6TH INTERNATIONAL WORKSHOP ON RELIABLE ENGINEERING COMPUTING
Reliability and Computations of Infrastructures
May 25-28, 2014 | IIT, Chicago, IL, USA
REC 2014

ILLINOIS INSTITUTE OF TECHNOLOGY

A generalized numerical framework of imprecise probability to propagate epistemic uncertainty

Marco de Angelis,
E. Patelli, M. Beer

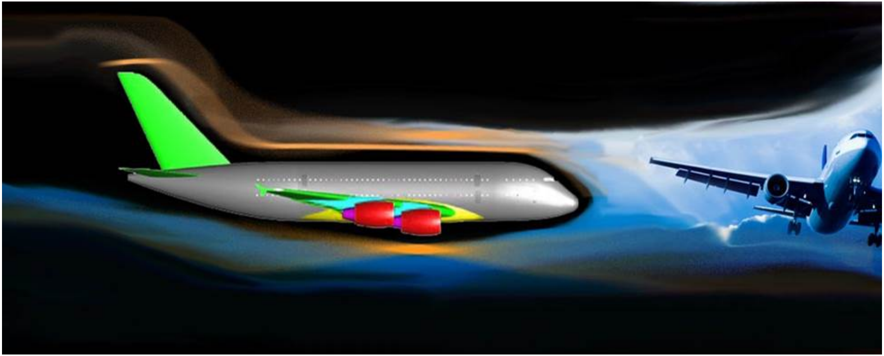
E: *marco.de-angelis@liverpool.ac.uk*
W: www.liv.ac.uk/risk-and-uncertainty
T: +44 01517945481

UNIVERSITY OF LIVERPOOL

Outline

Problem statement

- Computer model
- Reliability assessment
- Decision making



Sources of uncertainty

1. Model inputs (such as distributional models, hyper-parameters etc.)
2. Numerical approximation (discretization, truncation and round-off errors)
3. Model form (mathematical & physical model representing the system)

Uncertainty framework

- Verification
  - Numerical approximation
    - Discretization errors
    - Round-off errors
- **Uncertainty propagation**
  - propagation of input uncertainties through the model
  - processing of output uncertainties
- Validation
  - check the model against experimental data
- Decision
  - identify the worst case scenario
  - decide whether worst case is acceptable

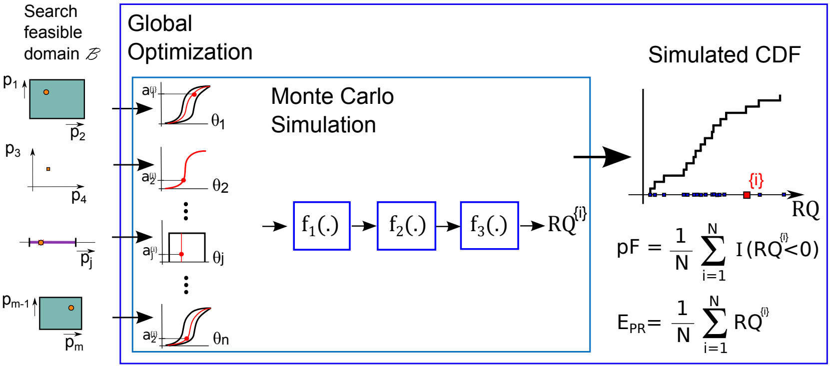
Aspects of propagation

Number of individual evaluations of the model needed depends on:

- Nonlinearity of the model
- Dependency structure between the input quantities
- Type of uncertainty, *aleatory*, *epistemic*, or *mixed*
- The numerical method used to perform the mapping

