REC2021 9TH INTERNATIONAL WORKSHOP ON RELIABLE ENGINEERING COMPUTING
Risk and Uncertainty in Engineering Computations
Virtual Conference: May 17-20, 2021

MUTUAL INFORMATION FOR GLOBAL SENSITIVITY ANALYSIS AND ADAPTIVE LEARNING SURROGATE MODELLING

AARHUS UNIVERSITY DEPARTMENT OF ENGINEERING — UMBERTO ALIBRANDI — *Aarhus University*

AARHUS UNIVERSITY DEPARTMENT OF ENGINEERING — LARS V. ANDERSEN — *Aarhus University*

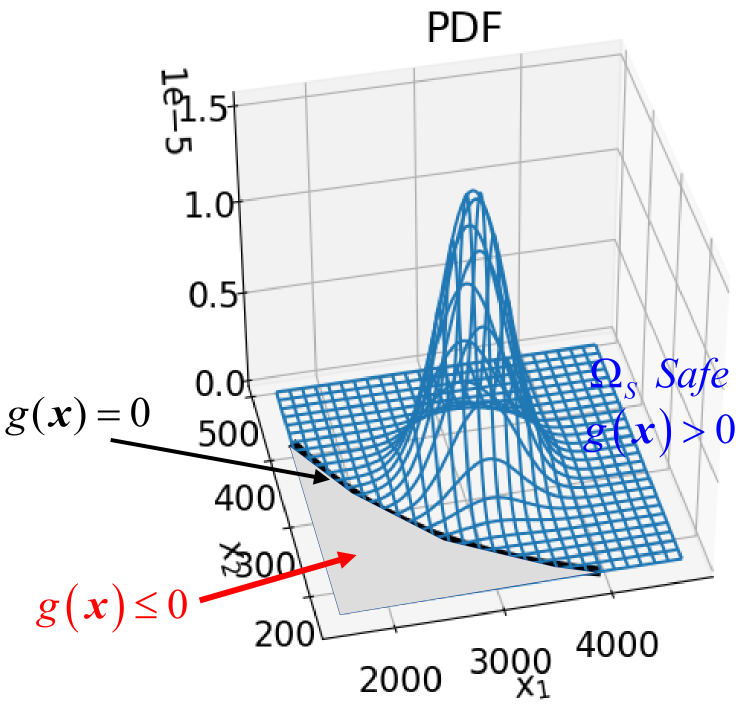
MINES ParisTech, POLITECNICO MILANO 1863 — ENRICO ZIO — *MINES ParisTech, Politecnico di Milano*

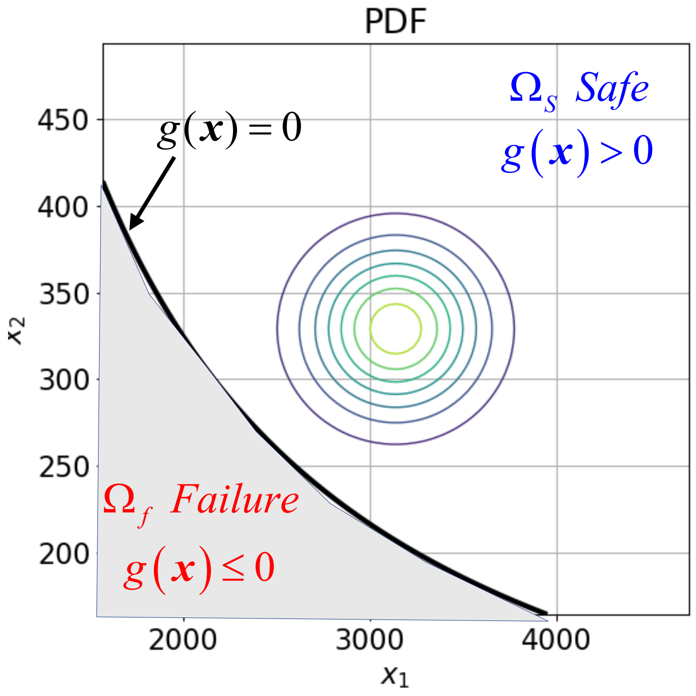
STRUCTURAL RELIABILITY

- **Introduction**
- Informational Correlation
- Sensitivity Analysis
- AK Entropy
- ALMI
- Conclusions

Limit State Function: $G(\boldsymbol{x}) = R(\boldsymbol{x}) - S(\boldsymbol{x})$

The space of the basic variables is divided into two distinct domains, the *safe set* and the *failure set*, separated by the *Limit State Surface*.





$$P_f = P[G(\boldsymbol{x}) \leq 0] = \int_{G(\boldsymbol{x})\leq 0} f_X(\boldsymbol{x})d\boldsymbol{x}$$

STUCTURAL RELIABILITY

$$P_f = P[G(\boldsymbol{x}) \leq 0] = \int_{G(\boldsymbol{x})\leq 0} \boxed{f_X(\boldsymbol{x})} d\boldsymbol{x}$$

1) Limit State Function $G(\boldsymbol{x})$

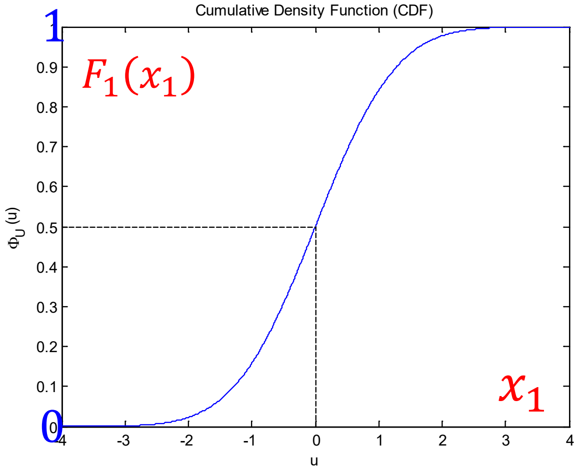
2) Probability distributions of the basic random variables $f_X(\boldsymbol{x})$:

How to model the dependence between random variables?

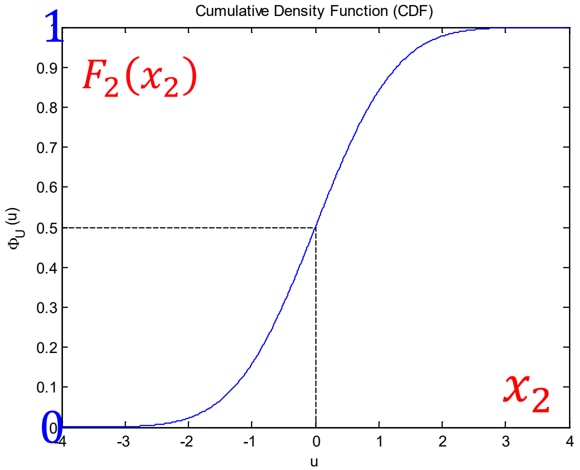
COPULA

- Introduction
- **Informational Correlation**
- Sensitivity Analysis
- AK Entropy
- ALMI
- Conclusions

Correlation $\rho(x_1, x_2)$ is not a proper measure of dependence.



$$V_1 = F_1(x_1)$$

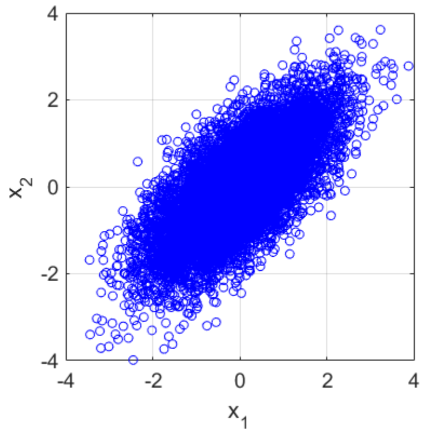


$$V_2 = F_2(x_2)$$

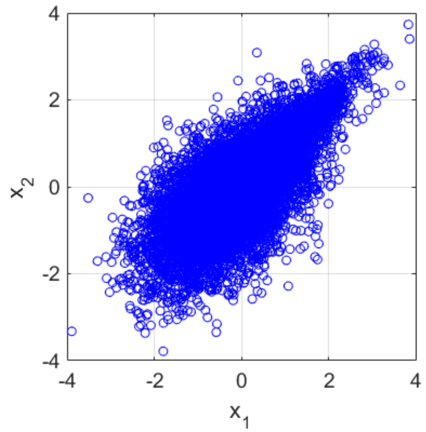
Copula $c(x_1, x_2)$ is a proper measure of dependence.

$$C(v_1, v_2) = Prob\left[\bigcap (V_k \leq v_k)\right] \equiv F(x_1, x_2)$$

$$f(x_1, x_2) = \underbrace{c(v_1, v_2; \vartheta)}_{\text{Copula Density}} \underbrace{f_1(x_1) f_2(x_2)}_{\text{Marginal Distributions}}$$



Gaussian Copula



Gumbel Copula

MUTUAL INFORMATION

- Introduction
- **Informational Correlation**
- Sensitivity Analysis
- AK Entropy
- ALMI
- Conclusions

It measures how much knowing one of these variables reduces uncertainty about the other

Mutual Information

$$MI(X_1,X_2) = \iint f(x_1,x_2)\, log\left(\frac{f(x_1,x_2)}{f_1(x_1)f_2(x_2)}\right) dx_1 dx_2 \geq 0$$

Copula Density

$$f(x_1,x_2) = \boxed{c(v_1,v_2)}\, f_1(x_1)f_2(x_2)$$

$$I(X_1,X_2) = E[\log c(x_1,x_2)]$$

The mutual information takes into account *the whole probabilistic dependence of* ***X***

