Advanced Line Sampling for Efficient Robust Reliability Analysis

Marco de Angelis*, Edoardo Patelli, Michael Beer

Institute for Risk and Uncertainty, University of Liverpool

Liverpool, United Kingdom

Abstract

In this paper, a novel strategy is designed to efficiently estimate set-valued failure probabilities, coupling Monte Carlo sampling-based with optimization methods. The notion of uncertainty is generalized to include both aleatory and epistemic uncertainties, and capture gaps of knowledge and scarcity of data. The proposed formulation of the generalized uncertainty model allows for sets of probability distribution functions, also known as credal sets, and sets of bounded variables. An Advanced Line Sampling method is developed and combined with the generalized uncertainty model, not only to reduce the time needed for a single reliability analysis, but also to increase the efficiency of the search for lower and upper bounds of the failure probability. The proposed strategy knocks down the computational barrier of computing interval failure probabilities, and reduces the cost of a robust reliability analysis by many orders of magnitude. The solution strategy is integrated into the open-source software for uncertainty quantification and risk analysis OpenCossan, allowing its application on large-scale engineering problems as well as broadening its spectrum of potential applications. The efficiency and applicability of the developed method is demonstrated via numerical examples.

Keywords: Structural Reliability, Monte Carlo Methods, Line Sampling, Credal Sets, Bounded Sets, Set-valued Probabilities, Generalized Uncertainty

*Corresponding author

Email address: marco.de-angelis@liverpool.ac.uk (Marco de Angelis)

1. Introduction

Engineering structures and systems, such as bridges, buildings, aircraft, offshore platforms, nuclear power plants, transmission towers and pipelines, are designed to fulfil specific requirements, and they should be able to deal with possible changes of loads and conditions. However, the design context is often characterized by partial knowledge and limited access to information. In such a context, in order to be able to bypass the difficulties in quantifying vague information, decisions often rely on experts opinions, rather than rigorous analyses. In this paper a generalized model of uncertainty is proposed and used for reliability assessment to address this issue. Within this model the risk is evaluated treating gaps of knowledge and scarcity of data as a key source of uncertainty. To translate this into practice, computational models that consider imprecision are proposed.

Risk is conventionally expressed as the product between the failure probability of the system and the consequences caused by the system's failure. While the consequences are quantified in monetary units, the failure probability is calculated, within a reliability assessment, in a rigorous probabilistic framework [1]. Commonly, this requires the specification of precise distributional models (of probability), including dependencies for the input variables.

Among the numerical methods proposed to assess reliability, simulation methods [2] have attracted significant attention. Simulation methods are generally applicable, but require a compromise between efficiency and accuracy. Simulation methods proposed in literature include Monte Carlo Simulation [3, 4], Importance Sampling [5, 6], Directional Sampling [7, 8], Line Sampling [9, 10], Subset Simulation [11] etc. The individual developments possess different performance features for different classes of problems. Herein, we target at high numerical efficiency assuming that the limit state surfaces only show moderate non-linearities. Since the latter applies to the majority of practical cases [12],

this does not impose a strong restriction. Hence, Line Sampling is selected as the basis for our development, which has been extended to deal with generalized probabilistic models.

The paper is organized as follows: In Section 2 a generalized uncertainty model is introduced. In Section 3, an Advanced Line Sampling method and adaptive algorithm is developed and, in Section 4, it is implemented in the generalized uncertainty framework. In Section 5, the integration in the general computational toolbox OpenCossan is briefly explained. In Section 6, numerical examples are given to demonstrate the efficiency of the method. Final remarks and conclusions are provided in Section 7.

2. A generalized uncertainty model

Traditionally, the assessment of structural reliability is based on well-defined (precise) probabilistic models [1]. Probabilistic models are constructed from data that, in a design context, are often scarce and not available to a sufficient extent [13]. In such a context, it is advisable to relax the assumption of a precise probabilistic model. A detailed reasoning and discussion in this direction with an overview on available generalized models is provided in [14].

In essence, a generalized model of uncertainty shall allow for imprecision in both the state variables of the system, denoted as $\boldsymbol{\theta}$, and the parameters of the probabilistic model $\boldsymbol{p}$. Depending on the amount of information available about the variables and parameters, imprecision can be modelled in different ways, for example by means of Intervals [15, 16, 17], Convex Models [18] and Fuzzy Sets [19, 20]. Intervals are used when variables are only known to be bounded within lower and upper limits, while Convex Models are used when variables are known to be bounded and also show some dependences. Fuzzy Sets allow the simultaneous analysis of different bounded sets, which is helpful if the bounds are not known precisely and to explore sensitivity with respect to the bounds of the inputs. When imprecision is also present within the probabilistic model, the uncertainty model shall include sets of probability distribution functions.

This is the case, for example, when statistical distributions given along with their confidence intervals are considered, or when data complying with several statistical distribution models are processed. Credal Sets [21] provide a quite general pathway to express and analyse sets of probability distributions. Thus, we utilize credal sets in combination with bounded sets for the subsequent development.

2.1. Credal sets and bounded sets

The generalized model of uncertainty, denoted by $\mathcal{M}$, defines type and extent of uncertainty in the state variables $\boldsymbol{\theta}$. The model $\mathcal{M}$ may represent a credal set $\mathcal{C}$ (see e.g. [21]), a bounded (interval) set $\mathcal{Q}$ (see e.g. [18]), or both at the same time.

A credal set of category I, namely $\mathcal{C}_I$, is a set of probability distribution functions, where the imprecision is defined in the distribution parameters. A credal set of category II, $\mathcal{C}_{II}$, is a set of probability distribution functions, where the imprecision is in the type of distribution functions (e.g. Normal, Log-normal, Gamma, Beta), whilst a credal set of category III, $\mathcal{C}_{III}$, has both imprecise distribution parameters and function type.

A bounded set of category I, namely $\mathcal{Q}_I$, is obtained by the Cartesian product $\bigtimes_i^b \underline{\overline{x}}_i$ of interval variables $\underline{\overline{x}}_i = \{x_i \mid x_i \in [\underline{x}_i,\ \overline{x}_i] \subset \mathbb{R}\}$, $i = 1, ..., b$. A bounded set of category II, namely $\mathcal{Q}_{II}$, is obtained from interval variables $\underline{\overline{x}}_i$, taking into account dependencies between the variables. This may be done in many different ways, for example using convex sets, i.e. by constructing the enclosing ellipsoid (see e.g. [22]), or using other types of sets (e.g. convex hulls, polytopes).

In this paper, without limiting generality but providing a basic development, only uncertainties where the imprecision is of category I are considered, thus the sets $\mathcal{C}_I$ and $\mathcal{Q}_I$ will be simply denoted as $\mathcal{C}$ and $\mathcal{Q}$ respectively.

2.2. Problem formulation

In performance-based engineering the structural system is considered as a collection of performance variables $g_i, i = 1, 2, ..., N_g$, which are functions of

the state variables $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^n$ (see e.g. [23]). Typically, the state variables are the inputs that defines the structural system, such as material properties, shape and size of structural elements, and load magnitudes, whilst the performances express specific structural responses, such as frequency and amplitude of vibrations, stresses, deflections and so forth.

The performance function $g : \mathbb{R}^n \to g_i \in \mathbb{R}$ maps values from the state space Θ to the performance variables of interest. For given criteria on the performance variables, g defines the failure domain $\Theta_F = \{\boldsymbol{\theta} \in \Theta \mid g(\boldsymbol{\theta}) \leq 0\}$, which is identified by the limit state surface $\tilde{\Theta} = \{\boldsymbol{\theta} \mid g(\boldsymbol{\theta}) = 0\}$. Points $\tilde{\boldsymbol{\theta}}$ on the limit state surface are referred to as limit state points. An important feature for our development is that the limit state is invariant to the uncertainty model $\mathcal{M}$, because it is intrinsic in the structural system, i.e. depends solely on the performance function g. The uncertainty model only determines the probability over the state space, but does not influence location of limit state points $\tilde{\boldsymbol{\theta}}$.

In our study, $\mathcal{M}$ is represented by both credal sets and bounded sets of category I. In the credal set $\mathcal{C}$, imprecision is considered in the distribution parameters of n_1 imprecise random variables, which is expressed with the bounded set $\mathcal{Q}_1$ of distribution parameters. A second bounded set $\mathcal{Q}_2$ is used to describe imprecision in the structural parameters, which are not associated with any distribution model.

The state variables are, thus, split into n_1 imprecise random variables $\boldsymbol{\xi} \in \Omega \subseteq \mathbb{R}^{n_1}$ belonging to $\mathcal{C}$ and n_2 interval variables $\boldsymbol{x} \in X \in \mathbb{R}^{n_2}$ belonging to $\mathcal{Q}_2$, where $n_1 + n_2 = n$. The credal set is defined as $\mathcal{C} = \{h_{\mathcal{D}}(\boldsymbol{\xi}; \boldsymbol{p}) \mid \boldsymbol{p} \in \mathbb{R}^m,\ \boldsymbol{p} \in \mathcal{Q}_1\}$, where $\mathcal{Q}_1$ is the bounded set $\mathcal{Q}_1 = \bigtimes_i^m [\underline{p_i},\ \overline{p_i}]$, $h_{\mathcal{D}}$ is the joint probability distribution function of random variables $\boldsymbol{\xi}$, $\mathcal{D}$ is the distributional model, and $\boldsymbol{p}$ are the distributional parameters. The bounded set of the remaining state variables is expressed as the Cartesian product $\mathcal{Q}_2 = \bigtimes_i^{n_2} [\underline{x_i},\ \overline{x_i}]$.

