

Structural reliability assessment through surrogate based importance sampling with dimension reduction

Miguel Munoz Zuniga, A. Murangira, Timothée Perdrizet

▶ **To cite this version:**

Miguel Munoz Zuniga, A. Murangira, Timothée Perdrizet. Structural reliability assessment through surrogate based importance sampling with dimension reduction. Reliability Engineering and System Safety, 2021, 207, pp.107289. 10.1016/j.ress.2020.107289 . hal-03159217

HAL Id: hal-03159217
https://ifp.hal.science/hal-03159217

Submitted on 4 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Reliability Engineering & System Safety 00 (2020) 1–30

Structural reliability assessment through surrogate based importance sampling with dimension reduction

M. Munoz Zuniga[a], A. Murangira[b], T. Perdrizet[a]

[a]*IFP Energies nouvelles 1 & 4, avenue de Bois-Préau 92852 Rueil-Malmaison Cedex - France*
[b]*La Javaness 19 rue Martel, 75010 Paris*

Abstract

We present a method for reliability assessment in extreme conditions from a numerical simulator through surrogate based importance sampling. As proposed in recent works in the literature, a Kriging surrogate is used to build an approximation of the limit state function and the optimal importance density. Our contribution is then the use of a sufficient dimension reduction method which enables the construction of the limit state function metamodel in lower dimension. The so called augmented failure probability and correction factor are recast in this dimension reduction framework. Simple strategies for metamodel refinement in the dimension reduction subspace are described and, in the case of Gaussian inputs, a computationally efficient MCMC scheme aimed at sampling the quasi-optimal importance density is presented. The case of non-Gaussian inputs is also laid out and it is argued and demonstrated through simulations that this approach can reduce the number of calls to the computer model, which is a crucial factor in reliability analysis. Advantages of this method are also supported by numerical simulations carried on an industrial case study concerned with the extreme response prediction of a wind turbine under wind loading.

Keywords: Reliability analysis, Surrogate model, Importance sampling, Sufficient dimension reduction, Kernel dimension reduction, time dependent response, rare event

1. Introduction

Structural reliability analysis often aims at assessing the probability of occurrence of an extreme event related to a given structure. In the usual setting, we are given a limit state function g which describes the performance of a system or structure with respect to some failure criterion or set of criteria for a given input vector $\boldsymbol{x} \in \mathbb{R}^d$. The failure domain, F, corresponds to the set of inputs for which the performance function g is negative, i.e. $F = \{\boldsymbol{x} \in \mathbb{R}^d \mid g(\boldsymbol{x}) \leq 0\}$. In our framework these inputs are uncertain and modeled as squared integrable random variables with know joint distribution. Thus, the failure probability boils down to:

$$P_f = \mathbb{P}(g(\boldsymbol{X}) \leq 0) = \mathbb{E}(\mathbb{1}_{g(\boldsymbol{X})\leq 0}) \tag{1}$$

with $\mathbb{1}_A$ the indicator function taking value one if A is fullfiled and 0 otherwise. This probability may readily be estimated through standard Monte Carlo simulation, however since P_f is often less than 10^{-4} it takes roughly $10^6 - 10^7$ evaluations of g to obtain an estimate with a coefficient of variation (c.o.v.) less than 5 %. This is clearly problematic for many engineering applications since most of the time g depends on the output of time-expensive to evaluate computer model. Analytical approximations based on the most probable failure point, also called design point (DP), have long been the practical alternative to Monte Carlo sampling. This has led to the popular FORM/SORM estimations (Ditlevsen and Madsen (13)) later adapted to deal with multiple design points (Kiureghian and Dakessian (28)). However these approximations are often poor whenever the non-linearity of g are non-negligible, the input dimension is high (Katafygiotis and Zuev (27)) and do not provide any confidence bounds on the failure probability estimate. Accelerated Monte Carlo techniques have been proposed in the literature to overcome these drawbacks (Morio and Balesdent (40), Caron et al. (8)). In particular, Importance Sampling (IS) (Hammersley and Handscomb (24)) is a standard variance reduction Monte Carlo method that can be used to estimate P_f by sampling from an importance density whose main contribution takes place near the limit state surface. A preliminary design point or multiple failure mode computation can be used to design an importance density with techniques such as a mixture of standard distributions centred on the failure modes (Au et al. (3), Melchers (39)). In the same vein, Yun et al. (54) proposed an IS estimator introducing a safety sphere and an IS density based criterion to reduce the number of costly model evaluations. To further reduce the estimation cost, Yao et al. (52) improved the latter by identifying the "critical region" which contains input samples with both high occurrence probability and high misjudgment risk. Still, these estimators rely strongly on the DP point(s) estimation and the number of evaluations of the limit state function required to reach a given accuracy may still be important specially in an high dimensional input setting (Katafygiotis and Zuev (27)), limiting the applicability of such a scheme. Subset simulation (Au and Beck (2)) eliminates the need to design an importance density by estimating the failure probability as a product of intermediate probabilities which are evaluated by Monte Carlo Markov Chain simulations. However, the incurred computing cost is still usually prohibitive in many industrial reliability cases. As a result, methods based on surrogate modelling have been devised to limit the number of computer model evaluations. Nevertheless, using a too crude approximation of the limit state function can lead to non-consistent failure probability estimators (Dubourg et al. (16)). Moreover, replacing the true model by an accurate surrogate can still lead to an inaccurate probability estimation (Li and Xiu (34)). In fact, a good practice consists in using the surrogate only/mostly as a guide for the selection of input values to be evaluated with the real model in order to get a consistent estimator and a precise failure probability estimation. In this latter framework, Echard et al. (20), Echard et al. (21) and Li et al. (36) proposed methods based on the active kriging principle, in which the performance function g is replaced by a Gaussian process regression metamodel. This latter model is iteratively

