

# Meta-model-based importance sampling for reliability sensitivity analysis

V. Dubourg<sup>a</sup>, B. Sudret<sup>b,\*</sup>

<sup>a</sup>*Phimeca Engineering, Centre d’Affaires du Zénith, 34 rue de Sarliève, F-63800 Cournon d’Auvergne, France*

<sup>b</sup>*ETH Zürich, Institute of Structural Engineering, Chair of Risk, Safety & Uncertainty Quantification, Wolfgang-Pauli-Strasse 15., CH-8093 Zürich, Switzerland*

---

## Abstract

Reliability sensitivity analysis aims at studying the influence of the parameters in the probabilistic model onto the probability of failure of a given system. Such an influence may either be quantified on a given range of values of the parameters of interest using a parametric analysis, or only locally by means of its partial derivatives. This paper is concerned with the latter approach when the limit-state function involves the output of an expensive-to-evaluate computational model. In order to reduce the computational cost it is proposed to compute the failure probability by means of the recently proposed meta-model-based importance sampling method. This method resorts to the adaptive construction of a Kriging meta-model which emulates the limit-state function. Then, instead of using this meta-model as a surrogate for computing the probability of failure, its probabilistic nature is used in order to build an **quasi-optimal** instrumental density function for accurately computing the actual failure probability through importance sampling. The proposed estimator of the failure probability recasts as a product of two terms. The augmented failure probability is estimated using the emulator only, while the correction factor is estimated using both the actual limit-state function and its emulator in order to quantify the substitution error. This estimator is then differentiated by means of the score function approach which enables the estimation of the gradient of the failure probability without any additional call to the limit-state function (nor its Kriging emulator). The approach is validated on three structural reliability examples.

**Improved**

*Keywords:* Kriging, structural reliability, meta-model-based importance sampling, sensitivity analysis

---

## 1. Introduction

Modern engineering has to cope with uncertainty at the various stages of the design, manufacturing and operating of systems and structures. Such uncertainty either arise from the observed scattering of the environmental conditions in which products and structures will evolve, or from a lack of knowledge that results in the formulation of hopefully conservative assumptions. No matter their source, aleatory (observed) and epistemic (reducible) uncertainties can be dealt with in the unified framework of probabilistic methods for uncertainty quantification and risk-based engineering.

In particular, reliability analysis is the discipline which aims at quantifying the level of safety of a system in terms of a probability of failure. From now on, it is assumed that the uncertain parameters of the problem at hand are modelled by a random vector  $\mathbf{X}$  whose joint probability distribution is explicitly known and dependent on a certain number of design parameters grouped in the vector  $\mathbf{d}$ . In practice these design parameters are considered as mean values or, more generally, characteristic values of the random variables gathered in  $\mathbf{X}$ . This assumption corresponds to the common situation where  $\mathbf{d}$  gathers “ideal” dimensions whereas the randomness in  $\mathbf{X}$  models the aleatory uncertainty in the manufacturing process due to tolerancing. **Added**

It is also assumed that there exists a deterministic computational model  $\mathcal{M}$  which enables the assessment of the system’s performance through a so-called limit-state function  $\mathbf{g}$ . According to this setup, the failure probability is defined by the following integral:

$$p_f(\mathbf{d}) = \mathbb{P}[\mathbf{g}(\mathbf{X}, \mathcal{M}(\mathbf{X})) \leq 0 \mid \mathbf{d}] = \int_{\mathbb{F}} f_{\mathbf{X}}(\mathbf{x} \mid \mathbf{d}) \, d\mathbf{x}, \quad (1)$$

where  $\mathbb{F} = \{\mathbf{x} \in \mathbb{X} : \mathbf{g}(\mathbf{x}, \mathcal{M}(\mathbf{x})) \leq 0\}$  is the failure domain and  $f_{\mathbf{X}}$  is the joint probability density function of the random vector  $\mathbf{X}$ . The dependence of  $\mathbf{g}$  on the output of  $\mathcal{M}$  will now be dropped for the sake of clarity in the notation, but it is important to remember that each

---

\*Corresponding author

*Email addresses:* [dubourg@phimeca.com](mailto:dubourg@phimeca.com) (V. Dubourg), [sudret@ibk.baug.ethz.ch](mailto:sudret@ibk.baug.ethz.ch) (B. Sudret)

evaluation of  $\mathbf{g}$  implies a run of the possibly expensive-to-evaluate computational model  $\mathcal{M}$ . Note that in general the design parameters  $\mathbf{d}$  could also affect the limit state function itself. **Added** This case is not addressed in the present paper though. We assume here that parameters  $\mathbf{d}$  only affect  $\mathbf{X}$ 's distribution as in [1, 2, 3, 4] because it considerably eases the reliability sensitivity analysis through the use of the importance sampling trick initially proposed by Rubinstein [5].

Introducing the failure indicator function  $\mathbb{1}_{\mathbb{F}}$ , the failure probability easily rewrites as its mathematical expectation:

$$p_f(\mathbf{d}) = \int_{\mathbf{X}} \mathbb{1}_{\mathbb{F}}(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x} | \mathbf{d}) d\mathbf{x} = \mathbb{E}_{\mathbf{X}} [\mathbb{1}_{\mathbb{F}}(\mathbf{X})]. \quad (2)$$

This enables the evaluation of the failure probability by the so-called Monte Carlo estimator [1]:

$$\hat{p}_{f\text{MCS}} = \frac{1}{N} \sum_{i=1}^N \mathbb{1}_{\mathbb{F}}(\mathbf{X}^{(i)}), \quad (3)$$

where  $\{\mathbf{X}^{(i)}, i = 1, \dots, N\}$  is a sample of  $N$  independent copies of the random vector  $\mathbf{X}$ . Due to the central limit theorem, this estimator is unbiased and convergent and its coefficient of variation is defined as follows (provided  $p_f \neq 0$ ): **Improved**

$$\delta_{\text{MCS}} = \sqrt{\frac{1 - p_f}{N p_f}}. \quad (4)$$

From this expression, it appears that the lower the probability  $p_f$ , the greater the number  $N$  of evaluations of  $\mathbb{1}_{\mathbb{F}}$  (hence, the number of runs of  $\mathcal{M}$ ). As an order of magnitude, one should expect a minimum sample size of  $N = 10^{k+2}$  for estimating a failure probability of  $10^{-k}$  with a 10% coefficient of variation. This clearly becomes intractable for expensive-to-evaluate failure indicator functions and low failure probabilities, which are both the trademark of engineered systems. Nowadays there exists a number of techniques to evaluate the failure probability at a far reduced computational cost. **Improved** **Improved** **Removed** They are divided here into two categories in the sequel.

On the one hand, variance reduction techniques [1] aim at rewriting Eqn. (2) in order to derive new Monte-Carlo-sampling-based estimators that feature a lower coefficient of varia-

tion than the one given in Eqn. (4). Importance sampling [1], directional sampling [6], line sampling [7, 8] and **subset simulation** [9] all enable a great reduction of the computational cost compared to crude Monte Carlo sampling. Importance sampling and **subset simulation** are certainly the most widely applicable techniques because they are not based on any geometrical assumptions about the topology of the failure domain  $\mathbb{F}$ . Nonetheless, importance sampling is only a concept and still requires the choice for an instrumental density function which is not trivial and strongly influences both the accuracy and the computational cost. **Improved Improved**

On the other hand, approximation techniques make use of meta-models that imitates the limit-state function (or at least the limit-state surface  $\{\mathbf{x} \in \mathbb{X} : \mathbf{g}(\mathbf{x}) = 0\}$ ) in order to reduce the computational cost. These meta-models are built from a so-called design of experiments  $\{\mathbf{x}^{(i)}, i = 1, \dots, m\}$  whose size  $m$  does not depend on the order of magnitude of the failure probability but rather on the nonlinearity of the performance function  $\mathbf{g}$  and the dimension  $n$  of the input space  $\mathbb{X} \subseteq \mathbb{R}^n$ . For instance, quadratic response surfaces [10], artificial neural networks [11], support vector machines [12], Kriging surrogates [13] and **polynomial (resp. sparse polynomial) chaos expansions** [14, 15, 16] have been used for surrogate-based reliability analysis [see 17, for a review]. **Improved**

The most efficient variance reduction techniques (namely subset simulation) still require rather large sample sizes that can potentially be reduced when **some knowledge about the shape of failure domain exists**. Despite the increasing accuracy of meta-models, surrogate-based (also called *plug-in*) approaches that consists in using emulators instead of the actual limit-state functions lacks an error measure (alike the historical first- and second-order reliability methods (FORM/SORM)). Starting from these two premises, a novel hybrid technique named meta-model-based importance sampling was proposed by Dubourg et al. [18, 19]. This technique makes use of Kriging predictors in order to approximate the optimal instrumental density function in an importance sampling scheme that theoretically reduces the estimation variance to zero. **Improved**

This paper is not only concerned with the evaluation of the failure probability in Eqn. (1) for a single value of the parameters  $\mathbf{d}$ , but also with the analysis of its sensitivity with respect to the latter vector. Within the structural reliability community, this type of analysis is

referred to as reliability sensitivity analysis [20, 21]. It provides an important insight on system failure for risk-based decision making (*e.g.* robust control, design or reliability-based design optimization).

However, it was previously recalled that the accurate estimation of a single value of the failure probability was already computationally costly. Hence, assessing the failure probability sensitivity by means of repeated reliability analyses is absolutely not affordable. Starting from this premise, Au [22] proposed to consider the parameters  $\mathbf{d}$  as artificially uncertain, and then use a conditional sampling technique in order to assess reliability sensitivity within a single simulation. The authors conclude that their approach reveals efficient up to 2–3 parameters in  $\mathbf{d}$ . Based on a similar idea, Taflanidis and Beck [23] developed an algorithm which enables the identification of a reduced set of the parameters  $\mathbf{d}$  that minimizes the failure probability. It is applied to the robust control of the dynamic behaviour of structural systems.