2.3. Failure probability for generalized uncertainty

When the uncertainty model comprises only precisely defined probability distributions, i.e. $m = 0$, $b = 0$ and $\mathcal{C}$ degenerates in one distribution function,

structural reliability is assessed in terms of a precise failure probability. Precise measures of failure probability are obtained as $p(\Theta_F, \mathcal{D}, \boldsymbol{p}) = \int_{\Theta_F} h_{\mathcal{D}}(\boldsymbol{\xi}; \boldsymbol{p})\, d\Theta$, where $d\Theta$ is the Lebesgue measure of an elementary portion of Θ. For simplicity $p(\Theta_F, \mathcal{D}, \boldsymbol{p})$ is subsequently denoted by p_F. Operating with the generalized uncertainty model $\mathcal{M}$ leads to imprecise measures of failure probability. The failure domain is split into two separate domains as $\Theta_F = \Omega_F \times X_F$, where $\Omega_F(\boldsymbol{x}) = \{\boldsymbol{\xi} \in \mathbb{R}^{n_1} \mid g(\boldsymbol{\xi}, \boldsymbol{x}) \leq 0\}$ and $X_F(\boldsymbol{\xi}) = \{\boldsymbol{x} \in \mathbb{R}^{n_2} \mid g(\boldsymbol{\xi}, \boldsymbol{x}) \leq 0\}$. Provided the definition of $\mathcal{C}$, the imprecise failure probability is expressed as the interval $\underline{\overline{p}}_F(\mathcal{C}, \mathcal{Q}_2) = \Big[\underline{p}_F(\mathcal{C}, \mathcal{Q}_2),\ \overline{p}_F(\mathcal{C}, \mathcal{Q}_2)\Big]$. The lower and upper bound of the imprecise failure probability are

$$\underline{p}_F(\mathcal{C}, \mathcal{Q}_2) = \inf_{\boldsymbol{x} \in \mathcal{Q}_2} \inf_{\boldsymbol{p} \in \mathcal{Q}_1} \int_{\Omega_F(\boldsymbol{x})} h_{\mathcal{D}}(\boldsymbol{\xi}; \boldsymbol{p})\, d\Omega; \qquad \overline{p}_F(\mathcal{C}, \mathcal{Q}_2) = \sup_{\boldsymbol{x} \in \mathcal{Q}_2} \sup_{\boldsymbol{p} \in \mathcal{Q}_1} \int_{\Omega_F(\boldsymbol{x})} h_{\mathcal{D}}(\boldsymbol{\xi}; \boldsymbol{p})\, d\Omega, \tag{1}$$

where, the order to which the operations of infimum and supremum are performed can be changed. The inner operand searches the bounds of p_F within $\mathcal{C}$, while the outer one searches the bounds of p_F within $\mathcal{Q}_2$.

Upper and lower bounds of failure and survival probabilities show a dual relationship. This can be seen clearly in the special case that the uncertainty model is restricted to $\mathcal{C}$ only. The probability function $h^{\circ}_{\mathcal{D}}$ that yields the lower bound $\underline{p}(\Omega_F)$, satisfies the equation $\int_{\Omega_F} h^{\circ}_{\mathcal{D}}(\boldsymbol{\xi})\, d\Omega + \int_{\Omega_S} h^{\circ}_{\mathcal{D}}(\boldsymbol{\xi})\, d\Omega = 1$, where Ω_S denotes the survival domain (complementary to the failure domain). Therefore, h° is also the function for which the upper bound $\overline{p}(\Omega_S)$ is obtained. Thus, the Equation $\underline{p}(\Omega_F) = 1 - \overline{p}(\Omega_S)$ establishes a dual (or conjugate) relationship between lower and upper probability functions. This relationship allows to identify the upper probability function when the lower probability function is known and vice versa. Note, however, that the complete function, which may also have an infinite support, is needed in order for the relationship to be used. From the definition of lower and upper probability follows that $\underline{p}_F \leq \overline{p}_F$. When $\mathcal{C}$ degenerates into a single probability distribution function, precise measures of probability $p_F = \underline{p}_F = \overline{p}_F$ are obtained.

3. Advanced Line Sampling

The computation of failure probabilities can be associated with quite a significant numerical effort. In cases where the number of random variables is high and the limit state surface is nonlinear, methods based on the computation of the Hessian become impractical. In these cases, advanced simulation methods, represent a useful alternative. Here, a new method that extends the concept of Line Sampling is presented. The method, named Advanced Line Sampling (ALS), does not only increase the efficiency of single reliability analysis but it proves to be essential for finding the lower and upper bound of the failure probability.

3.1. Concept of Line Sampling

Line Sampling, introduced in [9], and recently applied in [24], is an advanced simulation method developed to efficiently compute small failure probabilities for high dimensional problems. The method requires the knowledge of the so-called "important direction", $\boldsymbol{\alpha} \in \mathbb{R}^n$, which is defined as pointing towards the failure region. An initial approximation for the important direction is commonly obtained by computing the gradient of the performance function in the origin of the Standard Normal Space (SNS). Simulation methods estimate the failure probability by computing the integral

$$p_F = \int_{-\infty}^{\infty} \mathcal{I}_F(\boldsymbol{u})\ h_{\mathcal{N}}(\boldsymbol{u})\ d\boldsymbol{u}, \tag{2}$$

where, $\mathcal{I}_F : \mathbb{R}^n \to \{0, 1\}$ is the indicator function, $\boldsymbol{u} = T(\boldsymbol{\theta})$ are standard normal variables, $T : \mathbb{R}^n \to \mathbb{R}^n$ maps variables $\boldsymbol{\xi}$ from the original space to the SNS, and $h_{\mathcal{N}}(\boldsymbol{u}) = \prod_{i=1}^{n} \phi(u_i)$ is the standard normal PDF. Provided that $h_{\mathcal{N}}(\boldsymbol{u})$ is invariant to rotation of the coordinate axes, Equation 2 can be written in the form

$$p_F = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} \mathcal{I}_F(\boldsymbol{u}) \phi(u_1) du_1 \right) \prod_{i=2}^{n} \phi(u_i) du_i \tag{3}$$

for convenient evaluations. With u_1 pointing orthogonally towards the failure domain, the expansion $w(\boldsymbol{u}_{2:n}) = \int_{-\infty}^{\infty} \mathcal{I}_F(\boldsymbol{u})\ \phi(u_1)\ du_1$ from Equation 3 is a

function of the $n-1$ remaining standard variables $\boldsymbol{u}_{2:n} \in \mathbb{R}^{n-1}$ and provides a measure of likelihood for the variable $\boldsymbol{u}_{2:n}$ to be in the failure domain. All of the points with coordinates $\boldsymbol{u}^{\perp} = \{0, \boldsymbol{u}_{2:n}\}$ lie on the hyperplane orthogonal to the first coordinate u_1. Variable w can be calculated as $w\left(\boldsymbol{u}_{2:n}\right) = \Psi(F_1)$, where $\Psi(A) = \int_{-\infty}^{\infty} \mathcal{I}_A(y)\phi(y)dy$ is the Gaussian measure of a subset of $A \subset \mathbb{R}$. Let the scalar c^* be the smallest (in magnitude) value of the coordinate u_1 where the function $\mathcal{I}_F(\boldsymbol{u})$ steps from zero to one. This enables w to be approximately calculated as $w\left(\boldsymbol{u}_{2:n}\right) = \Phi(-|c^*|)$, where Φ is the standard normal CDF. Therefore, the failure probability can be obtained as the expected value

$$p_F = E[w\left(\boldsymbol{u}_{2:n}\right)] = \int_{-\infty}^{\infty} w\left(\boldsymbol{u}_{2:n}\right) \prod_{i=2}^{n} \phi(u_i) du_i. \tag{4}$$

Note that considering the standard normal CDF $\Phi(-|c^*|)$ in place of the Gaussian measure $\Psi(F_1)$, the probability w can only be overestimated, because it assumes that no further failure regions can be found on the line beyond c^*. LS provides an estimation of $E[w]$ by repeatedly generating points $\boldsymbol{u}_{2:n}$ from the standard normal PDF in $\mathbb{R}^{n-1}$, and computing the respective partial probabilities $w\left(\boldsymbol{u}_{2:n}\right)$. For example, generating N_L points $\boldsymbol{u}_{2:n}^{\{j\}},\ j = 1, 2, .., N_L$, an estimate of the failure probability is obtained computing the average

$$\hat{p}_F = \frac{1}{N_L} \sum_{j}^{N_L} w(\boldsymbol{u}_{2:n}^{\{j\}}). \tag{5}$$

Despite the important direction $\boldsymbol{\alpha}$ is not oriented as the first coordinate u_1, the above integrals can still be calculated exploiting the geometric features of the SNS. Standard normal points on the hyperplane orthogonal to $\boldsymbol{\alpha}$ can be generated from any standard normal point $\boldsymbol{u}$ as $\boldsymbol{u}_{\alpha}^{\perp} = \boldsymbol{u} - (\boldsymbol{u} \cdot \boldsymbol{\alpha})\boldsymbol{\alpha}$. In this way, the search for the limit state, for each random point $\{j\}$, can be set as $\boldsymbol{u}_{\alpha}^{\{j\}}(c) = \boldsymbol{u}_{\alpha}^{\perp\{j\}} + c\ \boldsymbol{\alpha}$. Standard implementation of LS operates with a fixed, initially determined important direction $\boldsymbol{\alpha}$. For each random point $\boldsymbol{u}^{\{j\}}$, the distance from the hyperplane to the performance function in the direction of $\boldsymbol{\alpha}$ is identified searching along the lines $\boldsymbol{u}_{\alpha}^{\{j\}}(c)$. The line search is conducted evaluating the performance function g on the support sequence $c = \{c_1, ..., c_{N_c}\}$,

to find the value c^* by means of interpolation, usually requiring $6-8$ model evaluations per line.

