Adaptive kriging meta-models for the simulation of rare events by importance sampling

V. Dubourg$^{1,2}$, B. Sudret$^{1,2}$, J.-M. Bourinet$^1$, F. Deheeger$^2$

$^1$Clermont Université, IFMA, EA 3867, Laboratoire de Mécanique et Ingénieries, BP 10448, F-63000 Clermont-Ferrand

$^2$Phimeca Engineering, Centre d’Affaires du Zénith, 34 rue de Sarliève, F-63800 Cournon d’Auvergne

GdR Mascot’NUM – Journée des doctorants
Villard-de-Lans, March 23, 2011
Buckling is a structural instability phenomenon

- triggered by some excessive load \( \textit{to be determined} \);
- depending on varied initial conditions \textit{unknown to some extent} (e.g. geometry, boundary conditions and material properties);
- affecting \textit{slender} (optimally designed) structures

Axially compressed beam  
Railway track  
Silo
Buckling is the major failure scenario for submarines pressure hulls.
Deterministic design optimization

\[
\theta^* = \arg \min_{\theta \in D_\theta} c(\theta) : \begin{cases} 
  f_i(\theta) \leq 0, & i = 1, \ldots, n_c \\
  g_l(x, \theta) \geq 0, & l = 1, \ldots, n_p
\end{cases}
\]
Problem formulation

Reliability-based design optimization

\[ \theta^* = \arg \min_{\theta \in D_\theta} c(\theta) : \begin{cases} 
  f_i(\theta) \leq 0, i = 1, \ldots, n_c \\
  P(g_l(X(\theta)) \leq 0) \leq \Phi(-\beta^0_l), l = 1, \ldots, n_p 
\end{cases} \]
Reliability-based design optimization

\[ X \sim f_X(\bullet, \theta^{(k)}) \]

**Probabilistic model**

\[ \mathcal{M}(\mathbf{x}) \]

**Physical model**

\[ P_f(\theta^{(k)}) \equiv \mathbb{P}(X \in D_f \mid \theta^{(k)}) \]

\[ D_f = \{ \mathbf{x} : g(\mathbf{x}, \mathcal{M}(\mathbf{x})) \leq 0 \} \]

**Failure scenario**

\[ \nabla_{\theta} P_f(\theta^{(k)}) \]

**Reliability sensitivity analysis**

\[ \theta^{(k+1)} = \text{Improve}(\theta^{(k)}) \]

**Optimizer**

\[ k = k + 1 \]

**Reliability analysis**

\[ P_f(\theta) \equiv \int_{D_f} f_X(\mathbf{x} \mid \theta) \, d\mathbf{x} = \mathbb{E}_{f_X(\bullet \mid \theta)} \left[ \mathbb{1}_{g \leq 0}(X) \right] \]
Reliability-based design optimization

\[ X \sim f_X(\bullet, \theta^{(k)}) \]

**Probabilistic model**

\[ \mathcal{M}(\mathbf{x}) \]

**Physical model**

Reliability analysis

\[ P_f(\mathbf{x} | \theta^{(k)}) = \mathbb{P}(X \in D_f | \theta^{(k)}) \]

**Failure scenario**

\[ D_f = \{ \mathbf{x} : g(\mathbf{x}, \mathcal{M}(\mathbf{x})) \leq 0 \} \]

Reliability sensitivity analysis

\[ \nabla_\theta P_f(\theta^{(k)}) \]

**Definition**

\[ P_f(\mathbf{x} | \theta) = \int_{D_f} f_X(\mathbf{x} | \theta) \, d\mathbf{x} = \mathbb{E} f_X(\bullet | \theta) \left[ I_{g \leq 0}(X) \right] \]
Outline

1. The kriging predictor
2. Adaptive probabilistic classification using kriging
3. Meta-model-based reliability analysis
4. Examples
Outline

1. The kriging predictor
   - What is a predictor?
   - Illustration of the kriging predictor

2. Adaptive probabilistic classification using kriging

3. Meta-model-based reliability analysis

4. Examples
What is a predictor?

Meta-model

- A *meta-model* is for a *model*, what the *model* is itself for the *real world*.
- It is *built from*:
  - a set of observations named a *design of experiments*;
  - and *statistical considerations*.

Predictor (a confidence meta-model)

- A *predictor* is able to give a *confidence level* onto its prediction.
- This level depends on the *available knowledge*: it is an *epistemic uncertainty*.
- Ex: (probabilistic) support vector machines, *kriging*.

Interest for reliability-based design

Predictors are *much faster to evaluate* than the original physical model, *flexible* and come with a certain *confidence measure*. 
What is a predictor?

Meta-model

- A **meta-model** is for a **model**, what the **model** is itself for the **real world**.
- It is **built from**:
  - a set of observations named a **design of experiments**;
  - and **statistical considerations**.
- **Ex**: quadratic response surfaces, linear functional regression, neural networks, support vector machines, **kriging**.

Predictor (a confidence meta-model)

- A **predictor** is able to give a **confidence level** onto its prediction.
- This level depends on the **available knowledge**: it is an **epistemic uncertainty**.
- **Ex**: (probabilistic) support vector machines, **kriging**.

Interest for reliability-based design

Predictors are **much faster to evaluate** than the original physical model, **flexible** and come with a certain **confidence measure**.
What is a predictor?

Meta-model

- A meta-model is for a model, what the model is itself for the real world.
- It is built from:
  - a set of observations named a design of experiments;
  - and statistical considerations.

Predictor (a confidence meta-model)

- A predictor is able to give a confidence level onto its prediction.
- This level depends on the available knowledge: it is an epistemic uncertainty.
- Ex: (probabilistic) support vector machines, kriging.

Interest for reliability-based design

Predictors are much faster to evaluate than the original physical model, flexible and come with a certain confidence measure.
The kriging predictor is an exact interpolator:
- the prediction variance is zero on the DOE \( \forall x \in X \);
- the predictor matches the observations \( \hat{M}(x_i) = M(x_i), i = 1, \ldots, m \).

The prediction variance increases as the prediction is made far from the observations.
Outline

1. The kriging predictor

2. Adaptive probabilistic classification using kriging
   - From regression to probabilistic classification
   - The probabilistic classification function
   - The proposed refinement strategy
   - Illustration

3. Meta-model-based reliability analysis

4. Examples
From regression to probabilistic classification

Classification ($g \leq 0$ vs. $g > 0$)
The probabilistic classification function

\( \mathcal{P} (\neq \mathbb{P}) \) is the probability \textit{w.r.t.} the kriging epistemic uncertainty.

- Let \( \pi \) denote \textit{the probabilistic classification function} defined as:
  \[
  \pi(x) = \mathcal{P} \left[ \hat{G}(x) \leq 0 \right] = \Phi \left( \frac{0 - \mu_{\hat{G}}(x)}{\sigma_{\hat{G}}(x)} \right)
  \]

- Let \( M \) denote the \textit{margin of uncertainty} defined as:
  \[
  M \equiv \left\{ x : 0 - k \sigma_{\hat{G}}(x) \leq \hat{G}(x) \leq 0 + k \sigma_{\hat{G}}(x) \right\}
  \]
  where \( k = \Phi^{-1} (97.5\%) = 1.96 \) if a 95% confidence interval is chosen.

- The probability that any point \( x \in \mathcal{D} \) lies in \( M \) has a \textit{closed-form expression}:
  \[
  \mathcal{P}[x \in M] = \Phi \left( \frac{k \sigma_{\hat{G}}(x) - \mu_{\hat{G}}(x)}{\sigma_{\hat{G}}(x)} \right) - \Phi \left( \frac{-k \sigma_{\hat{G}}(x) - \mu_{\hat{G}}(x)}{\sigma_{\hat{G}}(x)} \right)
  \]
The probabilistic classification function

\( P (\neq \mathbb{P}) \) is the probability \textit{w.r.t.} the kriging epistemic uncertainty.

