau)} d\tau \right) \right] \tag{29}$$

6. Poissonian Loading

A Poissonian white noise process is a type of delta-correlated process (Di Paola et al., 1995), which is "simple, robust and gives accurate results" (Di Paola and Ricciardi, 1992) when used to model loading caused by free flowing traffic, i.e. traffic unencumbered by an unusually heavy volume causing a continuous series of moving loads due to jams and tailbacks. Poissonian processes are most commonly defined as (Ricciardi, 1994):

$$S_P(t) = \sum_{P=1}^{N(t)} Y_P \delta(t - T_P) \tag{30}$$

where $N(t)$ is a counting function giving the number of impulses in the time interval (0,t), Y_P is the random amplitude (Di Paola and Pirrotta, 1999) of the forcing action, and $\delta(t - T_P)$ is a series of Dirac delta impulses Furtmüller et al. (2019) occurring at independent random times (T_P) following a Poissonian distribution.

When considering a moving load, this characterisation of the Poissonian load must be altered, it is also assumed that the loads will have a constant and equal velocity:

$$S_P(t) = \sum_{P=1}^{N(t)} Y_P \delta\left[x - \left(t - T_P\right) V_0 \right] W(t - T_P, t_L) \tag{31}$$

where $\delta\left[x - \left(t - T_P\right) V_0 \right]$ is a modification of the Dirac delta function from equation (30) in which moving loads arriving at random times with random amplitudes are considered, t_L is the time taken for the load to traverse the beam, length divided by the velocity of the moving load, L / V_0 and $W(t - t_p, t_L)$ is a window function which removes the force after it has traversed the beam; here $U(\cdot)$ is a unit step function: $W(t - T_P, t_L) = U(\tau)[1 - U(\tau - t_L)]$.

Following the method proposed by Ricciardi (1994), and Di Paola and Ricciardi (1992) the Poisson process is filtered to ensure that it is applicable to the beam's characteristics, this filtering causes Eq. (31) to take the following form:

$$S_k\left(t\right) = \sum_{P=1}^{N(t)} Y_P \psi_k\left(V_0\, \tau\right) W(t - T_P, t_L) \tag{32}$$

Substituting this into the original EoM gives:

$$EI \frac{\tilde{\partial}^4 w(x,t)}{\partial x^4} + \bar{m} \frac{\partial^2 w(x,t)}{\partial t^2} + R(x,t) = \sum_{\mathrm{P}\,=\,1}^{N(t)} Y_P \psi_k\left(V_0\ t\right) W\left(t - T_P, t_L\right) \tag{33}$$

Considering the theory of separable variable, this can also take the following form in the time domain:

$$\ddot{g}_k\left(t\right) + {\omega_k}^2 g_k(t) = \frac{2}{\Xi_k} S_k(t) \tag{34}$$

where $S_k(t)$ is the random forcing action at the k$^{\text{th}}$ mode, this takes the form:

$$S_k(t) = \sum_{P=1}^{N(t)} Y_P \psi_k(V_0 \tau) W(\tau, t_L) \tag{35}$$

7. Numerical Application

In this section a brief numerical application of the method presented in this paper. A bare beam, a beam fitted with a standard TMD and a beam fitted with a TMD in which the spring mass is considered are subjected to Poissonian loading and their mean and standard deviation responses in terms of midspan displacement are compared.

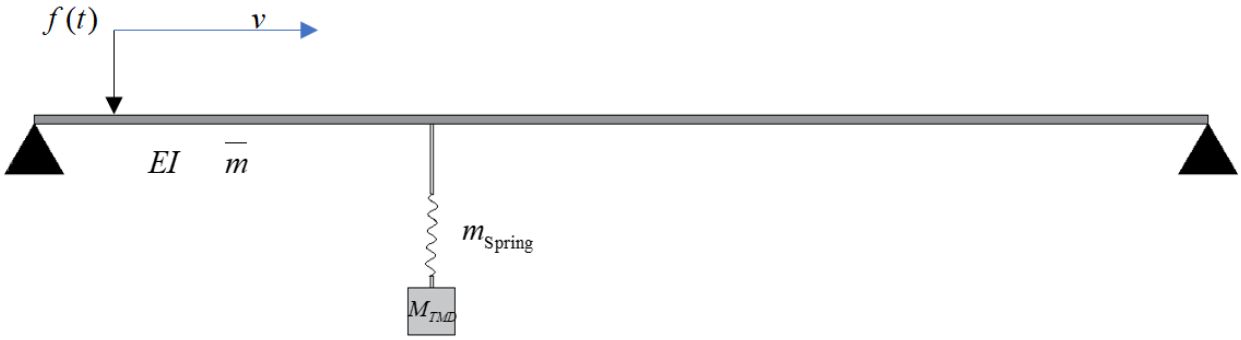


Figure 1. Euler-Bernoulli beam fitted with a TMD.

Figure 1 shows the beam configuration with one spring-mass attachment. The beam has a total length, L, of 6 m, a cross-sectional area, A, of 0.0625 m^2, the Young's modulus, E, is equal to 210 GPa, the second moment of area , $I = 0.000326\ m^4$, the mass per unit length $\overline{m} = 487.5\ Kg$, and the density $\rho = 7800\ kg/m^3$. The TMD tuning parameters selected for both the traditional and helical models are identical, however it should be noted that their distribution of masses will vary, both TMD models are attached at one third of the beam length, 2 m. In the traditional model the spring stiffness, k_{TMD} = 13.125 MN/m and the lumped mass, M_{TMD} = 812.5 kg were selected based on the mechanical characteristics chosen in the helical spring model in which the spring had a length, h = 1 m, stiffness, given by the Young's modulus multiplied by the cross sectional area, EA = 13.125 MN/m, a lumped mass kg and a distributed mass, m_{Spring} = 312.5 kg and the lumped mass M_{TMD} = 500 kg.

Firstly, the natural frequencies were obtained comparing the classical model of the TMD to the helical model, then a finite element model in which the mass of the spring was considered was used to validate the helical model. The results are reported below in Table 1:

Natural Frequency	FEM	Helical Spring Model	Lumped Mass Model
1st	85.978 Rad/s	85.978 Rad/s	85.976 Rad/s
2nd	360.889 Rad/s	360.888 Rad/s	360.810 Rad/s
3rd	923.942 Rad/s	923.956 Rad/s	923.956 Rad/s
4th	1453.414 Rad/s	1453.410 Rad/s	1449.149 Rad/s
5th	2351.407 Rad/s	2351.838 Rad/s	2343.061 Rad/s

As shown, the helical model is validated by the finite element model where a very good agreement is seen between both sets of natural frequencies. It is also evident that only a very small error exists between the classical TMD model and that of the helical model. It should be noted however that a general trend emerges showing that at higher frequencies, the error between the classical TMD and helical models grows.

To obtain the results from dynamic testing, a Monte-Carlo simulation was run 2000 times for each mode, 10000 times in total. The arrival rate set for these loads was 0.375 loads per second, the velocity, v, was 34 m/s, and the magnitude was between the range of 40000 and 240000 N.
The figures below show the responses obtained.

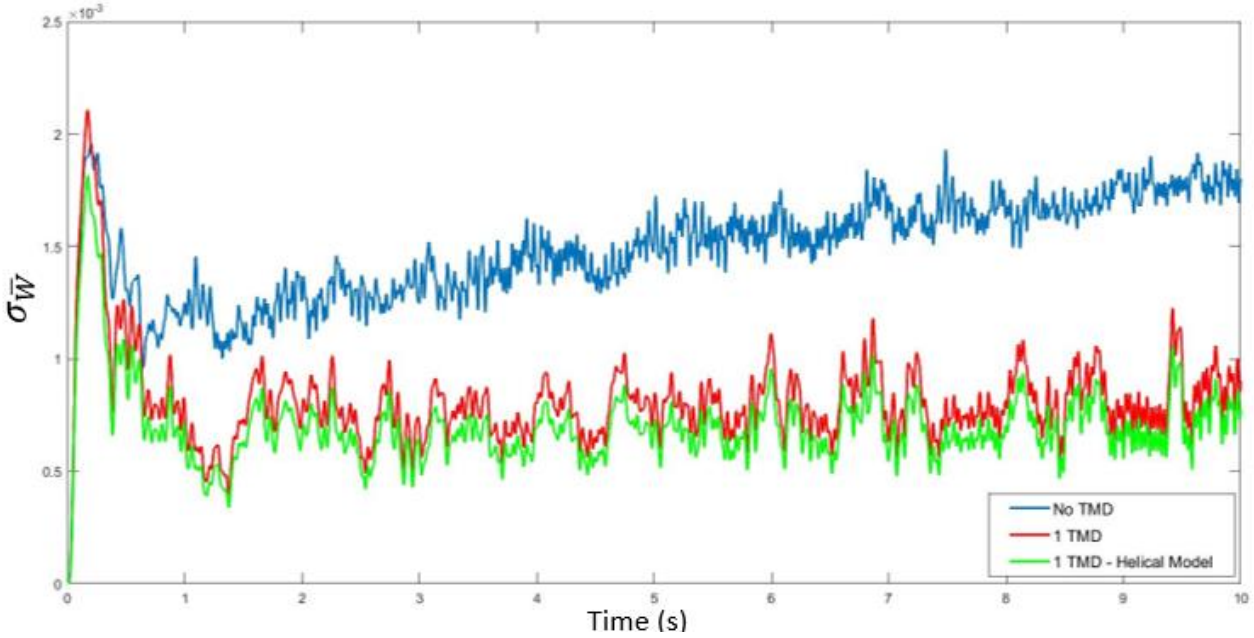


Figure 2. Standard Deviation of Displacement.

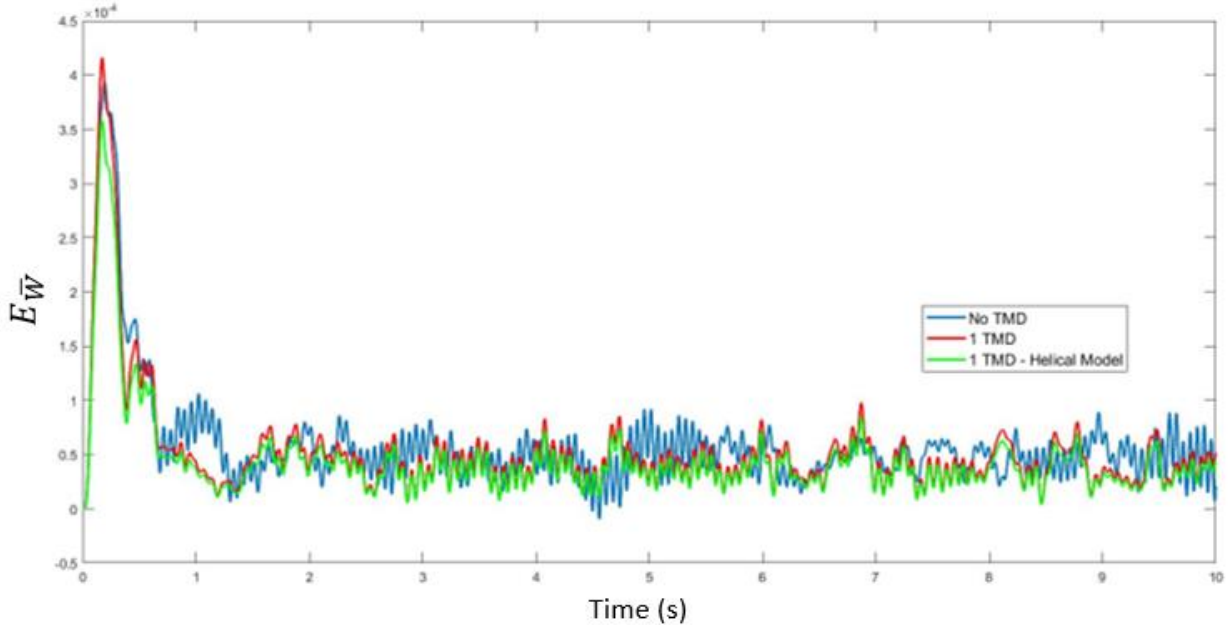


Figure 3. Mean Displacement.

Figures 2 and 3 allow two conclusive conclusions to be drawn, firstly, the traditional method used to model TMDs is highly accurate and the results obtained correlate well with those from the more complex and, therefore more accurate, helical model. The helical model does appear to slightly outperform the traditional TMD model however, the two models have very similar dynamic responses. This, correlation is observed in both the standard deviation and the mean displacement figures where the helical model and the traditional model have very similar response curves but with a slightly lower magnitude in the helical response. Secondly, it can be clearly observed that the tuning parameters selected are insufficient to provide a meaningful reduction in the mean displacement of the bridge structure, however, having observed a reduction in the standard deviation of the displacement, it is clear that some benefit is derived from the application of a TMD.

8. Conclusions

A novel method has been presented to find the response of an Euler-Bernoulli beam fitted with TMDs subjected to Poissonian loading in which the spring inertia effects are considered. This method yields a number of advantages over traditional methods, principally, a reduction in the computational power required to find an exact solution, with respect to the classical mathematical method. This was achieved by viewing an attachment, in this case a spring-mass damper with spring inertial effects, not as a discontinuity, wherein the beam would have to be split at each attachment point, considered as two separate but coupled continuous beams which are related through their

boundary conditions, but rather, as a point load. By considering attachments as reactionary forces creating loading at specific points, the number of unknown constants is reduced which greatly simplifies the classical matrix method solution, as, using the proposed method, there will only ever be 4 unknown constants which must be computed as opposed to the classical case in which 4(N+1) constants for N number of attachments must be found. The use of "smart" dampers was also discussed, these are still relatively new and their application is not yet widespread. This means that this is still a market ripe for expansion and with more efficient and accurate dynamic analysis, it may be easier to promote the introduction of these novel types of dampers to existing structures. The introduction of these novel types of dampers to the proposed analytical method has also been discussed and provided that a robust mathematical method has been developed for each type of damper, this should be more than possible.

Finally, a Monte-Carlo simulation was run in which three beam configurations were subjected to Poissonian loading. The results then showed an excellent correlation between the two TMD models despite a relatively minor improvement in the reduction of the maximum displacement with respect to the bare beam. The method proposed herein is shown to be applicable to cases involving Poissonian loading and can therefore be extended to consider other cases of stochastic loadings and cases involving smart damping devices.

Acknowledgements

The research presented in this project was carried out as part of the H2020-MSCA-ETN-2016. This project has received funding from the European Union's H2020 Programme for research, technological development and demonstration under grant agreement number 721493



References

Adam, C., S. Di Lorenzo, G. Failla, and A. Pirrotta. On the moving load problem in beam structures equipped with tuned mass dampers. *Meccanica*, 52(13): 3101–3115, 2017.

Bilello, C., M. Di Paola, and A. Pirrotta. Time delay induced effects on control of non-linear systems under random excitation. *Meccanica*, 37(1–2): 207–220, 2002.

Biondi, B. and S. Caddemi. Euler–Bernoulli beams with multiple singularities in the flexural stiffness. *European Journal of Mechanics - A/Solids*, (26) 5: 789–809, 2007.

Bucher, C. and M. Di Paola. Efficient solution of the first passage problem by Path Integration for normal and Poissonian white noise. *Probabilistic Engineering Mechanics*, 41: 121–128, 2015.

Di Lorenzo, S., M. Di Paola, G. Failla, and A. Pirrotta. On the moving load problem in Euler–Bernoulli uniform beams with viscoelastic supports and joints. *Acta Mechanica*, 228(3): 805–821, 2017.

Di Matteo, A., I. P. Dunn, G. Failla, and A. Pirrotta. A novel solution to find the dynamic response of an Euler-Bernoulli beam fitted with intraspan TMDs under Poisson type loading. *Infrastructures*, 5(5), 2020.

Di Matteo, A., M. Di Paola, and A. Pirrotta. Path integral solution for nonlinear systems under parametric Poissonian white noise input. *Probabilistic Engineering Mechanics*, 44: 89–98, 2016.

Di Paola, M. and A. Pirrotta. Non-linear systems under impulsive parametric input. *International Journal of Non-Linear Mechanics*, 34(5): 843–851, 1999.

Di Paola, M. and G. Ricciardi. Vibration of a bridge under a random train of moving loads. In *Probabilistic Mechanics and Structural and Geotechnical Reliability, Proceedings of the Specialty Conference*, 136–139, 1992.

Di Paola, M., G. Falsone, and A. Pirrotta. Stochastic response analysis of nonlinear systems under Gaussian inputs. *Probabilistic Engineering Mechanics*, 7(1): 15–21, 1992.

Di Paola, M., G. Ricciardi, and M. Vasta. A method for the probabilistic analysis of nonlinear systems. *Probabilistic Engineering Mechanics*, 10(1): 1–10, 1995.

Dunn, I., A. Di Matteo, G. Failla, A. Pirrotta, and A. F. Russillo. Stochastic response of beams equipped with tuned mass dampers subjected to Poissonian loads. In Proceedings of the *13th International Conference on Applications of Statistics and Probability in Civil Engineering, ICASP*, 2019.

Failla, G. An exact generalised function approach to frequency response analysis of beams and plane frames with the inclusion of viscoelastic damping. *Journal of Sound and Vibration*, 360: 171–202, 2016.

Failla, G. On the dynamics of viscoelastic discontinuous beams. *Mechanics Research Communications*, 60: 52–63, 2014.

Failla, G., M. Di Paola, A. Pirrotta, A. Burlon, and I. Dunn. Random vibration mitigation of beams via tuned mass dampers with spring inertia effects. *Meccanica*, 54(9): 1365–1383, 2019.

Falsone, G. The Use of Generalised Functions in the Discontinuous Beam Bending Differential Equations. *International Journal of Engineering Education*, 18(3): 337-343, 2002.

Furtmüller, T., A. Di Matteo, C. Adam, and A. Pirrotta. Base-isolated structure equipped with tuned liquid column damper: An experimental study. *Mechanical Systems and Signal Processing*, 116: 816–831, 2019.

Oliveto, G., A. Santini, and E. Tripodi. Complex Modal Analysis Of A Flexural Vibrating Beam With Viscous End Conditions. *Journal of Sound and Vibration*, 200(3): 327–345, 1997.

Pirrotta, A. Non-linear systems under parametric white noise input: Digital simulation and response. *International Journal of Non-Linear Mechanics*, 40(8): 1088–1101, 2005.

Ricciardi, G. Random Vibration of Beam under Moving Loads. *Journal of Engineering Mechanics*, 120(11): 2361–2380, 1994.

Explicit sensitivities of the stochastic response of structural systems under spectrum compatible fully non-stationary seismic excitations

T. Alderucci[1], F. Genovese[1], and G. Muscolino[1]
[1]*Department of Engineering, University of Messina, Messina, Italy. E-mail talerucci@unime.it, fedgenovese@unime.it, gmuscolino@unime.it*

Abstract: The sensitivity analysis is a suitable vehicle to evaluate the variation of structural responses under the influence of changes of structural parameters. It has been recently recognized that ground motion accelerations change both in their amplitude and frequency content and can be classified as non-stationary processes. It follows that the sensitivity of the response *evolutionary power spectral density* (*EPSD*) of structures subjected to non-stationary stochastic processes is an essential information and, consequently, it plays a fundamental role in structural design.
In this study handy expressions for the evaluation of sensitivities of stochastic response characteristics of structural systems with damping devices subjected to seismic excitations, modeled as fully non-stationary Gaussian stochastic spectrum compatible processes, are evaluated. Since the structural systems with damping devices are non-classically damped, first, the *time-frequency varying response* (*TFR*) function vector for non-classically damped systems is obtained in explicit form. Then, closed form solutions for the sensitivities of the *TFR*, as well as of the one-sided *EPSD* of the structural response, with respect to device parameters are evaluated.

Keywords: Sensitivity analysis, Complex modal analysis, Stochastic analysis, Fully non-stationary processes.

1. Introduction

The sensitivity analysis (i.e. the evaluation of partial derivatives of a performance measure with respect to system parameters) is a suitable vehicle to evaluate the variation of structural responses under the influence of changes of structural parameters (Arora and Haug, 1979). It follows that the sensitivity analysis plays a fundamental role in structural design.

Strong motion earthquakes are certainly the main critical actions for structures located in the seismically active regions of the earth. The analysis of recorded accelerograms after earthquakes evidence that different earthquakes produce ground motions with different characteristics; that is, ground motions with different intensity, duration, dominant periods and frequency content. It follows that in order to guaranty a good performance of structures in seismic areas a correct characterization of the ground motion acceleration is necessary. The analysis of recorded accelerograms in different sites shows that earthquake ground motion time-histories can be considered as sample of a zero mean Gaussian non-stationary processes in both amplitude and frequency content: the so-called *fully non-stationary processes*. It follows that the structural response is a stochastic process which must be coherently evaluated by applying the tools of stochastic dynamics. It is fully characterized by the so-called *evolutionary power spectral density* (*EPSD*) function matrix.

© 2021 *by T. Alderucci, F. Genovese, G. Muscolino*

The stochastic sensitivity, that is the variation of statistics of structural stochastic responses as a consequence of structural parameters modifications, was pioneering studied by Szopa (1984) and Socha (1986). Afterward several papers have been devoted to this subjected. As an example, Benfratello et al. (2000) proposed a procedure, in the time domain, to evaluate the sensitivity of the statistical moments of the response for stationary Gaussian and non-Gaussian white input processes. Chaudhuri and Chakraborty (2004) developed the formulation in double frequency domain to tackle non-stationary excitation for obtaining the analytical sensitivity statistics of various dynamic response quantities with respect to structural parameters. Cacciola et al. (2005) proposed a numerical procedure for the determination of the evolution of the response statistics sensitivity for both classically and non-classically damped structural systems subjected to non-stationary non-white input processes, by solving set of differential equations once the Kronecker algebra was applied. Liu (2012) proposed numerical methods for the calculation of the sensitivity and Hessian matrix of the response *power spectral density* (*PSD*) matrix function of structures subjected to uniformly modulated evolutionary random seismic excitation. The methods were formulated by accompanying the pseudo-excitation method (Lin et al, 1994) with the Gauss precise time step method or the Newmark method, respectively. Yan et al (2017), utilizing the pseudo-excitation method, implemented a procedure to evaluate the sensitivity of first and second order of response *PSD* functions once the derivatives of eigenpair are evaluated.

In this study handy expressions for the evaluation of sensitivities of stochastic response characteristics of structural systems with damping devices subjected to seismic excitations, modelled as fully non-stationary Gaussian stochastic spectrum compatible processes, are evaluated. Since the structural systems with damping devices are non-classically damped, first, according to the formulation recently proposed by Alderucci and Muscolino (2018), the *time-frequency varying response* (*TFR*) function for non-classically damped systems is evaluated in explicit form. Then, closed form solutions for the first-order derivatives of the *TFR* as well as of the one-sided *EPSD* of the structural response, with respect to device parameters are evaluated.

Numerical applications show the computational efficiency of the proposed approach.

2. Problem Formulation

2.1 Equations of motion

Let us consider a structural system subjected to seismic excitations whose configuration could be modified for design reasons introducing seismic devices. It follows that the equations of motion of an *n*-degree of freedom (*n*-DOF) structural linear system, quiescent at time $t=0$, can be written in the form

$$\mathbf{M}\ddot{\mathbf{U}}(t,\boldsymbol{\alpha})+\mathbf{C}(\boldsymbol{\alpha}_C)\dot{\mathbf{U}}(t,\boldsymbol{\alpha})+\mathbf{K}(\boldsymbol{\alpha}_K)\mathbf{U}(t,\boldsymbol{\alpha})=-\mathbf{M}\boldsymbol{\tau}\ddot{U}_{\mathrm{g}}(t);\quad \mathbf{U}(t_0,\boldsymbol{\alpha})=\mathbf{0}\,, \tag{1}$$

where $\mathbf{M}$, $\mathbf{C}(\boldsymbol{\alpha}_C)$, and $\mathbf{K}(\boldsymbol{\alpha}_K)$ are the $n\times n$ mass, damping, and stiffness matrices of the structure, $\mathbf{U}(\boldsymbol{\alpha},t)$ is the *n*-dimensional vector of nodal displacements relative to the ground; τ is the *n*-dimensional array listing the influence coefficients of the ground shaking; $\ddot{U}_{\mathrm{g}}(t)$ is the seismic acceleration; a dot over a variable denotes differentiation with respect to time. In Eq.(1), the dependence of the damping and stiffness matrices of the structure, as well as of the response vector, on the unknown *r*-order parameter vector α,

characterizing seismic devices, is stressed. The vector $\boldsymbol{\alpha}^T = \left[\boldsymbol{\alpha}_C^T \quad \boldsymbol{\alpha}_K^T\right]$, of order $r\ (r = r_c + r_k)$, collects seismic device parameters, which must be evaluated by the design procedure. It can be split as:

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}_0 + \Delta\boldsymbol{\alpha}, \tag{2}$$

where $\Delta\boldsymbol{\alpha}^T = \left[\Delta\boldsymbol{\alpha}_C^T \quad \Delta\boldsymbol{\alpha}_K^T\right]$ is assumed to be a vector collecting small parameter variations with respect to the nominal parameter vector $\boldsymbol{\alpha}_0^T = \left[\boldsymbol{\alpha}_{C,0}^T \quad \boldsymbol{\alpha}_{K,0}^T\right]$. It follows that the $n \times n$ damping, and stiffness matrices of the structure, defined in Eq.(1), can be split as follows:

$$\mathbf{K}(\boldsymbol{\alpha}_K) = \mathbf{K}_S + \mathbf{K}_D(\boldsymbol{\alpha}_K); \quad \mathbf{C}(\boldsymbol{\alpha}_C) = \mathbf{C}_S + \mathbf{C}_D(\boldsymbol{\alpha}_C) \tag{3}$$

in which $\mathbf{K}_S$ and $\mathbf{C}_S$ are the stiffness and damping matrices of the structure without devices, respectively; $\mathbf{K}_D(\boldsymbol{\alpha}_K) = \mathbf{K}_D(\boldsymbol{\alpha}_{K,0}) + \Delta\mathbf{K}_D(\boldsymbol{\alpha}_K)$ and $\mathbf{C}_D(\boldsymbol{\alpha}_C) = \mathbf{C}_D(\boldsymbol{\alpha}_{C,0}) + \Delta\mathbf{C}_D(\boldsymbol{\alpha}_C)$ are the additional stiffness and damping matrices due to the installation of devices. They are composed by $\mathbf{K}_D(\boldsymbol{\alpha}_{K,0})$ and $\mathbf{C}_D(\boldsymbol{\alpha}_{C,0})$, evaluated in correspondence of the nominal parameter vector $\boldsymbol{\alpha}_0$ of seismic devices, and by $\Delta\mathbf{C}_D(\boldsymbol{\alpha}_C) = \mathbf{C}_D(\boldsymbol{\alpha}_C) - \mathbf{C}_D(\boldsymbol{\alpha}_{C,0})$ and $\Delta\mathbf{K}_D(\boldsymbol{\alpha}_K) = \mathbf{K}_D(\boldsymbol{\alpha}_K) - \mathbf{K}_D(\boldsymbol{\alpha}_{K,0})$, their deviations with respect to the additional stiffness and damping matrices evaluated at nominal seismic device parameters.

Due to the presence of seismic devices, the structural system generally could become non-classically damped, it follows that to evaluate the structural response, Eq.(1) has to be written in state-variables:

$$\dot{\mathbf{Z}}(t,\boldsymbol{\alpha}) = \mathbf{D}(\boldsymbol{\alpha})\mathbf{Z}(t,\boldsymbol{\alpha}) + \mathbf{w}\ddot{U}_g(t); \quad \mathbf{Z}(t_0,\boldsymbol{\alpha}) = \mathbf{0} \tag{4}$$

where $\mathbf{Z}(t,\boldsymbol{\alpha})$ is the 2*n*-state-variable vector, while the matrix $\mathbf{D}(\boldsymbol{\alpha})$, of order $2n \times 2n$, and the vector $\mathbf{w}$, of order $2n$, are defined, respectively, as:

$$\mathbf{Z}(t,\boldsymbol{\alpha}) = \begin{bmatrix} \mathbf{U}(t,\boldsymbol{\alpha}) \\ \dot{\mathbf{U}}(t,\boldsymbol{\alpha}) \end{bmatrix}; \quad \mathbf{D}(\boldsymbol{\alpha}) = \begin{bmatrix} \mathbf{O}_{n,n} & \mathbf{I}_n \\ -\mathbf{M}^{-1}\mathbf{K}(\boldsymbol{\alpha}_K) & -\mathbf{M}^{-1}\mathbf{C}(\boldsymbol{\alpha}_C) \end{bmatrix}; \quad \mathbf{w} = \begin{bmatrix} \mathbf{O}_{n,1} \\ -\boldsymbol{\tau} \end{bmatrix} \tag{5}$$

with $\mathbf{I}_n$ the *n*-order identity matrix and $\mathbf{O}_{n,s}$ the zero matrix of order $n \times s$.

2.2 DETERMINISTIC SENSITIVITY FUNCTION

The sensitivity vector of the structural response, $\mathbf{s}_{\mathbf{Z},i}(\boldsymbol{\alpha}_0, t)$, with respect to *i*-th parameter α_i, *i*-th element of the *r*-order parameter vector $\boldsymbol{\alpha}$, is defined as follows:

$$\mathbf{s}_{\mathbf{Z},i}(t,\boldsymbol{\alpha}_0) = \left.\frac{\partial\mathbf{Z}(t,\boldsymbol{\alpha})}{\partial\alpha_i}\right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} \tag{6}$$

It follows that by performing the differentiation of Eq.(4), with respect to *i*-th parameter α_i, and setting $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0$, a differential equation governing the evolution of state-variable sensitivity vector is obtained (Cacciola et al, 2005), whose solution can be written as:

$$\mathbf{s}_{\mathbf{Z},i}(t,\boldsymbol{\alpha}_0)=\int_{t_0}^{t}\boldsymbol{\Theta}(t-\tau,\boldsymbol{\alpha}_0)\mathbf{D}_i'(\boldsymbol{\alpha}_0)\boldsymbol{\Psi}(\boldsymbol{\alpha}_0)\mathbf{X}(t,\boldsymbol{\alpha}_0)\,\mathrm{d}\tau \tag{7}$$

in which the matrix $\mathbf{D}_i'(\boldsymbol{\alpha}_0)$ is the differentiation of the matrix $\mathbf{D}(\boldsymbol{\alpha})$ with respect to *i*-th parameter, α_i, and setting $\boldsymbol{\alpha}=\boldsymbol{\alpha}_0$; $\boldsymbol{\Theta}(\boldsymbol{\alpha},t)$ is the transition matrix which can be evaluated once the following eigenproblem is solved:

$$\mathbf{D}^{-1}(\boldsymbol{\alpha})\boldsymbol{\Psi}(\boldsymbol{\alpha})=\boldsymbol{\Psi}(\boldsymbol{\alpha})\boldsymbol{\Lambda}^{-1}(\boldsymbol{\alpha});\quad \boldsymbol{\Psi}^T(\boldsymbol{\alpha})\mathbf{A}(\boldsymbol{\alpha}_C)\boldsymbol{\Psi}(\boldsymbol{\alpha})=\mathbf{I}_{2m};\quad \mathbf{A}(\boldsymbol{\alpha}_C)=\begin{bmatrix}\mathbf{C}(\boldsymbol{\alpha}_C) & \mathbf{M}\\ \mathbf{M} & \mathbf{O}_{n,n}\end{bmatrix}. \tag{8}$$

where the superscript T denotes the transpose operator, $\boldsymbol{\Lambda}(\boldsymbol{\alpha})$ is a diagonal matrix collecting the first $2m$ complex eigenvalues ($m\le n$ is the number of *complex modes* selected for the analysis), and $\boldsymbol{\Psi}(\boldsymbol{\alpha})$ is a complex matrix, of order $(2n\times 2m)$, collecting the corresponding $2m$ complex eigenvectors. Once the eigenproblem (8) is solved, the transition matrix can be evaluated as follows:

$$\boldsymbol{\Theta}(t,\boldsymbol{\alpha})=\exp\left[t\,\mathbf{D}(\boldsymbol{\alpha})\right]=\boldsymbol{\Psi}(\boldsymbol{\alpha})\exp\left[t\,\boldsymbol{\Lambda}(\boldsymbol{\alpha})\right]\boldsymbol{\Psi}^T(\boldsymbol{\alpha})\mathbf{A}(\boldsymbol{\alpha}_C) \tag{9}$$

Alternatively, the state-variable sensitivity vector (7), with respect to the *i*-th parameter, can be evaluated as:

$$\mathbf{s}_{\mathbf{Z},i}(t,\boldsymbol{\alpha}_0)=\boldsymbol{\Psi}(\boldsymbol{\alpha}_0)\mathbf{Y}_i(t,\boldsymbol{\alpha}_0) \tag{10}$$

where $\mathbf{Y}_i(t,\boldsymbol{\alpha}_0)$ is the sensitivity vector of the response, with respect to the parameter α_i, into the complex modal subspace, given as:

$$\mathbf{Y}_i(t,\boldsymbol{\alpha}_0)=\int_{t_0}^{t}\exp\left[(\tau-\rho)\boldsymbol{\Lambda}(\boldsymbol{\alpha}_0)\right]\mathbf{B}_i(\boldsymbol{\alpha}_0)\mathbf{X}(\tau,\boldsymbol{\alpha}_0)\,\mathrm{d}\tau \tag{11}$$

and:

$$\mathbf{B}_i(\boldsymbol{\alpha}_0)=\boldsymbol{\Psi}^T(\boldsymbol{\alpha}_0)\mathbf{A}(\boldsymbol{\alpha}_{C,0})\mathbf{D}_i'(\boldsymbol{\alpha}_0)\boldsymbol{\Psi}(\boldsymbol{\alpha}_0);\quad \mathbf{v}(\boldsymbol{\alpha}_0)=\boldsymbol{\Psi}^T(\boldsymbol{\alpha}_0)\mathbf{A}(\boldsymbol{\alpha}_{C,0})\mathbf{w} \tag{12}$$

Notice that for deterministic excitation the state-variable sensitivity vectors (7) and (11) with respect to the *i*-th uncertain parameter, can be easily evaluated by step-by-step procedures (Cacciola et al, 2005).

2.3 DEFINITION OF SEISMIC ACCELERATIONS AS FULLY NON-STATIONARY RANDOM PROCESSES

In this paper it is assumed that the ground motion acceleration, $\ddot{U}_g(t)$, is a zero-mean Gaussian fully non-stationary random process. In order to define this process, here the Priestley spectral representation of non-stationary processes is adopted (Priestley, 1965; 1967). Moreover, in the stochastic analysis the one-sided *Power Spectral Density* (*PSD*) function is generally used to characterize the input process. It has been demonstrated that, since the one-sided *PSD* function is not symmetric (Di Paola, 1985; Di Paola and Petrucci, 1990; Muscolino, 1991), the corresponding autocorrelation function is a complex function having real part coincident with the autocorrelation function corresponding to the two-sided *PSD*. This implies that,

from a mathematical point of view, the zero-mean Gaussian fully non-stationary random process is a complex process. It can be defined by means of the following *Fourier-Stieltjes integral*:

$$\ddot{U}_{\mathrm{g}}(t) = \sqrt{2}\int_0^{\infty} \exp(\mathrm{i}\omega t)\, a(\omega,t)\, \mathrm{d}N(\omega) \tag{13}$$

where $\mathrm{i} = \sqrt{-1}$ is the imaginary unit; $a(\omega,t)$ is a slowly varying complex deterministic time-frequency modulating function which has to satisfy the condition: $a(\omega,t) \equiv a^*(-\omega,t)$; $N(\omega)$ is a zero-mean process with orthogonal increments satisfying the condition:

$$\mathrm{E}\left\langle \mathrm{d}N(\omega_1)\, \mathrm{d}N^*(\omega_2) \right\rangle = \delta(\omega_1 - \omega_2) G_0(\omega_1)\, \mathrm{d}\omega_1\, \mathrm{d}\omega_2 \tag{14}$$

where the asterisk * denotes the complex conjugate quantity; the operator $\mathrm{E}\langle\bullet\rangle$ denotes the stochastic average; $\delta(\bullet)$ is the Dirac delta, and $G_0(\omega)$ is the one-sided *PSD* function of the "embedded" stationary counterpart process (Michealov et al, 1999), which is a real function for $\omega \geq 0$, while $G_0(\omega) = 0$ for $\omega < 0$.

Notice that, since a one-sided *PSD* has been herein assumed, the zero-mean Gaussian non-stationary random process $\ddot{U}_{\mathrm{g}}(t)$ is a complex one (Di Paola, 1985; Di Paola and Petrucci, 1990; Muscolino, 1991; Muscolino and Alderucci, 2015). This process can be completely defined, in the time domain, by the knowledge of its complex autocorrelation function:

$$R_{\ddot{U}_{\mathrm{g}}\ddot{U}_{\mathrm{g}}}(t_1,t_2) \equiv \mathrm{E}\left\langle \ddot{U}_{\mathrm{g}}(t_1)\, \ddot{U}_{\mathrm{g}}(t_2) \right\rangle = \int_0^{\infty} \exp\left[\mathrm{i}\omega(t_1 - t_2)\right] a(\omega,t_1)\, a^*(\omega,t_2) G_0(\omega)\, \mathrm{d}\omega \tag{15}$$

The complex process, $\ddot{U}_{\mathrm{g}}(t)$, having complex autocorrelation function (15) has been called *pre-envelope process* by Di Paola (1985). In the Priestley evolutionary process model, the integrand function

$$G_{\ddot{U}_{\mathrm{g}}\ddot{U}_{\mathrm{g}}}(\omega,t) = \left|a(\omega,t)\right|^2 G_0(\omega) \tag{16}$$

is called one-sided *evolutionary power spectral density* (*EPSD*) function of the non-stationary process $\ddot{U}_{\mathrm{g}}(t)$. This processes is called *fully non-stationary* random processes, since both time and frequency content change. If the modulating function is a time dependent function, $a(\omega,t) \equiv a(t)$, the non-stationary process is called *uniformly modulated* (or *quasi-stationary*) random process. In the latter case the *EPSD* function assumes the expression: $G_{\ddot{U}_{\mathrm{g}}\ddot{U}_{\mathrm{g}}}(\omega,t) = a(t)^2 G_0(\omega)$.

In this paper the seismic excitations, is modeled as fully non-stationary Gaussian stochastic spectrum compatible processes. It is well known that, on the contrary of stationary random model where the spectrum-compatible *PSD* function is univocally determined, for fully non-stationary random model the *EPSD* function cannot be defined univocally (Cacciola 2010). Here the iterative procedure recently proposed by (Alderucci et al, 2019b) is adopted.

3. Explicit sensitivities of the evolutionary power spectral density response function

3.1 Closed form solution for the sensitivity time-frequency varying response vector function

It is well known that the *time-frequency varying response* (*TFR*) vector function of the response plays a central role in the evaluation of the statistics of the response for both classically and non-classically damped structural systems subjected to fully non-stationary stochastic input. In the presence of the unknown *r*-order parameter vector $\boldsymbol{\alpha}$, the *TFR* vector function of nodal response, $\mathbf{Z}(\omega,t,\boldsymbol{\alpha})$, can be evaluated as follows:

$$\mathbf{Z}(\omega,t,\boldsymbol{\alpha}) = \boldsymbol{\Psi}(\boldsymbol{\alpha})\,\mathbf{X}(\omega,t,\boldsymbol{\alpha}) \tag{17}$$

where $\mathbf{X}(\omega,t,\boldsymbol{\alpha})$ is the *TFR* vector function of the modal complex response, given by:

$$\mathbf{X}(\omega,t,\boldsymbol{\alpha}) = \int_{t_0}^{t} \exp\left[(t-\tau)\,\boldsymbol{\Lambda}(\boldsymbol{\alpha})\right]\mathbf{v}(\boldsymbol{\alpha})\exp(\mathrm{i}\omega\tau)\,a(\omega,\tau)\,\mathrm{d}\tau \tag{18}$$

This vector function, in the following denoted by the acronym *MTFR* (*modal time-frequency varying response*), can be evaluated in closed form for the commonly adopted modulating functions (Muscolino and Alderucci, 2015; Alderucci and Muscolino, 2018). In particular, here the Spanos and Solomos (1983) model for the fully non-stationary seismic excitation is adopted, whose modulating time-frequency functions can be written as:

$$a(\omega,t) = \varepsilon(\omega)\,(t-t_0)\exp\left[-\alpha_a(\omega)(t-t_0)\right]\mathbb{U}(t-t_0); \tag{19}$$

where $\varepsilon(\omega)$ and $\alpha_a(\omega)$ could be complex functions. For quiescent structural systems at time $t_0 = 0$, the vector $\mathbf{X}(\omega,t,\boldsymbol{\alpha})$, defined in Eq.(18), can be evaluated in explicit form as (Muscolino and Alderucci, 2015; Alderucci and Muscolino, 2018):

$$\mathbf{X}(\omega,t,\boldsymbol{\alpha}) = -\varepsilon(\omega)\left\{\exp(-\beta(\omega)\,t)\left[\boldsymbol{\Gamma}^2(\omega,\boldsymbol{\alpha}) + t\,\boldsymbol{\Gamma}(\omega,\boldsymbol{\alpha})\right] - \exp\left[t\,\boldsymbol{\Lambda}(\boldsymbol{\alpha})\right]\boldsymbol{\Gamma}^2(\omega,\boldsymbol{\alpha})\right\}\mathbf{v}(\boldsymbol{\alpha})\,\mathbb{U}(t) \tag{20}$$

where $\beta(\omega) = \alpha_a(\omega) - \mathrm{i}\omega$ and $\boldsymbol{\Gamma}(\omega,\boldsymbol{\alpha})$ is a diagonal matrix defined as:

$$\boldsymbol{\Gamma}(\omega,\boldsymbol{\alpha}) = \left[\boldsymbol{\Lambda}(\boldsymbol{\alpha}) + \beta(\omega)\,\mathbf{I}_{2m}\right]^{-1}. \tag{21}$$

According to Eqs.(10) and (17), the *sensitivity* of the *TFR* vector function, with respect to *i*-th parameter, can be also evaluated as

$$\mathbf{s}_{\mathbf{Z},i}(\omega,t,\boldsymbol{\alpha}_0) = \boldsymbol{\Psi}(\boldsymbol{\alpha}_0)\,\mathbf{Y}_i(\omega,t,\boldsymbol{\alpha}_0) \tag{22}$$

where $\mathbf{Y}_i(\omega,t,\boldsymbol{\alpha}_0)$ is the *sensitivity* of the *TFR* vector function, with respect to the parameter α_i, projected into the complex modal subspace. It can be evaluated as follows:

$$\mathbf{Y}_i(\omega,t,\boldsymbol{\alpha}_0) = \int_0^t \exp\left[(t-\tau)\,\boldsymbol{\Lambda}(\boldsymbol{\alpha}_0)\right]\mathbf{B}_i(\boldsymbol{\alpha}_0)\,\mathbf{X}(\omega,\tau,\boldsymbol{\alpha}_0)\,\mathrm{d}\tau. \tag{23}$$

In the following, this vector will be synthetically denoted by the acronym *MSTFR* (*modal sensitivity time-frequency vector*). After very simple algebra it can be shown that this vector function can be evaluated as solution of the following differential equation, with zero start conditions at time $t_0 = 0$:

$$\dot{\mathbf{Y}}_i(\omega,t,\boldsymbol{\alpha}_0) = \boldsymbol{\Lambda}(\boldsymbol{\alpha}_0)\,\mathbf{Y}_i(\omega,t,\boldsymbol{\alpha}_0) + \mathbf{B}_i(\boldsymbol{\alpha}_0)\,\mathbf{X}(\omega,\tau,\boldsymbol{\alpha}_0)\,\mathbb{U}(t-t_0); \quad \mathbf{Y}_i(\omega,0,\boldsymbol{\alpha}_0) = \mathbf{0}. \tag{24}$$

To perform the solution of this set of differential equations the *MTFR* vector function vector, defined in Eq.(20), is rewritten as (Alderucci et al, 2019a):

$$\mathbf{X}(\omega,t,\boldsymbol{\alpha}_0) = \mathbf{X}_1(\omega,t,\boldsymbol{\alpha}_0) + \mathbf{X}_2(\omega,t,\boldsymbol{\alpha}_0) \tag{25}$$

where:

$$\begin{aligned} \mathbf{X}_1(\omega,t,\boldsymbol{\alpha}_0) &= -\varepsilon(\omega)\exp(-\beta(\omega)\,t)\,\left[\boldsymbol{\Gamma}_0^2(\omega) + t\,\boldsymbol{\Gamma}_0(\omega)\right]\mathbf{v}_0\,\boldsymbol{\mathcal{U}}(t); \\ \mathbf{X}_2(\omega,t,\boldsymbol{\alpha}_0) &= \varepsilon(\omega)\exp(t\,\boldsymbol{\Lambda}_0)\,\boldsymbol{\Gamma}_0^2(\omega)\mathbf{v}_0\,\boldsymbol{\mathcal{U}}(t). \end{aligned} \tag{26}$$

Since the *MTFR* vector function has been split as the sum of two contributions, the *MSTFR* vector, solution of Eq.(24), can be split as the sum of two vectors too, solutions of the following two sets of differential equations, with zero start initial conditions at time $t_0 = 0$:

$$\begin{aligned} \mathbf{Y}_i(\omega,t,\boldsymbol{\alpha}_0) = \mathbf{Y}_{i,1}(\omega,t,\boldsymbol{\alpha}_0) + \mathbf{Y}_{i,2}(\omega,t,\boldsymbol{\alpha}_0) = \big\{ & \mathbf{Y}_{i,1,\mathrm{p}}(\omega,t,\boldsymbol{\alpha}_0) + \mathbf{Y}_{i,2,\mathrm{p}}(\omega,t,\boldsymbol{\alpha}_0) \\ & - \exp(t\,\boldsymbol{\Lambda}_0)\left[\mathbf{Y}_{i,1,\mathrm{p}}(\omega,0,\boldsymbol{\alpha}_0) + \mathbf{Y}_{i,2,\mathrm{p}}(\omega,0,\boldsymbol{\alpha}_0)\right]\big\}; \quad t > 0 \end{aligned} \tag{27}$$

where the particular solution vectors of Eq.(24), can be evaluated, after some algebra, as follows:

$$\begin{aligned} \mathbf{Y}_{i,1,\mathrm{p}}(\omega,t,\boldsymbol{\alpha}_0) &= \varepsilon(\omega)\exp(-\beta(\omega)\,t)\,\boldsymbol{\Gamma}(\omega,\boldsymbol{\alpha}_0)\left[\boldsymbol{\Gamma}(\omega,\boldsymbol{\alpha}_0)\mathbf{B}_i(\boldsymbol{\alpha}_0) + \mathbf{B}_i(\boldsymbol{\alpha}_0)\boldsymbol{\Gamma}(\omega,\boldsymbol{\alpha}_0) + t\,\mathbf{B}_i(\boldsymbol{\alpha}_0)\right]\boldsymbol{\Gamma}(\omega,\boldsymbol{\alpha}_0)\mathbf{v}(\boldsymbol{\alpha}_0); \\ \mathbf{Y}_{i,2,\mathrm{p}}(\omega,t,\boldsymbol{\alpha}_0) &= \varepsilon(\omega)\mathbf{P}_i(t,\boldsymbol{\alpha}_0)\exp\left[t\,\boldsymbol{\Lambda}(\boldsymbol{\alpha}_0)\right]\boldsymbol{\Gamma}^2(\omega,\boldsymbol{\alpha}_0)\mathbf{v}(\boldsymbol{\alpha}_0); \end{aligned} \tag{28}$$

where $\mathbf{P}_i(\boldsymbol{\alpha}_0,t)$ is a matrix of order $(2m\times 2m)$ whose elements, $P_{i,jk}(\boldsymbol{\alpha}_0,t)$, are defined as follows:

$$P_{i,jj}(t,\boldsymbol{\alpha}_0) = t\,B_{i,jj}(\boldsymbol{\alpha}_0); \quad P_{i,jk}(t,\boldsymbol{\alpha}_0) = \frac{B_{i,jk}(\boldsymbol{\alpha}_0)}{\lambda_k - \lambda_j},\ j \neq k \tag{29}$$

with $B_{i,jk}(\boldsymbol{\alpha}_0)$ elements of the matrix $\mathbf{B}_i(\boldsymbol{\alpha}_0)$.