refined so as to provide accurate predictions in the vicinity of the limit state surface and suggests inputs values to be evaluated by the real model. In the subset simulation framework, a similar approach has been proposed by Bourinet et al. (7), where a SVM classifier is built to emulate the intermediate limit state surfaces. Alternative surrogate based approaches emcompass the work of Sun et al. (49) involving a surrogate based proposal function incorporating the inputs probability density. Li and Xiu (34) used a generalized polynomial chaos surrogate as the guideline for an MC estimator and finally Au (4) proposed a stratified, with respect to a given surrogate output, estimator of the seeked failure probability.

On the other hand, to ensure consistent importance sampling estimators, the Meta-IS (Dubourg et al. (15)) method was introduced, where the sought probability is evaluated via an approximation of the Optimal Importance Density (OID) based on a kriging emulator. The corresponding IS estimator of P_f is cast as a product of the failure probability relative to the metamodel - the augmented failure probability - and a correction factor based on the limit state function g. Simulations on mechanical reliability problems have demonstrated the efficiency of this methodology both in terms of accuracy and simulation budget.

However, it is known that fitting a kriging metamodel can break down in high dimensional settings ($d \geq 20-30$) if no stronger assumptions are imposed on the function to approximate (our working hypothesis), or demand an unreasonable amount of code evaluations to obtain sufficient accuracy in the failure region. Furthermore, in the methodology of Meta-IS, sampling from the approximate OID is accomplished by resorting to an MCMC algorithm, most of which are known to suffer from the curse of dimensionality. On the other hand, in the last couple of decades, dimension reduction (DR) techniques have been an intense subject of research in computational statistics. For instance, to tackle the ”curse of dimensionnality” within a kriging model, Lelièvre et al. (32), Lelievre (31) proposed a reduction dimension strategy to enable the construction of a kriging metamodel in high dimension by applying a PCA to the set of hyper-parameters before their optimization but based its failure probability estimator solely on the surrogate prediction. Another strategy was proposed in Jiang and Li (26) where an active subspace approach (Constantine (10)) is proposed to reduce the input dimension. This latter DR technique is appealing but requires the knowledge or the estimation of the gradient of the limit state function. Nevertheless, when dealing with an expensive numerical code with no gradient information and an high dimensional input vector, the gradient estimation cost can become prohibitive. The sliced inverse regression (SIR) method as a DR technique have also been coupled with a polynomial chaos expansion to overcome the dimensionality issue (Xu and Wang (51)). The SIR method is an efficient DR approach but its application is circumscribes to Gaussian (or more generally to Elliptical) inputs distributions. To overcome these latter limitations, methods such as kernel dimension reduction (KDR) (Fukumizu et al. (23)) and gradient-based KDR (gKDR) (Fukumizu and Leng (22)) have been developped. The method based on gKDR presents practical advantages

that will be discussed latter on the paper and, despite its name, does not require directly the limit state function gradient estimation because of its kernel based approach. We can ultimatly mention the work of Lataniotis et al. (30) which proposes a general framework to couple a surrogate estimation with a DR technique by minimizing alternatively a given loss function with respect to, on the one hand, the surrogate (hyper-)parameters and, on the other hand, meta-parameters of the selected DR algorithm. We can notice that our approach almost lie within this latter framework.

In this paper, we lay out a methodology for importance sampling based on the sufficient dimension reduction (SDR) framework (Chiaromonte and Cook (9)) for regression. SDR is based on the assumption that the output statistical dependency on the input $\boldsymbol{X}$ can be described entirely by projecting $\boldsymbol{X}$ on a lower dimensional subspace. Based on this assumption, we propose to build a kriging metamodel in the reduced dimension subspace, making the Meta-IS algorithm tractable. Assuming that the input vector is Gaussian, we show that the augmented failure probability can be estimated through Monte Carlo simulation in the reduced dimension subspace. As for the correction factor, which is based on a sample from an approximation of the optimal importance density, we demonstrate how the sample generation can be achieved efficiently since MCMC sampling is only performed in the reduced dimension subspace. We also discuss the case of application of SDR for Gaussian process emulation of complex models with non-Gaussian inputs. The paper is organized as follows. Section 2 recalls the basics of metamodel-based importance sampling. Section 3 gives an overview of two dimension reduction tools: KDR and gKDR. Section 4 is devoted to metamodel based importance sampling with sufficient dimension reduction. Finally, in section 5 we numerically illustrate these ideas on an academic example and on a relevant industrial case study which consists in the assessment of the structural reliability of a wind turbine, where the wind is modelled as a Gaussian process leading to a high dimensional ($d \geq 100$) case study.