Here, the objective is to get a more local guess of the influence of  $\mathbf{d}$  on the failure probability through the calculation of the gradient of the failure probability. This quantity then enables the use of gradient-based nonlinear constrained optimization algorithms for solving the reliability-based design optimization problem [24, 25, 26]. This topic has already been given a quite significant interest. For instance, Bjerager and Krenk [27] differentiated the Hasofer-Lind reliability index which itself enables the calculation of the gradient of the failure probability. However, FORM may suffer from incorrect assumptions (namely, the linearity of the limit-state surface in the standard space and the uniqueness of the most probable failure point) that are hard to check in practice. Valdebenito and Schuëller [28] propose a parametric approximation of the failure probability which then enables its differentiation. However the accuracy of this approach is **conditional** on the ability of the proposed parametric model to fit the actual failure probability function. **Improved**

The *score function* approach that is used here was initially proposed by Rubinstein [5] [see also 1, Chapter 7]. It features the double advantage that (*i*) it is a simple post-processing of a sampling-based reliability analysis (*i.e.* it does not require any additional calls to the limit-state function), and (*ii*) it can be applied to any Monte-Carlo-sampling-

based technique (either on the actual limit-state function or on a surrogate). For instance, it has already been applied to importance sampling [1] and subset simulation [3].

The goal of this paper is to show how the score function approach may be coupled with meta-model-based importance sampling for efficient reliability sensitivity analysis. The paper is divided into three sections. First, the basics of meta-model-based importance sampling are briefly summarized. Then, the score function approach is applied to the proposed estimator of the failure probability and the whole approach is eventually tested on three structural reliability examples.

## 2. Meta-model-based importance sampling for reliability analysis

This section recalls the basics of meta-model-based importance sampling. For a more exhaustive introduction, the interested reader is referred to the papers by Dubourg et al. [18, 19] and the Ph.D thesis of Dubourg [29].

### 2.1. Probabilistic classification using Kriging

Generally speaking, meta-modelling (*a.k.a.* supervised learning) techniques aim at building an emulator of a given function (here, the limit-state function  $\mathbf{g}$ ) from a dataset  $\mathcal{D} = \{(\mathbf{x}^{(i)}, y_i), i = 1, \dots, m\}$  where  $y_i = \mathbf{g}(\mathbf{x}^{(i)})$ . Kriging starts with the prior assumption that the function to imitate is a sample path of a Gaussian process [30]. This Gaussian process  $Y$  is further defined as follows:

$$Y(\mathbf{x}) = \mathbf{f}(\mathbf{x})^\top \boldsymbol{\beta} + U(\mathbf{x}), \quad \mathbf{x} \in \mathbb{X}, \quad (5)$$

where the mean  $\mathbf{f}(\mathbf{x})^\top \boldsymbol{\beta}$  is modelled as a generalized linear model whose functional basis  $\mathbf{f} = \{f_j, j = 1, \dots, P\}$  must be chosen and coefficients  $\boldsymbol{\beta}$  must be calculated. In the above equation  $U$  is a zero-mean **strongly** stationary Gaussian process, with variance  $\sigma_Y^2$  **Removed** and whose autocorrelation  $R$  is chosen and parametrized with  $\boldsymbol{\theta}$ . The most widely used class of autocorrelation function is the anisotropic squared exponential one which reads:

$$R(\mathbf{x} - \mathbf{x}') = \exp \left[ \sum_{i=1}^n - \left( \frac{x_i - x'_i}{\theta_i} \right)^2 \right], \quad (\mathbf{x}, \mathbf{x}') \in \mathbb{X}^2. \quad (6)$$

Hence, the observations vector  $\mathbf{Y} = (Y(\mathbf{x}^{(i)}), i = 1, \dots, m)$  and some unobserved value  $Y(\mathbf{x})$  at a new point  $\mathbf{x}$  are jointly distributed as follows:

$$\begin{Bmatrix} Y(\mathbf{x}) \\ \mathbf{Y} \end{Bmatrix} \sim \mathcal{N}_{1+m} \left( \begin{Bmatrix} \mathbf{f}(\mathbf{x})^\top \boldsymbol{\beta} \\ \mathbf{F} \boldsymbol{\beta} \end{Bmatrix}, \sigma^2 \begin{bmatrix} 1 & \mathbf{r}(\mathbf{x})^\top \\ \mathbf{r}(\mathbf{x}) & \mathbf{R} \end{bmatrix} \right), \quad \mathbf{x} \in \mathbb{X} \quad (7)$$

where  $\mathbf{F} = [f_j(\mathbf{x}^{(i)}), i = 1, \dots, m, j = 1, \dots, P]$  is the regression matrix,  $\mathbf{R} = [R(\mathbf{x}^{(i)} - \mathbf{x}^{(j)}), i, j = 1, \dots, m]$  is the correlation matrix and  $\mathbf{r}(\mathbf{x}) = (R(\mathbf{x} - \mathbf{x}^{(i)}), i = 1, \dots, m)^\top$  is the cross-correlation vector. The Kriging predictor basically exploits this non-zero cross-correlation in order to predict the unobserved value  $Y(\mathbf{x})$  given the observations in  $\mathbf{Y}$  through the calculation of the conditional (or posterior) distribution of  $[Y(\mathbf{x}) \mid \mathbf{Y} = \mathbf{y}]$ , **Improved** which reads: “the distribution of the unobserved value  $Y(\mathbf{x})$  given the observations  $\mathbf{y}$  in the dataset  $\mathcal{D}$ ”. Note that throughout this paper, random quantities are denoted with upper case letters while deterministic realizations are denoted with the corresponding lower case letters. Hence  $\mathbf{Y}$  is used for describing the (unconditional) joint model, and  $\mathbf{y}$  will now be used for denoting the set of model responses that have been evaluated.

Santner et al. [30] eventually prove that for given values of  $\boldsymbol{\theta}$  and  $\sigma_Y^2$ , this conditional distribution remains Gaussian with mean and variance:

$$\mu_{\hat{Y}}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^\top \hat{\boldsymbol{\beta}} + \mathbf{r}(\mathbf{x})^\top \mathbf{R}^{-1} (\mathbf{y} - \mathbf{F} \hat{\boldsymbol{\beta}}), \quad (8)$$

$$\sigma_{\hat{Y}}^2(\mathbf{x}) = \sigma_Y^2 \left( 1 - \begin{bmatrix} \mathbf{f}(\mathbf{x}) \\ \mathbf{r}(\mathbf{x}) \end{bmatrix}^\top \begin{bmatrix} \mathbf{0} & \mathbf{F}^\top \\ \mathbf{F} & \mathbf{R} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{f}(\mathbf{x}) \\ \mathbf{r}(\mathbf{x}) \end{bmatrix} \right), \quad \mathbf{x} \in \mathbb{X}, \quad (9)$$

where  $\hat{\boldsymbol{\beta}} = (\mathbf{F}^\top \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^\top \mathbf{R}^{-1} \mathbf{y}$  is the generalized least-squares solution of the underlying **Improved** linear regression problem.

The accuracy of the predictor  $\hat{Y}$  obviously depends on the choice of  $\boldsymbol{\theta}$  and  $\sigma_Y^2$ . State-of-the-art practices for determining the best values of these parameters are *maximum likelihood estimation* (MLE) according to the marginal distribution of  $\mathbf{Y}$  in Eqn. (7), and *cross-validation* (CV) techniques [see 31, Chapter 5]. In particular the variance  $\sigma_Y^2$  is estimated from the regression residuals as follows  $\hat{\sigma}_Y^2 = \frac{1}{m} (\mathbf{y} - \mathbf{F} \hat{\boldsymbol{\beta}})^\top \mathbf{R}^{-1} (\mathbf{y} - \mathbf{F} \hat{\boldsymbol{\beta}})$ . The applications presented hereafter resort to MLE.

The Gaussian probability measure attached to the Kriging predictor models the uncertainty about the predicted values depending on the distance from the test point  $\mathbf{x} \in \mathbb{X}$  to the points in the *design of experiments* (DOE)  $\mathcal{X} = \{\mathbf{x}^{(i)}, i = 1, \dots, m\}$ . This uncertainty is of epistemic nature, meaning that it can be reduced by making the DOE more dense in the input space  $\mathbb{X}$ . For usual covariance functions such as the one given in Eqn. (6), the Kriging prediction is consistent at all points in the DOE (*i.e.*  $\mu_{\hat{Y}}(\mathbf{x}^{(i)}) = y_i$  and  $\sigma_{\hat{Y}}(\mathbf{x}^{(i)}) = 0$  for  $i = 1, \dots, m$ ). Asymptotic consistency cannot be guaranteed though because it actually depends on the adequacy of the chosen stationary correlation model to the data [see 32, pp. 132–156].

If the sign of the emulated function is the only quantity of interest (as for reliability analysis), it can be predicted in a probabilistic manner due to the Gaussianity of the predictor  $\hat{Y}$ . In this respect, the *probabilistic classification function* is introduced:

$$\pi(\mathbf{x}) = \mathcal{P} \left[ \hat{Y}(\mathbf{x}) \leq 0 \right] = \Phi \left( \frac{0 - \mu_{\hat{Y}}(\mathbf{x})}{\sigma_{\hat{Y}}(\mathbf{x})} \right), \quad \mathbf{x} \in \mathbb{X}, \quad (10)$$

where  $\Phi$  denotes the *cumulative density function* (CDF) of the standard Gaussian distribution, and  $\mathcal{P}$  is the epistemic probability measure introduced by the Kriging prediction that shall not be confused with the probability measure  $\mathbb{P}$  attached to the random vector  $\mathbf{X}$ .