X_1, X_2 Independent $\rightarrow$ $MI(X_1,X_2) = 0$

U. Alibrandi & K.M. Mosalam, 2020

AARHUS UNIVERSITY
DEPARTMENT OF ENGINEERING

INFORMATIONAL COEFFICIENT OF CORRELATION

Introduction

Informational Correlation

Sensitivity Analysis

AK Entropy

ALMI

Conclusions

Coefficient of Correlation

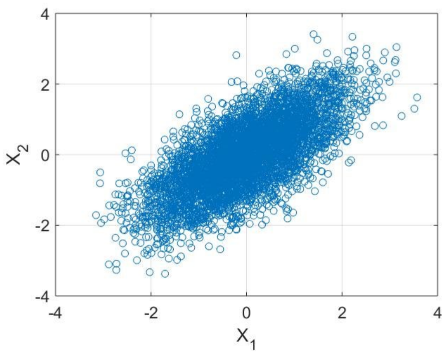
$$\rho(X_1, X_2) = \frac{1}{\sigma_1 \sigma_2} \int_0^1 \int_0^1 [C(v_1, v_2) - v_1 v_2] dF_1^{-1} dF_2^{-1}$$

It depends from copula and marginals

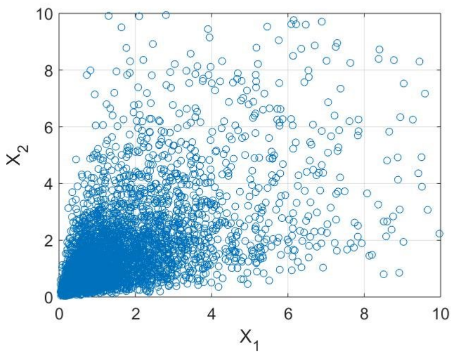
Informational Coefficient of Correlation

$$\rho_{MI}(X_1, X_2) = \sqrt{1 - e^{-2MI(X_1, X_2)}}$$

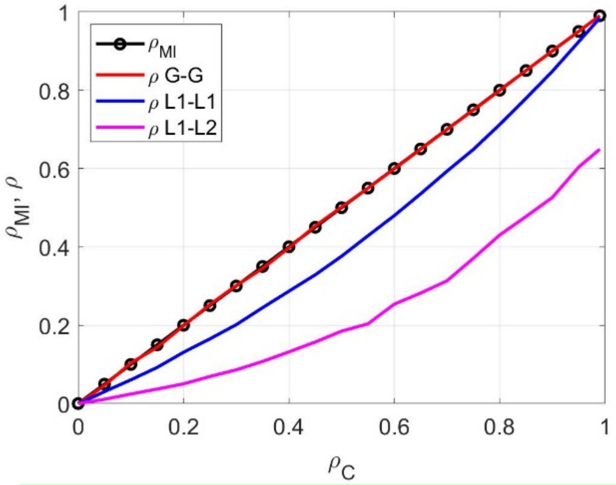
It depends from copula **only**



Normal Marginals



Lognormal Marginals



The correlation underestimates the dependence

INFORMATIONAL COEFFICIENT OF CORRELATION

Coefficient of Correlation

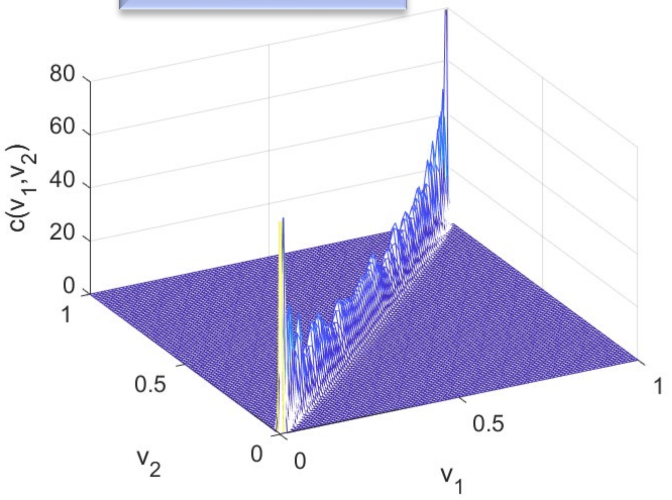
$$\rho(g, u) = 1$$

Informational Coefficient of Correlation

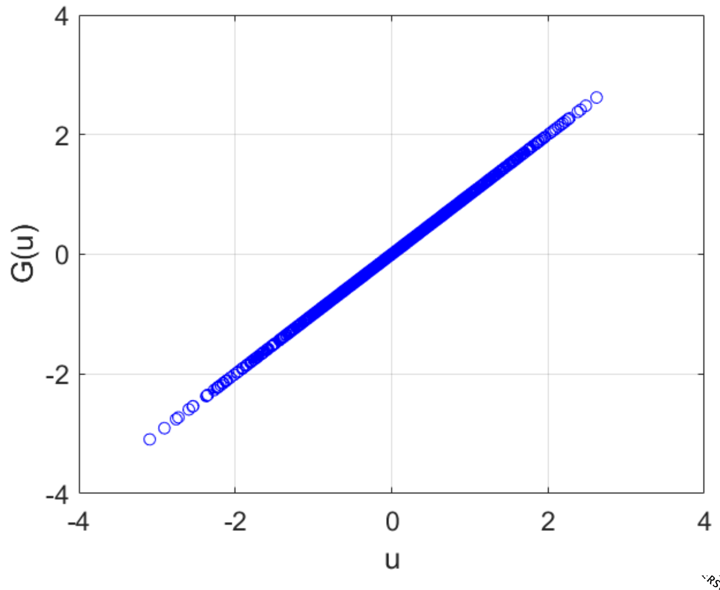
$$\rho_{MI}(g, u) = 1$$

$$g(u) = u$$

Copula



If $y = g(u)$ then (y, u) are fully correlated



INFORMATIONAL COEFFICIENT OF CORRELATION

Coefficient of Correlation

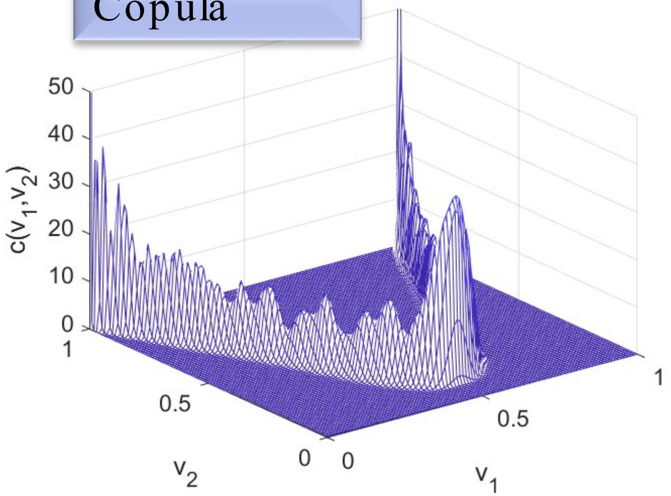
$$\rho(g,u) = 0$$

Informational Coefficient of Correlation

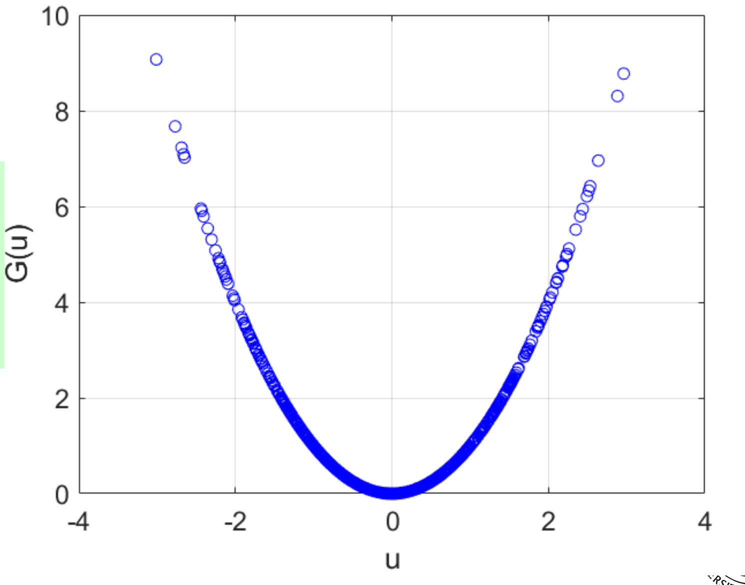
$$\rho_{MI}(g,u) = 1$$

$$g(u) = u^2$$

Copula



If $y = g(u)$ then (y, u) are fully correlated



INFORMATIONAL SENSITIVITY ANALYSIS

- Introduction
- Informational Correlation
- **Sensitivity Analysis**
- AK Entropy
- ALMI
- Conclusions

$$P_f = P[G(\boldsymbol{x}) \leq 0] = \int_{G(\boldsymbol{x})\leq 0} f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x}$$

How to rank the most important variables to be included in the reliability problem?