3.2. Adaptive Algorithm

An Adaptive algorithm, is developed in order to further improve the numerical efficiency for implementation in the generalized uncertainty environment. The improvement concerns the efficiency in evaluating Equation 5. In contrast to the standard algorithm, ALS uses a support sequence that is dynamically generated to adapt to the shape of the limit state surface. This makes the algorithm significantly faster and capable of recognising the level of non-linearity of the performance function. Moreover, ALS not only allows for variations in the important direction but is also capable of identifying new important directions to be updated during the simulation. Hence, only a very rough estimation of the important direction is required at the start of the simulation.

The main features of ALS are: (i) it minimizes the number of samples along the lines $\{j\}$ to identify $c^{*\{j\}}$, (ii) it adapts the important direction to the shape of the limit state surface. The first feature is achieved developing an efficient line search procedure and line selection. The second feature is achieved computing weights to each working direction.

As in standard implementations, ALS algorithm operates setting an (initial) important direction $\boldsymbol{\alpha}$ and generating a number N_L of points $\boldsymbol{u}_\alpha^{\{j\}}$. First, a line j is deployed from the origin of the SNS approximately towards the failure region as $\boldsymbol{u}_\alpha^{\{0\}}(c) = c\ \boldsymbol{\alpha}$. Then, the value $c^\circ = \{c \in \mathbb{R} \mid g(c^\circ \boldsymbol{\alpha}) = 0\}$, is used for the starting point on the first line as $\boldsymbol{u}_\alpha^{\{1\}}(c_0) = \boldsymbol{u}^{\perp\{1\}} + c_0\ \boldsymbol{\alpha}$. A line search procedure, based on a Newton-Raphson iterative procedure, is applied to identify the root $c^{*\{j\}}$, which is used to compute the partial probability $p_F^{\{j\}} = w(\boldsymbol{u}_\alpha^{\{j\}}) = \Phi(-|c^{*\{j\}}|)$. Using the identified root, the procedure is repeated as $\boldsymbol{u}_\alpha^{\{j\}}(c_0) = \boldsymbol{u}^{\perp\{j\}} + c^{*\{j-1\}}\ \boldsymbol{\alpha}$, until all lines are processed. To increase the efficiency, the algorithm does not process the lines randomly as they are generated. Lines are selected according to a criterion based on the metric space that recognizes the nearest line to one being currently processed.

Hence, in case of slightly non-linear limit state surface the distances $c^{*\{j\}}$ and $c^{*\{j-1\}}$, for lines j and $j-1$ respectively, are expected to have approximately the same value. To identify the neighbouring lines, the index of the line closest to the origin is computed as $k_1 = \arg\min_j \left\| \boldsymbol{u}^{\perp\{j\}} - 0 \right\|$. Subsequently, all the other indices are calculated as $k_{i+1} = \arg\min_{j \neq k_i} \left\| \boldsymbol{u}^{\perp\{j\}} - \boldsymbol{u}^{\perp\{k_i\}} \right\|$, and as illustrated in the pseudo code of the algorithm in Figure 1.

3.3. Adaptation of the important direction

ALS allows to change the important direction without re-evaluating the performance function along the processed lines. This feature is useful when there is only little evidence of the optimal important direction, so that an approximate direction can be set at the beginning of the simulation and a better direction can be obtained during the simulation. An optimal direction generally provides a more accurate estimate of the failure probability. The important direction is usually associated with the design point $\tilde{\boldsymbol{u}}^* = \min_{\boldsymbol{u} \in \{\boldsymbol{u} | g=0\}} \|\boldsymbol{u}\|$, i.e. the point on the limit state that carries the highest probability density. As the ALS proceeds, the norm of the new state points $\tilde{\boldsymbol{u}} = \boldsymbol{u}^{\perp} + c^* \boldsymbol{\alpha}$ can be computed with nearly no cost, thus a new direction can be set as a more probable point is identified on the limit state. Thus, if a point $\tilde{\boldsymbol{u}}_{\alpha}^{\{j\}}$ is found, such that $||\tilde{\boldsymbol{u}}_{\alpha}^{\{j-1\}}|| > ||\tilde{\boldsymbol{u}}_{\alpha}^{\{j\}}||$, then the new important direction can be set as $\boldsymbol{\alpha}_{new} = \tilde{\boldsymbol{u}}_{\alpha}^{\{j\}} / ||\tilde{\boldsymbol{u}}_{\alpha}^{\{j\}}||$. Changing the important direction does not affect the expected value of the failure probability. However, an improvement of the important direction reduces the variance of the estimation.

3.4. Efficiency and accuracy

Adaptive Line Sampling shows an improvement in efficiency and accuracy above the standard version. This is elucidated in a comparative study with a reference solution obtained by Monte Carlo simulation. An explicit performance function is used to test the methods, which is expressed as $g(\boldsymbol{x}) = -\sqrt{\boldsymbol{x}^T \boldsymbol{x}} + a$, where $\boldsymbol{x}$ are N_{RV} independent normal random variables and a is a constant. Firstly, the test is run with just two random variables, but with decreasing

```
begin
α = α_1;                                  % set initial direction
u^{{j}_1^N} ~ N(0,1);
u^{⊥{j}_1^N} = u^{{j}_1^N} − (u^{{j}_1^N} · α) α;
k_1 = min_j ‖u^{⊥{j}} − 0‖                % get the first line index
find c° such that g(T(c° α)) = 0;
c_0 = c°;                                 % initialize distance from hyperplane
for i = 1 → N do
    u_α^{{k_i}}(c_0) = u^{⊥{k_i}} + c_0 α;
    find c* such that g(T(u_α^{{k_i}}(c*)) = 0;
    k_{i+1} = min_{j≠k_{1,...,i}} ‖u^{⊥{j}} − u^{⊥{k_i}}‖;        % get the next line index
    c_0 = c*;
    p_F^{{i}} = Φ(−|c*|);                 % compute partial probability
    if ||u_α^{{k_i}}(c*)|| < c° then
        c° = ||u_α^{{k_i}}(c*)||;
        α = u_α^{{k_i}}(c*)/c°;           % update direction
    end if
end for
p̂_F = (1/N) Σ_i^N p_F^{{i}};              % failure probability
end
```

Figure 1: Pseudocode for Adaptive Line Sampling

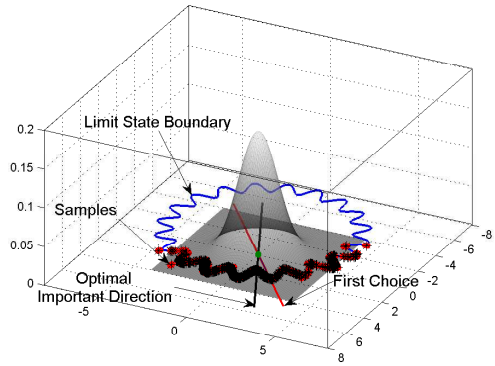
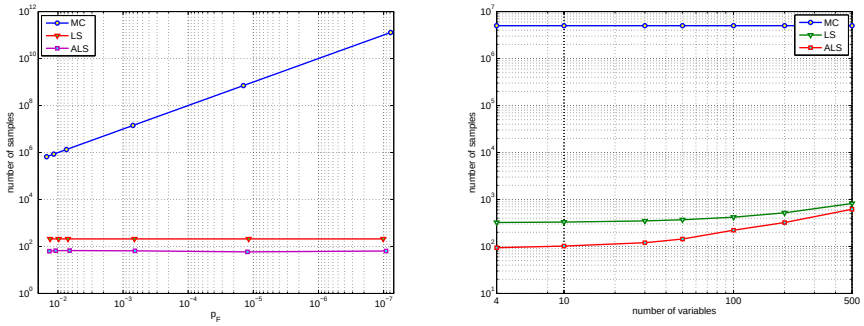


Figure 2: Directional change in standard normal space with ALS over the non-linear limit state boundary defined in the original space by the performance function: $g(\boldsymbol{\theta}) = -(\theta_1+\theta_2)+d^2\ (1+a\ \sin(b\ \tan^{-1}(\theta_1,\theta_2)))$, where $\theta_1 \sim N(5, 2^2)$, $\theta_2 \sim N(2, 2^2)$, $d = 10$, $a = 0.2$ and $b = 20$.

probability targets, by selecting different values of a in the performance function g, as shown in Table 1. Note that, in this case, the values of probability $p_F = \Phi(-\beta)$ obtained by First Order Reliability Method [25] are biased because of the concave shape of the limit state surface. An illustration of the performance of the methods is shown in Figure 3a. A satisfactory level of accuracy ($CoV = 5 \cdot 10^{-2}$) is achieved with just 65 samples using ALS compared to a necessary sample size of 210 samples using LS. Secondly, the test is run fixing probability targets (approximately to 10^{-3}), while progressively increasing the number of random variables, as shown in Table 2. The results of this second test, as illustrated in Figure 3b, show that Monte Carlo, is insensitive to the number of variables, whilst the other methods show some sensitivity to the number of variables but require significantly less samples to achieve the same level of accuracy. As expected, in this second test, the probability of failure computed with the First Order Reliability Method is inaccurate, as also shown in Table 2. In both cases, ALS demonstrated to be the most efficient, and $3-4$ times faster than LS.