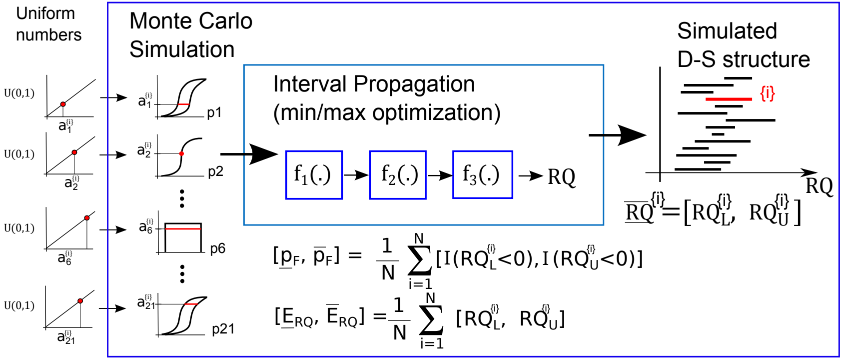
Parametric approach

Standard approach



Non-parametric approach

Counter approach



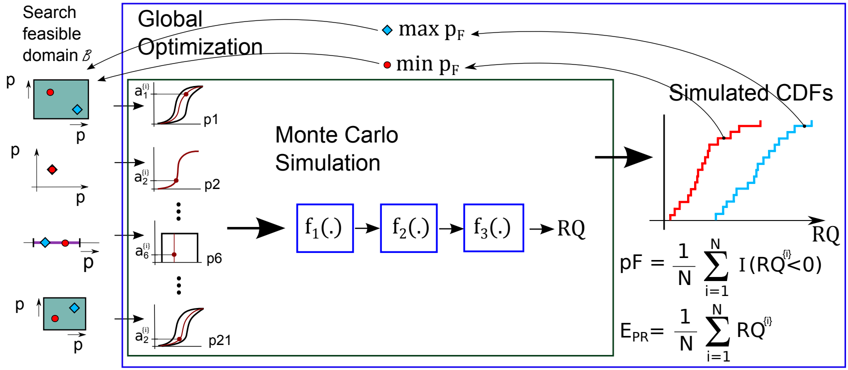
$$\overline{\underline{RQ}}^{\{i\}} = [RQ_L^{\{i\}},\ RQ_U^{\{i\}}]$$

$$[\underline{p}_F,\ \overline{p}_F] = \frac{1}{N}\sum_{i=1}^{N}[I(RQ_L^{\{i\}}<0), I(RQ_U^{\{i\}}<0)]$$

$$[\underline{E}_{RQ},\ \overline{E}_{RQ}] = \frac{1}{N}\sum_{i=1}^{N}[RQ_L^{\{i\}},\ RQ_U^{\{i\}}]$$

Back propagation problem (Tracking problem)

Is it possible in a non-parametric approach?



Propagation of failure probabilities

Problem statement

Characterize the problem:

- What is the targeted failure probability?
- How many state variables?
- Is the response quantity (RQ) monotonic?
- Does the RQ display a single mode?

Propagation of failure probabilities

Problem statement

Characterize the problem:

- What is the targeted failure probability?
- How many state variables?
- Is the response quantity (RQ) monotonic?
- Does the RQ display a single mode?

Traditional reliability assessment:

$$\Theta_F = \{\boldsymbol{\theta} \in \Theta \mid g(\boldsymbol{\theta}) \leq 0\} \tag{1}$$

$$p_F = \int_{\Theta_F} h_{\mathcal{D}}(\boldsymbol{\theta}; \boldsymbol{p})\, d\Theta \tag{2}$$

Outline

Generalized numerical framework of uncertainties

A bounded set $\mathcal{B}$ is defined by a vector of intervals $\underline{\overline{x}}_i,\ i = 1, ..., n_x$ and a dependence function $\Phi(x)$

$$\mathcal{B} = \times_{i=1}^{n_x} [\underline{\Phi}(x_i),\ \overline{\Phi}(x_i)] \tag{3}$$

A credal set $\mathcal{C}$ is the set of distribution functions

$$\mathcal{C} = \{h_{\mathcal{D}}\,(\xi; \boldsymbol{p}) \mid \boldsymbol{p} \in \mathcal{B}_p\}\,,\ \mathcal{B}_p = \times_i^{n_p}\ [\underline{p_i},\ \overline{p_i}] \tag{4}$$

Generalized numerical framework of uncertainties

A bounded set $\mathcal{B}$ is defined by a vector of intervals $\overline{\underline{x}}_i, \; i = 1, ..., n_x$ and a dependence function $\Phi(x)$

$$\mathcal{B} = \times_{i=1}^{n_x} [\underline{\Phi}(x_i), \; \overline{\Phi}(x_i)] \tag{3}$$

A credal set $\mathcal{C}$ is the set of distribution functions

$$\mathcal{C} = \left\{ h_{\mathcal{D}} \left(\xi; \boldsymbol{p} \right) \;\middle|\; \boldsymbol{p} \in \mathcal{B}_p \right\}, \; \mathcal{B}_p = \times_i^{n_p} \; [\underline{p_i}, \; \overline{p_i}] \tag{4}$$

Redefinition of failure domain:

$$\Theta_F = \Omega_F \times X_F \tag{5}$$

where,

$$\Omega_F(\boldsymbol{x}) = \left\{ \boldsymbol{\xi} \in \mathbb{R}^{n_\xi} \mid g(\boldsymbol{\xi}, \boldsymbol{x}) \leq 0 \right\}, \tag{6}$$

$$X_F(\boldsymbol{\xi}) = \left\{ \boldsymbol{x} \in \mathbb{R}^{n_x} \mid g(\boldsymbol{\xi}, \boldsymbol{x}) \leq 0 \right\}. \tag{7}$$