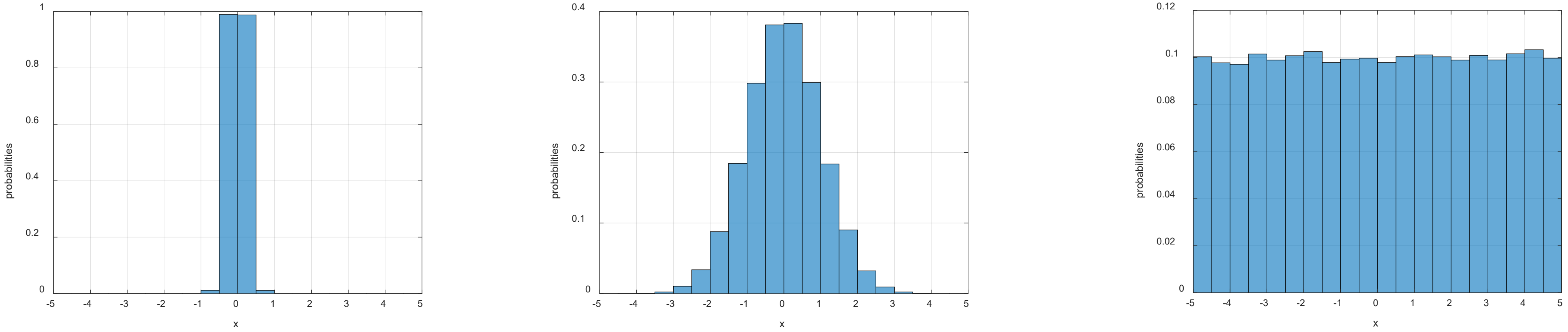
Our approach: **Probabilistic Sensitivity Analysis**

ENTROPY AS A MEASURE OF UNCERTAINT

The Shannon's entropy functional is its expected information content

$$H(X) = -\int f(x) \log f(x)\, dx = -E[\log f(x)]$$

The **entropy is a measure of uncertainty**



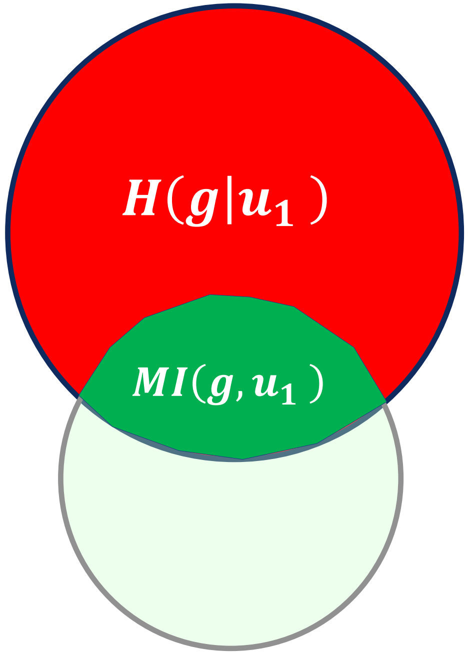
$H = 0.75$ $H = 2.12$ $H = 3.00$

VALUE OF INFORMATION

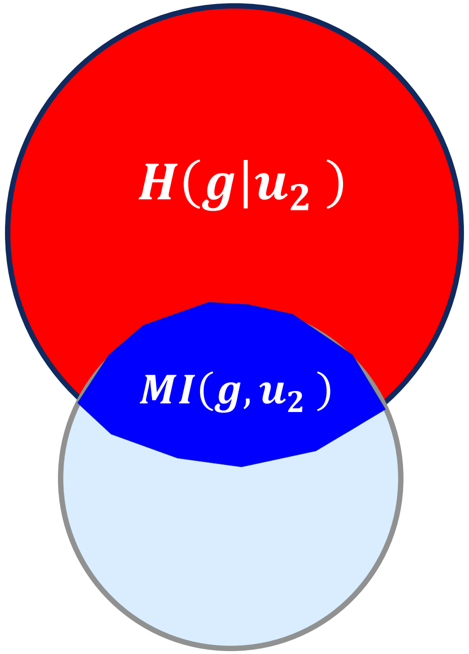
$$g = g(u_1, u_2)$$

$$VoI(u_1) = H(g) - H(g|u_1) \equiv MI(g, u_1)$$

$$VoI(u_2) = H(g) - H(g|u_2) \equiv MI(g, u_2)$$



The $\mathrm{VoI}(u_k)$ is the entropy difference of $\Delta H(g)$ between observing and not observing u_k



Informational Sensitivity Indices

$$\delta_{VoI}(u_1) = \rho_{MI}(g, u_1)$$

$$\delta_{VoI}(u_2) = \rho_{MI}(g, u_2)$$

INFORMATIONAL SENSITIVITY INDICES EXAMPLE

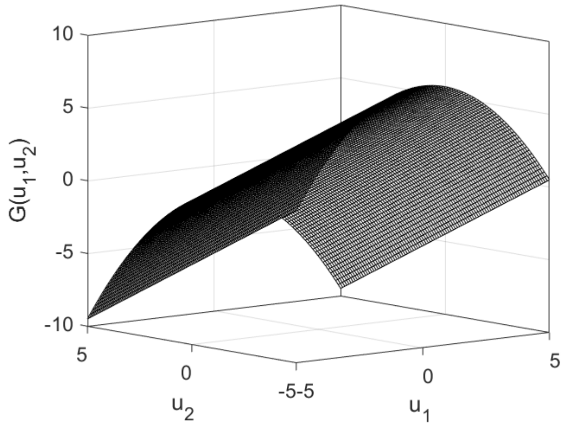
$\rho(g, u_1) = 0.00$

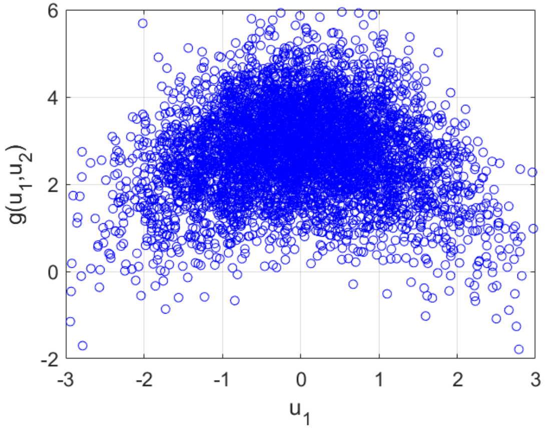
$\delta_{VoI}(u_1) = \rho_{MI}(g, u_1) = 0.41$

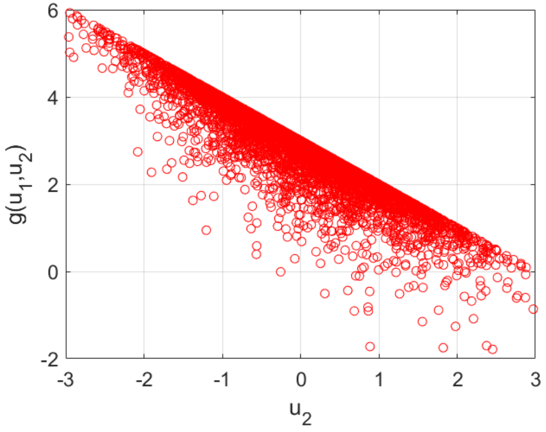
$\rho(g, u_2) = -0.92$

$\delta_{VoI}(u_2) = \rho_{MI}(g, u_2) = 0.96$

$g(u_1, u_2) = 3 - u_2 - 0.3u_1^2$







STRUCTURAL RELIABILITY

$$P_f = P[G(\boldsymbol{x}) \leq 0] = \int_{\boxed{G(\boldsymbol{x})\leq 0}} f_X(\boldsymbol{x})d\boldsymbol{x}$$

1) Limit State Function $G(\boldsymbol{x})$

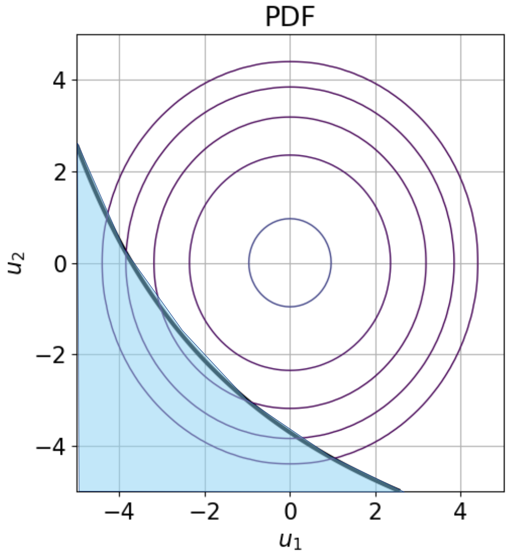
2) Probability distributions of the basic random variables $f_X(\boldsymbol{x})$

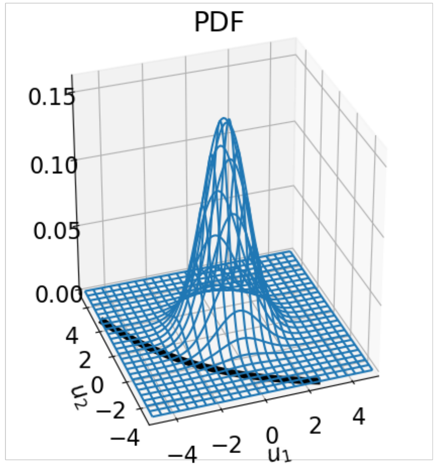
RESPONSE SURFACE METHOD

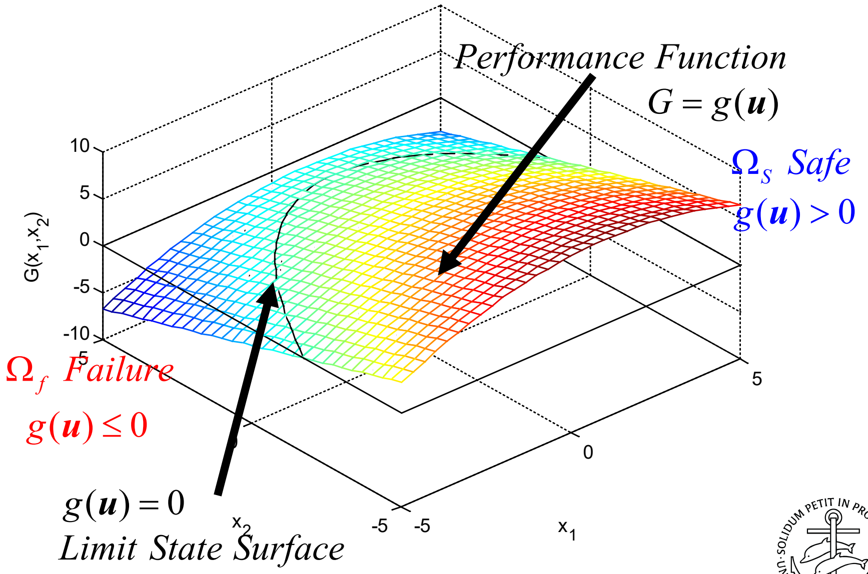
In the *Response Surface Method (RSM)* the *Limit State Function(LSF)* is replaced by a simple explicit function, called *Response Surface (RS)* whose values can be easily evaluated.