$g(\boldsymbol{x}) = -\sqrt{x_1^2 + x_2^2} + a;$ $\quad x_1 \sim N(5, 2^2),\ x_2 \sim N(2, 2^2);$ $\quad a = (10, 10.2, 10.5, 12, 14, 16)$								
	MC		ALS			LS		
$\lvert c^\circ \rvert$	$\hat{p}_F^{MC}$	CoV	$\hat{p}_F^{ALS}$	CoV	Ns	$\hat{p}_F^{LS}$	CoV	Ns
2.307	$1.49\ 10^{-2}$	$0.8\ 10^{-2}$	$1.35\ 10^{-2}$	$6.0\ 10^{-2}$	63	$1.32\ 10^{-2}$	$5.5\ 10^{-2}$	210
2.407	$1.16\ 10^{-2}$	$0.9\ 10^{-2}$	$1.08\ 10^{-2}$	$10.8\ 10^{-2}$	66	$9.75\ 10^{-3}$	$5.4\ 10^{-2}$	210
2.557	$7.40\ 10^{-3}$	$1.1\ 10^{-2}$	$6.60\ 10^{-3}$	$4.5\ 10^{-2}$	67	$7.00\ 10^{-3}$	$5.8\ 10^{-2}$	210
3.304	$7.06\ 10^{-4}$	$3.8\ 10^{-2}$	$6.58\ 10^{-4}$	$9.2\ 10^{-2}$	65	$6.69\ 10^{-4}$	$13.6\ 10^{-2}$	210
4.307	$1.42\ 10^{-5}$	$26.5\ 10^{-2}$	$1.23\ 10^{-5}$	$13.9\ 10^{-2}$	59	$1.18\ 10^{-5}$	$8.3\ 10^{-2}$	210
5.307	–	–	$9.18\ 10^{-8}$	$12.4\ 10^{-2}$	64	$10.14\ 10^{-8}$	$14.0\ 10^{-2}$	210

Table 1: Test of ALS and LS on the bidimensional performance $g(\boldsymbol{x}) = -\sqrt{x_1^2 + x_2^2} + a$; comparison with the reference solution obtained via MC with 10^6 samples.



(a) (b)

Figure 3: Number of samples required from ALS and LS compared to the reference solution obtained with MC and 10^6 samples (a) for a decreasing probability target, and (b) for increasing dimension.

$g(\boldsymbol{x}) = -\sqrt{\boldsymbol{x}^T\boldsymbol{x}} + a;$		$x_i \sim N(2,1);$		$a = (7.0, 9.3, 14.7, 18.1, 24.8, 34.0, 52.5)$				
		MC	ALS			LS		
N_{RV}	$\lvert c^\circ \rvert$	$\hat{p}_F^{MC}$	$\hat{p}_F^{ALS}$	CoV	Ns	$\hat{p}_F^{LS}$	CoV	Ns
4	3.000	$3.27\ 10^{-3}$	$3.52\ 10^{-3}$	$19.7\ 10^{-2}$	94	$2.75\ 10^{-3}$	$7.2\ 10^{-2}$	215
10	2.975	$8.37\ 10^{-3}$	$6.97\ 10^{-3}$	$11.1\ 10^{-2}$	102	$8.57\ 10^{-3}$	$16.9\ 10^{-2}$	221
30	3.745	$4.56\ 10^{-3}$	$4.13\ 10^{-3}$	$12.9\ 10^{-2}$	120	$4.27\ 10^{-3}$	$15.6\ 10^{-2}$	241
50	3.958	$7.44\ 10^{-3}$	$7.87\ 10^{-3}$	$17.8\ 10^{-2}$	144	$7.59\ 10^{-3}$	$16.5\ 10^{-2}$	261
100	4.800	$4.87\ 10^{-3}$	$5.27\ 10^{-3}$	$17.5\ 10^{-2}$	222	$5.92\ 10^{-3}$	$19.7\ 10^{-2}$	311
200	5.716	$5.85\ 10^{-3}$	$5.20\ 10^{-3}$	$18.3\ 10^{-2}$	323	$6.16\ 10^{-3}$	$17.0\ 10^{-2}$	411
500	7.778	$4.15\ 10^{-3}$	$3.44\ 10^{-3}$	$21.4\ 10^{-2}$	619	$3.56\ 10^{-3}$	$17.8\ 10^{-2}$	711

Table 2: Test of ALS and LS on the performance function $g(\boldsymbol{x}) = -\sqrt{\boldsymbol{x}^T\boldsymbol{x}} + a$; comparison with reference solution from MC with 10^6 samples and $CoV \leq 0.03$, and increasing dimension of the limit state.

4. Sampling-based estimation of set-valued reliability

When imprecision is considered, the failure probability is obtained as interval $\overline{p_F}$. In order to calculate the bounds of the failure probability, a global search in the bounded sets $\mathcal{Q}_1$ and $\mathcal{Q}_2$ is performed. A naive approach to the problem for searching in the above sets would be prohibitive in the majority of cases due to the numerical effort incurred. In fact, two nested loops are required, where the inner loop estimates the failure probability and the outer loop searches for the bounds of the probability. The ALS method not only makes the computation of probabilities faster compared with Monte Carlo, but most importantly, can be adopted to significantly ease the search procedure of failure probability.

4.1. The global search for lower and upper failure probabilities

The objective function for the global search in the sets $\mathcal{Q}_1$ and $\mathcal{Q}_2$ is given as the failure probability depending upon the coordinates of $\mathcal{Q}_1$ and $\mathcal{Q}_2$. The

search can be seen as an iterative procedure that converges after some steps, towards the sought lower and upper failure probability bounds.

4.1.1. The search in the bounded set of distribution parameters $\mathcal{Q}_1$

The set $\mathcal{Q}_1$ of distribution parameters defines the set of all probability distribution functions to be considered in the analysis. Any element of $\mathcal{Q}_1$ is associated with a different value of failure probability. Nonetheless, the limit state does not change as we search in $\mathcal{Q}_1$. This is because the limit state depends upon the structural system and not upon the uncertainty model that defines the probability distribution over the state variables. Since the important direction is defined as any direction pointing towards the failure domain, during the search in $\mathcal{Q}_1$, an approximate $\boldsymbol{\alpha}$ can be set for the entire analysis, independently from the distribution functions of the random variables. However, changing the distribution functions modifies the location of the most probable point on the limit state surface. Hence, the direction $\boldsymbol{\alpha}$, set at beginning of the analysis, might not be the optimal one for all the distributions analysed. This motivates the implementation of a flexible algorithm capable of searching and updating new optimal directions.

Each step of the search procedure requires the estimation of a failure probability. In the standard approach a completely new simulation would be carried out to find each of these failure probabilities. However, if the distribution functions do not significantly change, it is not necessary to run a whole new simulation. For this reason, the proposed strategy includes a verification of those changes during the search for the probability bounds. Taking advantage of the bijective mapping T of the random variables between original and standard normal space, and of the fact that the limit state does not change as we search in $\mathcal{Q}_1$, any point $\tilde{\boldsymbol{u}}$ on the limit state can be transformed back onto the original space, and then re-mapped to the SNS for the next simulation. When a new reliability analysis is started, the points on the limit state $\tilde{\boldsymbol{u}}$, previously found, can be used to feed further analyses. Let $\boldsymbol{\alpha_i}$ denote the direction of the current simulation and $\boldsymbol{\alpha}_{i-1}$ be the direction of the previous simulation. At the current

simulation, the failure probability

$$p_{F(i)} = \int_{\mathbb{R}^{n-1}} w(\boldsymbol{u}^{\perp}_{\alpha_i})\ h_{\mathcal{N}}(\boldsymbol{u}^{\perp}_{\alpha_i})\ du^{n-1} \tag{6}$$

can be computed using the limit state points from the previous simulation $w(\boldsymbol{u}^{\perp}_{\alpha_{i-1}}) \rightarrow w(\boldsymbol{u}^{\perp\text{remap}}_{\alpha_i})$. However, the new values of w obtained with the re-mapped points are no longer drawn from a probability distribution. Therefore, in order to be able to compute the failure probability using the points from previous simulations, a dummy probability density function $h_{\mathcal{X}}$ is constructed around the re-mapped points on the hyperplane. The density function $h_{\mathcal{X}}$ is a multi-modal distribution with density peaks centred on the re-mapped points and are weighed using the metric properties of the SNS. The failure probability can then be written as

$$p_{F(i+1)} = \int_{\mathbb{R}^{n-1}} w(\boldsymbol{u}^{\perp\text{remap}}_{\alpha_{i+1}})\ h_{\mathcal{N}}(\boldsymbol{u}^{\perp}_{\alpha_i}) \frac{h_{\mathcal{X}}(\boldsymbol{u}^{\perp\text{remap}}_{\alpha_{i+1}})}{h_{\mathcal{N}}(\boldsymbol{u}^{\perp}_{\alpha_i})}\ du^{n-1}. \tag{7}$$

By means of the ratio $q = h_{\mathcal{X}}(\boldsymbol{u}^{\perp\text{remap}}_{\alpha_i})/h_{\mathcal{N}}(\boldsymbol{u}^{\perp}_{\alpha_{i-1}})$, the probability $p_{F(i)}$ can now be computed using the information from simulation $i-1$ as

$$\tilde{p}_{F(i)} = \frac{1}{N}\sum_{j}^{N} q^{\{j\}}\ w(\boldsymbol{u}^{\perp\text{remap}}_{\alpha_i}) = \frac{1}{N}\sum_{j}^{N} q^{\{j\}}\ R(w(\boldsymbol{u}^{\perp}_{\alpha_{i-1}})), \tag{8}$$

where, $R(.)$ is a function that transforms variable $w(\boldsymbol{u}^{\perp}_{\alpha_{i-1}})$ from the standard space of simulation $i-1$ to the standard space of simulation i.