2. Importance sampling with a kriging metamodel

Throughout this article, the input is a real d dimensional random vector $\boldsymbol{X} = (X_1, \ldots, X_d)^T$ with density q. As previously stated g denotes the limit state function.

Kriging models are flexible and efficient surrogates to complex computer codes. As most metamodels, they rely on an initial design of experiments (DoE) $\mathcal{D} = \left\{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^{N_D}\right\}$. The popularity of kriging stems among other things from the availability of prediction uncertainty estimates, given by the kriging variance. This makes it possible to devise refinement strategies aimed at increasing some measure of accuracy in the regions of interest. For reliability purposes this region is in the vicinity of the limit state surface F.

2.1. Gaussian process based model

Kriging is based on the assumption that the performance function g is a sample from a Gaussian process (GP) G so that

$$G(\boldsymbol{x}) = \mathbf{f}^T(\boldsymbol{x})\boldsymbol{\beta} + Z(\boldsymbol{x}) \tag{2}$$

where $\mathbf{f}^T(\boldsymbol{x})\boldsymbol{\beta}$ is the GP mean and Z a zero-mean stationary Gaussian process. $\mathbf{f} = \left(f_1, \ldots, f_p\right)^T$ is a vector of basis functions $\in \mathcal{L}^2(\mathbb{R}^d, \mathbb{R})$ and β a vector in $\mathbb{R}^p$. Z is parametrized by its stationnary autocovariance function

$$C(\boldsymbol{x}, \boldsymbol{x}') = \sigma_G^2 R_\theta(\boldsymbol{x} - \boldsymbol{x}') \tag{3}$$

where σ_G^2 is the GP variance and θ is a vector of hyper-parameters of the autocorrelation function R. The autocorrelation function is given a priori and is a key modeling ingredient for the surrogate. As can be seen in (4) and (5), the prediction at any point $\boldsymbol{x}$ is written as the sum of a trend term and a linear combination of $\boldsymbol{r}(\boldsymbol{x}) = R_\theta(\boldsymbol{x} - \boldsymbol{x}_i)$, moreover, the variance predictor is deeply impacted by the kernel choice. Without prior information, we selected the anisotropic stationary Matern-5/2 autocorrelation function which offers enough flexibility to adequately capture the variability of numerous objective function depending on the choice of the hyper-parameters.

Prediction at a previously unobserved input $\boldsymbol{x}$ is based on the best linear unbiased prediction (BLUP) of $G(\boldsymbol{x})$ given the observations $\mathbf{y} = (g(\boldsymbol{x}^1), \ldots, g(\boldsymbol{x}^{N_D}))$ at the DoE $\mathcal{D}$. The BLUP at $\boldsymbol{x}$, denoted $\hat{G}(\boldsymbol{x})$, is a normal random variable $\mathcal{N}(m_{\hat{G}}, \sigma_{\hat{G}}^2)$ where

$$m_{\hat{G}}(x) = \mathbf{f}^T(\boldsymbol{x})\hat{\boldsymbol{\beta}} + \mathbf{r}^T(\boldsymbol{x})\mathbf{R}^{-1}(\mathbf{y} - \mathbf{F}\hat{\boldsymbol{\beta}}) \tag{4}$$

$$\sigma_{\hat{G}}^2(\boldsymbol{x}) = \sigma_G^2(1 - \mathbf{r}^T(\boldsymbol{x})\mathbf{R}^{-1}\mathbf{r}(\boldsymbol{x}) + \mathbf{v}^T(\boldsymbol{x})(\mathbf{F}^T\mathbf{R}^{-1}\mathbf{F})^{-1}\mathbf{v}(\boldsymbol{x})) \tag{5}$$

are respectively the prediction mean and variance, $\mathbf{R}$ is the correlation matrix of the DoE defined by $\mathbf{R}_{ij} = R_\theta(\boldsymbol{x}^i - \boldsymbol{x}^j)$, $i, j = 1, \ldots, N_D$. The term $\mathbf{r} = R_\theta(\boldsymbol{x} - \boldsymbol{x}^i)$, $i = 1, ..., N^D$, is the cross-correlation vector between the prediction and the observations while $\mathbf{F}$ is the matrix defined by $\mathbf{F}_{ij} = f_j(\boldsymbol{x}^i)$, $1 \leq i \leq N_D$, $1 \leq j \leq p$. The vector $\hat{\boldsymbol{\beta}}$ is the solution to a generalized least-squares problem $\hat{\boldsymbol{\beta}} = (\mathbf{F}^T\mathbf{R}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\mathbf{R}^{-1}\mathbf{y}$ and $\mathbf{v}(x) = \mathbf{F}^T\mathbf{R}^{-1}\mathbf{r}(\boldsymbol{x}) - \mathbf{f}(\boldsymbol{x})$. More details on Gaussian process modeling can be found in (45, 43).