Failure probability

Lower and upper bounds

Lower and upper bounds of failure probability are

$$\underline{p}_F(\mathcal{C}, \mathcal{B}_x) = \inf_{\boldsymbol{x}\in\mathcal{B}_x \; \boldsymbol{p}\in\mathcal{B}_p} \int_{\Omega_F(\boldsymbol{x})} h_{\mathcal{D}}(\xi; \boldsymbol{p}) \; d\Omega; \tag{8}$$

$$\overline{p}_F(\mathcal{C}, \mathcal{B}_x) = \sup_{\boldsymbol{x}\in\mathcal{B}_x \; \boldsymbol{p}\in\mathcal{B}_p} \int_{\Omega_F(\boldsymbol{x})} h_{\mathcal{D}}(\xi; \boldsymbol{p}) \; d\Omega. \tag{9}$$

Failure probability

Conjugate relationship

When the uncertainty set $\mathcal{M}$ is restricted to $\mathcal{C}$ only, i.e. $\mathcal{M} = \mathcal{C}$, the probability function h° that yields the lower bound $\underline{p}_F$ satisfies

$$\int_{\Omega_F} h^{\circ}_{\mathcal{D}}(\xi)\; d\Omega + \int_{\Omega_S} h^{\circ}_{\mathcal{D}}(\xi)\; d\Omega = 1, \tag{10}$$

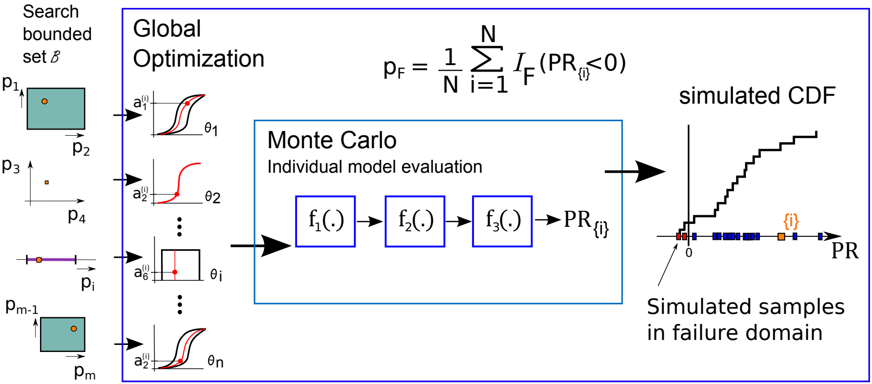
where, $\Omega_F \cup \Omega_S = \Omega$. Then, Eq. (9) establishes a *conjugate* (dual) relationship

$$\overline{p}(\Omega_S) = 1 - \underline{p}(\Omega_F), \tag{11}$$

with $\overline{p}(\Omega_S) = \int_{\Omega_S} h^{\circ}_{\mathcal{D}}(\xi)\; d\Omega$ and $\underline{p}(\Omega_F) = \int_{\Omega_F} h^{\circ}_{\mathcal{D}}(\xi)\; d\Omega$.

Global search for lower and upper bounds

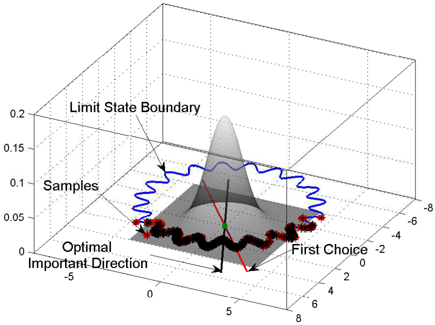
Nested loops



Proposed strategy

Epistemic propagation of failure probabilities

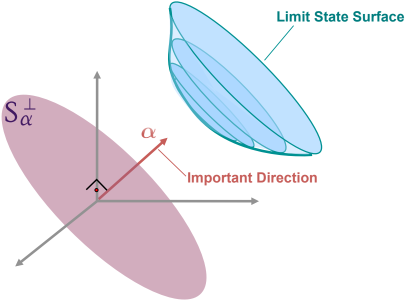
- Advanced Line Sampling
  - Efficient simulation method
  - Adaptive algorithm
- Driven optimization process
  - Exploit an averaged important direction
  - Identify the conjugate states



de Angelis, Patelli, Beer. "Advanced line sampling for efficient robust reliability analysis." Structural Safety, submitted, 2014

Line Sampling

- Important direction, α
- Generate points on the hyperplane
- Samples along lines
- Approximate point on the limit state