4.1.2. The search in the bounded set of structural parameters $\mathcal{Q}_2$

Imprecision of structural parameters, characterized by the bounded set $\mathcal{Q}_2$, requires an extension of the procedure developed so far. In fact, the bounded variables $\boldsymbol{x} \in \mathbb{R}^{n_2}$ change the shape of the limit state boundary, which needs to be addressed with a second search as described in Equation (1). In this section, we propose a strategy to include the variables $\boldsymbol{x} \in \mathcal{Q}_2$ in the numerical framework presented so far. The strategy consists of an extension to an augmented probability space, where the interval variables are treated as dummy normal random variables having imprecise mean values and fixed standard deviations.

In simple terms, this permits a combined consideration of the bounded set $\mathcal{Q}_2$ together with the bounded set $\mathcal{Q}_1$ in the same manner. Each dummy imprecise random variable has an interval mean value $\overline{\underline{\mu_x}} = \overline{\underline{x}}$, and a real-valued standard deviation σ_x to be fixed with some convenient value. By defining these dummy imprecise random variables a thorough search can be performed in both sets $\mathcal{Q}_1$ and $\mathcal{Q}_2$ simultaneously. The only requirement for the dummy imprecise random variables is that the chosen value of σ_x should neither be too large nor be too small to avoid numerical issues in computing the failure probability. The standard deviation σ_x can be set, for example, as a fraction of the interval radius $\sigma_x = k(\overline{x} - \underline{x})/2$, where k can be any value between 0 and 1. Once the argument optima in the sets $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are found, the associated bounds on the failure probability are also known. Two more reliability analyses at the end of the search, run on the argument optima, will be needed to find the failure probability bounds. Note that during this procedure sampling outside the intervals may occur. However, points outside the intervals are solely used to drive the search process. In cases where the physical model restricts the evaluation to the range of the intervals, truncated normal random variables are used for the dummy imprecise variables, which lower and upper limits are equal to the endpoints of the intervals.

When the limit state surface is only slightly non-linear the search procedure can be sensibly sped up. In fact, in this case the important directions in the original space are all oriented towards the same region of the state space. This implies that, as we search in $\mathcal{Q}_1$ and $\mathcal{Q}_2$, the coordinates of the important directions may vary but do not change in sign. Therefore, the sign of the coordinates of the important direction in the original space can be used to identify the mean states that are the nearest and furthest from the limit state surface. Let us denote these two states as conjugate states. Where the mean state is the nearest to the limit state surface (upper conjugate state), it is also where the failure probability attains its maximum. The contrary applies at the furthest mean state from the limit state surface (lower conjugate state), where the failure probability is minimum. However, at this stage, a distinction between

distributions defined in terms of moments (first and second) and in terms of parameters is necessary. If the probability distribution is defined in terms of moments, the argument optima of the failure probability are obtained at the conjugate states, by selecting the maximum and minimum variance, respectively for the minimum and maximum failure probability. This applies because we can pick values at the corners at the hyper-cube defined in terms of moments, without problems of dependency. If the probability distribution is defined in terms of parameters, the search domain in the space of the moments may no longer be a hyper-cube. In fact, in this case a conjugate state (either lower or upper) corresponds to only one value of variance, which may not be at a corner of the domain. However, lower and upper bounds of the failure probability can still be found selecting the conjugate states and the maximum/minimum variance independently, and find the associated parameter combinations.

5. Integration of the strategy in OpenCossan

The developed algorithm has been integrated into OpenCossan [26], which is an open-source integrated numerical framework for uncertainty quantification and risk analyses [24]. OpenCossan is coded exploiting the object-oriented Matlab® environment, which makes it highly flexible using a modular software architecture. Recalling that a class is an extensible case of objects and properties [27], the strategy takes advantage of three main new classes, namely *AdvancedLineSampling*, *LineSamplingData* and *ExtremeCase*.

5.1. Class Advanced Line Sampling

The class *AdvancedLineSampling* integrates the methods for estimating precise failure probabilities. Provided a ProbabilisticModel, containing a performance function, and an Input object, a reliability analysis can be performed invoking the method *computeFailureProbability*, as shown in Figure 4. In order to optimize the performance and increase the robustness of *AdvancedLineSampling* several new methods are developed. These methods include: *lineSearch*,

which implements a Newton-Raphson method to look for the roots of the limit state surface, *extractLineIndex*, which searches the index of the nearest line to the current one amongst the ones not already processed, and *computePartialProbabilities*, which computes the probabilities $p_F^{\{i\}} = \Phi(-|c^{*\{i\}}|)$ for every line $i = 1, ..., N$. The design of these methods ensure the accuracy and robustness of the algorithm. The method *computePartialProbabilities*, for example, is responsible for evaluating the expansions w and can be given the option of eliminating individual lines during the search process. The *lineSearch* method provides the user with the choice to adjust the control parameters, such as the tolerance on the values of g and the minimum step size, to control the accuracy of the algorithm if necessary.

5.2. Class Line Sampling Data

The class *LineSamplingData* is a key-component in the economy of the strategy. It stores the results obtained from every line in a structured and organized way, it can be used to plot the results, and feed further analyses. This class is essential also for the parallelization of the algorithm, in combination with the methods *merge* and *add*, which allow to gather results coming from different analyses.

5.3. Class Extreme Case Analysis

Eventually, in order to search for lower and upper bounds of the failure probability, the class *ExtremeCase* is created in OpenCossan. *ExtremeCase* connects the inner solver running the ALS simulations with an optimizer such as *Genetic Algorithms*. *ExtremeCase* makes use of the method *optimize* to deploy the optimization, and of the method *ConstructSolutionSequence* to formulate the optimization problem. *ConstructSolutionSequence* is a central method for the efficiency of the algorithm; it checks number and accuracy of simulations and controls whether more simulations are needed to complete the optimization.

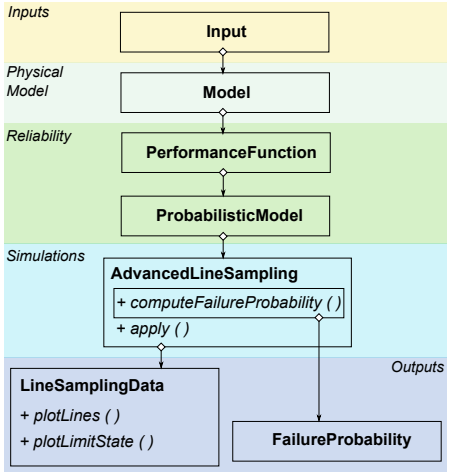


Figure 4: Simplified UML diagram for the implementation of the Advanced Line Sampling strategy in OpenCossan.

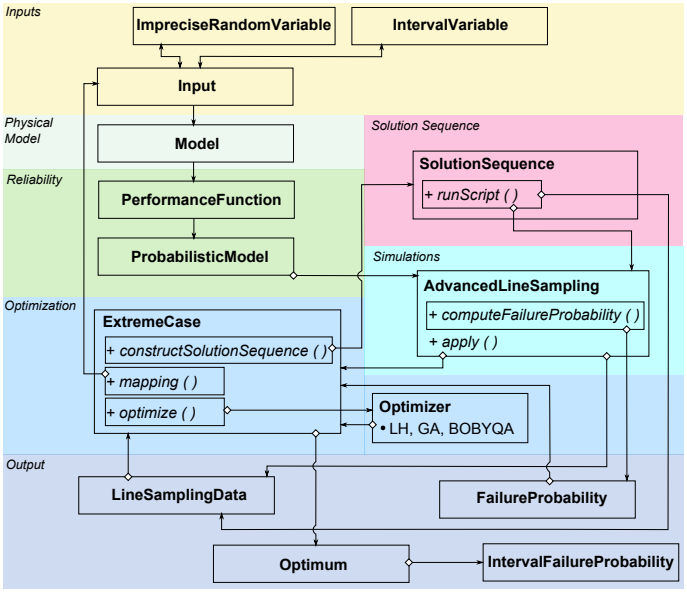


Figure 5: Simplified UML diagram with patterns for the implementation of the Extreme Case strategy in OpenCossan.

6. Examples and applications

6.1. Synthetic example

To demonstrate the performance of the proposed method a synthetic example is presented. The approach developed in this paper, denoted as approach A, is compared to a solution obtained through global optimization, denoted as approach B. Both approaches are applied to calculate the interval failure probability $\overline{\underline{p_F}}$. In approach A the argument optima are detected using the information of the important direction as explained in Section 4.1.2. The sign of the important direction in the original space identifies the conjugate states where the optima of the failure probability are located.

In approach B, the search procedure is driven by optimizers. The example is solved using both Genetic Algorithm (GA) according to [28] and BOBYQA from [29], as global and local searchers, respectively. With this approach a thorough search in the sets $\mathcal{Q}_1$ and $\mathcal{Q}_2$ is performed. The objective function is given by the failure probability, thus, at any iteration of GA/BOBYQA, a simulation with ALS is performed. This approach can be run only because ALS requires just few evaluations of the performance function to complete an iteration. Replacing ALS with Monte Carlo would lead to hundreds of evaluations of the performance function for each iteration, making approach B intractable.