3.3 CLOSED FORM SOLUTIONS FOR THE SENSITIVITIES OF THE EPSD RESPONSE MATRIX FUNCTION

Since it has been assumed that the ground motion acceleration, $\ddot{U}_{\mathrm{g}}(t)$, is a zero-mean Gaussian fully non-stationary random process, with one-sided *EPSD* function, $G_{\ddot{U}_{\mathrm{g}}\ddot{U}_{\mathrm{g}}}(\omega,t)$, the stochastic response is a zero-mean fully non-stationary stochastic vector process too, whose one-sided *EPSD* matrix function, $\mathbf{G}_{\mathbf{ZZ}}(\omega,t,\boldsymbol{\alpha})$, can be evaluated as follows (Alderucci et al, 2019a):

$$\mathbf{G}_{\mathbf{ZZ}}(\omega,t,\boldsymbol{\alpha}) = G_0(\omega)\,\mathbf{Z}^*(\omega,t,\boldsymbol{\alpha})\mathbf{Z}^T(\omega,t,\boldsymbol{\alpha}) = G_0(\omega)\,\boldsymbol{\Psi}^*(\boldsymbol{\alpha})\mathbf{X}^*(\omega,t,\boldsymbol{\alpha})\mathbf{X}^T(\omega,t,\boldsymbol{\alpha})\boldsymbol{\Psi}^T(\boldsymbol{\alpha}) \tag{30}$$

where $G_0(\omega)$ is the one-sided *PSD* function of the "embedded" stationary counterpart of the input process, while $\mathbf{Z}(\omega,t,\boldsymbol{\alpha})$ and $\mathbf{X}(\omega,t,\boldsymbol{\alpha})$ are the *TFR* vector responses, in state variables, into nodal and modal complex spaces, respectively.

Once the one-sided *EPSD* matrix function, $\mathbf{G}_{\mathbf{ZZ}}(\omega,t)$, is defined, it is possible to evaluate in compact form the statistics of the response as follows:

$$\boldsymbol{\Sigma}_{\mathbf{ZZ}}(t,\boldsymbol{\alpha}) = \boldsymbol{\Psi}^{*}(\boldsymbol{\alpha})\left[\int_0^{\infty} G_0(\omega)\,\mathbf{X}^{*}(\omega,t,\boldsymbol{\alpha})\,\mathbf{X}^{T}(\omega,t,\boldsymbol{\alpha})\,\mathrm{d}\omega\right]\boldsymbol{\Psi}^{T}(\boldsymbol{\alpha}) \tag{31}$$

This matrix is the so-called *pre-envelope covariance* (*PEC*) matrix function, in nodal space, it is a $2n\times 2n$ Hermitian matrix, whose real part coincides with the classical covariance matrix (Di Paola, 1985). By differentiating the *PEC* matrix function with respect to the *i*-th parameter, it is possible to evaluate its *sensitivity* function in the neighbourhood of nominal parameters, $\boldsymbol{\alpha}=\boldsymbol{\alpha}_0$ as follows:

$$\boldsymbol{\Sigma}_{\mathbf{s}_{\mathbf{Z},i}}(t,\boldsymbol{\alpha}_0) = \left.\frac{\partial\boldsymbol{\Sigma}_{\mathbf{ZZ}}(t,\boldsymbol{\alpha})}{\partial\alpha_i}\right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}_0} = \mathrm{E}\left\langle\mathbf{Z}(t,\boldsymbol{\alpha}_0)\mathbf{s}_{\mathbf{Z},i}^{*T}(t,\boldsymbol{\alpha}_0)\right\rangle + \mathrm{E}\left\langle\mathbf{Z}(t,\boldsymbol{\alpha}_0)\mathbf{s}_{\mathbf{Z},i}^{*T}(t,\boldsymbol{\alpha}_0)\right\rangle^{*T} \tag{32}$$

whose elements are the *sensitivity* of first three spectral moments with respect to the parameter α_i. In the previous equation the sensitivity vector $\mathbf{s}_{\mathbf{Z},i}(t,\boldsymbol{\alpha}_0)$ which has been defined in Eq.(7) appears. It follows that the following relationship holds:

$$\mathrm{E}\left\langle\mathbf{Z}(t,\boldsymbol{\alpha}_0)\mathbf{s}_{\mathbf{Z},i}^{*T}(t,\boldsymbol{\alpha}_0)\right\rangle = \boldsymbol{\Psi}^{*}(\boldsymbol{\alpha}_0)\left\{\int_0^{\infty}\mathbf{X}^{*}(\omega,t,\boldsymbol{\alpha}_0)\mathbf{Y}_i^{T}(\omega,t,\boldsymbol{\alpha}_0)\,G_0(\omega)\mathrm{d}\omega\right\}\boldsymbol{\Psi}^{T}(\boldsymbol{\alpha}_0) \tag{33}$$

where the vector $\mathbf{X}(\omega,t,\boldsymbol{\alpha}_0)$ and $\mathbf{Y}_i(\omega,t,\boldsymbol{\alpha}_0)$ are given in explicit form in Eqs. (20) and (27), respectively. Substituting Eq.(33) into Eq.(32) the so-called *sensitivity* of the *PEC* matrix function can be rewritten as:

$$\boldsymbol{\Sigma}_{\mathbf{s}_{\mathbf{Z},i}}(t,\boldsymbol{\alpha}_0) \equiv \int_0^{\infty}\mathbf{G}_{\mathbf{s}_{\mathbf{Z},i}}(\omega,t,\boldsymbol{\alpha}_0)\mathrm{d}\omega \tag{34}$$

where the matrix

$$\mathbf{G}_{\mathbf{s}_{\mathbf{Z},i}}(\omega,t,\boldsymbol{\alpha}_0) = G_0(\omega)\,\boldsymbol{\Psi}^{*}(\boldsymbol{\alpha}_0)\left[\mathbf{X}^{*}(\omega,t,\boldsymbol{\alpha}_0)\,\mathbf{Y}_i^{T}(\omega,t,\boldsymbol{\alpha}_0)+\mathbf{Y}_i^{*}(\omega,t,\boldsymbol{\alpha}_0)\,\mathbf{X}^{T}(\omega,t,\boldsymbol{\alpha}_0)\right]\boldsymbol{\Psi}^{T}(\boldsymbol{\alpha}_0). \tag{35}$$

can be interpreted as the *sensitivity* of the *EPSD* matrix function. This matrix function is evaluated by means of closed form solutions too.

4. Numerical Application

In order to show the effectiveness of the proposed method, four Single-Degree-of-Freedom (SDOF) systems are analysed. The selected time-frequency modulating function for the fully non stationary process

is the one proposed by Spanos and Solomos (1983), defined in Eq.(19), whose parameters herein selected are:

$$\alpha_a(\omega)=\frac{1}{2}\left(0.15+\frac{\omega^2}{225\pi^2}\right);\quad \varepsilon(\omega)=\frac{\sqrt{2}}{15\,\pi\,a_{\max}}\omega;\quad t_0=0 \tag{36}$$

where, in order to normalize to one the modulating function, the parameter $a_{\max}$ is herein set equal to 1.34.

For the SDOF system the *PEC* matrix function, in nodal space, it is a 2×2 Hermitian matrix, whose real part coincides with the classical covariance matrix. It can be evaluated as:

$$\mathbf{\Sigma}_{zz}(t,\boldsymbol{\alpha})=\begin{bmatrix}\lambda_{0,u}(t,\boldsymbol{\alpha}) & \mathrm{i}\,\lambda_{1,u}(t,\boldsymbol{\alpha})\\ -\mathrm{i}\,\lambda_{1,u}^{*T}(t,\boldsymbol{\alpha}) & \lambda_{2,u}(t,\boldsymbol{\alpha})\end{bmatrix} \tag{37}$$

where the function $\lambda_{i,u}(t,\boldsymbol{\alpha})$ is the so-called *non-geometric spectral moments* (*NGSM*) of *i*-th order of stochastic response (Michealov et al, 1999). The *EPSD* function, having the spectrum-compatible stationary counterpart *PSD* evaluated according to Alderucci et al. (2019b), is depicted in Figure 1.

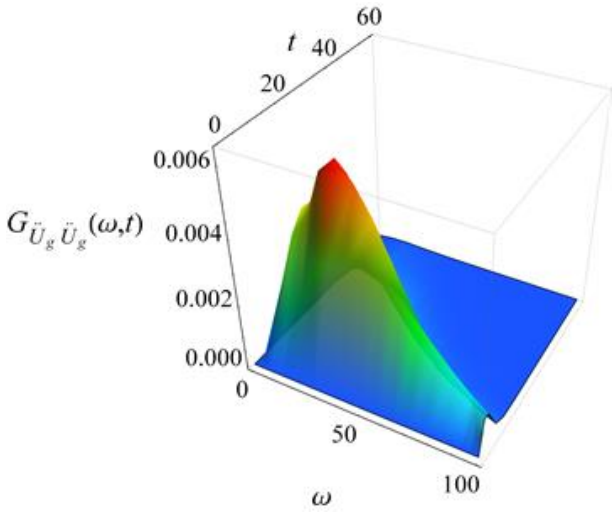


Figure 1. *EPSD* function of the input process according to Alderucci et al. (2019b) model.

The four SDOF systems have damping ratio equal to $\zeta_0=0.02$ and natural periods $T_0=0.1$ s , $T_0=0.2$ s, $T_0=0.6$ s, and $T_0=1$ s, respectively. A damper device, with damping coefficient c_d and stiffness k_d is connected to each SDOF system. The analysis is conducted varying the abovementioned damper parameters within $1\div 10^8$ Ns/m and $1\div 10^8$ N/m, respectively, in order to obtain the optimal design values.

Figure 2 depicts, for each SDOF system, the maximum values of the first NGSM, $\max\left\{\lambda_{0,u}(t)\right\}$, that is coincident with the maximum value of the variance of the response, versus the damper stiffness, for different values of the damping coefficients. From the analysis of these figures, it can be evidenced that changes in stiffness values don't modify significantly the response. Consequently, according to the usual characteristics of the commercial devices, the design stiffness value can be assumed equal to 30 N/m for all the analysed oscillators.

In Figure 3, for the four oscillators, the maximum values of the variance of the response, versus the damping coefficient, for the previously chosen optimal value of the stiffness $k_d = 30\,\text{N/m}$, is reported.

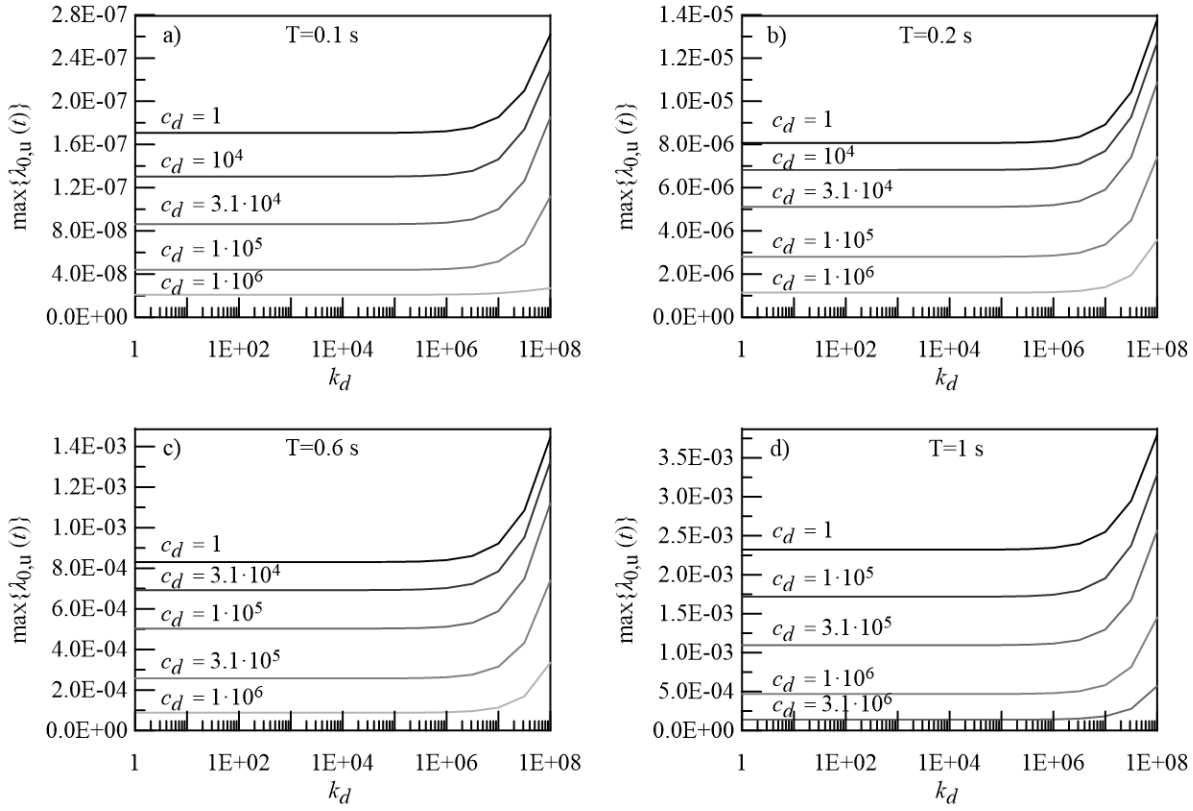


Figure 2. Maximum values of the variance of the response, $\max\left\{\lambda_{0,u}(t)\right\}$, for the four oscillators, versus the stiffness of the device, varying the damping coefficient: a) $T_0 = 0.1\,\text{s}$, b) $T_0 = 0.2\,\text{s}$, c) $T_0 = 0.6\,\text{s}$, d) $T_0 = 1.0\,\text{s}$.

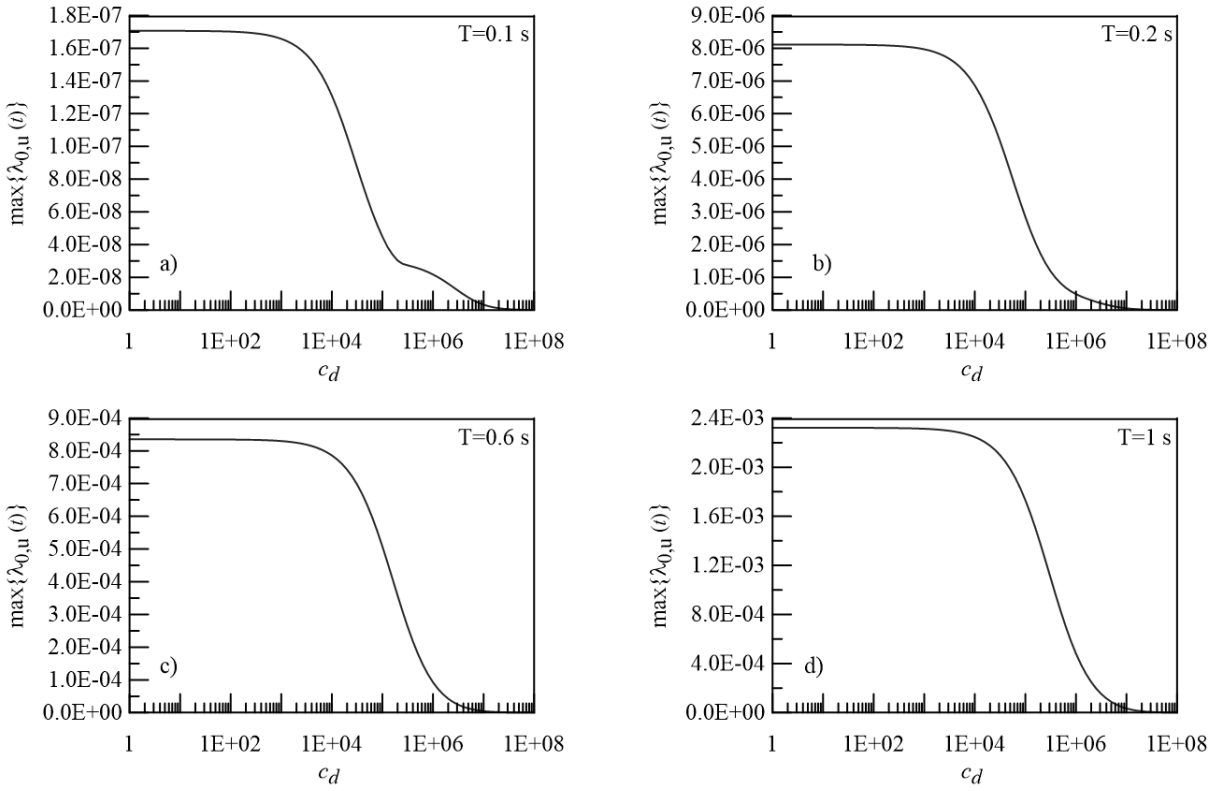


Figure 3. Maximum values of the variance of the response, $\max\left\{\lambda_{0,u}(t)\right\}$, for the four oscillators, versus the damping coefficient of the device: a) $T_0 = 0.1$ s, b) $T_0 = 0.2$ s, c) $T_0 = 0.6$ s, d) $T_0 = 1.0$ s.

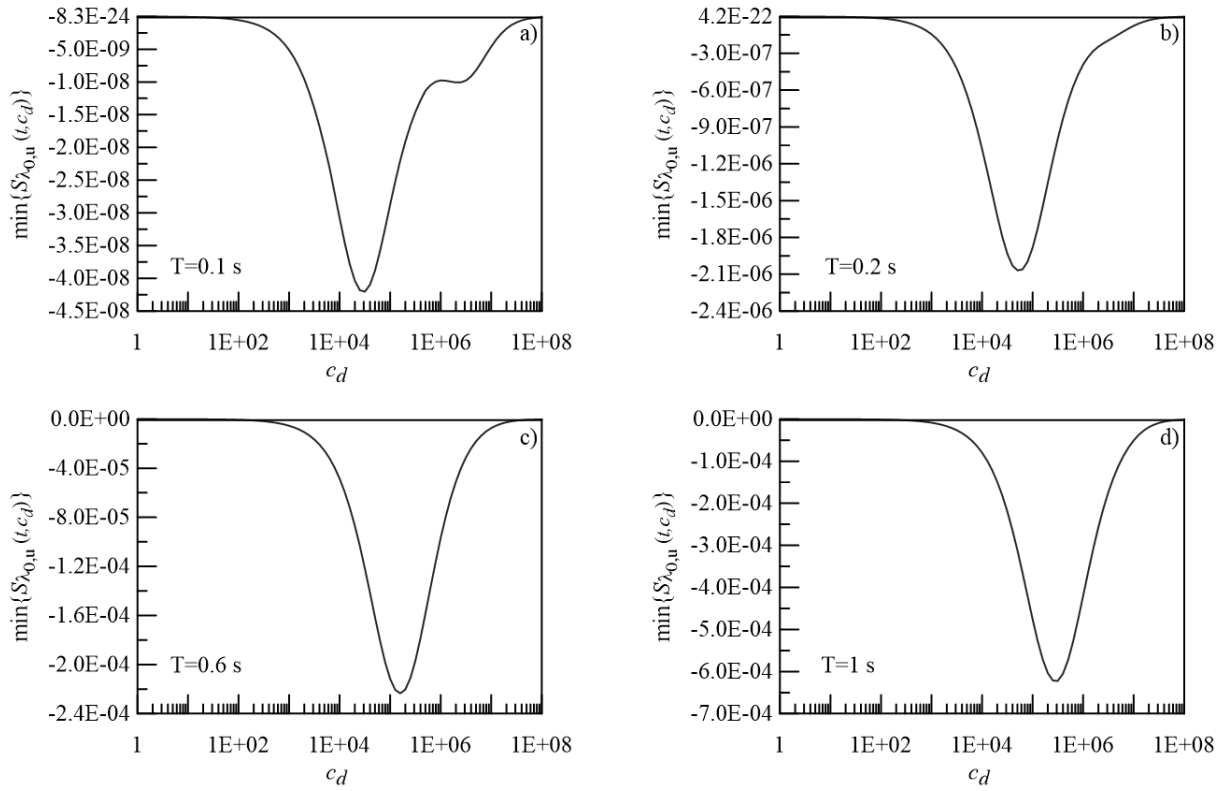


Figure 4. Minimum values of the sensitivity function of the variance of the response versus the damping coefficient , $\min\left\{S_{\lambda_0}(t,c_d)\right\}$, varying the damping coefficient of the device: a) $T_0 = 0.1$ s, b) $T_0 = 0.2$ s, c) $T_0 = 0.6$ s, d) $T_0 = 1.0$ s.

Figure 4 depicts the minimum values of the sensitivity of the variance of the response $\min\left\{S_{\lambda_0}(t,c_d)\right\} = \min\left\{\left|\partial\lambda_{0,u}(t)/\partial c_d\right|_{k_d=k_{d,0}}\right\}$, for the four analysed SDOF versus the damping coefficient of the damper device.

It is well known that for small variation of a parameter with respect the nominal one, it is possible to predict with good accuracy the variation of the response spectral moment by the knowledge of its sensitivity. Then, for negative sensitivities, the optimal value can be obtained when the sensitivity reaches its minimum value. It follows that, for the four SDOF systems, the optimal damping coefficients, reported in Table 1, are those corresponding to the points at which the minimum values of the sensitivity functions, depicted in Figure 4, assume the smallest values. Therefore, they can be selected as the optimal damping coefficients for the design dissipation devices.

Table I. Optimal damping values for each SDOF system

T [s]	c_d [Ns/m]
0.1	31 623
0.2	50 119
0.6	158 489
1.0	310 000

5. Summary and Conclusions

The present work aimed to define a new method to evaluate sensitivities of stochastic response characteristics of structural systems subjected to seismic excitations; the ground motion acceleration was herein modeled as fully non-stationary Gaussian stochastic processes.
Closed form solutions for the first-order derivatives of the *TFR* as well as of the one-sided evolutionary *PSD* (*EPSD*) of the structural response, with respect to damping parameters of devices, are evaluated. Numerical applications on different SDOF oscillators showed the accuracy and the computational efficiency of the proposed method.

References

Alderucci T. and G. Muscolino. Time-frequency varying response functions of non-classically damped linear structures under fully non-stationary stochastic excitations. *Probabilistic Eng Mech*, 54:95-109, 2018.

Alderucci T., F. Genovese and G. Muscolino. Response sensitivity of structural systems subjected to fully non-stationary random processes. *In Proceedings of the 3rd International Conference on Uncertainty Quantification in Computational Sciences and Engineering, UNCECOMP 2019*, 75-89, 2019 (a).

Alderucci T., G. Muscolino and S. Urso. Stochastic analysis of linear structural systems under spectrum and site intensity compatible fully non-stationary artificial accelerograms. *Soil Dynamics and Earthquake Engineering*, 126: 105762, 2019 (b).

Arora S. J. and EJ. Haug. Methods of design sensitivity analysis in structural optimization. *American Institute of Aeronautics and Astronautics Journal*, 17: 970-974, 1979.

Benfratello S. S. Caddemi and G. Muscolino. Gaussian and non-Gaussian stochastic sensitivity analysis of discrete structural system. *Computers and Structures*, 78:425-434, 2000.

Cacciola P., P. Colajanni and G. Muscolino. Combination of modal responses consistent with seismic input representation. *Journal of Structural Engineering (ASCE)*, 130: 47-55, 2004.

Cacciola P., P. Colajanni and G. Muscolino. A modal approach for the evaluation of the response sensitivity of structural systems subjected to non-stationary random processes. *Computer Methods in Applied Mechanics and Engineering*, 194:4344-4361, 2005.

Cacciola P. A stochastic approach for generating spectrum-compatible fully non-stationary earthquakes. *Computers and Structures*, 88: 889-901, 2010.

Chaudhuri A. and S. Chakraborty. Sensitivity evaluation in seismic reliability analysis of structures. *Computer Methods in Applied Mechanics and Engineering*, 193:59-68, 2004.

Di Paola M. and G. Petrucci. Spectral moments and pre-envelope covariances of nonseparable processes. *J Appl Mech*, 57:218-24, 1990.

Di Paola M. Transient spectral moments of linear systems. *SM Arch*; 10:225-43, 1985.

Liu Q. Sensitivity and Hessian matrix analysis of evolutionary PSD functions for nonstationary random seismic responses. *J. Eng. Mech. (ASCE)*, 138:716-720, 2012.

Michealov G., S. Sarkani and L.D. Lutes. Spectral Characteristics of Nonstationary Random Processes - A Critical Review. *Structural Safety*, 21:223-244, 1999.

Muscolino G. and T. Alderucci. Closed-form solutions for the evolutionary frequency response function of linear systems subjected to separable or non-separable non-stationary stochastic excitations. *Probabilistic Engineering Mechanics*, 40:75-89, 2015.

Muscolino G. Nonstationary pre-envelope covariances of nonclassically damped systems. *J Sound Vib*, 149:107-23, 1991.

Priestley M. Evolutionary spectra and non-stationary random processes. *Journal of the Royal Statistical Society*. *Series B (Methodological)*, 27:204-237, 1965.

Priestley M. Power spectral analysis of non-stationary random processes. *Journal of Sound and Vibration*, 6: 86-97, 1967.

Socha L. The sensitivity analysis of stochastic non-linear dynamic systems. *Journal of Sound and Vibration*, 110:271-296, 1986.

Spanos P. and G.P. Solomos. Markov approximation to transient vibration. *Journal of Engineering Mechanics (ASCE)*, 109:1134-1150, 1983.

Szopa J. Sensitivity of stochastic dynamic systems to initial conditions. *Journal of Sound and Vibration*, 97:645-649, 1984.

Yan W-J., H-P. Wan and W-X. Ren. Analytical local and global sensitivity of power spectrum density functions for structures subject to stochastic excitation. *Computers and Structures*, 182:325-336, 2017.

Comparison of fully non-stationary artificial accelerograms generation methods in non-linear dynamic analyses

Federica Genovese[1,2,*], Giuseppe Muscolino[2], Alessandro Palmeri[1]
[1]*Loughborough University, United Kingdom, Email: {F.Genovese, A.Palmeri}@Lboro.ac.uk*
[2]*University of Messina, Italy, Email: {fedgenovese, gmuscolino}@unime.it*

Abstract: Non-linear dynamic analyses are a powerful tool to assess the performance of earthquake-resistant structural and geotechnical systems. Inevitably, the validity of the predicted seismic responses depends on the fidelity of the computational model to the true behaviour of the assets being analysed and the representativeness of time histories of ground acceleration in relation to the actual seismic hazard for the site under consideration. The generation of artificial time histories is generally allowed by international seismic codes and represents a valid alternative to the use of real accelerograms, provided that the key features in the expected seismic input are preserved in the generated signals. As a consequence, different stochastic generation methods of fully non-stationary, spectrum-compatible accelerograms have been proposed in the technical literature. The effects of two alternative randomisation strategies are compared in this paper, based respectively on the use of *i*) circular wavelet transform and *ii*) evolutionary piecewise power spectral density functions. The case of a simple nonlinear structural system with an elastoplastic behaviour is addressed, as representative of a broad range of structural and geotechnical systems that experience yielding and plastic deformations under relatively intense seismic events.

Keywords: Artificial accelerogram, Ground motion record, Inelastic response spectrum, Non-stationary stochastic process, Signal processing, Spectrum-compatible accelerogram, Wavelets transform.

1. Introduction

A proper definition of the design seismic action is a fundamental step in the dynamic analysis of structural and geotechnical systems as well as in dynamic soil-structure interaction problems. Seismic design codes typically represent the earthquake-induced ground shaking in terms of pseudo-acceleration response spectrum. However, there are situations (e.g. design of structures with passive protection devices or site response analysis) in which the use of the elastic design response spectrum is not considered appropriate and time-history analyses are required.

The increasing availability of strong motion records makes the use of real accelerograms an attractive option for defining the dynamic excitation. In this respect, the selection of representative sets of accelerograms is a crucial issue as it is influenced by multiple sources of uncertainties related to the definition of the seismic hazard at the site of interest. Different procedures for the selection of proper sets of recorded accelerograms have been proposed in the literature (e.g. Katsanos et al., 2010; Bommer and Acevedo, 2004). However, due to mechanical properties and non-linear behaviour of soils, there are situations in which it is not possible to obtain the minimum number of spectrum-compatible accelerograms required by seismic codes to carry out fully dynamic analysis without applying large scale factors to each record of the set, which may distort the recorded accelerogram, leading to unrealistic input motions (Genovese et al., 2019). In these

© 2021 *F. Genovese, G. Muscolino, A. Palmeri*

situations, the use of artificial accelerograms could represent a possible alternative to realistically reproduce the seismic excitations. Since stationary artificial accelerograms are usually characterized by an excessive number of strong motion cycles and a high energy content, it becomes necessary to develop procedures to generate artificial accelerograms which suitably incorporate the large variability of the seismological parameters observed in real-life time-histories.

Over the years, procedures for the artificial generation of acceleration time histories have been proposed by several researchers. Saragoni and Hart (1974) simulates sample functions of a stochastic earthquake process, obtained by modulating contiguous regions of filtered Gaussian white noises using a deterministic time envelope function; Der Kiureghian and Crempien (1989) defined the generation process as a superposition of individually modulated stationary component processes, each representing the content in the motion in a distinct frequency band; Conte and Peng (1977) proposed a sigma oscillatory model in which each uniformly modulated process consists of the product of a real deterministic time modulating function and a stationary Gaussian sub-process.

This paper investigates two recently proposed methods for the simulation of artificial accelerograms, which both attempt to preserve the time variation of amplitude and frequency content of a target real strong motion record. The two alternative strategies are based respectively on the use of *i*) circular wavelet transform (*CWT*) and *ii*) evolutionary piecewise power spectral density (*EPSD*) function. In the first approach, the *CWT* method is used to decompose a real-valued parent (i.e. target) accelerogram into the superposition of complex-valued harmonic wavelets with complex-valued combination coefficients.

The second method, for a given target accelerogram, requires the following steps: *i*) to find a fully non-stationary model of earthquake ground motion such that the target accelerogram may be considered as one of its samples; *ii*) to evaluate the mean elastic spectrum of a set of generated fully non-stationary accelerogram samples; *iii*) to satisfy the compatibility with the elastic target response spectrum by means of an iterative procedure. In order to quantify the influence of these alternative probabilistic models on the seismic response of inelastic dynamic systems, constant-ductility response spectra have been computed for bilinear single-degree-of-freedom (SDoF) oscillators, considering different values of ductility.

2. Fully non-stationary generation methods

2.1. WAVELETS FORMULATION

The wavelet analysis consists in the expansion of a given signal in terms of “wavelets” which are generated by scaling and shifting a fixed function called “mother wavelet”. Among all different types of wavelets, the “harmonic wavelets” and “musical wavelets” proposed by Newland (1994) are particularly useful for dynamic analysis. These families of wavelets are complex-valued functions in the time domain, with a rectangular box-shaped Fourier transform in the frequency domain.

Another approach to decompose the real-valued signal $\ddot{U}_g(t)$ into the superposition of complex-valued wavelets $\psi_{\{m,n\},k}(t)$ having complex-valued combination coefficients $a_{\{m,n\},k}$ consists in the use of the circular wavelets:

$$\psi_{\{m,n\},k}(t)=\frac{1}{n-m}\sum_{j=m}^{n-1}\exp\left[\mathrm{i}\,2\pi j\left(t-\tau_{\{m,n\},k}\right)\right] \tag{1}$$

where $\tau_{\{m,n\},k} = k/(n-m)$ is a deterministic time shift of the wavelets belonging to the $\{m,n\}-$ indexed frequency band, while k is an integer number.

In the discrete wavelets transform, the complex-valued combination coefficients $a_{\{m,n\},k}$ of the parent function are calculated by a discrete convolution of the signal $\ddot{U}_{\mathrm{g}}(t)$ with the band wavelets $\psi_{\{m,n\},k}(t)$ (Cecini and Palmeri, 2015):

$$a_{\{m,n\},k} = \sum_{\ell=0}^{N} \ddot{U}_{\mathrm{g}}\left(\frac{t_\ell}{t_f}\right) \cdot \frac{\Delta t}{t_f} \cdot \psi_{\{m,n\},k}\left(t_\ell\right) \tag{2}$$

where k =0, ..., (n-m-1), $t_\ell = \ell \cdot \Delta t$ is the ℓth of the $N = t_f / \Delta t$ discrete time instants at which the signal is discretized, being Δt and t_f the sampling interval and the time duration of the signal, respectively.

In this paper, the randomisation of the target signal $\ddot{U}_{\mathrm{g}}(t)$ has been evaluated using the circular wavelet as mother wavelet.

The generation formula for the rth sample of the random process can be written as:

$$f^{(r)}(t) = 2 \sum_{\{m,n\}} \sum_{k=0}^{n-m-1} \sum_{j=m}^{n-1} \frac{\left|a_{\{m,n\},k}\right|}{n-m} \times \cos\left[\mathrm{i}\, 2\pi j\left(t - \tau_{\{m,n\},k}\right) + \theta_{\{m,n\},k} + \phi^{(r)}_{\{m,n\},k}\right] \tag{3}$$

where $\phi^{(r)}_{\{m,n\},k}$ is the rth realization of a random variable uniformly distributed over the interval $\left[0, 2\pi\right]$ while $\overline{\theta}_{\{m,n\},k} = \arg\left\{\overline{a}_{\{m,n\},k}\right\}$ is the corresponding deterministic phase of the complex-valued coefficient of the parent signal.

2.2. EVOLUTIONARY POWER SPECTRAL DENSITY FUNCTION METHOD

According to the model proposed by Muscolino et al. (2021), the second method for generating random samples of a fully non-stationary zero-mean Gaussian process considered in this paper, consists of a four-step procedure.

First, divide the time axis of the target accelerogram in n contiguous time intervals, in which a uniformly modulated process is introduced as the product of a deterministic modulating function, $a(t)$, times a stationary zero-mean Gaussian sub-process $X_k(t)$, whose power spectral density (*PSD*) function $G_{X_k}(\omega)$ is filtered by two Butterworth filters:

$$G_{X_k}(\omega) = \beta_k \left(\frac{\omega^2}{\omega^2 + \omega_{H,k}^2}\right)\left(\frac{\omega_{L,k}^4}{\omega^4 + \omega_{L,k}^4}\right)\frac{\rho_k}{\pi}\left(\frac{1}{\rho_k^2 + (\omega + \Omega_k)^2} + \frac{1}{\rho_k^2 + (\omega - \Omega_k)^2}\right); \quad k = 1, \ldots, n \tag{4}$$

where β_k is evaluated in such a way that the sub-process $X_k(t)$ possesses unit variance. All the parameters present in Eq. 4 depend on the occurrences of maxima P_k and of zero-level up-crossings $N_{0,k}^{+}$ of the target accelerogram, in the various k intervals:

$$\Omega_k \cong \frac{2\pi N_{0,k}^+}{\Delta T_k}; \quad \rho_k \simeq \frac{\pi N_{0,k}^+}{2\Delta T_k}\left[\pi - 2\frac{N_{0,k}^+}{P_k}\right] \tag{5}$$

Second, estimate, in each time interval, the parameters of modulating function $a(t)$ by least-square fitting the expected energy of the proposed model to the energy of the target accelerogram subdivided in three parts:

$$a(t) = \sum_{j=1}^{2} \bar{a}_j(t)\, \mathbb{W}(t_{j-1}, t_j) + a(t_2) \exp\left[\frac{(t-t_2)}{T_D - t_2} \ln\left(\frac{|\ddot{U}_g(T_f)|}{a(t_{n_a-1})}\right)\right] \mathbb{W}(t_2, t_3), \tag{6}$$

in which $\mathbb{W}(t_{k-1}, t_k) = \mathbb{U}(t-t_k) - \mathbb{U}(t-t_{k-1})$ is the window function, $\mathbb{U}(t)$ being the unit step function.

Third, generate the *i*th sample of the process via the formula:

$$F_0^{(i)}(t) = a(t)\sqrt{2\Delta\omega}\left[\sum_{k=1}^{n}\sum_{r=1}^{m_N} \mathbb{W}(t_{k-1}, t_k) \sin\left(r\Delta\omega t + \theta_r^{(i)}\right)\sqrt{G_{X_k}(r\Delta\omega)}\right] \tag{7}$$

$\theta_r^{(i)}$ being the random phase angles, uniformly distributed over the interval $[0, 2\pi]$; m_N is the number of parts in which the *PSD* function is discretized while $\Delta\omega$ is the frequency sampling interval.

Lastly, reduce the gap between the elastic response spectrum of target accelerogram $S_e^{(T)}$ and the mean elastic response spectrum of artificial accelerograms $\bar{S}_e^{(j-1)}$ through the introduction of a corrective iterative *PSD* function:

$$\bar{G}_{X_k}^{(j)}(\omega_k) = \bar{G}_{X_k}^{(j-1)}(\omega_k)\, \frac{S_e^{(T)}(\omega_\ell, \zeta_0)^2}{\bar{S}_e^{(j-1)}(\omega_\ell, \zeta_0)^2}$$

being $\bar{G}_0^{(0)}(\omega_k) = 1$ in the first iteration Vanmarcke and Gasparini (1977).

According to the formulation described by Genovese et al. (2020), a set of samples can be generated as:

$$\bar{F}_0^{(i)}(t) = a(t)\sqrt{2\Delta\omega}\left[\sum_{k=1}^{n}\sum_{r=1}^{m_N} \mathbb{W}(t_{k-1}, t_k) \sin\left(r\Delta\omega t + \theta_r^{(i)}\right)\sqrt{\bar{G}_0^{(j)}(r\Delta\omega) G_{X_k}(r\Delta\omega)}\right]. \tag{8}$$

3. Numerical application

In order to highlight the performance of the two proposed procedures to generate sets of accelerograms having main characteristics similar to those of the target one, and consequently useful for the non-linear time-histories analysis, constant ductility response spectra have been computed for bilinear and stiffness degrading load-deformation models, considering different displacement ductility values.

3.1. TARGET MOTION

The North-South component of the ground motion recorded at Vasquez Rocks Park during the 1994 Northridge earthquake has been downloaded from the PEER database (Ancheta et al., 2013) and has been used in the following as target accelerogram. The selected ground motion, with Moment Magnitude M_w = 6.7 and a site-source distance R_{JB} = 23.1 km (Joyner and Boore, 1979), has been recorded by a station with an average shear wave velocity in the upper 30 m equal to $V_{s,30}$ = 996 m/s (EC8, soil class "A"). The target accelerogram, having an overall duration T_f = 36.6 s and a sampling interval $\Delta t = 0.02\,\mathrm{s}$, is characterized by a peak ground acceleration a_{max} = 0.132 g, a total and Arias intensity equal to I_0 = 1.9 m^2/s^3 and I_A = 0.3 m/s, respectively, and a total number of zero-level up-crossings and of peaks equal to N_0^+= 196 and P_0 = 212, respectively. The recommended lowest usable frequency related to the filtering of the target record to remove low-frequency (long-period) noise is equal to 0.25 Hz.

3.2 WAVELETS-BASED METHOD

Due to the Heisenberg's uncertainly principle, it is not possible simultaneously localize a signal in both the time and frequency domains (Mallat, 2009).

In this paper, three different schemes of Circular Wavelet Transform (*CWT*) have been investigated to highlights how the trade-off between localizations in the two domains plays a fundamental role for the purpose of generating meaningful time histories of ground accelerations.

The three different configurations correspond to p=1, 32 and 915 frequency bands of equal bandwidth, and 50 acceleration time-histories have been generated by Eq. (3). Three generic samples together with the target accelerogram (black line) are plotted in Figure 1.