Let \( \pi \) denote the \textit{probabilistic classification function} defined as:

\[
\pi(x) = P[\hat{G}(x) \leq 0] = \Phi \left( \frac{0 - \mu_{\hat{G}}(x)}{\sigma_{\hat{G}}(x)} \right)
\]

Let \( M \) denote the \textit{margin of uncertainty} defined as:

\[
M \equiv \left\{ x : 0 - k \sigma_{\hat{G}}(x) \leq \hat{G}(x) \leq 0 + k \sigma_{\hat{G}}(x) \right\}
\]

where \( k = \Phi^{-1}(97.5\%) = 1.96 \) if a 95% confidence interval is chosen.

The probability that any point \( x \in D_X \) lies in \( M \) has a \textit{closed-form expression}:

\[
P[x \in M] = \Phi \left( \frac{k \sigma_{\hat{G}}(x) - \mu_{\hat{G}}(x)}{\sigma_{\hat{G}}(x)} \right) - \Phi \left( \frac{-k \sigma_{\hat{G}}(x) - \mu_{\hat{G}}(x)}{\sigma_{\hat{G}}(x)} \right)
\]
The probabilistic classification function

\( P (\neq P) \) is the probability w.r.t. the kriging epistemic uncertainty.

Let \( \pi \) denote the probabilistic classification function defined as:

\[
\pi(x) = P \left[ \hat{G}(x) \leq 0 \right] = \Phi \left( \frac{0 - \mu_{\hat{G}}(x)}{\sigma_{\hat{G}}(x)} \right)
\]

Let \( M \) denote the margin of uncertainty defined as:

\[
M \equiv \left\{ x : 0 - k \sigma_{\hat{G}}(x) \leq \hat{G}(x) \leq 0 + k \sigma_{\hat{G}}(x) \right\}
\]

where \( k = \Phi^{-1}(97.5\%) = 1.96 \) if a 95% confidence interval is chosen.

The probability that any point \( x \in D_X \) lies in \( M \) has a closed-form expression:

\[
P[x \in M] = \Phi \left( \frac{k \sigma_{\hat{G}}(x) - \mu_{\hat{G}}(x)}{\sigma_{\hat{G}}(x)} \right) - \Phi \left( \frac{-k \sigma_{\hat{G}}(x) - \mu_{\hat{G}}(x)}{\sigma_{\hat{G}}(x)} \right)
\]
The probabilistic classification function

\( \mathcal{P} (\neq \mathcal{P}) \) is the probability w.r.t. the kriging epistemic uncertainty.

Let \( \pi \) denote the probabilistic classification function defined as:

\[
\pi(x) = \mathcal{P} \left[ \hat{G}(x) \leq 0 \right] = \Phi \left( \frac{0 - \mu\hat{G}(x)}{\sigma\hat{G}(x)} \right)
\]

Let \( \mathbb{M} \) denote the margin of uncertainty defined as:

\[
\mathbb{M} \equiv \left\{ x : 0 - k \sigma\hat{G}(x) \leq \hat{G}(x) \leq 0 + k \sigma\hat{G}(x) \right\}
\]

where \( k = \Phi^{-1}(97.5\%) = 1.96 \) if a 95% confidence interval is chosen.

The probability that any point \( x \in \mathcal{D}_x \) lies in \( \mathbb{M} \) has a closed-form expression:

\[
\mathcal{P}[x \in \mathbb{M}] = \Phi \left( \frac{k \sigma\hat{G}(x) - \mu\hat{G}(x)}{\sigma\hat{G}(x)} \right) - \Phi \left( \frac{-k \sigma\hat{G}(x) - \mu\hat{G}(x)}{\sigma\hat{G}(x)} \right)
\]
Sequential refinement strategies

- Given a *design of experiments* $X = \{x_1, \ldots, x_m\}$, and a *prediction* $\hat{G}$,
- *the best improvement point* to reduce the margin of uncertainty *maximizes* the proposed criterion:
  $$x_{m+1} = \arg\max_{x \in D_X} P[x \in M]$$
- Other criteria: *expected feasibility function*, *one-step-look-ahead criterion*, *etc.*

Premise

- The proposed criteria are *highly multimodal*, therefore the global optimization problem is *ill-posed*;
- There does not exist a *single best point* (especially for the contour approximation problem);

Alternative (inspired from [Hurtado (2004), Deheeger & Lemaire (2007)])

- Let us consider $C(x) \propto P[x \in M] w(x)$ as a PDF for the improvement points;
- $w$ is a *weighting PDF* to ensure that $C$ is *proper*, *i.e.*:
  $$\int_{D_X} C(x) \, dx < +\infty$$
The proposed refinement strategy

State-of-the-art

[Bichon et al. (2008), Vazquez & Bect (2009), Echard et al. (2011)]

Sequential refinement strategies

- Given a design of experiments \( X = \{x_1, \ldots, x_m\} \), and a prediction \( \hat{G} \),
- the best improvement point to reduce the margin of uncertainty maximizes the proposed criterion:
  \[
  x_{m+1} = \arg \max_{x \in \mathcal{D}_X} P[x \in M]
  \]
- Other criteria: expected feasibility function, one-step-look-ahead criterion, etc.

Premise

- The proposed criteria are highly multimodal, therefore the global optimization problem is ill-posed;
- There does not exist a single best point (especially for the contour approximation problem);

Alternative (inspired from [Hurtado (2004), Deheeger & Lemaire (2007)])

- Let us consider \( C(x) \propto P[x \in M]w(x) \) as a PDF for the improvement points;
- \( w \) is a weighting PDF to ensure that \( C \) is proper, i.e.
  \[
  \int_{\mathcal{D}_X} C(x) \, dx < +\infty
  \]
The proposed refinement strategy

State-of-the-art

[\textit{Bichon et al. (2008), Vazquez & Bect (2009), Echard et al. (2011)}]

Sequential refinement strategies

- Given a \textit{design of experiments} $X = \{x_1, \ldots, x_m\}$, and a \textit{prediction} $\hat{G}$,
- \textit{the best improvement point} to reduce the margin of uncertainty \textit{maximizes} the proposed criterion:

$$x_{m+1} = \arg \max_{x \in D_X} P[x \in M]$$

- Other criteria: \textit{expected feasibility function}, \textit{one-step-look-ahead criterion}, \textit{etc.}

Premise

- The proposed criteria are \textit{highly multimodal}, therefore the global optimization problem is \textit{ill-posed};
- There does not exist a \textit{single best point} (especially for the \textit{contour approximation} problem);

Alternative (inspired from [\textit{Hurtado (2004), Deheeger & Lemaire (2007)}])

- Let us consider $C(x) \propto P[x \in M] w(x)$ as a PDF for the improvement points;
- $w$ is a \textit{weighting PDF} to ensure that $C$ is \textit{proper}, \textit{i.e.}:

$$\int_{D_X} C(x) \, dx < +\infty$$
The proposed refinement strategy

Step-by-step

0. Initialization

1. Define the refinement pseudo-PDF $C$

2. Sampling of $C$

3. Statistical reduction of the candidates

4. Evaluation & update
The proposed refinement strategy

Step-by-step

0. Initialization
- Space-filling design of experiments
- Build an initial kriging surrogate

1. Define the refinement pseudo-PDF $C$

2. Sampling of $C$

3. Statistical reduction of the candidates

4. Evaluation & update
The proposed refinement strategy

Step-by-step

0. Initialization

1. Define the refinement pseudo-PDF $C$
   
   $$C(x) = P[x \in M] w(x)$$

2. Sampling of $C$

3. Statistical reduction of the candidates

4. Evaluation & update
The proposed refinement strategy

Step-by-step

0. Initialization

1. Define the refinement pseudo-PDF $C$

2. Sampling of $C$
   
   *Sampling methods for PDF defined up to an unknown normalizing constant.*

   \textit{Ex: slice sampling (MCMC)}

3. Statistical reduction of the candidates

4. Evaluation & update
The proposed refinement strategy

Step-by-step

0. Initialization

1. Define the refinement pseudo-PDF \( C \)

2. Sampling of \( C \)

3. Statistical reduction of the candidates
   \( \text{Non-supervised classification techniques.} \)
   \( \text{Ex: K-means clustering} \)