2.2. Basics of metamodel based importance sampling (Meta-IS)

In this section, we give an overview of the Meta-IS algorithm originally due to Dubourg et al. (15). The method relies on the use of a quasi-optimal importance density which mimics the intractable optimal importance density.

Importance sampling is a well known variance reduction method where the quantity of interest is an expectation of a squared integrable function. In the context of reliability methods, the expectation of interest is P_f. Let $\tilde{q}$ be a probability density function with support containing the one of $\mathbb{1}_{g(\boldsymbol{x})\leq 0}q(\boldsymbol{x})$ and such that $\mathbb{E}_{\tilde{q}}(\mathbb{1}_{g(\boldsymbol{X})\leq 0}q(\boldsymbol{X})/\tilde{q}(\boldsymbol{X})) < \infty$. IS stems from the equality $P_f = \mathbb{E}_{\tilde{q}}(\mathbb{1}_{g(\boldsymbol{X})\leq 0}q(\boldsymbol{X})/\tilde{q}(\boldsymbol{X}))$. Given an i.i.d. sample from $\tilde{q}$, $\boldsymbol{X}^{(i)}$ $i = 1, \ldots, n$, the importance sampling estimator of P_f reads

$$\hat{P}_f^{IS} = n^{-1}\sum_{i=1}^{n} \mathbb{1}_{g(\boldsymbol{X}^{(i)})\leq 0}q(\boldsymbol{X}^{(i)})/\tilde{q}(\boldsymbol{X}^{(i)}) \tag{6}$$

The above estimator is unbiased and has a known variance which can be minimized with respect to $\tilde{q}$. An important result from Rubinstein and Kroese (44) is that $\mathrm{Var}_{\tilde{q}}(\hat{P}_f^{IS}) = 0$ for the following optimal proposal density

$$\tilde{q}_{opt}(\boldsymbol{x}) = \mathbb{1}_{g(\boldsymbol{x})\leq 0}q(\boldsymbol{x})/P_f \tag{7}$$

which is generally difficult to sample from since it involves the failure probability we are trying to estimate. In Meta-IS, samples are drawn from a quasi-optimal density obtained by replacing the indicator function with a continuous probabilistic classification function linked to the kriging predictor:

$$\pi(\boldsymbol{x}) = \Phi(-m_{\hat{G}}(\boldsymbol{x})/\sigma_{\hat{G}}(\boldsymbol{x})) \text{ if } \boldsymbol{x} \notin \mathcal{D} \text{ and } \mathbb{1}_{g(\boldsymbol{x})\leq 0} \text{ if } \boldsymbol{x} \in \mathcal{D} \tag{8}$$

where Φ is the standard normal cumulative distribution function (cdf). Now by swapping $\pi(\boldsymbol{x})$ and $\mathbb{1}_{g(\boldsymbol{x})\leq 0}$ in (7), one obtains the following quasi-optimal importance density (Dubourg et al. (15)):

$$\tilde{q}_*(\boldsymbol{x}) = \pi(\boldsymbol{x})q(\boldsymbol{x})/P_{f,\epsilon} \tag{9}$$

where $P_{f,\epsilon} = \int \pi(\boldsymbol{x})q(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$ is the augmented failure probability.

Considering the defined quasi-optimal instrumental density, the failure probability estimate may now be broken down as follows (Dubourg et al. (15))

$$P_f^{IS} = P_{f,\epsilon}\alpha_{\mathrm{corr}} \tag{10}$$

where $P_{f,\epsilon}$ is the augmented failure probability, which relies solely on the metamodel, and $\alpha_{\mathrm{corr}} = \mathbb{E}_{\tilde{q}_*}(\mathbb{1}_{g(\boldsymbol{X})\leq 0}/\pi(\boldsymbol{X}))$ is a factor that corrects for the bias. Let $\boldsymbol{X}^{(i)}$, $i = 1, \ldots, N_{\mathrm{meta}}$ be an i.i.d. sample from the prior density q, then

$$\hat{P}_{f,\epsilon} = N_{\mathrm{meta}}^{-1}\sum_{i=1}^{N_{\mathrm{meta}}} \pi(\boldsymbol{X}^{(i)}) \tag{11}$$

is an unbiased and consistent estimator of $P_{f,\epsilon}$. The estimation error is quantified by the Monte Carlo variance

$$\hat{\sigma}^2_{\epsilon} = [N_{\text{meta}}(N_{\text{meta}} - 1)]^{-1} \sum_{i=1}^{N_{\text{meta}}} \left(\pi(\boldsymbol{x}^{(i)}) - \hat{P}_{f,\epsilon}\right)^2 \tag{12}$$