Figure 1 illustrates the aforementioned Kriging properties on the following one-dimensional function:  $x \mapsto y = x \sin(x)$  on the range  $[0; 15]$ . The big red dots in the upper panel constitutes the dataset  $\mathcal{D}$  obtained from the original function (red dotted line). The Kriging prediction is represented here by its mean (blue line) and its 95% confidence interval (blue shade area) which is bounded by  $\mu_{\hat{Y}}(\mathbf{x}) \pm 1.96 \sigma_{\hat{Y}}(\mathbf{x})$ ,  $\mathbf{x} \in \mathbb{X}$ . The probabilistic classification function  $\pi$  is plotted as the blue solid line in the lower panel and it is opposed to the failure indicator function  $\mathbb{1}_{\mathbb{F}}$  (red dotted line) which should be seen as its target. **Figure 1 was Improved** procued with the matplotlib and scikit-learn Python modules [33].

## 2.2. Proposed importance sampling scheme

Once the Kriging model is built, surrogate-based reliability methods usually make use of the mean prediction only for estimating the failure probability: a proxy failure domain



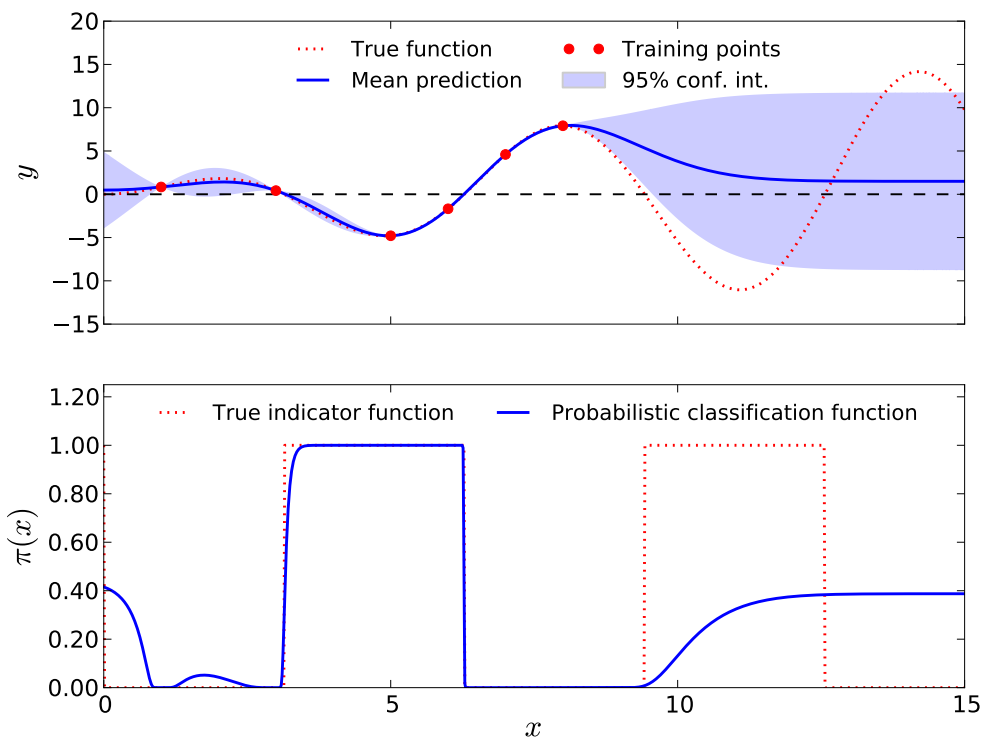


Figure 1: (Top) Kriging prediction for the function  $x \mapsto y = x \sin(x)$ . (Bottom) Probabilistic classification function  $\pi$ .

$\tilde{\mathbb{F}} = \{\mathbf{x} \in \mathbb{X} : \tilde{\mathbf{g}}(\mathbf{x}) \leq 0\}$  is used where the surrogate function (*i.e.* the Kriging mean prediction,  $\tilde{\mathbf{g}}(\mathbf{x}) \equiv \mu_{\hat{\gamma}}(\mathbf{x})$ ) is used instead of  $\mathbb{F}$  in Eqn. (1).

Picheny [34] proposes to use the probabilistic classification function  $\pi$  as a surrogate for the failure indicator function  $\mathbb{1}_{\mathbb{F}}$  in Eqn. (2). He thus defines the following quantity:

$$p_{f\varepsilon} = \int_{\mathbb{X}} \pi(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \mathbb{E}_{\mathbf{X}} [\pi(\mathbf{X})], \quad (11)$$

as an estimator of the failure probability. However, it is argued that this quantity does not equal the failure probability. Indeed, it actually sums both the Kriging epistemic uncertainty and the one attached to the random vector  $\mathbf{X}$ . This is the reason why Dubourg et al. [19] named this quantity the *augmented failure probability*.

In order to compute the *actual* failure probability, Dubourg et al. [19] propose to use an **proxy to the optimal importance sampling scheme, as summarized in the sequel**. First, recall **Improved** that importance sampling consists in using an instrumental density function which favours the outcome of the failure event  $\mathbb{F}$  while satisfying some requirements such as dominance over  $\mathbb{1}_{\mathbb{F}} f_{\mathbf{X}}$ , *i.e.* the instrumental density should be non-zero on the failure domain. Given such an admissible density  $h$ , the failure probability rewrites:

$$p_f = \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}}(\mathbf{x}) \frac{f_{\mathbf{X}}(\mathbf{x})}{h(\mathbf{x})} h(\mathbf{x}) d\mathbf{x} = \mathbb{E}_{\mathbf{Z}} \left[ \mathbb{1}_{\mathbb{F}}(\mathbf{Z}) \frac{f_{\mathbf{X}}(\mathbf{Z})}{h(\mathbf{Z})} \right], \quad (12)$$

where  $\mathbf{Z}$  is distributed according to  $h$ . It can easily be shown that the optimal instrumental density reads [1]:

$$h^*(\mathbf{x}) = \frac{\mathbb{1}_{\mathbb{F}}(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})}{\int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}}(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}} = \frac{\mathbb{1}_{\mathbb{F}}(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})}{p_f}, \quad (13)$$

and that the corresponding variance of estimation reduces to zero. However, the latter density cannot be used in practice because it involves both the failure probability (which is the sought quantity) and the supposedly expensive-to-evaluate failure indicator function.

Hence, Dubourg et al. [19] propose to approximate this optimal density using the probabilistic classification function as a surrogate for the failure indicator function. The **proposed Improved** instrumental density reads:

$$\hat{h}^* = \frac{\pi(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})}{\int_{\mathbb{X}} \pi(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}} = \frac{\pi(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})}{p_{f\varepsilon}}. \quad (14)$$

As opposed to its optimal counterpart, the latter approximation solely involves the Kriging predictor which is much less expensive to evaluate than the actual limit-state function.

Eventually, substituting  $\widehat{h}^*$  for  $h$  in Eqn. (12), it can be shown [see *e.g.* 19] that the failure probability recasts as follows:

$$p_f = p_{f\varepsilon} \alpha_{\text{corr}} \quad (15)$$

where  $\alpha_{\text{corr}}$  is the so-called correction factor defined as follows:

$$\alpha_{\text{corr}} = \mathbb{E}_{\mathbf{Z}} \left[ \frac{\mathbb{1}_{\mathbb{F}}(\mathbf{Z})}{\pi(\mathbf{Z})} \right], \quad (16)$$

where  $\mathbf{Z}$  is distributed according to  $\widehat{h}^*$ .

The latter correction factor measures the adequacy of the probabilistic classification function to the failure indicator function. The more accurate the Kriging predictor, the closer to unity  $\alpha_{\text{corr}}$ , and the smaller its estimation variance. In other words, it means that meta-model-based importance sampling does not require significantly more calls to the limit-state function when the Kriging predictor is accurate enough. But, in addition, it always provides unbiased estimates of the failure probability even if the Kriging predictor lacks accuracy as illustrated in Dubourg et al. [19].

Since both terms in Eqn. (15) can be estimated from two independent samples (see the next subsection), the overall coefficient of variation of the Meta-IS estimator reads as follows [see 19, Appendix 1]:

$$\delta_{\text{Meta-IS}} = \sqrt{\delta_\varepsilon^2 + \delta_{\text{corr}}^2 + \delta_\varepsilon^2 \delta_{\text{corr}}^2} \quad (17)$$

$$\underset{\delta_\varepsilon, \delta_{\text{corr}} \ll 1}{\approx} \sqrt{\delta_\varepsilon^2 + \delta_{\text{corr}}^2}, \quad (18)$$

where  $\delta_\varepsilon$  (resp.  $\delta_{\text{corr}}$ ) is the coefficient of variation of the augmented failure probability (resp. that of the correction factor).

### 2.3. Practical use: Meta-IS

First, in order to guarantee that the almost-optimal instrumental density  $\widehat{h}^*$  is sufficiently close to the optimal one, Dubourg et al. [18, 19] propose to use an adaptive refinement technique that consists in sequentially reducing the Kriging prediction variance in the vicinity of

the limit-state surface. Such strategies are already extensively used in the field of surrogate-based reliability analysis and other contour approximation problems [see *e.g.* 35, 13, 36]. **Improved** The particularity of the proposed refinement strategy lies in its ability to add batches of points to the design of experiments instead of a single (supposedly best) one. The adaptive refinement is stopped using a re-sampling estimate of the correction factor. Indeed, it stops when a kind of leave-one-out estimate of  $\alpha_{\text{corr}}$  is sufficiently close to 1 (within a maximal computational budget  $m \leq m_{\text{max}}$ ). **Improved** For further details the reader is referred to the original **Improved** paper [19].