Once that the RS is determined, it is no longer necessary to develop further Finite Element analyses.

$$P_f = P[g(\boldsymbol{u}) \leq 0] = \int_{g(\boldsymbol{u})\leq 0} \varphi_n(\boldsymbol{u})d\boldsymbol{u}$$







Normal standard space

ENTROPY-BASED ADAPTIVE KRIGING: AK-E

Mean prediction

$$g_{AK}(\boldsymbol{u}) = \mu(\boldsymbol{u}) + \varepsilon(\boldsymbol{u})$$

Zero-mean Gaussian Random field

$$= h(\boldsymbol{u}) \cdot \boldsymbol{w} + \varepsilon(\boldsymbol{u})$$

Covariance function

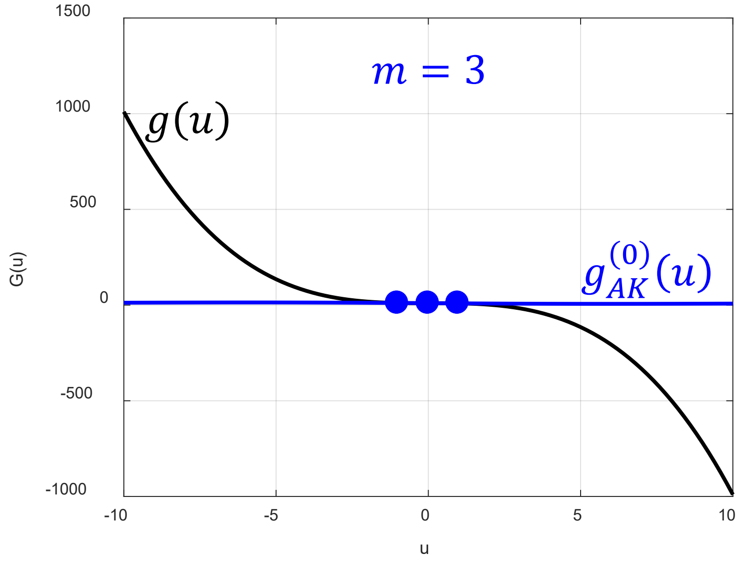
Correlation function

$$Cov_{\varepsilon}(\boldsymbol{u}, \boldsymbol{u}') = \sigma_{\varepsilon}^{2} R_{\varepsilon}(\boldsymbol{u}, \boldsymbol{u}'; \boldsymbol{\vartheta}_{\varepsilon})$$

$$= \sigma_{\varepsilon}^{2} exp\left(-\sum_{k=1}^{n} \vartheta_{\varepsilon,k} |u_k - u_k'|^2\right)$$

Square exponential model

$$g(u) = 10 - u^3$$

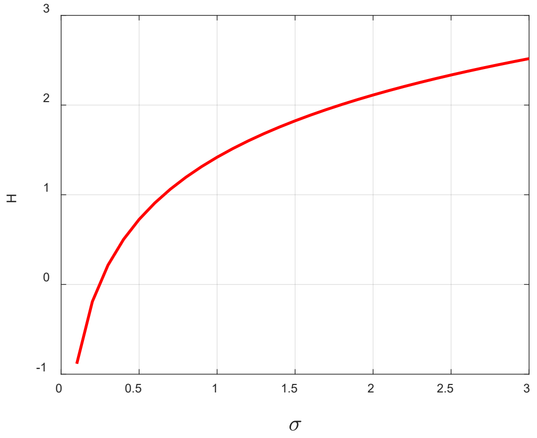


How to choose the new points of the sampling plan?

AK-H: LEARNING FUNCTION FOR SURROGATE MODELLING

$$g_{AK}(\boldsymbol{u}) = \mu(\boldsymbol{u}) + \varepsilon(\boldsymbol{u})$$

$$\mathcal{H}(\boldsymbol{u}) = \frac{1}{2}(1 + \log 2\pi) + \log \sigma(\boldsymbol{u})$$