From a practical standpoint, sampling from q is usually straightforward whereas obtaining draws from $\tilde{q}_*$ given in (9) requires adequate algorithms such as Monte Carlo Markov Chain simulation. Let $\boldsymbol{t}^{(j)}$, $j = 1, ..., N_{corr} + b$ be a chain with stationary distribution $\tilde{q}_*$, then from ergodicity

$$\hat{\alpha}_{\text{corr}} = N^{-1}_{\text{corr}} \sum_{j=b+1}^{N_{\text{corr}}+b} \mathbb{1}_{g(\boldsymbol{t}^{(j)})\leq 0}/\pi(\boldsymbol{t}^{(j)}) \tag{13}$$

is a biased finite sample estimate of α_{corr} where b is a burnin parameter. The standard squared error can be assessed as

$$\hat{\sigma}^2_{\text{corr}} = N^{-1}_{\text{corr}} \left(N^{-1}_{\text{corr}} \sum_{j=b+1}^{N_{\text{corr}}+b} h(\boldsymbol{t}^j) - \hat{\alpha}^2_{\text{corr}} \right) (1 + \hat{\gamma}_{\text{corr}}) \tag{14}$$

where $h(\boldsymbol{t}^j) = \mathbb{1}_{g(\boldsymbol{t}^{(j)})\leq 0}/\pi^2(\boldsymbol{t}^{(j)})$ and $\hat{\gamma}_{\text{corr}}$ is an estimator of $2\sum_{k=0}^{\infty} \text{Corr}(h(\boldsymbol{t}^{(0)}), h(\boldsymbol{t}^{(k)}))$, which can be estimated as detailed in Dubourg (14). In the original paper of Dubourg et al. (15), samples from the quasi-optimal density are obtained through a modified Metropolis-Hastings scheme (Au and Beck (2)). To mitigate the inflation of the estimation variance due to the dependence in the chain, thinning can be performed prior to the correction factor estimation, that is taking one in k draws in the simulated chain. This of course increases the length of the simulated chain which must be kN_{corr} long in order to retain N_{corr} samples. From a computational perspective, the MCMC sampler only evaluates the probabilistic classification function, and is therefore expected to be relatively efficient.

While the estimation of the augmented failure probability $P_{f,\epsilon}$ only resorts to the Kriging predictor, its computation in a high dimensional space through standard Monte Carlo simulation can still be improved specially when dealing with small probabilities. Therefore, we adopt a splitting estimator for $P_{f,\epsilon}$, much like in subset simulation, as introduced in Sudret et al. (48)

2.3. Kriging metamodel refinement

Before computing both quantities $\hat{P}_{f,\epsilon}$ and $\hat{\alpha}_{\text{corr}}$, it is paramount to have a sufficiently accurate Kriging metamodel. Otherwise, the approximation $\tilde{q}_*$ of the optimal importance density would be poor and yield high estimation variance. To obtain an accurate emulator, one can iteratively enrich the initial design DoE with new limit state function evaluations. This initial DoE can be any relevant space filling design. In reliability analysis, iterative refinement methods are usually geared towards accurate approximation of a target region as evidenced by Vasquez and Bect (50), Picheny

et al. (42) or Bect et al. (5). In our context the target region is the area where the sign of the model $\hat{G}$ is the most uncertain. The target region is formalized as an in-fill criterion function to be optimized in order to get the next point to be added to the doE (Bect et al. (5)). However, the in-fill criterion might feature multiple optima or cause the optimizer to be stuck on a local plateau. Moreover, unless one uses a multi-start scheme (which would not guarantee finding all local optima anyway), most iterative in-fill criteria do not allow for multiple points to be added to the DoE. This has motivated sampling based design enhancements as in Bourinet et al. (7) which rely on a so-called margin shrinking concept: initially these were based on adding multiple points in the margin of a support vector margin (SVM) classifier and can be adapted to the Kriging metamodel enhancement. The idea is to draw a large number of samples, from the weighted margin probability density defined as $h(\boldsymbol{x}) \propto w(\boldsymbol{x})q(\boldsymbol{x})$ where w is the margin probability (Dubourg (14)): the probability for the GP model to be within an interval centered on 0. Drawing samples from h yields points with a high uncertainty on the sign of their Kriging predictor: these are good candidates to be added to the DoE. In practice, an MCMC algorithm (e.g. slice sampling from Neal (41)) can provide a sufficiently large sample with stationary distribution h. The k-means algorithm, with the euclidean distance, then provides N_{add} clusters from these samples, which centers are evaluated on the performance function and added to the DoE. The Kriging model is then updated on the basis of this enhanced design.

The model refinement is iterated as long as a stopping criterion is not met. In the Meta-IS algorithm, the usual criterion is a leave-one-out estimate of the correction factor which is defined as

$$\alpha_{\text{LOO}} = N_D^{-1} \sum_{i=1}^{N_D} \mathbb{1}_{g(\boldsymbol{x}_i)\leq 0} / \pi_{-i}(\boldsymbol{x}_i) \tag{15}$$

where $\pi_{-i}(\boldsymbol{x}) = \Phi\left(-\hat{m}_{G_{\mathcal{D}\setminus \boldsymbol{x}_i}}(\boldsymbol{x}_i)/\hat{\sigma}_{G_{\mathcal{D}\setminus \boldsymbol{x}_i}}(\boldsymbol{x}_i)\right)$ is the classification function obtained by removing observation $\boldsymbol{x}_i$ from the DoE $\mathcal{D}$. A factor π close to 1 signifies accuracy of the classifier based on the Kriging classification function π. The DoE enhancement can therefore be terminated whenever $0.1 \leq \alpha_{\text{LOO}} \leq 10$ and the number of experiments N_D is higher than a pre-specified threshold (typically a few tenths/hundreds). It is also necessary to impose a maximum of computer model evaluations during the initial space-filling and refinement phase, as there is no guarantee that α_{LOO} will get closer to 1 with a reasonable amount of limit state function evaluations.