Then the augmented failure probability is estimated using intensive Monte Carlo sampling. It should be noted that an arbitrarily small coefficient of variation  $\delta_\varepsilon$  is affordable here because the augmented failure probability involves the Kriging predictor only. If the augmented failure probability is smaller than  $10^{-6}$ , crude Monte Carlo sampling may be intractable (even on Kriging predictors). In this respect, Dubourg [29] proposed an *adaptive splitting* technique which makes the approach scalable to small failure probabilities. The idea is similar to the one used for subset simulation (see Appendix A).

Eventually, according to Eqn. (16), the correction factor is estimated by averaging the ratio between the failure indicator function and the probabilistic classification function over  $N_{\text{corr}}$  realizations  $\{\mathbf{Z}^{(i)}, i = 1, \dots, N_{\text{corr}}\}$ . These  $N_{\text{corr}}$  copies of  $\mathbf{Z}$ , which are distributed according to the almost optimal instrumental distribution  $\widehat{h}^*$ , are sampled using *Markov chain Monte Carlo* (MCMC).

The modified Metropolis-Hastings algorithm proposed by Au and Beck [9] in the space of independent standard Gaussian random variables can be used for this purpose. It is currently seeded with the points in the sample set used for estimating the augmented failure probability for which  $\pi$  is not zero. Note that the dependency between two successive steps of the MCMC sampling procedure implies an increase of the estimation variance [9]. To circumvent this drawback, it is proposed to resort to the *thinning* procedure that consists in retaining one sample every  $t$  states of the Markov chains [37]. Such a procedure is not affordable within subset simulation, because the intermediate instrumental densities involve the costly-to-evaluate failure indicator function. It is affordable here because the almost-

optimal instrumental density only involves the Kriging predictor. *Burn-in*, which consists in discarding an arbitrarily large number of increments at the beginning of the chains, is also used to ensure that the chains have reached  $\mathbf{Z}$ 's distribution as their stationary distributions despite the arbitrary seeding. The interested reader is referred to the book by Robert and Casella [38] and the thesis by Dubourg [29, Chapter 3 & Appendix B] for an extended discussion on the use of MCMC (Central Limit theorem for MCMC, most common MCMC samplers, burn-in and thinning procedures). Improved Improved Improved

### 3. Meta-model-based importance sampling for reliability sensitivity analysis

This section deals with the calculation of the partial derivatives of the failure probability with respect to the parameters  $\mathbf{d}$  in the probabilistic model using meta-model-based importance sampling together with the score function approach.

#### 3.1. The score function approach

First, the partial derivative of the failure probability with respect to the  $k$ -th component of  $\mathbf{d}$  reads:

$$\frac{\partial p_f(\mathbf{d})}{\partial d_k} = \frac{\partial}{\partial d_k} \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}}(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}, \mathbf{d}) \, d\mathbf{x}. \quad (19)$$

Assuming that (i) the joint PDF  $f_{\mathbf{X}}$  is continuously differentiable with respect to  $d_k$  and that (ii) the integration range  $\mathbb{X}$  does not depend on  $d_k$ , the partial derivative of the failure probability recasts as follows:

$$\frac{\partial p_f(\mathbf{d})}{\partial d_k} = \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}}(\mathbf{x}) \frac{\partial f_{\mathbf{X}}(\mathbf{x}, \mathbf{d})}{\partial d_k} \, d\mathbf{x}. \quad (20)$$

Note that the latter equation does not hold for the uniform distribution as the bounds of its integration range depend on its statistical moments [24, 25].

Then, in order to compute this integral as an expectation with respect to the same probability measure as the one used for estimating the failure probability itself, Rubinstein [5] genuinely proposed to use an *importance sampling trick* as in the previous section [see

also 1, Chapter 7]. It proceeds as follows:

$$\begin{aligned}
\frac{\partial p_f(\mathbf{d})}{\partial d_k} &= \int_{\mathbf{X}} \mathbb{1}_{\mathbb{F}}(\mathbf{x}) \frac{\partial f_{\mathbf{X}}(\mathbf{x}, \mathbf{d}) / \partial d_k}{f_{\mathbf{X}}(\mathbf{x}, \mathbf{d})} f_{\mathbf{X}}(\mathbf{x}, \mathbf{d}) \, d\mathbf{x} \\
&= \int_{\mathbf{X}} \mathbb{1}_{\mathbb{F}}(\mathbf{x}) \frac{\partial \log f_{\mathbf{X}}(\mathbf{x}, \mathbf{d})}{\partial d_k} f_{\mathbf{X}}(\mathbf{x}, \mathbf{d}) \, d\mathbf{x} \\
&= \mathbb{E}_{\mathbf{X}} [\mathbb{1}_{\mathbb{F}}(\mathbf{X}) \kappa_k(\mathbf{X}, \mathbf{d})],
\end{aligned} \tag{21}$$

where the so-called *score function*:

$$\kappa_k(\mathbf{X}, \mathbf{d}) = \frac{\partial \log f_{\mathbf{X}}(\mathbf{x}, \mathbf{d})}{\partial d_k} \tag{22}$$

has been introduced. Hence, given a sample  $\{\mathbf{X}^{(i)}, i = 1, \dots, N\}$  of  $N$  independent copies of the random vector  $\mathbf{X}$ , the following estimator:

$$\frac{\widehat{\partial p_f(\mathbf{d})}}{\partial d_k}_{\text{MCS}} = \frac{1}{N} \sum_{i=1}^N \mathbb{1}_{\mathbb{F}}(\mathbf{X}^{(i)}) \kappa_k(\mathbf{X}^{(i)}, \mathbf{d}) \tag{23}$$

is unbiased and asymptotically convergent according to the central limit theorem.

In addition to these first convenient properties, it can also be seen that the estimation of the failure probability (see Eqn. (3)) and that of its gradient can be done with the same  $N$ -sample. Hence, reliability sensitivity analysis through the score function approach is a *simple post-processing* of a Monte-Carlo-sampling-based reliability analysis. It does not require any additional sampling of the indicator function  $\mathbb{1}_{\mathbb{F}}$ . It should also be noticed that the approach extends to both (i) higher order derivatives, provided the joint PDF is sufficiently differentiable [1], (ii) and statistical moments of any order and of any variable [*i.e.* not only indicator functions, see 1, 2].

### 3.2. Application to Meta-IS

The score function approach has been successfully applied to other Monte-Carlo-sampling-based estimators such as the subset simulation estimator by Song et al. [3] [see also 29, pp. 161–163]. It is applied here to the meta-model-based importance sampling estimator.

First, recall that the failure probability was recast as the product of the augmented failure probability in Eqn. (11) and the correction factor in Eqn. (16). Then, the partial

derivative of the failure probability in Eqn. (15) with respect to the  $k$ -th design variable can easily be derived as that of a **sum** of two terms: Improved

$$\frac{\partial p_{f\varepsilon}(\mathbf{d})}{\partial d_k} = \frac{\partial \alpha_{\text{corr}}(\mathbf{d})}{\partial d_k} p_{f\varepsilon}(\mathbf{d}) + \alpha_{\text{corr}}(\mathbf{d}) \frac{\partial p_{f\varepsilon}(\mathbf{d})}{\partial d_k}. \quad (24)$$

If crude Monte Carlo sampling is used for estimating the augmented failure probability, its partial derivative can easily be obtained by substituting the probabilistic classification function  $\pi$  for the failure indicator function in Eqn. (21):

$$\frac{\partial p_{f\varepsilon}(\mathbf{d})}{\partial d_k} = \mathbb{E}_{\mathbf{X}} [\pi(\mathbf{X}) \kappa_k(\mathbf{X}, \mathbf{d})], \quad (25)$$

whose Monte Carlo estimator reads (see Eq. (23)):

$$\frac{\widehat{\partial p_{f\varepsilon}(\mathbf{d})}}{\partial d_k}_{\text{MCS}} = \frac{1}{N} \sum_{i=1}^N \pi(\mathbf{X}^{(i)}) \kappa_k(\mathbf{X}^{(i)}, \mathbf{d}). \quad (26)$$

If  $p_{f\varepsilon}$  is estimated by means of the adaptive splitting technique **detailed in Appendix A**, Improved its partial derivatives can be obtained in a similar fashion to the case of subset simulation [see 3, 29, Chapter 4].

The  $k$ -th partial derivative of the correction factor is obtained as follows. First, let us get back to its integral definition in Eqn. (16) and replace the almost-optimal instrumental density  $h^*$  by its expression:

$$\frac{\partial \alpha_{\text{corr}}(\mathbf{d})}{\partial d_k} = \int_{\mathbb{X}} \frac{\mathbb{1}_{\mathbb{F}}(\mathbf{x})}{\pi(\mathbf{x})} \frac{\partial}{\partial d_k} \left( \frac{\pi(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})}{p_{f\varepsilon}} \right) d\mathbf{x}. \quad (27)$$

Since the probabilistic classification function  $\pi$  does not depend on the parameters in  $\mathbf{d}$ , it further rewrites:

$$\frac{\partial \alpha_{\text{corr}}(\mathbf{d})}{\partial d_k} = \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}}(\mathbf{x}) \frac{\frac{\partial f_{\mathbf{X}}(\mathbf{x}, \mathbf{d})}{\partial d_k} p_{f\varepsilon} - f_{\mathbf{X}}(\mathbf{x}, \mathbf{d}) \frac{\partial p_{f\varepsilon}}{\partial d_k}}{p_{f\varepsilon}^2} d\mathbf{x}. \quad (28)$$