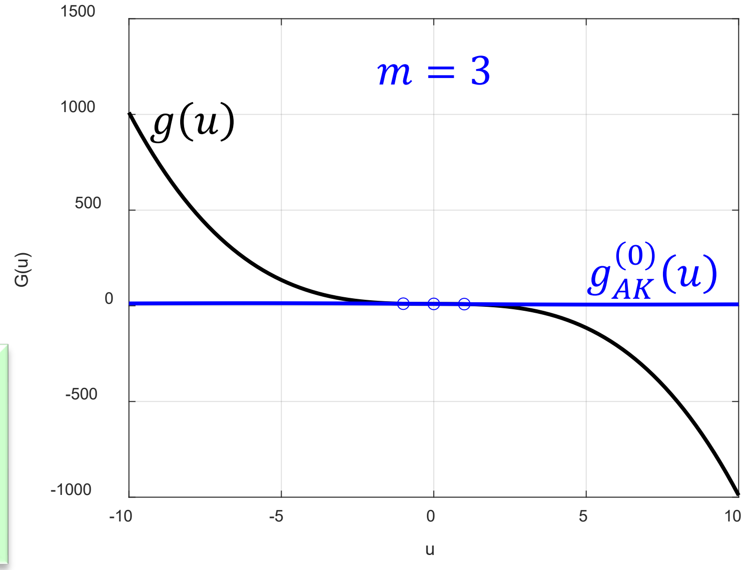


$$\max_{u} \mathcal{H}(\boldsymbol{u})$$

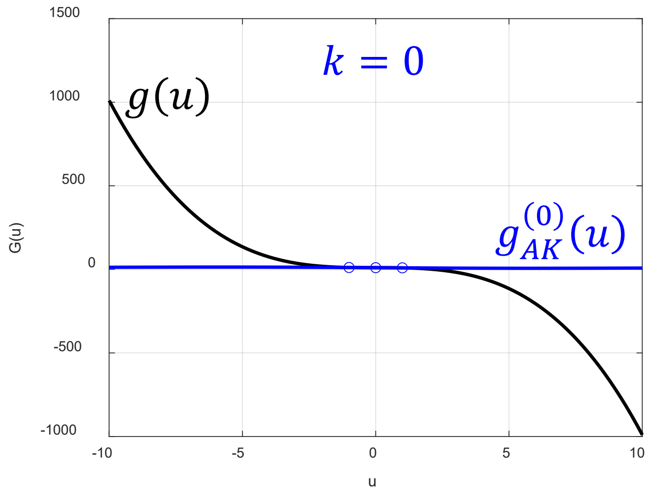
High epistemic uncertainty

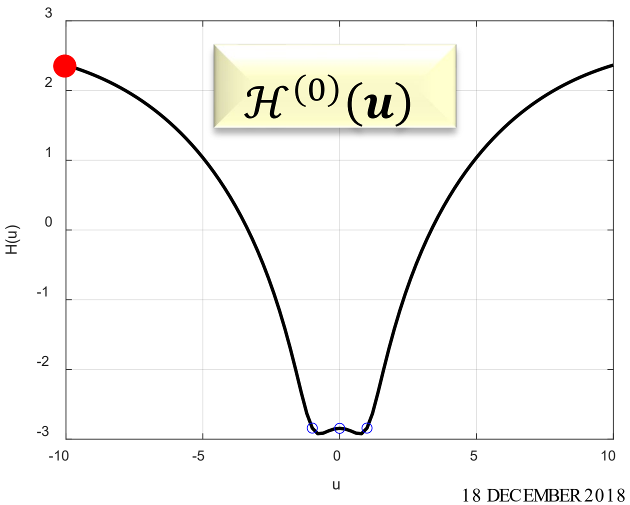
Surrogate modelling

$$\mathcal{H}(\boldsymbol{u}) = H_g(\boldsymbol{u})$$

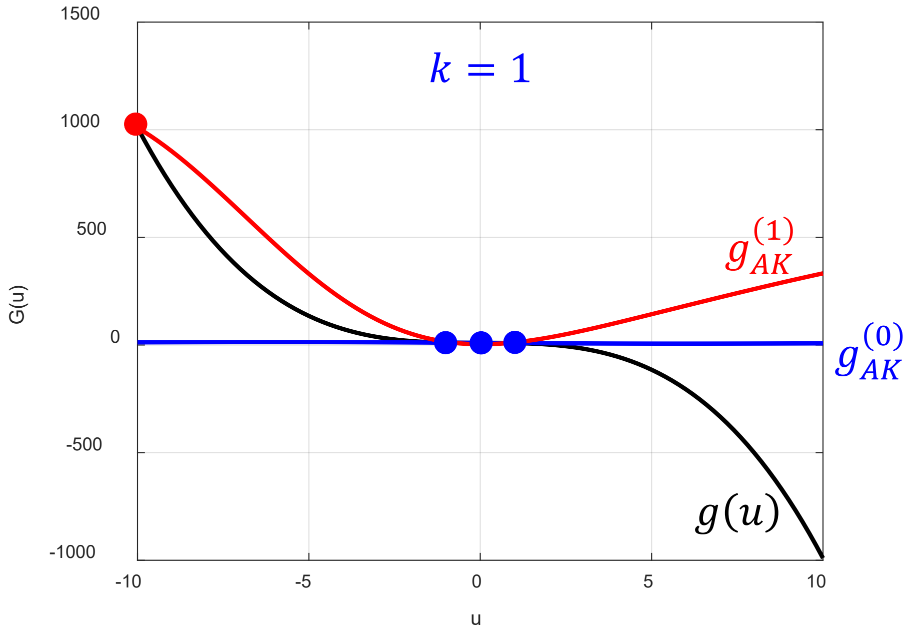


ADAPTIVE KRIGING AKH: ITERATION 1





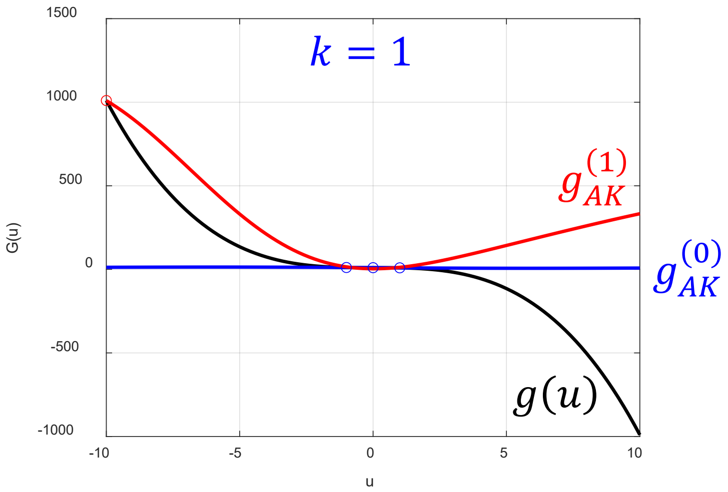
$$g(u) = 10 - u^3$$

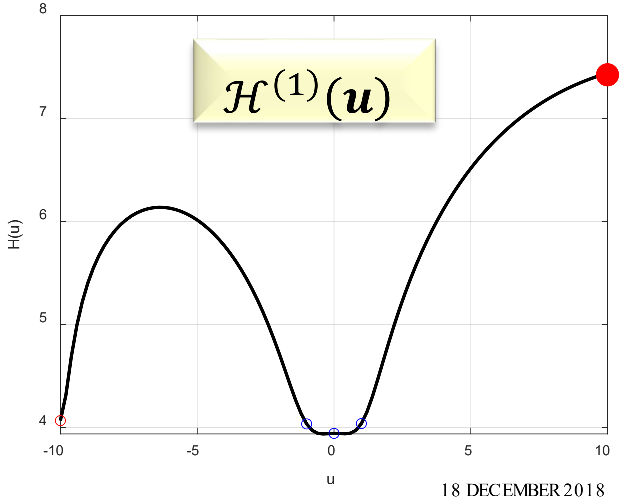


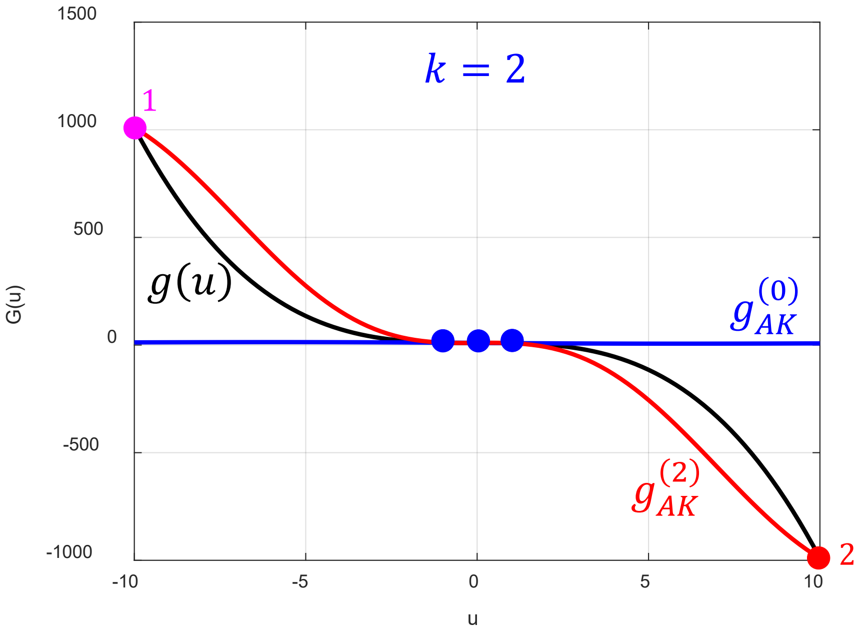
$$\boldsymbol{u}^{(1)} = \min_{u} \mathcal{H}^{(0)}(\boldsymbol{u})$$