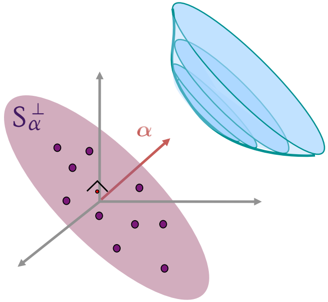


$$P_F = \int \mathbb{I}_{\mathcal{F}}(\mathbf{x})\, h_{\mathcal{N}}(\mathbf{x})\, d\mathbf{x} = \int_{n-1} \prod_{j=1}^{n-1} \phi(x_j) dx_j \int \mathbb{I}_{\mathcal{F}} \phi(x_n) dx_n$$

$$\hat{P_F} = \frac{1}{N_L} \sum_{j=1}^{N_L} p_F^{(j)} = \frac{1}{N_L} \sum_{j=1}^{N_L} \Phi(-\bar{\ell}^{(j)})$$

Line Sampling

- Important direction, α
- Generate points on the hyperplane
- Samples along lines
- Approximate point on the limit state

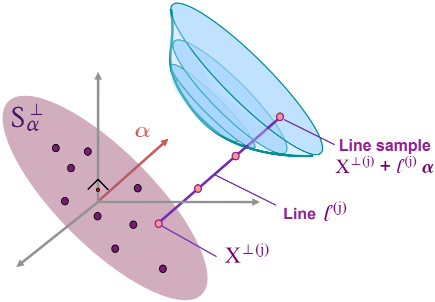


$$P_F = \int \mathbb{I}_{\mathcal{F}}(\mathbf{x})\, h_{\mathcal{N}}(\mathbf{x})\, d\mathbf{x} = \int_{n-1} \prod_{j=1}^{n-1} \phi(x_j) dx_j \int \mathbb{I}_{\mathcal{F}} \phi(x_n) dx_n$$

$$\hat{P_F} = \frac{1}{N_L} \sum_{j=1}^{N_L} p_F^{(j)} = \frac{1}{N_L} \sum_{j=1}^{N_L} \Phi(-\bar{\ell}^{(j)})$$

Line Sampling

- Important direction, α
- Generate points on the hyperplane
- Samples along lines
- Approximate point on the limit state

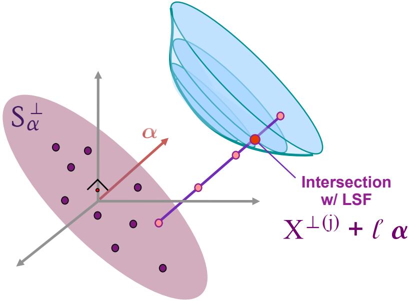


$$P_F = \int \mathbb{I}_{\mathcal{F}}(\mathbf{x})\, h_{\mathcal{N}}(\mathbf{x})\, d\mathbf{x} = \int_{n-1} \prod_{j=1}^{n-1} \phi(x_j) dx_j \int \mathbb{I}_{\mathcal{F}} \phi(x_n) dx_n$$

$$\hat{P}_F = \frac{1}{N_L} \sum_{j=1}^{N_L} p_F^{(j)} = \frac{1}{N_L} \sum_{j=1}^{N_L} \Phi(-\bar{\ell}^{(j)})$$

Line Sampling

- Important direction, α
- Generate points on the hyperplane
- Samples along lines
- Approximate point on the limit state



$$P_F = \int \mathbb{I}_{\mathcal{F}}(\mathbf{x})\ h_{\mathcal{N}}(\mathbf{x})\ d\mathbf{x} = \int_{n-1} \prod_{j=1}^{n-1} \phi(x_j) dx_j \int \mathbb{I}_{\mathcal{F}} \phi(x_n) dx_n$$

$$\hat{P}_F = \frac{1}{N_L} \sum_{j=1}^{N_L} p_F^{(j)} = \frac{1}{N_L} \sum_{j=1}^{N_L} \Phi(-\bar{\ell}^{(j)})$$

The search in $\mathcal{B}_p$

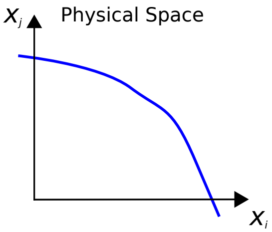
Imprecision in the distribution parameters

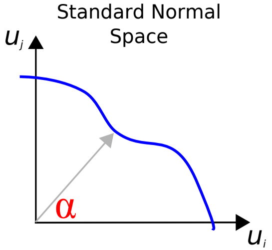
- Aim is identifying the elements of $\mathcal{B}_p$ producing lower and upper bounds
- The failure domain Θ_F does not change as the search is performed in $\mathcal{B}_p$
- An averaged important direction can be set for each iteration
- The failure domain is invariant to the uncertainty set $\mathcal{M}$
- The adaptive algorithm updates the important direction to increase the accuracy of estimations