3. Sufficient dimension reduction for regression and classification

3.1. Motivations and principle of sufficient dimension reduction

While flexible enough to cover a wide range of reliability analysis problems, the metamodel based importance sampling procedure can sometimes be intractable when dealing with high dimensional inputs. Indeed, the Kriging meta-

model usually needs a DoE with sufficient size to cover the input space. Furthermore, the Gaussian process training which requires learning the kernel hyperparameters might be faced with matrix singularity during marginal likelihood optimization, especially when the kernel is parametrized by a scalar parameter for each input variable.

Attempts have been made recently to propose adequate representations for high dimensional Kriging. For instance, additive Kriging (Durrande et al. (17)) can be used by assuming an additive model for the emulator, i.e. that model is a sum of univariate metamodels. This was shown to be more effective than using standard separable kernels on a few test cases involving between 10 and 50 variables. Another relevant approach is to incorporate variable selection into the Gaussian process regression framework as suggested by Yi et al. (53).

In this work, we take another approach which is motivated by the literature in supervised dimension reduction, specifically the sufficient dimension reduction (SDR) framework (Chiaromonte and Cook (9)). SDR focuses on the discovery of a subspace S, of the initial input space, such that the projection of the inputs in S explains the behaviour of the output. This latter is equivalent to say that the distribution of $Y|\boldsymbol{X}$ is the same as the one of $Y|\text{Projected}_S(\boldsymbol{X})$. To do so a low rank matrix B spanning S is estimated which enables, in our context, to perform the regression of the output Y on $\boldsymbol{X}$ only with respect to the projected inputs: $B^T\boldsymbol{X}$. Indeed, for many regression models, it is frequent for the output to be the combination of non-linear functions depending on linear combinations of the inputs (or a subset) as well as an optional noise independent of the inputs.

More formally, $S_B = \text{span}(B)$ is a dimension reduction subspace if

$$Y \perp \boldsymbol{X} \mid B^T\boldsymbol{X} \tag{16}$$

where $B \in \mathbb{R}^{d\times r}$ and $\perp$ stands for independence. In our case, (16) could be rewritten as $Y = g_r(B^T\boldsymbol{X}, \varepsilon)$ where ε is a noise term independent of $\boldsymbol{X}$ and g_r an unknown function. The interest of SDR for metamodeling is obvious: if we could find a matrix B with sufficiently low rank r, a surrogate model of the link function g_r can be built. Typically, in structural reliability of offshore structures we might be dealing with over 100 input variables. If $r \leq 10$ a usual GP metamodel can then be applied to model the response $Y = g(\boldsymbol{X})$ but in the subspace S_B. Because, we have no prior knowledge on the noise term, we may obtain a coarse metamodel by using the simplified model

$$Y = g_r(B^T\boldsymbol{X}) \tag{17}$$

In this case, the surrogate model is fitted to the link function g_r. To deal with the fact that model (17) does not take into account the noise in the original dimension reduction model, we suggest to use a nugget term in the corresponding metamodel which value will be estimated jointly with the other hyper-parameters by maximizing the GP model

likelihood. The fact that we are fitting a metamodel not to the true limit state function but to the link function g_r in the simplified regression model (17) should not necessarily hamper the metamodel's performance if the regression noise ε is not predominant. We call the importance sampling procedure with SDR subspace metamodelling Meta-ISDR, where DR stands for dimension reduction. It will be outlined in 4.

A similar approach has been laid out in Xu and Wang (51) where the SIR dimension reduction approach is used before constructing a polynomial chaos expansion of the limit state function with respect to the reduced variables. The SDR approach we couple with a GP model is more general in the sense that, unlike in SIR, it does not impose any linear assumption on the mean of the input variables conditionned on the output one neither any assumption on the inputs distributions. Nonetheless, in the same manner as in Xu and Wang (51), the SDR method is used as a screening approach. Moreover, we are interested in reducing the dimension for the regression of the output Y in a region that is large enough to contain part of the failure domain. Therefore, it is necessary to obtain a training set consisting in a random samples that fill the input space but also populate the failure domain. For this reason, the initial design of experiment performed for the SDR purpose is a mixture of a global and a local design: a global uniform distribution within a given maximal global hyper-sphere and local uniform distributions within spheres (of given radius) centered at the discovered design points (included in the global sphere).

For the selection of variables important for the reliability analysis, the same approach could be directly applied to the indicator output $\mathbb{1}_{g(\boldsymbol{X})\leq 0}$ since the kernel dimension reduction method we use has the particularity of handling binary output. The difficulty is then to start with a sample having enough 0-1 information for the SDR analysis since the event considered is rare. Instead of performing the DR study before the reliability analysis, it might be possible to incorporate the SDR approach within this meta-IS since the method generates sample points in the vicinity of the limite state. However, the latter strategy will impose to work, at least at the beginning, in the high dimensional space. This idea will be further investigated in another paper.