Using the same importance sampling trick as in Eqn. (21), one gets:

$$\frac{\partial \alpha_{\text{corr}}(\mathbf{d})}{\partial d_k} = \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}}(\mathbf{x}) \frac{\frac{\partial f_{\mathbf{X}}(\mathbf{x}, \mathbf{d})}{\partial d_k} p_{f\varepsilon} - f_{\mathbf{X}}(\mathbf{x}, \mathbf{d}) \frac{\partial p_{f\varepsilon}}{\partial d_k}}{p_{f\varepsilon}^2} \frac{\widehat{h}^*(\mathbf{x})}{\widehat{h}^*(\mathbf{x})} d\mathbf{x}. \quad (29)$$

This expression can be further simplified by replacing again the density  $\widehat{h}^*$  in the denominator on the right with its expression:

$$\frac{\partial \alpha_{\text{corr}}(\mathbf{d})}{\partial d_k} = \int_{\mathbb{X}} \mathbb{1}_{\mathbb{F}}(\mathbf{x}) \frac{\frac{\partial f_{\mathbf{X}}(\mathbf{x}, \mathbf{d})}{\partial d_k} p_{f_\varepsilon} - f_{\mathbf{X}}(\mathbf{x}, \mathbf{d}) \frac{\partial p_{f_\varepsilon}}{\partial d_k}}{p_{f_\varepsilon}^2} \frac{p_{f_\varepsilon}}{\pi(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}, \mathbf{d})} \widehat{h}^*(\mathbf{x}) \, d\mathbf{x}. \quad (30)$$

It finally reduces to:

$$\begin{aligned} \frac{\partial \alpha_{\text{corr}}(\mathbf{d})}{\partial d_k} &= \int_{\mathbb{X}} \frac{\mathbb{1}_{\mathbb{F}}(\mathbf{x})}{\pi(\mathbf{x})} \left( \kappa_k(\mathbf{x}, \mathbf{d}) - \frac{\frac{\partial p_{f_\varepsilon}(\mathbf{d})}{\partial d_k}}{p_{f_\varepsilon}} \right) \widehat{h}^*(\mathbf{x}) \, d\mathbf{x}, \\ &= \mathbb{E}_{\mathbf{Z}} \left[ \frac{\mathbb{1}_{\mathbb{F}}(\mathbf{Z})}{\pi(\mathbf{Z})} \left( \kappa_k(\mathbf{Z}, \mathbf{d}) - \frac{\frac{\partial p_{f_\varepsilon}(\mathbf{d})}{\partial d_k}}{p_{f_\varepsilon}} \right) \right], \end{aligned} \quad (31)$$

where  $\mathbf{Z}$  is distributed according to  $\widehat{h}^*$  as for  $\alpha_{\text{corr}}$  (see Eqn. (16)). This means that the partial derivative of the correction factor can be estimated using the same sample  $\{\mathbf{Z}^{(i)}, i = 1, \dots, N_{\text{corr}}\}$  as when evaluating  $\alpha_{\text{corr}}$ .

To conclude, Eqn. (24) can be estimated without any additional call to the failure indicator function (nor to the Kriging predictor) provided one has saved the failed samples and the score function can be calculated.

### 3.3. Calculation of the score function

Let us denote by  $F_{\mathbf{X}}$  the joint cumulative distribution function (CDF) of the input random vector  $\mathbf{X}$ , which is parametrized by some deterministic vector  $\mathbf{d}$ . In order to derive the score function in the general case, we use the *copula* representation of  $F_{\mathbf{X}}$  which has been recently popularized in the field of structural reliability by Lebrun and Dutfoy [39, 40].

According to Sklar's theorem [41], and omitting the dependency in  $\mathbf{d}$  for the time being, the joint CDF may be cast as a combination of a *copula function*  $C : \mathbf{u} \in [0, 1]^n \mapsto \mathbb{R}$  and the set of *marginal distributions*  $\{F_{X_i}, i = 1, \dots, n\}$  as follows:

$$F_{\mathbf{X}}(\mathbf{x}) = C(F_{X_1}(x_1), \dots, F_{X_n}(x_n)). \quad (32)$$

Accordingly the joint probability density function reads:

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}) &\equiv \frac{\partial^n F_{\mathbf{X}}(\mathbf{x})}{\partial x_1 \dots \partial x_n} \\ &= c(F_{X_1}(x_1), \dots, F_{X_n}(x_n)) \prod_{i=1}^n f_{X_i}(x_i), \end{aligned} \quad (33)$$



where  $c : \mathbf{u} \in [0, 1]^n \mapsto \mathbb{R}$  is the *copula density function* defined by:

$$c(\mathbf{u}) = \frac{\partial^n C(\mathbf{u})}{\partial u_1 \dots \partial u_n} \quad \mathbf{u} = \{u_1, \dots, u_n\} \quad (34)$$

and the marginal PDFs are denoted by  $\{f_{X_i}, i = 1, \dots, n\}$ . Note that in the case when the components of  $\mathbf{X}$  are independent, the joint CDF reduces to the simple product of the marginal distributions. In this case the copula function is the so-called *independent copula*  $C_{\text{ind}}(u_1, \dots, u_n) = u_1 \dots u_n$  and the copula density is constant and equal to  $c_{\text{ind}}(u_1, \dots, u_n) = 1$  over the whole  $n$ -hypercube  $[0; 1]^n$ .

We now assume that vector  $\mathbf{d}$  only affects the marginal distribution, *i.e.* that the copula function  $C(\cdot)$  is not directly dependent on  $\mathbf{d}$ . Moreover we assume that the  $k$ -th component of vector  $\mathbf{d}$ , say  $d_k$ , is involved in the  $k$ -th marginal distribution *only*. Indeed, as observed in the introduction,  $d_k$  is usually the mean value of the marginal distribution of  $X_k$ .

Re-introducing the dependency in  $\mathbf{d}$  for the sake of clarity, the  $k$ -th score function reads:

$$\begin{aligned} \kappa_k(\mathbf{x}, \mathbf{d}) &= \frac{\partial \log f_{\mathbf{X}}(\mathbf{x}, \mathbf{d})}{\partial d_k} \\ &= \frac{\partial \log c(F_{X_1}(x_1), \dots, F_{X_n}(x_n))}{\partial u_k} \frac{\partial F_{X_k}(x_k, d_k)}{\partial d_k} + \frac{\partial \log f_{X_k}(x_k, d_k)}{\partial d_k}. \end{aligned} \quad (35)$$

The latter equation simply reduces to the derivative of the logarithm of the  $k$ -th marginal PDF in the case of independence (since  $c_{\text{ind}}$  is constant over the whole unit  $n$ -hypercube) as intuited.

The expressions for the partial derivatives of usual parametric copulas densities, marginal PDFs and CDFs can often be obtained analytically [see *e.g.* 24, 25, 29, pp. 164–167]. If analytical expressions are not available for these quantities, they can be evaluated numerically using a finite difference scheme.

#### 4. Application Examples

The meta-model-based importance sampling procedure and the score function approach are now applied to a number of examples in structural reliability for the sake of validation and illustration. The results are compared to that obtained by FORM [27] or subset simulation [9, 3]. Some reliability sensitivity results are compared here in terms of the so-called

*elasticities* [see *e.g.* 42]. The elasticity of the failure probability with respect to a non-zero parameter  $d$  is defined as follows:

$$e_d = \frac{\partial p_f}{\partial d} \frac{d}{p_f}, \quad d \neq 0. \quad (36)$$

This dimensionless sensitivity measure enables a more objective conclusion when the random variables parameters are heterogeneous in dimension. This measure is also less sensitive the potential bias in the failure probability estimate (*e.g.* when the FORM approximation is not correct). **Improved**

#### 4.1. Example 1: capacity & demand

The proposed technique is first applied to a simple capacity & demand example for which both the failure probability and its partial derivatives are known analytically. The performance function reads:

$$\mathbf{g}(r, s) = r - s, \quad (37)$$

where  $r$  is the outcome of the random capacity  $R$ , and  $s$  is that of the random demand  $S$ . Capacity and demand are grouped in the input random vector  $\mathbf{X} = (R, S)^\top$ .  $R$  (resp.  $S$ ) is lognormally distributed with mean  $\mu_R = 7$  (resp.  $\mu_S = 1$ ) and standard deviation  $\sigma_R = 0.5$  (resp.  $\sigma_S = 0.5$ ). Both variables are independent.

Thanks to the linearity of the performance function  $\mathbf{g}$  and the independence of  $R$  and  $S$ , the Hasofer-Lind reliability comes easily and reads as follows:

$$\beta_{\text{HL}} = \frac{\mu_R - \mu_S}{\sqrt{\sigma_R^2 + \sigma_S^2}}. \quad (38)$$

The corresponding exact failure probability is  $p_f = \Phi(-\beta_{\text{HL}})$ . The partial derivatives of the failure probability with respect to any parameter  $d$  in the probabilistic model is given by:

$$\frac{\partial p_f}{\partial d} = -\frac{\partial \beta_{\text{HL}}}{\partial d} \varphi(-\beta_{\text{HL}}), \quad (39)$$

where  $\varphi$  denotes the standard Gaussian PDF and the partial derivatives of the Hasofer-Lind reliability index can easily be obtained from Eqn. (38).

Method	Analytical	FORM	Monte Carlo	Subset	Meta-IS
$p_f$	$1.10 \times 10^{-5}$	$1.20 \times 10^{-5}$	$1.05 \times 10^{-5}$	$1.12 \times 10^{-5}$	$1.07 \times 10^{-5}$
c.o.v.	–	–	< 5%	< 5%	< 5%
$\partial p_f / \partial \mu_R$	$-3.48 \times 10^{-5}$	$-3.75 \times 10^{-5}$	$-3.31 \times 10^{-5}$	$-3.53 \times 10^{-5}$	$-3.40 \times 10^{-5}$
$\partial p_f / \partial \mu_S$	$3.48 \times 10^{-5}$	$3.75 \times 10^{-5}$	$3.33 \times 10^{-5}$	$3.55 \times 10^{-5}$	$3.37 \times 10^{-5}$
$\partial p_f / \partial \sigma_R$	$1.04 \times 10^{-4}$	$1.12 \times 10^{-4}$	$9.91 \times 10^{-5}$	$1.06 \times 10^{-4}$	$1.04 \times 10^{-4}$
$\partial p_f / \partial \sigma_S$	$1.04 \times 10^{-4}$	$1.12 \times 10^{-4}$	$9.97 \times 10^{-5}$	$1.07 \times 10^{-4}$	$1.03 \times 10^{-4}$
# <b>g</b> -calls	–	61	$38 \times 10^6$	451,721	32 + 100

Table 1: Results for Example 1.