4. Evaluation & update
The proposed refinement strategy

Step-by-step

0. Initialization

1. Define the refinement pseudo-PDF $C$

2. Sampling of $C$

3. Statistical reduction of the candidates

4. Evaluation & update
The proposed refinement strategy

Illustration

\[ C(u) = P[u \in M] \varphi_n(u, 0, \text{Id}_n) \]
Outline

1. The kriging predictor
2. Adaptive probabilistic classification using kriging
3. **Meta-model-based reliability analysis**
   - State-of-the-art
   - Importance sampling
   - Meta-model-based importance sampling
   - Stopping the adaptive refinement strategy
4. Examples
Substitution of the failure subset $\mathcal{D}_f$

[Deheeger & Lemaire (2007), Bichon et al. (2008), Vazquez & Bect (2009), Echard et al. (2011)]

Let $\hat{\mathcal{D}}^{i}_{-1}$, $\hat{\mathcal{D}}^{0}_{f}$ and $\hat{\mathcal{D}}^{1+}_{f}$ denote the following subsets of $\mathcal{D}_X$:

$$\hat{\mathcal{D}}^{i}_{f} = \{ x : \mu_{\hat{G}}(x) + i \cdot 1.96 \sigma_{\hat{G}}(x) \leq 0 \}, \quad i = -1, 0, +1$$

Due to the **positiveness** of standard deviation:

$$\hat{\mathcal{D}}^{+1}_{f} \subseteq \hat{\mathcal{D}}^{0}_{f} \subseteq \hat{\mathcal{D}}^{-1}_{f}$$

$$\Rightarrow \mathbb{P}(X \in \hat{\mathcal{D}}^{+1}_{f}) \leq \mathbb{P}(X \in \hat{\mathcal{D}}^{0}_{f}) \leq \mathbb{P}(X \in \hat{\mathcal{D}}^{-1}_{f})$$

Unfortunately there is no proof that:

$$\hat{\mathcal{D}}^{+1}_{f} \subseteq \mathcal{D}_f \subseteq \hat{\mathcal{D}}^{-1}_{f}?$$
Substitution of the failure subset $\mathcal{D}_f$

[Deheeger & Lemaire (2007), Bichon et al. (2008), Vazquez & Bect (2009), Echard et al. (2011)]

$\hat{P}_f = \frac{N_f}{N}$

Let $\hat{\mathcal{D}}_{f}^{-1}, \hat{\mathcal{D}}_{f}^{0}$ and $\hat{\mathcal{D}}_{f}^{+1}$ denote the following subsets of $\mathcal{D}_X$:

$\hat{\mathcal{D}}_{f}^{i} = \{x : \mu_{\hat{G}}(x) + i \cdot 1.96 \cdot \sigma_{\hat{G}}(x) \leq 0\}, \quad i = -1, 0, +1$

Due to the **positiveness** of standard deviation:

$\hat{\mathcal{D}}_{f}^{+1} \subseteq \hat{\mathcal{D}}_{f}^{0} \subseteq \hat{\mathcal{D}}_{f}^{-1}$

$\Rightarrow \mathbb{P}(X \in \hat{\mathcal{D}}_{f}^{+1}) \leq \mathbb{P}(X \in \hat{\mathcal{D}}_{f}^{0}) \leq \mathbb{P}(X \in \hat{\mathcal{D}}_{f}^{-1})$

Unfortunately there is **no proof** that:

$\hat{\mathcal{D}}_{f}^{+1} \subseteq \mathcal{D}_f \subseteq \hat{\mathcal{D}}_{f}^{-1}$
Substitution of the failure subset $\mathcal{D}_f$

[Deheeger & Lemaire (2007), Bichon et al. (2008), Vazquez & Bect (2009), Echard et al. (2011)]

Let $\hat{\mathcal{D}}_{f_1}^{-1}$, $\hat{\mathcal{D}}_{f_0}$ and $\hat{\mathcal{D}}_{f_1}^{+1}$ denote the following subsets of $\mathcal{D}_X$:

$$\hat{\mathcal{D}}_{f_i} = \{ x : \mu_{\hat{G}}(x) + i \cdot 1.96 \sigma_{\hat{G}}(x) \leq 0 \}, \quad i = -1, 0, +1$$

Due to the *positiveness* of standard deviation:

$$\hat{\mathcal{D}}_{f_1}^{+1} \subseteq \hat{\mathcal{D}}_{f_0} \subseteq \hat{\mathcal{D}}_{f_1}^{-1}$$

$$\Rightarrow P(X \in \hat{\mathcal{D}}_{f_1}^{+1}) \leq P(X \in \hat{\mathcal{D}}_{f_0}) \leq P(X \in \hat{\mathcal{D}}_{f_1}^{-1})$$

Unfortunately there is no proof that:

$$\hat{\mathcal{D}}_{f_1}^{+1} \subseteq \mathcal{D}_f \subseteq \hat{\mathcal{D}}_{f_1}^{-1}$$
Substitution of the failure subset $\mathcal{D}_f$

[Deheeger & Lemaire (2007), Bichon et al. (2008), Vazquez & Bect (2009), Echard et al. (2011)]

Let $\mathcal{D}_f^{-1}$, $\mathcal{D}_f^{0}$ and $\mathcal{D}_f^{+1}$ denote the following subsets of $\mathcal{D}_{\mathbf{x}}$:

$$\mathcal{D}_f^{i} = \left\{ \mathbf{x} : \mu_{\hat{G}}(\mathbf{x}) + i \cdot 1.96 \sigma_{\hat{G}}(\mathbf{x}) \leq 0 \right\}, \quad i = -1, 0, +1$$

Due to the *positiveness* of standard deviation:

$$\mathcal{D}_f^{+1} \subseteq \mathcal{D}_f^{0} \subseteq \mathcal{D}_f^{-1}$$

$$\Rightarrow \mathbb{P}(\mathbf{X} \in \mathcal{D}_f^{+1}) \leq \mathbb{P}(\mathbf{X} \in \mathcal{D}_f^{0}) \leq \mathbb{P}(\mathbf{X} \in \mathcal{D}_f^{-1})$$

Unfortunately there is *no proof* that:

$$\mathcal{D}_f^{+1} \subseteq \mathcal{D}_f \subseteq \mathcal{D}_f^{-1}$$
Substitution of the failure subset $D_f$

[Deheeger & Lemaire (2007), Bichon et al. (2008), Vazquez & Bect (2009), Echard et al. (2011)]

Let $\hat{D}_f^{-1}$, $\hat{D}_f^0$ and $\hat{D}_f^{+1}$ denote the following subsets of $D_X$:

$$\hat{D}_f^i = \{ \mathbf{x} : \mu_G(\mathbf{x}) + i \cdot 1.96 \sigma_G(\mathbf{x}) \leq 0 \}, \quad i = -1, 0, +1$$

Due to the *positiveness* of standard deviation:

$$\hat{D}_f^{+1} \subseteq \hat{D}_f^0 \subseteq \hat{D}_f^{-1}$$

$$\Rightarrow P(\mathbf{X} \in \hat{D}_f^{+1}) \leq P(\mathbf{X} \in \hat{D}_f^0) \leq P(\mathbf{X} \in \hat{D}_f^{-1})$$

Unfortunately there is no proof that:

$$\hat{D}_f^{+1} \subseteq D_f \subseteq \hat{D}_f^{-1}$$
Substitution of the failure subset $\mathcal{D}_f$  

[Deheeger & Lemaire (2007), Bichon et al. (2008), Vazquez & Bect (2009), Echard et al. (2011)]

Let $\hat{\mathcal{D}}^{-1}_f$, $\hat{\mathcal{D}}^0_f$ and $\hat{\mathcal{D}}^{+1}_f$ denote the following subsets of $\mathcal{D}_X$:

$$\hat{\mathcal{D}}^i_f = \{ \mathbf{x} : \mu_{\hat{G}}(\mathbf{x}) + i \cdot 1.96 \sigma_{\hat{G}}(\mathbf{x}) \leq 0 \}, \quad i = -1, 0, +1$$