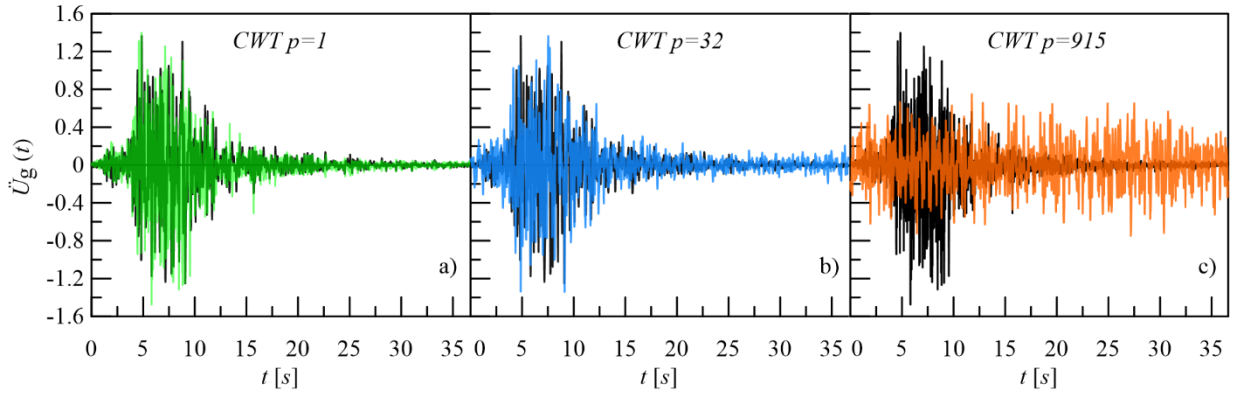


Figure 1. Comparison among the selected accelerogram (black line) and the *i*th generated sample by the proposed circular wavelets method, considering a subdivision of the frequency domain into a) p=1 (green line), b) p=32 (blue line) and c) p=915 (orange line), frequencies bands.

As the number of bands increase, the fidelity in representing the frequency content is increases but at the same time the generated samples tend to lose the fidelity in terms of non-stationary characteristics, until they become realizations of a stationary random process (see Figure 1c, with p= 915). This effect is also highlighted in Figure 2 in which the mean values of the modules of the Fourier spectra, obtained for the three analysed configurations, are compared with the target one.

In Figures 3a and 3b, a further comparison is represented in terms of cumulative energy functions $I_0(t)$ and cumulative zero-level up crossing functions $N_0^+(t)$, respectively. For p=1, the mean trend of the energy function is in a good agreement with the target one (dashed black line) while the zero-level up crossing function is very far from the real trend. The opposite happens in the third case, p=915.

Consequently, it is necessary to find, a central configuration between the 2 extreme cases, to obtain samples having characteristics sufficiently close to those of the target one both in the time and frequency domain. The case of p=32 allows to obtain a result, with a sufficient degree of accuracy, in both domains.

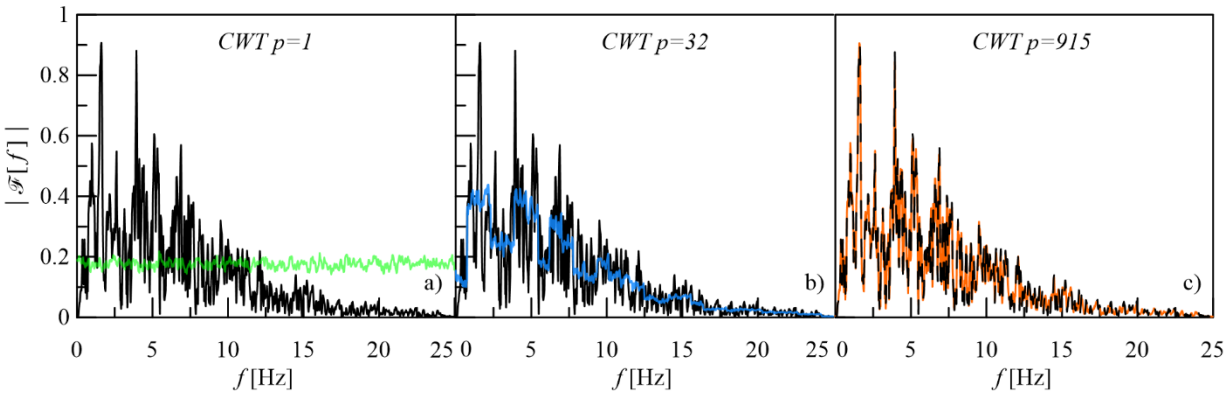


Figure 2. Comparison among the target Fourier spectrum module (black line) and the *i*th Fourier spectrum module of the generated sample by the *CWT* method, considering a) p=1 (green line), b) p=32 (blue line) and c) p=915 (orange line) frequencies bands.

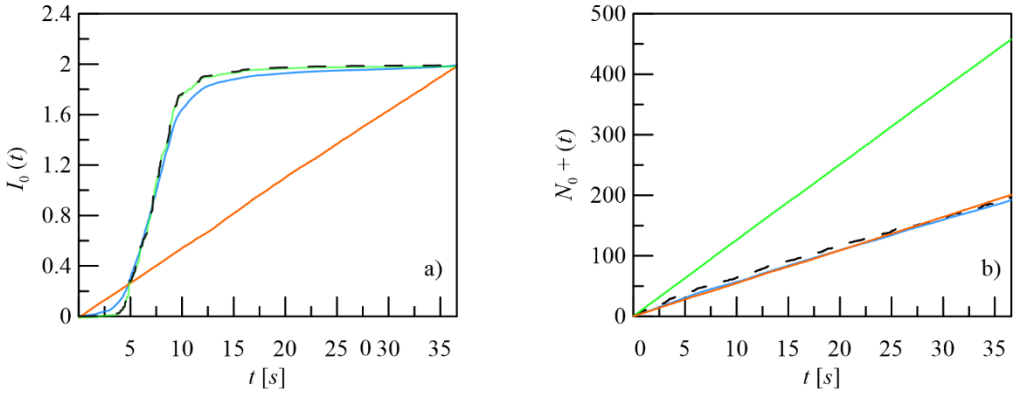


Figure 3. Comparison between the target (black, dashed lines) and the averages a) cumulative energy functions, b) zero level up crossing functions, considering p=1 (green lines), p=32 (blue lines) and p=915 (orange lines) frequencies bands.

3.3 EVOLUTIONARY POWER SPECTRAL DENSITY FUNCTION (*EPSD*) METHOD

Using the iterative procedure described in Section 2.2, a set of 50 artificial accelerograms, after 4 iterations, has been generated using appropriate modulating and *PSD* functions which allowed to preserve the amplitude and the frequency content of the target ground motion. Further details about the parameters that characterize these functions can be found in (Genovese et al., 2020).

3.4. NON-LINEAR DYNAMIC ANALYSES

In this section, inelastic response spectra have been computed to reproduce actual nonlinear structural response by means of an elastoplastic representation of a SDoF oscillator.

Specifically, the post-yield kinematic hardening ratio has been set equal to 0, i.e. an elastic-perfectly plastic system has been analyzed.

All the spectra have been obtained using the SeismoSpect 2021 Software, with a linear baseline correction, for a 2% damping value and three different ductility factors μ equal to 2, 5, and 8, respectively.

In Figure 4, the velocity spectra of the target accelerogram (red lines) are compared to that obtained as the mean value of 50 samples (black continues lines), for μ=5, considering the three different schemes analysed with the Circular Wavelet Transform method and samples generated by the Evolutionary Power spectral density function method.

The confidence intervals evaluated as the means values plus/minus the corresponding standard deviations (black dotted lines) and the velocity spectra of each sample (gray lines) are also reported in Figure 4.

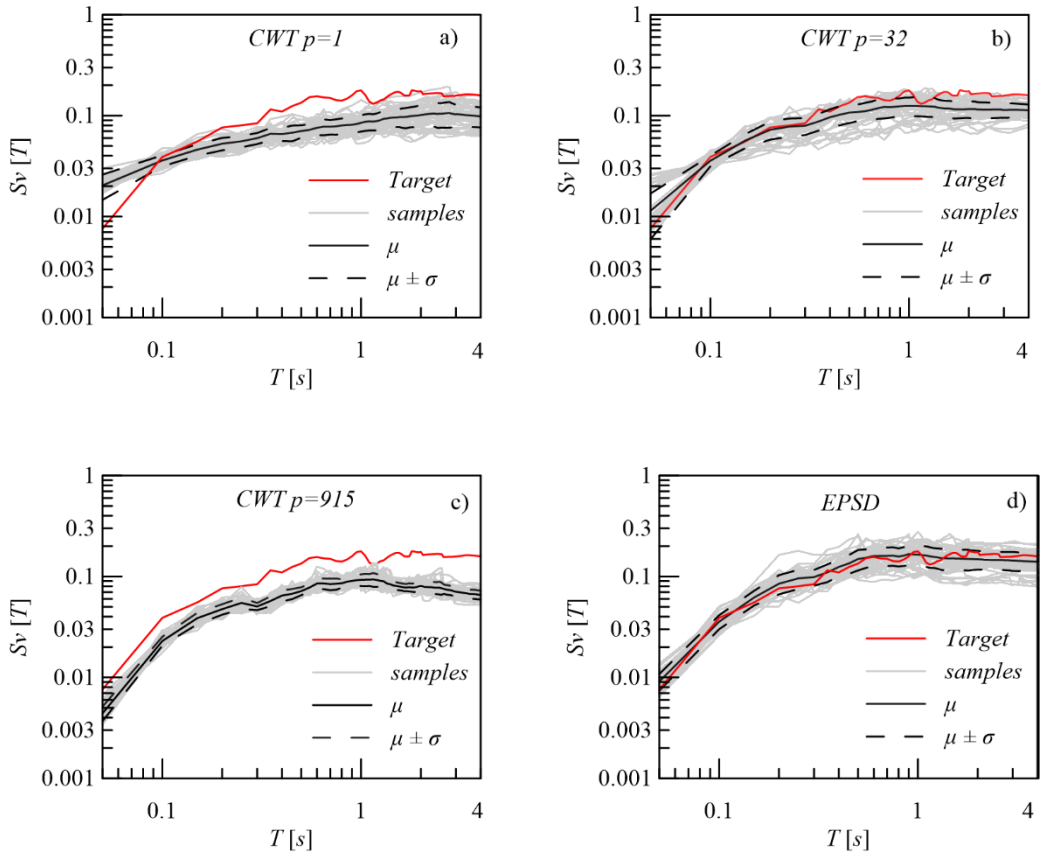


Figure 4. Comparison among the target velocity spectrum (red line) and the mean one (black, continue line), of the generated samples (gray lines) together with the corresponding mean value plus/minus standard deviation functions (black dashed lines) for a) p=1, b) p=32, c) p=915 frequencies bands, d) *EPSD* function method.

A comparison between the mean inelastic velocity spectra of the generated samples by *CWT* transform method (case p=32) and by the *EPSD* function one, for three different ductility factors, are shown in Figure 5a and b, respectively.

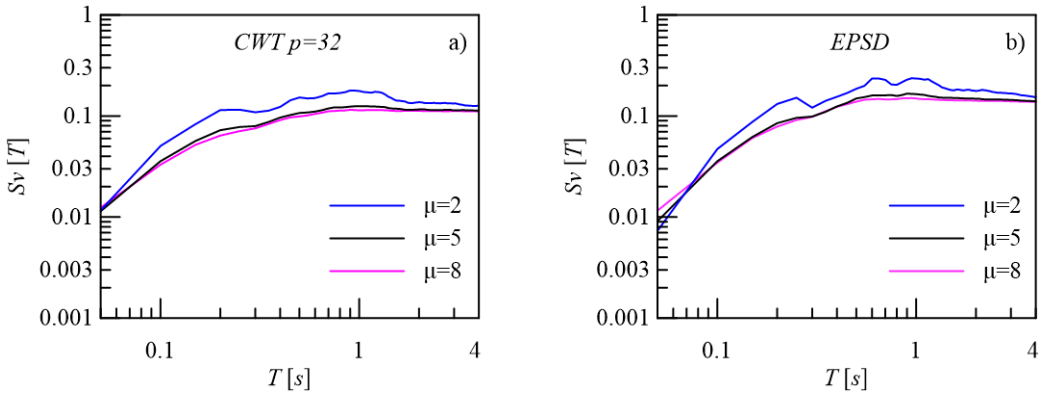


Figure 5. Comparison between mean velocity spectra evaluated for different ductility factors: a) *CWT* method p=32, b) *EPSD* function method.

Conclusions

In this paper, two recently proposed methods for generating samples of fully non-stationary zero-mean Gaussian processes, having a target acceleration time-history as one of its own samples, have been compared.

The *CWT*-method, consisting of a phase angle rotation of the circular wavelets, allows generating the required number of fully non-stationary samples without the need of defying the evolutionary power spectral density function of the ground acceleration.

The correct choice of the number of bands in which to divide the frequency domain is an important step to generate samples with the desired time-variation of amplitude and frequency content.

Further work may be needed to optimize the selection of the frequency bands, e.g. with non-uniform bandwidths.

The *EPSD* function method allows to obtain samples with characteristics closer to the target event than those generated with *CWT* method. However, the latter procedure tends to be more complex and requires several iterative steps.

The numerical results show that the average values of the inelastic velocity spectra of the generated, samples obtained by the *CWT* method, are close to the target one in the case of the subdivision of the frequency domain in 32 parts.

The other 2 configurations studied lead to results totally different from the real one. This is due to the loss of fidelity in wither the time or frequency domain, thus confirming the importance of accurately representing the seismic input in both domains.

By contrast, the application of iterative corrections in the *EPSD* method allows the target velocity spectrum to completely fall into the confidence interval evaluated as the mean value plus/minus standard deviation of the velocity spectrum of the generated samples.

References

Ancheta, T.D. et al. *PEER NGA-West2 database*. Pacific Earthquake Engineering Research Center/03; California; 2013.

Bommer, J.J. and B. Acevedo. The use of real earthquake accelerograms as input to dynamic analysis. *Journal of Earthquake Engineering*, 8, Special issue, 1, pp. 43-91. 2004

Cecini, D. and A. Palmeri. Spectrum-compatible accelerograms with harmonic wavelets. *Computers and Structures*. Vol. 147, pp.26-35, 2015.

Conte, J.P., Peng, B.-F. (1997). Fully nonstationary analytical earthquake ground-motion model. *Journal of Engineering Mechanics* (ASCE), 123(1), pp. 15-24, 1997.

Der Kiureghian A. and J. Crempien J. An evolutionary model for earthquake ground motion. *Structural Safety*, 6 , pp. 235-246, 1989.

Genovese, F., D., Aliberti, G. Biondi and E. Cascone. A procedure for the selection of input ground motion for 1D seismic response analysis. Earthquake Geotechnical Engineering for Protection and Development of Environment and constructions – Silvestri and Moraci (Eds). *Proceedings of the 7th ICEGE*, 17-20 June 2019, Rome, Italy, pp. 2591-2598, 2019.

Genovese, F., G. Muscolino, G. Biondi and E. Cascone. Generation of artificial accelerograms consistent with earthquake-induced ground motions. *Proceedings of the 11th International Conference on Structural Dynamic*, EURODYN, Virtual, Athens; Greece; 23-26 November 2020; Code 165382, Vol. 2, Pages 3027-3042, 2020.

Joyner, W.B., and D.M. Boore. Peak horizontal acceleration and velocity from strong motion records including records from the 1979 Imperial Valley, California, earthquake. *Bulletin of the Seismological Society of America* 1981; 71: 2011–2038.

Katsanos, E.I., A.G. Sextos and G.D. Manolis. Selection of earthquake ground motion records: A state-of-the-art review from a structural engineering perspective 2010. *Soil Dynamics and Earthquake Engineering*, 30 (4), pp. 157-169, 2010.

Mallat, S. A Wavelet tour of signal processing: the sparse way. *Elsevier, Academic*, 2009.

Newland D. Harmonic and musical wavelet. *Proceedings. Royal Society A*, Vol. 444, pp.605-620, 1994.

Muscolino, G., F. Genovese, G. Biondi and E. Cascone. Generation of Fully Non-Stationary Random Processes Consistent with Target Seismic Accelerograms. *Soil Dynamics and Earthquake Engineering*, 141(106467):1-14, 2021.

Saragoni, G. R. and G.C. Hart. Simulation of earthquake ground motions. *Earthquake engineering and structural dynamics*, 2, pp. 249-267, 1974.

SeismoSpect. *Seismosoft: Earthquake engineering software products*, 2021.

Vanmarcke, E.H. and D.A. Gasparini. Simulated earthquake ground motions. *Proceedings SMIRT-4 Conference*, San Francisco, paper K1/9, 1977.

Interval analysis of the forced vibration of beams with uncertain damage

F. Cannizzaro, N. Impollonia, G. Cocuzza Avellino, S. Caddemi, and I. Caliò
Department of Civil Engineering and Architecture, University of Catania, Catania 95125, Italy,
{francesco.cannizzaro, nicola.impollonia, giuseppe.cocuzzaavellino, salvatore.caddemi, ivo.calio}@unict.it

Abstract: This paper presents a study of the influence of damage intensities on the dynamic response of beams with an arbitrary number of cracks. Crack locations are considered as deterministic whereas their depths are assumed as interval uncertain but bounded variables. Explicit although approximated expressions of the main modal parameters related as a function of the crack severities are provided and the reliability of the proposed formulas is verified with the exact solutions. The methodology relies on modelling the local stiffness reduction at cracks by means of generalised functions (distributions), which allows the formulation of closed form expressions of the mode shapes and the relevant frequency equation in presence of an arbitrary number of cracks. The proposed methodology, which keeps the size of the problem constant as the number of cracks increases and avoids the need to perform re-analysis of the problem, is here adopted to assess upper and lower bounds of the response of beams with multiple cracks subjected to deterministic loads by making use of both time and frequency domain analyses. The efficacy of the proposed approach is corroborated by numerical examples on simple damaged straight beams and allows future advancements for similar analyses on damaged frames in presence of crack uncertainties.

Keywords: Dynamic response, Generalised functions, concentrated damage, Sherman-Morrison formula, interval analysis

1. Introduction

Structural deterministic analyses lead to a deterministic response. However, such an approach can be considered reliable only in presence of structures and load scenarios free of uncertainties. Nevertheless, it is widely accepted that uncertainties in the mechanical properties or in the loads acting on a structure strongly affect the expected response. This simple principle is accounted for in most of the structural codes (Eurocode 1, 1991, FEMA 356, 2000) adopted by practitioners, to assure a reasonable safety factor in the structural design and assessment.

One of the most dangerous aspects for the safety of structures is the presence of damage. The presence of impairments in a structure is often hidden and its magnitude can be mostly considered uncertain. For such a reason investigating how the structural response can be affected by a damage intensity represents a challenging issue whose solution can have a significant practical impact.

Damage can occur in structures in a diffused or concentrated pattern. Usually, especially in presence of an early detection, damage can be approximately considered concentrated. One of the most widely adopted model to simulate the presence of damage in a structure is the so called equivalent rotational hinge model (Irwin, 1957a and 1957b), according to which the presence of a damage in a straight or curved beam can be modelled by disconnecting the two sides of the beam and inserting a rotational spring whose stiffness can be related to the crack depth according to different models (Paipetis and Dimarogonas, 1986, Gudmnundson, 1983, Bilello 2001). Classical applications of structural theories follow the strategy of modelling beam-like

© 2021 *by* F. Cannizzaro, N. Impollonia, G. Cocuzza Avellino, S. Caddemi, and I. Caliò
.

structures according to a Euler-Bernoulli or Timoshenko theory, and require the enforcement of continuity conditions at the cracked sections; however, the main drawback is that the problem size increases with the number of cracked sections. An effective way to treat discontinuities due to concentrated cracks is to adopt a mathematical distributional model which avoids enforcing continuity conditions at the cracked sections. Accordingly, boundary conditions are required at the two ends of the beam only, irrespectively of the along-axis cracks present in the beam. Such an approach has been successfully applied to deterministic problems in the static, stability and dynamic fields, also with reference to frames and curved beams and to inverse problems as well (Caddemi et al., 2015, Palmeri and Cicirello, 2011, Yavari et al., 2000, Wang and Qiao, 2007, Cannizzaro et al, 2017, Di Lorenzo et al., 2017, Failla and Impollonia, 2012). However, it has never been applied to problems with uncertainties.

The modelling of uncertainties can be performed according to different models. One of the most complete and satisfactory widely used strategy is to adopt a stochastic model where a given parameter is described according to a classical probabilistic approach (Muscolino et al., 2000, Impollonia and Ricciardi, 2006, Falsone and Impollonia, 2002). In this case the output of the structural analysis is characterised by a probability density function. Other strategies make use of the fuzzy logic (Moens and Vandepitte, 2002) which assumes that a given value can be associated to an uncertain parameter with a certain level of 'truth' ranging from 0 to 1.

In the last years an alternative method which gained importance in the relevant literature is the interval analysis. Such a strategy is based on the assumption that an uncertain parameter can vary within a given range, without assuming any probability content; under such hypotheses the goal of the interval analysis is to infer the corresponding bounds of a response parameter (Alefeld and Herzberger, 1984, Moore et al. 2009). This methodology has a significant importance for the engineering practice since it provides a tool to explore the variability of the response due to the input parameter uncertainties and is able to provide the maximum and minimum values of the response, which have the greatest relevance for practical purposes. The interval analysis has been applied to several problems, mainly related to statics, dynamic frequency response of structures and transient analysis (Impollonia and Muscolino, 2011, Muscolino and Sofi, 2013, Xia et al., 2010). With regard to damaged structures, the interval analysis has been mainly employed for inverse problems (Campi et al., 2009). The direct problem for damaged beams with uncertain but bounded intensities, at least in the dynamic field, was treated in (Muscolino and Santoro, 2019) within the framework of a classic FEM approach.

In this paper, beam structures assumed as continuous Euler-Bernoulli models, are treated by making use of the generalized approach previously mentioned to model the presence of concentrated cracks, and introducing a model of uncertainty for the crack severities (Cannizzaro et al. 2020). The main modal parameters (frequencies, generalized masses and load terms in case of forced vibrations), are then approximated with explicit expressions, whose reliability is duly verified. Then, according to this approximated approach, the forced vibrations of cracked beams are assessed, and the bounds are computed in a simplified manner. In spite of the limited computational effort needed for the interval analysis, it is shown how the proposed approach is able to reproduce the bounds of the response computed according to an exact re-analysis (Kirsch, 2008) associated to a significant computational burden. Remarkably, the proposed procedure is able to account for non-monotonic relationship between input and output variables. Then, for forced vibration the extreme values of the response can be detected even at an internal point of the input variable interval, when convex analysis (Hu and Qiu, 2010) or other simplified procedures fail.

2. The adopted model

When a multi-cracked beam is considered, classic approaches require FEM analyses or the enforcement of continuity conditions at the cracked sections to obtain its response. To make more effective this procedure, a solution strategy of the governing equations of a multi-cracked beam, based on generalised functions, is here employed (Cannizzaro et al., 2018). The solution, proposed in explicit closed form and here briefly recalled, is unique and is provided in terms of four (boundary conditions dependent) integration constants only, and, unlike classic approaches, does not require any additional node, or continuity condition, at the cracked sections.

Let us consider an Euler-Bernoulli beam, with a dimensionless spatial abscissa x spanning from 0 to the length L and distributed mass m, in presence of multiple cracks at x_i=1,…,n, and a generic transverse load $\bar{q}(x,t)$, as shown in Figure 1.

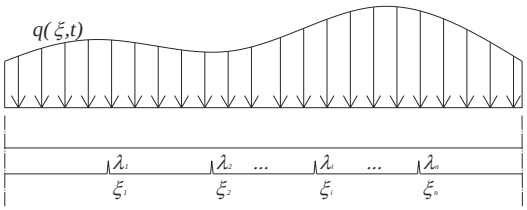


Figure 1. Damaged beam subjected to a generic time-dependent external transverse load.

By introducing the normalised abscissa $\xi=x/L$, the dimensionless governing equation of the vibratory motion of the deterministic beam in terms of normalized deflection function $u(\xi,t)=v(x,t)/L$ with t time variable, is written as follows:

$$u^{IV}(\xi,t)+\frac{mL^4}{E_oI_o}\ddot{u}(\xi,t)=\sum_{i=1}^{n}\Delta u'(\xi_i,t)\delta''(\xi-\xi_i)+q(\xi,t) \tag{1}$$

being $q(\xi,t)=\dfrac{\bar{q}(x,t)L^3}{E_oI_o}$. By introducing the dimensionless crack compliance $\lambda_i=\dfrac{E_oI_o}{K_iL}$ (being K_i the stiffness of the equivalent rotational spring), the following formula, which relates the unknown rotations at the cracked sections with the relevant bending moments, holds:

$$\Delta u'(\xi_i,t)=\lambda_i u''(\xi_i,t) \tag{2}$$

The deterministic eigenproperties associated to Eq.(1) have already been inferred in (Cannizzaro et al., 2018). In the following the damage intensities at the cracked sections will be considered as uncertain parameters collected in the vector $\boldsymbol{\lambda}=[\lambda_1, \lambda_2,\ldots, \lambda_i,\ldots, \lambda_n]$. Thus, the eigenproperties of the beam as well as its response will be dependent on the crack severities. Therefore, the generic p-th mode shape ϕ_p can be expressed as:

$$\phi_p(\xi;\boldsymbol{\lambda})=\sum_{j=1}^{4}C_{j,p}(\boldsymbol{\lambda})\,f_{j,p}(\xi;\boldsymbol{\lambda}) \tag{3}$$

where

$$f_{j,p}(\xi;\boldsymbol{\lambda})=h_{j,p}(\xi;\boldsymbol{\lambda})+\sum_{i=1}^{n}\bar{h}_{i,p}(\xi;\boldsymbol{\lambda})\lambda_i f''_{j,p}(\xi_i;\boldsymbol{\lambda}) \quad , \quad j=1,\ldots,4 \tag{4}$$

and

$$h_{1,p}(\xi;\boldsymbol{\lambda})=\sin\alpha_p\xi;\quad h_{2,p}(\xi;\boldsymbol{\lambda})=\cos\alpha_p\xi;\quad h_{3,p}(\xi;\boldsymbol{\lambda})=\sinh\alpha_p\xi;\quad h_{4,p}(\xi;\boldsymbol{\lambda})=\cosh\alpha_p\xi;$$
$$\bar{h}_{i,p}(\xi;\boldsymbol{\lambda})=\frac{1}{2\alpha_p}\left[\sin\alpha_p(\xi-\xi_i)+\sinh\alpha_p(\xi-\xi_i)\right]U(\xi-\xi_i) \tag{5}$$

being $\alpha_p^4(\boldsymbol{\lambda})=\omega_p^2 mL^4/EI$ is the p-th frequency parameter depending on the end condition of the beam and on its damage configuration. The functions $h_{j,p}(\xi,\boldsymbol{\lambda})$, j=1,…4, represent the solution of the homogeneous beam, while $\bar{h}_{i,p}(\xi;\boldsymbol{\lambda}), i=1,\ldots,n,$ represent additional functions able to account for the presence of the cracks.

The response of the beam can be obtained through modal superposition, that is once the eigenvalue problem of the multi-cracked beam has been solved for the natural frequencies and modes, the displacement is given by a linear combination of the modes:

$$u(\xi,t;\boldsymbol{\lambda})=\sum_{r=1}^{\infty}\phi_p(\xi;\boldsymbol{\lambda})z_p(t)\cong\sum_{r=1}^{n_m}\phi_p(\xi;\boldsymbol{\lambda})z_p(t) \tag{6}$$

being $z_p(t)$ the p-th modal coordinate and n_m the number of considered modes. In the modal space, by exploiting the orthogonality properties of the mode shapes and considering a viscous damping, the governing equations are

$$\ddot{z}_p(t)+2\zeta_p\alpha_p^2(\boldsymbol{\lambda})\dot{z}_p(t)+\alpha_p^4(\boldsymbol{\lambda})z_p(t)=\frac{Q_p(t;\boldsymbol{\lambda})}{M_p(\boldsymbol{\lambda})}\qquad p=1,\ldots,n_m \tag{7}$$

being ζ_p the modal damping associated to the generic p-th mode. The generalized modal mass and the modal load term appearing in Eq.(7) are given by the following formulas

$$M_p(\boldsymbol{\lambda})=m\int_0^1\phi_p(\xi;\boldsymbol{\lambda})\phi_p(\xi;\boldsymbol{\lambda})d\xi;\quad Q_p(t;\boldsymbol{\lambda})=\int_0^1\phi_p(\xi;\boldsymbol{\lambda})q(\xi,t)d\xi \tag{8}$$

The integration constants appearing in Eq.(3) can be obtained by enforcing the relevant boundary conditions. Without loss of generality, here two cases corresponding to clamped-clamped and pinned-pinned conditions are elucidated in Table I.

Table I. Boundary conditions and determinantal equations

Clamped-clamped	$\phi(0;\boldsymbol{\lambda})=0;\quad \phi'(0;\boldsymbol{\lambda})=0;$ $\phi(1;\boldsymbol{\lambda})=0;\quad \phi'(1;\boldsymbol{\lambda})=0$	$\left[f_{3,p}(1;\boldsymbol{\lambda})-f_{1,p}(1;\boldsymbol{\lambda})\right]\left[f'_{4,p}(1;\boldsymbol{\lambda})-f'_{2,p}(1;\boldsymbol{\lambda})\right]+$ $-\left[f_{4,p}(1;\boldsymbol{\lambda})-f_{2,p}(1;\boldsymbol{\lambda})\right]\left[f'_{3,p}(1;\boldsymbol{\lambda})-f'_{1,p}(1;\boldsymbol{\lambda})\right]=0$
Pinned-pinned	$\phi(0;\boldsymbol{\lambda})=0;\quad \phi''(0;\boldsymbol{\lambda})=0;$ $\phi(1;\boldsymbol{\lambda})=0;\quad \phi''(1;\boldsymbol{\lambda})=0$	$f_{1,p}(1;\boldsymbol{\lambda})\,f''_{3,p}(1;\boldsymbol{\lambda})-f_{3,p}(1;\boldsymbol{\lambda})\,f''_{1,p}(1;\boldsymbol{\lambda})=0$

3. Interval uncertainty model

An explicit approximated formulation is here proposed to assess the variability of the dynamic response of a damaged beam associated to a non-deterministic parameter in the beam configuration free of any computational burden usually required by the previously addressed exact formulation. In particular, the crack severities will be considered as uncertain variables in given intervals. According to the classical interval arithmetic, the generic variable λ_i can range in the interval $\lambda_i \in \left[\underline{\lambda}_i, \overline{\lambda}_i\right]$ characterized by the midpoint value $\lambda_{o,i}$.Given the fluctuation amplitude $\Delta\lambda_i$ the following relations hold:

$$\lambda_{o,i} = \frac{1}{2}\left(\underline{\lambda}_i + \overline{\lambda}_i\right); \qquad \Delta\lambda_i = \frac{1}{2}\left(\overline{\lambda}_i - \underline{\lambda}_i\right) \tag{9}$$

where the symbols $\underline{\bullet}$ and $\overline{\bullet}$ indicate the lower and the upper bounds, respectively. A generic value of the i-th crack severity λ_i can be then expressed as:

$$\lambda_j = \lambda_{o,j} + \beta_j, \qquad \beta_i \in \left[-\Delta\lambda_i, \Delta\lambda_i\right] \tag{10}$$

Depending on the cracks severities, the modal parameters change consequently, and need to be assessed to compute the response of the beam. In particular, the needed parameters for the generic p-th mode are the frequency parameter $\alpha_p^4(\boldsymbol{\lambda})$, the load term $Q_p(t;\boldsymbol{\lambda})$, and the generalized modal mass $M_p(\boldsymbol{\lambda})$, see Eqs. (8). Since the exact computation of such parameters requires to find the roots of the characteristic equation for the frequency parameters $\alpha_p^4(\boldsymbol{\lambda})$, and the computation of integrals depending on the p-th mode shape, in the following an approximated but explicit approach to evaluate the mentioned modal parameters is introduced.
Let us consider the case of a generic cracked beam with uncertain severities $\boldsymbol{\lambda}$ according to the model introduced in Eqs. (9). In order to provide an approximated evaluation of the main modal parameters of the damaged beam affected by uncertainty in the damage intensities, $2n$ significant damage configurations will be considered, and the correspondence between exact and approximated computations of the frequency parameter $\alpha_p^4(\boldsymbol{\lambda})$ for such configurations will be enforced. Each of the n pairs of these configurations can be defined as follows:

$$\begin{aligned} \overline{\boldsymbol{\lambda}}_s &= \left[\lambda_{o,1}, \lambda_{o,2}, ..., \overline{\lambda}_s, ...\lambda_{o,n}\right] \\ \underline{\boldsymbol{\lambda}}_s &= \left[\lambda_{o,1}, \lambda_{o,2}, ..., \underline{\lambda}_s, ...\lambda_{o,n}\right] \end{aligned} \qquad s = 1, ..., n \tag{11}$$

The following interpolating formula, proposed in (Sherman and Morrison, 1949) for the inversion of an invertible matrix and already employed in (Impollonia, 2006), is adapted for the treated case and here proposed to provide an estimation of the interval frequency parameters as functions of the uncertain damage configuration:

$$\begin{aligned} \alpha_p^4\left(\boldsymbol{\lambda}\right) &\cong \alpha_{o,p}^4 + \sum_{i=1}^{n} \frac{\beta_i a_{p,i}}{1+\beta_i b_{p,i}} \\ M_p\left(\boldsymbol{\lambda}\right) &\cong M_{o,p} + \sum_{i=1}^{n} \frac{\beta_i c_{p,i}}{1+\beta_i d_{p,i}} \qquad p = 1, ..., n_m, \quad \beta_i \in \left[-\Delta\lambda_i, \Delta\lambda_i\right] \\ Q_p\left(t;\boldsymbol{\lambda}\right) &\cong Q_{o,p}\left(t\right) + \sum_{i=1}^{n} \frac{\beta_i e_{p,i}\left(t\right)}{1+\beta_i f_{p,i}\left(t\right)} \end{aligned} \tag{12}$$

being $\alpha_{o,p}^4$, $M_{o,p}$ and $Q_{o,p}(t)$ the p-th modal parameters associated to the reference distribution $\boldsymbol{\lambda}_o=[\lambda_{o,1}, \lambda_{o,2},\ldots, \lambda_{o,n}]$ of damage, whereas $a_{p,i}$, $b_{p,i}$, $c_{p,i}$, $d_{p,i}$ and $e_{p,i}(t)$, $f_{p,i}(t)$ represent appropriate sets of coefficients evaluated explicitly by enforcing the correspondence between exact and approximate expression of the frequency parameters at the $2n$ considered crack configurations $\underline{\boldsymbol{\lambda}}_s, \overline{\boldsymbol{\lambda}}_s$, $s=1,\ldots,n$ previously defined in Eq. (11), which can be explicitly evaluated with the following formulas:

$$
\begin{aligned}
a_{p,s} &= \frac{2}{\Delta\lambda_s}\frac{\left[\alpha_p^4\left(\overline{\boldsymbol{\lambda}}_s\right)-\alpha_{o,p}^4\right]\left[\alpha_{o,p}^4-\alpha_p^4\left(\underline{\boldsymbol{\lambda}}_s\right)\right]}{\alpha_p^4\left(\overline{\boldsymbol{\lambda}}_s\right)-\alpha_p^4\left(\underline{\boldsymbol{\lambda}}_s\right)};\quad b_{p,s}=\frac{1}{\Delta\lambda_s}\frac{2\alpha_{o,p}^4-\alpha_p^4\left(\underline{\boldsymbol{\lambda}}_s\right)-\alpha_p^4\left(\overline{\boldsymbol{\lambda}}_s\right)}{\alpha_p^4\left(\overline{\boldsymbol{\lambda}}_s\right)-\alpha_p^4\left(\underline{\boldsymbol{\lambda}}_s\right)} \\
c_{p,s} &= \frac{2}{\Delta\lambda_s}\frac{\left[M_p\left(\overline{\boldsymbol{\lambda}}_s\right)-M_{o,p}\right]\left[M_{o,p}-M_p\left(\underline{\boldsymbol{\lambda}}_s\right)\right]}{M_p\left(\overline{\boldsymbol{\lambda}}_s\right)-M_p\left(\underline{\boldsymbol{\lambda}}_s\right)};\qquad d_{p,s}=\frac{1}{\Delta\lambda_s}\frac{2M_{o,p}-M_p\left(\underline{\boldsymbol{\lambda}}_s\right)-M_p\left(\overline{\boldsymbol{\lambda}}_s\right)}{M_p\left(\overline{\boldsymbol{\lambda}}_s\right)-M_p\left(\underline{\boldsymbol{\lambda}}_s\right)} \\
e_{p,s}(t) &= \frac{2}{\Delta\lambda_s}\frac{\left[Q_p\left(t;\overline{\boldsymbol{\lambda}}_s\right)-Q_{o,p}(t)\right]\left[Q_{o,p}(t)-Q_p\left(t;\underline{\boldsymbol{\lambda}}_s\right)\right]}{Q_p\left(t;\overline{\boldsymbol{\lambda}}_s\right)-Q_p\left(t;\underline{\boldsymbol{\lambda}}_s\right)};\ f_{p,s}(t)=\frac{1}{\Delta\lambda_s}\frac{2Q_{o,p}(t)-Q_p\left(t;\underline{\boldsymbol{\lambda}}_s\right)-Q_p\left(t;\overline{\boldsymbol{\lambda}}_s\right)}{Q_p\left(t;\overline{\boldsymbol{\lambda}}_s\right)-Q_p\left(t;\underline{\boldsymbol{\lambda}}_s\right)}
\end{aligned}
\tag{13}
$$

with p=1,…,n_m and s=1,…,n. In the following Figure 2, a validation in terms of accuracy of the proposed procedure is presented. A clamped-clamped beam, with a crack located at ξ_1=0.2 associated to a reference severity $\lambda_{o,1}$=0.15 in investigated; the considered amplitude for the damage severity interval is $\Delta\lambda_1$=0.15. A concentrated transversal unit step load is considered at the abscissa ξ_o=0.7. In particular, in Figures 2b, 2d, 2f the comparisons between exact and approximated parameters are shown for the first five frequencies, that is frequency parameters (Figure 2b), generalized modal mass (Figure 2d) and load term (Figure 2f); all the terms are normalized by the corresponding reference value (see Table 2). In Figure 2c, 2e, 2g the corresponding relative errors are reported. The continuous lines correspond to the exact properties, whereas the dashed lines are relative to the approximated values computed according to Eqs. (12).

Table II. Reference values of the main modal properties for the beam reported in Figure 2a

p	$\alpha_{o,p}^4$	$M_{o,p}$	$Q_{o,p}$
1	498.34	1.074	1.143
2	3257.52	0.850	-1.311
3	12315.67	1.108	1.051
4	35523.70	1.769	0.349
5	86618.26	1.645	-1.769

The goal of the present formulation is to explicitly assess the bounds of a generic response term $r(\xi_r,t;\boldsymbol{\lambda})$ at a desired abscissa ξ_r and, as a particular case, the corresponding bounds within given intervals of the uncertain parameters. Consistently with the model adopted for the uncertain parameters in Eq. (10) the minimum and maximum bounds of the response will be indicated in the following as $\underline{r}\left(\xi_r,t;\boldsymbol{\lambda}\right)$ and $\overline{r}\left(\xi_r,t;\boldsymbol{\lambda}\right)$, respectively. In order to obtain such bounds of the response, a standard procedure would require to compute $r(\xi_r,t;\boldsymbol{\lambda})$ according to the exact procedure for a finite number n_λ of configurations associated to the uncertain parameters $\boldsymbol{\lambda}$ in the given intervals, for example at given equally spaced steps. Such a procedure requires n_λ

computations of the fundamental frequencies, as well as solving the integrals in Eqs. (8); in addition, unless $n_\lambda \to \infty$ the bounds of the response can be evaluated only in an approximate way.

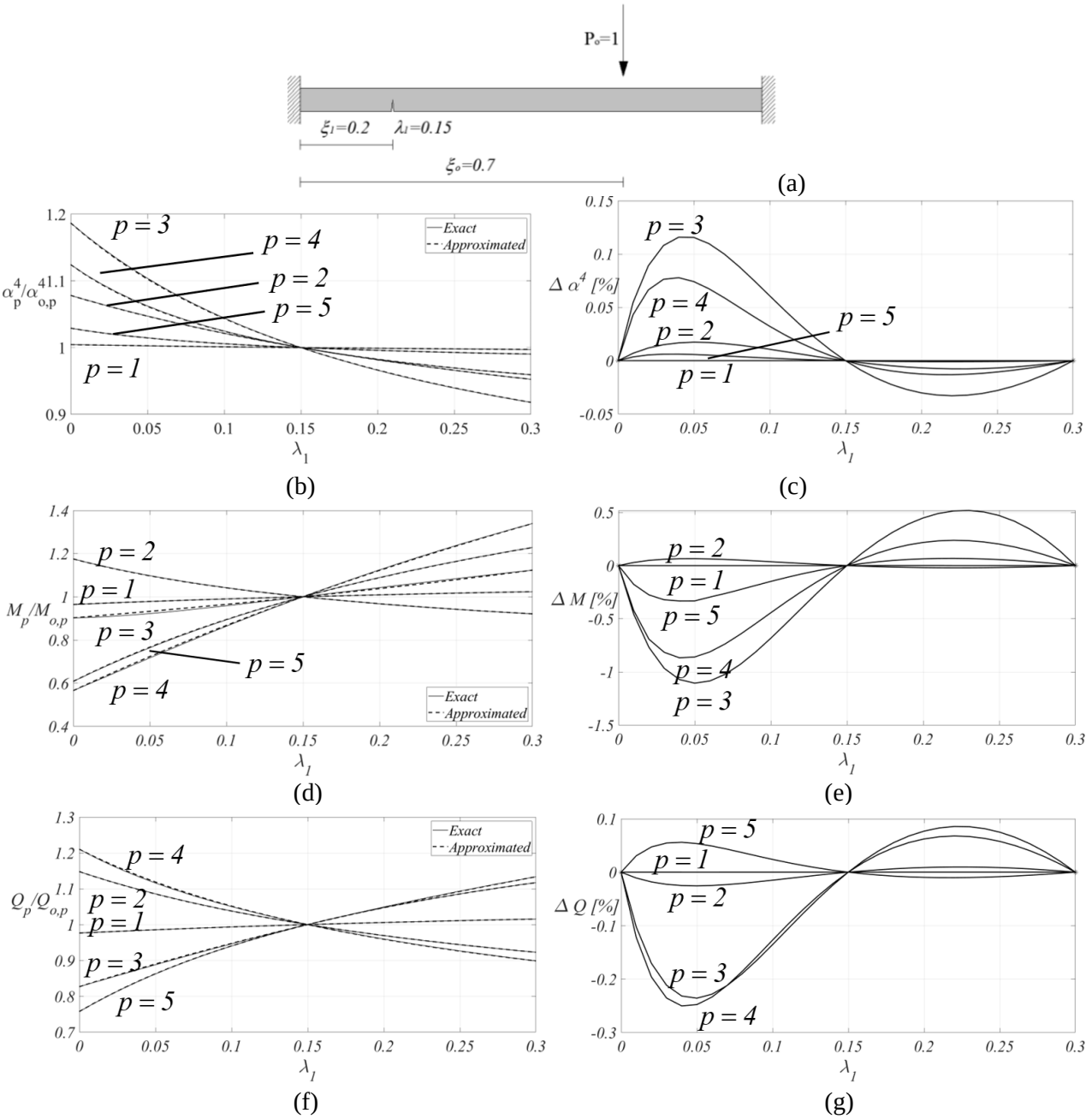


Figure 2. Comparison between exact and approximated modal properties for a (a) single cracked beam: (b) frequency parameter and (c) relative error, (d) generalized modal mass and (e) relative error, (f) load term and (g) relative error

An alternative procedure is to express the response of the beam according to the approximated but explicit formulation in terms of fundamental frequencies $\alpha_p^4(\boldsymbol{\lambda})$, generalized masses $M_p(\boldsymbol{\lambda})$ and load terms $Q_p(t;\boldsymbol{\lambda})$. The generic response of the beam can be expressed as:

$$r\left(\xi_r, t;\boldsymbol{\lambda}\right) = \sum_{p=1}^{n_m} \phi_p^k\left(\xi_r;\boldsymbol{\lambda}\right) z_p\left(t\right) \tag{14}$$

being k the order of the derivative associated to desired response $r(\xi_r,t;\boldsymbol{\lambda})$. Depending on the value assumed by k, the response in terms of deflection ($k = 0$), rotation ($k = 1$), bending moment ($k = 2$), shear force ($k = 3$) can be inferred. The mode shapes and their derivatives can be obtained in an approximated way by replacing in Eqs. (3)-(5) the approximated frequency parameter computed according to Eqs. (12); the approximated modal coordinates, considering the entire response or only the steady-state response, will require the approximated computation of the modal generalized mass and of the load term according to the scheme proposed in Eqs. (12).