It can be seen from the numerical results in the first column of Table 1, that an increase of the mean resistance  $\mu_R$  would decrease the failure probability while an increase of the mean demand would increase the failure probability. An increase of uncertainty through either of the two standard deviations would also increase the failure probability. It can also be noted from the derivatives absolute values that  $R$  and  $S$  have exactly the same impact on the failure probability since the problem is symmetric. Decreasing the standard deviations seems more efficient than changing the means, but this conclusion depends on the associated cost of improvement. Indeed, decreasing tolerances might be more expensive than changing means.

The results obtained by the three reliability methods are given in Table 1. FORM results are of course in agreement with the analytical reference solution because the limit-state surface is linear in the standard space. The insignificant discrepancy on  $\partial p_f / \partial \mu_S$  is only due here to the iHLRF algorithm [43] which is used to find the most probable failure point in FORM. The number of **g**-calls accounts for the calculation of the gradient of the limit-state function using finite differences. Sampling methods all yield results in reasonable agreement with the reference. Despite the score function approach applies to Monte Carlo sampling, it can be seen that its convergence requires a large sample size which makes it prohibitive for expensive-to-evaluate limit-state functions. Subset simulation enables a significant reduction of this computational cost but still requires a few hundreds of thousands of calls to achieve

the same coefficient of variation than Meta-IS (the subset simulation coefficient of variation given here is the lower bound provided by Au and Beck [9]). Meta-IS enables an even larger variance reduction in this simple (linear, two-dimensional) case as it can be shown from the total number of  $\mathbf{g}$ -calls ( $m + N_{\text{corr}}$ ). The Kriging predictor emulates accurately the actual limit-state function as measured by the correction factor which equals 1 here. The proposed Meta-IS estimator may seem expensive on this simple example (with respect to FORM or plug-in estimators), but it used here as if it was not so simple, assuming the user has no expertise on the problem at hand.

#### 4.2. Example 2: A roof truss

This second example is taken from the paper by Song et al. [3, Example 2]. It consists in studying the serviceability limit-state of a roof truss with respect to its maximal vertical displacement (see Figure 2). The limit-state function is defined as follows:

$$\mathbf{g}(q, l, A_c, E_c, A_s, E_s) = 0.03 - \frac{ql^2}{2} \left( \frac{3.81}{A_c E_c} + \frac{1.13}{A_s E_s} \right), \quad (40)$$

where  $q$  is the magnitude of a uniformly distributed load,  $l$  is the roof span and  $(E_s, A_s)$  and  $(E_c, A_c)$  denote the Young's modulus and the cross section areas of the steel and concrete beams respectively. All these quantities are modelled by independent normal random variables whose distributions are given in Table 2, and the admissible deflection is arbitrarily fixed to 0.03 m as in the reference article. Note that the probabilistic modelling is, strictly speaking, not appropriate since  $l, A_c, E_c, A_s$  and  $E_s$  should remain positive. Despite log-normal distributions would be more relevant here, we use the same Gaussian model as in the reference paper for the sake of comparison.

The results are given in Table 3. Despite the limit-state function is not linear in the standard space, the limit-state surface is close to an hyperplane so that the FORM approximation reveals rather accurate. Since the failure probability is quite large here, subset simulation does not reveal efficient compared to crude Monte Carlo sampling. Meta-IS is again very efficient for estimating the failure probability up to the required 2% coefficient of variation because the Kriging predictor accurately fits the limit-state surface. One could

Variable	Mean	C.o.v.
$q$ (N/m)	20,000	7%
$l$ (m)	12	1%
$A_s$ (m <sup>2</sup> )	$9.82 \times 10^{-4}$	6%
$A_c$ (m <sup>2</sup> )	$400 \times 10^{-4}$	12%
$E_s$ (N/m <sup>2</sup> )	$1 \times 10^{11}$	6%
$E_c$ (N/m <sup>2</sup> )	$2 \times 10^{10}$	6%

Table 2: Probabilistic model for Example 2.

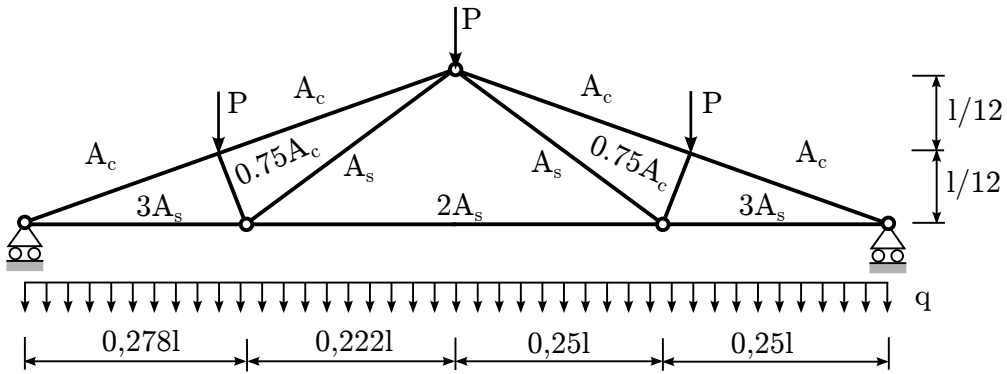


Figure 2: Example 2: the roof truss reliability problem according to Song et al. [3].

argue that the elasticities in column “Meta-IS #1” are not in agreement with that obtained by the other reliability methods. Indeed, despite the signs and ranking are well captured, the values are somewhat different. In this respect, we proved convergence by raising the sample size for estimating the correction factor to  $N_{\text{corr}} = 1,000$  and its sensitivities in the second stage. The results in column “Meta-IS #2” validate the approach. The corresponding results are in average closer to those obtained by Monte Carlo simulation than those obtained by FORM. This shows that despite it is hard to calculate, the coefficient of variation of the partial derivatives of the failure probability is higher than that of the failure probability itself.

Failure appears most sensitive to the mean of the roof span  $\mu_l$  and the load  $\mu_q$ . An increase of these two parameters would increase the failure probability since their elasticities are positive. Then the parameters with significant sensitivity are the mean values of the cross sections and Young’s moduli. The corresponding four elasticities are negative, meaning that an increase (*e.g.* a stiffer structure) would decrease the failure probability. This makes sense regarding what these two variables physically mean. **Added**

### 4.3. Example 3: A system reliability problem

This example was first proposed by Waarts [44, AE12]. It consists in a two-dimensional four-branch series system reliability problem whose limit-state function is defined as follows:

$$\mathbf{g}(\mathbf{x}) = \min \begin{pmatrix} 3 + (x_1 - x_2)^2/10 - (x_1 + x_2)/\sqrt{2} \\ 3 + (x_1 - x_2)^2/10 + (x_1 + x_2)/\sqrt{2} \\ x_1 - x_2 + 7/\sqrt{2} \\ x_2 - x_1 + 7/\sqrt{2} \end{pmatrix}, \quad \mathbf{x} \in \mathbb{X}. \quad (41)$$

where  $\mathbf{x}$  is the realization of a standard Gaussian random vector  $\mathbf{X} \sim \mathcal{N}_2(\mathbf{0}, \mathbf{I}_2)$ . This problem is illustrated in Figure 3. From this figure, it can be seen that the failure probability should feature zero derivatives with respect to the means since the failure domain surrounds the mean. The zero mean design minimizes the failure probability. The only way to reduce it would be to decrease the standard deviations.

Method	FORM	Monte Carlo	Subset	Meta-IS #1	Meta-IS #2
$p_f$	$7.58 \times 10^{-3}$	$9.44 \times 10^{-3}$	$9.48 \times 10^{-3}$	$9.54 \times 10^{-3}$	$9.41 \times 10^{-3}$
c.o.v.	–	2%	2%	2%	2%
$e_{\mu_q}$	24.44	23.94	23.71	23.30	23.93
$e_{\mu_t}$	53.77	49.46	51.26	54.24	50.48
$e_{\mu_{A_s}}$	–20.34	–19.09	–19.38	–19.65	–19.13
$e_{\mu_{A_c}}$	–9.02	–9.12	–9.02	–5.82	–8.76
$e_{\mu_{E_s}}$	–20.34	–19.63	–19.61	–20.15	–19.90
$e_{\mu_{E_c}}$	–8.20	–7.82	–7.98	–11.33	–8.12
$e_{\sigma_q}$	2.58	2.46	2.38	2.06	2.45
$e_{\sigma_t}$	0.25	0.20	0.21	0.07	0.18
$e_{\sigma_{A_s}}$	1.31	1.22	1.24	1.49	1.15
$e_{\sigma_{A_c}}$	1.03	1.27	1.28	0.41	1.22
$e_{\sigma_{E_s}}$	1.31	1.30	1.29	1.49	1.35
$e_{\sigma_{E_c}}$	0.21	0.21	0.27	0.80	0.32
# <b>g</b> -calls	112	262 350	299, 977	84 + 50	84 + 1, 000

Table 3: Results for Example 2.