ADAPTIVE KRIGING AKH: ITERATION 2

$$g(u) = 10 - u^3$$



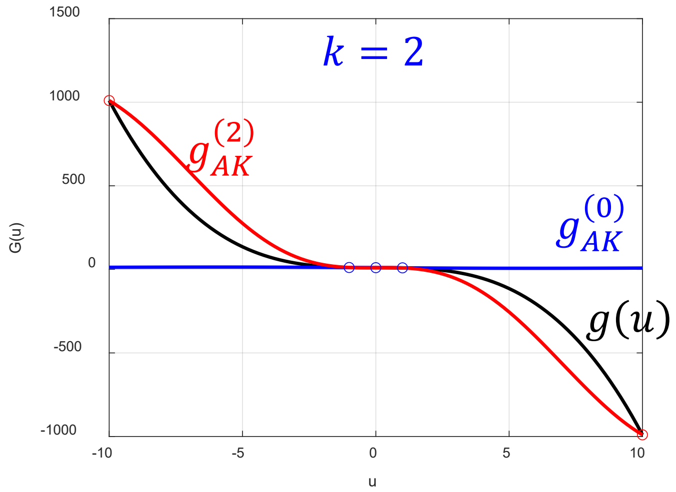


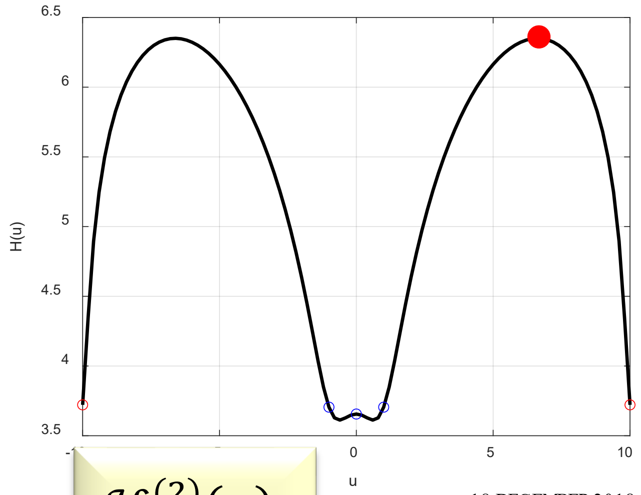


$$\boldsymbol{u}^{(2)} = \min_{u} \mathcal{H}^{(1)}(\boldsymbol{u})$$

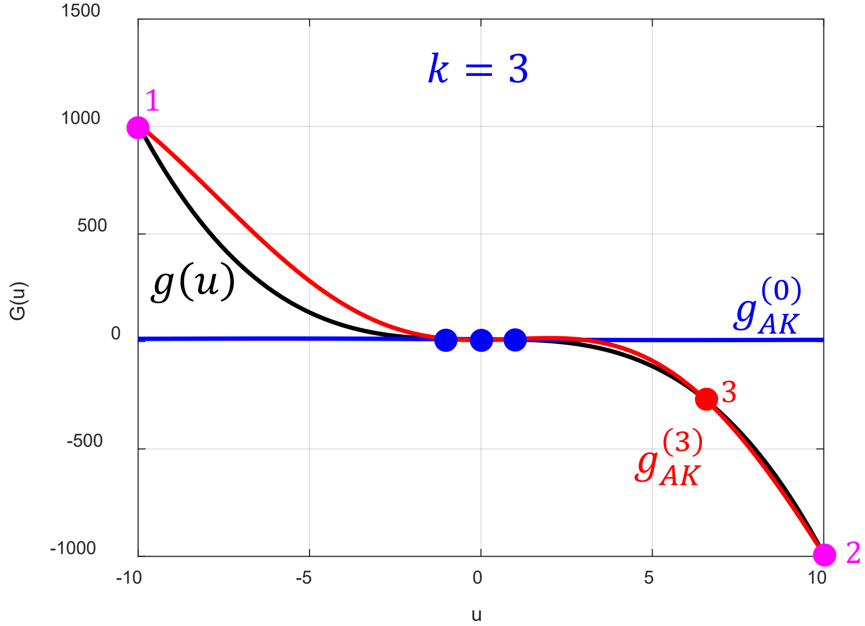
ADAPTIVE KRIGING AKH: ITERATION 3

$$g(u) = 10 - u^3$$



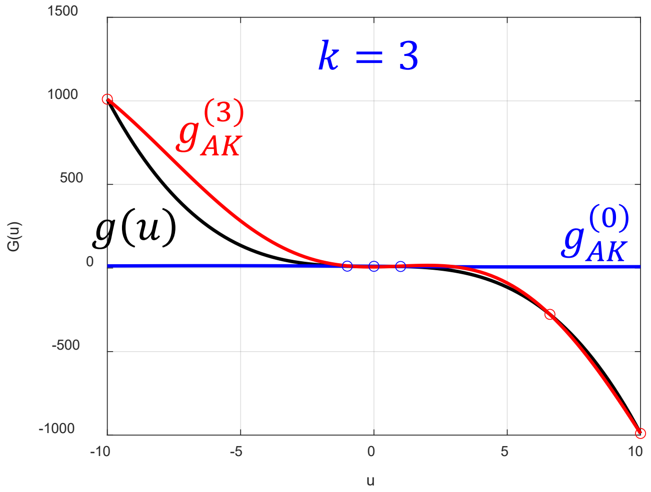


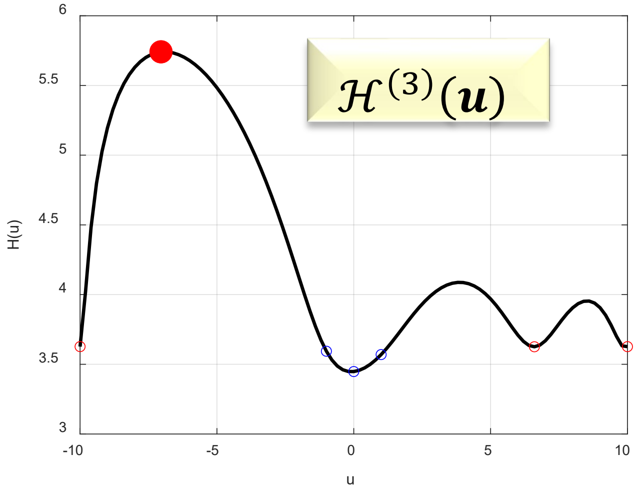
$$\mathcal{H}^{(2)}(\boldsymbol{u})$$



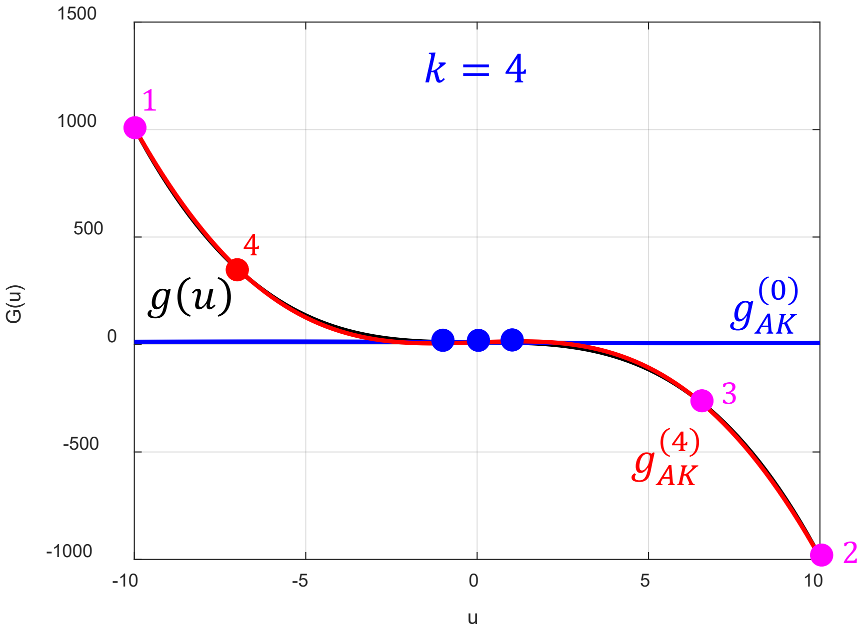
$$\boldsymbol{u}^{(3)} = \min_{u} \mathcal{H}^{(2)}(\boldsymbol{u})$$

ADAPTIVE KRIGING AKH: ITERATION 4





$$g(u) = 10 - u^3$$



$$\boldsymbol{u}^{(4)} = \min_{u} \mathcal{H}^{(3)}(\boldsymbol{u})$$

ACTIVE LEARNING (AL) OR ADAPTIVE KRIGING (AK)?

Mean prediction

$$g_{AK}(\boldsymbol{u}) = \mu(\boldsymbol{u}) + \varepsilon(\boldsymbol{u})$$

Zero-mean Gaussian Stationary process

$$= h(\boldsymbol{u}) \cdot \boldsymbol{w} + \varepsilon(\boldsymbol{u})$$

What about different surrogate models?

Support Vector Machines?
Polynomial Chaos Expansion?
Neural Networks?
……

ACTIVE LEARNING MUTUAL INFORMATION: AL-MI

- Introduction
- Informational Correlation
- Sensitivity Analysis
- AK Entropy
- **AL MI**
- Conclusions

Mean prediction

$$g_{AK}(\boldsymbol{u}) = \mu_g(\boldsymbol{u}) + \varepsilon(\boldsymbol{u})$$

Zero-mean Gaussian Random Field

$$= \boldsymbol{h}(\boldsymbol{u}) \cdot \boldsymbol{w} + \varepsilon(\boldsymbol{u})$$

$$g(u_1, u_2) = 3 - u_2 - 0.3u_1^2$$

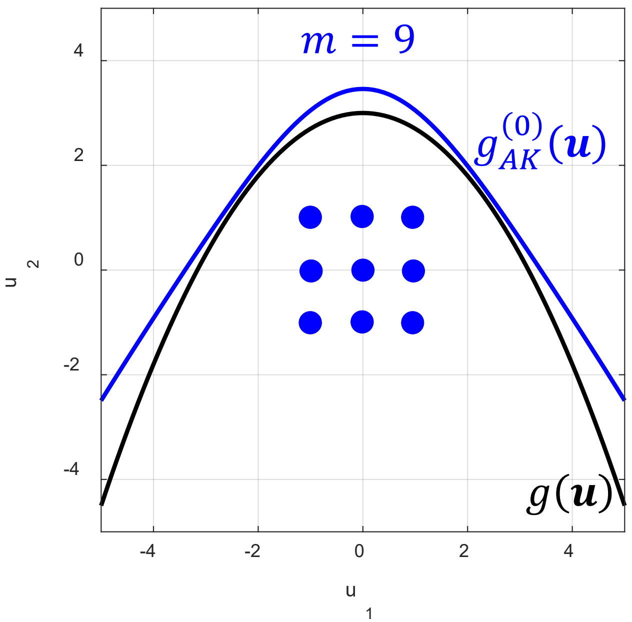
Covariance function

Correlation function

$$Cov_\varepsilon(\boldsymbol{u}, \boldsymbol{u}') = \sigma_\varepsilon^2 R_\varepsilon(\boldsymbol{u}, \boldsymbol{u}'; \boldsymbol{\vartheta}_\varepsilon)$$

$$= \sigma_\varepsilon^2 exp\left(-\sum_{k=1}^{n} \vartheta_{\varepsilon,k} |u_k - u_k'|^2\right)$$

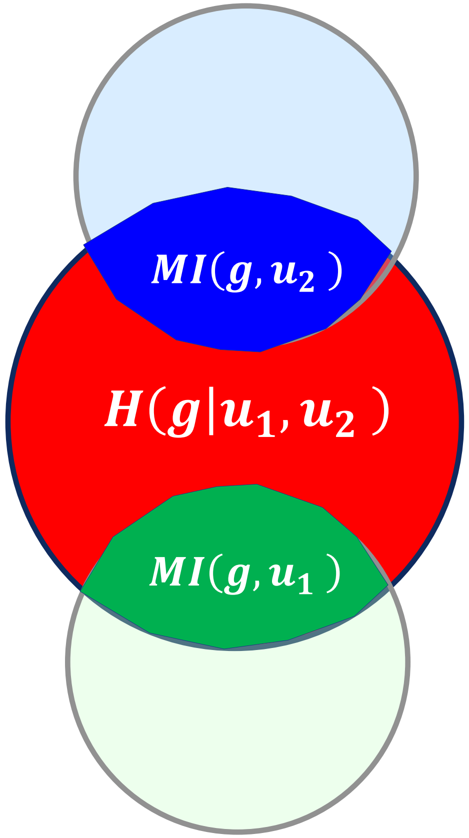
Square exponential model



How to choose the new points of the sampling plan?

AL-MI: LEARNING FUNCTION

$$\max_u H(g|u_1,u_2) = \min_u \sum_{k=1} MI(g,u_k)$$

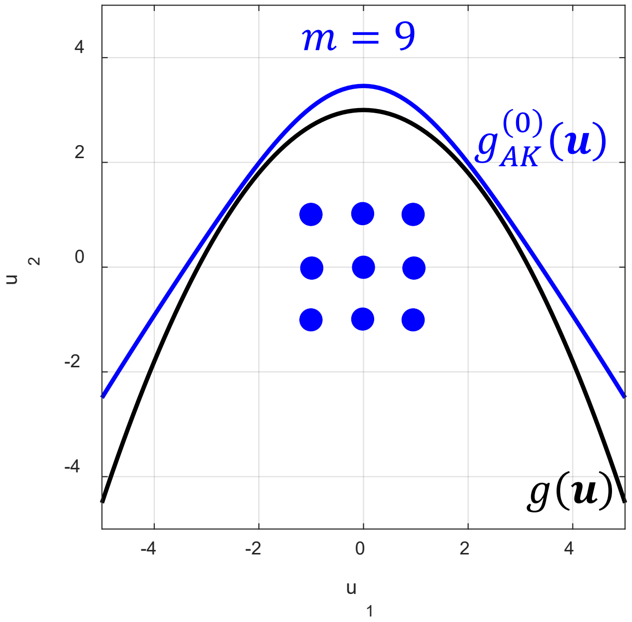


- close to the limit state
- High epistemic uncertainty

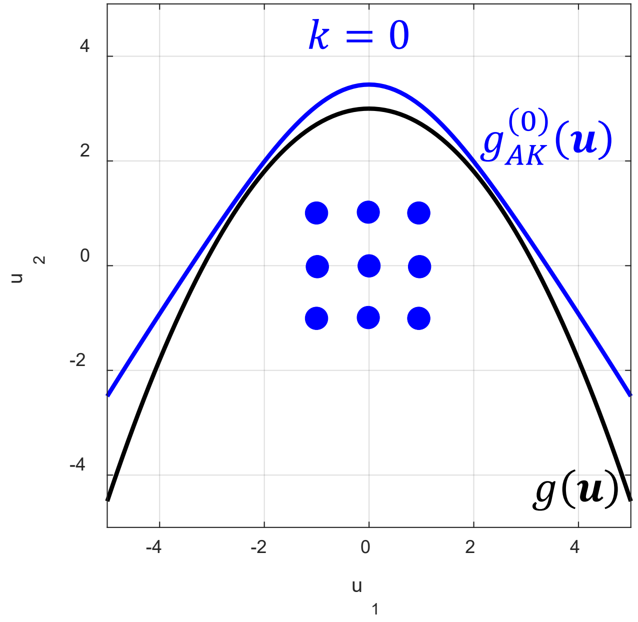
$$\min_u \mathcal{M}I(\boldsymbol{u})$$

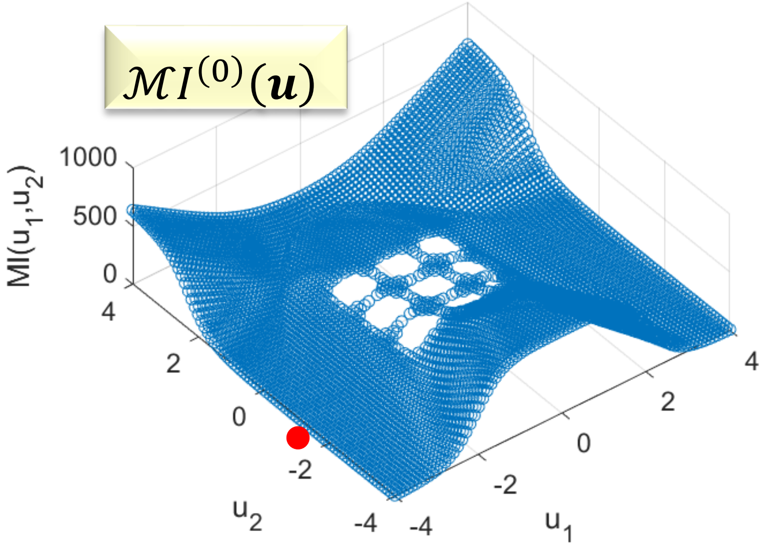
Learning function

$$\mathcal{M}I(\boldsymbol{u}) = \frac{|\mu_g(\boldsymbol{u})|}{1-\rho_{MI}(\boldsymbol{u})}$$