The search in $\mathcal{B}_p$

Imprecision in the distribution parameters

- Aim is identifying the elements of $\mathcal{B}_p$ producing lower and upper bounds
- The failure domain Θ_F does not change as the search is performed in $\mathcal{B}_p$
- An averaged important direction can be set for each iteration
- The failure domain is invariant to the uncertainty set $\mathcal{M}$
- The adaptive algorithm updates the important direction to increase the accuracy of estimations

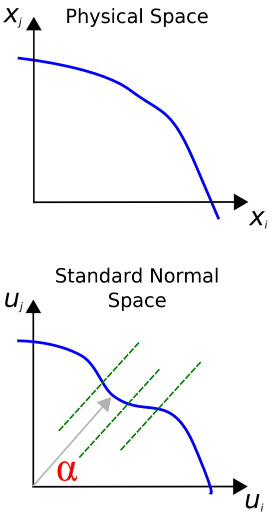




The search in $\mathcal{B}_p$

Imprecision in the distribution parameters

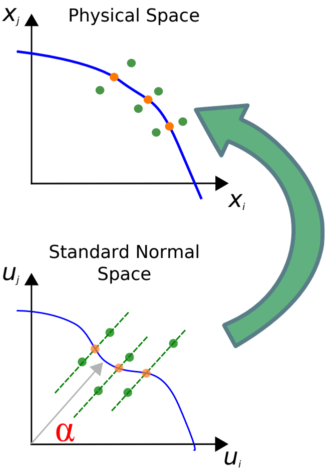
- Aim is identifying the elements of $\mathcal{B}_p$ producing lower and upper bounds
- The failure domain Θ_F does not change as the search is performed in $\mathcal{B}_p$
- An averaged important direction can be set for each iteration
- The failure domain is invariant to the uncertainty set $\mathcal{M}$
- The adaptive algorithm updates the important direction to increase the accuracy of estimations



The search in $\mathcal{B}_p$

Imprecision in the distribution parameters

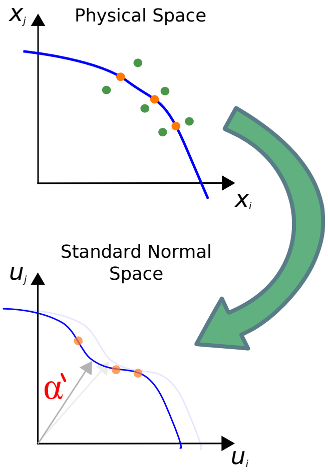
- Aim is identifying the elements of $\mathcal{B}_p$ producing lower and upper bounds
- The failure domain Θ_F does not change as the search is performed in $\mathcal{B}_p$
- An averaged important direction can be set for each iteration
- The failure domain is invariant to the uncertainty set $\mathcal{M}$
- The adaptive algorithm updates the important direction to increase the accuracy of estimations



The search in $\mathcal{B}_p$

Imprecision in the distribution parameters

- Aim is identifying the elements of $\mathcal{B}_p$ producing lower and upper bounds
- The failure domain Θ_F does not change as the search is performed in $\mathcal{B}_p$
- An averaged important direction can be set for each iteration
- The failure domain is invariant to the uncertainty set $\mathcal{M}$
- The adaptive algorithm updates the important direction to increase the accuracy of estimations



The search in $\mathcal{B}_x$

Imprecision in the structural parameters

- Recall the definition of failure domain from Eq. (4):

$$\Omega_F(\boldsymbol{x}) = \{\boldsymbol{\xi} \in \mathbb{R}^{n_\xi} \mid g(\boldsymbol{\xi}, \boldsymbol{x}) \leq 0\},$$

- Intervals make the failure domain no longer invariant to the uncertainty set $\mathcal{M}$,
- For the sake of searching in $\mathcal{B}_x$, assume the intervals are imprecise Gaussian random variables

$$\underline{\overline{\boldsymbol{x}}} \rightarrow \boldsymbol{\eta} \in \mathcal{C}_x = \left\{ h_{\mathcal{N}}(\boldsymbol{\eta}; \underline{\overline{\boldsymbol{\mu}}}_x, \boldsymbol{\sigma}_x) \mid \underline{\overline{\boldsymbol{\mu}}}_x = \underline{\overline{\boldsymbol{x}}},\ \boldsymbol{\sigma}_x \in [0, \boldsymbol{x}_r] \right\}, \quad (12)$$

- The uncertainty set is now the credal set $\mathcal{M}' = \mathcal{C} \cup \mathcal{C}_x$

The search in $\mathcal{B}_x$

Imprecision in the structural parameters

- The search in $\mathcal{M}'$ allows to identify the conjugate states

$$\boldsymbol{\mu}_* = \arg\min_{\mathcal{M}'} p_F, \quad \boldsymbol{\mu}^* = \arg\max_{\mathcal{M}'} p_F \tag{13}$$

- To these states it is associated the corresponding argument minimum and maximum to be held within the intervals

$$(\boldsymbol{\mu}_{x*}, \ \boldsymbol{\mu}_x^*) \rightarrow (\boldsymbol{x}_*, \ \boldsymbol{x}^*) \tag{14}$$

- As the argument optima $(\boldsymbol{x}_*, \ \boldsymbol{x}^*)$ are identified in $\mathcal{B}_x$, just two more reliability analyses are needed to estimate lower and upper failure probabilities.