Under these hypotheses, which require the exact solution of the problem at $2n+1$ given configurations of the damage pattern, namely $\boldsymbol{\lambda}_{\mathbf{o}}; \overline{\boldsymbol{\lambda}}_s, \underline{\boldsymbol{\lambda}}_s, s = 1,...,n$, the dependency of the dynamic response on the damage intensities can be explicitly provided; thus, an approximated but explicit assessment of the dynamic response variability is obtained. A further resulting important advantage of the proposed procedure lies in the possibility of inferring the bounds of a certain response parameter; to this purpose, from the symbolic response expression, which is function of the damage intensities, the bounds of the response, associated to uncertain but bounded intervals of the damage severities, can be simply obtained by solving two problems of maximum/minimum of constrained nonlinear functions.

4. Numerical applications

The explicit approach described in the previous sections is here applied with reference to two meaningful applications for the forced vibration analysis of damaged beams. The presented applications concern with the search for the bounds of response parameters within given intervals of the crack depths both in terms of response spectra and time histories.

In the first subsection, with reference to the steady-state response only, spectra of the response for a single cracked beam subjected to pulsating loads will be proposed and compared again with the results obtained through the deterministic response associated to the scanning of the damage configurations at given steps. Then, in second subsection, the bounds of the response of a double cracked beam will be assessed and compared with the response computed by scanning the damage intensity interval and applying the exact deterministic analysis for each value. In this case both the transient and the steady-state responses will be accounted for.

It will be shown how the bounds of the response are not always associated to the bounds of the intervals of the crack depth parameter, thus proving that the procedure is more robust than other approaches.

4.1. INTERVAL FREQUENCY-RESPONSE CURVES

In the dynamics of structures, a useful tool to summarize the behaviour of a structure subjected to frequency-dependent loads is the frequency-response curve. Basically, given a pulsating load with frequency ω applied to a structure, a frequency-response curve collects the magnification factor of the steady-state response with respect to the corresponding static response (Chopra, 2001). In an undamped deterministic analysis the

frequency-response curve tends to diverge when ω assumes a value corresponding to each of the natural frequencies of the system (resonance phenomenon); even in presence of damping the peaks of the responses (not divergent) can be encountered in correspondence of the damped natural frequencies. When an uncertainty of the crack severity in a beam is considered, the natural frequencies can change significantly, as already shown in Figure 2; therefore, it is expected that an interval frequency-response band will show that the bounds of the response can be associated to values of the damage severities which do not coincide with the bounds of the considered interval.

To this purpose an application is here proposed with reference to the beam reported in Figure 2a and subjected to a concentrated pulsating load at the midspan of the beam. The reference output parameter is the transversal displacement at the abscissa corresponding to the midspan of the beam. The first five frequencies are accounted for and the re-analysis is associated to 31 analyses (considering a damage step size $\Delta\tilde{\lambda}_1$ =0.01). In Figure 3 the results are shown for the case of undamped system. The steady-state response $u_{ss}(\xi_r)$/ $u_{st}(\xi_r)$ at the desired abscissa divided by the corresponding response of the beam when subjected to the static load and then normalized by the corresponding response of the beam associated to the reference damage configuration $u_{ss,o}(\xi_r)$ versus the normalized frequency parameter of the pulsating load $\eta = \alpha_\omega^2/\alpha_{o,1}^2$ is reported. The exact responses associated to three significant damage configurations are reported in the same plots, namely the reference configuration (i.e. λ_1=0.15), and those associated to the healthy beam (i.e. λ_1=0) and to the most severe configuration (i.e. λ_1=0.3) with solid, dashed, and dashdot lines, respectively, Figure 3a. Such responses are not able to envelope correctly the grey area (see for example the second and the third resonance frequencies) and the classical vertex method does not seem suitable to assess the bounds of the response when an uncertainty in the damage intensities is introduced. In Figure 3b the bounds of the response obtained with the proposed approach are superimposed to the re-analysis envelope.

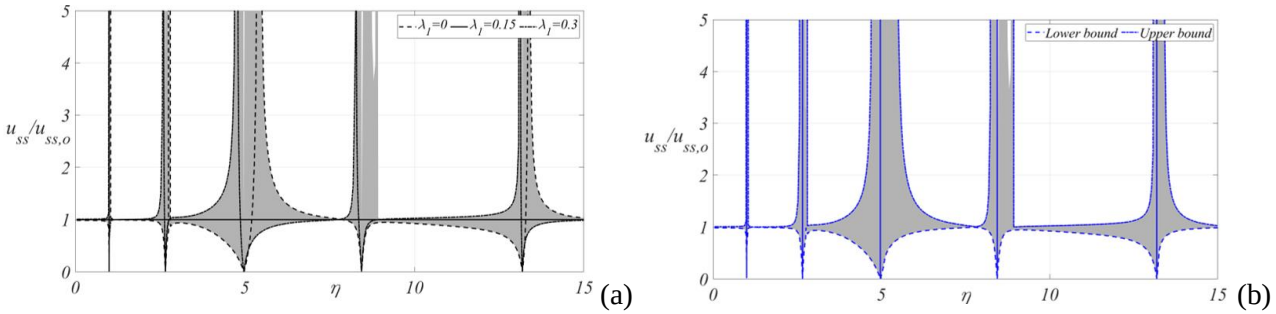


Figure 3. Normalized frequency-response interval spectrum for the undamped beam reported in Figure 2a: comparison between the exact results associated to a re-analysis (grey band) and (a) exact analyses associated to boundary intensities of damage and (b) response bounds obtained with the proposed approach

Three main remarks can be made observing the latter results:

i) the proposed approximated explicit approach is able to bind the envelope obtained with the re-analysis for the whole considered range of the load frequency;

ii) the envelope obtained through re-analysis, being based on a discrete evaluation of the response for a finite number of values of the damage severity of the crack, is itself an approximated way to assess the actual envelope of the dynamic response; theoretically, the re-analysis is able to regain the exact response variability only when an infinite number of values of the damage severities is considered;

iii) in this regard, the proposed approximated solution appears to be even more reliable than a re-analysis associated to a finite number of damage configurations (see for example the range $8 \leq \eta \leq 9$).
The dependency of the steady state response on the crack intensity is further investigated considering two values of the pulsation, namely η=4.5 (Figure 4a) and η=5.0 (Figure 4b). In the first case a monotonic trend of the monitored response parameter is encountered as the damage intensity increases; in the second case a non-monotonic trend is observed, and the response tends to diverge for a damage intensity $\lambda \approx 0.1278$, thus showing that a vertex method is not suitable for the treated case.

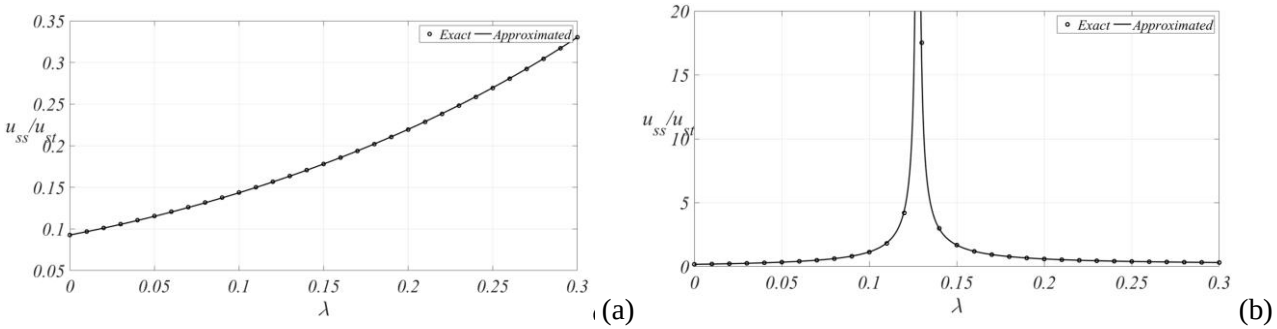


Figure 4. Crack intensity vs steady state response for the beam reported in Figure 2a for specific frequency ratios: (a) η =4.5 and (b) η =5.0

4.2. INTERVAL TIME-DOMAIN RESPONSE

The effectiveness of the proposed approximated solution is assessed in this subsection by comparing the bounds of the response computed according to a scanning exact procedure with those obtained through the procedure proposed in this paper.
Aiming at highlighting the advantages offered by the proposed procedure when multiple cracks are considered, in Figure 5 the results relative to a pinned-pinned double cracked beam as reported in Figure 5a are presented; the beam is subjected to a concentrated pulsating load acting on the midpoint of the beam, $q(\xi,t)=P_o\delta(\xi-\xi_L)\sin\omega t$ (P_o=1, ξ_L =0.5, ω=10 rad/s). The two cracks are located at ξ_1=0.2 and ξ_2=0.7, respectively, and are both characterized by the reference intensity $\lambda_1 = \lambda_2$=0.05. The approximated modal parameters are computed considering a fluctuation amplitude $\Delta\lambda_1 = \Delta\lambda_2$=0.05. Again, the grey area represents the envelope of the exact responses computed scanning the damage intensities at evenly spaced steps $\Delta\tilde{\lambda}_1 = \Delta\tilde{\lambda}_2 = 0.01$; therefore, 121 exact analyses were performed, that is the computational burden increases with the second power of the number of cracks. On the other hand, the results relative to the interval analysis obtained with the proposed approach require the computation of the exact modal parameters for five configurations, that is $\boldsymbol{\lambda}_o = [0.05, 0.05]$, $\overline{\boldsymbol{\lambda}}_1 = [0.1, 0.05]$, $\underline{\boldsymbol{\lambda}}_1 = [0, 0.05]$, $\overline{\boldsymbol{\lambda}}_2 = [0.05, 0.1]$, $\underline{\boldsymbol{\lambda}}_2 = [0.05, 0]$, and then solving two maximum/minimum conditioned problems for each time step. The monitored output parameter is again the transversal displacement at the midpoint of the beam, and the first three frequencies are considered with an associated damping equal to 0.05 for all the modes.
In Figure 5b and 5c the grey area represents the ensemble of the exact responses computed through re-analysis. For convenience, the exact responses for the reference damage, and those associated to the lowest

and the highest crack severities (i.e. $\lambda_1 = \lambda_2 = 0$ and $\lambda_1 = \lambda_2 = 0.1$) are reported in Figure 5b with continuous, dashed, and dashdot lines, respectively. It may be noticed that the bounds of the responses identified by the contour of the grey area cannot be identified by simply running the three mentioned exact analyses, and even running a larger number of analyses for given values of damage severities (in this case 121 exact analyses were run), the correct bounds of the response can be identified only in an approximate way. Theoretically the correct bounds of the response would require a very large number of analyses. Conversely, in Figure 5c the bounds of the response are computed according to the proposed procedure. The relevant lower and upper bounds are reported with the blue dashed and dashdot lines, respectively. The approach here proposed is able to effectively assess the bounds of the response for all the time history analysis, and it is also associated to a significantly lower computational burden with respect to a re-analysis approach.

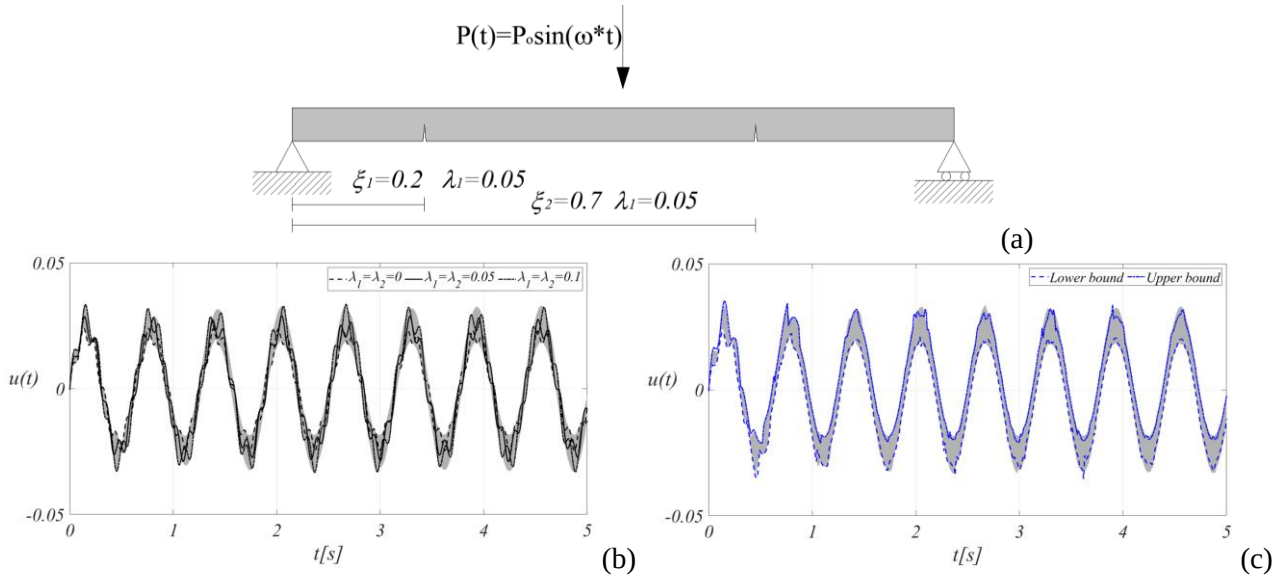


Figure 5. Interval time-domain response for the (a) beam subjected to a concentrated pulsating load: comparison between the exact results associated to a re-analysis (grey band) and (b) exact analyses associated to boundary intensities of damage and (c) response bounds obtained with the proposed approach

5. Conclusions

In this paper uncertainties in the dynamic response of damaged beams are treated within the framework of the interval analysis; in particular, damage severities are considered uncertain. To this purpose, a model previously introduced for the deterministic dynamic analysis of multi-cracked beams, able to avoid the enforcing of continuity conditions at the cracked sections, is employed. Here, taking advantage of an approximated approach to express main modal parameters (frequency parameters, generalized modal mass and load term), the response to the forced vibration of cracked beam is inferred explicitly, thus making possible to easily infer the bounds of the response associated to given intervals of the damage severities. The accuracy of the approximated approach for the evaluation of the modal parameters is compared with reference solutions; then, the method is applied to single and multi-cracked beams, both with reference to transient time history analyses and to the steady-state response. The proposed approach is much faster with respect to a

classic scanning approach and brought to light that for the forced vibrations of damaged structures the bounds of the response are not necessarily associated to the most and less severe configurations, but might involve intermediate damage levels.

Acknowledgements

The authors gratefully acknowledge the financial support of the Ministero dell'Istruzione, dell'Università e della Ricerca (National Research Project PRIN 2015JW9NJT "Advanced mechanical modeling of new materials and structures for the solution of 2020 Horizon challenges").

References

Alefeld, G., Herzberger, J. Introduction to interval computation, *Academic press*, 1984.

Bilello, C. Theoretical and experimental investigation on damaged beams under moving systems, Ph.D. Thesis, University of Palermo, Italy, 2001.

Caddemi, S., Caliò, I., Cannizzaro, F. Influence of an elastic end support on the dynamic stability of Beck's column with multiple weak sections. *International Journal of Nonlinear Mechanics* 69:14-28, 2015.

Campi, M.C., Calafiore, G., Garatti, S. Interval predictor models: Identication and reliability. *Automatica* 45(2):382-392, 2009.

Cannizzaro, F., Greco, A., Caddemi, S., Caliò, I. Closed form solutions of a multi-cracked circular arch under static loads. *International Journal of Solids and Structures* 121(15):191-200, 2017.

Cannizzaro, F., De Los Rios, J., Caddemi, S., Caliò, I., Ilanko, S. On the use of a roving body with rotary inertia to locate cracks in beams. *Journal of Sound and Vibration* 425(7):275-300, 2018.

Cannizzaro, F., Impollonia, N., Caddemi, S., Caliò, I. Explicit dynamic response of damaged beams with application to uncertain and identification problems, *Journal of Sound and Vibration*, 487, 2020.

Chopra, A.K. Dynamics of Structures. *Prentice Hall*, 2001.

Di Lorenzo, S., Di Paola, M., Failla, G., Pirrotta, A. On the moving load problem in Euler–Bernoulli uniform beams with viscoelastic supports and joints. *Acta Mechanica* 228(3):805-821, 2017.

ENV 1991-1, Eurocode 1-Basis of Design and Actions on Structures-Part 1: Basis of Design, 1994.

Falsone, G., Impollonia, N. A new approach for the stochastic analysis of finite element modelled structures with uncertain parameters. *Computer Methods in Applied Mechanics and Engineering* 191(44):5067-5085, 2002.

FEMA 356 Federal Emergency Management Agency (2000), Prestandard and Commentary for the Seismic Rehabilitation of Buildings.

Gudmnundson, P. The dynamic behaviour of slender structures with cross-sectional cracks. *Journal of Mechanical Physics and Solids* 31:329–345, 1983.

Hu, J., Qiu, Z. Non-probabilistic convex models and interval analysis method for dynamic response of a beam with bounded uncertainty. *Applied Mathematical Modelling* 34(3):725-734, 2010.

Impollonia, N. A method to derive approximate explicit solutions for structural mechanics problems. *International Journal of Solids and Structures* 43:7082–7098, 2006.

Impollonia, N., Muscolino, G. Interval analysis of structures with uncertain-but-bounded axial stiffness. *Computer Methods in Applied Mechanics and Engineering* 200(21-22):1945-1962, 2011.

Impollonia, N., Ricciardi, G. Explicit solutions in the stochastic dynamics of structural systems. *Probabilistic Engineering Mechanics* 21:171–181, 2006.

Irwin, G. R. Analysis of stresses and strains near the end of a crack traversing a plate. *Journal of Applied Mechanics* 24:361–364, 1957a.

Irwin, G.R. Relation of stresses near a crack to the crack extension force. In *Proceedings of the 9th Congress on Applied Mechanics*, Brussels, 1957b.

Kirsch, U. Reanalysis of Structures. In: *Reanalysis of Structures. Solid Mechanics and Its Applications*, vol 151. Springer, Dordrecht, 2008.

Moens, D., Vandepitte, D. Fuzzy Finite Element Method for Frequency Response Function Analysis of Uncertain Structures. *AIAA Journal* 40(1):126-136, 2002.

Moore, R.E., Kearfott, R.B., Cloud, M.J. Introduction to interval analysis. *Society for Industrial and Applied Mathematics*, 2009.

Muscolino, G., Ricciardi, G., Impollonia, N. Improved dynamic analysis of structures with mechanical uncertainties under deterministic input. *Probabilistic Engineering Mechanics* 15:199–212, 2000.

Muscolino, G., Santoro, R. Dynamics of multiple cracked prismatic beams with uncertain-but-bounded depths under deterministic and stochastic loads. *Journal of Sound and Vibration* 443:717-731, 2019.

Muscolino, G., Sofi, A. Bounds for the stationary stochastic response of truss structures with uncertain-but-bounded parameters. *Mechanical Systems and Signal Processing* 37(1):163-181, 2013.

Paipetis, S.A., Dimarogonas, A.D. Analytical Methods in Rotor Dynamics, *Elsevier Applied Science*, London, 1986.

Palmeri, A., Cicirello, A. Physically-based Dirac's delta functions in the static analysis of multi-cracked Euler-Bernoulli and Timoshenko beams. *International Journal of Solids and Structures* 48(14-15):2184-2195, 2011.

Sherman, J., Morrison, W.J. Adjustment of an Inverse Matrix Corresponding to Changes in the Elements of a Given Column or a Given Row of the Original Matrix (abstract). *Annals of Mathematical Statistics* 20:621, 1949.

Yavari, A., Sarkani, S., Moyer Jr., E.T. On applications of generalized functions to beam bending problems. *International Journal of Solids and Structures* 37(40):5675-5705, 2000.

Wang, J., Qiao, P. Vibration of beams with arbitrary discontinuities and boundary conditions. *Journal of Sound and Vibration*, 308(1-2):12-27, 2007.

On the guarantees derived from a possibilistic interpretation of ensemble predictions and their operational use

Noémie Le Carrer
Institute for Risk and Uncertainty, Dep. of Engineering, University of Liverpool, United Kingdom

Abstract.
Ensemble forecasting has become popular in weather prediction to reflect the uncertainty about high-dimensional, nonlinear systems with extreme sensitivity to initial conditions. By means of small strategical perturbations of the initial conditions, sometimes accompanied with stochastic parameterisation schemes of the atmosphere-ocean dynamical equations, ensemble forecasting aims at sampling possible future scenarii and ideally at interpreting them in a Monte-Carlo-like approximation. Traditional probabilistic interpretations of ensemble forecasts do not take epistemic uncertainty into account, nor the fact that ensemble predictions cannot always be interpreted in a density-based manner due to the strongly nonlinear dynamics of the atmospheric system. As a result, probabilistic predictions are not always reliable, especially in the case of extreme events. In this work, we investigate whether relying on possibility theory can circumvent these limitations. We show how it can be used to compute confidence intervals with guaranteed reliability, when a classical probabilistic postprocessing technique fails to do so in the case of extreme events. We empirically illustrate our approach with an imperfect version of the Lorenz 96 model, and demonstrate that it is promising for risk-averse decision-making.

Keywords: Possibility theory; Decision-making under uncertainty; Reliability; Statistical postprocessing; Extreme event prediction; Imprecise probabilities; Risk-averse decision-making

1. Introduction

In weather forecasting, it is acknowledged that by design (limited size of set of ensemble predictions — EPS, targeted sampling of initial conditions — ICs) and by context (flow-dependent regime error, strongly nonlinear system), raw ensemble forecasts generally do not provide reliable probabilistic predictions (Bröcker and Smith, 2008; Gneiting and Katzfuss, 2014). This is especially the case for extreme events (Legg and Mylne, 2004). The latter result from nonlinear interactions at small scales, which implies that they generally cannot be associated with a high density of ensemble members (Mylne et al., 2002). Ensemble forecasts are made more reliable and operational via recalibration (Buizza, 2018), whose aim can be summarized as "finding the transformation that, applied to the raw ensemble, leads to the probability distribution that will maximise a performance metric on a training set". In spite of the diversity of approaches developed in the literature (Buizza, 2018) and their technical success for improving the prediction skills when it comes to common events, the actionability of probabilistic predictions often remains problematic (Smith, 2016). In

© 2021 *Noémie Le Carrer.*

particular, the probabilistic prediction of extreme events often needs a development on its own (Friederichs and Hense, 2007; Friederichs et al., 2018).

(Bröcker and Smith, 2008) questioned whether probability distributions constitute *the best representation of the valuable information contained in an EPS*. We advance convincing arguments that possibility theory, "a weaker theory than probability [...] also relevant in non-probabilistic settings where additivity no longer makes sense" (Dubois et al., 2004), is an interesting alternative. Our investigation is particularly relevant since conceptual and practical limitations restrict the applicability of a density-based (i.e. additive) interpretation of EPSs. We show how interpreting EPSs in a possibilistic way brings useful formal guarantees on the derived confidence intervals, even in the case of extreme events.

Section 2 summarises the basics of possibility theory, Section 3 presents our possibilistic framework and discusses the theoretical guarantees that can be associated with its outputs. Section 4 introduces the synthetic experiments on the Lorenz 96 system (Lorenz, 1996) (L96) which allow us to assess these guarantees and their operational cost for both common and extreme events. We compare them with the outputs of a classical probabilistic interpretation of EPSs, and discuss our results in Section 5.

2. Possibility theory

Possibility theory is an uncertainty theory developed from fuzzy set theory by (Zadeh, 1978) and (Dubois and Prade, 2012). It is designed to handle incomplete information and represent ignorance. Considering a system whose state is described by a variable $x \in \mathcal{X}$, the possibility distribution $\pi : \mathcal{X} \to [0, 1]$ represents the available information about the current state of the system. Given an event $A = \{x \in S_A\}$, where S_A is a subset of $\mathcal{X}$, the possibility and necessity measures are defined respectively as: $\Pi(A) = \sup_{x \in S_A} \pi(x)$ and $N(A) = 1 - \Pi(\bar{A})$ where $\bar{A}$ represents the complementary

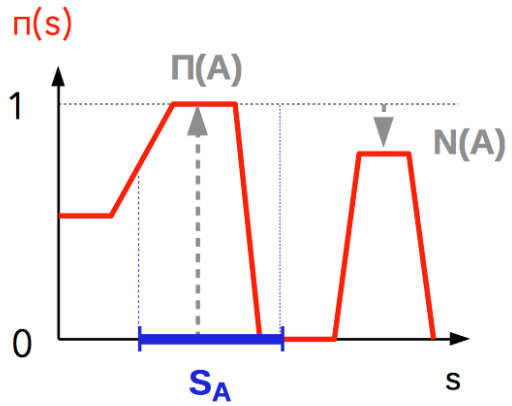


Figure 1. Possibility distribution $\pi(s)$ where for an event of interest $A =$ "$s \in S_A$", the possibility $\Pi(A)$ and necessity $N(A) = 1 - \Pi(\bar{A})$ measures are represented.

event of A (see Figure 1 for a visual understanding of these quantities). Both measures satisfy the following axioms and conventions (Cayrac et al., 1994):

1. $\Pi(\mathcal{X}) = 1$ and $\Pi(\varnothing) = 0$
2. $\Pi(A \cup B) = \max(\Pi(A), \Pi(B))$
3. $N(A) = 1 \Leftrightarrow \Pi(\bar{A}) = 0$ indicates that A has to happen, it is necessary: $\bar{A}$ is impossible;
4. $0 < N(A) < 1$ is a tentative acceptance of A to a degree $N(A)$;
5. $(\Pi(A) = \Pi(\bar{A}) = 1) \Leftrightarrow (N(A) = N(\bar{A}) = 0)$ represents total ignorance: the evidence doesn't allow us to conclude whether A is rather true or false.

Possibility and probability distributions are interconnected through the concept of imprecise probabilities (Dempster, 2008). A probability measure P and possibility measure Π are consistent iff (Dubois et al., 2004):

$$P(A) \leq \Pi(A), \quad \forall\, A \tag{1}$$

The definition of necessity implies that in these conditions:

$$N(A) \leq P(A) \leq \Pi(A), \quad \forall\, A \tag{2}$$

.

2.0.0.1. *From data to possibility distribution* Let $x \in \mathcal{X}$ be a stochastic variable for which we try to make a prediction. The evidence about x_t is a set $S = \{x_1, ..., x_{N_s}\}$ of N_s samples of x, which we assume has been randomly generated from an unknown probability distribution P. To turn this information into a possibility distribution describing the knowledge on the actual value of x, we use the technique developed by (Masson and Denœux, 2006). Their methodology is specifically designed to derive a possibility distribution from scarce data. The idea is, after binning the x-axis into n bins, to recover the simultaneous confidence intervals at level β on the true probability $P(x \in b_i)$ for each bin b_i. From these confidence intervals and considerations about Equation (1), the procedure allows us to compute a possibility distribution $\pi(x)$ that *dominates* with confidence β the true probability distribution (i.e. Equation (1) is verified in $100\beta\%$ of the cases). The simultaneous confidence intervals for multinomial proportions are computed by means of Goodman's formulation (Goodman, 1965). This procedure takes into account the uncertainty on the multinomial proportions that is due to the limited size of S. This is fundamental for our application, which is to seek guarantees on the possibility of observing a given event.

As shown by Equation (2), a possibility distribution can be seen as a complete and consistent framework to deal with imprecise probabilities. Although the above procedure for computing a possibility distribution mostly relies on probabilities, its result contains more information than a purely probabilistic distribution *in the situation of incompleteness* (typically implied by a small dataset S). Indeed, the interval on the true probability allows the incompleteness of data or knowledge to be accounted for, while a point probability hides the fact that the said probability cannot be fully trusted (e.g. due to epistemic uncertainty). Figure 2 illustrates the results of this

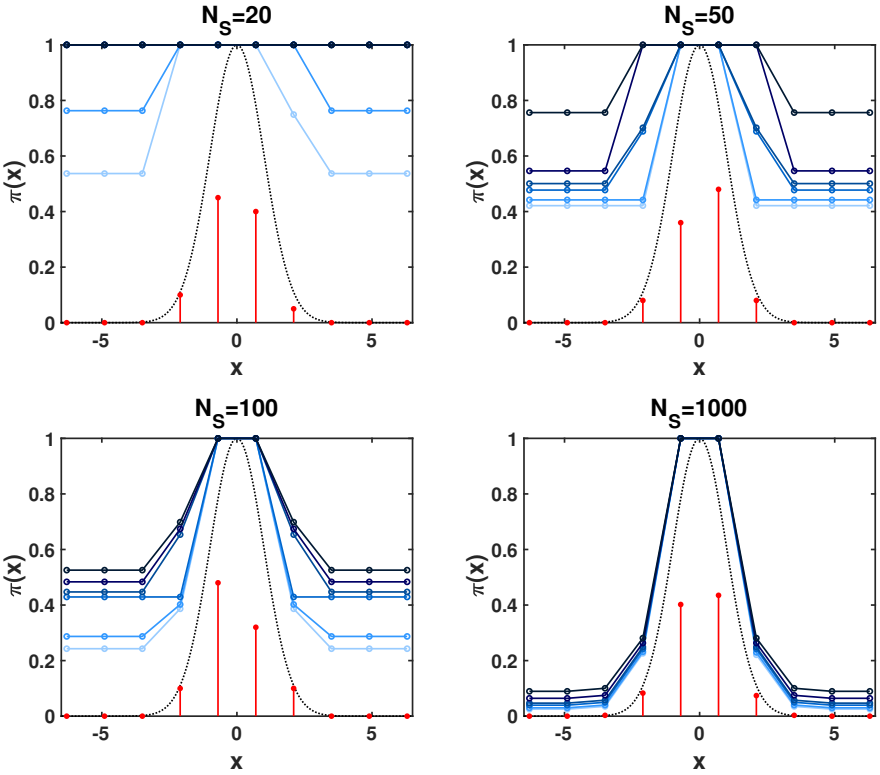


Figure 2. Possibility distributions (solid lines) derived from datasets of N_s elements sampled from a standard normal distribution. This derivation requires the computation of simultaneous confidence intervals for multinomial proportions over the x-axis binned into $n = 10$ bins. The effect of the confidence level $\beta = \{0.6, 0.75, 0.9, 0.95, 0.99, 1\}$ of the Goodman's formulation is shown (larger β are plotted darker). Vertical red lines represent a frequency histogram of the same datasets and the normalised underlying Gaussian distribution is represented as a dotted line.

methodology applied to datasets sampled from a normal distribution, for various levels of β and N_s. For a given N_s, the larger β is, the more conservative is the distribution: γ such as $\pi(x) \geq \gamma \; \forall x$ is larger, which implies that for any event $A \subset \mathcal{X}$: $\Pi(\bar{A}) \geq \gamma$. This also reads: $N(A) \leq 1 - \gamma$, meaning that the confidence level associated with any A cannot reach high values. Increasing N_s reduces the relative effect of β and all distributions tend in shape towards the underlying probability distribution, even if the tails remains more conservative for larger β.

3. Proposed framework

We are interested in the prediction of the state variable x_{t_0+t} of a dynamical system at lead time t, starting from the IC x_{t_0}. For simplicity, we omit the reference to t_0 and note x_t the *verification*. In the EPS context, given a numerical prediction model $\mathcal{M}$, the elements of information at hand are:

1. An ensemble of M predictions at lead time t, the ensemble members or EPS, obtained by means of $\mathcal{M}$ applied to slightly perturbed ICs around t_0: $\tilde{\boldsymbol{x}}_t = \{\tilde{x}_t^1, ..., \tilde{x}_t^M\}$.

2. An archive $\mathcal{I}_t$ containing the pairs $(\tilde{\boldsymbol{x}}_{t_0+t}, x_{t_0+t})$ for the lead time t of interest and N_I different instances of t_0. These instances are chosen so that the initial points of two successive trajectories are statistically independent from each other.

3.1. Deriving possibility distributions from EPSs

The objective of our possibilistic interpretation of EPSs is to derive from an EPS $\tilde{\boldsymbol{x}}_t$ and the archive $\mathcal{I}_t$ a possibility distribution $\pi(x_t|\tilde{\boldsymbol{x}}_t, \mathcal{I}_t)$, that encodes the knowledge derived from $\tilde{\boldsymbol{x}}_t$ about the verification x_t. The procedure described in this section is summarised and illustrated in the steps 1—5 of Figure 3.

Both system and model being (to a certain extent) deterministic and (close to) stationary, the past behaviour of the couple {system, model} is representative of its future behaviour. Consequently, if we are able to enumerate the possible values (already seen in $\mathcal{I}_t$ or not) for the verification x_t associated with a small range S_{x_t} of the values taken by ensemble members, then a future observation x_t should belong to that set of possible values when an ensemble member $\tilde{x}_t^m$ falls within S_{x_t}. Beyond that, we would like to know which ones of these values are more possible than others for x_t. In other words, we want to estimate the possibility distribution $\pi(x_t|\tilde{x}_t^m \in S_{x_t})$. Because there is no notion of 'density' of the evidence in the possibilistic perspective (at least in our rationale for choosing this framework), the number of ensemble members falling in S_{x_t} will not affect the resulting possibility distribution for x_t.

To make use of the full set of ensemble members, we first partition the x-axis into n bins b_i, take the subset B of bins occupied by at least one ensemble member of the EPS, and compute $|B|$ possibility distributions $\pi(x_t|\tilde{x}_t^m \in b_j)$ where $b_j \in B$. Namely, for each bin $b_j \in B$ occupied by at least one ensemble member $\tilde{x}_t^m \in \tilde{\boldsymbol{x}}_t$, we retrieve the N_s ensemble members $\tilde{x}_t^m \in b_j$ in the archive $\mathcal{I}_t$ and build a histogram of the set of corresponding verifications (so-called *analogs*) over the same binned x-axis. We then derive $\pi(x_t|\tilde{x}_t^m \in b_j)$ following the methodology presented in Section 2.0.0.1.

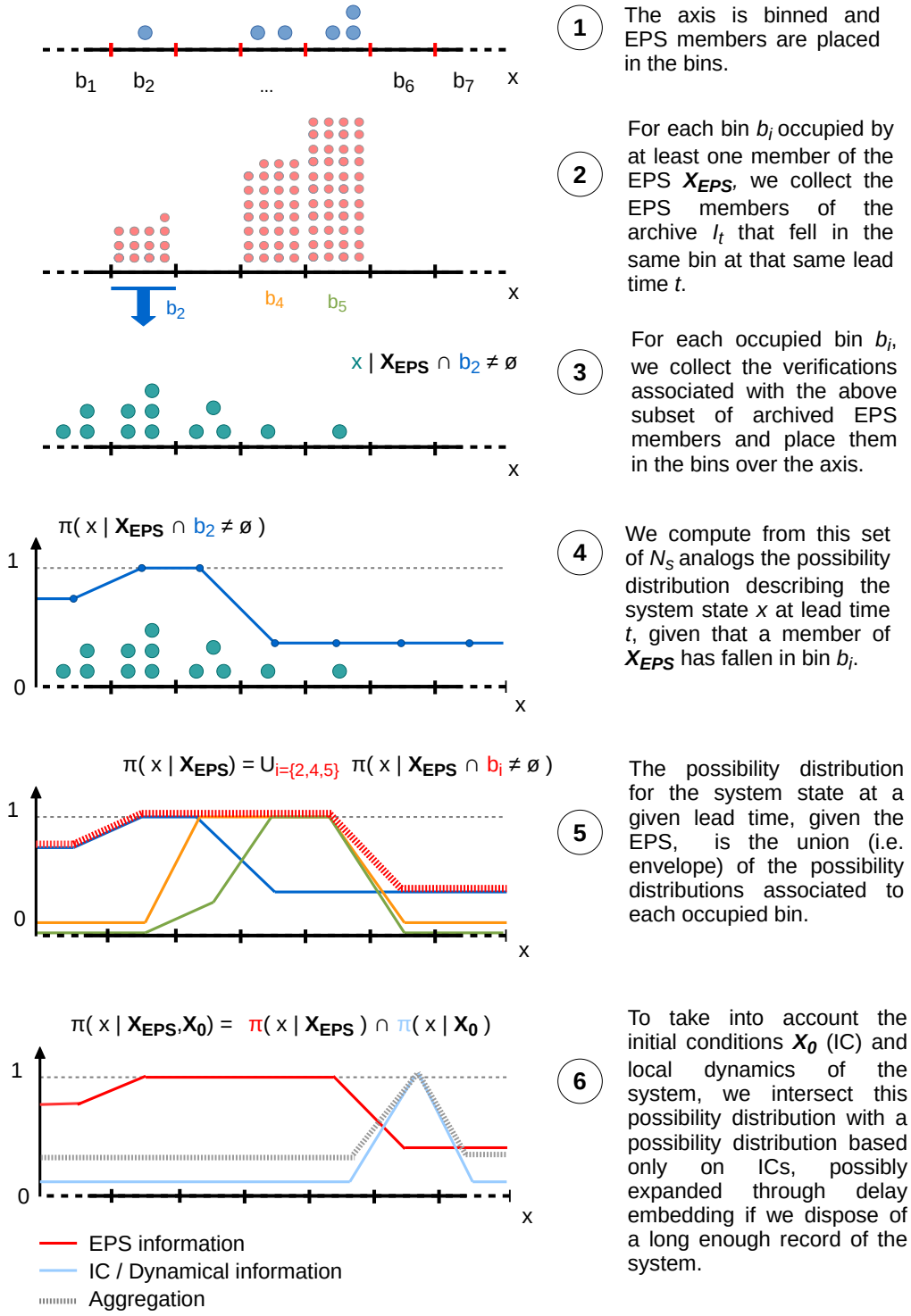
We obtain $|B|$ possibility distributions $\pi(x_t|\tilde{x}_t^m \in b_j)$, each dominating with confidence β the true probability distribution $P(x_t|\tilde{x}_t^m \in b_j)$. Each possibility distribution provides the possibilities for the verification x_t given the presence of one or more ensemble members in bin b_j and is thus a partial view on the state x_t. Since there is only one truth for x_t, we can merge them through a union operator. Fuzzy set theory offers several definitions for computing the distribution resulting from the union of two fuzzy distributions. We adopt here the standard definition for its intuitive rationale: $\pi_{A\cup C}(x) = \max(\pi_A(x), \pi_C(x))$.

We construct the resulting possibility distribution as:

$$\pi_{EPS}(x_t \in b_i|\tilde{\boldsymbol{x}}_t) = \cup_{j|b_j\in B}\ \pi(x_t|\tilde{x}_t^m \in b_j) = \sup_{j|b_j\in B} \pi(x_t \in b_i|\tilde{x}_t^m \in b_j)\ ,\ i = 1, ..., n. \tag{3}$$

3.2. From possibility distribution to prediction

We focus on the continuous interpretation of π_{EPS} and now turn to our approach for producing confidence intervals for the future value x_t, and on the associated formal guarantees.



1 The axis is binned and EPS members are placed in the bins.

2 For each bin b_i occupied by at least one member of the EPS $\boldsymbol{X_{EPS}}$, we collect the EPS members of the archive l_t that fell in the same bin at that same lead time t.

3 For each occupied bin b_i, we collect the verifications associated with the above subset of archived EPS members and place them in the bins over the axis.

4 We compute from this set of N_S analogs the possibility distribution describing the system state x at lead time t, given that a member of $\boldsymbol{X_{EPS}}$ has fallen in bin b_i.

5 The possibility distribution for the system state at a given lead time, given the EPS, is the union (i.e. envelope) of the possibility distributions associated to each occupied bin.

6 To take into account the initial conditions $\boldsymbol{X_0}$ (IC) and local dynamics of the system, we intersect this possibility distribution with a possibility distribution based only on ICs, possibly expanded through delay embedding if we dispose of a long enough record of the system.

Figure 3. Methodology of the possibilistic interpretation of EPSs developed in this paper.

As can be easily derived from Equation (2), a possibility density π is consistent with the associated probability measure P iff its $\alpha-$cuts $C_\pi^\alpha = \{x,\ \pi(x) \geq \alpha\}$ satisfy:

$$P(x \in C_\pi^\alpha) = P(C_\pi^\alpha) \geq 1 - \alpha\ ,\ \forall\, \alpha \in [0, 1]. \tag{4}$$

This constitutes an easily verifiable consistency criterion (Hose and Hanss, 2019).

The possibility distribution satisfying this criterion is not unique. Beyond consistency, the choice of a possibility distribution to model the knowledge at hand is driven by the principle of maximum specificity (Dubois et al., 2004). If π_1 and π_2 are two possibility distributions such that $\pi_1(x) \leq \pi_2(x)\ \forall x \in \mathcal{X}$, then π_1 is said more specific than π_2 and is more informative (i.e. less conservative). Maximum specificity w.r.t. the probabilistic information (*a priori* unknown) is achieved when the possibility distribution is probabilistically calibrated[1]:

$$P(C_\pi^\alpha) = 1 - \alpha\ ,\ \forall \alpha \in [0, 1]. \tag{5}$$

This means that each $\alpha-$cut represents a frequentist confidence interval at level $1 - \alpha$ for the variable of interest and π is a consonant confidence structure (Balch, 2020).

By construction, the individual possibility distributions $\pi(x_t|\tilde{x}_t^m \in b_j)$ verify Equation (1) with a guaranteed confidence level β. π_{EPS} being made of their envelope, it cannot be more specific than any single one of them and consequently the same guarantee applies. In the case of its $\alpha-$cuts, this reads:

$$P\Big(P(x_t \in C_\pi^\alpha) \geq 1 - \alpha\Big) \geq \beta. \tag{6}$$

(Masson and Denœux, 2006) show empirically that their data-to-possibility transformation is rather conservative and provides a possibility distribution that actually dominates the true probability distribution with a rate much higher than the guaranteed β. Even for small sample sizes, the choice of β is not critical and quasi perfect coverage rate is obtained: $\beta \geq 0.8$, ensures that $P\Big(P(x \in C_\pi^\alpha) \geq 1 - \alpha\Big) \to 1$. Under this assumption, the $(1 - \alpha)$-cuts can be used as candidate confidence intervals of *guaranteed* level α. Ideally, we are looking for $(1-\alpha)$-cuts verifying Equation (5), which ensures optimal specificity of π_{EPS} and thus maximally informative confidence intervals.

4. Experiments

4.1. Experimental setting

We reproduce the experiment designed by (Williams et al., 2014), who used an imperfect L96 model to investigate the performances of ensemble postprocessing for the prediction of extreme events. The system dynamics is governed by the following system of coupled equations, where the

[1] Indeed, any conservative statement such as $\exists\, \gamma \mid \pi(x) \geq \gamma,\ \forall x$ implies that $P(C_\pi^\alpha) = 1\ \forall \alpha \leq \gamma$. Equation (5) ensures that a possibility distribution showing such conservative properties is discarded when compared to a possibility distribution that does not show them.

X variables represent slow-moving, large-scale processes, while Y variables represent small-scale, possibly unresolved, physical processes:

$$\frac{dX_j}{dt} = X_{j-1}(X_{j+1} - X_{j-2}) - X_j + F - \frac{hc}{b}\sum_{k=1}^{K} Y_{j,k} \tag{7}$$

$$\frac{dY_{j,k}}{dt} = cbY_{j,k+1}(Y_{j,k-1} - Y_{j,k+2}) - cY_{j,k} + \frac{hc}{b}X_j \tag{8}$$

where $j = 1, \ldots, J$ and $k = 1, \ldots, K$. The parameters are set to: $J = 8$, $K = 32$, $h = 1$, $b = 10$, $c = 10$ and $F = 20$. This perfect model is randomly initialised and then integrated forward in time by means of a Runge-Kutta 4th-order method with time step $dt = 0.002$ (model time units) until enough trajectories of duration 1.4, starting every 1.5 time units, are recorded for our analysis. A lead time $t = 1$ corresponds to 0.2 model time units after initialisation and can be associated with approximately 1 day in the real world (Lorenz, 1996). We are interested in predicting the variable X_1.

An imperfect version of the L96 system is implemented to generate predictions for the X_j variables. In Equation (7), $-\frac{hc}{b}\sum_{k=1}^{K} Y_{j,k}$ is replaced with:

$$0.262 - 1.262X_j + 0.004608X_j^2 + 0.007496X_j^3 - 0.0003226X_j^4 \tag{9}$$

To reproduce the perturbation of the ICs, M perturbed members $\tilde{X}_j$ are sampled independently around the true value of each variable X_j following a normal distribution $\tilde{X}_j \sim \mathcal{N}(X_j, 0.1^2)$. These ensemble sets are initialised each time a new trajectory record starts, and integrated forward in time up to lead time 1.4 by means of a Runge-Kutta 4th-order method with lower time resolution ($\tilde{dt} = 0.02$ model time units). The size of the ensemble is set to $M = 24$, a value comparable to operational weather forecasting schemes (e.g. $M = 17$ for the Met Office Global and Regional Ensemble Prediction System).