In the sufficient dimension reduction literature, the kernel dimension reduction method is relatively complex and require a non-negligible time of process. We clearly can not enter into all details of the theoretical background necessary but hope to give the reader the main motivations and key ideas supporting its use. Refer to the literature cited for details on the methodology.

Recently, kernel methods have lead to the kernel dimension reduction (KDR) algorithm (Fukumizu et al. (23)) and its gradient based counterpart (gKDR, Fukumizu and Leng (22)): these algorithms can cope with a scalar, vectorial or binary output and impose no strict conditions on the distribution of the input vector X. As such, they lends themselves well to the task at hand since the user has much more latitude in the choice of the sampling distribution. For this

same reason we dismissed methods such as the one proposed in Li (35), Cook (11) and Li et al. (33) which impose restrictions on the input distribution.

To our knowledge, all supervised dimension reduction techniques offering a mapping between the initial space and the projected one, rely on projection on linear sub-spaces. We have seen few, yet un-published, attempts to extend the approach to non-linear spaces but, as far as we know, they will rely on gradient informations which is to costly to get in our computer experiment context. Therefore, in the paper we discard the possibility of discovering non-linear functions $g_1, ..., g_M$ from the initial space into $\mathbb{R}$ such that $g(\boldsymbol{x}) = g_r(g_1(\boldsymbol{x}), ..., g_M(\boldsymbol{x}), \epsilon)$.

Nevertheless, even if the KDR approach can only discover linear functions g_1,..., g_M, represented by the projection matrix B, it is not restricted to a linear function g_r of the projected inputs $B^t\boldsymbol{x}$. Indeed, KDR will seek for the best linear combinations of the inputs with respect to an output that can be complex combination of non-linear functions of these projected inputs as well as an optional noise independent of the inputs. In very loose terms, a space of functions (characterized by a given kernel) is considered as the set of non-linear functions, defined in the projected space, that can approximate functions (from another kernel generated space) of the output. The optimal projection matrix B is the one that minimizes the approximation error of the function of the output by an optimal function of $B^t\boldsymbol{x}$ selected from the kernel-generated set. This latter explanation can be summarized by the following optimization problem:

$$\min_{B, B^T B = I_r} \sum_{a=1}^{\infty} \min_{f \in \mathcal{H}_X^B} \mathbb{E}\Big[(\xi_a(Y) - \mathbb{E}(\xi_a(Y)) - (f(\boldsymbol{X}) - \mathbb{E}(f\boldsymbol{X}))\Big]^2,$$

where the ξ_as form a complete orthonormal system of the reproducing kernel Hilbert space (RKHS) $\mathcal{H}_\mathcal{Y}$ associated to the kernel $k_\mathcal{Y}$ and $\mathcal{H}_X^B$ is the RKHS associated to the kernel k_X^B. The advantage of the KDR approach is to propose an equivalent but much less expensive formulation of this latter optimization task. The kernels and more details are presented in the next section.

For instance, assuming no noise, if $g(\boldsymbol{x}) = x_1x_4 + \sin(x_2 + x_3)$, $\boldsymbol{x} \in \mathbb{R}^4$, then the projection matrix would be the 4×3 matrix B with column vectors the normalized versions of: $(1, 0, 0, 1)$, $(0, 1, 1, 0)$ and $(-1, 0, 0, 1)$.

As stated previously, we notice that the method can not project on non-linear subspaces and interaction terms, such as x_1x_4 above, cannot be detected as one new variable unless you specificaly introduce this new variable in the $\boldsymbol{x}$ input vector, for instance such that $x_5 = x_1x_4$ and perform the dimension reduction analysis on the random vector $(X_1, ..., X_5)$. This latter fact introduce the possibility to a priori incorporate non-linear functions of the inputs in the random vector $\boldsymbol{X}$ used in the SDR step.

3.2. Kernel dimension reduction

The basis of KDR is to express the conditional independence criterion (16) as an optimization problem that reaches its minimum for the appropriate dimension reduction matrix. The optimization problem is defined through conditional covariance operators on reproducing kernel Hilbert spaces (RKHSs) that capture conditional independence. From a practical standpoint these operators are estimated from Gram matrices of the data and a specific functional of these Gram matrices is minimized to yield an estimate of a SDR matrix.