Due to the *positiveness* of standard deviation:

$$\hat{\mathcal{D}}^{+1}_f \subseteq \hat{\mathcal{D}}^0_f \subseteq \hat{\mathcal{D}}^{-1}_f$$

$$\Rightarrow \mathbb{P}(\mathbf{X} \in \hat{\mathcal{D}}^{+1}_f) \leq \mathbb{P}(\mathbf{X} \in \hat{\mathcal{D}}^0_f) \leq \mathbb{P}(\mathbf{X} \in \hat{\mathcal{D}}^{-1}_f)$$

Unfortunately there is no proof that:

$$\hat{\mathcal{D}}^{+1}_f \subseteq \mathcal{D}_f \subseteq \hat{\mathcal{D}}^{-1}_f$$
Substitution of the failure subset $\mathcal{D}_f$

[Deheeger & Lemaire (2007), Bichon et al. (2008), Vazquez & Bect (2009), Echard et al. (2011)]

Let $\hat{D}^{-1}_f$, $\hat{D}^0_f$ and $\hat{D}^{+1}_f$ denote the following subsets of $\mathcal{D}_X$:

$$\hat{D}^i_f = \{ x : \mu_\hat{G}(x) + i \cdot 1.96 \cdot \sigma_\hat{G}(x) \leq 0 \}, \quad i = -1, 0, +1$$

Due to the \textit{positiveness} of standard deviation:

$$\hat{D}^{+1}_f \subseteq \hat{D}^0_f \subseteq \hat{D}^{-1}_f$$

$$\Rightarrow \mathbb{P}(X \in \hat{D}^{+1}_f) \leq \mathbb{P}(X \in \hat{D}^0_f) \leq \mathbb{P}(X \in \hat{D}^{-1}_f)$$

Unfortunately there is no proof that:

$$\hat{D}^{+1}_f \subseteq \mathcal{D}_f \subseteq \hat{D}^{-1}_f$$
Substitution of the failure subset $\mathcal{D}_f$

[[Deheeger & Lemaire (2007), Bichon et al. (2008), Vazquez & Bect (2009), Echard et al. (2011)]]

- Let $\hat{\mathcal{D}}_{f}^{i}$, $\hat{\mathcal{D}}_{f}^{0}$ and $\hat{\mathcal{D}}_{f}^{+1}$ denote the following subsets of $\mathcal{D}_X$:
  \[
  \hat{\mathcal{D}}_{f}^{i} = \{ x : \mu_{\hat{G}}(x) + i \cdot 1.96 \sigma_{\hat{G}}(x) \leq 0 \}, \quad i = -1, 0, +1
  \]
- Due to the **positiveness** of standard deviation:
  \[
  \hat{\mathcal{D}}_{f}^{+1} \subseteq \hat{\mathcal{D}}_{f}^{0} \subseteq \hat{\mathcal{D}}_{f}^{-1}
  \]
  \[
  \Rightarrow P(X \in \hat{\mathcal{D}}_{f}^{+1}) \leq P(X \in \hat{\mathcal{D}}_{f}^{0}) \leq P(X \in \hat{\mathcal{D}}_{f}^{-1})
  \]
- Unfortunately there is no proof that:
  \[
  \hat{\mathcal{D}}_{f}^{+1} \subseteq \mathcal{D}_f \subseteq \hat{\mathcal{D}}_{f}^{-1}
  \]
Importance sampling

[Rubinstein & Kroese (2008)]

**Principle**

Back to *the integral definition* of the failure probability:

\[ P_f = \int_{\mathcal{D}_X} \mathbb{1}_{g \leq 0} (x) f_X(x) \, dx \]

\[ = \int_{\mathcal{D}_X} \mathbb{1}_{g \leq 0} (x) \frac{f_X(x)}{h(x)} h(x) \, dx, \quad \text{so that:} \quad \mathcal{D} \left( \mathbb{1}_{g \leq 0} f_X \right) \subseteq \mathcal{D}(h) \]

\[ = \mathbb{E}_h \left[ \mathbb{1}_{g \leq 0} (X) \ell (X) \right], \quad \text{where:} \quad \ell (x) = \frac{f_X(x)}{h(x)} \]

**Choice of the instrumental PDF \( h \)**

The *optimal PDF reduces* the estimation variance to zero:

\[ h^* (x) = \frac{\mathbb{1}_{g \leq 0} (x) f_X(x)}{\int_{\mathcal{D}_X} \mathbb{1}_{g \leq 0} (x) f_X(x) \, dx} = \frac{\mathbb{1}_{g \leq 0} (x) f_X(x)}{P_f} \]

but \( P_f \) is *the quantity of interest*!

**Objective**: Find a *good approximation* of \( h^* \).
Importance sampling

[Rubinstein & Kroese (2008)]

Principle
Back to the integral definition of the failure probability:

\[
P_f = \int_{D_X} \mathbb{1}_{g \leq 0}(\mathbf{x}) f_X(\mathbf{x}) \, d\mathbf{x}
\]

\[
= \int_{D_X} \mathbb{1}_{g \leq 0}(\mathbf{x}) \frac{f_X(\mathbf{x})}{h(\mathbf{x})} h(\mathbf{x}) \, d\mathbf{x}, \quad \text{so that: } D \left( \mathbb{1}_{g \leq 0} f_X \right) \subseteq D(h)
\]

\[
= \mathbb{E}_h \left[ \mathbb{1}_{g \leq 0}(\mathbf{X}) \ell(\mathbf{X}) \right], \quad \text{where: } \ell(\mathbf{x}) = \frac{f_X(\mathbf{x})}{h(\mathbf{x})}
\]

Choice of the instrumental PDF \( h \)
The optimal PDF reduces the estimation variance to zero:

\[
h^*(\mathbf{x}) = \frac{\mathbb{1}_{g \leq 0}(\mathbf{x}) f_X(\mathbf{x})}{\int_{D_X} \mathbb{1}_{g \leq 0}(\mathbf{x}) f_X(\mathbf{x}) \, d\mathbf{x}} = \frac{\mathbb{1}_{g \leq 0}(\mathbf{x}) f_X(\mathbf{x})}{P_f}
\]

but \( P_f \) is the quantity of interest!

Objective: Find a good approximation of \( h^* \).
Importance sampling

[Rubinstein & Kroese (2008)]

**Principle**

Back to *the integral definition* of the failure probability:

\[
P_f = \int_{D_X} \mathbb{1}_{g \leq 0}(\mathbf{x}) f_X(\mathbf{x}) \, d\mathbf{x}
\]

\[
= \int_{D_X} \mathbb{1}_{g \leq 0}(\mathbf{x}) \frac{f_X(\mathbf{x})}{h(\mathbf{x})} h(\mathbf{x}) \, d\mathbf{x}, \quad \text{so that: } D \left( \mathbb{1}_{g \leq 0} f_X \right) \subseteq D(h)
\]

\[
= \mathbb{E}_h \left[ \mathbb{1}_{g \leq 0}(\mathbf{X}) \ell(\mathbf{X}) \right], \quad \text{where: } \ell(\mathbf{x}) = \frac{f_X(\mathbf{x})}{h(\mathbf{x})}
\]

**Choice of the instrumental PDF \( h \)**

The *optimal* PDF *reduces* the estimation variance *to zero*:

\[
h^*(\mathbf{x}) = \frac{\mathbb{1}_{g \leq 0}(\mathbf{x}) f_X(\mathbf{x})}{\int_{D_X} \mathbb{1}_{g \leq 0}(\mathbf{x}) f_X(\mathbf{x}) \, d\mathbf{x}} \equiv \frac{\mathbb{1}_{g \leq 0}(\mathbf{x}) f_X(\mathbf{x})}{P_f}
\]

but \( P_f \) is *the quantity of interest*!