Outline

Linear performance function with noise

case (a): description

$$
\begin{aligned}
g(\boldsymbol{\xi}, \boldsymbol{x}) &= -500 + \xi_1 + 2\xi_2 + 2\xi_3 + \xi_4 - 5x_1 - 5x_2 + \\
&\quad + 0.001 \sum_{i=1}^{4} \sin(100\xi_i) + 0.001 \sum_{j=1}^{2} \sin(100x_j);
\end{aligned}
$$

SV	Symbol	Uncert. type	Mean/Interval	Stand. dev.
1	ξ_1	LN($\underline{\overline{\mu_1}}$, $\underline{\overline{\sigma_1}}$)	$\underline{\overline{\mu_1}} = [110,\ 125]$	$\underline{\overline{\sigma_1}} = [10,\ 14]$
2	ξ_2	LN($\underline{\overline{\mu_2}}$, $\underline{\overline{\sigma_2}}$)	$\underline{\overline{\mu_2}} = [115,\ 130]$	$\underline{\overline{\sigma_2}} = [10,\ 14]$
3	ξ_3	LN($\underline{\overline{\mu_3}}$, $\underline{\overline{\sigma_3}}$)	$\underline{\overline{\mu_3}} = [115,\ 130]$	$\underline{\overline{\sigma_3}} = [10,\ 14]$
4	ξ_4	LN($\underline{\overline{\mu_4}}$, $\underline{\overline{\sigma_4}}$)	$\underline{\overline{\mu_4}} = [115,\ 130]$	$\underline{\overline{\sigma_4}} = [10,\ 14]$
5	x_1	Interval $\underline{\overline{x_1}}$	$\underline{\overline{x_1}} = [45,\ 52]$	-
6	x_2	Interval $\underline{\overline{x_2}}$	$\underline{\overline{x_2}} = [35,\ 43]$	-

Linear performance function with noise

case (a): solution

The sign of the averaged important direction

$$\mathrm{sign}(\alpha) = (-1, -1, -1, -1, 1, 1),$$

allows us to identify the argument optima

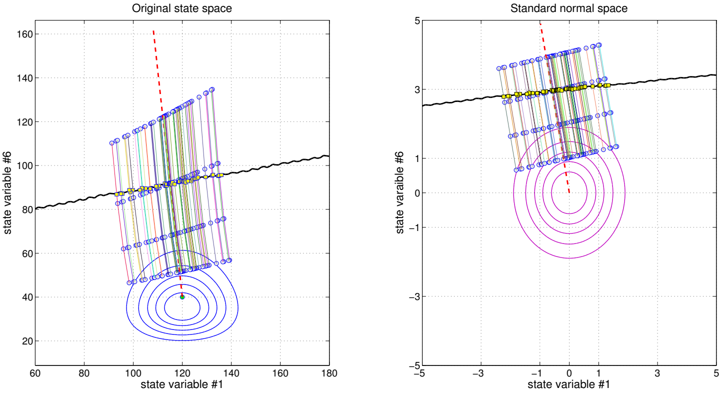
$$\underset{\boldsymbol{p}\in\mathcal{B}_\xi \; \boldsymbol{x}\in\mathcal{B}_x}{\arg\min} \; p_F = (\overline{\mu_1}, \overline{\sigma_1}, \overline{\mu_2}, \overline{\sigma_2}, \overline{\mu_3}, \overline{\sigma_3}, \overline{\mu_4}, \overline{\sigma_4}, \underline{x_1}, \underline{x_2}),$$

$$\underset{\boldsymbol{p}\in\mathcal{B}_\xi \; \boldsymbol{x}\in\mathcal{B}_x}{\arg\max} \; p_F = (\underline{\mu_1}, \underline{\sigma_1}, \underline{\mu_2}, \underline{\sigma_2}, \underline{\mu_3}, \underline{\sigma_3}, \underline{\mu_4}, \underline{\sigma_4}, \overline{x_1}, \overline{x_2}).$$

Here the search domain has 10 dimensions, thus more than 1024 reliability analyses (iterations) would have been required to find an approximation of the failure probability bounds with a full approach.