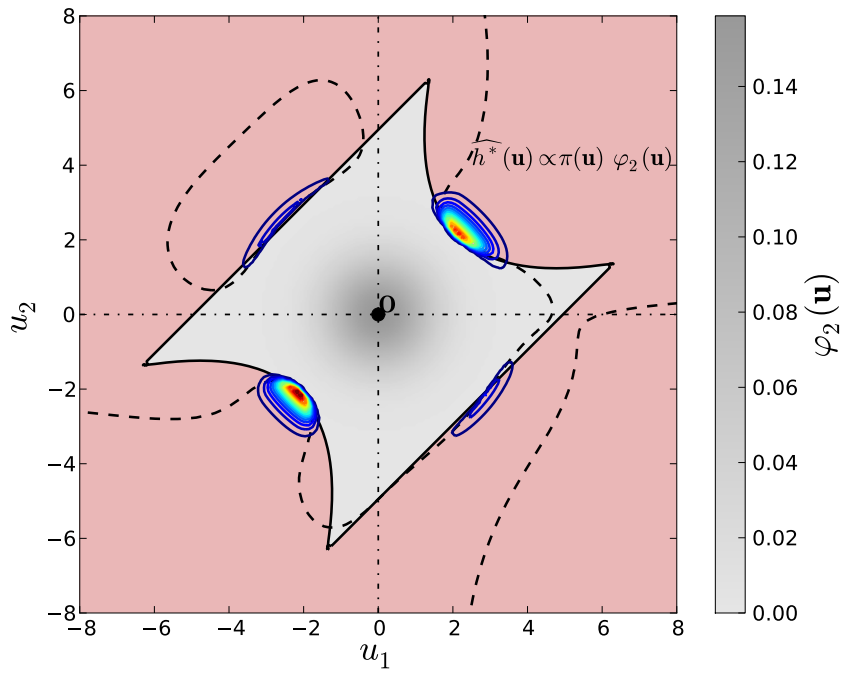


Figure 3: Example 3: the four-branch series system reliability problem by Waarts [44].



Method	FORM	Monte Carlo	Subset	Meta-IS
$p_f$	$1.35 \times 10^{-3}$	$2.17 \times 10^{-3}$	$2.29 \times 10^{-3}$	$2.75 \times 10^{-3}$
c.o.v.	–	< 5%	< 5%	< 5%
$\partial p_f / \partial \mu_1$	$3.13 \times 10^{-3}$	$1.04 \times 10^{-4}$	$5.64 \times 10^{-5}$	$-5.84 \times 10^{-5}$
$\partial p_f / \partial \mu_2$	$3.13 \times 10^{-3}$	$1.75 \times 10^{-4}$	$1.86 \times 10^{-4}$	$-4.24 \times 10^{-4}$
$\partial p_f / \partial \sigma_1$	$6.64 \times 10^{-3}$	$1.15 \times 10^{-2}$	$1.20 \times 10^{-2}$	$1.39 \times 10^{-2}$
$\partial p_f / \partial \sigma_2$	$6.64 \times 10^{-3}$	$1.11 \times 10^{-2}$	$1.22 \times 10^{-2}$	$1.53 \times 10^{-2}$
# <b>g</b> -calls	13	190 000	277, 346	64 + 500

Table 4: Results for Example 3.

Results are given in Table 4. Even if FORM is inapplicable there due to the 4 design points, the bias in the failure probability is not very strong because the contribution of the three secondary design points is not that significant for this standard normal probability distribution. Reducing the standard deviations would make the probability lower and emphasize the bias. However, even with this unit-variance distribution, it can be seen that FORM-based sensitivities are not able to discriminate the influence of the means and that of the standard deviations while the other sampling-based techniques (including Meta-IS) do. Indeed, sampling-based techniques did not find out that the partial derivatives of the failure probability are zero, but they did find that the means are 100 times less influent than the standard deviations, a feature that FORM was not able to investigate. Contours of the final Meta-IS instrumental PDF are illustrated in Figure 3 together with the mean surrogate limit-state surface (dashed black line). Despite the Kriging approximation does not exactly fit the true limit-state surface, it enables a good approximation of the optimal instrumental PDF that explores all the four failure modes.

## 5. Conclusion

The score function approach proposed by Rubinstein [5] reveals an efficient tool for reliability sensitivity analysis. First, using an importance-sampling-like trick, he showed Improved that the gradient of the failure probability turns out to be the expectation of the gradient

of the log-likelihood of the failed samples with respect to the distribution of the random vector  $\mathbf{X}$ . It means that reliability sensitivities can readily be obtained after a Monte-Carlo-sampling-based reliability at a negligible computational cost (no additional call the limit-state function).

Then, Song et al. [3] showed that the score function approach applies to any other sampling-based reliability methods such as subset simulation. In the present work, the authors derived the gradient of their proposed Meta-IS estimator. This enables a great variance reduction in the estimation of both the failure probability and its gradient as illustrated through the examples. Improved

These conclusions imply that reliability-based design optimization using Meta-IS is now conceivable. Indeed, as Meta-IS is able to calculate both the failure probability and its gradient, it can be used within an optimization loop for automating the search of the best compromise between cost and failure probability. However, in order to get the best efficiency (minimize the number of calls to the limit-state function), one should first think about an efficient recycling of the design of experiments from one nested reliability analysis to the other (if the change in the design is not too big). Indeed, building the Kriging surrogate from scratch at each iteration of the optimizer would be computationally inefficient [26]. Improved

## References

- [1] Rubinstein, R., Kroese, D.. Simulation and the Monte Carlo method. Wiley Series in Probability and Statistics; Wiley; 2008.
- [2] Millwater, H.. Universal properties of kernel functions for probabilistic sensitivity analysis. Prob Eng Mech 2009;24(1):89 – 99.
- [3] Song, S., Lu, Z., Qiao, H.. Subset simulation for structural reliability sensitivity analysis. Reliab Eng Sys Safety 2009;94(2):658–665.
- [4] Yuan, X.K.. Local estimation of failure probability function by weighted approach. Prob Eng Mech 2013;34(1):1–11.
- [5] Rubinstein, R.. The score function approach for sensitivity analysis of computer simulation models. Mathematics and Computers in Simulation 1986;28(5):351–379.
- [6] Bjerager, P.. Probability integration by directional simulation. J Eng Mech 1988;114(8):1285–1302.

- [7] Koutsourelakis, P., Pradlwarter, H., Schuëller, G.. Reliability of structures in high dimensions, part I: algorithms and applications. *Prob Eng Mech* 2004;19:409–417.
- [8] Pradlwarter, H., Schuëller, G., Koutsourelakis, P., Charmpis, D.. Application of line sampling simulation method to reliability benchmark problems. *Structural Safety* 2007;29(3):208–221.
- [9] Au, S.K., Beck, J.. Estimation of small failure probabilities in high dimensions by subset simulation. *Prob Eng Mech* 2001;16(4):263–277.
- [10] Bucher, C., Bourgund, U.. A fast and efficient response surface approach for structural reliability problems. *Structural Safety* 1990;7(1):57–66.
- [11] Hurtado, J.E., Alvarez, D.A.. Reliability assessment of structural systems using neural networks. In: *Proc. European Congress on Computational Methods in Applied Sciences and Engineering, ECCOMAS 2000 (Barcelona, 11-14 Sept. 2000)*. 2000, Paper #290.
- [12] Hurtado, J.. *Structural reliability – Statistical learning perspectives; vol. 17 of Lecture notes in Applied and Computational Mechanics*. Springer; 2004.
- [13] Bichon, B., Eldred, M., Swiler, L., Mahadevan, S., McFarland, J.. Efficient global reliability analysis for nonlinear implicit performance functions. *AIAA Journal* 2008;46(10):2459–2468.
- [14] Berveiller, M., Sudret, B., Lemaire, M.. Stochastic finite elements: a non intrusive approach by regression. *Eur J Comput Mech* 2006;15(1-3):81–92.
- [15] Blatman, G., Sudret, B.. Sparse polynomial chaos expansions and adaptive stochastic finite elements using a regression approach. *Comptes Rendus Mécanique* 2008;336(6):518–523.
- [16] Blatman, G., Sudret, B.. An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. *Prob Eng Mech* 2010;25:183–197.
- [17] Sudret, B.. Meta-models for structural reliability and uncertainty quantification. In: Phoon, K., Beer, M., Quek, S., Pang, S., editors. *Proc. 5th Asian-Pacific Symp. Struct. Reliab. (APSSRA’2012)*, Singapore. 2012, p. 53–76. Keynote lecture.
- [18] Dubourg, V., Deheeger, F., Sudret, B.. Metamodel-based importance sampling for the simulation of rare events. In: Faber, M., Köhler, J., Nishijima, K., editors. *Proc. 11th Int. Conf. on Applications of Stat. and Prob. in Civil Engineering (ICASP11)*, Zurich, Switzerland. 2011,.
- [19] Dubourg, V., Sudret, B., Deheeger, F.. Metamodel-based importance sampling for structural reliability analysis. *Prob Eng Mech* 2013;33:47–57.
- [20] Wu, Y.. Computational methods for efficient structural reliability and reliability sensitivity analysis. *AIAA Journal* 1994;32(8):1717–1723.
- [21] Valdebenito, M., Schuëller, G.. A survey on approaches for reliability-based optimization. *Struct Multidisc Optim* 2010;42:645–663.
- [22] Au, S.K.. Reliability-based design sensitivity by efficient simulation. *Computers & Structures*