**Objective**: Find a *good approximation* of \( h^* \).
Let us replace the **failure indicator function** \( \mathbb{1}_{g \leq 0} \) by the **probabilistic classification function** \( \pi \) :

\[
\hat{h}^*(\mathbf{x}) \equiv \frac{\pi(\mathbf{x}) f_X(\mathbf{x})}{P_{f\epsilon}}
\]

with

\[
P_{f\epsilon} = \int_{D_X} \pi(\mathbf{x}) f_X(\mathbf{x}) \, d\mathbf{x}
\]

Recall that the **probabilistic classification function** reads:

\[
\pi(\mathbf{x}) \equiv P\left[ \hat{G}(\mathbf{x}) \leq 0 \right]
\]
Let us replace the failure indicator function $\mathbb{1}_{g \leq 0}$ by the probabilistic classification function $\pi$:

$$\hat{h}^*(x) \equiv \frac{\pi(x) f_X(x)}{P_{f\varepsilon}}$$

with

$$P_{f\varepsilon} = \int_{D_X} \pi(x) f_X(x) \, dx$$
Meta-model-based importance sampling
Work on the definition...

Replacing the optimal instrumental PDF by its approximation, the *failure probability* eventually reads:

\[
P_f = \int_{D_X} \mathbb{1}_{g \leq 0}(x) \frac{f_X(x)}{h^*(x)} h^*(x) \, dx
\]

\[
= \int_{D_X} \mathbb{1}_{g \leq 0}(x) \frac{f_X(x)}{\pi(x) f_X(x)} h^*(x) \, dx
\]

\[
= P_f \varepsilon \int_{D_X} \mathbb{1}_{g \leq 0}(x) \frac{h^*(x)}{\pi(x)} \, dx \equiv P_f \varepsilon \alpha_{corr}
\]

where we have introduced:

\[
\alpha_{corr} = \mathbb{E}_{h^*} \left[ \frac{\mathbb{1}_{g \leq 0}(X)}{\pi(X)} \right]
\]

and recall that:

\[
P_f \varepsilon = \int_{D_X} \pi(x) f_X(x) \, dx \equiv \mathbb{E}_{f_X} [\pi(X)]
\]
Replacing the optimal instrumental PDF by its approximation, the \textit{failure probability} eventually reads:

\[
P_f = \int_{\mathcal{D}_X} \mathbb{1}_{g \leq 0}(x) \frac{f_X(x)}{\hat{h}^*(x)} \hat{h}^*(x) \, dx
\]

\[
= \int_{\mathcal{D}_X} \mathbb{1}_{g \leq 0}(x) \frac{f_X(x)}{\pi(x) f_X(x)} \hat{h}^*(x) \, dx
\]

\[
= P_{f \epsilon} \int_{\mathcal{D}_X} \frac{\mathbb{1}_{g \leq 0}(x)}{\pi(x)} \hat{h}^*(x) \, dx \equiv P_{f \epsilon} \alpha_{corr}
\]

where we have introduced:

\[
\alpha_{corr} \equiv \mathbb{E}_{\hat{h}^*} \left[ \frac{\mathbb{1}_{g \leq 0}(X)}{\pi(X)} \right]
\]

and recall that:

\[
P_{f \epsilon} = \int_{\mathcal{D}_X} \pi(x) f_X(x) \, dx \equiv \mathbb{E}_{f_X} [\pi(X)]
\]
Meta-model-based importance sampling

Work on the definition...

Replacing the optimal instrumental PDF by its approximation, the \textit{failure probability} eventually reads:

\[
P_f = \int_{D_X} \mathbb{1}_{g \leq 0}(x) \frac{f_X(x)}{h^*(x)} \widehat{h}^*(x) \, dx
\]

\[
= \int_{D_X} \mathbb{1}_{g \leq 0}(x) \frac{f_X(x)}{p_f^\varepsilon(x)} \widehat{h}^*(x) \, dx
\]

\[
= P_f^\varepsilon \int_{D_X} \frac{\mathbb{1}_{g \leq 0}(x)}{\pi(x)} \widehat{h}^*(x) \, dx \equiv P_f^\varepsilon \alpha_{\text{corr}}
\]

where we have introduced:

\[
\alpha_{\text{corr}} \equiv \mathbb{E}_{\widehat{h}^*} \left[ \frac{\mathbb{1}_{g \leq 0}(X)}{\pi(X)} \right]
\]

and recall that:

\[
P_f^\varepsilon = \int_{D_X} \pi(x) f_X(x) \, dx \equiv \mathbb{E}_{f_X} [\pi(X)]
\]
Meta-model-based importance sampling

Work on the definition...

Replacing the optimal instrumental PDF by its approximation, the \textit{failure probability} eventually reads:

\begin{align*}
P_f &= \int_{\mathcal{D}_X} \mathbb{1}_{g \leq 0}(\mathbf{x}) \frac{f_X(\mathbf{x})}{\hat{h}^*(\mathbf{x})} \hat{h}^*(\mathbf{x}) \, d\mathbf{x} \\
    &= \int_{\mathcal{D}_X} \mathbb{1}_{g \leq 0}(\mathbf{x}) \frac{f_X(\mathbf{x})}{\pi(\mathbf{x}) f_X(\mathbf{x})} \hat{h}^*(\mathbf{x}) \, d\mathbf{x} \\
    &= P_{f,\varepsilon} \int_{\mathcal{D}_X} \frac{\mathbb{1}_{g \leq 0}(\mathbf{x})}{\pi(\mathbf{x})} \hat{h}^*(\mathbf{x}) \, d\mathbf{x} \equiv P_{f,\varepsilon} \alpha_{\text{corr}}
\end{align*}

where we have introduced:

\[ \alpha_{\text{corr}} \equiv \mathbb{E}_{\hat{h}^*} \left[ \frac{\mathbb{1}_{g \leq 0}(X)}{\pi(X)} \right] \]

and recall that:

\[ P_{f,\varepsilon} = \int_{\mathcal{D}_X} \pi(\mathbf{x}) f_X(\mathbf{x}) \, d\mathbf{x} \equiv \mathbb{E}_{f_X} [\pi(X)] \]
The two terms’ *Monte-Carlo estimators* read:

\[
\hat{P}_{f\epsilon} = \frac{1}{N_{\epsilon}} \sum_{k=1}^{N_{\epsilon}} \pi(x^{(k)})
\]

\[
\hat{\alpha}_{\text{corr}} = \frac{1}{N_{\text{corr}}} \sum_{k=1}^{N_{\text{corr}}} \frac{1_{g \leq 0}(x^{(k)})}{\pi(x^{(k)})}
\]

where:

- \(\{x^{(1)}, \ldots, x^{(N_{\epsilon})}\}\) is distributed according to \(f_X\);
- \(\{x^{(1)}, \ldots, x^{(N_{\text{corr}})}\}\) is distributed according to \(\hat{h}^\ast\) (*using MCMC techniques*).

The final *estimator* is: \(\hat{P}_{f\text{ meta IS}} = \hat{P}_{f\epsilon} \hat{\alpha}_{\text{corr}}\).

The final *coefficient of variation* is proven to read as follows:

\[
\delta \equiv \frac{\sigma}{\hat{P}_{f\text{ meta IS}}} = \sqrt{\delta_{\epsilon}^2 + \delta_{\text{corr}}^2 + \delta_{\epsilon}^2 \delta_{\text{corr}}^2}
\]

\(\delta_{\epsilon}, \delta_{\text{corr}} \ll 1\)
The two terms’ *Monte-Carlo estimators* read:

\[
\hat{P}_{f\varepsilon} = \frac{1}{N_{\varepsilon}} \sum_{k=1}^{N_{\varepsilon}} \pi(x^{(k)})
\]

\[
\hat{\alpha}_{corr} = \frac{1}{N_{corr}} \sum_{k=1}^{N_{corr}} \frac{\mathbb{I}_{g \leq 0} (x^{(k)})}{\pi(x^{(k)})}
\]

where:

- \(\{x^{(1)}, \ldots, x^{(N_{\varepsilon})}\}\) is distributed according to \(f_X\);
- \(\{x^{(1)}, \ldots, x^{(N_{corr})}\}\) is distributed according to \(\hat{h}^*\) (*using MCMC techniques*).