4.2. Reference model: Gaussian ensemble dressing

We compare the performances of our approach (POSS hereafter) to those of a classical probabilistic framework for interpreting EPSs, namely a Gaussian ensemble dressing (GEB hereafter). Its predictive probability distribution reads (Roulston and Smith, 2003):

$$p(x_t|\tilde{\boldsymbol{x}}_t)_\theta = \frac{1}{M}\sum_{i=1}^{M} \mathcal{N}(a\tilde{x}_t^i + \omega, \sigma^2) \tag{10}$$

We infer the parameters $\theta = \{a, \omega, \sigma\}$ through the optimisation of the ignorance score (Roulston and Smith, 2002) over the archive $\mathcal{I}_t$ used in the possibilistic framework. To that end, we use the nonlinear programming solver provided by MATLAB® and apply the guidance developed in (Bröcker and Smith, 2008) to provide robust solutions.

Confidence intervals at level α on x_t are obtained from p by a method that provides the desired intervals associated with the highest-density regions (Hyndman, 1996). We also report in the next

section the performances of the confidence intervals similarly extracted from the unprocessed probability density (hereafter RAW) associated with the EPS (a histogram of the EPS normalised to represent a probability density).

4.3. Evaluation criteria

We aim at answering the questions:

(a) Can a possibilistic treatment of the EPS provide more guarantees than a probabilistic interpretation?

(b) If yes, at what cost?

To that end, we compare the performances of the confidence intervals at level α, noted I^α, extracted from the methodologies POSS, GEB and RAW as described in the previous sections. We say that a confidence interval is *guaranteed at level* α if the coverage probability verifies $P(x \in I^\alpha) \geq \alpha$. We use the term guaranteed in the sense that such an interval is associated with a lower bound on the (frequentist) probability that the verification falls within it. Such guarantees are sought e.g. in risk-averse decision-making. We say that it is *reliable*, or *probabilistically calibrated*, when $P(x \in I^\alpha) \approx \alpha$. We call it all the more *conservative* than $P(x \in I^\alpha) - \alpha$ is large, which is associated with non optimal interval *precision*.

4.4. Experiments

All results presented here use $n = 30$ bins of similar width to partition the $x-$axis [2]. The test set consists in 40,000 independent trajectories of length $t = 7$ days and the corresponding EPS predictions. All EPSs have beforehand been preprocessed to remove the constant bias. We consider a range of archive size $N_I \in \{156, 1560, 5 \times 10^3, 15 \times 10^3, 30 \times 10^3\}$. In particular, $N_I = 156$ corresponds to 3 years of model archive, whereas $N_I = 1560$ amounts to 30 years, which corresponds to the standard length of a historical re-forecast dataset (Hamill et al., 2004; Hagedorn, 2008). The two latter N_I are operational figures, unlike larger values that we present to study the asymptotic properties of our framework.

We define two types of events: an extreme event, "$x \leq q_5$" (EE), and a common event, "$q_{50} < x \leq q_{55}$" (NEE) where q_i represents the percentile of level i of the climatic distribution of x (i.e. global distribution), plotted in Figure 4 along with both events. This will allow us to use test sets of similar sizes[3] in order to position our approach against the generic probabilistic postprocessing techniques that are known to weakly address such extreme events.

A preliminary assessment (Figure 5) of the effect of the parameter β of Goodman's model on the probabilistic reliability of the $(1-\alpha)$-cuts derived from π_{EPS} shows that varying β from 0.6 to 1 does not impact guarantees at any given N_I for the events of interest. It only impacts precision

[2] This choice is based on the range covered by the climatology of x and the fact that x can be associated to a physical quantity of the atmosphere, e.g. temperature, which leads to bins of width ≈ 2 degrees. For other systems and applications, the bins can be for instance partitioned so that the distribution of the climatology is homogeneous over the bins.

[3] About 2×10^3 elements.

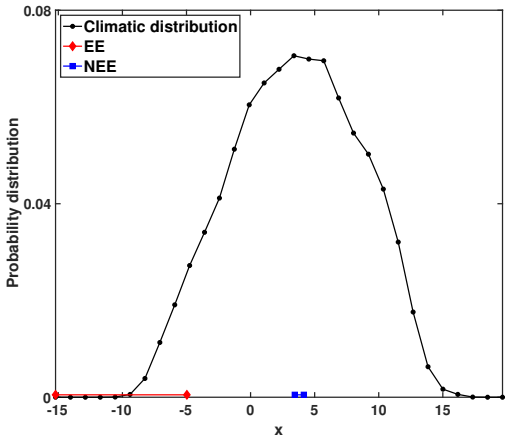


Figure 4. Climatic distribution of the L96's variable of interest X_1 (x for simplicity) where the 'extreme' event "$x \leq q_5$" (EE) and 'common' event "$q_{50} < x \leq q_{55}$" (NEE) are reported.

and its effect is only visible for small archives ($N_I \leq 156$) or large lead times, especially in the EE case. We consequently use $\beta = 0.9$ in our experiments, which allows to improve specificity while maintaining guarantees on confidence intervals.

5. Results

5.1. Empirical assessment of formal guarantees

Figure 6 reports the coverage probability of the confidence intervals I^α extracted for $\alpha \in \{0, 0.05, 0.1, ..., 1\}$ for all evaluated methodologies at lead times $t \in \{1, 3, 5, 7\}$ days. We first note that using RAW leads to confidence intervals that are not guaranteed for $t > 1$ day for both EE and NEE. Postprocessing (here GEB) allows to make them guaranteed at all lead times for the NEE and for $t \leq 3$ days for the EE. The effect of the training set size for the probabilistic treatment does not appear to be significant. Conversely, the confidence intervals derived using POSS are globally guaranteed for both events and at all lead times for operational archives ($N_I < 5 \times 10^3$). Interestingly, when the archive grows significantly, confidence intervals with large α are not guaranteed anymore for the larger lead times in the EE case. The effect appears all the earlier (in terms of lead time) than N_I is large.

We observe here a limitation of possibility theory: its strength lies in incomplete information. As shown in Figure 2, the larger the datasets used to derive possibility distributions, the closer the possibility distribution is in shape to the underlying probability distribution. In particular, the level γ such as $\pi(x) \geq \gamma \ \forall x$ tends towards zero. In other words, such possibility distributions tend to conceal the possibility of rare events.

We illustrate this phenomenon in Figure 7, where we represent the average density of analogs used to compute the individual $\pi(x_t | \tilde{x}_t^m \in b_j)$ (see step 3, Figure 3). In the EE case, as the lead time increases, this average density decreases by several orders of magnitude for the more extreme

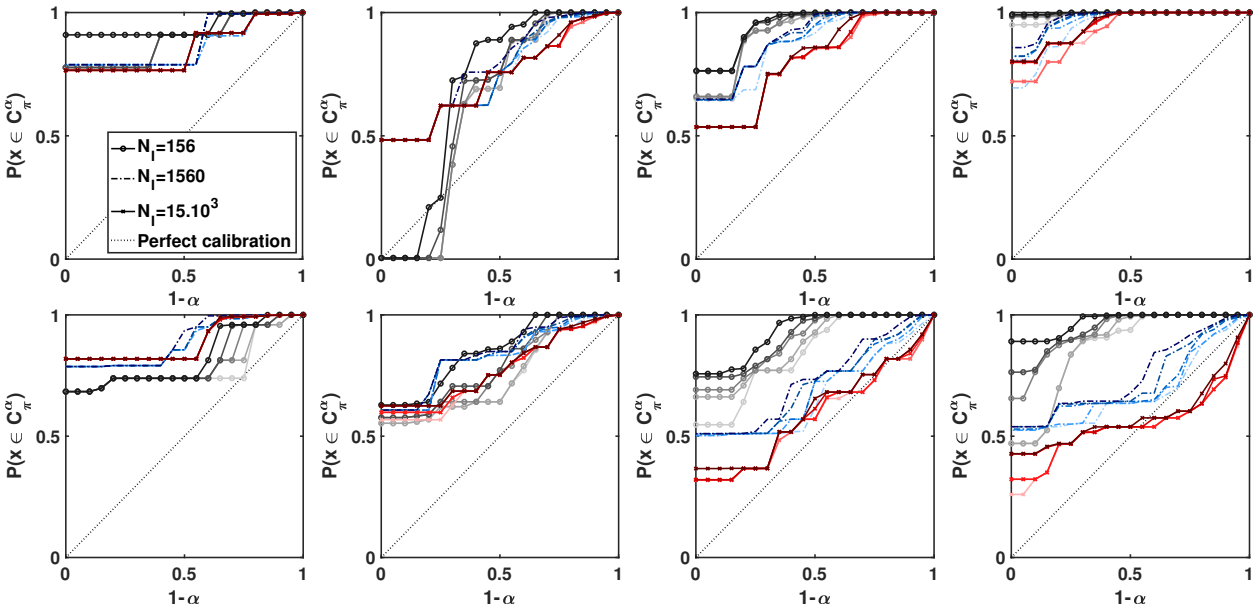


Figure 5. Coverage probability of the $\alpha-$cuts of π_{EPS} at lead time $t \in \{1, 3, 5, 7\}$ days (left to right), in the case of the NEE (top) and EE (bottom). Goodman models with parameter $\beta \in \{0.6, 0.75, 0.9, 0.95, 0.99\}$ (the darker the line, the larger β) are compared in the case of three archives of respective size $N_I \in \{156, 1560, 15 \times 10^3\}$ (grey, blue and red color scales respectively).

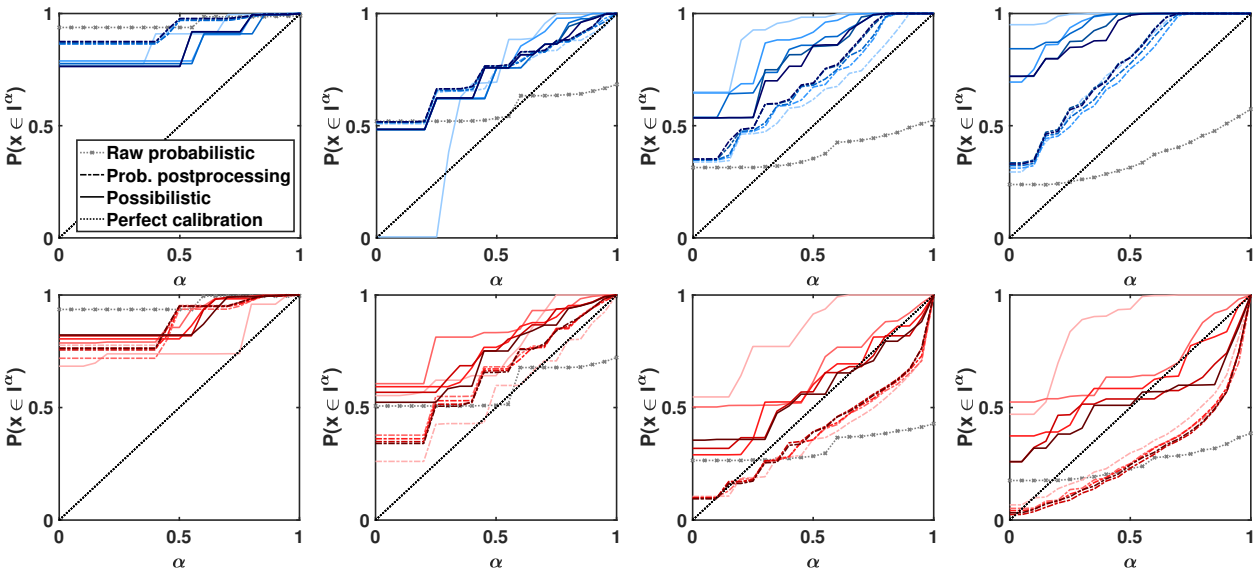


Figure 6. Coverage probability of the $(1 - \alpha)$-cuts of π_{EPS} used as confidence intervals of level α at lead time $t \in \{1, 3, 5, 7\}$ days (left to right), in the case of the NEE (top) and EE (bottom). The EPS archive size is $N_I \in \{156, 1560, 5 \times 10^3, 15 \times 10^3, 30 \times 10^3\}$ (the larger the darker the line). The coverage probability of the confidence intervals of level α derived from the raw EPS's probability density and from the postprocessed density (with the same training set of size N_I as used in the possibilistic framework) is reported as well.

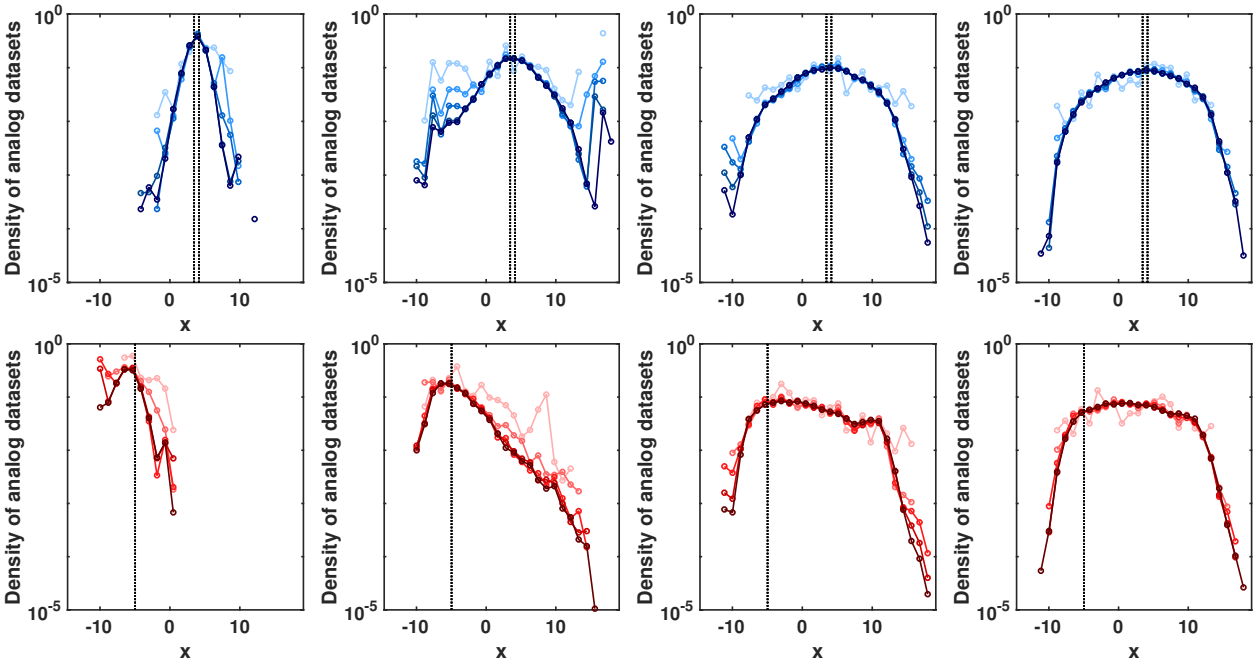


Figure 7. Average density of the analog datasets used to derive π_{EPS}, for sizes $N_I \in \{156, 1560, 5 \times 10^3, 15 \times 10^3, 30 \times 10^3\}$ (the larger N_I the darker the line) and lead time $t \in \{1, 3, 5, 7\}$ days (left to right), in the case of the NEE (top) and the EE (bottom). Only densities above 0 are represented. Vertical dotted lines allow to visualise the events of interest (note that the EE is only defined by its upper bound).

bins ($x \to \inf \mathcal{X}$). This drop is all the more significant than N_I is large. For small $N_I \leq 1560$, the more extreme bins are, as expected, not represented but the intermediary bins are and their density remains above $\frac{1}{100}$. For very large $N_I \geq 5 \times 10^3$, the more extreme bins are represented however their density drops below $\frac{1}{1000}$. In other words, the rarest events part of EE are represented only for extremely large archives, where they will be part of large analog sets, which implies, given the asymptotic behaviour illustrated in Figure 2, that they will be concealed from the associated possibility distributions. More precisely, the level γ such as $\pi_{EPS}(x) \geq \gamma \ \ \forall x \in \mathcal{X}$ remains strictly positive so $P(x \in I^\alpha) = 1$ remains valid for $\alpha \approx 1$ (that is the large scale $(1-\alpha)$-cuts where $1 - \alpha \to 0$). However for intermediate α, the $(1-\alpha)$-cuts may not extend enough towards extreme bins, which negatively impacts the coverage rate. This trend is only observed for sufficiently large α, as possibility distributions remain globally more conservative than the EPS-based probability distributions (see next Section), and consequently provide I^α that encompass more observations than the frequentist calibration requires in the case of smaller α (i.e. for the upper part of the distribution). The "sufficiently large α" decreases with increasing lead times and archive sizes, following the effect described in Figure 7. Figure 8 illustrates our point by breaking down the coverage probability for three subsets of the EE: large archives lead to POSS-based confidence intervals that are all the more guaranteed as the event of interest is not too extreme. Probabilistic calibration for the more extreme part of EE can be improved by increasing the parameter β, however this has no effect in the case of large archives (see Figure 9).

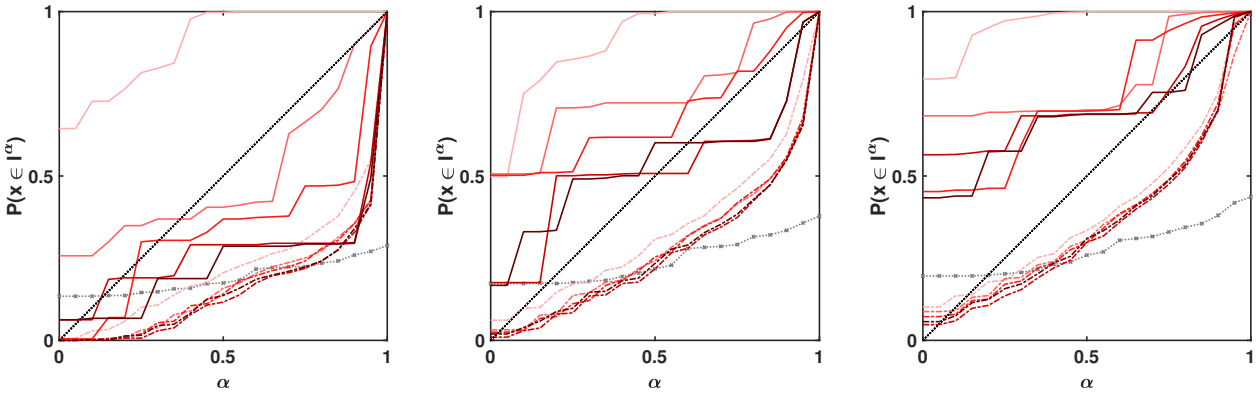


Figure 8. Coverage probability of the $(1-\alpha)$-cuts of π_{EPS} at lead time $t = 7$ days, in the case of events belonging to a partition of subsets of EE (from left to right: $x \leq q_1$, $q_1 < x \leq q_3$ and $q_3 < x \leq q_5$). The EPS archive size varies: $N_I \in \{156, 1560, 5 \times 10^3, 15 \times 10^3, 30 \times 10^3\}$ (the larger the darker the line). The probabilistic calibration of the confidence intervals of level α derived from the raw EPS's probability density and from the postprocessed density (with the same training set of size N_I as the possibilistic framework) is also reported. See Figure 6 for legend.

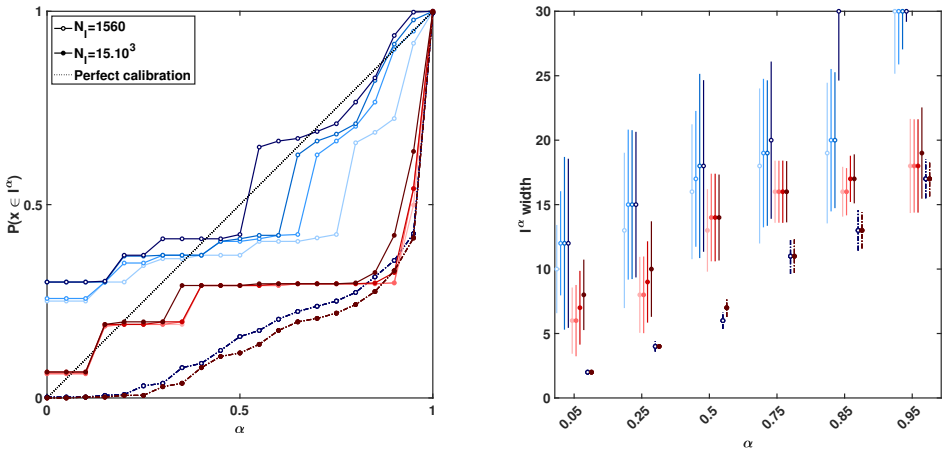


Figure 9. Coverage probability (left) and associated width distributions (right) of the confidence intervals of level α at lead time $t = 7$ days, in the EE case, for two archive sizes $N_I \in \{1560, 15 \times 10^3\}$ (blue and red color scale respectively). POSS results (solid line) for increasing Goodman's parameter $\beta \in \{0.6, 0.9, 0.95, 0.99\}$ (the larger the darker the line) are compared to GEB results (dotted line). The width distribution is represented through its mean and one standard deviation above and below.

The NEE case study does not suffer from this limitation as the density of analogs falling in the NEE bins remains around $\frac{1}{10}$ at all lead times. In comparison to GEB, POSS improves the reliability of confidence intervals for very short lead times while they remain more conservative for large lead times.

Provided that N_I is not too large (which we assume is always the case for operational archives), Figures 6 and 8 clearly show the added value of treating the EPS in a possibilistic manner in terms of guarantees for the EE at large lead times, or in terms of reliability for the NEE at very small lead

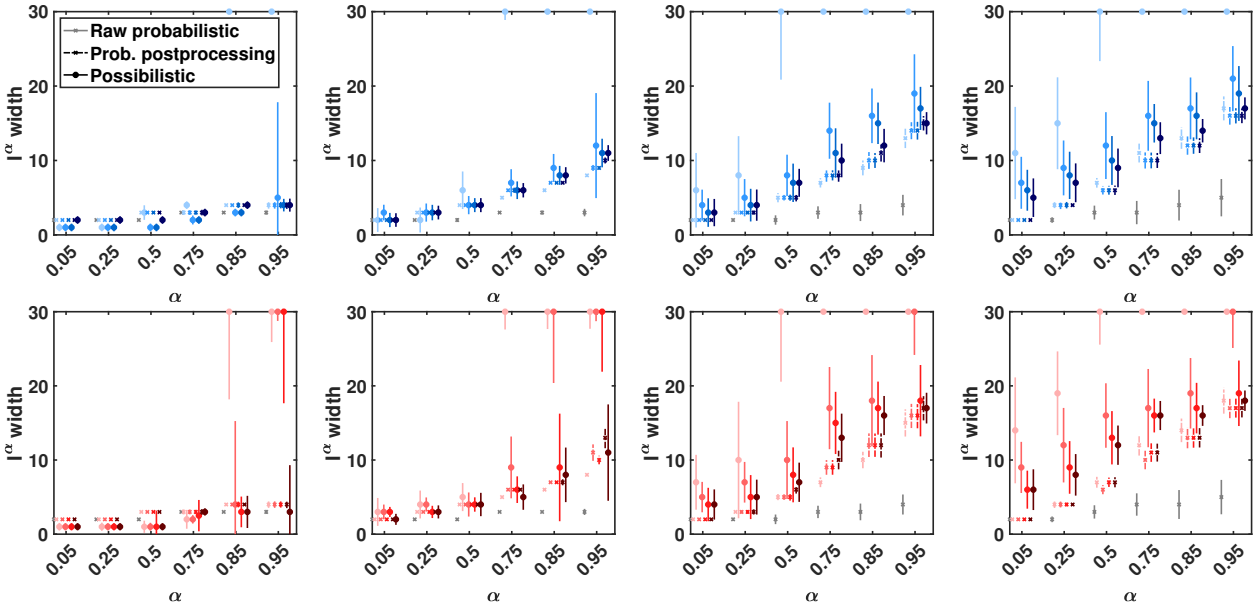


Figure 10. Distribution (mean ± standard deviation) of the width of the possibility and probability-based confidence intervals described in the legend of Figure 6 for lead time $t \in \{1, 3, 5, 7\}$ days (left to right), in the case of the NEE (top) and EE (bottom). Only the cases $N_I \in \{156, 1560, 5 \times 10^3, 30 \times 10^3\}$ are represented (the larger N_I, the darker the line).

times. However, we can wonder what is the cost of such improvements. How do the possibility-based confidence intervals compare to their probability-based counterparts, in terms of precision?

5.2. Interval precision

Figure 10 compares the average width of the confidence intervals derived from the three methodologies. For both EE and NEE, N_I affects the width of the possibilistic I^α significantly, making them narrower with larger N_I, all the more than the lead time increases. Their probabilistic counterparts are generally much smaller, except when $N_I \approx 30 \times 10^3$.

For NEE and level $\alpha < 0.9$, POSS brings more information at very short lead times ($t = 1$ day) than the probabilistic approaches: intervals are smaller or equal in size and remain guaranteed. This is all the more true that the archive is of intermediate size ($N_I = 1560$). Increasing the lead time beyond $t = 3$ days favors the probabilistic approach, which is more reliable with narrower intervals.

For EE, the added value of POSS over GEB is observed on two occasions: 1) intervals are as reliable yet narrower for very small lead times and $\alpha < 0.9$, whatever the archive size ; 2) for large lead times and intermediary-sized archives ($N_I \in \{1560, 5 \times 10^3\}$), possibility-based confidence intervals are both guaranteed, reliable and operational (i.e. not too wide compared to GEB's results, contrary to what $N_I = 156$ produces), while the probabilistic intervals are narrower yet not guaranteed at all. In the case of particularly rare events, as represented in Figure 8, an intermediary archive such as $N_I = 1560$ is able to produce confidence intervals close to perfect reliability even

for large lead times, as long as the parameter β is increased towards 1. Such reliability is reached at the expense of the interval width, which is significantly increased (w.r.t. smaller β) for the largest $\alpha \geq 0.85$.

6. Conclusion

We introduced a novel framework to interpret EPSs where a possibility distribution π_{EPS} is derived from the EPS at hand and an archive of (EPS; verification). We showed how to use the $(1-\alpha)$-cuts of a continuous interpretation of π_{EPS} to produce confidence intervals at level α about the future value of the variable of interest. Our possibility-based confidence intervals come with formal guarantees, and experimental results show that they overpass probability-based ones in two situations: 1) at very small lead times for both common and extreme events, where they are as reliable yet narrower; 2) more blatantly, at intermediate and large lead times for extreme events, where they remain guaranteed and can be brought close to perfect reliability even for particularly rare events, yet at the expense of precision. These results can be reached with operational archive like the $20-30$-year reforecast datasets. The guarantees are retained for smaller archives, which however lead to more conservative intervals and thereby impede operationality.

Possibility theory is a promising tool for the prediction of extreme events, given a limited and imperfect amount of information on the system's dynamics. Beyond the results presented in this article, further developments by the author (Le Carrer and Ferson, 2020) show how π_{EPS} can be combined with additional possibility distributions constructed from alternative sources of information such as the IC or dynamical information (see step 6 of Figure 3). Therein, the concept of ignorance briefly introduced in Section 2 is developed and presented as a promising tool for risk communication.

Acknowledgements

This research is funded by the Engineering & Physical Sciences Research Council (EPSRC) and the Economic & Social Research Council (ESRC) with grant no. EP/L015927/1. We are very grateful to N. Berthier for his careful reviewing.

References

Balch, M. S.: 2020, 'New Two-Sided Confidence Intervals for Binomial Inference Derived Using Walley's Imprecise Posterior Likelihood as a Test Statistic'. *International Journal of Approximate Reasoning*.

Bröcker, J. and L. A. Smith: 2008, 'From ensemble forecasts to predictive distribution functions'. *Tellus A: Dynamic Meteorology and Oceanography* **60**(4), 663–678.

Buizza, R.: 2018, 'Ensemble Forecasting and the Need for Calibration'. In: *Statistical Postprocessing of Ensemble Forecasts*. Elsevier, pp. 15–48.

Cayrac, D., D. Dubois, M. Haziza, and H. Prade: 1994, 'Possibility theory in "Fault mode effect analyses". A satellite fault diagnosis application'. In: *Proceedings of 1994 IEEE 3rd International Fuzzy Systems Conference*. pp. 1176–1181.

Dempster, A. P.: 2008, 'Upper and Lower Probabilities Induced by a Multivalued Mapping'. In: R. R. Yager and L. Liu (eds.): *Classic Works of the Dempster-Shafer Theory of Belief Functions*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 57–72.

Dubois, D., L. Foulloy, G. Mauris, and H. Prade: 2004, 'Probability-Possibility Transformations, Triangular Fuzzy Sets, and Probabilistic Inequalities'. *Reliable computing* **10**(4), 273–297.

Dubois, D. and H. Prade: 2012, *Possibility theory: an approach to computerized processing of uncertainty*. Springer Science & Business Media.

Friederichs, P. and A. Hense: 2007, 'Statistical Downscaling of Extreme Precipitation Events Using Censored Quantile Regression'. *Monthly Weather Review* **135**(6), 2365–2378.

Friederichs, P., S. Wahl, and S. Buschow: 2018, 'Postprocessing for Extreme Events'. In: S. Vannitsem, D. S. Wilks, and J. W. Messner (eds.): *Statistical Postprocessing of Ensemble Forecasts*. Elsevier, pp. 127–154.

Gneiting, T. and M. Katzfuss: 2014, 'Probabilistic forecasting'. *Annual Review of Statistics and Its Application* **1**, 125–151.

Goodman, L. A.: 1965, 'On simultaneous confidence intervals for multinomial proportions'. *Technometrics* **7**(2), 247–254.

Hagedorn, R.: 2008, 'Using the ECMWF reforecast dataset to calibrate EPS forecasts'. *ECMWF Newsletter* **117**, 8–13.

Hamill, T. M., J. S. Whitaker, and X. Wei: 2004, 'Ensemble reforecasting: Improving medium-range forecast skill using retrospective forecasts'. *Monthly Weather Review* **132**(6), 1434–1447.

Hose, D. and M. Hanss: 2019, 'Possibilistic calculus as a conservative counterpart to probabilistic calculus'. *Mechanical Systems and Signal Processing* **133**, 106290.

Hyndman, R. J.: 1996, 'Computing and graphing highest density regions'. *The American Statistician* **50**(2), 120–126.

Le Carrer, N. and S. Ferson: 2020, 'Beyond probabilities: A possibilistic framework to interpret ensemble predictions and fuse imperfect sources of information'. Under review.

Legg, T. P. and K. R. Mylne: 2004, 'Early Warnings of Severe Weather from Ensemble Forecast Information'. *Weather and Forecasting* **19**(5), 891–906.

Lorenz, E. N.: 1996, 'Predictability: A problem partly solved'. In: *Proc. Seminar on predictability*, Vol. 1.

Masson, M.-H. and T. Denœux: 2006, 'Inferring a possibility distribution from empirical data'. *Fuzzy Sets and Systems* **157**(3), 319–340.

Mylne, K., C. Woolcock, J. Denholm-Price, and R. Darvell: 2002, 'Operational calibrated probability forecasts from the ECMWF ensemble prediction system: implementation and verification'. In: *Preprints of the Symposium on Observations, Data Asimmilation and Probabilistic Prediction*. pp. 113–118.

Roulston, M. S. and L. A. Smith: 2002, 'Evaluating Probabilistic Forecasts Using Information Theory'. *Monthly Weather Review* **130**(6), 1653–1660.

Roulston, M. S. and L. A. Smith: 2003, 'Combining dynamical and statistical ensembles'. *Tellus A: Dynamic Meteorology and Oceanography* **55**(1), 16–30.

Smith, L. A.: 2016, 'Integrating Information, Misinformation and Desire: Improved Weather-Risk Management for the Energy Sector'. In: P. J. Aston, A. J. Mulholland, and K. M. Tant (eds.): *UK Success Stories in Industrial Mathematics*. Cham: Springer International Publishing, pp. 289–296.

Williams, R. M., C. A. T. Ferro, and F. Kwasniok: 2014, 'A comparison of ensemble post-processing methods for extreme events'. *Quarterly Journal of the Royal Meteorological Society* **140**(680), 1112–1120.

Zadeh, L.: 1978, 'Fuzzy sets as a basis for a theory of possibility'. *Fuzzy Sets and Systems* **1**(1), 3–28.

Why Ellipsoids in Mechanical Analysis of Wood Structures

F. Niklas Schietzold[1], Julio Urenda[2,3], Vladik Kreinovich[3], Wolfgang Graf[1], and Michael Kaliske[1]

[1] *Institute for Structural Analysis, Technische Universität Dresden*
01062 Dresden, Germany, {niklas.schietzold, wolfgang.graf, michael.kaliske}@tu-dresden.de

[2,3] *Departments of* [2] *Mathematical Sciences and* [3] *Computer Science*
University of Texas at El Paso, El Paso, TX 79968, USA, {jcurenda, vladik}@utep.edu

Abstract. Wood is a very mechanically anisotropic material. At each point on the wooden beam, both average values and fluctuations of the local mechanical properties corresponding to a certain direction depend, e.g., on whether this direction is longitudinal, radial or tangential with respect to the grain orientation of the original tree. This anisotropy can be described in geometric terms, if we select a point x and form *iso-correlation* surfaces – i.e., surfaces formed by points y with the same level of correlation $\rho(x, y)$ between local changes in the vicinities of the points x and y. Empirical analysis shows that for each point x, the corresponding surfaces are well approximated by concentric homothetic ellipsoids. In this paper, we provide a theoretical explanation for this empirical fact.

Keywords: wood mechanical properties, ellipsoids, anisotropy

1. Formulation of the Problem: Need for a Theoretical Explanation of an Empirical Fact

Many constructions are made of wood. Wood is one of the oldest materials used in construction. During the past millennia, people have developed a lot of skills for working with wood. However, in spite of this experience, wood remains one of the most difficult materials to handle. The main reason for this difficulty is that, in contrast to many other construction materials which are mostly homogeneous and isotropic, wood is highly inhomogeneous and anisotropic. At each point in the wooden beam, both the average values and fluctuations of the local mechanical properties corresponding to a certain direction depend, e.g., on whether this direction is longitudinal, radial or tangential with respect to the grain orientation of the original tree. In designing wooden constructions, it is important to properly describe and to properly take into account this inhomogeneity and anisotropy; see, e.g., (Vorreiter, 1949; Kollman, 1982; Reuter, 2009; Ulrich and Seim, 2014; Jenkel et al., 2015; Leichsenring et al., 2018; Schietzold, Graf, and Kaliske, 2018; Fleischhauer et al., 2019).

How can we describe local fluctuations of mechanical characteristics? These fluctuations are caused by many different relatively small factors. It is known that the distribution of the joint effect of a large number of relatively small factors is close to Gaussian – this follows from the Central Limit Theorem, according to which this distribution tends to Gaussian when the number of factors increases; see, e.g., (Sheskin, 2011). To describe a Gaussian distribution, it is sufficient

© 2021 *by F. N. Schietzold et al.*

to describe its first and second moments. For a general random field $f(x)$, this means that we need to describe its mean values $E[f(x)]$ (where $E[\cdot]$ denotes the expected value) and its covariances $E[f(x)\cdot f(y)]$. For fluctuations, the mean is 0, so we only need to describe covariances. In statistics, it is often convenient, instead of explicitly describing covariances, to describe the standard deviations $\sigma[f(x)] \stackrel{\text{def}}{=} \sqrt{E[(f(x)^2]}$ and correlations $\rho(x,y) \stackrel{\text{def}}{=} \dfrac{E[f(x)\cdot f(y)]}{\sigma[f(x)]\cdot\sigma[f(y)]}$. Then, covariances can be reconstructed as $E[f(x)\cdot f(y)] = \sigma[f(x)]\cdot\sigma[f(y)]\cdot\rho(x,y)$.

An interesting property of the corresponding correlation functions was recently empirically found; see, e.g., (Schietzold, Graf, and Kaliske, 2019) and references therin. This property is about *iso-correlation* surfaces corresponding to each spatial location x, i.e., surfaces formed by all the points y for which the correlation $\rho(x,y)$ is equal to a constant ρ_0. Empirical analysis shows that for each point x, the corresponding surfaces are well approximated by concentric homothetic ellipsoids; see, e.g., (Schietzold, Graf, and Kaliske, 2019). This property helps narrow down possible functions $\rho(x,y)$ when we analyze mechanical properties of wood – and thus, has a potential to make mechanical analysis of wooden structures more efficient.

The problem is that so far, this property was purely empirical, it had no theoretical justification, and thus, engineers were reluctant to use it: it is known that sometimes, empirical properties found under some conditions do not work well when conditions change. To make this property more reliable and thus, more practically useful, it is therefore desirable to come up with a theoretical explanation.

In this paper, we provide a desired theoretical explanation for this empirical fact.

2. Our Explanation: Main Idea

The main ideas behind our explanation are similar to the ideas used in (Finkelstein, Kosheleva, and Kreinovich, 1996; Li, Ogura, and Kreinovich, 2002) to explain efficiency of ellipsoid approximation in numerical analysis (see, e.g., (Schweppe, 1968; Schweppe, 1973; Fogel and Huang, 1982; Belforte and Bona, 1985; Norton, 1985; Chernousko, 1988; Soltanov, 1990; Utyubaev, 1990; Filippov, 1992; Chernousko, 1994)); the main difference is that now we consider:

- *not* classes of sets (such as the class of all ellipsoids), but
- classes of *families* of sets (e.g., the class of all families of concentric homothetic ellipsoids).

Specifically, we show that for the smallest dimension d for which it is possible to have an affine-invariant optimality criterion on the space of all such d-dimensional classes, for any such criterion, the optimal family consists of concentric homothetic ellipsoids. Thus, such families of ellipsoids provide the optimal approximation to the actual surfaces – at least in the *first* approximation, i.e., approximation corresponding to the smallest possible number of parameters.

3. Our Explanation: Details

Family of sets: towards a precise definition. For each spatial point x, we would like to describe, for each possible value ρ_0 of the correlation $\rho(x, y)$, the set $S_{\rho_0}(x)$ of all the points y for which the correlation $\rho(x, y)$ between the values at x and y is greater than or equal to ρ_0.

What are the natural properties of these families of sets?

First property: coverage. For each y, there is some value of $\rho(x, y)$, so for this x, the union of all these sets $S_{\rho_0}(x)$ coincides with the whole space.

Second property: monotonicity. Of course, if $\rho(x, y) \geq \rho_0$ and $\rho_0 \geq \rho_0'$, then $\rho(x, y) \geq \rho_0'$. So, the sets $S_{\rho_0}(x)$ should be inclusion-monotonic: if $\rho_0 \leq \rho_0'$, then $S_{\rho_0'}(x) \subseteq S_{\rho_0}(x)$.

Third property: boundedness. From the physical viewpoint, the further away is the point y from the point x, the less the physical quantities corresponding to these points are correlated. As the distance increases, this correlation should tend to 0. Thus, each set $S_{\rho_0}(x)$ is *bounded*.

Fourth property: continuity. In physics, most processes are continuous – with the exception of processes like fracturing, which we do not consider here. We can therefore conclude that the correlation $\rho(x, y)$ continuously depends on y. So, if we have $\rho(x, y_n) \geq \rho_0$ for some sequence of points y_n that converges to a point y ($y_n \to y$), then we should have $\rho(x, y) = \lim\limits_{n\to\infty} \rho(x, y_n) \geq \rho_0$. In other words, if $y_n \in S_{\rho_0}(x)$ and $y_n \to y$, then $y \in S_{\rho_0}(x)$, i.e., each set $S_{\rho_0}(x)$ is *closed*.

Similarly, it is reasonable to conclude that the set $S_{\rho_0}(x)$ should continually depend on ρ_0, i.e., that of the two values ρ_0 and ρ_0' are close, then the corresponding sets $S_{\rho_0}(x)$ and $S_{\rho_0'}(x)$ should also be close. A natural way to describe closeness between (bounded closed) sets is to use the so-called Hausdorff distance. In precise terms, for any $\varepsilon > 0$, we say that the sets A and B are *ε-close* if:

- every point $a \in A$ is ε-close to some point $b \in B$ (in the sense that the distance $d(a, b)$ does not exceed ε: $d(a, b) \leq \varepsilon$), and
- every point $b \in B$ is ε-close to some point $a \in A$.

The Hausdorff distance $d_H(A, B)$ between the sets A and B is then defined as the smallest ε for which the sets A and B are ε-closed. It can be shown that this distance can be equivalently defined as follows:

$$d_H(A, B) = \max\left(\sup_{a\in A} d(a, B), \sup_{b\in B} d(b, A)\right),$$

where $d(a, B) \stackrel{\text{def}}{=} \inf\limits_{b\in B} d(a, b)$.

Fifth property: what is the set of possible values of the parameter? In this family of sets, correlation value is a parameter. What are the possible values of correlation? In general, correlations can take any value from -1 to 1. When $y = x$, the correlation is clearly equal to 1. When $y \to \infty$, we get values close to 0. Since the function $\rho(x, y)$ is continuous, this function takes all intermediate values. So, the possible values of the correlation form some interval. In some cases, we may have all possible negative values, in other cases, only some negative values, in yet other cases, we only

have non-negative values. So, in general, we will consider all possible intervals of possible value of ρ_0. This interval may be closed – if there are points with - correlation, or is can be open.

Resulting definition. So, we arrive at the following definition.

Definition 1. *Let $N \geq 2$ be an integer. By a* family of sets*, we mean a set $\{S_c : c \in I\}$ of bounded closed sets $S_c \subseteq \mathbb{IR}^N$ obtained by applying, to each real number c from a non-degenerate interval I (open or closed, finite or infinite), a mapping $c \to S_c$ that has the following properties:*

- *the dependence of S_c on c is continuous: if $c_n \to c$, then $d_H(S_{c_n}, S_c) \to 0$,*
- *the family S_c is monotonic: if $c < c'$, then $S_{c'} \subseteq S_c$, and*
- *the union of all the sets S_c coincides with the whole space.*

Comments.

- According to this definition, the family remains the same if we simply re-parameterize the family: e.g., if instead of the original parameter c, we use a new parameter $c' = c + c_0$ or $c' = \lambda \cdot c$ for some constants c_0 and λ.
- In our specific problem, we are interested in the 3-D case $N = 3$. However, since we can envision similar problem in the plane $N = 2$ or in higher-dimensional spaces – and since the proof of our main result does not depend on any specific value N – in this paper, we consider the general case $N \geq 2$.
- We are specifically interested in *concentric homothetic families of ellipsoids*, i.e., in families of the type $S_c = c \cdot E + a$, where a is a given vector, and E is an ellipsoid with a center at 0.

Class of families of sets. For different situation, in general, we get different correlations and thus, different families of sets. We would like to find general class of such families that would, ideally, cover all such situations. We can use different parameters to differentiate different families from this class. In other words, a class can be described as a method for assigning, to each possible combination of values of these parameters, a specific family. As before, it makes sense to require that the resulting mapping is continuous.

Definition 2. *Let $N \geq 2$ and $r > 0$ be integers. By an r-*parametric class of families of sets*, we mean a mapping that assigns, to each element $p = (p_1, \ldots, p_r)$ from an open r-dimensional set $D \subseteq \mathbb{IR}^r$, a family $\{S_c(p)\}$ so that the dependence of $S_c(p)$ on c and p is continuous.*

Optimality criteria: general idea. Out of all possible classes, we want to select a class which is, in some reasonable sense, optimal. For this, we need to be able to describe when some classes are better than others. In other words, we need to have an *order* on the set of all the classes. It would be nice to have a *total* (*linear*) order, in the sense that for every two classes, we should be able to tell which one is better, but it may be sufficient to have a *partial* order – as long as this order enables us to select the best class, it is OK if some not-best classes, we do not have an opinion of which of them is better.

In practice, usually, optimality criteria are described in numerical form: we have an objective function $f(a)$ that assigns a numerical value to each possible alternative a, and we want to select an alternative for which this value is the largest possible (or, depending on the context, the smallest possible). For example, a company wants to maximize its profit, a city wants to upgrade its road system so as to minimize the average travel time, etc.