- 2005;83(14):1048–1061.
- [23] Taflanidis, A., Beck, J.. Stochastic subset optimization for reliability optimization and sensitivity analysis in system design. *Computers & Structures* 2009;87(5-6):318–331.
  - [24] Lee, I., Choi, K., Noh, Y., Zhao, L., Gorsich, D.. Sampling-based stochastic sensitivity analysis using score functions for RBDO problems with correlated random variables. *J Mech Des* 2011;133(2):10.
  - [25] Lee, I., Choi, K., Zhao, L.. Sampling-based RBDO using the stochastic sensitivity analysis and Dynamic Kriging method. *Struct Multidisc Optim* 2011;44(3):299–317.
  - [26] Dubourg, V., Sudret, B., Bourinet, J.M.. Reliability-based design optimization using kriging and subset simulation. *Struct Multidisc Optim* 2011;44(5):673–690.
  - [27] Bjerager, P., Krenk, S.. Parametric sensitivity in first order reliability theory. *J Eng Mech* 1989;115(7):1577–1582.
  - [28] Valdebenito, M., Schuëller, G.. Efficient strategies for reliability-based optimization involving non-linear, dynamical structures. *Computers & Structures* 2011;89(19–20):1797–1811.
  - [29] Dubourg, V.. Adaptive surrogate models for reliability analysis and reliability-based design optimization. Ph.D. thesis; Université Blaise Pascal, Clermont-Ferrand, France; 2011. URL <http://tel.archives-ouvertes.fr/tel-00697026>.
  - [30] Santner, T., Williams, B., Notz, W.. The design and analysis of computer experiments. Springer series in Statistics; Springer; 2003.
  - [31] Rasmussen, C., Williams, C.. Gaussian processes for machine learning. Adaptive computation and machine learning; Cambridge, Massachusetts: MIT Press; Internet ed.; 2006.
  - [32] Vazquez, E.. Modélisation comportementale de systèmes non-linéaires multivariés par méthodes à noyaux et applications. Ph.D. thesis; Université Paris XI - UFR scientifique d’Orsay; 2005.
  - [33] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 2011;12:2825–2830.
  - [34] Picheny, V.. Improving accuracy and compensating for uncertainty in surrogate modeling. Ph.D. thesis; University of Florida; 2009.
  - [35] Au, S.K.. Augmenting approximate solutions for consistent reliability analysis. *Prob Eng Mech* 2007;22(1):77–87.
  - [36] Picheny, V., Ginsbourger, D., Roustant, O., Haftka, R.. Adaptive designs of experiments for accurate approximation of a target region. *J Mech Des* 2010;132(7):071008.
  - [37] Miao, F., Ghosn, M.. Modified subset simulation method for reliability analysis of structural systems. *Structural Safety* 2011;33:251–260.
  - [38] Robert, C., Casella, G.. Monte Carlo statistical methods (2<sup>nd</sup> Ed.). Springer Series in Statistics; Springer Verlag; 2004.

- [39] Lebrun, R., Dutfoy, A.. An innovating analysis of the Nataf transformation from the copula viewpoint. *Prob Eng Mech* 2009;24(3):312–320.
- [40] Lebrun, R., Dutfoy, A.. Do Rosenblatt and Nataf isoprobabilistic transformations really differ? *Prob Eng Mech* 2009;24(4):577–584.
- [41] Sklar, A.. Fonctions de répartition à  $n$  dimensions et leurs marges. *Publications de l’Institut de Statistique de L’Université de Paris* 8 1959;8(1):11.
- [42] Lemaire, M.. *Structural Reliability*. Wiley; 2009.
- [43] Zhang, Y., Der Kiureghian, A.. Two improved algorithms for reliability analysis. In: Rackwitz, R., Augusti, G., Bori, A., editors. *Proc. 6th IFIP WG7.5 on Reliability and Optimization of Structural systems*, Assisi, Italy. Chapman & Hall, London; 1995,.
- [44] Waarts, P.H.. *Structural reliability using finite element methods: an appraisal of DARS: Directional Adaptive Response Surface Sampling*. Ph.D. thesis; Technical University of Delft, The Netherlands; 2000.

## Appendix A. Estimation of the augmented failure probability using an adaptive splitting technique

As previously stated in this paper, crude Monte Carlo sampling may reveal rather inefficient if the augmented failure probability is low (even on Kriging predictors). Instead, it is proposed to adapt the subset simulation scheme proposed by Au and Beck [9] to the problem of estimating the augmented failure probability in Eqn. (11).

Given a strictly decreasing sequence of quantiles  $q_1 > \dots > q_s = 0$ , Eqn. (11) rewrites in the following split form:

$$p_{f\varepsilon} = p_{1\varepsilon} \prod_{i=2}^s p_{i|i-1\varepsilon}. \quad (\text{A.1})$$

where the intermediate *augmented* probabilities read as follows:

$$p_{1\varepsilon} = \int_{\mathbb{X}} \pi_1(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}, \quad (\text{A.2})$$

$$p_{i|i-1\varepsilon} = \int_{\mathbb{X}} \frac{\pi_i(\mathbf{x})}{\pi_{i-1}(\mathbf{x})} h_{i\varepsilon}^*(\mathbf{x}) d\mathbf{x}, \quad i = 2, \dots, s. \quad (\text{A.3})$$

The latter expression involves the collection of *approximate* intermediate indicator functions that are defined as follows:

$$\pi_i(\mathbf{x}) = \mathcal{P} \left[ \widehat{Y}(\mathbf{x}) \leq q_i \right] = \Phi \left( \frac{q_i - \mu_{\widehat{Y}}(\mathbf{x})}{\sigma_{\widehat{Y}}(\mathbf{x})} \right), \quad i = 1, \dots, s, \quad (\text{A.4})$$

thanks to the normality of the Kriging predictor  $\widehat{Y}(\mathbf{x})$ , as well as the *approximate* instrumental PDFs that read:

$$h_{i\varepsilon}^*(\mathbf{x}) = \frac{\pi_{i-1}(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})}{\int_{\mathbb{X}} \pi_{i-1}(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x}}, \quad i = 2, \dots, s. \quad (\text{A.5})$$

Then, practical use of this splitting is as follows.

- (i) The optimal decreasing sequence of quantiles  $q_1 > \dots > q_s = 0$  is determined by applying the classical subset simulation approach to the mean  $\mu_{\widehat{Y}}$  of the Kriging predictor with intermediate probabilities set to  $p_0$  (say  $p_0 = 10\%$ ).
- (ii) The first term of the product in Eqn. (A.1) is estimated using the previously determined quantile  $q_1$ :

$$\widehat{p}_{1\varepsilon} = \frac{1}{N} \sum_{k=1}^N \pi_1(\mathbf{X}^{(k)}), \quad (\text{A.6})$$

$$\delta_{1\varepsilon} = \sqrt{\frac{1 - p_{1\varepsilon}}{N p_{1\varepsilon}}}, \quad (\text{A.7})$$

where  $\mathcal{X}^{[1]} = \{\mathbf{X}^{(k)}, k = 1, \dots, N\}$  is a sample of  $N$  independent copies of the random vector  $\mathbf{X}$ .

- (iii) The other  $s - 1$  terms are eventually estimated by sampling  $N$  copies of the random vectors  $\mathbf{X}_\varepsilon^{[i]}$  from their respective approximate instrumental PDFs  $\{h_{i\varepsilon}^*, i = 2, \dots, s\}$ . This sampling resorts to the modified Metropolis-Hastings algorithm of Au and Beck [9]. Note that the Markov chains can be seeded with the points in the  $i$ -th sample generated in step (i). Indeed, these samples are such that  $h_{i\varepsilon}^*$  is non-zero because of the following statement:

$$\mu_{\widehat{Y}}(\mathbf{x}) \leq q_{i-1} \quad \Rightarrow \quad \pi_{i-1}(\mathbf{x}) = \Phi\left(\frac{q_{i-1} - \mu_{\widehat{Y}}(\mathbf{x})}{\sigma_{\widehat{Y}}(\mathbf{x})}\right) \geq 0.5, \quad i = 2, \dots, s. \quad (\text{A.8})$$

The estimation of the intermediate augmented probabilities uses the quantiles  $\{q_i, i = 2, \dots, s\}$  determined in the preliminary step (i). These probabilities (resp. their

coefficient of variation) are computed as follows:

$$\widehat{p}_{i|i-1\varepsilon} = \frac{1}{N} \sum_{k=1}^N \frac{\pi_i(\mathbf{X}_\varepsilon^{[i](k)})}{\pi_{i-1}(\mathbf{X}_\varepsilon^{[i](k)})}, \quad (\text{A.9})$$

$$\delta_{i\varepsilon} \approx \frac{1}{p_{i|i-1\varepsilon}} \sqrt{\frac{1}{N} \left[ \frac{1}{N} \sum_{k=1}^N \frac{\pi_i(\mathbf{X}_\varepsilon^{[i](k)})^2}{\pi_{i-1}(\mathbf{X}_\varepsilon^{[i](k)})^2} - p_{i|i-1\varepsilon}^2 \right]} (1 + \gamma_{i\varepsilon}), \quad (\text{A.10})$$

where  $\{\mathbf{X}_\varepsilon^{[i](k)}, k = 1, \dots, N\}$  is a sample of  $N$  copies of the  $i$ -th instrumental random vector  $\mathbf{X}^{[i]}$ , and  $\gamma_{i\varepsilon}$  is the coefficient that accounts for the correlation in the MCMC sample [see 9, 29, Appendix B].

Eventually, the subset sampling estimator of the augmented failure probability is evaluated as follows:

$$\widehat{p}_{f\varepsilon\text{SS}} = \widehat{p}_{1\varepsilon} \prod_{i=2}^s \widehat{p}_{i|i-1\varepsilon}, \quad (\text{A.11})$$

and its coefficient of variation satisfies (as proved by Au and Beck [9] for subset simulation):

$$\sqrt{\sum_{i=1}^s \delta_{i\varepsilon}^2} \leq \delta_{\varepsilon\text{SS}} \leq \sqrt{\sum_{i=1}^s \sum_{j=1}^s \delta_{i\varepsilon} \delta_{j\varepsilon}}. \quad (\text{A.12})$$