The final *estimator* is: \(\hat{P}_{f\text{ meta IS}} = \hat{P}_{f\varepsilon} \hat{\alpha}_{corr}\).

The final *coefficient of variation* is proven to read as follows:

\[
\delta \equiv \frac{\sigma}{\hat{P}_{f\text{ meta IS}}} = \sqrt{\delta_{\varepsilon}^2 + \delta_{corr}^2 + \delta_{\varepsilon}^2 \delta_{corr}^2} \approx \sqrt{\delta_{\varepsilon}^2 + \delta_{corr}^2}
\]

\(\delta_{\varepsilon}, \delta_{corr} \ll 1\)
Meta-model-based importance sampling
The meta-model-based importance sampling estimator

- The two terms’ **Monte-Carlo estimators** read:

\[
\hat{P}_{f\varepsilon} = \frac{1}{N_{\varepsilon}} \sum_{k=1}^{N_{\varepsilon}} \pi(x^{(k)}) \\
\hat{\alpha}_{\text{corr}} = \frac{1}{N_{\text{corr}}} \sum_{k=1}^{N_{\text{corr}}} \frac{\mathbb{1}_{g \leq 0}(x^{(k)})}{\pi(x^{(k)})}
\]

where:

- \(\{x^{(1)}, \ldots, x^{(N_{\varepsilon})}\}\) is distributed according to \(f_X\);
- \(\{x^{(1)}, \ldots, x^{(N_{\text{corr}})}\}\) is distributed according to \(\hat{h}^*\) (using MCMC techniques).

- The final **estimator** is: \(\hat{P}_{f\text{ meta IS}} = \hat{P}_{f\varepsilon} \hat{\alpha}_{\text{corr}}\).
- The final **coefficient of variation** is proven to read as follows:

\[
\delta \equiv \frac{\sigma}{\hat{P}_{f\text{ meta IS}}} = \sqrt{\delta_{\varepsilon}^2 + \delta_{\text{corr}}^2 + \delta_{\varepsilon}^2 \delta_{\text{corr}}^2} \approx \sqrt{\delta_{\varepsilon}^2 + \delta_{\text{corr}}^2}
\]

\(\delta_{\varepsilon}, \delta_{\text{corr}} \ll 1\)
Stopping the adaptive refinement strategy
What is a “good kriging prediction” for our purpose?

The *adaptive refinement procedure* should be stopped if:

- the correction factor is *sufficiently close to 1*, meaning:

\[ P_f = P_f \varepsilon \alpha_{\text{corr}} \xrightarrow{\alpha_{\text{corr}} \rightarrow 1} P_f \varepsilon \]

- the *latest improvement* brought to the correction factor is less than some tolerance:

\[ \frac{\left| \alpha_{\text{corr}}^{(k+1)} - \alpha_{\text{corr}}^{(k)} \right|}{\alpha_{\text{corr}}^{(k)}} \leq \epsilon_{\alpha} \]

- some *maximum number of points in the DOE* is exceeded:

\[ m \geq N_{\text{DOE max}} \]

Problem

Estimating \( \alpha_{\text{corr}} \) is rather *expensive* so that it should be estimated *only once*!
Stopping the adaptive refinement strategy

What is a “good kriging prediction” for our purpose?

The *adaptive refinement procedure* should be stopped if:

- the correction factor is *sufficiently close to 1*, meaning:

\[
P_f = P_f \varepsilon \alpha_{\text{corr}} \xrightarrow{\alpha_{\text{corr}} \to 1} P_f \varepsilon
\]

- the *latest improvement* brought to the correction factor is less than some tolerance:

\[
\frac{|\alpha^{(k+1)}_{\text{corr}} - \alpha^{(k)}_{\text{corr}}|}{\alpha^{(k)}_{\text{corr}}} \leq \epsilon\alpha
\]

- some *maximum number of points in the DOE* is exceeded:

\[
m \geq N_{\text{DOE max}}
\]

**Problem**

*Estimating* $\alpha_{\text{corr}}$ is rather *expensive* so that it should be estimated *only once!*
Since $\hat{\alpha}_{\text{corr}}$ is expensive to evaluate, it is proposed to use its leave-one-out estimate instead to assess the prediction accuracy:

- At some point, the design of experiments $X = \{x_1, \ldots, x_m\}$ contains $m$ observations;
- We may compute the $m$ leave-one-out predictions of $\{g(x_i), i = 1, \ldots, m\}$ denoted by $\{\hat{G}_{X \setminus x_i}(x_i), i = 1, \ldots, m\}$
- And then, the following score:

$$\hat{\alpha}_{\text{corr LOO}} = \frac{1}{m} \sum_{i=1}^{m} \frac{\mathbb{1}_{g \leq 0}(x_i)}{P[\hat{G}_{X \setminus x_i}(x_i)]}$$

gives an idea about the order of magnitude of the correction factor.
Outline

1. The kriging predictor
2. Adaptive probabilistic classification using kriging
3. Meta-model-based reliability analysis
4. Examples
   - A first simple example
   - Buckling of imperfect shells
Influence of the dimension

Problem definition

\[ X \sim \mathcal{L} \mathcal{N}(1, 0.2 \text{Id}_n), \quad G = g(X) \equiv (n + 0.6 \sqrt{n}) - \sum_{i=1}^{n} X_i \]

Results

<table>
<thead>
<tr>
<th></th>
<th>(n)</th>
<th>2</th>
<th>10</th>
<th>50</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{p}_f)  (\text{MC})</td>
<td>(4.78 \times 10^{-3})</td>
<td>(2.70 \times 10^{-3})</td>
<td>(1.91 \times 10^{-3})</td>
<td>(1.73 \times 10^{-3})</td>
<td></td>
</tr>
<tr>
<td>(\delta_{\text{MC}})</td>
<td>(\leq 2%)</td>
<td>(\leq 2%)</td>
<td>(\leq 2%)</td>
<td>(\leq 2%)</td>
<td></td>
</tr>
<tr>
<td>(N)</td>
<td>522 000</td>
<td>925 000</td>
<td>1 100 000</td>
<td>1 450 000</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Meta-model-based importance sampling</th>
<th>(n)</th>
<th>2</th>
<th>10</th>
<th>50</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{p}_f)  (\text{meta IS})</td>
<td>(5.03 \times 10^{-3})</td>
<td>(2.66 \times 10^{-3})</td>
<td>(1.93 \times 10^{-3})</td>
<td>(1.70 \times 10^{-3})</td>
<td></td>
</tr>
<tr>
<td>(\delta_{\text{meta IS}})</td>
<td>(\leq 1.41%)</td>
<td>(\leq 2%)</td>
<td>(\leq 2%)</td>
<td>(\leq 2%)</td>
<td></td>
</tr>
<tr>
<td>(N_{\text{DOE}})</td>
<td>6 \times 2</td>
<td>6 \times 10</td>
<td>6 \times 50</td>
<td>6 \times 100</td>
<td></td>
</tr>
<tr>
<td>(\hat{p}_f)  (\epsilon)</td>
<td>(5.03 \times 10^{-3})</td>
<td>(2.26 \times 10^{-3})</td>
<td>(1.95 \times 10^{-3})</td>
<td>(1.83 \times 10^{-3})</td>
<td></td>
</tr>
<tr>
<td>(\delta_{\epsilon})</td>
<td>(\leq 1.41%)</td>
<td>(\leq 1.41%)</td>
<td>(\leq 1.41%)</td>
<td>(\leq 1.41%)</td>
<td></td>
</tr>
<tr>
<td>(\hat{\alpha}_{\text{corr}})</td>
<td>1</td>
<td>0.98</td>
<td>0.99</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>(\delta_{\text{corr}})</td>
<td>0%</td>
<td>(\leq 1.41%)</td>
<td>(\leq 1.41%)</td>
<td>(\leq 1.41%)</td>
<td></td>
</tr>
<tr>
<td>(N_{\text{corr}})</td>
<td>100</td>
<td>800</td>
<td>1 500</td>
<td>2 100</td>
<td></td>
</tr>
</tbody>
</table>
The Scordelis-Lo shell roof
Mechanical model: determine the critical buckling load