However, often, we need to go somewhat beyond this approach. Indeed, for example, a company may have two (or more) different projects that lead to the same expected profit. In this case, we can use this non-uniqueness to optimize something else – e.g., out of all most-profitable projects, select the one that leads to the smallest possible long-term environmental impact. In this case, we have a more complex criterion for comparing alternatives: instead of saying that an alternative a is better than the alternative a' if $f(a) > f(a')$, we say that a is better if:

- either $f(a) > f(a')$
- or $f(a) = f(a')$ and $g(a) > g(a')$, for some other numerical criterion $g(a)$.

If this still does not select us a unique alternative, we can optimize yet something else, etc. In view of this possibility, in this paper, we do not restrict ourselves to numerical optimization criteria and use the most general definition of the optimality criterion, when:

- for some pairs of alternatives a and a', we know that a is better (we will denote it by $a' < a$),
- for some pairs of alternatives a and a', we know that a' is better ($a < a'$), and
- for some pairs of alternatives a and a', a and a' are of the same value (we will denote it by $a \sim a'$).

Clearly, if a' is better than a, and a'' is better than a', then a'' should be better than a, etc. Thus, we arrive at the following definition:

Definition 3. *Let A be a set; elements of this set will be called* alternatives*. By an* optimality criterion*, we mean a pair of binary relations $(<, \sim)$ on the set A for which the following properties hold:*

- *if $a < a'$ and $a' < a''$, then $a < a''$;*
- *if $a < a'$ and $a' \sim a''$, then $a < a''$;*
- *if $a \sim a'$ and $a' < a''$, then $a < a''$;*
- *if $a \sim a'$ and $a' \sim a''$, then $a \sim a''$;*
- *if $a \sim a'$, then $a' \sim a$;*
- *if $a < a'$, then we cannot have $a' < a$ or $a \sim a'$.*

Comment. Such a pair of relations is sometimes called a *partial pre-order*.

Definition 4. *Let $(<, \sim)$ be an optimality criterion on a set A. An alternative a_{opt} is called* optimal *with repect to this criterion if for every alternative $a \in A$, we have $a < a_{\text{opt}}$ or $a \sim a_{\text{opt}}$.*

We need a final optimality criterion. If an optimality criterion does not select any alternative as optimal, this means that this criterion still needs work – this may happen if for most pairs of alternatives, this criterion does not tell us which alternative is better. So, for the optimality criterion to be useful, it must select *at least one* optimal alternative.

If the criterion selects *several* alternatives as optimal, this means – as we have mentioned earlier – that this criterion is not final: we can use the resulting non-uniqueness to optimize something else, i.e., in effect, to come up with a better optimality criterion. If for this better criterion, we still have several optimal alternatives, we can (and should) modify this criterion even further, etc., until we finally get a criterion for which there is exactly one optimal alternative. Thus, we arrive at the following natural definition.

Definition 5. *We say that an optimality criterion is* final *if there exists exactly one alternative which is optimal with respect to this criterion.*

For our problem, an optimality criterion must be affine-invariant. In our case, we want to compare different classes (of families of sets). In selecting optimality criteria, it is reasonable to take into account that while we want to deal with sets of points in physical space, from the mathematical viewpoint, we deal with sets of tuples of real numbers. Real numbers (coordinates) describing each point depend on what coordinate system we use: if we select a different starting point, then all the coordinates are shifted $x_i \to x_i + a_i$; if we select different axes for the coordinates, we get a rotation $x_i \to \sum\limits_{j=1}^{N} r_{ij} \cdot x_j$ for an appropriate matrix r_{ij}, etc.

These transformations make sense for the *isotropic* case, when all the properties of a material are the same in all directions. Wood is an example of an *anisotropic* material: e.g., it is easier to cut it along the orientation of the original tree than across that orientation. It is known that in many cases, the description of an anisotropic material can be reduced to the isotropic case if we apply an appropriate affine transformation. This usually comes from the fact that, e.g., mechanical properties of a body can be described by a symmetric matrix, and each symmetric matrix can become a unit matrix if we use its eigenvectors as the base for the new coordinate system.

In view of this, it is reasonable to require that our optimality criterion is invariant not only with respect to shifts and rotations, but also with respect to all possible affine (linear) transformations. Thus, we arrive at the following definitions.

Definition 6. *Let $N > 2$ be an integer. By an* affine transformation*, we mean a transformation $T : \mathbb{IR}^N \to \mathbb{IR}^N$ of the type $(Tx)_i = a_i + \sum\limits_{j=1}^{N} b_{ij} \cdot x_j$ for some reversible matrix b_{ij}. Let T be an affine transformation.*

- *Let $S \subseteq \mathbb{IR}^N$ be a set. By the* result $T(S)$ *of applying T to S, we mean the set $\{T(s) : s \in S\}$.*
- *Let $F = \{S_c : c \in I\}$ be a family of sets. By the* result $T(F)$ *of applying T to F, we mean the family $\{T(S_c) : c \in I\}$.*

– *Let $C = \{S_c(p)\}$ be class of families. By the* result *$T(C)$ of applying T to C, we mean the class $\{T(S_c(p))\}$.*

Definition 7. *Let A be a set of alternatives, let $(<, \sim)$ be an optimality criterion of the set A, and let $\mathcal{T}$ be a class of transformations $A \to A$. We say that the optimality criterion $(<, \sim)$ is* $\mathcal{T}$-invariant *if for every $T \in \mathcal{T}$ and for all $a, a' \in A$, the following two properties hold:*

– *if $a < a'$ then $T(a) < T(a')$, and*

– *If $a \sim a'$, then $T(a) \sim T(a)$.*

Proposition 1. *Let $N > 0$ and $r > 0$ an integers, and let $(<, \sim)$ be a final affine-invariant optimality criterion on the set of all r-parametric classes of families of sets in $\mathbb{R}^N$. Then:*

– $r \geq \dfrac{N \cdot (N+3)}{2} - 1$*; and*

– *for $r = \dfrac{N \cdot (N+3)}{2} - 1$, the optimal class consists of concentric homothetic families of ellipsoids.*

Comment. This result indeed shows that the class of concentric homothetic families of ellipsoids is the simplest of all possible optimal classes – simplest in the sense that it requires the smallest number of parameters to describe.

4. Proof

1°. Since the optimality criterion is final, there exists exactly one class C_{opt} which is optimal with respect to this criterion, i.e., for which $C < C_{\text{opt}}$ or $C \sim C_{\text{opt}}$ for all other classes C. Let us first prove that the optimal class C_{opt} is itself affine-invariant, i.e., that $T(C_{\text{opt}}) = C_{\text{opt}}$ for each affine transformation T.

Indeed, due to optimality, for each class C and for each affine transformation class T, for the class $T^{-1}(C)$, i.e., we have either $T^{-1}(C) < C_{\text{opt}}$ or $T^{-1}(C) \sim C_{\text{opt}}$.

Since the criterion is affine-invariant, we have either $T(T^{-1}(C)) < T(C_{\text{opt}})$ or $T(T^{-1}(C)) \sim T(C_{\text{opt}})$. Here, by the definition of the inverse transformation, $T(T^{-1}(C)) = C$, so we conclude that for every class C, we have either $C < T(C_{\text{opt}})$ or $C \sim T(C_{\text{opt}})$. By definition of optimality, this means that the class $T(C_{\text{opt}})$ is optimal. However, our optimality criterion is final, which means that there is only one optimal class. Thus, indeed, $T(C_{\text{opt}}) = C_{\text{opt}}$.

Since the optimal class is affine-invariant, with each family F this class also contains the family $T(F)$. This means, in its turn, that for each set S_c from each family, some family from the optimal class also contains the set $T(S_c)$.

2°. Let us show that $r \geq \dfrac{N \cdot (N+3)}{2} - 1$. Indeed, it is known (see, e.g., (Busemann, 1955)) that for every non-degenerate bounded set S (i.e., for every bounded set which is not contained in a

proper subspace), among all ellipsoids that contain S, there exists a unique ellipsoid of the smallest volume. It is also known that this correspondence between a set and the corresponding ellipsoid is affine-invariant: if an ellipsoid E corresponds to the set S_c, then, for each affine transformation T, to the set $T(S_c)$ there corresponds the ellipsoid $T(E)$.

It is known that every two ellipsoids can be obtained from each other by an appropriate affine transformation. Thus, the family of all ellipsoids corresponding to all the sets from all the families consists of all the ellipsoids. How many ellipsoids are there? A general ellipsoid can be determine by a quadratic formula $\sum\limits_{ij} a_{ij} \cdot x_i \cdot x_j + \sum\limits_{i=1}^{N} a_i \cdot x_i \le 1$ for some symmetric matrix a_{ij} and a vector a_i – and it is easy to see that different combinations of the matrix and the vector lead to different ellipsoids. We need N values $a_1, \ldots, a_N$ to describe a vector. Out of N^2 elements of the matrix, we need N values to describe its diagonal values a_{ii} and we need $\dfrac{N^2-N}{2}$ to describe non-diagonal elements: we divide by two since the matrix is symmetric $a_{ij} = a_{ji}$. Thus, overall, we need

$$N + N + \frac{N^2-N}{2} = \frac{N \cdot (N+3)}{2}$$

values.

Thus, the set of all ellipsoids is $\dfrac{N \cdot (N+3)}{2}$-dimensional. Since to each set S_c from families from the optimal class, we assign an ellipsoid, the dimension of the set of such sets should also be at least $\dfrac{N \cdot (N+3)}{2}$-dimensional. These sets are divided into 1-parametric families, so the dimension r of the class of such families cannot be smaller than the above dimension minus 1. Thus, indeed,

$$r \ge \frac{N \cdot (N+3)}{2} - 1.$$

3°. Let us now prove that for the smallest possible dimension $r = r_{\min} \stackrel{\text{def}}{=} \dfrac{N \cdot (N+3)}{2} - 1$, all the sets S_c from the each family of the optimal class are ellipsoids.

In Part 2 of this proof, we showed that each ellipsoid is associated with some set S_c from one of these families. The unit ball with a center at 0 is clearly an ellipsoid. Let us consider the set S_c which is associated with this unit ball. A unit ball is invariant with respect to all the rotations around its center. If the associated set S_c is not equal to the unit ball, this means that this set is not invariant with respect to at least some rotations. In other words, the group of all rotations that leave this set invariant is a proper subgroup of the group of all rotations. This implies that the dimension of this group is smaller than the dimension of the group of all rotations – and thus, that there exists at least 1-parametric family $\mathcal{R}$ of rotations R with respect to which the set S_c is not invariant.

Since the optimal class is affine-invariant, all the sets $R(S_c)$ are also sets from some family from the optimal class – and for all of them, the same unit ball is the smallest-volume ellipsoid. Thus, for this particular ellipsoid – the unit ball, we have at least a 1-dimensional family of sets S_c associated with this same ellipsoid. By applying a generic affine transformation, we can find a

similar at-least-1-dimensional family of sets corresponding to each ellipsoid. Thus, the dimension of the set of all sets S_c is at least one larger than the dimension of the family of all ellipsoids, i.e. at least $\frac{N \cdot (N+3)}{2} + 1 = r_{\min} + 2$. However, we have a $r_{\min}$-dimensional class of 1-dimensional families of sets, so the overall dimension of the set of all the sets S_c cannot be larger than $r_{\min} + 1$. This contradiction shows that the set S_c cannot be different from the enclosing minimal-volume ellipsoid. Thus, indeed, each set from each family from the optimal class is an ellipsoid.

4°. To complete the proof, we need to prove that ellipsoids in each family are concentric and homothetic.

We have proven that each ellipsoid appears as an appropriate smallest-volume set. Now that we know that each set S_c coincides with its smallest-volume enclosure, we can thus conclude that each ellipsoid appears as one of the sets S_c from one of the families from the optimal class. Similarly to Part 3 of this proof, let us consider the unit ball centered at 0. If the 1-dimensional family F_0 containing this ball is not invariant with respect to all possible rotations around the ball's center, then we have at least a 1-dimensional group of different families containing the same ellipsoid – the unit ball. However, the only way for an $r_{\min}$-dimensional class of 1-dimensional families to cover the whole $(r_{\max} + 1)$-dimensional family of ellipsoids is when all elements of all families are different. So we cannot have several families containing the same ellipsoid.

This argument shows that the family F_0 containing the unit ball *should be* rotation-invariant. Since all the sets from this family are included in each other and thus, cannot be transformed into each other by rotations – this means that each ellipsoid from this family F_0 must be rotation-invariant. This means that each ellipsoid from this family must be a ball concentric with our selected unit ball – and thus, homothetic to this ball.

For any other family F, by selecting any ellipsoid E from this family and applying the affine transformation that transforms the above unit ball into E, we get a new family $T(F_0)$ of concentric homothetic ellipsoids. Since an ellipsoid can only belong to one family, we thus conclude that the family F also consists of concentric homothetic ellipsoids.

The proposition is proven.

5. Conclusions

Wood is one the oldest construction materials. However, in spite of several thousand years of experience with wooden constructions, predicting and estimating mechanical properties of wooden constructions remains a difficult problem. One of the main reasons for this difficulty is that, in contrast to many other constructions materials which are largely homogeneous and isotropic, wood is highly inhomogeneous and anisotropic. Recently, a new property of wooden materials was discovered that has a potential to make mechanical analysis of wooden structures more efficient: that for wood, iso-correlation surfaces (i.e., surfaces of equal correlation) are well-approximated by concentric homothetic ellipsoids. The problem is that this property is purely empirical, it has no theoretical explanation and thus, engineers are understandably reluctant to rely on it in their estimates. In this papet, we provide a theoretical explanation for this empirical fact and thus, make this property more reliable and therefore more useable.

Acknowledgements

This work was supported in part by the US National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science) and HRD-1242122 (Cyber-ShARE Center of Excellence).

The authors are thankful to the anonymous referees for valuable suggestions.

References

Belforte, G., B. Bona. An improved parameter identification algorithm for signal with unknown-but-bounded errors. *Proceeding of the 7th IFAC Symposium on Identification and Parameter Estimation*, York, U.K., 1985.

Busemann, H. *The Geometry of Geodesics*. Academic Press, New York, 1955.

Chernousko, F. L. *Estimation of the Phase Space of Dynamic Systems*. Nauka publ., Moscow, 1988 (in Russian).

Chernousko, F. L. *State Estimation for Dynamic Systems*. CRC Press, Boca Raton, Florida, 1994.

Filippov, A. F. Ellipsoidal estimates for a solution of a system of differential equations. *Interval Computations*, No. 2(4):6–17, 1992.

Finkelstein, A., O. Kosheleva, and V. Kreinovich. Astrogeometry, error estimation, and other applications of set-valued analysis. *ACM SIGNUM Newsletter*, 31(4):3–25, 1996.

Fleischhauer, R., J. U. Hartig, P. Haller, and M. Kaliske. Moisture-dependent thermomechanical constitutive modeling of wood. *Engineering Computations*, 36(1), 2019.

Fogel, E., Y. F. Huang. On the value of information in system identification. Bounded noise case. *Automatica*, 18:229–238, 1982.

Jenkel, C., F. Leichsenring, W. Graf, and M. Kaliske. Stochastic modelling of uncertainty in timber engineering. *Engineering Structures*, 99:296–310, 2015.

Kollmann, F. *Technologie des Holzes und der Holzwerkstoffe*. Springer, Berlin, 1982.

Leichsenring, F., C. Jenkel, W. Graf, and M. Kaliske. Numerical simulation of wooden structures with polymorphic uncertainty in material properties. *International Journal of Reliability and Safety*, 12 (1/2), 24–45, 2018.

Li, S., Y. Ogura, and V. Kreinovich. *Limit Theorems and Applications of Set Valued and Fuzzy Valued Random Variables*. Kluwer Academic Publishers, Dordrecht, 2002.

Norton, J. P. Identification and application of bounded parameter models. *Proceeding of the 7th IFAC Symposium on Identification and Parameter Estimation*, York, U.K., 1985.

Reuter, U. *Modellierung ungewisser Daten des Bauingenieurwesens und Anwendung der Modelle in Statik und Dynamik*. Habilitation, Faculty of Civil Engineering – Technische Universität Dresden, 2009.

Schietzold, F. N., W. Graf, and M. Kaliske. Polymorphic uncertainty modeling for optimization of timber structures. *Proceedings of the 8th International Workshop on Reliable Engineering Computing REC'2018*, Liverpool, UK, 2018, pp. 63–78.

Schietzold, F. N., W. Graf, and M. Kaliske. Optimization of tree trunk axes locations in polymorphic uncertain modeled timber structures. In: M. Beer and E. Zio, editors, *Proceedings of the 29th European Safety and Reliability Conference ESREL'2019*, Hannover, Germany, Research Publishing, Singapore, pp. 3201–3208, 2019.

Schweppe, F. C. Recursive state estimation: unknown but bounded errors and system inputs. *IEEE Transactions on Automatic Control*, 13:22, 1968.

Schweppe, F. C. *Uncertain Dynamic Systems*. Prentice Hall, Englewood Cliffs, New Jersey, 1973.

Sheskin, D. J. *Handbook of Parametric and Non-Parametric Statistical Procedures*. Chapman & Hall/CRC, London, UK, 2011.

Soltanov, S. T. Asymptotics of the function of the outer estimation ellipsoid for a linear singularly perturbed controlled system. In: S. P. Shary and Yu. I. Shokin (eds.), *Interval Analysis*, Krasnoyarsk, Academy of Sciences Computing Center, Technical Report No. 17, 1990, pp. 35–40 (in Russian).

Ulrich, H. and W. Seim. *Experimentelle Prüfungen an Kleinproben aus Fichtenholz*. Report G01-2014, Fachbereich Bauingenieur- und Umweltingenieurwesen, Fachgebiet Bauwerkserhaltung und Holzbau, Universität Kassel, 2014.

Utyubaev, G. S. On the ellipsoid method for a system of linear differential equations. In: S. P. Shary (ed.), *Interval Analysis*, Krasnoyarsk, Academy of Sciences Computing Center, Technical Report No. 16, 1990, pp. 29–32 (in Russian).

Vorreiter, L. *Holztechnologisches Handbuch, Band 1*. Verlag Georg Fromme & Co., Vienna, 1949.

Imprecise subset simulation for reliability analysis

Dimitris G. Giovanis, Michael D. Shields
Johns Hopkins University, Department of Civil and Systems Engineering, dgiovan1@jhu.edu, michael.shields@jhu.edu

Abstract. The objective of this work is to quantify the uncertainty in structural reliability that results from lack of available data, necessary to precisely identify probability distributions for the model input parameters. To this end, we propose a computationally efficient framework that incorporates multimodel Bayesian/information theoretic method for probabilistic inference in the context of the popular Subset simulation (SuS), in order to estimate the uncertainty in the probability of failure estimates for a given structure. The proposed method consists of two steps: First, a multi-model Bayesian/information theoretic framework is utilized in order to assess the model-form and parametric uncertainty of the input random variables. Subsequently, imprecise SuS is conducted for the identified family of distributions. The outcome of the proposed method is imprecise probabilities of given response quantities in the form of probabilities of probabilities that allows us to bound response quantities of interest in a probabilistic manner and therefore assess confidence in our probabilistic estimates.

Keywords: Subset simulation, imprecise probabilities, structural reliability

1. Introduction

Various methods have been developed over the last decades to estimate the reliability of an engineering system, such as the widely-used First and Second Order Reliability methods (FORM/SORM) (Rackwitz and Fiessler, 1978; Breitung, 1989), response surface methods (Faravelli, 1989), as well as improved simulation approaches such as the robust Monte Carlo simulation (MCS) (Fishman, 1996; Rubinstein, 1981), variance reduction methods based on Markov Chain Monte Carlo (MCMC) algorithms (e.g. line sampling (Pradlwarter et al., 2007), importance sampling (Schuëller et al., 2004; Koutsourelakis et al., 2004), and subset simulation (Au and Beck, 2001)) to name just a few. However, the success of these reliability methods is conditional on exact knowledge of the probability distributions of the random variables, which allows mapping the problem in the standard normal space.

In reality, exact distributions for the random variables of a structural system are rarely known. This lack of knowledge limits the applicability of even the most robust reliability methods, despite their many advances. The challenge is thus to infer the distribution models from data that describe the system. However, when the available data are scant, it is often impossible to assign a unique probability distribution model for the random variables, which introduces uncertainty (or imprecision) into the estimate. Along those lines, over the last several years significant attention has been placed on the impact of imperfect knowledge about the probability distributions on the

© 2021 *by D. G. Giovanis and M. D. Shields.*

evaluation of structural reliability by the scientific community (Hurtado et al., 2012; Fetz and Oberguggenberger, 2016; Zhang and Shields, 2018a; Wang et al., 2012).

In this work we address the challenge of quantifying uncertainty in probability of failure estimates using subset simulation (SuS) when data for inferring distributions are scarse. SuS has proven very successful for wide-ranging problems when the probability models (and hence the conditional levels) are uniquely prescribed. However, when the probability models are uncertain and represented by multiple models, the location of each conditional level is uncertain, which induces uncertainty in probability of failure estimates. The paper is organized as follows: Section 2 presents an overview of subset simulation for structural reliability analysis. In Section 3, we discuss the tools necessary for quantifying model-form and parametric uncertainty from sparse data. In Section 4, the proposed imprecise subset simulation is introduced. Last but not least, in Section 5 we utilize the proposed methodology to conduct uncertainty quantification on the probability of failure estimates using imprecise SuS for a stochastic single degree freedom system under dynamic excitation.

2. Review of subset simulation (SuS)

Consider that the behavior of a structural system is reflected in a performance function $g(\mathbf{x})$, where $\mathbf{x} = [X_1, \ldots, X_n] \in \mathbb{R}^n$ is a vector of the n uncertain parameters of the system. Structural reliability is interested in determining the probability of failure of the system given by

$$P_F = \mathbb{P}(g(\mathbf{x}) \leq 0) = \int_F \mathrm{I}_F(\mathbf{x}) f(\mathbf{x}) \mathrm{d}\mathbf{x} \tag{1}$$

where F is the failure region, $f(\mathbf{x})$ is the joint probability density of $\mathbf{x}$, and I_F is the indicator function: $\mathrm{I}_F(\mathbf{x}) = \mathbf{1}$ if $g(\mathbf{x}) \leq 0$ and $\mathrm{I}_F(\mathbf{x}) = 0$ otherwise. For practical issues, it is preferable to transfer the problem to the standard normal space (with the one-to-one mapping $\mathbf{u} = \mathcal{T}(\mathbf{x})$) where calculations are easier. The probability of failure can be then expressed as

$$P_F = \mathbb{P}(G(\mathbf{u}) \leq 0) = \int_F \mathrm{I}_F(\mathbf{u}) \varphi_n(\mathbf{u}) \mathrm{d}\mathbf{u} \tag{2}$$

where $\varphi_n(\mathbf{u}) = \prod_{j=1}^n \phi_j(u_j)$ and $\phi_j(u_j)$ is the standard normal probability density function and $G(\mathbf{u}) = g(\mathcal{T}^{-1}(\mathbf{u}))$ is the performance function in the standard normal space.

Subset simulation (Au and Beck, 2001) is a variance reduction method that has drawn significant attention because of its efficiency to estimate small failure probabilities in high-dimensional spaces. In SuS, variance reduction is achieved through representation of the failure event in the standard normal space, $F = \{\mathbf{u} \in \mathbb{R}^n : G(\mathbf{u}) \leq 0\}$, as a sequence of M partial failure events $F_1, F_2, \cdots, F_M$, such that $F_1 \supset F_2 \supset \cdots \supset F_M$, and $F = \cap_{i=1}^M F_i$. This division into subsets offers the possibility to transform the simulation of rare events into a set of simulations of more frequent events and evaluate the failure probability P_F as a product of probabilities

$$P_F = \mathbb{P}[F] = \mathbb{P}[\cap_{i=1}^m F_i] = \mathbb{P}[F_1] \prod_{i=2}^m \mathbb{P}(F_i | F_{i-1}) = P_1 \prod_{i=2}^m P_i \tag{3}$$

Each intermediate failure event is defined as $F_i = \{G(\mathbf{u}) \leq b_i\}$, where $b_i > 0$ is a threshold value selected in a way that nearly equal partial failure probabilities $P_i,\ i = 2, \cdots, M$ are obtained for each subset. However, it is difficult to specify in advance the limit values b_i according to a prescribed probability P_i. Therefore the limit values have to be determined adaptively within the simulation. Usually, each conditional probability P_i equals a target value typically chosen between 0.1 and 0.3 (Zuev et al., 2012).

Although the probability $P_1 = \mathbb{P}[F_1]$ can be easily computed through Monte Carlo simulation, calculating the subsequent conditional probabilities is not trivial. It requires generation of samples from the conditional distribution $\varphi_n(\mathbf{u}|F_i)$ given that $\mathbf{u}$ lies in F_i. To this end, several MCMC algorithms have been proposed to drawn samples from $\varphi_n(\mathbf{u}|F_i)$. In the original SuS work (Au and Beck, 2001), a component-wise modified Metropolis-Hasting (MMH) algorithm was introduced. Since that time, A large number of new MCMC algorithms have been introduced. These include algorithms based on repeated generations of the candidate state until the first acceptance criterion is satisfied (Santoso et al., 2011), the related delayed rejection method (Miao and Ghosh, 2011; Au et al., 2012; Zuev et al., 2012), methods that translate gradually samples from the prior to samples from the posterior distribution through a combination of importance sampling and MCMC (Ching and Chen, 2007) and methods to adaptively update the proposal density in each conditional level (Papaioannou et al., 2015). In (Shields et al., 2020) the authors proposed to use the affine invariant ensemble sampler with stretch moves proposed by Goodman and Weare (Goodman and Weare, 2010), to perform SuS for problems where distributions or conditional distributions are strongly non-Gaussian, highly dependent, and/or degenerate. The main benefit of this algorithm is that it allows SuS to be conducted for problems of arbitrary complexity in the original parameter space of random variables (i.e. without transformation of variables to uncorrelated standard normal). For this reason, in this work we utilize the affine invariant ensemble sampler-based SuS.

3. Quantifying model-form and parametric uncertainty from small data

In order to perform reliability analysis of a structural system, knowledge of the distribution model $\mathcal{M}_j$ (with probability density function (pdf) $p_j(x;\theta_j)$ and corresponding parameters θ_j) of each random variable X_j is required. However, in order to be able to identify a unique "best" model among a set of candidate models (selected by the user), inference of collected data $\mathbf{d}$ is necessary.

Consider a set of candidate probability models $\{\mathcal{M}_l\}_{l=1}^{N_d}$ that reflects a range of plausible distributions of the random variable X_j. The multi-model inference framework, first presented by (Burnham and Anderson, 2004) assigns a probability π_l to model $\mathcal{M}_l$ which reflects our confidence that this model is the "best" model (among the candidates models) to represent the data. In this setting, the information theoretic multi-model selection approach, introduced in (Zhang and Shields, 2018a) will be used to calculate these probabilities by measuring the Kullback-Leibler (K-L) information loss ((Kullback and Leibler, 1951)). Alternatively, Bayesian multi-model selection can be used as presented in (Zhang and Shields, 2018b). More specifically, the Akaike Information Criterion (AIC) ((Akaike, 1974)), defined as

$$\mathrm{AIC}^{(l)} = -2\log(\mathcal{L}(\theta_l^\star|\mathbf{d},\mathcal{M}_l)) + 2k_l \tag{4}$$

is used to quantify the information loss from representing the data with candidate model $\mathcal{M}_l$. In this equation, $\mathcal{L}(\cdot)$ is the likelihood function for model $\mathcal{M}_l$ evaluated at the maximum likelihood parameters $\theta_l^\star$ given the data $\mathbf{d}$. Moreover, k_l is the dimension of the parameter vector θ_l and the term $2k_l$ a the bias correction factor. Since the AIC is an asymptotic quantity relying on large data sets, an extension of the AIC introduced by (Hurvich and Tsai, 1989), termed AIC_c, is utilized when dealing with small data sets. The modified criterion incorporates a second-order bias correction term as

$$\text{AIC}_c^{(l)} = -2\log(\mathcal{L}(\theta^\star|\mathbf{d},\mathcal{M}_l)) + 2k_l + \frac{(2k_l^2+k_l)}{n-k_l-1} \tag{5}$$

which is a function of is the size n of the data set and the dimension k_l of the parameter vector.

Model probabilities can be derived from the AIC_c, under the assumption that all prior models have equal probability $1/N_d$, as

$$\pi_l = p(\mathcal{M}_l|\mathbf{d}) = \frac{\exp\left(\frac{-\Delta_A^{(l)}}{2}\right)}{\sum_{l=1}^{N_d}\exp\left(\frac{-\Delta_A^{(l)}}{2}\right)} \tag{6}$$

where $\Delta_A^{(l)} = \text{AIC}_c^{(l)} - \min_j(\text{AIC}_c^{(l)})$ is the relative information loss (in this scale the most probable model has $\Delta_A^{(j)} = 0$).

After quantifying model-form uncertainty, we also need to quantify uncertainties associated with the parameters θ_r of model $\mathcal{M}_r$ where $r = 1,\ldots,m$, where m is the number of models from the original model set having non-neglibile model probability. This is done using standard Bayesian parameter estimation where a prior pdf $p(\theta_r;\mathcal{M}_r)$ is assigned that reflects our current belief about the values of the parameters. Appropriate selection of a non-informative prior distribution is very important when datasets are small (Zhang and Shields, 2018b). A popular and simple choice, though not necessarily the best choice, is the uniform prior $p(\theta_r;\mathcal{M}_r) = 1/m$ which favors all parameters equally within a specified range. The posterior pdf is estimated using Bayes' Rule as:

$$p^\star(\theta_r|\mathbf{d},\mathcal{M}_r) = \frac{p(\mathrm{d}|\theta_r,\mathcal{M}_r)p(\theta_r;\mathcal{M}_r)}{p(\mathrm{d};\mathcal{M}_r)} \propto \mathcal{L}(\theta_r|\mathrm{d},\mathcal{M}_r)p(\theta_r;\mathcal{M}_r) \tag{7}$$

where $\mathcal{L}(\theta_r|\mathrm{d};\mathcal{M}_r) = p(\mathrm{d}|\theta_r;\mathcal{M}_r)$ is the likelihood of the parameters θ_r given the observed data d, and $p(\mathrm{d};\mathcal{M}_r)$ is the evidence estimated by

$$p(\mathbf{d};\mathcal{M}_r) = \int p(\mathbf{d}|\theta_r;\mathcal{M}_r)p(\theta_r;\mathcal{M}_r)d\theta_r \tag{8}$$

Since computation of the evidence is often not straightforward, a Markov Chain Monte Carlo (MCMC) approach is applied to draw samples from the posterior $p^\star(\theta_r|\mathrm{d};\mathcal{M}_r)$.

This two-step multi-model selection process provides set of models $\{\mathcal{M}_r\}_{r=1}^m$ with associated model probabilities π_r and the joint posterior pdf of each model's parameters. In order to use these multimodel probabilities, we randomly sample from the model set and model parameters by means of Monte Carlo sampling to establish a finite set of n_c probabilities models to be used for

reliability analysis, as follows. First, we randomly pick a probability model $\mathcal{M}_r$ according to the model probabilities π_r. We then randomly sample the parameters of that model from the joint posterior density obtained from MCMC. For uncertainty propagation it is shown in (Zhang and Shields, 2018a) that n_c can be made arbitrarily large without adding significant computational expense.

4. Imprecise subset simulation (iSuS)

It is widely accepted that subset simulation is efficient and precise when the probability model on the parameter space is uniquely prescribed. However, when the probability model is uncertain and represented by multiple models, there is currently no means by which to incorporate this uncertainty in reliability estimates short of conducting repeated subset simulations, which comes at extremely large computational cost. That is, if conditional probabilities in SuS are held constant, then threshold value b_i (delineating the conditional levels) changes for distribution model $\mathcal{M}_r$ as illustrated in Fig.1). Thus, a new SuS is required for each of the model, which is clearly intractable, especially for real-world structures.

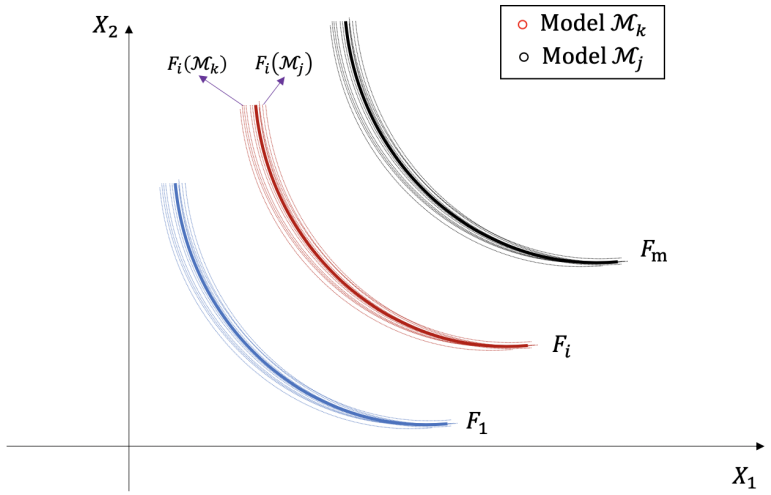


Figure 1. For distribution models $\mathcal{M}_j$ and $\mathcal{M}_k$ running SuS will result in two distinct conditional levels $F_i(\mathcal{M}_j)$ and $F_i(\mathcal{M}_k)$, respectively, if the conditional probability P_i is kept constant.

To overcome this challenge, we propose an approach in which a single SuS is conducted and reliability analysis is conducted using all candidate models are simultaneously. This is achieved by fixing the conditional performance levels F_i, and allowing the conditional probabilities to vary with each probability distribution in the multi-model set, i.e. $P_i \to P_{ir} \equiv P_i(\mathcal{M}_r)$. To this end, an optimal joint sampling density $q_n^\star(\cdot|F_i)$ is identified in order to generate samples that lie in the conditional levels F_i. Then, importance sampling is used to re-weight the samples according to each

distribution in the set and recalculate the conditional probabilities P_{ir} for each model $\mathcal{M}_r$. The probability of failure for model $\mathcal{M}_r$ is then calculated as:

$$P_F(\mathcal{M}_r) = P_1(\mathcal{M}_r)\prod_{i=2}^{m} P_{ir}) \tag{9}$$

A key-point of the proposed approach is to identify the optimal sampling density $q_n^\star(\cdot|F_i)$ from which to sample. This optimal density must be representative of all plausible models and their associated probabilities. (Zhang and Shields, 2018a) derived this optimal sampling density analytical as a probability weighted mixture of the $\{\mathcal{M}_r\}_{r=1}^m$ distribution models. This mixture model is used herein.

5. Steps of the proposed method

The precise steps for the proposed method are: For each variable X_j, $j = 1, \ldots, n$

1. Identify the set of plausible models and their associated model probabilities $\{\mathcal{M}_r, \pi_r\}_{r=1}^m$ using information theoretic multi-model selection.
2. For each model $\mathcal{M}_r$: i) identify the joint posterior parameter probability density $p(\theta_r|\mathbf{d}; \mathcal{M}_r)$ using Bayesian inference and, ii) sample ($n_\theta > 10,000$) from this density using MCMC.
3. Select n_c models from the pool of $m \times n_\theta$ candidate models.
4. Construct the optimal marginal sampling density

$$\hat{q}_j^\star(\mathbf{x}) = \sum_{t=1}^{n_c} \pi_t \mathbb{E}_\theta[p_t(\mathbf{x}|\theta_t)] \tag{10}$$

where $\mathbb{E}_\theta[p_t(\vec{x}|\theta_t]$ is the expectation of all distributions of type $\mathcal{M}_t$ with respect to it's model parameters θ, evaluated by simple Monte Carlo.

The next step is to run a single (optimal) SuS with the following target distribution

$$q_n^\star(\mathbf{x}) = \prod_{j=1}^{n} \hat{q}_j^\star(\mathbf{x}) \tag{11}$$

in order to identify the conditional performance levels F_i^{opt} and obtain samples $\mathbf{x}_i^{opt}$ in each conditional level. Then, we use importance sampling (IS) to re-weight the conditional probabilities P_i for each distribution model $\mathcal{M}_t$. More specifically, at subset i, for each model we calculate the importance weight w_t at each sample point $\mathbf{x}_i^{opt}$ as

$$w_t(\mathbf{x}_i^{opt}) = \frac{p_t(\mathbf{x}_i^{opt})}{q_n^\star(\mathbf{x}_i^{opt})}, \quad t = 1, \ldots, n_c \tag{12}$$

where $p_t(\cdot) = p(\theta_t|\mathbf{d}, \mathcal{M}_t)$. As a result, the conditional probability of failure P_i corresponding to model $\mathcal{M}_t$ is estimated as

$$P_i^t = \frac{1}{N} \sum_{l=1}^{N_c} w_t(\mathbf{x}_l) \mathrm{I}_{F_i^{opt}}(\mathbf{x}_l) \tag{13}$$

Calculate the probability of failure for each model $\mathcal{M}_j$ according to

$$P_F^t = \prod_{i=1}^{m} P_i^t \tag{14}$$

The result is a stochastic set of failure probabilities $\{P_F^t\}_{t=1}^{n_c}$ that incorporate the effects of model form and parametric uncertainty on the true probability of failure.

The SuS with the affine invariant MCMC sampler and the multi-model selection framework used in this work are available in the open-source UQpy Python toolbox (Olivier et al., 2020).

6. Numerical example

A single degree of freedom system having stiffness k and mass m is excited by a sinusoidal load $x(t) = \sin(\omega t)$ with frequency ω rad/sec (see Fig.2). Both the stiffness and mass are consider to be stochastic, normally distributed according to $N(\mu_k, \sigma_k)$ and $N(\mu_m, \sigma_m)$, respectively. The system's equation of motion given by:

$$m\ddot{x} + kx = x(t) \tag{15}$$

In this system, resonance occurs when the natural frequency of the system $\omega_n = \sqrt{\frac{k}{m}}$ becomes equal to the excitation frequency ω of the periodically applied force $x(t)$. To avoid resonance, we consider failure of the system to be associated with the natural frequency being within a threshold ϵ of the excitation frequency ω. That is, failure of the system occurs when $\omega - \epsilon \leq \omega_n \leq \omega + \epsilon$. Hence, the probability of failure is defined as:

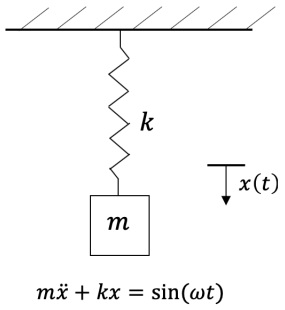


Figure 2. A single degree stochastic system under sinusoidal excitation.

$$P_F = P\left(\omega - \epsilon \leq \omega_n \leq \omega + \epsilon\right) \tag{16}$$

$$= P\left(k - (\omega - \epsilon)^2 m \leq 0 \cup (\omega + \epsilon)^2 m - k \leq 0\right) \tag{17}$$

For the needs of this study we selected $\mu_k = 125$, $\sigma_k = 5$, $\mu_m = 5$, $\sigma_m = 1$, $\omega = 6$ and $\epsilon = 0.0001$. Reliability analysis of the model with means of brute-force Monte Carlo simulation with 10^6 samples gives a probability of failure $P_F = 3.2 \times 10^{-5}$. We also performed 20 independent runs of SuS for the and we calculated a probability of failure $P_F = 3.0 \times 10^{-5}$ with coefficient of variation $c.o.v = 0.40$. Figure 3 shows the samples drawn using a single SuS with 1000 samples per subset level with conditional probability equal to 0.1.

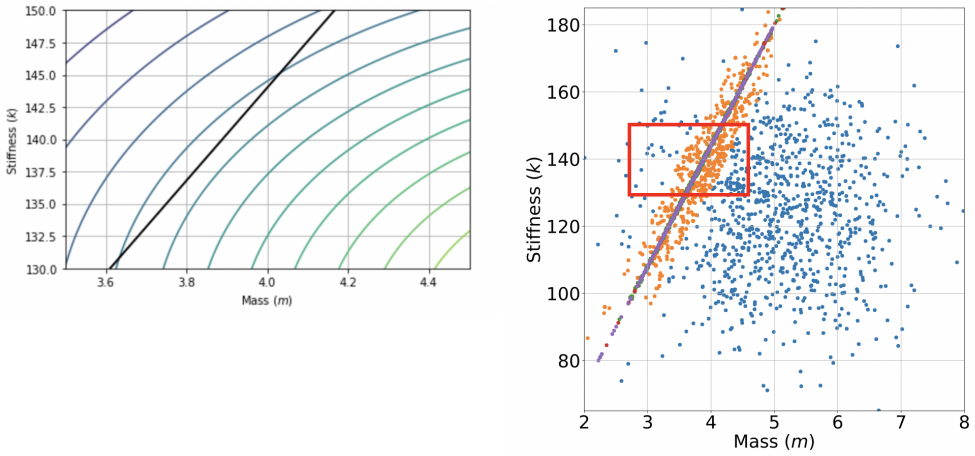


Figure 3. Limit-state function and samples drawn with SuS.

Given limited data from which to infer the distributions of k and m, the multi-model selection framework is applied to each random variable in order to select the "best" probability models from the following distribution families: *lognormal, logistic, inverse Gaussian, Maxwell, Levy* and *normal*. We consider data sets of size 20, 50, 100, 1000 and 10000 in order to showcase the sensitivity of the methods to the number of available data. Figure 4 depicts the probability π of each candidate model as a function of data set size. From this figure we can observe that at least 100 data points are required in order to uniquely identify the true normal distribution for the stiffness and more than 1000 data points are required for the mass.

Next, the joint parameter distribution for each one of the selected models is determined through Bayesian inference. However, for brevity these distributions are not shown. Figure 5 depicts 1000 sampled candidate distributions for each random variable and for increasing data set size. Notice that the clouds of possible distributions narrow with more available data.

To assess the influence of model-form and parametric uncertainty on P_F estimates, we apply the proposed iSuS method. First, we identify the joint optimal sampling density function $q_n^\star(\mathbf{x})$, shown in Figure 6, for different data set sizes. As can be seen from this figure, for increasing number of available data the probability density is converging to the true joint normal distribution.

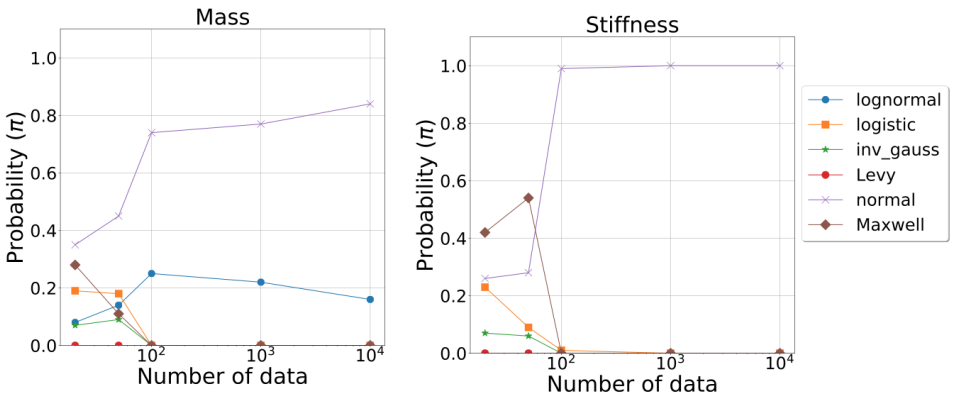


Figure 4. Model probabilities as a function of data set size.