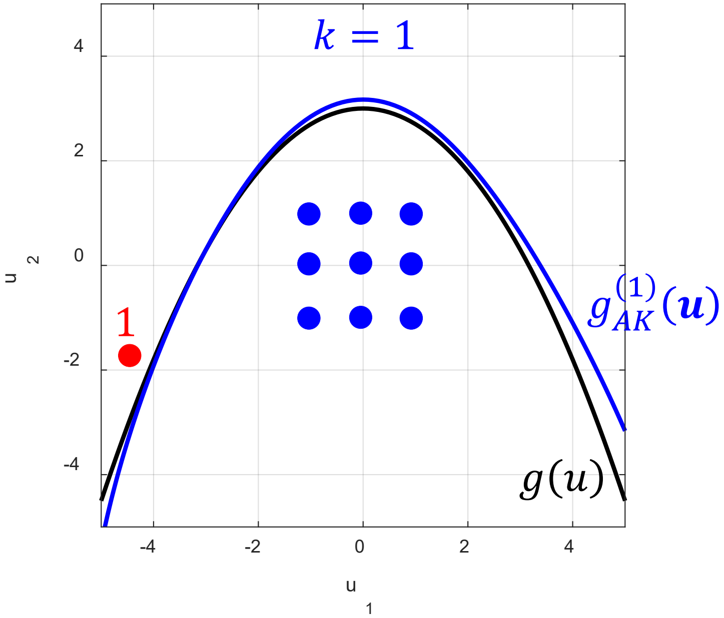


ACTIVE LEARNING AL-MI: ITERATION 1





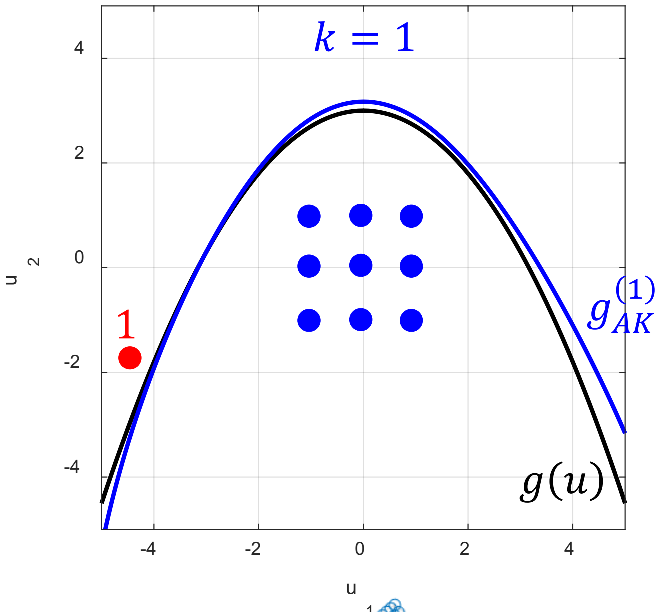
$$g(u_1, u_2) = 3 - u_2 - 0.3u_1^2$$

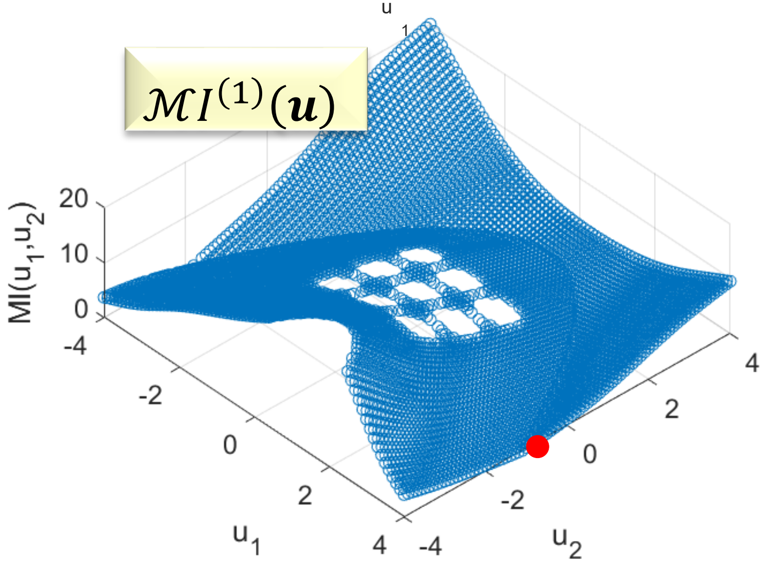


$$\boldsymbol{u}^{(1)} = \min_{u} \mathcal{M}I^{(0)}(\boldsymbol{u})$$

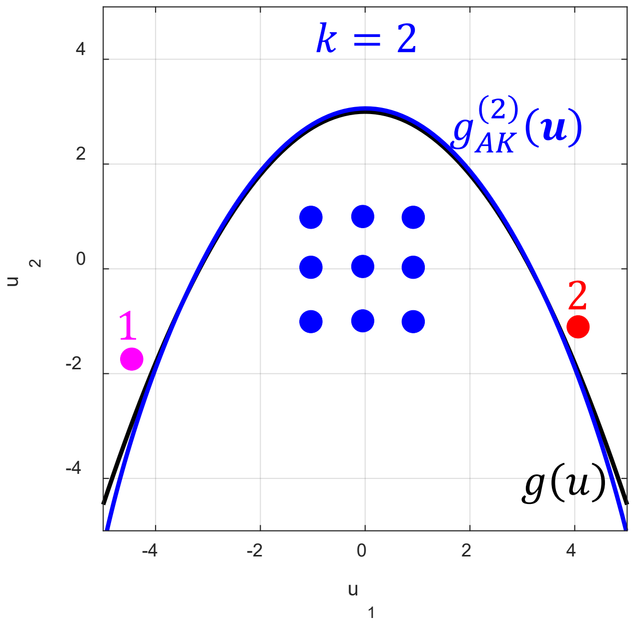
AARHUS UNIVERSITY
DEPARTMENT OF ENGINEERING

ACTIVE LEARNING AL-MI: ITERATION 2



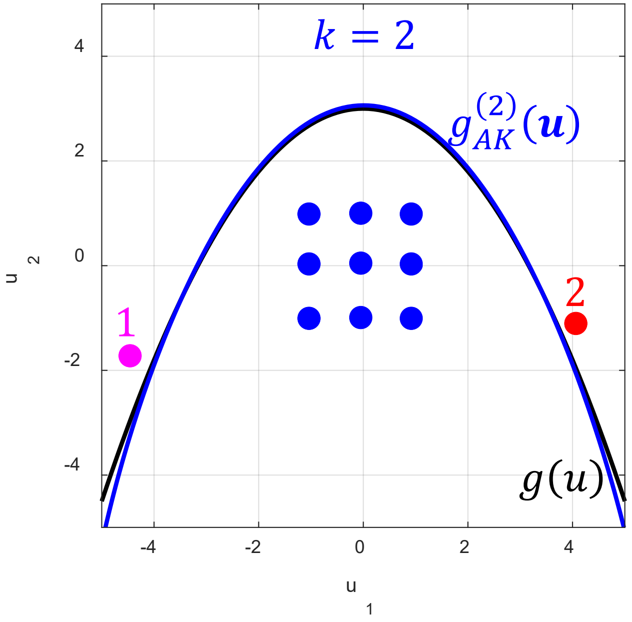


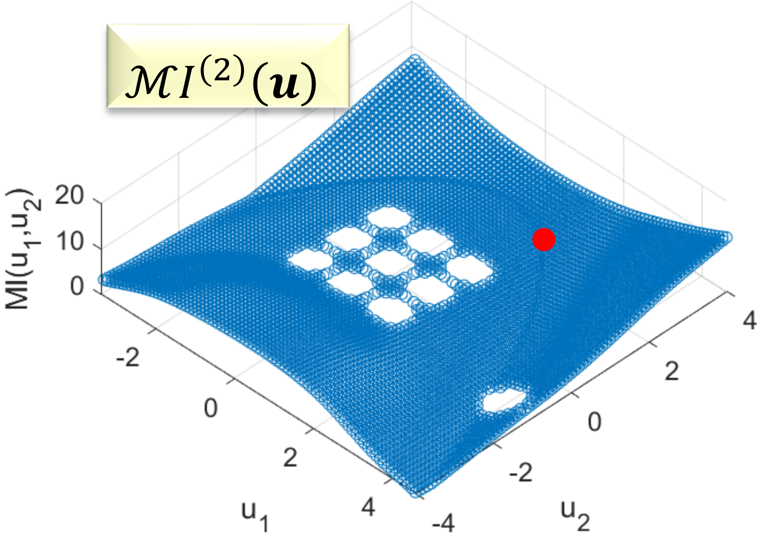
$$g(u_1, u_2) = 3 - u_2 - 0.3u_1^2$$



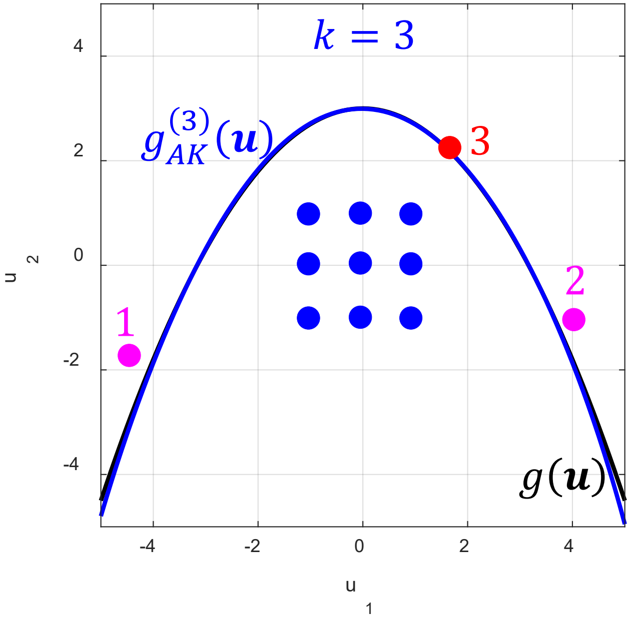
$$\boldsymbol{u}^{(2)} = \min_{u} \mathcal{M}I^{(1)}(\boldsymbol{u})$$