Structure of interest

Equilibrium path

Deformed shape at instability

Failure scenario

\[ D_f = \{ \xi \in \mathcal{D}_\Xi : g(\xi) = \lambda_{cr}(\xi) q - q_{service} \leq 0 \} \]
The Scordelis-Lo shell roof
Mechanical model: determine the critical buckling load

Structure of interest

Equilibrium path

Deformed shape at instability

Failure scenario

$$\mathcal{D}_f \equiv \{ \xi \in \mathcal{D}_\Xi : g(\xi) = \lambda_{cr}(\xi) q - q_{service} \leq 0 \}$$
The kriging predictor
Adaptive probabilistic classification using kriging
Meta-model-based reliability analysis
Examples

The Scordelis-Lo shell roof
Probabilistic model: 4 independent random fields

- The lognormal RFs are obtained by means of a translation of the sample paths of standard Gaussian RFs with squared exponential covariance;
- The Karhunen-Loève disc. error of the underlying Gaussian RFs is less than 5%;
- The imperfection RF is simulated from 3 i.i.d. Gaussian RVs whose standard deviation has been reasonably tuned.

<table>
<thead>
<tr>
<th>Field</th>
<th>Model</th>
<th>Mean</th>
<th>Std Dev.</th>
<th>RVs ξ</th>
</tr>
</thead>
<tbody>
<tr>
<td>E (MPa)</td>
<td>Lognormal RFs</td>
<td>210 000</td>
<td>6 300</td>
<td>30 × (N(0,1)) i.i.d.</td>
</tr>
<tr>
<td>(\sigma_y) (MPa)</td>
<td>covariance exp(^2)</td>
<td>390</td>
<td>27.3</td>
<td>30 × (N(0,1)) i.i.d.</td>
</tr>
<tr>
<td>e (mm)</td>
<td>(\ell_1 = \ell_2 = 3 500) mm</td>
<td>76</td>
<td>3.8</td>
<td>30 × (N(0,1)) i.i.d.</td>
</tr>
<tr>
<td>(\zeta) (mm)</td>
<td>Random. lin. comb. (3 modes)</td>
<td>0</td>
<td>(±3\sigma \approx \mu_e/2)</td>
<td>3 × (N(0,\sigma)) i.i.d.</td>
</tr>
</tbody>
</table>

93 RVs
The Scordelis-Lo shell roof
Probabilistic model: 4 independent random fields

- **$E(x, \xi_E)$**
- **$e(x, \xi_e)$**
- **$\zeta(x, \xi_\zeta)$**
- **$\sigma_y(x, \xi_{\sigma_y})$**

- The lognormal RFs are obtained by means of a translation of the sample paths of standard Gaussian RFs with squared exponential covariance;
- The Karhunen-Loève disc. error of the underlying Gaussian RFs is less than 5%;
- The imperfection RF is simulated from 3 i.i.d. Gaussian RVs whose standard deviation has been reasonably tuned.

<table>
<thead>
<tr>
<th>Field</th>
<th>Model</th>
<th>Mean</th>
<th>Std Dev.</th>
<th>RVs $\xi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E$ (MPa)</td>
<td>Lognormal RFs</td>
<td>210 000</td>
<td>6 300</td>
<td>$30 \times \mathcal{N} (0, 1)$ i.i.d.</td>
</tr>
<tr>
<td>$\sigma_y$ (MPa)</td>
<td>covariance $\exp^2$</td>
<td>390</td>
<td>27,3</td>
<td>$30 \times \mathcal{N} (0, 1)$ i.i.d.</td>
</tr>
<tr>
<td>$e$ (mm)</td>
<td>$\ell_1 = \ell_2 = 3500$ mm</td>
<td>76</td>
<td>3,8</td>
<td>$30 \times \mathcal{N} (0, 1)$ i.i.d.</td>
</tr>
<tr>
<td>$\zeta$ (mm)</td>
<td>Random. lin. comb. (3 modes)</td>
<td>0</td>
<td>$\pm 3 \sigma \approx \mu_e / 2$</td>
<td>$3 \times \mathcal{N} (0, \sigma)$ i.i.d.</td>
</tr>
</tbody>
</table>

93 RVs
The Scordelis-Lo shell roof
Probabilistic model: 4 independent random fields

- The lognormal RFs are obtained by means of a translation of the sample paths of standard Gaussian RFs with squared exponential covariance;
- The Karhunen-Loève disc. error of the underlying Gaussian RFs is less than 5%;
- The imperfection RF is simulated from 3 i.i.d. Gaussian RVs whose standard deviation has been reasonably tuned.

<table>
<thead>
<tr>
<th>Field</th>
<th>Model</th>
<th>Mean</th>
<th>Std Dev.</th>
<th>RVs</th>
<th>93 RVs</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E$ (MPa)</td>
<td>Lognormal RFs</td>
<td>210 000</td>
<td>6 300</td>
<td>30 $\times \mathcal{N} (0, 1)$ i.i.d.</td>
<td></td>
</tr>
<tr>
<td>$\sigma_y$ (MPa)</td>
<td>covariance $\exp^2$</td>
<td>390</td>
<td>27.3</td>
<td>30 $\times \mathcal{N} (0, 1)$ i.i.d.</td>
<td></td>
</tr>
<tr>
<td>$e$ (mm)</td>
<td>$l_1 = l_2 = 3 500$ mm</td>
<td>76</td>
<td>3.8</td>
<td>30 $\times \mathcal{N} (0, 1)$ i.i.d.</td>
<td></td>
</tr>
<tr>
<td>$\zeta$ (mm)</td>
<td>Random. lin. comb. (3 modes)</td>
<td>0</td>
<td>$\pm 3 \sigma \approx \mu_e / 2$</td>
<td>3 $\times \mathcal{N} (0, \sigma)$ i.i.d.</td>
<td></td>
</tr>
</tbody>
</table>
The Scordelis-Lo shell roof
Probabilistic model: 4 independent random fields

- The lognormal RFs are obtained by means of a translation of the sample paths of standard Gaussian RFs with squared exponential covariance;
- The Karhunen-Loève disc. error of the underlying Gaussian RFs is less than 5%;
- The imperfection RF is simulated from 3 i.i.d. Gaussian RVs whose standard deviation has been reasonably tuned.

<table>
<thead>
<tr>
<th>Field</th>
<th>Model</th>
<th>Mean</th>
<th>Std Dev.</th>
<th>RVs $\xi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E$ (MPa)</td>
<td>Lognormal RFs</td>
<td>210 000</td>
<td>6 300</td>
<td>$30 \times \mathcal{N} (0, 1) \text{ i.i.d.}$</td>
</tr>
<tr>
<td>$\sigma_y$ (MPa)</td>
<td>covariance exp$^2$</td>
<td>390</td>
<td>27.3</td>
<td>$30 \times \mathcal{N} (0, 1) \text{ i.i.d.}$</td>
</tr>
<tr>
<td>$e$ (mm)</td>
<td>$l_1 = l_2 = 3500$ mm</td>
<td>76</td>
<td>3.8</td>
<td>$30 \times \mathcal{N} (0, 1) \text{ i.i.d.}$</td>
</tr>
<tr>
<td>$\zeta$ (mm)</td>
<td>Random. lin. comb. (3 modes)</td>
<td>0</td>
<td>$\pm 3 \sigma \approx \mu_e / 2$</td>
<td>$3 \times \mathcal{N} (0, \sigma) \text{ i.i.d.}$</td>
</tr>
</tbody>
</table>
The Scordelis-Lo shell roof
Probabilistic model: 4 independent random fields

- The lognormal RFs are obtained by means of a translation of the sample paths of standard Gaussian RFs with squared exponential covariance;
- The Karhunen-Loève disc. error of the underlying Gaussian RFs is less than 5%;
- The imperfection RF is simulated from 3 i.i.d. Gaussian RVs whose standard deviation has been reasonably tuned.