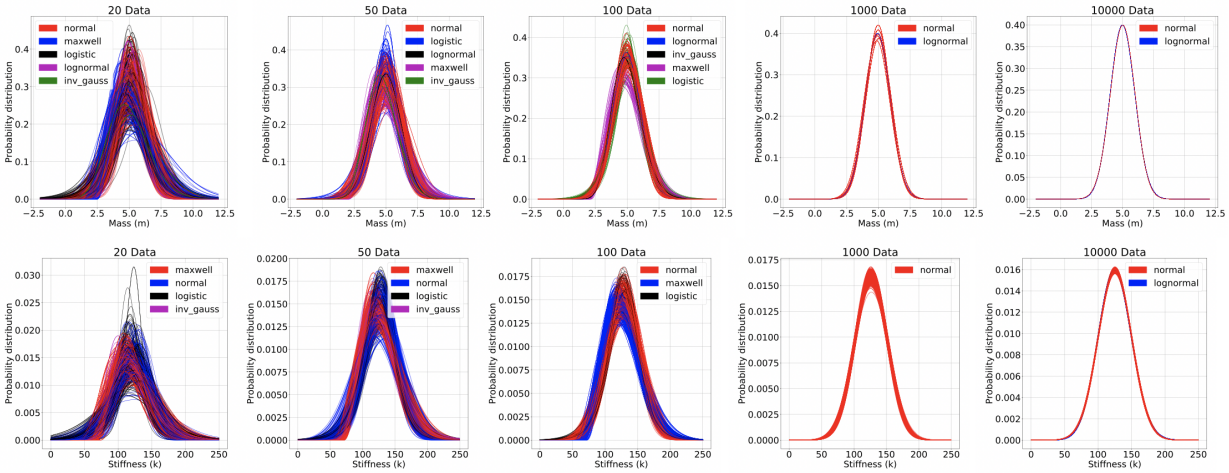


Figure 5. Example 1: Sample sets of probability distributions for different data set sizes for mass m (top) and stiffness k (bottom).

After identifying the optimal sampling distribution we run SuS once, in order to identify the optimal conditional performance levels F_i^{opt} and the conditional samples. Then, the importance weight of each sample is used to weight the conditional probabilities. As a result, each conditional probability of failure is re-evaluated and the probability of failure for the specific model is calculated. Figure 7 shows the resulting mean empirical cumulative distribution functions (from 20 repeated iSuS evaluations) for the probability of failure for the different data set sizes. From this figure we can identify a trend that results in large uncertainty in P_F estimats for small data sets and convergence towards the true P_F for increasing data set size.

We highlight at this point that, with the proposed method assessing uncertainty in P_F estimates requires a single iSuS evaluation, which is equivalent to a single SuS evaluation plus the nominal cost

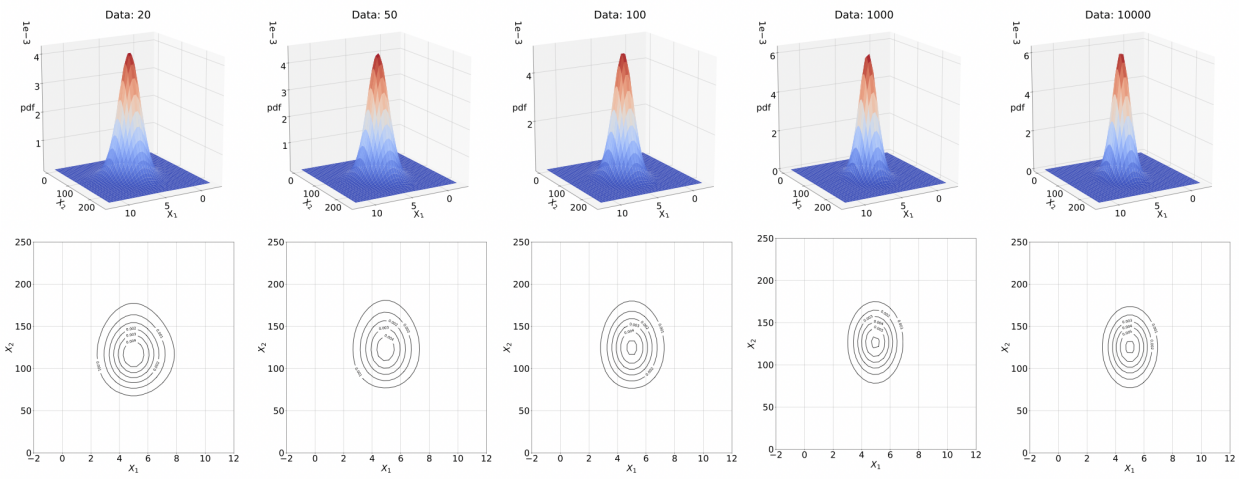


Figure 6. Example 1: Joint probability density $q_n^\star(\mathbf{x})$ (top) with the respective contour (bottom) that is used as the target density in the framework of SuS, for different size of the data set.

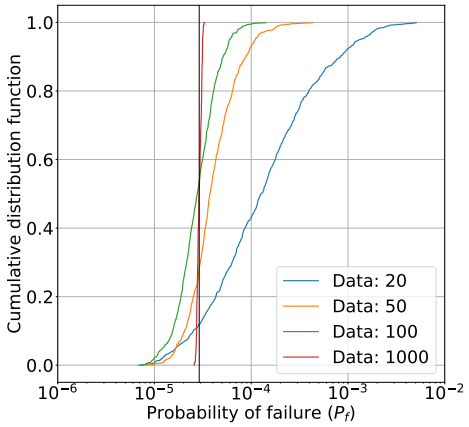


Figure 7. Emperical cdfs of the probability of failure P_F for different data set sizes resulting from 20 iSuS runs. The true $P_F = 2.9 \times 10^{-5}$ is indicated by a vertical black line.

of model reweighting. Accessing the same information using a brute force Monte Carlo approach, requires a new SuS for each model $\mathcal{M}_j$, or a total of 1000 SuS analyses in this case.

7. Conclusions

We propose a method to quantify the uncertainty in structural reliability estimates using subset simulation that results from lack of available data. Our motivation stems from the fact that subset

simulation is very efficient and precise when the probability model on the parameter space (and hence the conditional levels) is uniquely prescribed. However, when the probability model is uncertain and represented by multiple models the location of each conditional level is uncertain which induces uncertainty in the probability of failure estimate. We propose a framework build upon multimodel Bayesian/information theoretic methods and Subset simulation. The outcomes of the proposed method are imprecise probabilities of given response quantities in the form of probabilities of failure probabilities that allows us to bound the probability of failure in a probabilistic manner and therefore assess confidence in our estimates.

Acknowledgements

This work has been supported by the Office of Naval Research under award number N00014- 510 18-1-2644 with Dr. Paul Hess as program officer.

References

Zhang, J. and M. D. Shields On the quantification and propagation of imprecise probabilities resulting from small datasets. *Mechanical Systems and Signal Processing*, 98:465–483, 2018.

Kullback, S. and R. A. Leibler. On information and sufficiency. *Ann. Math. Stat.*, 22(1):79–86, 1951.

Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control*, 19(6):716–723, 1974.

Hurvich, C. M. and C. L. Tsai. Regression and time series model selection in small samples. *Biometrika*, 76(2):297–307, 1989.

Au, S. and J. Beck. Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4):263–277, 2001.

Papaioannou, I. and W. Betz and K. Zwirglmaier and D. Straub. MCMC algorithms for Subset Simulation. *Probabilistic Engineering Mechanics*, 41:89–103, 2015.

Shields, M. D. and D. G. Giovanis and V. Sundar. Subset Simulation for problems with strongly non-Gaussian, highly anisotropic, and degenerate distributions. *Computers and Structures*, Accepted, 2020.

Rackwitz, R. and B. Fiessler. Structural reliability under combined load sequences. *Computers and Structures*, 9:489–494, 1978.

Breitung, K. Asymptotic approximations for probability integrals. *Prob. Eng. Mech.*, 4(4):187–190, 1989.

Guan, X. L. and R. E. Melchers. Effect of response surface parameter variation on structural reliability estimates . *Structural Safety*, 23:429–444, 2001.

Fishman, G. S. Monte Carlo: concepts, algorithms, and applications. *Springer*, New York, 1996.

Rubinstein, R. Y. Simulation and the Monte-Carlo method. *Wiley*, New York, 1981.

Pradlwarter, H. J. and G. I. Schuëller and P. S. Koutsourelakis D. C. Charmpis. Application of line sampling simulation method to reliability benchmark problems. *Structural Safety*, 29:208-221, 2007.

Schuëller, G. I. and H. J. Pradlwarter and P. S. Koutsourelakis. A critical appraisal of reliability estimation procedures for high dimensions. *Structural Safety*, 19:463-474, 2004.

Koutsourelakis, P. S. and H. J. Pradlwarter and G. I. Schuëller. Reliability of structures in high dimensions, part I: algorithms and applications. *Structural Safety*, 19:409-417, 2004.

Goodman,J. and J. Weare. Ensemble samplers with affine invariance. *Commun. Appl. Math. Comput. Sci*, 5(1):65-80, 2010.

Miao, F. and M. Ghosh. Modified subset simulation method for reliability analysis of structural systems. *Structural Safety*, 4:251-260, 2011.

Zuev, K. M. and J. L. Beck and S. K. Au and L. S. Katafygiotis. Bayesian post-processor and other enhancements of Subset Simulation for estimating failure probabilities in high dimensions. *Computers & structures*, 92:283-296, 2012.

Au, S. K. and J. L. Beck and K. M. Zuev and L. S. Katafygiotis. Discussion of paper by F. Miao and M. Ghosn "Modified subset simulation method for reliability analysis of structural systems, Structural Safety, 33, 251-260, 2011. *Structural Safety*, 34(1):379-380, 2012.

Olivier, A. and D. G. Giovanis and B. S. Aakash and M. Chauhan and L. Vandanapu and M. D. Shields. UQpy: A general purpose Python package and development environment for uncertainty quantification. *Journal of Computational Science.*, Vol. 47, 101204, 2020.

Santoso, A. M. and K. K. Phoon and S. T. Quek. Modified Metropolis–Hastings algorithm with reduced chain correlation for efficient subset simulation. *Probabilistic Engineering Mechanics*, 26(2):331-341, 2011.

Ching, J. and Y. Chen. Transitional Markov Chain Monte Carlo method for Bayesian model updating, model class selection, and model averaging. *J. Eng. Mech.*, 133:816-832, 2007.

Beran, R. Minimum Hellinger distance estimates for parametric models. *Ann. Stat.*, 445-463, 1977.

Lindsay, B. G. Efficiency versus robustness: the case for minimum Hellinger distance and related methods. *Ann. Stat.*, 1081-1114, 1994.

Lindsay, B. G. Efficiency versus robustness: the case for minimum Hellinger distance and related methods. *Ann. Stat.*, 1081-1114, 1994.

Hess, P. E. and . Bruchman and I. A. Assakaf and B. M. Ayyub. Uncertainties in material and geometric strength and load variables. *Naval Eng. J.*, 2, 2002.

Faravelli, L. Response surface approach for reliability analysis. *J. of Eng. Mech.*, 2763-2781, 1989.

Fetz, T., M. Oberguggenberger. Imprecise random variables, random sets, and Monte Carlo simulation. International Journal of Approximate Reasoning, 78:252-264, 2016.

Hurtado, J. E. and D. A. Alvarez and J. Ramírez. Fuzzy structural analysis based on fundamental reliability concepts. Computers & Structures, 112:183-192, 2012.

Wang, C. and H. Zhang and M. Beer. Computing tight bounds of structural reliability under imprecise probabilistic information. Computers & Structures, 208:92-104, 2018.

Burnham, K. P. and D. R. Anderson. Multimodel inference: understanding AIC and BIC in model selection. Sociological methods & research, 33:261–304, 2004.

Zhang, J. and M. D. Shields. The effect of prior probabilities on quantification and propagation of imprecise probabilities resulting from small datasets. *Computer Methods in Applied Mechanics and Engineering*, 334:483–506, 2018.

Enabling the evidence theory through non-intrusive parametric model order reduction for crash simulations

J. S. Jehle[1),2)], V. A. Lange[1),3)], M. Gerdts[2)]
[1)] *Research and Innovation Center, BMW Group, D-80788 Munich, Germany, jonas.jehle@bmw.de, volker.vl.lange@bmw.de*
[2)] *Institute of Applied Mathematics and Scientific Computing, Department of Aerospace Engineering, Universität der Bundeswehr München, D-85577 Neubiberg, Germany, jonas.jehle@unibw.de, matthias.gerdts@unibw.de*
[3)] *Associate Professorship of Computational Mechanics, Department of Civil, Geo and Environmental Engineering, Technical University of Munich, D-80333 Munich, Germany, volker.vl.lange@tum.de*

Abstract. To improve the crashworthiness of cars, it is crucial for vehicle safety departments to consider uncertainties. In addition to hardware crash tests, finite element simulations (FES) serve to assess the behavior in a crash scenario. As these black-box simulations are computationally expensive, e.g. due to the huge number of finite elements, uncertainty quantification has prohibitively high cost. We therefore propose to approximate the FES using non-intrusive parametric model order reduction (MOR). The use of MOR requires model outputs of training simulations—called snapshots—resulting from different input parameter sets. Here, snapshots consist of the node displacements of all finite elements at specific time states. MOR strongly reduces the high dimension of the problem to a lower one by performing a singular value decomposition. Afterwards, the lower dimensional subspace is approximated by metamodels—one for each dimension. This study applies Gaussian process regressions. Combining MOR with metamodels allows the reconstruction of the full FES. New parameter sets can be quickly evaluated, i.e. uncertainties can be rapidly propagated through the reduced model. This enables the use of the evidence theory for each key result of the FES. The resulting plausibility and belief curve help vehicle safety engineers to improve the crashworthiness under uncertainty.

Keywords: non-intrusive parametric model order reduction, metamodel, uncertainty quantification, evidence theory, crash simulation

1. Introduction

Hardware crash tests of vehicles are subject to various uncertainties that influence the test result. It is common to distinguish between the epistemic and aleatory category of uncertainties in literature (Der Kiureghian and Ditlevsen, 2009; Beyer and Sendhoff, 2007). Epistemic uncertainties are caused by lack of knowledge when the proper value for a certain quantity can not be fixed. They can be reduced by clearer insight through additional or improved data (Swiler and Giunta, 2007). Aleatory uncertainties are irreducible and imply that the system response under investigation is

© 2021 *by J. S. Jehle, V. A. Lange, M. Gerdts.*

intrinsically random (Oberkampf et al., 2004). Both forms of uncertainty are encountered during vehicle development.

Component properties such as sheet metal thickness, elasticity modules, or bumper height are not specified in early development phases. Later on, they are fixed and these epistemic uncertainties vanish. Test scatter like impact angle, velocity, or the position of the dummy is always existent and thus a representative of the aleatory class.

Apart from the hardware tests, vehicle safety engineers work with computer simulations to study the behavior of the car in a collision. These simulations are often finite element simulations (FES). Compared to hardware tests, they are cheap with respect to monetary cost but nonetheless they incur immense computational effort, i.e. time and mobilization of high-performance computing hardware. The expensive calculation costs are due to nonlinear phenomena and the necessary detail of the model, which is reflected by the high number of finite elements. To gain an overview of possible test results, the uncertainties occurring in hardware tests must be incorporated into simulations.

We suggest an approach to propagate these uncertainties, epistemic and aleatory, through automotive crash FES and quantify the outcomes. The main part of this approach is an extended version of the Evidence Theory (ET) based on the work of (Oberkampf et al., 2001) and (Oberkampf and Helton, 2002). This method propagates the predetermined uncertain inputs through the model. The uncertainty of each input is modeled through the use of (sub-)intervals. Experts are responsible for subdividing the intervals and assigning a probability of occurrence to each subinterval. Next, ET considers all possible combinations of subintervals between the different inputs and calculates the composite probabilities for each combination.

Then, optimization strategies are used to find the infimum and supremum values of the model for each subinterval combination. Afterwards, all infimum and supremum values are aggregated, sorted in ascending order, and stored together with their corresponding composite probabilities. Cumulating the probabilities yields two curves belonging to the infimum and supremum values. They have the form of a cumulative distribution function. On the basis of these curves, it is possible to identify with two different measures which outputs of the model are likely to occur.

Executing these optimizations can form a bottleneck in the application of ET. As FES of vehicle crash models are regarded as black-boxes, optimization requires many evaluations. This makes larger analyses with ET prohibitively expensive in terms of time. This bottleneck can be overcome by using faster models.

We thus propose to use non-intrusive parametric Model Order Reduction (MOR) (Guo and Hesthaven, 2018; Le Guennec et al., 2018; Swischuk et al., 2019) for approximating the computer simulation. MOR requires model outputs of training simulations that are generated from various input parameter sets. Here, the model outputs comprise the node displacements of all finite elements at particular time states. A singular value decomposition is performed on the training data to generate a lower dimensional subspace. Metamodels are trained to approximate each of the remaining dimensions. The complete FES can then be reconstructed from the output of a small number of metamodels. As these can be evaluated efficiently, this reduced model is an attractive choice to replace the FES in ET.

The proposed combined approach of non-intrusive parametric MOR and ET is described in more detail in the following sections. The first part of Section 2 guides through the different steps of ET. The second part presents the elements of the non-intrusive MOR approach. The crashbox simulation

characterized in the first part of Section 3 serves as proof of concept. The deformation of the finite element crashbox is simulated with the software (LS-DYNA, 2018). The second part of Section 3 demonstrates the application steps of the new approach needed for the crashbox. We implement our combined method in (MATLAB, 2018). Section 4 analyzes the quality of the approach and discusses results by interpreting the ET curves. Section 5 concludes the paper.

2. Developing the Combined Approach of Non-intrusive Parametric MOR and ET

2.1. Uncertainty Propagation through the System via Evidence Theory

The Evidence Theory (ET) devised by the mathematicians Dempster and Shafer (Shafer, 1976)—also known as Dempster-Shafer Theory or Theory of the Belief Functions—was originally invented to combine information from different sources to a comprehensive statement. Each of these sources has a credibility allocation which is taken into acount during the calculation. In (Nikolaidis et al., 2004; Swiler et al., 2009; Eldred and Swiler, 2009; Auer et al., 2010), ET is extended to propagate interval-based uncertainties of input parameters through the system function. As a result, one obtains lower and upper limits for the cumulative distribution function of the output. These are called belief and plausibility curve, respectively.

To use the ET, it is not necessary to know the form of the system function, f, i.e. we consider a general input-output model,

$$y = f(\mu) \in \mathbb{R}, \tag{1}$$

for the inputs $\mu \in \mathbb{R}^n, n \in \mathbb{N}$, without any further knowledge about f. Hence, f can also be a black-box model. The components, μ_i, $i \in \{1, \dots, n\}$, of the input vector, μ, are viewed as independent probability variables. We assume that each component, μ_i, has a probability density whose support is contained in the connected interval $I_i \subset \mathbb{R}$, that is, I_i covers the range of possible values of μ_i. To start the ET analysis, experts assess these overall intervals in two steps. They are illustrated in Fig. 2 for the exemplary application.

1. The intervals I_i are split into $k_i \in \mathbb{N}$ subintervals,

$$I_i^j \subset I_i, \qquad j \in \{1, \dots, k_i\}, \tag{2}$$

for each input dimension, $i \in \{1, \dots, n\}$. These subsets are called focal elements. By forming their union, a new input set,

$$\hat{I}_i := \bigcup_{j=1}^{k_i} I_i^j, \tag{3}$$

for μ_i is obtained. Note that it is admissible for these subintervals to be overlapping, connected, or disjoint. Clearly, $\hat{I}_i$ is a subset of the overall interval I_i. It is, however, possible that $\hat{I}_i$ is not connected, i.e. it can contain gaps.

2. The experts proceed with a so-called Basic Probability Assignment (BPA) where they allocate probabilities to all focal elements, I_i^j. In mathematical terms, the experts define a BPA operator,

m_i, for every μ_i that maps the focal elements, I_i^j, to a probability. These operators can be expressed as

$$m_i(J) \begin{cases} \in [0,1], & \text{for } J \in \mathcal{I}_i := \left\{ I_i^j \middle| j \in \{1,\dots,k_i\} \right\}, \\ = 0, & \text{for } J \in \mathcal{B}(\mathbb{R}) \backslash \mathcal{I}_i, \end{cases} \tag{4}$$

where $\mathcal{B}(\mathbb{R})$ is the Borel σ-algebra of the real numbers. For each m_i, it must hold that the sum of probabilities distributed to the focal elements equals one,

$$\sum_{J \in \mathcal{I}_i} m_i(J) = 1, \qquad i \in \{1,\dots,n\}. \tag{5}$$

The assessment does not have to be unanimous between the experts. If there are differences, one can combine deviating opinions by the traditional ET approach, e.g. via Dempster's rule of combination (Sentz and Ferson, 2002). Note that the expert assessments can be based on their knowledge, existing data, physical models, related information, etc. For simplicity, only one fictive expert is considered here.

In case the expert assumes an input to be a continuous random variable, its probability density function can be represented approximately by a histogram. For this purpose, there exist different methods to choose this histogram appropriately, see (Knuth, 2019; Scott, 1979; Freedman and Diaconis, 1981). Histograms can be converted into ET structures. This allows the theory to deal roughly with continuous random variables.

After generating these initial data, the ET uncertainty propagation is accomplished in four steps:

3. We combine all possible focal elements between input dimensions to n-dimensional hypercubes called interval cells. There are in total $k := \prod_{i=1}^{n} k_i$ interval cells of the form

$$C^{j_1,j_2,\dots,j_n} := I_1^{j_1} \times I_2^{j_2} \times \cdots \times I_n^{j_n}, \tag{6}$$

for $j_i \in \{1,\dots,k_i\}$ and $i \in \{1,\dots,n\}$. We can simplify the notation by C^l, $l \in \{1,\dots,k\}$, instead of writing $C^{j_1,j_2,\dots,j_n}$, $j_i \in \{1,\dots,k_i\}$, respectively, since there is a one-to-one mapping between $\{1,\dots,k\}$ and $\prod_{i=1}^{n} \{1,\dots,k_i\}$. Furthermore, we denote the set of resulting interval cells as $\mathcal{C}$.

4. The composite BPAs, p^l, for C^l are computed by

$$p^l := p^{j_1,\dots,j_n} := \prod_{i=1}^{n} m_i\left(I_i^{j_i}\right). \tag{7}$$

Herein, we exploited the independence of the probability variables, $\mu_i, i \in \{1,\dots,n\}$. In this way, the single BPAs corresponding to composite focal elements are multiplied together. Note that the probabilities of all interval cells sum up to one, i.e.

$$\sum_{l=1}^{k} p^l = 1. \tag{8}$$

Steps 3 and 4 are depicted in Fig. 3 for the exemplary application, whereas the result of the following steps 5 and 6 are shown in Fig. 4(c).

5. Every interval cell, C^l, has to be propagated through the model. To obtain lower and upper limits for the cumulative distribution function, we solve two optimization problems

$$\overline{y}^l = \sup_{\mu \in C^l} f(\mu), \tag{9}$$

$$\underline{y}^l = \inf_{\mu \in C^l} f(\mu), \tag{10}$$

for each interval cell. In total, we have to solve $2k$ optimization problems. We remark that other works (Helton et al., 2006; Tian et al., 2018) use sampling strategies to find these values where their accuracy strongly depends on the size of the sample.

6. We aggregate all supremum and infimum values, $\overline{y}^l$ and $\underline{y}^l$, along with their composite BPAs, p^l. The belief and plausibility curve result from sorting $\overline{y}^l$ and $\underline{y}^l$ in ascending order and cumulating their probabilities, respectively. Expressed as formulas, we define the belief curve as

$$\mathrm{Bel}(y) = \sum_{\left\{l | y \geq \overline{y}^l\right\}} p^l \tag{11}$$

and the plausibility curve as

$$\mathrm{Pl}(y) = \sum_{\left\{l | y \geq \underline{y}^l\right\}} p^l. \tag{12}$$

From a probabilistic point of view, since μ is considered as a random variable, $y = f(\mu)$ is a random variable as well. The cumulative distribution function of y is unknown but it can be limited using the ET curves. For an arbitrary but fixed $\tilde{y}$, the belief curve forms a lower limit of the cumulative distribution function of y, i.e. it holds

$$\mathrm{Bel}(\tilde{y}) \leq P(y \leq \tilde{y}), \tag{13}$$

where $P(A)$ denotes the probability of event A. Beyond, the plausibility curve provides an upper limit, i.e.

$$P(y \leq \tilde{y}) \leq \mathrm{Pl}(\tilde{y}). \tag{14}$$

Note that computational costs of ET increase, when the number of inputs or the number of focal elements grow, since one has to run more and more optimizations. In addition, when the optimization of the system function is expensive due to nonlinearities or its black-box character, applying ET gets prohibitively costly. Different works therefore propose to couple ET with metamodels to accelerate the optimization, see for example the Multi-Point Approximation Method (Bae et al., 2004) and Stochastic Expansion Methods (Eldred et al., 2011). The metamodel has to cover the

input space $\prod_{i=1}^{n} \hat{I}_i$ and builds a fast-responding mathematical model for the output based on training sample points.

2.2. Non-intrusive Parametric Model Order Reduction

We suggest another approach to overcome the problem of solving the costly ET optimizations for FES. It is similar to traditional metamodels in that we want to quickly produce a scalar output—referred to as the key result here. Conventional metamodels only train on the key results themselves. In contrast, non-intrusive parametric Model Order Reduction (MOR) allows to reconstruct the whole simulation by approximating the node displacements of the finite elements.

Key results can then be extracted from the approximated simulation. When multiple key results are considered, they can be derived from a single model rather than creating separate metamodels. Furthermore, the outputs of finite element crash simulations are vectors that together with the FES time steps form a curve, e.g. the displacement curve for the center node of the bumper. The key results are then extracted from these curves. These extractions are based on known mathematical functions that may be nonlinear or discontinuous. Common metamodels have to model these functions as well. MOR does not have to consider these extractions. The extractions can be applied on the MOR approximations afterwards which may yield better results.

The theory for non-intrusive parametric MOR was introduced in (Guo and Hesthaven, 2018) for quasi-static problems. The same authors extended this approach for time-dependent problems in (Guo and Hesthaven, 2019). We summarize the important points and modify the theory for our problem.

We consider a model expressed by a time-dependent parametrized vector function

$$u(t, \mu), \quad (t, \mu) \in \mathcal{T} \times \mathcal{M} \tag{15}$$

where $\mathcal{T} = [0, T], T \in \mathbb{R}_{>0}$, denotes the time interval and $\mathcal{M} \subset \mathbb{R}^n$ is called the parameter space. In our crash test application, the function u describes the displacements of all nodes of the finite element in x-, y-, and z-direction at a specific time, $t \in \mathcal{T}$. The x-, y-, and z-displacements of each node are understood as three individual degrees of freedom. The total number of degrees of freedom, $N_h \in \mathbb{N}$, is thus obtained by multiplying the number of nodes, $N_{\mathrm{nd}} \in \mathbb{N}$, by three, i.e. $N_h = 3N_{\mathrm{nd}}$. Evaluated at $\left(\tilde{t}, \tilde{\mu}\right) \in \mathcal{T} \times \mathcal{M}$, $u\left(\tilde{t}, \tilde{\mu}\right)$ is a vector of length N_h. We assume that the vector $u\left(\tilde{t}, \tilde{\mu}\right)$ first contains all x-displacements, followed by the y-displacements, and finally the z-displacements; node-by-node in a predetermined order, i.e.

$$u\left(\tilde{t}, \tilde{\mu}\right) = \left(u_{\mathrm{x}_1}, \ldots, u_{\mathrm{x}_{N_{\mathrm{nd}}}}, u_{\mathrm{y}_1}, \ldots, u_{\mathrm{y}_{N_{\mathrm{nd}}}}, u_{\mathrm{z}_1}, \ldots, u_{\mathrm{z}_{N_{\mathrm{nd}}}}\right)^T \left(\tilde{t}, \tilde{\mu}\right) \in \mathbb{R}^{N_h}. \tag{16}$$

In FES, the time space is represented by the discrete set $\mathcal{T}_d := \{t_1, \ldots, t_{N_t}\} \subset \mathcal{T}$ of size $N_t \in \mathbb{N}$. We assume equidistant points in time for simplicity. For an arbitrary but fixed parameter configuration $\tilde{\mu} \in \mathcal{M}$, the whole simulation can then be described by the matrix

$$U(\tilde{\mu}) := \left[u\left(t_1, \tilde{\mu}\right), \ldots, u\left(t_{N_t}, \tilde{\mu}\right)\right] \in \mathbb{R}^{N_h \times N_t}. \tag{17}$$

The non-intrusive parametric MOR approach approximates this matrix using dimensionality reduction and regression in decoupled offline and online stages. First, in an offline fashion, the high number of dimensions in the FES is reduced in three steps:

1. We collect $N_s \in \mathbb{N}$ snapshots, i.e. evaluations of U, by running FES with different parameter constellations, $\mu^1, \ldots, \mu^{N_s} \in \mathcal{M}$, and record them in the snapshot matrix

$$S = (S_1, \ldots, S_{N_s}) := \left[U\left(\mu^1\right), \ldots, U\left(\mu^{N_s}\right)\right] \in \mathbb{R}^{N_h \times N_t N_s}. \tag{18}$$

   We assume

$$N_h > N_t N_s \tag{19}$$

   which is typically fulfilled for FES due to the high number of nodes.

2. The singular values of S are found via singular value decomposition, i.e. we decompose S into

$$S = Q \Sigma Z^T \tag{20}$$

   with the left- and right-singular factors $Q \in \mathbb{R}^{N_h \times N_h}$ and $Z \in \mathbb{R}^{N_t N_s \times N_t N_s}$ being orthogonal matrices. The diagonal matrix, $\Sigma \in \mathbb{R}^{N_h \times N_t N_s}$, moreover contains the singular values, $\sigma_1, \ldots, \sigma_{N_t N_s}$, of S in decreasing order. Here, the MATLAB routine `svd` is chosen for the singular value decomposition.

3. The projection matrix $V \in \mathbb{R}^{N_h \times L}, L \in \mathbb{N}$, reduces the dimensionality of the full space model from N_h to $L \ll N_h$. It is used to map from the full space to the reduced space and back. After finding the lowest natural number L, fulfilling the ratio

$$\sum_{i=1}^{L} \sigma_i^2 \Bigg/ \sum_{j=1}^{N_t N_s} \sigma_j^2 > \epsilon, \tag{21}$$

   for a fixed $\epsilon \in [0.9, 1)$, we obtain V by setting it equal to the first L columns of the left-singular factor Q. This procedure is meant to only use those columns for the reduced space that correspond to dominant modes.

After this reduction procedure, we are able to project a full simulation, $U(\tilde{\mu})$, for $\tilde{\mu} \in \mathcal{M}$, into the reduced space by multiplying with V^T, i.e. we compute

$$R := V^T U(\tilde{\mu}) \in \mathbb{R}^{L \times N_t}. \tag{22}$$

Projecting back into the full space, in turn, is performed by multiplying by V, i.e.

$$VR = VV^T U(\tilde{\mu}) \in \mathbb{R}^{N_h \times N_t} \tag{23}$$

lies again in the full space. The projection error between $U(\tilde{\mu})$ and $VV^T U(\tilde{\mu})$ is measured through the Schmidt-Eckart-Young theorem (Guo and Hesthaven, 2019).

Next, we perform the regression in the reduced space to make predictions in online stages. For that, we require another $N_r \in \mathbb{N}$ snapshots, $U\left(\hat{\mu}^1\right), \ldots, U\left(\hat{\mu}^{N_r}\right) \in \mathbb{R}^{N_h \times N_t}$, for $\hat{\mu}^1, \ldots, \hat{\mu}^{N_r} \in \mathcal{M}$, and project them into the reduced space to train our regression model. The remaining procedure proposes a tensor-decomposition-based regression including the following four steps:

4. We write parts of the projected training snapshots $V^TU\left(\hat{\mu}^1\right),\ldots,V^TU\left(\hat{\mu}^{N_r}\right)\in\mathbb{R}^{L\times N_t}$ in the tensor-decomposition matrices

$$D^l:=\left[\left(V^TU\left(\hat{\mu}^1\right)\right)_l^T,\ldots,\left(V^TU\left(\hat{\mu}^{N_r}\right)\right)_l^T\right]=\left[\left(V_l^TU\left(\hat{\mu}^1\right)\right)^T,\ldots,\left(V_l^TU\left(\hat{\mu}^{N_r}\right)\right)^T\right]\in\mathbb{R}^{N_t\times N_r},\tag{24}$$

for $l\in\{1,\ldots,L\}$, where, as notation, A_l declares the l-th column of a matrix A.

5. By singular value decompositions

$$D^l\approx\hat{D}^l=\sum_{k=1}^{d_l}\lambda_k^l\psi_k^l\left(\phi_k^l\right)^T,\tag{25}$$

of the tensor-decomposition matrices, for $l\in\{1,\ldots,L\}$, time and parameter spaces are decomposed. Here, the numbers $d_l\in\mathbb{N}$ denote the corresponding ranks of truncation calculated by the same ratio used in (21) with the singular values λ_k^l of $\hat{D}^l$. The columns $\psi_k^l\in\mathbb{R}^{N_t}$ and $\phi_k^l\in\mathbb{R}^{N_r}$, $k\in\{1,\ldots,d_l\}$, of the left- and right-singular factors correspond to the time and the parameter spaces, respectively.

6. For the predictions, we construct $2\sum_{l=1}^{L}d_l$ regression models to approximate all ψ_k^l,ϕ_k^l for a new set of times and parameters. Practically, we apply Gaussian process regressions to calculate the regression functions

$$\hat{\psi}_k^l\ :\ \mathcal{T}\to\mathbb{R},t\mapsto\hat{\psi}_k^l(t),\quad\text{from the training set}\left\{\left(t_i,\left(\psi_k^l\right)_i\right)\middle|i\in\{1,\ldots,N_t\}\right\},\tag{26}$$

$$\hat{\phi}_k^l\ :\ \mathcal{M}\to\mathbb{R},\mu\mapsto\hat{\phi}_k^l(\mu),\quad\text{from the training set}\left\{\left(\hat{\mu}^j,\left(\phi_k^l\right)_j\right)\middle|j\in\{1,\ldots,N_r\}\right\},\tag{27}$$

for $k\in\{1,\ldots,d_l\}$ and $l\in\{1,\ldots,L\}$. Gaussian process regression—also known as Kriging—interpolates the training data and builds a fast-responding metamodel for a scalar output, see (Santner, 2003). The MATLAB routines `fitrgp` and `predict` serve to create the Gaussian process regressions and to make the predictions.

7. After these preparations, we can approximate new evaluations of u—and consequently U—through computing

$$u(t,\mu)\approx\hat{u}(t,\mu):=V\begin{pmatrix}\sum_{k=1}^{d_1}\lambda_k^1\hat{\psi}_k^1(t)\hat{\phi}_k^1(\mu)\\ \vdots\\ \sum_{k=1}^{d_L}\lambda_k^L\hat{\psi}_k^L(t)\hat{\phi}_k^L(\mu)\end{pmatrix}\in\mathbb{R}^{N_h}\tag{28}$$

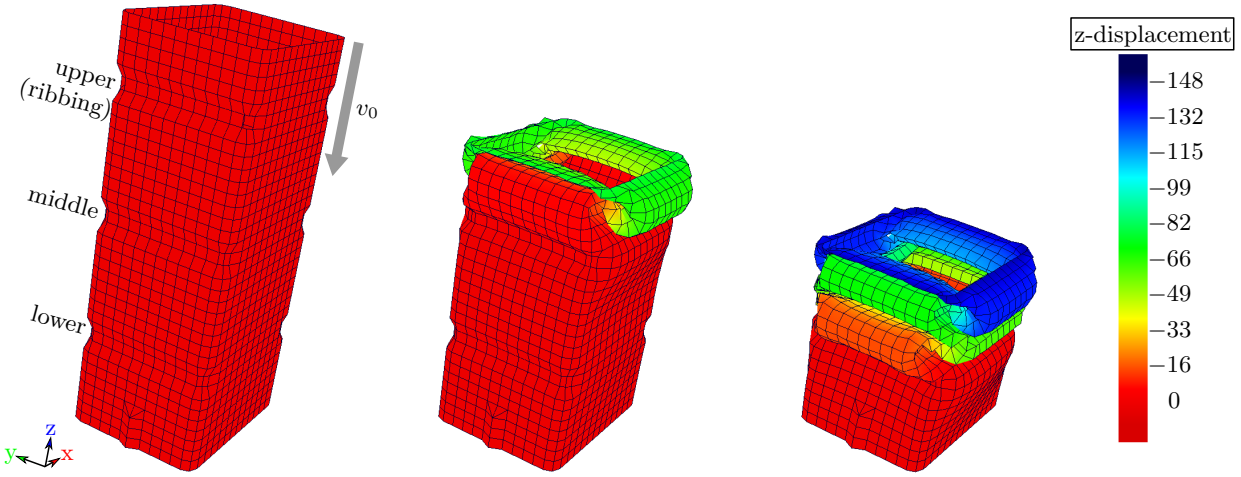
for $(t,\mu)\in\mathcal{T}\times\mathcal{M}$. Multiplying by V in (28) projects the approximated solution into the full space.

Following these seven steps, we are able to reconstruct a FES. The desired key results—further processed scalar outputs—can then be derived from the approximations in a fast manner. For the purpose of ET, this enables to quickly solve the optimizations (9) and (10) which lets the bottleneck disappear. In total $N_r + N_s$ snapshots were needed to run the non-intrusive parametric MOR approach.

3. Application of the MOR-ET Approach

3.1. Finite Element Simulation of a Crashbox Deformation Model

Using the resulting model for the system function (1) of ET leads to an accelerated and thus feasible execution of ET. To investigate this approach, we employ the benchmark problem of a crashbox deformation. This FES is publicly accessible on the LS-DYNA homepage, see (LS-DYNA, 2020), and run in the LS-DYNA crash solver.



(a) Time state 1 (b) Time state 11 (c) Time state 22

Figure 1. Different time states of the LS-DYNA crashbox simulation with input $\mu = (2.7, 8.8, 0.7, -0.6, -1.2)$. The colors indicate the amount of the nodes z-displacement.

In the LS-DYNA crashbox model, a symmetric tube-shaped crashbox with six ribbings is impacted by a plate moving in negative z-direction, see Fig. 1(a)-(c). The crashbox is about 272 mm tall, 116 mm wide, and 96 mm long. Each two ribbings are located opposite one another at the same height. The model consists of $N_{\text{nd}} = 1925$ nodes, i.e. there are $N_h = 3 \cdot N_{\text{nd}} = 5775$ degrees of freedom. The LS-DYNA solver writes $N_t = 22$ time states of the interval $\mathcal{T} = [0, 20]$ in the .d3plot output file.

With the post-processor software (Animator4, 2019), one can view the simulation video as well as convert the .d3plot output file into a .a4db output format that is readable in MATLAB. To load the required N_h x-,y-,z-displacements of the nodes for every time state from the .a4db file into

Table I. Inputs of the LS-DYNA crashbox model.

Input	Interval	Unit	Uncertainty type	Description
μ_1	$[2, 3]$	mm	epistemic	wall thickness
μ_2	$[8.5, 9]$	$\frac{\text{m}}{\text{s}}$	aleatory	initial velocity
μ_3	$[-3, 1]^*$	mm	epistemic	relative upper ribbings depth
μ_4	$[-3, 1]^*$	mm	epistemic	relative middle ribbings depth
μ_5	$[-3, 1]^*$	mm	epistemic	relative lower ribbings depth

* The interval will be trimmed later by the expert, see Fig. 2.

MATLAB, we use the routines `h5info` and `h5read`. Time is measured in milliseconds, displacements in millimeters, the velocity in meters per second, and the mass in kilograms.

We consider $n = 5$ inputs, $\mu = (\mu_1, \mu_2, \mu_3, \mu_4, \mu_5)$, see Table I. The first input dimension characterizes the wall thickness of the crashbox which varies between two and three millimeters, i.e. $\mu_1 \in I_1 = [2, 3]$. The initial velocity of the moving plate lies in the interval $I_2 = [8.5, 9]$ described by the second input μ_2. Entries μ_3, μ_4, and μ_5 represent the depths of the respective opposite upper, middle, and lower ribbings, see Fig. 1, relative to the original configuration. They can be varied within $I_i = [-3, 1], i \in \{3, 4, 5\}$. Negative values of $\mu_i, i \in \{3, 4, 5\}$, indicate that the respective ribbings do not sink as deep as the ribbings of the default crashbox. Positive values have the reverse effect. Based on the uncertain variables $\mu_1, \mu_3, \mu_4, \mu_5$, statements can be made about the epistemic component robustness. The velocity, μ_2, is considered to be an irreducible aleatory uncertainty since it scatters in each experiment.

3.2. Application of the MOR-ET Approach to a Crashbox Simulation

We apply our proposed MOR-ET approach to the maximal z-deformation of the crashbox simulation. This scalar key result is measured at the last time step by one specific node corresponding to the plate—in our FES, this node defines the plate. More precisely, the maximal z-deformation of the crashbox is equal to the final z-displacement of the plate node plus the initial z-distance between the plate and the crashbox of 1.517 mm. Adding the initial distance is correct since the plate is moving in negative z-direction and therefore, all displacements are negative in the example. In this case, the extraction function for the key result is straightforward. Since we consider only one key result for this purpose, conventional metamodels could also be applied. However, we will take this example as a proof of concept.

The maximal z-deformation of the crashbox simulation is taken as the ET system function f. The function f has a black-box character as it is derived by the FES. We follow the six points explained in Section 2.1 to obtain the plausibility and belief curve:

1. & 2. First, a fictive expert assesses the intervals $I_1, \ldots, I_5$ by splitting them into focal elements and providing BPAs, see Fig. 2. Regarding the ribbings, the expert wants the crashbox to deform first above the upper ribbing, second between upper and the middle ribbing, and third at the remaining parts of the crashbox. For that reason, the intervals I_3, I_4, and I_5 are cut.

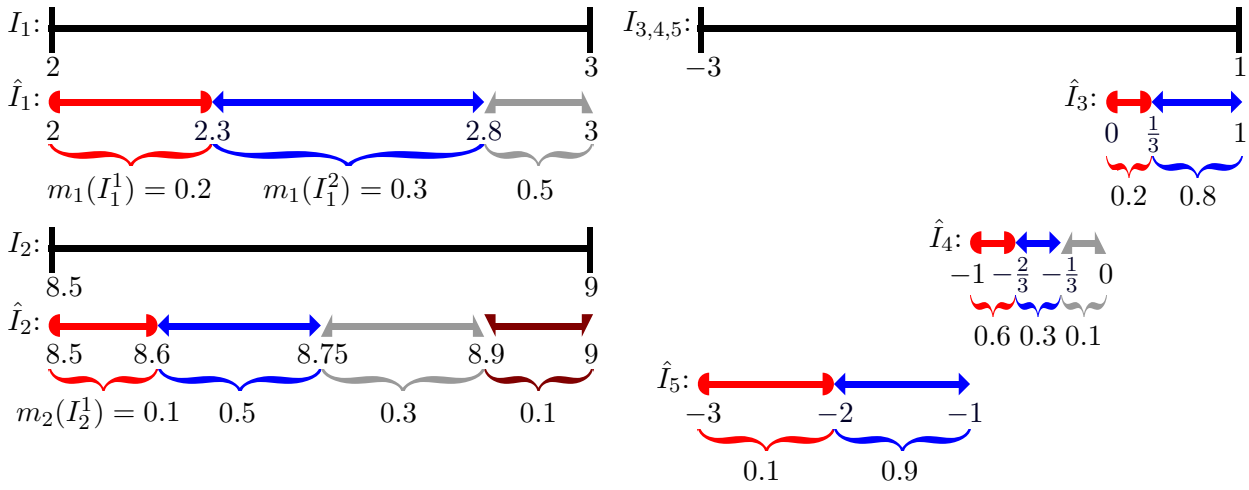


Figure 2. An expert divides the intervals into focal elements and allocates BPAs.

3. & 4. Interval cells are formed and the corresponding composite BPAs are computed, see Fig. 3 for some examples. In total, there are $k = 3 \cdot 4 \cdot 2 \cdot 3 \cdot 2 = 144$ interval cells.