Linear performance function with noise

case (a): results



Linear performance function with noise

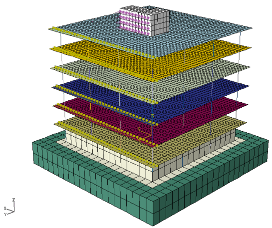
case (a): results

- Approach A: proposed strategy
- Approach B: full approach based on Line Sampling and Latin Hypercube Sampling (LHS)

Approach A		Approach B (LHS)	
$\overline{p_F}$	N_s	$\overline{p_F}$	N_s
$[1.4\ 10^{-10},\ 0.43]$	252	$[3.2\ 10^{-6},\ 8.4\ 10^{-2}]$	2.1×10^6

Six-level building

imprecision in distribution parameters

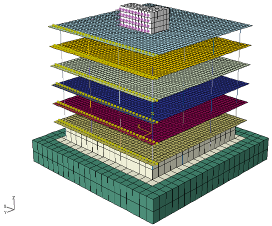


- 244 State Variables, 8200 elements, 66300 degrees of freedom
- Columns' depth and breadth are the interval $[0.36,\ 0.44]$m
- Performance function: $g(\boldsymbol{\theta}) = \left|\sigma_I(\boldsymbol{\theta}) - \sigma_{III}(\boldsymbol{\theta})\right|/2 - \sigma_y,$

- Investigate sensitivity of failure probabilities against imprecision
- Identify the worst case scenario
- Robust realibility analysis

Six-level building

imprecision in distribution parameters

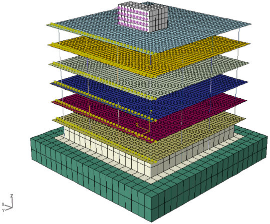


- 244 State Variables, 8200 elements, 66300 degrees of freedom
- Columns' depth and breadth are the interval $[0.36,\ 0.44]$m
- Performance function: $g(\boldsymbol{\theta}) = \left|\sigma_I(\boldsymbol{\theta}) - \sigma_{III}(\boldsymbol{\theta})\right|/2 - \sigma_y,$

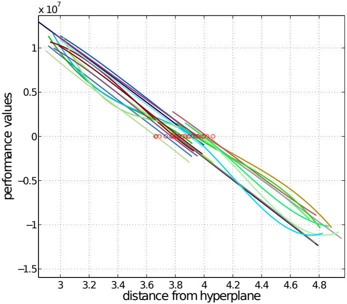
- Investigate sensitivity of failure probabilities against imprecision
- Identify the worst case scenario
- Robust realibility analysis

Six-level building

imprecision in distribution parameters



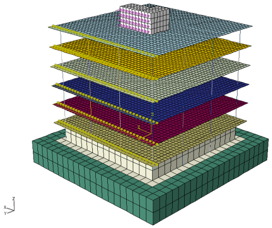
- 244 State Variables, 8200 elements, 66300 degrees of freedom
- Columns' depth and breadth are the interval $[0.36,\ 0.44]$m
- Performance function:
  $g(\boldsymbol{\theta}) = \left|\sigma_I(\boldsymbol{\theta}) - \sigma_{III}(\boldsymbol{\theta})\right|/2 - \sigma_y,$



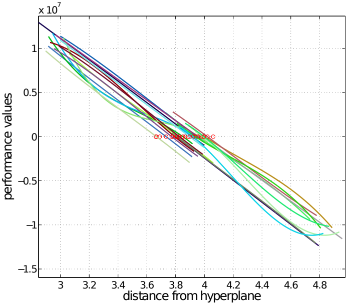
- Investigate sensitivity of failure probabilities against imprecision
- Identify the worst case scenario
- Robust realibility analysis

Six-level building

imprecision in distribution parameters



- 244 State Variables, 8200 elements, 66300 degrees of freedom
- Columns' depth and breadth are the interval [0.36, 0.44]m
- Performance function:
  $g(\boldsymbol{\theta}) = \left|\sigma_I(\boldsymbol{\theta}) - \sigma_{III}(\boldsymbol{\theta})\right|/2 - \sigma_y,$



Objectives:

- Investigate sensitivity of failure probabilities against imprecision
- Identify the worst case scenario
- Robust realibility analysis

Large scale finite element model

Uncertainty set $\mathcal{M}_{II} = \mathcal{C} \cup \mathcal{B}_x$

$$\mathcal{C}\left\{h_{\mathcal{D}}\left(\zeta;\boldsymbol{p}\right) \mid \boldsymbol{p} \in \mathbb{R}^{104},\ \boldsymbol{p} \in \mathcal{B}_{\xi}\right\}, \quad \mathcal{B}_x = \times_i^{192}\underline{\overline{x}}_i.$$