<table>
<thead>
<tr>
<th>Field</th>
<th>Model</th>
<th>Mean</th>
<th>Std Dev.</th>
<th>RVs $\xi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E$ (MPa)</td>
<td>Lognormal RFs</td>
<td>210 000</td>
<td>6 300</td>
<td>$30 \times \mathcal{N}(0, 1)$ i.i.d.</td>
</tr>
<tr>
<td>$\sigma_y$ (MPa)</td>
<td>covariance $\exp^2$</td>
<td>390</td>
<td>27,3</td>
<td>$30 \times \mathcal{N}(0, 1)$ i.i.d.</td>
</tr>
<tr>
<td>$e$ (mm)</td>
<td>$\ell_1 = \ell_2 = 3 500$ mm</td>
<td>76</td>
<td>3,8</td>
<td>$30 \times \mathcal{N}(0, 1)$ i.i.d.</td>
</tr>
<tr>
<td>$\zeta$ (mm)</td>
<td>Random. lin. comb. (3 modes)</td>
<td>$0 \pm 3 \sigma \approx \mu_e / 2$</td>
<td></td>
<td>$3 \times \mathcal{N}(0, \sigma)$ i.i.d.</td>
</tr>
</tbody>
</table>

93 RVs
The Scordelis-Lo shell roof
Results: multiple most-probable failure configurations

Most probable failure configuration #1

Importance factors \( \sum \alpha_i^2 \)

\[
\begin{align*}
E(x, \xi_E^{(1)}) & \quad e(x, \xi_e^{(1)}) \\
\zeta(x, \xi_\zeta^{(1)}) & \quad \sigma_y(x, \xi_{\sigma_y}^{(1)})
\end{align*}
\]

\[
\beta^{(1)} \approx 4.01 \\
p_{f1}^{(1)} \approx 3.02 \times 10^{-5}
\]
The Scordelis-Lo shell roof

Results: multiple most-probable failure configurations

Importance factors ($\sum \alpha_i^2$)

$\beta^{(1)} \approx 4.01$
$\beta^{(2)} \approx 4.01$

$p_{f1}^{(1)} \approx 3.02 \times 10^{-5}$
$p_{f1}^{(2)} \approx 3.02 \times 10^{-5}$
The Scordelis-Lo shell roof
Results: multiple most-probable failure configurations

Most probable failure configuration #3

\[ E(x, \xi_E^{(3)}) \quad e(x, \xi_e^{(3)}) \]

\[ \zeta(x, \xi_\zeta^{(3)}) \quad \sigma_y(x, \xi_{\sigma y}^{(3)}) \]

Importance factors (\( \sum \alpha_i^2 \))

\[ \beta^{(1)} \approx 4.01 \]
\[ p_{f1}^{(1)} \approx 3.02 \times 10^{-5} \]

\[ \beta^{(2)} \approx 4.01 \]
\[ p_{f1}^{(2)} \approx 3.02 \times 10^{-5} \]

\[ \beta^{(3)} \approx 4.00 \]
\[ p_{f1}^{(3)} \approx 3.11 \times 10^{-5} \]
The Scordelis-Lo shell roof

Results: multiple most-probable failure configurations

Most probable failure configuration #4

\[ E(x, \xi_{E}) \]
\[ e(x, \xi_{e}) \]
\[ \zeta(x, \xi_{\zeta}) \]
\[ \sigma_{y}(x, \xi_{\sigma_{y}}) \]

Importance factors (\( \sum \alpha_{i}^{2} \))

\[ \beta^{(1)} \approx 4.01 \]
\[ p_{f1}^{(1)} \approx 3.02 \times 10^{-5} \]
\[ \beta^{(2)} \approx 4.01 \]
\[ p_{f1}^{(2)} \approx 3.02 \times 10^{-5} \]
\[ \beta^{(3)} \approx 4.00 \]
\[ p_{f1}^{(3)} \approx 3.11 \times 10^{-5} \]
\[ \beta^{(4)} \approx 4.01 \]
\[ p_{f1}^{(4)} \approx 3.04 \times 10^{-5} \]
The Scordelis-Lo shell roof
Results: multiple most-probable failure configurations

Most probable failure configuration #4

Importance factors \( \sum_i \alpha_i^2 \)

| \( \beta^{(1)} \) | \( \approx 4.01 \) |
| \( p_{f1}^{(1)} \) | \( \approx 3.02 \times 10^{-5} \) |
| \( \beta^{(2)} \) | \( \approx 4.01 \) |
| \( p_{f1}^{(2)} \) | \( \approx 3.02 \times 10^{-5} \) |
| \( \beta^{(3)} \) | \( \approx 4.00 \) |
| \( p_{f1}^{(3)} \) | \( \approx 3.11 \times 10^{-5} \) |
| \( \beta^{(4)} \) | \( \approx 4.01 \) |
| \( p_{f1}^{(4)} \) | \( \approx 3.04 \times 10^{-5} \) |

Serial system reliability: \( 1.22 \times 10^{-4} \leq p_{f1} \Sigma \leq 1.22 \times 10^{-4} \)
### Examples

#### A first simple example

**Buckling of imperfect shells**

The Scordelis-Lo shell roof

Results: probability estimates

<table>
<thead>
<tr>
<th>DOE</th>
<th>MPFP search</th>
<th>Simulations</th>
<th>$P_f$ est.</th>
<th>C.o.V.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subset (réf.)</td>
<td>-</td>
<td>20 000</td>
<td>$1.26 \times 10^{-4}$</td>
<td>12%</td>
</tr>
<tr>
<td>multi-FORM</td>
<td>$\approx 10 000$</td>
<td>-</td>
<td>$1.22 \times 10^{-4}$</td>
<td>-</td>
</tr>
<tr>
<td>meta-IS</td>
<td>$6 \times 93$</td>
<td>9 464</td>
<td>$1.32 \times 10^{-4}$</td>
<td>14%</td>
</tr>
</tbody>
</table>

\[ (2.06 \times 10^{-4} \times 0.64) \]

- The **CDF tail** can be reconstructed from a subset simulation analysis;
- This is a nice feature for **probabilistic buckling analyses**;
- *e.g.* for submarines pressure hulls: collapse probability *v.s.* diving depth.

![Graph showing the CDF tail reconstruction](image-url)
Conclusions

- The proposed meta-model-based importance sampling strategy allows one to quantify and eliminate the substitution error;

- It is closely related with other:
  - well-known strategies such as control variate or regression sampling;
  - two-step conditional sampling strategies: [Piera-Martinez et al. (2007)].

- It has also been successfully applied to basic RBDO examples with a few more points that were not developed here:
  - deriving the sensitivities of the estimator w.r.t. the design parameters;
  - estimating $P_{f_\epsilon}$ efficiently using a modified subset sampling strategy;
  - recycling the DOE from one RBDO iteration to the other.
Estimation of small failure probabilities in high dimensions by subset simulation.

Efficient global reliability analysis for nonlinear implicit performance functions.

AK-MCS : an Active learning reliability method combining Kriging and Monte Carlo Simulation.

*Structural reliability – Statistical learning perspectives.* Lecture notes in applied and computational mechanics. Springer.

Support vector machine for efficient subset simulations : 2SMART method.
*Proc. 10th Int. Conf. on Applications of Stat. and Prob. in Civil Engineering (ICASP10).*

Ditlevsen, O., H. Madsen (1996)
*Structural reliability methods.* Wiley.

RKHS classification for multivariate extreme-value analysis.
*Proc. 7th IASC - Statistics for Data Mining, Learning and Knowledge Extraction.*

Reliability analysis – a review and some perspectives.


The design and analysis of computer experiments.

A sequential Bayesian algorithm to estimate a probability of failure.
*Proc. 15th IFAC Symposium on System Identification.*