Two cases are considered in this study, namely case (a) and case (b).

Case (a): The linear performance function $g(\xi, x) = 7 + \xi - 2x$, is evaluated. It includes the imprecise random variable $\xi \in \mathcal{C}$, where

$\mathcal{C} = \{h_{\mathcal{N}}(\xi; \mu, \sigma) \mid \mu \in [0.9, 1.3],\ \sigma \in [0.7, 2.1]\}$, and the interval variable $\underline{\overline{x}} = [1,\ 3]$. In this illustrative case the gradient $\nabla g = (1, -2)$, suggests the initial important direction $\boldsymbol{\alpha} = (1, -2)/\sqrt{5}$. Approach A leads to the bounds of the failure probability and the associated argument optima $(x_*, x^*) = (\underline{x}, \overline{x})$, and $\boldsymbol{p}_* = (\overline{\mu}, \underline{\sigma})$, $\boldsymbol{p}^* = (\underline{\mu}, \overline{\sigma})$, as shown in Table 3. With approach B, using GA with a population size of 50 individuals, an approximation of the of lower and upper bound was obtained after 52 iterations, while BOBYQA delivered a slightly less accurate estimate. In this case approach A coincides with the

$g = 7 + \xi - 2x$;			
	Approach A	Approach B (GA)	Approach B (BOBYQA)
$(\mu_*,\ \mu^*)$	$(2,\ 0)$	$(2,\ 0.04)$	$(2,\ 0.12)$
$(\sigma_*,\ \sigma^*)$	$(1.2,\ 2.3)$	$(1.28,\ 2.3)$	$(1.32,\ 2.3)$
$(x_*,\ x^*)$	$(1,\ 3)$	$(1,\ 3)$	$(1,\ 2.92)$
$\overline{\underline{p_F}}$	$[2.717\ 10^{-9},\ 0.332]$	$[2.702\ 10^{-8},\ 0.322]$	$[4.371\ 10^{-7},\ 0.134]$

Table 3: Results from Example I, case (a), argument optima and associated failure probabilities for approaches A and B respectively.

$g = 9 + \boldsymbol{\xi}^T \boldsymbol{a}_1 - \boldsymbol{x}^T \boldsymbol{a}_2$;			
$\boldsymbol{a}_1 = (1, 4, 2, 0.1, 0.2, 0.6, 5, 0.01, 0.2, 0.3, 0.25, 0.14, 0.8, 3),\ \boldsymbol{a}_2 = (-2, 0.1, 1)$			
	Approach A	Approach B (GA)	Approach B (BOBYQA)
$\overline{\underline{p_F}}$	$[1.795\ 10^{-9},\ 0.1452]$	$[7.302\ 10^{-6},\ 0.0053]$	$[2.538\ 10^{-5},\ 0.0046]$

Table 4: Results from Example I, case (b), interval failure probability for approach A and B.

closed-form solution and it is clearly advantageous above approach B.

Case (b): The multidimensional linear performance function $g(\boldsymbol{\xi}, \boldsymbol{x}) = 9 + \boldsymbol{\xi}^T \boldsymbol{a}_1 - \boldsymbol{x}^T \boldsymbol{a}_2$, where $\boldsymbol{a}_1 \in \mathbb{R}^{14}$, and $\boldsymbol{a}_2 \in \mathbb{R}^3$, is considered. The imprecise random variables $\boldsymbol{\xi} \in \mathbb{R}^{14}$ are defined by the credal set $\mathcal{C} = \left\{ h_{\mathcal{N}}(\boldsymbol{\xi}; \boldsymbol{\mu}, \boldsymbol{\sigma}) \mid \boldsymbol{\mu} \in \underline{\overline{\boldsymbol{\mu}}}, \boldsymbol{\sigma} \in \underline{\overline{\boldsymbol{\sigma}}} \right\}$, where $\underline{\overline{\boldsymbol{\mu}}} = [0.1,\ 1]^{14}$, and $\underline{\overline{\boldsymbol{\sigma}}} = [1.2,\ 2.3]^{14}$, while the interval variables $\boldsymbol{x} \in \mathbb{R}^{14}$ are defined by the bounded set $\underline{\overline{\boldsymbol{x}}} = [1,\ 3]^3$. Again, because of the linearity, approach A delivers numerically exact results for the failure probability (equal to the closed-form solution). As expected, approach B provides only a rough approximation of the solution, as shown in Table 4. The global search becomes inefficient when the dimensionality of the search domain grows, in the example up to 31.

6.2. Large scale finite element model of a six-storey building

In this example the reliability analysis of a six-story building subject to wind load is carried out. Three different models of uncertainty are considered with increasing level of generality. Firstly, a standard reliability analysis, where the inputs are modelled by precise probability distribution functions, is performed. Secondly, the structural parameters are modelled as imprecise random variables with the credal set $\mathcal{C}$. In the third analysis both imprecise random variables and intervals are considered for structural parameters.

An ABAQUS finite element model (FEM) is built for the six-story building, as illustrated in Figure (6), which includes beam, shell and solid elements. The load is considered as combination of a (simplified) lateral wind load and the self-weight, which are both modelled by deterministic static forces acting on nodes of each floor. The magnitude of the wind load increases with the height of the building. The FEM of the structure involves approximately 8200 elements and $66,300$ DOFs. A total of 244 independent random variables are considered to account for the uncertainty of the structural parameters. The material strength (capacity) is represented by a normal distribution, while log-normal distributions are assigned to the Young's modulus, the density and the Poisson ratio. In addition, the cross-sectional width and height of the columns are modelled by independent uniform distributions. A summary of the distribution models is reported in Table 5.

Component failure for the columns of the 6th storey is considered as failure criterion. The performance function is defined as

$$f(\boldsymbol{\theta}) = |\sigma_I(\boldsymbol{\theta}) - \sigma_{III}(\boldsymbol{\theta})| \, /2 - \sigma_y, \tag{9}$$

i.e. as the difference between the maximum Tresca stress, where $\sigma_{III} \leq \sigma_{II} \leq \sigma_I$ are the principal stresses, and σ_y is the yield stress.

Standard reliability analysis. A reliability analysis is carried out with the precise distribution models from Table 5, and using LS and ALS for comparison of efficiency. The initial important direction is selected based on the gradient in

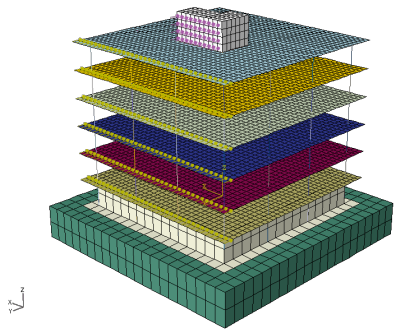


Figure 6: FE-model of the six-story building.

# U.V.	Probability dist.	Distribution	Description	Units
1	N(0.1, 0.001)	Normal	Column's strength	GPa
$2-193$	Unif$(0.36, 0.44)$	Uniform	Sections size	m
$194-212$	LN(35.0, 12.25)	Log-normal	Young's modulus	GPa
$213-231$	LN(2.5, 0.0625)	Log-normal	Material's density	kg/dm^3
$232-244$	LN(0.25, 0.000625)	Log-normal	Poisson's ratio	-

Table 5: Precise distribution models for the input structural parameters.

the origin of the SNS. The identified important direction is displayed in figure 7, where the first coordinate (the material's strength) appears to be the most important one. The other coordinates refer to the size of the cross-sections, the Young's modulus, the density, and the Poisson's ratio, respectively (see Table 5). As illustrated in Figure (7), only a few uncertain variables (U.V.) dominate the important direction; these are the Young's modulus of columns of floor 6 (U.V. #199) and the density of the columns of floors 5 and 6 (U.V. #223 and #224), along with the yield strength (U.V. #1). In this example, performing LS with 30 lines (180 samples) leads to the failure probability of $\hat{p}_F = 1.30 \cdot 10^{-4}$ and a coefficient of variation of $CoV = 0.076$. ALS leads to the probability of failure $\hat{p}_F = 1.42 \cdot 10^{-4}$ with a coefficient of variation of $CoV = 0.092$, but

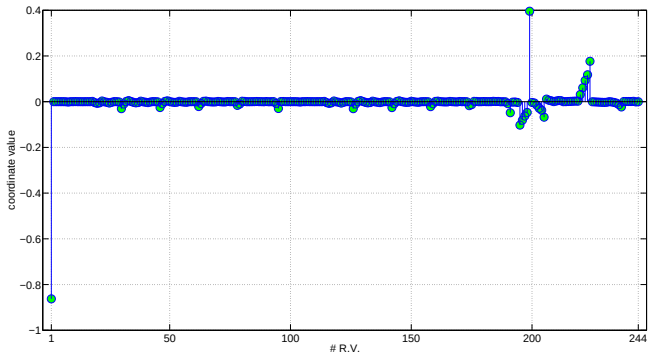


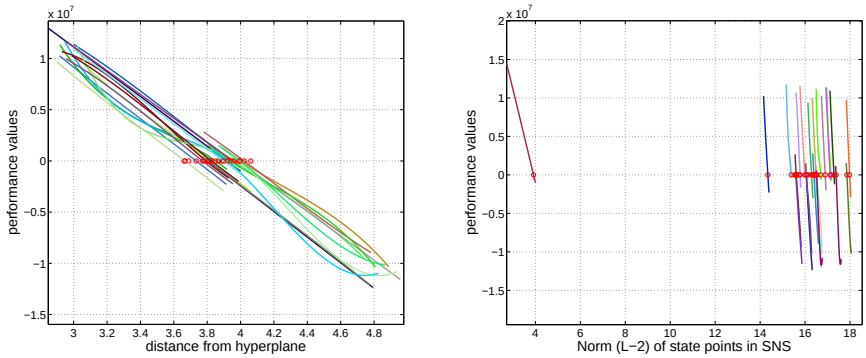
Figure 7: Values of the 244 coordinates of the initial important direction in the standard normal space.

with only 62 samples. Both methods estimate approximately the same value of failure probability, but with quite a smaller number of model evaluations were required by ALS.