ACTIVE LEARNING AL-MI: ITERATION 3





$$g(u_1, u_2) = 3 - u_2 - 0.3u_1^2$$



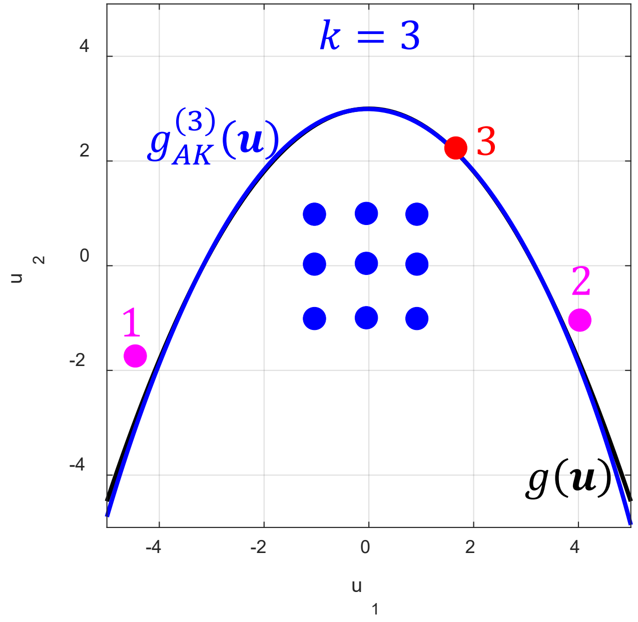
$$\boldsymbol{u}^{(3)} = \min_{u} \mathcal{M}I^{(2)}(\boldsymbol{u})$$

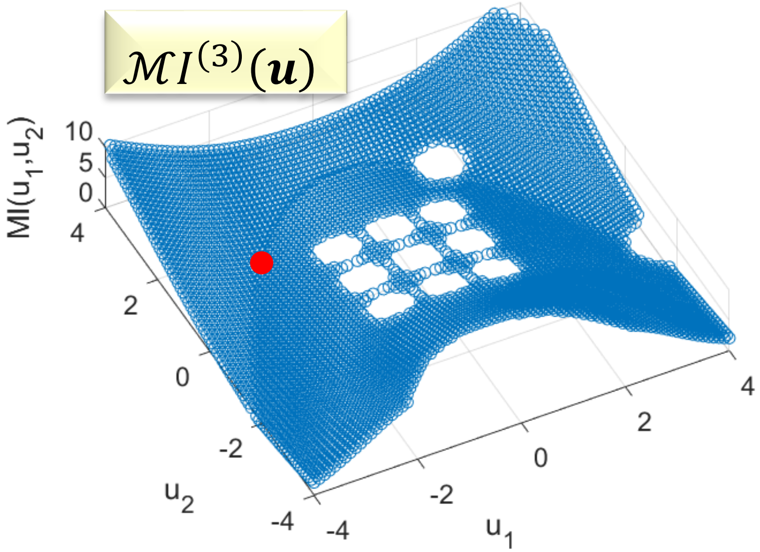
NDI
SOR

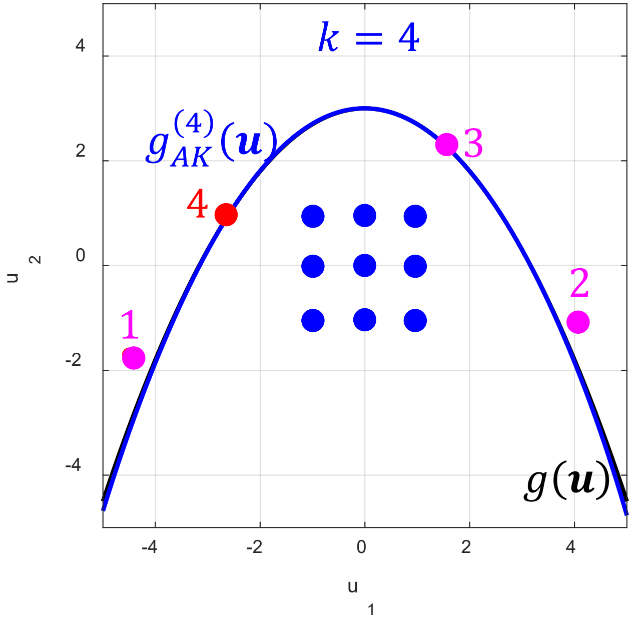
AARHUS UNIVERSITY
DEPARTMENT OF ENGINEERING

ACTIVE LEARNING AL-MI: ITERATION 4

$$g(u_1, u_2) = 3 - u_2 - 0.3u_1^2$$







$$\boldsymbol{u}^{(4)} = \min_{u} \mathcal{M}I^{(3)}(\boldsymbol{u})$$

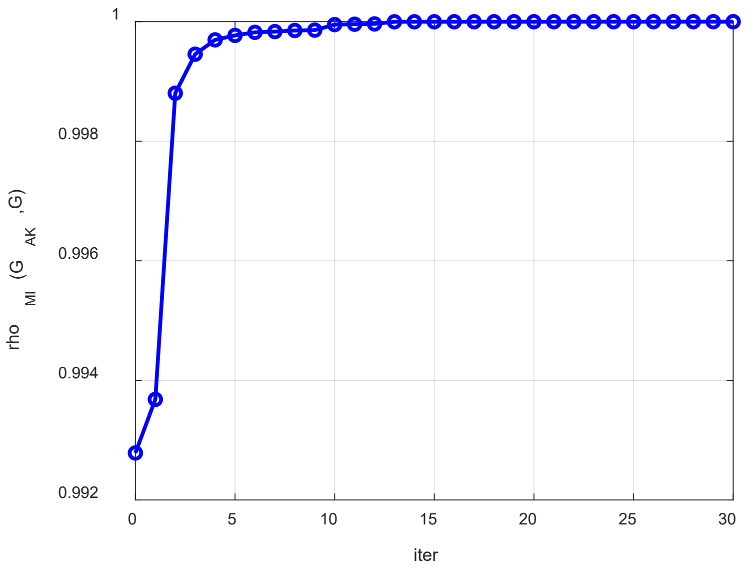
AARHUS UNIVERSITY
DEPARTMENT OF ENGINEERING

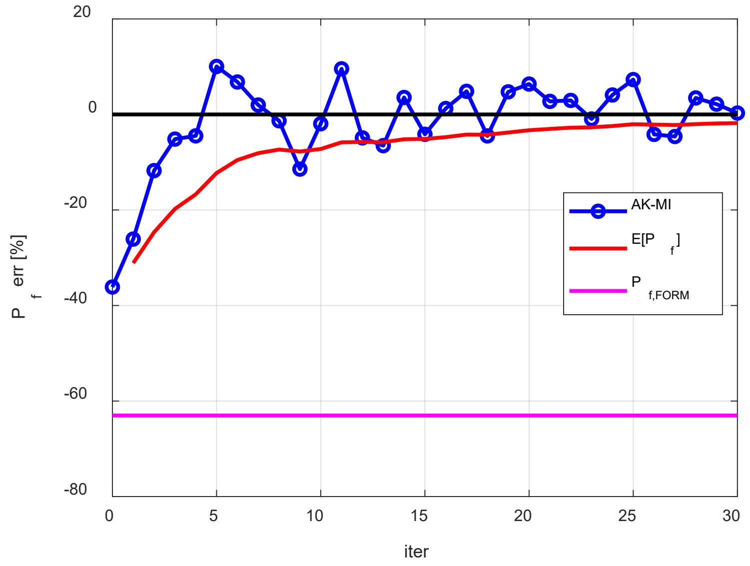
ACTIVE LEARNING AL-MI: EXAMPLE

$$g_{AK}(u_1,u_2) = \mu_g(u_1,u_2) + \varepsilon(u_1,u_2)$$

$$g(u_1,u_2) = 3 - u_2 - 0.3u_1^2$$

$$\rho_{MI}(g, g_{AK})$$





CONCLUSIONS

- It is proposed to adopt as a measure of dependence the Informational Coefficient of Correlation, instead of the correlation; the correlation is not a proper measure of dependence and always underestimates the dependence between random variables

- The informational coefficient of correlation is used to define informational sensitivity indices, ranking the importance of the variables through the Value Of Information

- For surrogate modelling it is proposed a procedure of Adaptive Kriging where the epistemic uncertainty is described through the Entropy (AK-H) (and not the variance)

- For surrogate modelling it is proposed a procedure of Active Learning Surrogate Modelling based on the Mutual Information (AL-MI): different from the existing literature it can be implemented to any kind of metamodel (e.g. PCE, SVM, Neural Networks) and it is not limited to Kriging only

- The research is promising in exploring novel research directions in data-driven Uncertainty Quantification and Risk Analysis for Risk-informed Digital Twin, including Decision Making under Uncertainty, Bayesian experimental design, optimal sensor placement

Thank you - Grazie

AARHUS
UNIVERSITY