SV	Probability dist.	Distribution	Description	Units
1	$N(0.1,\ 10^{-4})$	Normal	Column's strength	GPa
2 – 193	Unif$(0.36, 0.44)$	Uniform	Sections size	m
194 – 212	$LN(35.0,\ 12.25)$	Lognormal	Young's modulus	GPa
213 – 231	$LN(2.5,\ 6.25\ 10^{-2})$	Lognormal	Material's density	kg/dm^3
232 – 244	$LN(0.25,\ 6.25\ 10^{-4})$	Lognormal	Poisson's ratio	-

Large scale finite element model

Uncertainty set $\mathcal{M}_{II} = \mathcal{C} \cup \mathcal{B}_x$

$$\mathcal{C}\left\{h_{\mathcal{D}}\left(\zeta;\boldsymbol{p}\right)\mid \boldsymbol{p}\in\mathbb{R}^{104},\ \boldsymbol{p}\in\mathcal{B}_{\xi}\right\},\quad \mathcal{B}_x = \times_i^{192}\underline{\overline{x}}_i.$$

SV	Uncertainties type		$\underline{\overline{p}} = p_c\ [1-\epsilon,\ 1+\epsilon],\ \underline{\overline{x}} = [\underline{x},\ \overline{x}]$	
1	distribution	$\mathrm{N}(\underline{\overline{\mu}},\ \underline{\overline{\sigma}}^2)$	$\mu_c = 0.1$	$\sigma_c = 0.01$
$2-193$	interval	$\underline{\overline{x}}$	$\underline{x} = 0.36$	$\overline{x} = 0.44$
$194-212$	distribution	$\mathrm{LN}(\underline{\overline{m}},\ \underline{\overline{v}})$	$m_c = 35$	$v_c = 12.25$
$213-231$	distribution	$\mathrm{LN}(\underline{\overline{m}},\ \underline{\overline{v}})$	$m_c = 2.5$	$v_c = 6.25\ 10^{-2}$
$232-244$	distribution	$\mathrm{LN}(\underline{\overline{m}},\ \underline{\overline{v}})$	$m_c = 0.25$	$v_c = 6.25\ 10^{-4}$

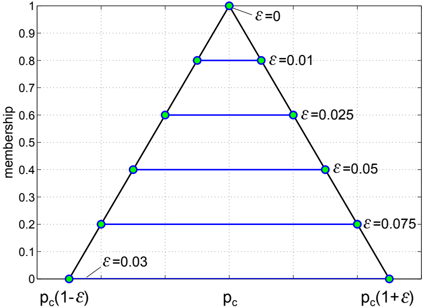
Large scale finite element model

Results

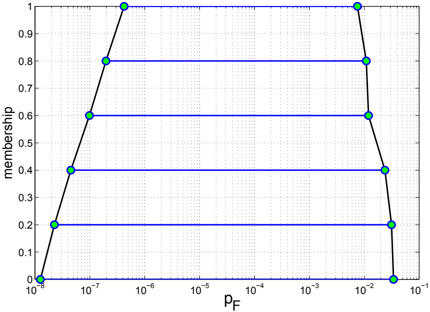
	Lower Bound		Upper Bound		
ϵ	$\underline{p}_F$	CoV	$\overline{p}_F$	CoV	Ns
0.000	$4.70\ 10^{-7}$	$10.2\ 10^{-2}$	$6.73\ 10^{-3}$	$11.5\ 10^{-2}$	259
0.010	$2.28\ 10^{-7}$	$13.4\ 10^{-2}$	$9.71\ 10^{-3}$	$12.2\ 10^{-2}$	247
0.015	$1.10\ 10^{-7}$	$10.3\ 10^{-2}$	$1.11\ 10^{-2}$	$7.6\ 10^{-2}$	255
0.020	$5.19\ 10^{-8}$	$13.1\ 10^{-2}$	$2.08\ 10^{-2}$	$14.6\ 10^{-2}$	255
0.025	$2.51\ 10^{-8}$	$9.97\ 10^{-2}$	$2.72\ 10^{-2}$	$15.3\ 10^{-2}$	249
0.030	$1.40\ 10^{-8}$	$9.94\ 10^{-2}$	$3.21\ 10^{-2}$	$6.5\ 10^{-2}$	254

Large scale finite element model

Results



(a)



(b)

Outline

Final remarks

- A strategy to propagate epistemic uncertainty with the failure probability was proposed
- *Bounded* and *Credal sets* are formulated as a sound way to account for epistemic uncertainty in a parametric sense
- With the proposed numerical strategy, uncertainty propagation's efficiency can be significantly increased
- When the underlying model displays monotonic, time of the analysis is comparable to a single Monte Carlo
- The strategy is generally applicable so far as the model displays a single failure mode
- The solution strategy is integrated in the general purpose software **Open Cossan**