Imprecision in distribution parameters ($\boldsymbol{\mathcal{M}}_I$). The model of uncertainty is extended to include the credal set

$\mathcal{C}\left\{h_{\mathcal{D}}\left(\boldsymbol{\theta};\boldsymbol{p}\right) \mid \boldsymbol{p}\in\mathbb{R}^{488},\ \boldsymbol{p}\in\mathcal{Q}_1\right\}$, where $\mathcal{D}$ are the probability distribution models from Table 5, and

$\boldsymbol{p}=(\mu_1,\ \sigma_1,...,m_{244},v_{244})$ are the distribution parameters of these models specified by the bounded set $\mathcal{Q}_1=\bigtimes_i^{488}\ \underline{\overline{p}}_i$. The interval parameters are represented as $\underline{p}=p_c\left(1-\epsilon\right)$, $\overline{p}=p_c\left(1+\epsilon\right)$, using the interval center $p_c=(\overline{p}+\underline{p})/2$ and the relative radius of imprecision ϵ. These intervals $[\underline{p},\ \overline{p}]$ are summarized in the bounded set $\mathcal{Q}_1$. In the example, all interval parameters, are modeled with the same relative imprecision ϵ. In order to explore the effects of ϵ on the results, we use a fuzzy set to consider a nested set of intervals $\tilde{p}=\left\{[\underline{p},\ \overline{p}]\right\}$ for the parameters in one analysis. The amplitude (width) of the intervals is controlled by ϵ to obtain fuzzy sets $\tilde{p}$ as shown in Figure 9. An upper limit for the relative uncertainty is set as $\overline{\epsilon}=0.075$. Specifically, the intervals for $\epsilon=\{0,0.005,0.01,0.025,0.05,0.075\}$ are considered. The reliability analysis



(a) (b)

Figure 8: Values of the performance function along the lines in SNS for one reliability analysis of the multi-storey building. In Figure (a), the lines are plotted aginst the distance from the hyperplane c, while in Figure (b) the lines are plotted against the L-2 norm $||\tilde{\boldsymbol{u}}||$ of the state points $\tilde{\boldsymbol{u}} = T(\tilde{\boldsymbol{\theta}})$.

with the generalized model of uncertainty is performed using the important direction determined in the original space.

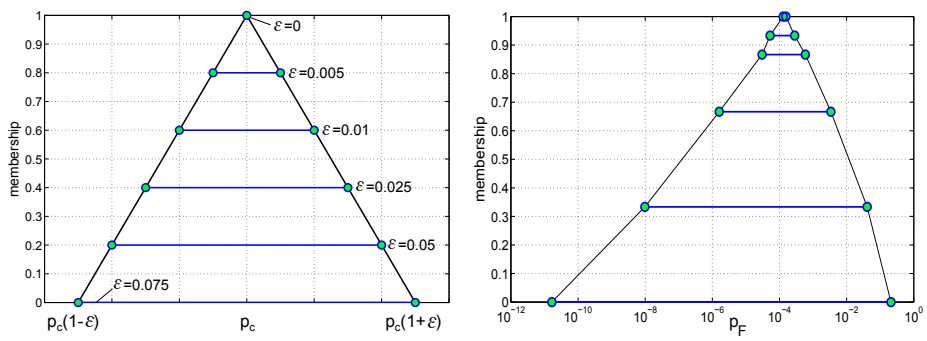
From a rough search in the set $\mathcal{Q}_1$, it was found that the important direction did not significantly change in the original space. This allowed us to identify the argument optima in the bounded set $\mathcal{Q}_1$ as combination of extreme moments as described in Section 4.1.2. These upper and lower conjugate states are also associated with the maximum and minimum of the failure probability, respectively. The result of the robust reliability analysis is shown in Figure 9b and in Table 6. From Table 6 can be appreciated that the number of samples required by one robust reliability analysis, on average, is approximately 254, which is even less than number of samples required by two standard reliability analyses using Line Sampling (~ 360 samples). This is an astounding results considering that a standard approach, driven by two nested loops, would have required several hundreds of thousands of samples. The failure probability is obtained as a fuzzy set, which includes the standard reliability analysis as special case with $\epsilon = 0$. Each interval for p_F corresponds to the respective interval $\overline{p} = [\underline{p},\ \overline{p}]$ in the

	Lower Bound		Upper Bound		
ϵ	$\underline{p}_F$	CoV	$\overline{p}_F$	CoV	Ns
0.000	$1.42\ 10^{-4}$	$9.2\ 10^{-2}$	$1.42\ 10^{-4}$	$9.2\ 10^{-2}$	126
0.005	$5.75\ 10^{-5}$	$8.7\ 10^{-2}$	$2.63\ 10^{-4}$	$7.1\ 10^{-2}$	257
0.010	$4.57\ 10^{-5}$	$33.6\ 10^{-2}$	$5.30\ 10^{-4}$	$11.5\ 10^{-2}$	250
0.025	$1.75\ 10^{-6}$	$8.8\ 10^{-2}$	$3.22\ 10^{-3}$	$5.3\ 10^{-2}$	253
0.050	$2.27\ 10^{-8}$	$57.0\ 10^{-2}$	$3.88\ 10^{-2}$	$5.4\ 10^{-2}$	255
0.075	$1.88\ 10^{-11}$	$12.2\ 10^{-2}$	$2.02\ 10^{-1}$	$3.5\ 10^{-2}$	254

Table 6: Results of the robust reliability analysis of the multi-storey building from model $\boldsymbol{\mathcal{M}}_I$, obtained in terms of lower and upper bounds of the failure probability.

input for the same membership level, and each membership level is associated with a different value ϵ, see Figure 9b. In a design context, this result can be used to identify a tolerated level of imprecision for the inputs given a constrain on the failure probability. For example, fixing an allowable failure probability of 10^{-3}, the maximum level of imprecision for the distribution parameters is limited to 1%, see Figure 9.

Imprecision in both distribution parameters and structural parameters ($\boldsymbol{\mathcal{M}}_{II}$). In this example the section sizes $\boldsymbol{x} \in \mathbb{R}^{192}$ are considered as interval variables, while the remaining structural parameters $\boldsymbol{\zeta} \in \mathbb{R}^{52}$ are considered as imprecise random variables. The model of uncertainty comprises the set $\mathcal{C}\left\{h_{\mathcal{D}}\left(\boldsymbol{\zeta};\boldsymbol{p}\right) \mid \boldsymbol{p} \in \mathbb{R}^{104},\ \boldsymbol{p} \in \mathcal{Q}_1\right\}$, and the set $\mathcal{Q}_2 = \bigtimes_i^{192} \overline{\underline{x}}_i$. The imprecise distribution parameters are modeled using the radius of imprecision ϵ, as in model case $\boldsymbol{\mathcal{M}}_I$, see Table 8. An upper limit for the relative radius of imprecision is set to $\overline{\epsilon} = 0.03$. In the analysis, a rough search in the sets $\mathcal{Q}_1$ and $\mathcal{Q}_2$ allowed us again to identify a main important direction for determining the argument optima associated with the minimum and maximum value of failure probability. The result is shown in Table 9 and Figure 10b. From Table 9 can be appreciated that the number of



(a) (b)

Figure 9: (a) Fuzzy paramters $\tilde{p} = \{p_c\ [1-\epsilon_j,\ 1+\epsilon_j]\}_{j=1}^{6}$ and (b) fuzzy failure probability obtained with model $\boldsymbol{\mathcal{M}}_I$ as set of results for different levels of imprecision.

samples required by one robust reliability analysis, on average, is approximately 254. Again, it is necessary to point out that a standard approach, driven by two nested loops, would have required several hundreds of thousands of samples to compute the interval failure probability.

To explore the sensitivity against imprecision of the uncertain parameters, the failure probability is obtained as a fuzzy set. The relative radii of imprecision $\epsilon = \{0, 0.01, 0.015, 0.020, 0.025, 0.03\}$ are considered to construct a fuzzy model for all parameters, see Figure 10a. The intervals for the structural parameters $\underline{\overline{x}}$ in $\mathcal{Q}_2$, describing the size of the cross-sections, are independent of ϵ, see Table 8. Once more, the analysis may serve as a design tool to find the tolerable level of imprecision provided a threshold of allowable probability.

Here, the uncertainty due to imprecision is larger, because the whole range of the intervals is taken into account for the cross-sections. As in the previous case, a rough search in the sets $\mathcal{Q}_1$ and $\mathcal{Q}_2$ allowed us to identify a main important direction for selecting the argument optima producing minimum and maximum value of failure probability. Values of failure probability, obtained with $\epsilon = \{0, 0.01, 0.015, 0.020, 0.025, 0.03\}$, are shown in Figure 10 (b).