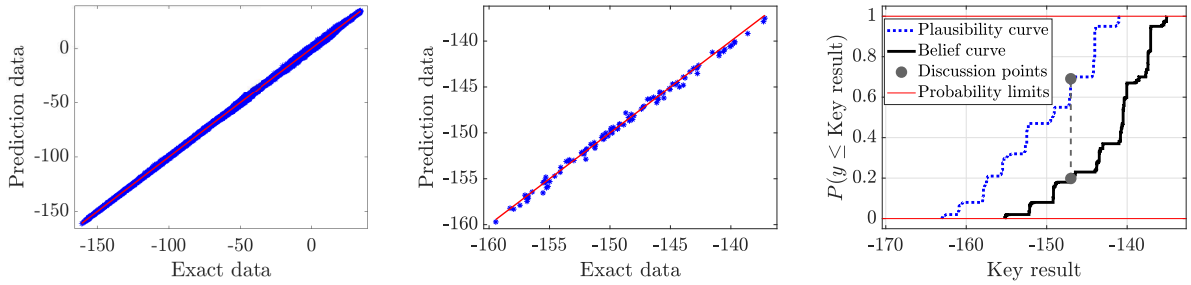
$C^1 = I_1^1 \times I_2^1 \times I_3^1 \times I_4^1 \times I_5^1 =$ ● × ● × ● × ● × ● with $p^1 = 0.2 \cdot 0.1 \cdot 0.2 \cdot 0.6 \cdot 0.1 = 2.4 \cdot 10^{-4}$,
$C^2 = I_1^1 \times I_2^1 \times I_3^1 \times I_4^1 \times I_5^2 =$ ● × ● × ● × ● × ◆ with $p^2 = 0.2 \cdot 0.1 \cdot 0.2 \cdot 0.6 \cdot 0.9 = 0.0022$,
⋮
$C^{144} = I_1^3 \times I_2^4 \times I_3^2 \times I_4^3 \times I_5^2 =$ ▲ × ▼ × ◆ × ▲ × ◆ with $p^{144} = 0.5 \cdot 0.1 \cdot 0.8 \cdot 0.1 \cdot 0.9 = 0.0036$.

Figure 3. Some interval cells and their composite BPAs.

5. To accelerate the optimizations (9) and (10), we approximate the system function f by non-intrusive MOR as introduced in Section 2.2. In total, 200 simulations of the full crash simulation with different input parameters μ are run. We use $N_s = 20$ simulations to set up the snapshot matrix S in (18). After the singular value composition and choosing $\epsilon = 0.9999$, $L = 8$ is calculated by (21). A lower value of ϵ leads to a lower value of L, e.g. $\epsilon = 0.999$ implies $L = 4$, which can yield a lower approximation quality. Another $N_r = 80$ simulations train the tensor-decomposition-based regressions, compare to (26) and (27). The remaining 100 full-scale simulations are taken for validation. Each of the three samples is created in Latin Hypercube style (McKay et al., 1979).

   The MOR approach predicts the N_h degrees of freedoms for all N_t time states with a high approximation quality for the validation sample. Fig. 4(a) shows the regression plot with $N_h \cdot N_t \cdot 100 = 12\,705\,000$ points. A second regression plot, see Fig. 4(b), focuses on the key result of the validation sample, i.e. the maximal z-deformation. The $2k$ optimization problems for the different interval cells are now quickly solved by applying MATLAB's routine `particleswarm` on the MOR surrogate model and its negative counterpart. These took about seven hours wall time on a common office laptop.

6. Now, it is a straightforward task to compile the plausibility and belief curve as well as plot them, see Fig. 4(c).



(a) Regression plot for all degrees of freedom.

(b) Regression plot for the considered key result.

(c) Plausibility and belief curve.

Figure 4. Regression plots for the 100 validation simulations and the resulting plausibility and belief curve.

4. Application Results and Discussion

The first four steps of the ET remain unchanged: Experts specify and assess the uncertainties and—according to their opinions—interval cells are obtained. In general, the more subintervals and the more finely defined these intervals are, the more the belief curve and the plausibility curve converge against each other. In other words, if the experts choose to subdivide the initial intervals in Fig. 2 more detailed, the ET curves in Fig. 4(c) slide closer together. However, the number of optimization problems to be solved would increase strongly.

We focus on the regression process in step five. The reconstruction of the crashbox simulation by non-intrusive MOR is considered sufficiently good. The whole simulation is approximated by a coefficient of determination equal to 0.9998—calculated from the 100 validation simulations. The coefficient of determination for the key result has the value 0.9952. The slightly lower number can be explained by the different ranges of the considered outputs. Note that the regression plot, Fig. 4(b), shows a much smaller range than the plot in Fig. 4(a). This causes deviations to appear larger and the coefficient of determination hence is lower.

To compute the infimum and supremum values of the k interval cells, the MATLAB routine `particleswarm` calls the MOR key result function 345 450 times. If we assume that we require the same number of full-scale simulations to find these values, the LS-DYNA simulation would have needed to be run over 345 thousand times. This would have taken longer than 2100 hours since one simulation has an approximate CPU time of 22 seconds. Instead, we only ran 100 full FES to train the MOR key result function and took another 100 to validate our approximation. Together with the seven hours of solving the optimizations, this adds up to less than nine hours for running the proposed MOR-ET approach.

We now proceed to interpret the results of our uncertainty analysis. Since a deformation of less than one millimeter is not relevant for us, we leave out the decimal places and round to integers.

The smallest point of the plausibility curve is located on the far left at $-163\,\text{mm}$ in Fig. 4(c). The largest point of the belief curve is located on the far right at $-135\,\text{mm}$. These two values define the target set $[-163, -135]$ of the key result due to the uncertain inputs. This means that the key result can take on values within this interval. The courses of the plausibility and belief curve state how likely certain outputs are—with two different measures.

Let us exemplarily consider the key result $-147\,\text{mm}$ on the x-axis in Fig. 4(c), see the marked discussion points. On the one hand, the plausibility curve has the y-value 0.69 at $-147\,\text{mm}$. Equation (14) thus yields the upper bound 0.69 for the cumulative distribution function of the key result at $-147\,\text{mm}$, i.e. it holds

$$P(y \leq -147) \leq \text{Pl}(-147) = 0.69. \tag{29}$$

On the other hand, the y-value of the belief curve is equal to 0.20 at $-147\,\text{mm}$. Equation (13) therefore defines a lower bound for the cumulative distribution function, i.e. one obtains

$$0.20 = \text{Bel}(-147) \leq P(y \leq -147). \tag{30}$$

Together, this limits the value of the cumulative distribution function at $-147\,\text{mm}$ to be between 0.20 and 0.69.

For a clearer understanding, we introduce the terms plausible and believable. Concerning the exemplary data points just mentioned, we define these terms as follows. The plausibility curve says it is 69% plausible that the key result is smaller than or equal to $-147\,\text{mm}$. In contrast, the belief curve states it is 20% believable that the key result is smaller than or equal to $-147\,\text{mm}$. We regard a smaller key result as worse than a larger. Depending on the measure, it is 69% plausible or 20% believable that the highest—and therefore best—crashbox z-deformation will not be larger—and hence not be better—than $-147\,\text{mm}$. In summary, the plausibility curve here is the worst case approximation of the key result cumulative distribution function whereas the belief curve is the best case approximation.

As a practical interpretation, note that the $-272\,\text{mm}$ high crashbox suffers a deformation between $-163\,\text{mm}$ and $-135\,\text{mm}$. Although it is 0% believable that the deformation is less than $-157\,\text{mm}$, it is still 20% plausible that it is less than or equal to $-157\,\text{mm}$, see Fig. 4(c). Therefore, we can not limit the target set by excluding its boundary values. For vehicle safety engineers, this means that the deformation varies by $28\,\text{mm}$ due to the uncertainties which is more than 10% of the total height of the crashbox. Because of this high percentage, we advise engineers to reduce the epistemic uncertainties.

5. Conclusion

Uncertainties are inherent in automotive hardware crash tests. Especially in early phase development, finite element models are used to simulate different crash scenarios. The uncertainties of the hardware tests must be taken into account in the simulations.

The Evidence Theory (ET) was utilized to propagate uncertain inputs through an LS-DYNA finite element crash simulation. As inputs, we chose the wall thickness of the crashbox, the initial

velocity of the moving plate and the depths of three ribbings. The wall thickness and the ribbing depths are epistemic uncertainties that are eliminated in later stages of development. The initial velocity is an aleatory input due to deviations in the test procedure.

To enable the computationally expensive optimization step of the ET, the time-dependent crash simulation was approximated by a non-intrusive parametric Model Order Reduction (MOR) approach. This regression method reconstructs the simulation for new input parameter sets in a fast and efficient way. Using this reduced model, the plausibility and belief curve of ET were computed. Output uncertainties as well as bounds of their distribution were derived by these curves. This information can then be used by vehicle safety engineers to assess the performance of the crashbox.

While the approximation results were satisfactory for the purpose of this study, it remains an open question how the regression quality of the MOR approach compares to conventional metamodels. A comparative study is hence proposed for future research. The MOR approach learns the behavior of the whole simulation and then extracts the key result. The basic physics of the system is therefore integrated in the regression model. In contrast, usual metamodels only process the key result itself, without any knowledge about the underlying system. This may has a positive effect for the approximation quality of the MOR approach. Moreover, if several key results are of interest, they can be extracted from one central model instead of creating different individual metamodels—one for each key result.

We are aware that the example simulation of this study is smaller in scope than present industrial problems. Engineers face total vehicle simulations that include millions of finite elements, complex material behavior, and element failures. We think that our work is an important step in tackling this challenge.

References

Animator4. Version 2.4.1. Braunschweig, Niedersachsen: Gesellschaft für numerische Simulation mbH, 2019.

Auer, E., W. Luther, G. Rebner, and P. Limbourg. A Verified MATLAB Toolbox for the Dempster-Shafer Theory. *Proceedings of the Workshop on the Theory of Belief Functions*, 2010.

Bae, H.-R., R. V. Grandhi, and R. A. Canfield. An Approximation Approach for Uncertainty Quantification Using Evidence Theory. *Reliability Engineering & System Safety*, 86(3):215–225, 2004.

Beyer, H.-G. and B. Sendhoff. Robust Optimization – A Comprehensive Survey. *Computer Methods in Applied Mechanics and Engineering*, 196(33—34):3190–3218, 2007.

Der Kiureghian, A. and O. Ditlevsen. Aleatory or Epistemic? Does It Matter?. *Structural Safety*, 32(2):105–112, 2009.

Eldred, M. S. and L. P. Swiler. Efficient Algorithms for Mixed Aleatory-epistemic Uncertainty Quantification with Application to Radiation-hardened Electronics, Part I : Algorithms and Benchmark Results. Technical Report SAND2009-5805, Sandia National Laboratories, 2009.

Eldred, M. S., L. P. Swiler, and G. Tang. Mixed Aleatory-Epistemic Uncertainty Quantification with Stochastic Expansions and Optimization-based Interval Estimation. *Reliability Engineering & System Safety*, 96(9):1092–1113, 2011.

Freedman, D. and P. Diaconis. On the Histogram as a Density Estimator: L_2 Theory. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 57:453–476, 1981.

Guo, M. and J. S. Hesthaven. Data-driven Reduced Order Modeling for Time-dependent Problems. *Computer Methods in Applied Mechanics and Engineering*, 345:75–99, 2019.

Guo, M. and J. S. Hesthaven. Reduced Order Modeling for Nonlinear Structural Analysis Using Gaussian Process Regression. *Computer Methods in Applied Mechanics and Engineering*, 341:807–826, 2018.

Helton, J. C., J. D. Johnson, W. L. Oberkampf, and C. B. Storlie. A Sampling-Based Computational Strategy for the Representation of Epistemic Uncertainty in Model Predictions with Evidence Theory. Technical Report SAND2006-5557, Sandia National Laboratories, 2006.

Knuth, K. H. Optimal Data-based Binning for Histograms and Histogram-based Probability Density Models. *Digital Signal Processing*, 95:102581, 2019.

Le Guennec, Y., J.-P. Brunet, F.-Z. Daim, M. Chau, and Y. Tourbier. A Parametric and Non-intrusive Reduced Order Model of Car Crash Simulation. *Computer Methods in Applied Mechanics and Engineering*, 338:186–207, 2018.

LS-DYNA. Examples Homepage, https://www.dynaexamples.com. Livermore, California: Livermore Software Technology Corporation, 2020.

LS-DYNA. Version R9.3. Livermore, California: Livermore Software Technology Corporation, 2018.

MATLAB. Version 9.5.0.944444 (R2018b). Natick, Massachusetts: The MathWorks Inc., 2018.

McKay, M. D., R. J. Beckman, and W. J. Conover. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*, 21(2):239—245, 1979.

Nikolaidis, E., D. M. Ghiocel, and S. Singhal. Engineering Design Reliability Handbook. CRC Press, Boca Raton, Florida, 2004.

Oberkampf, W. L. and J. C. Helton. Investigation of Evidence Theory for Engineering Applications. *43rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, 2002.

Oberkampf, W. L., J. C. Helton, and K. Sentz. Mathematical Representation of Uncertainty. *19th AIAA Applied Aerodynamics Conference*, 2001.

Oberkampf, W. L., J. C. Helton, C. A. Joslyn, S. F. Wojtkiewicz, and S. Ferson. Challenge Problems: Uncertainty in System Response Given Uncertain Parameters. *Reliability Engineering & System Safety*, 85(1–3):11–19, 2004.

Santner, T. J., B. J. Williams, and W. I. Notz. The Design and Analysis of Computer Experiments. Vol. 1, Springer, New York City, 2003.

Scott, D. W. On Optimal and Data-based Histograms. *Biometrika*, 66:605-–610, 1979.

Sentz, K. and S. Ferson. Combination of Evidence in Dempster-Shafer Theory. Technical Report SAND2002-0835, Sandia National Laboratories, 2002.

Shafer, G. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, 1976.

Swiler, L. P. and A. A. Giunta. Aleatory and Epistemic Uncertainty Quantification for Engineering Applications. *Proceedings of the Joint Statistical Meetings, American Statistical Association*, 2007.

Swiler, L. P., T. L. Paez, and R. L. Mayes. Epistemic Uncertainty Quantification Tutorial. *Proceedings of the IMAC-XXVII*, 2009.

Swischuk, R., L. Mainini, B. Peherstorfer, and K. Willcox. Projection-based Model Reduction: Formulations for Physics-based Machine Learning. *Computers & Fluids*, 179:704–717, 2019.

Tian, W., P. de Wilde, Z. Li, J. Song, and B. Yin. Uncertainty and Sensitivity Analysis of Energy Assessment for Office Buildings Based on Dempster-Shafer Theory. *Energy Conversion and Management*, 174:705–718, 2018.

Computation of Temporal Uncertainties in Construction Project Management Using Imprecise Probability

Mehdi Modares[1*], Michael Desch[1], Rafi L. Muhanna[2], and Robert L. Mullen[3]

[1]*Department of Civil, Architectural and Environmental Engineering, Illinois Institute of Technology, USA,*
[2]*School of Civil and Environmental Engineering, Georgia Institute of Technology, USA*
[3]*Department of Civil and Environmental Engineering, University of South Carolina, USA*
[*]*Corresponding author: mmodares@iit.edu*

Abstract. In construction project management, scheduling plays a crucial role in a project's success. Construction scheduling and monitoring schemes are used to predict the project's total duration. For this prediction, the Critical Path Method (CPM) and the Program Evaluation and Review Technique (PERT) are widely used. These methods predict the overall project's completion time based on the duration of each task and the allowable float in each task's starting time thru a network of tasks.

However, there exist variations and uncertainties in the duration of each event. As such, deterministic construction scheduling schemes are incapable of quantifying and computing these uncertainties. Although precise probability approaches are utilized for the enumeration and analysis of uncertainties in construction scheduling, their accuracy highly depends on the availability of significant data. Due to limitations of the available data, those precise probability approaches may yield erroneous results.

In this work, a new forward-backward formulation for CPM and PERT is introduced, through which, uncertainties are defined based on the concepts of imprecise probability. To represent the imprecise probability structures, a probability box (P-box) formulation is utilized. Thru the network, in the forward formulation, the uncertainty in the duration of each task is represented by an independent P-box, leading to a determination of the resulting P-box for the project's total time. Conversely, in the backward path, given the P-box for the project's total time, the float time P-box for each task is determined. The forward path calculation is performed using P-box arithmetic; whereas, the backward path calculation is achieved through solving an inverse problem. For illustrating the applicability of this method, an example problem is presented and compared with conventional CPM, PERT, and interval approaches.

Keywords: Imprecise Probability, Construction Scheduling, Critical Path Method

1. Introduction

A network for planning, scheduling, and monitoring construction and industrial projects is based on a set of activities, with prescribed durations, that constructs a directed acyclic graph (Fortin et al. 2010). The Critical Path Method (CPM) is a widely accepted method utilized to determine the characteristics of these networks. Using CPM, the total duration of a project, as well as the project's critical activities, are determined.

© 2021 *by Modares, Desch, Muhanna, and Mullen.*

The numerical calculations performed in conventional CPM are deterministic. However, if there are uncertainties in the duration of activities, the conventional CPM scheme does not quantify or compute their effects. Methods that consider the uncertainties in scheduling are mainly based on probabilistic approaches (Abourizk et al. 1991, 1994; Ahuja et al. 1994). As such, they require sufficient information to choose the Probability Density Function (PDF) for each uncertain duration. For example, in Program Evaluation and Review Technique (PERT), the duration of each activity is defined in terms of the minimum, expected, and maximum durations (Moder et al. 1983; Cottrell 1999). PERT is performed presuming the duration of each activity is related to the Beta distribution. Additionally, possibilistic methods are used to enumerate the uncertainties in scheduling. For example, in the fuzzy approach to scheduling, the durations of activities are defined as fuzzy variables with predefined membership functions (Lessmann et al. 1994; Fetz et al. 1999; Lu and AbouRizk 2000).

The selection of a probability distribution for the duration of each activity needs inter-connected, multi-activity data. Alternatively, if sufficient data to determine these distributions are unavailable, a range of probability distributions that represent the possible activity durations may be considered. Imprecise probability theory is one of the approaches for considering uncertainty in a system without the assumption of a prescribed PDF. Imprecise probability theory is useful when the available information about a system is not sufficient enough such that the PDF of input parameters cannot be accurately chosen. Contrary to fuzzy set theory, using the imprecise probability approach can sharpen the description of its parameters as more information becomes available.

In this work, a method is developed that is capable of treating temporal uncertainties in project scheduling. Imprecise probability theory is used to describe the duration of each activity in a project. As input parameters, the duration of each activity is defined as a Probability-box (P-box). These are sets of cumulative distribution functions (CDF)s, defined by the upper and lower bounding CDFs. For the forward path analysis thru the network, direct P-box arithmetic is used. Conversely, for the backward path thru the network, the float times for each task are calculated based on the solution of an inverse-like problem.

2. Background

This section provides the background information that is foundational to the method developed in this work. This includes overviews of the CPM, PERT, and P-box based Imprecise Probability.

2.1. Critical Path Method

The critical path method is an approach to project scheduling that identifies: 1) the total project duration, 2) a ‘critical path’ of activities that cannot be delayed without increasing the total project duration, 3) the earliest and latest times each activity can be scheduled for without increasing the total project duration (early start/finish and late start/finish), 4) how long each activity can be delayed without affecting other activities (free float), and 5) how long each activity can be delayed without affecting the critical path (total float). To perform this method, first, a list of project tasks is developed, and it is determined which tasks are reliant upon each other. A network of the activities is then developed. Activities are considered nodes and the links between nodes represent the reliance between tasks. The duration of each task is also determined. The following steps are followed to perform the analysis (Hinze 2012):

Forward Pass

- The early start (ES) time of each starting node, ES_S, is set as 0
  $ES_S = 0$
- The early finish (EF) time of each starting node, EF_S, is set as the early start time plus the duration (Dur) of the starting node, Dur_S.
  $EF_S = ES_S + Dur_S$
- The early start time of all other nodes in the network is defined as the maximum value of the early finish time of all directly preceding nodes.
  $ES = \max(EF\ of\ each\ predecessor)$
- The early finish of each node is then defined as the early start plus the duration of the node.
  $EF = ES + Dur$

Backward Pass

- A lag value is calculated for each link in the network. For a link showing node A precedes node B, the lag can be calculated as the early start of node B minus the early finish of node A
  $L_{A \to B} = ES_B - EF_A$
- The free float (FF) of each node is then calculated as the minimum value of the lags of each of the links leaving the node.
  $FF = \min(lag\ of\ each\ link\ leaving\ node)$
- The free float and total float (TF) of the end node is set to 0.
  $FF_E = 0\ \ and\ \ TF_E = 0$
- The total float of all other nodes is defined as the lag to a decedent node plus the free float of the descendent node. If a node has multiple descendants, then the minimum sum is taken.
  $TF = \min(Lag\ to\ descendant + FF\ of\ descendant)$
- The late start (LS) time of a node is then calculated as the early start plus the total float
  $LS = ES + TF$
- The late finish (LF) time of a node is then calculated as the early finish plus the total float
  $LF = EF + TF$

In this approach, the end of day convention is used where all values represent the end of a day. Therefore, the early start time of the starting node is set to zero, $ES_S = 0$. The total project duration time will be the finish day (EF or LF of the end node). The critical path is found as the path from the network start to end with zero total float. At least one such path will exist. If any task on this path is delayed or extended, the project time will increase.

2.2. PROGRAM EVALUATION AND REVIEW TECHNIQUE

Program Evaluation and Review Technique (PERT) is also used for project scheduling. PERT has a similar approach to the CPM; however, using PERT, the duration of each activity is defined in terms of the minimum, expected and maximum durations.

PERT has an underlying probabilistic assumption related to the Beta distribution. The minimum, expected, and maximum durations for each activity are used to calculate the probabilistic parameters for each activity. The mean (μ_{PERT}) and standard deviation (σ_{PERT}) are calculated as:

$$\mu_{PERT} = \frac{\text{minimum} + 4 \times \text{expected} + \text{maximum}}{6} \tag{1}$$

$$\sigma_{PERT} = \frac{\text{maximum} - \text{minimum}}{6} \tag{2}$$

2.3. IMPRECISE PROBABILITY STRUCTURES BASED ON P-BOXES

2.3.1. *Continuously Bounded P-boxes*

Define the CDF for the random variable X as $F(x)$. If the distribution parameters are uncertain, for every x, an interval $\left[\underline{F}(x), \overline{F}(x)\right]$ generally can be constructed that is capable of bounding the possible values of $F(x)$, therefore, $\underline{F}(x) \leq F(x) \leq \overline{F}(x)$. The two CDFs $\underline{F}(x)$ and $\overline{F}(x)$ are known as "probability bounds" (Williamson 1990, Ferson 2003). Figure 1 shows a continuous probability box schematically.

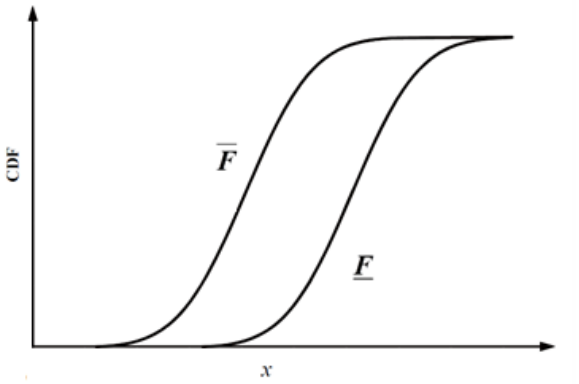


Figure 1. A general continuous probability box

2.3.2. *Discrete Bounded P-boxes*

Aternatively, the bounded P-boxes can be constructed by the use of discrete, interval based P-box structures. This discrete P-box structure includes a collection of interval values, with corresponding associated probability. Figure 2 shows a uniformly discretized P-box that encloses the original continuous P-box schematically.

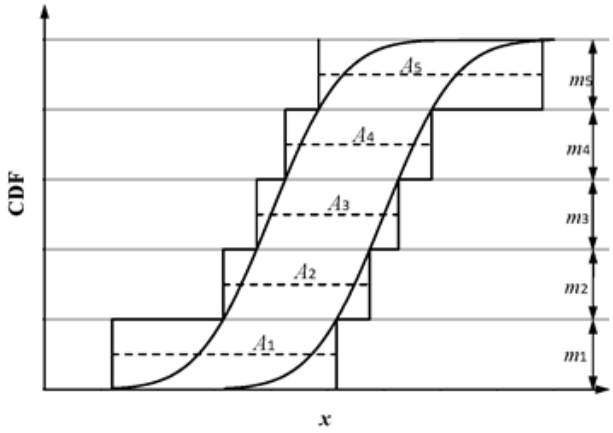


Figure 2. A general uniformly discretized probability box (Enclosure Discretization)

3. Methodology

3.1. Formulation of Imprecise Probability Project Scheduling

The general algorithm for Imprecise Probability Project Scheduling (IPPS), which is based on the CPM, is given below.

1. Determine the imprecise probability structure, based on P-boxes, for each activity. For each activity:

    a) Construct an independent imprecise probability structure for the uncertain duration.
    b) Construct a P-box structure for the uncertain duration.

2. Perform forward pass and backward pass analyses on the network using the constructed P-box structures for the uncertain duration of each activity using either a combinatorial or simulation approach.

    a) In the forward pass, for each activity, determine the P-box structures of 1) the early start time and 2) the early finish time by using P-box addition.
    b) In the backward pass, for each activity, determine the P-box structures of 1) the lag time(s) associated with the activity, 2) the free float time, 3) the total float time, 4) the late start time, and 5) the late finish time.

3. Identify the critical path(s) as the path(s) from the start through the end of the network where every activity has zero total float.

4. Identify the P-box structure for the total project duration.

3.2. IPPS COMPUTATIONS PROCEDURE

Using IPPS, the interval-based computations for P-box structures are performed using interval arithmetic operations for all combinations of intervals corresponding to the P-boxes. For one combination and considering two P-box structures $\boldsymbol{X}$ and $\boldsymbol{Y}$ with intervals $\boldsymbol{X_i} = \left[\underline{x_i}, \overline{x_i}\right]$ and $\boldsymbol{Y_j} = \left[\underline{y_j}, \overline{y_j}\right]$ corresponding to their i and j discretization levels, respectively:

1. In the forward path for addition:

$$\boldsymbol{X_i} + \boldsymbol{Y_j} = \left[\underline{x_i}, \overline{x_i}\right] + \left[\underline{y_j}, \overline{y_j}\right] = \left[\underline{x_i} + \underline{y_j}, \overline{x_i} + \overline{y_j}\right] \tag{3}$$

2. In the backward path for subtraction :

$$\boldsymbol{X_i} - \boldsymbol{Y_j} = \left[\underline{x_i}, \overline{x_i}\right] - \left[\underline{y_j}, \overline{y_j}\right] = \left[\underline{x_i} - \overline{y_j}, \overline{x_i} - \underline{y_j}\right] \tag{4}$$

As the backward path is an inverse problem of the forward path, the input and output interval variables are dependent. As such, the application of Eq. 4 may yield results with overestimation. To reduce this overestimation, the interval variables are defined based on their midpoints and radii, as:

$$\boldsymbol{X_i} = mid(\boldsymbol{X_i}) + rad(\boldsymbol{X_i}) \times \boldsymbol{\varepsilon} \tag{5}$$

$$\boldsymbol{Y_j} = mid(\boldsymbol{Y_j}) + rad(\boldsymbol{Y_j}) \times \boldsymbol{\varepsilon} \tag{6}$$

in which, $\boldsymbol{\varepsilon} = [-1,1]$. Therefore, Eq. 4 can be rewritten as:

$$\boldsymbol{X_i} - \boldsymbol{Y_j} = mid(\boldsymbol{X_i}) - mid(\boldsymbol{Y_j}) + \left(rad(\boldsymbol{X_i}) - rad(\boldsymbol{Y_j})\right) \times \boldsymbol{\varepsilon} \tag{7}$$

in which, the overestimation is reduced because of the single representation of $\boldsymbol{\varepsilon}$ in Eq. 7.

IPPS Summary

In IPPS, the interval solution approach is used to compute the P-box results of the CPM procedure. An interval is chosen from each of the input duration P-boxes, and the network is solved. Eqs. 3 and 7 are used where additions and subtractions are needed. This gives each network output as an interval value. All possible combinations of interval network results are computed and then used to formulate each output as a P-box.

4. Numerical Example

4.1. PROBLEM DEFINITION

The numerical example, based on a model from Chu (2008), is used to illustrate the developed method. This model, a network with seven tasks and three parallel paths, is defined by Table I. Each task has three durations attributed to it, an optimistic (minimum) time to completion, an expected (mean) time to completion, and a pessimistic (maximum) time to completion.

Table I. Network Definition

Task	*Minimum Duration*	*Expected Duration*	*Maximum Duration*	*Preceding Task(s)*
A	7	10	16	-
B	4	5	6	-
C	26	30	50	A
D	35	40	58	B
E	8	12	14	B
F	30	36	48	C and D
G	8	9	12	E

Figure 3 depicts this network schematically.

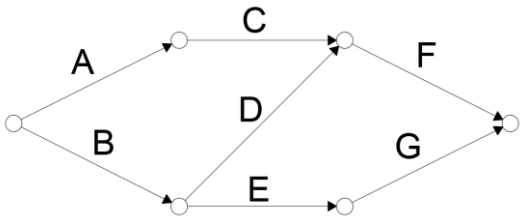


Figure 3. Network Graph, as defined by Chu (2008), with Tasks on Links

4.2. PROBLEM SOLUTION

The network graph in Figure 3 depicts the tasks on the links. As the algorithm used for the developed method places tasks on the nodes, Figure 4 is drawn to depict the same network but with tasks on the nodes. A dummy "END" node with zero duration is added to satisfy the requirement that the network terminates on a single ending node.

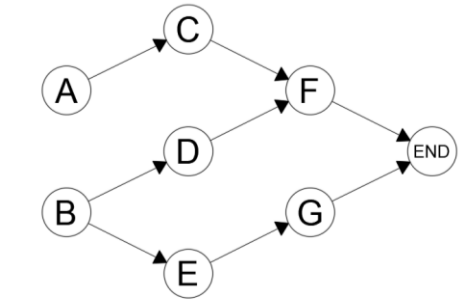


Figure 4. Network Graph – Tasks on Nodes

The network is solved using the following methods:

1. CPM - The duration of each task is the expected duration. $Dur_i = expected_i$
2. PERT - The duration of each task is the PERT mean. $Dur_i = \mu_{PERT_i}$
3. IPPS - The duration of each task is a normal distribution with mean and standard deviation equal to the PERT mean and PERT standard deviation, respectively $Dur_i = N(\mu_{PERT_i}, \sigma_{PERT_i})$.
4. Interval Analysis - The duration of each task is the interval constructed between the task's minimum duration and maximum duration $Dur_i = [min_i, max_i]$.

IPPS Analysis

In the IPPS method, for each task, no variation is considered in either the mean or standard deviation, $\overline{F}(x)$ coincides with $\underline{F}(x)$. The imprecise probability structure of each activity is discretized into ten P-box intervals with equal probability mass. A combinatorial analysis (10^7 combinations) is performed (seven-node network with each node with ten P-box intervals). As the normal CDF function is continuous, the tails are truncated at $P = 0.005, P = 0.995$ in order to establish finite bounds for the lower-most and upper-most P-box intervals. Figure 5 shows the seven P-box interval envelopes for the imprecise probability durations.

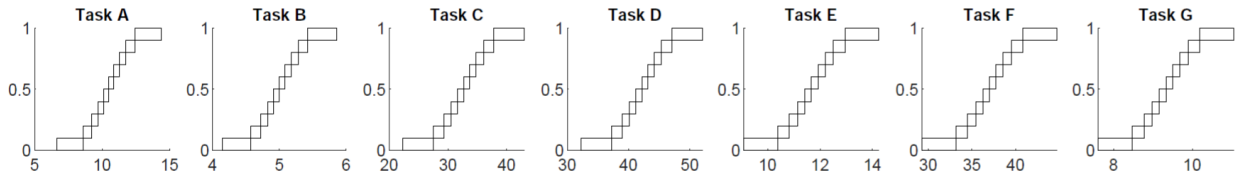


Figure 5. P-box durations

4.3. Analytical Results

The problem is solved using the four analysis schemes and the detailed results are presented in Table II. The results displayed for the IPPS is the interval corresponding to a probability $P = 0.5$.

TABLE II. Results for each task, comparing CPM, PERT, IPPS, and Interval approaches

Task	*Approach*	*Duration*	*Early Start*	*Early Finish*	*Free Float*	*Total Float*	*Late Start*	*Late Finish*
A	CPM	10	0	10	0	5	5	15
	PERT	10.50	0.00	10.50	0.00	4.00	4.00	14.50
	IPPS	[10.50, 10.50]	[0.00, 0.00]	[10.50, 10.50]	[0.00, 0.00]	[3.67, 4.33]	[3.67, 4.33]	[13.83, 15.17]
	Interval	[7, 16]	[0, 0]	[7, 16]	[0, 0]	[0, 6]	[0, 6]	[7, 22]
B	CPM	5	0	5	0	0	0	5
	PERT	5.00	0.00	5.00	0.00	0.00	0.00	5.00
	IPPS	[5.00, 5.00]	[0.00, 0.00]	[5.00, 5.00]	[0.00, 0.00]	[0.00, 0.00]	[0.00, 0.00]	[5.00, 5.00]
	Interval	[4, 6]	[0, 0]	[4, 6]	[0, 0]	[0, 0]	[0, 0]	[4, 6]
C	CPM	30	10	40	5	5	15	45
	PERT	32.67	10.50	43.17	4.00	4.00	14.50	47.17
	IPPS	[32.67, 32.67]	[10.50, 10.50]	[42.33, 44.00]	[3.67, 4.33]	[3.67, 4.33]	[13.83, 15.17]	[46.31, 49.76]
	Interval	[26, 50]	[7, 16]	[33, 66]	[0, 6]	[0, 6]	[7, 22]	[33, 72]
D	CPM	40	5	45	0	0	5	45
	PERT	42.17	5.00	47.17	0.00	0.00	5.00	47.17
	IPPS	[42.17, 42.17]	[5.00, 5.00]	[46.62, 47.71]	[0.00, 0.00]	[0.00, 0.00]	[5.08, 5.17]	[47.17, 49.00]
	Interval	[35, 58]	[4, 6]	[39, 64]	[0, 2]	[0, 2]	[4, 8]	[39, 66]
E	CPM	12	5	17	0	55	60	72
	PERT	11.67	5.00	16.67	0.00	58.17	63.17	74.83
	IPPS	[11.67, 11.67]	[5.00, 5.00]	[16.49, 16.84]	[0.00, 0.00]	[58.02, 60.09]	[62.93, 65.18]	[74.36, 77.08]
	Interval	[8, 14]	[4, 6]	[12, 20]	[0, 0]	[49, 82]	[53, 88]	[61, 102]
F	CPM	36	45	81	0	0	45	81
	PERT	37.00	47.17	84.17	0.00	0.00	47.17	84.17
	IPPS	[37.00, 37.00]	[47.19, 48.75]	[83.54, 86.60]	[0.00, 0.00]	[0.00, 0.00]	[47.19, 48.75]	[83.54, 86.60]
	Interval	[30, 48]	[39, 66]	[69, 114]	[0, 0]	[0, 0]	[39, 66]	[69, 114]
G	CPM	9	17	26	55	55	72	81
	PERT	9.33	16.67	26.00	58.17	58.17	74.83	84.17
	IPPS	[9.33, 9.33]	[16.49, 16.84]	[25.61, 26.39]	[58.02, 60.09]	[58.02, 60.09]	[74.36, 77.08]	[83.53, 86.60]
	Interval	[8, 12]	[12, 20]	[20, 32]	[49, 82]	[49, 82]	[61, 102]	[69, 114]
END	CPM	0	81	81	0	0	81	81
	PERT	0.00	84.17	84.17	0.00	0.00	84.17	84.17
	IPPS	[0.00, 0.00]	[83.54, 86.60]	[83.54, 86.60]	[0.00, 0.00]	[0.00, 0.00]	[83.54, 86.60]	[83.54, 86.60]
	Interval	[0, 0]	[69,114]	[69, 114]	[0, 0]	[0, 0]	[69, 114]	[69, 114]

Figure 6 shows the P-box output for project duration calculated as the Early Finish of task F.

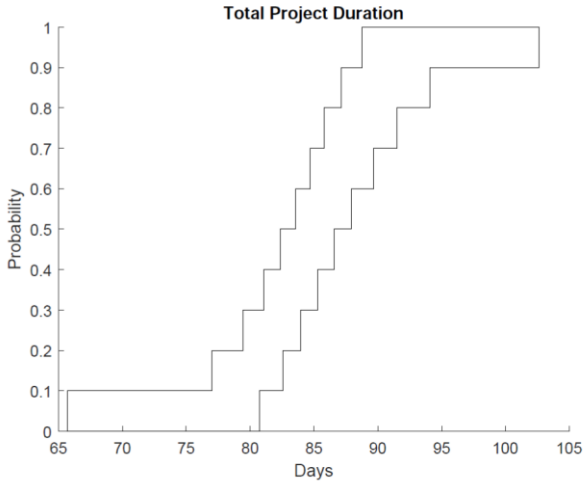


Figure 6. P-box output for project total duration (early finish of activity F)

4.4. OBSERVATIONS

For the CPM, PERT, and IPPS, path B-D-F is immediately identified at the critical path. In the interval approach, it can be seen that node D has the possibility of some total float. This indicates that with no assumption on the distribution of the uncertainty, there is sufficient uncertainty in the network tasks to have multiple possible critical paths.

It can be seen that in every case in Table 2 the interval results enclose the IPPS results. As well, the IPPS results enclose the deterministic PERT results. This outcome is expected as imprecise probabilities make fewer presumptions than deterministic values, and intervals make no presumptions within their bounds at all. Differences between the interval bounds and the first/last p-box bounds can be observed (as seen in Figure 6). This can be accounted for as the effect of the difference between the interval bounds and the truncation points of the imprecise probability CDFs. The truncations at $P = 0.005, 0.995$ correspond to ± 2.576 standard deviations, while the interval solution corresponds to ± 3 standard deviations.

4.5. COMPUTATION TIME

The network results are obtained using MATLAB R2019a running on a processor clocked at 1.70GHz with 8Gb of memory. The computation time for CPM was 0.02 sec, for PERT was 0.02 sec, for IPPS was 918.25 sec, and for the interval method was 0.03 sec. The significantly longer computation time for IPPS is attributed to the combinatorial solution. Using a simulation approach for IPPS will significantly reduce the computational effort.

5. Summary and Conclusions

In this work, a new formulation for construction project scheduling for a network with temporal uncertainty based on the concepts of imprecise probability (called IPPS) is introduced. IPPS can obtain the P-box structures for the early start, early finish, free float, total float, late start, and late finish for each activity as well as the critical path and the overall project duration. The results obtained from IPPS have a higher level of confidence and robustness because of the objective determination of uncertainties in the parameter distributions. Although using the combinatorial approach in IPPS method yields exact results, this approach can be extremely computationally expensive. Therefore, for larger networks, using a simulation approach in IPPS is more desirable. The IPPS capability to more robustly enumerate uncertainties makes it attractive to introduce imprecise probability concepts in the field of construction project scheduling and management.

References

Abourizk, S.M., D. W. Halpin, and J. R. Wilson. "Visual interactive fitting of beta distributions", Journal of Construction Engineering and Management, 117, pp. 589-605. (1991).

Abourizk, S.M., D. W. Halpin, and J. R. Wilson, "Fitting beta distributions based on sample data", Journal of Construction Engineering and Management, 120, pp. 288-305. (1994).

Ahuja, H.N., S. P. Dozzi, and S. M. Abourizk, "Project management: techniques in planning and controlling construction projects", John Wiley, (1994).

Chu, C. "Analysis of effect factors-based stochastic network planning model," Journal of Systems Engineering and Electronics, 19(1), pp. 112-118, (2008), DOI: 10.1016/S1004-4132(08)60054-7.

Cottrell, W. D. "Simplified program evaluation and review technique (PERT)", Journal of Construction Engineering and Management 125, pp. 16-22 (1999).

Ferson, S., V. Kreinovick, L. Ginzburg, and F. Sentz. "Constructing probability boxes and Dempster-Shafer structures" (2003).

Fetz, T., M. Oberguggenberger, J. Jager, D. Koll, G. Krenn, H. Lessmann, and R.F. Stark. "Fuzzy Models in Geotechnical Engineering and Construction Management," Computer-Aided Civil and Infrastructure Engineering, 14 pp. 93–106. (1999).

Fortin, J., P. Zielinski, D. Dubois, and H. Fargier. "Criticality analysis of activity networks under interval uncertainty" J. Sched, 13 pp. 609–627. (2010), DOI: 10.1007/s10951-010-0163-3.

Hinze, J.W. *Construction Planning and Scheduling*. 4th Ed. Prentice Hall, Upper Saddle River, NJ. (2012).

Lessmann, H., J. Muhlogger, and M. Oberguggenberger. "Netzplantechnik mit unscharfen Methoden (Network techniques with fuzzy methods)" Bauingenieur, 69 pp. 469-478. (1994).

Lu, M. and S. M. Abourizk, "Simplified CPM/PERT simulation model", Journal of Construction Engineering and Management 126, 219-226 (2000).

Moder, J.J., C. R. Phillips, and E. W. Davis. "Project management with CPM, PERT and precedence diagramming" Van Nostrand Reinhold, (1983).

Williamson, R. and T. Downs. "Probabilistic arithmetic: numerical methods for calculating convolutions and dependency bounds", International Journal of Approximate Reasoning, 4 pp. 89-158 (1990).

AUTHOR INDEX

A
Ahrens, A., 435
Alderucci, T., 552
Alibrandi, U., 164
Andersen, L. V., 164
Auer, E., 435

B
Beer, M., 39, 140, 178, 380, 408
Behrendt, M., 39
Betancourt, D.,11, 102
Boccamazzo, A., 495
Böttcher, M., 449
Broggi, M., 140, 380

C
Caddemi, S., 575
Caliò, I., 575
Callens, R. R. P., 53
Cannizzaro, F., 575
Carboni, B., 495
Cavaleri, L., 510
Chiaia, B., 350
Ciano, M., 195
Cocuzza Avellino, G., 575
Comerford, L., 39

D
Daub, M., 365
De Angelis, M., 27, 39
De Domenico, D., 524
De Gregorio, D., 209
Desch, M.,1, 642
Di Matteo, A., 538
Di Paola, M., 510
Dong, B., 476
Dunn, I. P., 538

E
Edler, P., 68

F
Faes, M. G. R., 53, 140, 187, 380
Failla, G., 538
Falsone, G., 339
Fedele, F., 86
Ferrotto, M.,F., 510
Ferson, S., 27, 265, 395
Franci, A., 225
Freitag, S., 68, 102

G
Gao, W., 476
Genovese, F., 552, 566
Gerdts, M., 627
Gioffrè, M., 195
Giovanis, D. G., 615
Giunta, F., 250
Graf, W., 449, 604
Gray, A., 27, 395
Gray, N., 265
Grigoriu, M., 195

H
Hanss, M., 295
Hose, D., 295
Huang, Y. Q., 486

I
Iacuaniello, M., 209
Impollonia, N., 575

J
Jang, B., 310
Jehle, J. S., 627
Jensen, H. A., 178
Jerez, D. J., 178

K
Kaliske, M., 449, 604
Karakašević, J., 279
Kosheleva, O., 408
Kreinovich, V., 265, 408, 604

L
Lacarbonara, W., 495
Lahmer, T., 118
Lange, V. A., 627
Laudani, R., 339
Le Carrer, N., 588
Luo, D. H., 486

M
Marasco, G., 350
Marelli, S., 365
Meschke, G., 68
Metsis, P., 127
Misraji, M., 178
Modares, M., 1, 642
Moens, D., 53
Mohammadi, J., 238, 310
Montanino, A., 209, 225
Muhanna, R. L., 11, 86, 102, 642
Mukunoki, D., 419
Mullen, R. L., 642
Muscolino, G., 250, 552, 566

O
Oberguggenberger, M., 279
Ogita, T., 419
Ozaki, K., 419

P
Palmeri, A., 566
Patelli, E., 27
Pellecchia, D., 328
Peters, S., 68
Phoon, K.-K., 380
Pirrotta, A., 538
Proppe, C., 154

Q
Quaranta, G., 495

R
Ricciardi, G., 524
Rosati, L., 328

S
Saleh, E., 238
Santoro, R., 465
Savvas, D., 127
Schietzold, F. N., 604
Schmidt, A., 118
Sessa, S., 328
Shahi, S., 86
Shields, M. D., 615
Sofi, A., 250, 465
Spanos, P. D., 140
Stefanou, G., 127
Sucato, V., 510
Sudret, B., 365

T
Tubino, F., 465

U
Urenda, J., 604

V
Vaiana, N., 328
Valdebenito, M. A., 178, 187
Ventura, G., 350

W
Wang, Q., 476

Y
Yu, Y., 476

Z
Zhang, Y., 39
Zio, E., 164
Zuccaro, G., 209, 225