# U.V.	Prob. dist.	$\underline{\overline{p}} = p_c\ [1-\epsilon,\ 1+\epsilon]$		Description	Units
1	$\mathrm{N}(\underline{\overline{\mu}},\ \underline{\overline{\sigma}})$	$\mu_c = 0.1$	$\sigma_c = 0.01$	Columns' strength	GPa
$2-193$	$\mathrm{Unif}(\underline{\overline{a}}, \underline{\overline{b}})$	$a_c = 0.36$	$b_c = 0.44$	Sections' size	m
$194-212$	$\mathrm{LN}(\underline{\overline{m}},\ \underline{\overline{v}})$	$m_c = 35$	$v_c = 12.25$	Young's modulus	GPa
$213-231$	$\mathrm{LN}(\underline{\overline{m}},\ \underline{\overline{v}})$	$m_c = 2.5$	$v_c = 0.0625$	Material's density	$\mathrm{kg/dm^3}$
$232-244$	$\mathrm{LN}(\underline{\overline{m}},\ \underline{\overline{v}})$	$m_c = 0.25$	$v_c = 0.000625$	Poisson's ratio	-

Table 7: Inputs definition from uncertainty model $\boldsymbol{\mathcal{M}}_I$; the relative radius of imprecision for this model is set as $\epsilon = \{0, 0.005, 0.0.01, 0.025, 0.05, 0.075\}$.

# U.V.	Uncertainties type		$\underline{\overline{p}} = p_c\ [1-\epsilon,\ 1+\epsilon],\ \underline{\overline{x}} = [\underline{x},\ \overline{x}]$	
1	distribution	$\mathrm{N}(\underline{\overline{\mu}},\ \underline{\overline{\sigma}}^2)$	$\mu_c = 0.1$	$\sigma_c^2 = 0.001$
$2-193$	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	$\mathrm{LN}(\underline{\overline{m}},\ \underline{\overline{v}})$	$m_c = 2.5$	$v_c = 0.0625$
$232-244$	distribution	$\mathrm{LN}(\underline{\overline{m}},\ \underline{\overline{v}})$	$m_c = 0.25$	$v_c = 0.000625$

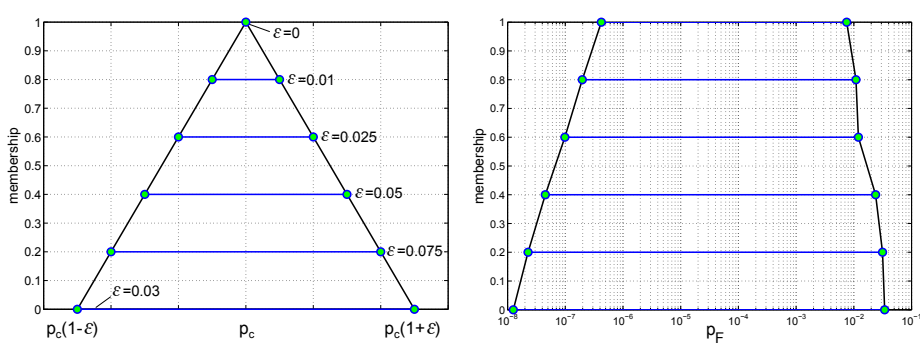
Table 8: Inputs definition from uncertainty model $\boldsymbol{\mathcal{M}}_{II}$; the relative radius of imprecision for this model is set as $\epsilon = \{0, 0.01, 0.015, 0.020, 0.025, 0.03\}$.

7. Conclusions

In this paper an efficient computational strategy for computing set-valued failure probabilities was presented. The approach couples advanced sampling-based methods with optimization procedures. An Adaptive algorithm was developed and implemented into the broader Advanced Line Sampling method. The global search for lower and upper bounds of the failure probability was driven using the information provided by an averaged important direction, obtained in the original space of the state variables, to identify the conjugate states. It was shown, by means of examples, how the advanced search strategy dramatically

	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

Table 9: Results of the robust reliability analysis of the multi-storey building from model $\boldsymbol{\mathcal{M}}_{II}$, obtained in terms of lower and upper bounds of the failure probability.



(a) (b)

Figure 10: (a) Fuzzy distribution parameters $\tilde{p} = \{p_c\ [1-\epsilon_j,\ 1+\epsilon_j]\}_{j=1}^{6}$ and (b) fuzzy failure probability from model $\boldsymbol{\mathcal{M}}_{II}$ obtained as set of results for different levels of imprecision

reduces the computational time of robust reliability analysis without compromising the accuracy of results. The efficiency of the proposed method allows its application on real scale engineering problems, while its accuracy guarantees the computation of informative intervals of failure probability.

[1] O. Ditlevsen, H. O. Madsen, Structural reliability methods, Vol. 178, Citeseer, 2007.

[2] P. Bjerager, On computation methods for structural reliability analysis, Structural Safety 9 (2) (1990) 79–96.

[3] J. Hurtado, A. Barbat, Monte carlo techniques in computational stochastic mechanics, Archives of Computational Methods in Engineering 5 (1) (1998) 3–29.

[4] N. Metropolis, S. Ulam, The monte carlo method, Journal of the American statistical association 44 (247) (1949) 335–341.

[5] R. Melchers, Importance sampling in structural systems, Structural safety 6 (1) (1989) 3–10.

[6] S. Engelund, R. Rackwitz, A benchmark study on importance sampling techniques in structural reliability, Structural Safety 12 (4) (1993) 255–276.

[7] O. Ditlevsen, P. Bjerager, R. Olesen, A. Hasofer, Directional simulation in gaussian processes, Probabilistic Engineering Mechanics 3 (4) (1988) 207–217.

[8] O. Ditlevsen, R. E. Melchers, H. Gluver, General multi-dimensional probability integration by directional simulation, Computers & Structures 36 (2) (1990) 355–368.

[9] P. Koutsourelakis, H. Pradlwarter, G. Schuëller, Reliability of structures in high dimensions, part i: algorithms and applications, Probabilistic Engineering Mechanics 19 (4) (2004) 409–417.

[10] P. Koutsourelakis, Reliability of structures in high dimensions. part ii. theoretical validation, Probabilistic engineering mechanics 19 (4) (2004) 419–423.

[11] S.-K. Au, J. L. Beck, Estimation of small failure probabilities in high dimensions by subset simulation, Probabilistic Engineering Mechanics 16 (4) (2001) 263–277.

[12] A. Der Kiureghian, Analysis of structural reliability under parameter uncertainties, Probabilistic engineering mechanics 23 (4) (2008) 351–358.

[13] A. Der Kiureghian, P.-L. Liu, Structural reliability under incomplete probability information, Journal of Engineering Mechanics 112 (1) (1986) 85–104.

[14] M. Beer, S. Ferson, V. Kreinovich, Imprecise probabilities in engineering analyses, Mechanical Systems and Signal Processing 37 (1) (2013) 4–29.

[15] R. Moore, W. Lodwick, Interval analysis and fuzzy set theory, Fuzzy Sets and Systems 135 (1) (2003) 5–9.

[16] S. Ferson, V. Kreinovich, J. Hajagos, W. Oberkampf, L. Ginzburg, Experimental uncertainty estimation and statistics for data having interval uncertainty, Sandia National Laboratories, 2007.

[17] D. Moens, D. Vandepitte, A survey of non-probabilistic uncertainty treatment in finite element analysis, Computer methods in applied mechanics and engineering 194 (12) (2005) 1527–1555.

[18] Y. Ben-Haim, A non-probabilistic concept of reliability, Structural Safety 14 (4) (1994) 227–245.

[19] B. Möller, W. Graf, M. Beer, Fuzzy structural analysis using α-level optimization, Computational Mechanics 26 (6) (2000) 547–565.

[20] L. A. Zadeh, Fuzzy sets, Information and control 8 (3) (1965) 338–353.

[21] M. Zaffalon, The naive credal classifier, Journal of statistical planning and inference 105 (1) (2002) 5–21.

[22] C. Jiang, X. Han, G. Lu, J. Liu, Z. Zhang, Y. Bai, Correlation analysis of non-probabilistic convex model and corresponding structural reliability technique, Computer Methods in Applied Mechanics and Engineering 200 (33) (2011) 2528–2546.

[23] M. Valdebenito, H. Pradlwarter, G. Schuëller, The role of the design point for calculating failure probabilities in view of dimensionality and structural nonlinearities, Structural Safety 32 (2) (2010) 101–111.

[24] E. Patelli, H.M.Panayirci, M. Broggi, B. Goller, P. Beaurepaire, H. J. Pradlwarter, G. I. Schuëller, General purpose software for efficient uncertainty management of large finite element models, Finite Elements in Analysis and Design 51 (2012) 31–48.

[25] K. Breitung, Asymptotic approximations for multinormal integrals, Journal of Engineering Mechanics 110 (3) (1984) 357–366.

[26] E. Patelli, M. Broggi, M. de Angelis, M. Beer, OpenCossan: An efficient open tool for dealing with epistemic and aleatory uncertainties, ACSE, Proccedings of ICVRAM-ISUMA, Liverpool 14-16 July 2014. In press.

[27] C. Larman, Applying UML and Patterns: An Introduction to Object-Oriented Analysis and Design and Iterative Development, 3/e, Pearson Education India, 2012.

[28] D. E. Goldberg, et al., Genetic algorithms in search, optimization, and machine learning, Vol. 412, Addison-wesley Reading Menlo Park, 1989.

[29] M. J. Powell, The bobyqa algorithm for bound constrained optimization without derivatives, Cambridge NA Report NA2009/06 (2009), University of Cambridge, Cambridge.