Let $(\boldsymbol{X}, Y)$ be a random vector on $\mathcal{X} \times \mathcal{Y}$ with probability distribution probability P_{XY} where $(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$ and $(\mathcal{Y}, \mathcal{B}_{\mathcal{Y}})$ are measured spaces. Let $k_{\mathcal{X}}$ and $k_{\mathcal{Y}}$ be measurable positive definite kernels on $\mathcal{X}$ and $\mathcal{Y}$ respectively, with associated reproducing kernel Hilbert spaces $\mathcal{H}_{\mathcal{X}}$ and $\mathcal{H}_{\mathcal{Y}}$. Furthermore, assume $\mathbb{E}(k_{\mathcal{X}}(\boldsymbol{X}, \boldsymbol{X})) < \infty$ and $\mathbb{E}(k_{\mathcal{Y}}(Y, Y)) < \infty$. The conditional independence criterion of (16) in KDR relies on the cross-covariance operator $\Sigma_{YX} : \mathcal{H}_{\mathcal{X}} \mapsto \mathcal{H}_{\mathcal{Y}}$ defined as in Fukumizu et al. (23) by

$$\langle h, \Sigma_{YX} f \rangle_{\mathcal{H}_{\mathcal{Y}}} = \mathbb{E}_{XY}\left((f(\boldsymbol{X}) - \mathbb{E}_X(f(\boldsymbol{X}))(h(Y) - \mathbb{E}_Y(h(Y)))\right) \tag{18}$$

for all $f \in \mathcal{H}_{\mathcal{X}}$, $h \in \mathcal{H}_{\mathcal{Y}}$. The covariance operator $\Sigma_{\boldsymbol{X}\boldsymbol{X}}$ is defined similarly. The conditional covariance operator $\Sigma_{YY|X}$ follows as combination of $\Sigma_{\boldsymbol{X}\boldsymbol{X}}$ and Σ_{YY} and some known bounded operators. Let $B \in \mathbb{R}^{d \times r}$ a matrix such that span(B) is a dimension reduction subspace and $B^T B = I_r$. Letting k_r be a positive definite kernel and $k_{\mathcal{X}}^B$ a positive definite kernel defined by $k_{\mathcal{X}}^B(\boldsymbol{x}, \boldsymbol{x}') = k_r(B^T \boldsymbol{x}, B^T \boldsymbol{x}')$, we define the (cross-)covariance operators w.r.t. kernel $k_{\mathcal{X}}^B$ as Σ_{YX}^B, $\Sigma_{\boldsymbol{X}Y}^B$, Σ_{YY}^B and $\Sigma_{\boldsymbol{X}\boldsymbol{X}}^B$. The conditional cross covariance operator $\Sigma_{YY|\boldsymbol{X}}^B$ associated to $k_{\mathcal{X}}^B$ is defined similarly as $\Sigma_{YY|\boldsymbol{X}}$. Under mild conditions that are satisfied in particular if all involved kernels are Gaussian RBFs, the following fundamental relationship links the dimension reduction subspace and the conditional cross-covariance operators (Fukumizu et al. (23)).

$$\Sigma_{YY|\boldsymbol{X}} = \Sigma_{YY|\boldsymbol{X}}^B \Leftrightarrow Y \perp \boldsymbol{X} \mid B^T \boldsymbol{X} \tag{19}$$

In addition to (19), we have for all B: $\Sigma_{YY|\boldsymbol{X}}^B \geq \Sigma_{YY|\boldsymbol{X}}$ (with the order of self-adjoint operators). Thus, in order to find the SDR matrix B, the KDR algorithm minimizes $\mathrm{Tr}(\Sigma_{YY|\boldsymbol{X}}^B)$ subject to $B^T B = I_r$. A reformulation of the latter optimization problem leads to the practical algorithm which takes as inputs n i.i.d. samples $(\boldsymbol{x}^i, y^i)$ from P_{XY} and solves

$$\min_{B^T B = I_r} \mathrm{Tr}\left[G_Y (G_{\boldsymbol{X}}^B + n\varepsilon_n I_n)^{-1}\right] \tag{20}$$

where G_Y and $G_{\boldsymbol{X}}^B$ are the centred Gram matrices defined by $G_{\boldsymbol{X}}^B = H K_{\boldsymbol{X}}^B H$, $G_Y = H K_Y H$, $(K_{\boldsymbol{X}}^B)_{ij} = k_{\mathcal{X}}^B(\boldsymbol{x}^i, \boldsymbol{x}^j)$, $(K_Y)_{ij} = k_{\mathcal{Y}}(y^i, y^j)$, $1 \leq i, j \leq n$ and $H = I_n - \frac{\mathbb{1}\mathbb{1}^T}{n}$. The parameter ε_n is a regularization term that facilitates matrix

inversion. In KDR this non-convex optimization problem is solved by a steepest-descent with algorithm line search (see Fukumizu et al. (23) for more details).

The considered kernels $k_{\mathcal{Y}}$ and k_r are both isotropic Gaussian kernel involving one parameter each. The median of pairwise distances of the $\boldsymbol{X}$-data is estimated and denoted σ_{median}. Then, during the steepest-descent optimization, for a fixed number of iteration l_{max}, the hyper-parameter of the kernel k_r is sequentially increased taking the values $2\sigma^2_{median}(1 + cl)$ for $l = 0, ..., l_{max}$ and c a fixed step size. The same is done with the Y-data: the hyper-parameter of the kernel $k_{\mathcal{Y}}$ is incremented with the same step size c. This procedure imply imposing more regularization as the optimization proceeds. Fukumizu et al. (23) proposed to do the opposite by decreasing the regularization along the optimization iterations. We believe that starting with less regularization would benefit by enabling to select a better direction in the optimization at the beginning and then when more advanced in the optimization we can be less demanding and increase the regularization without loosing much in optimality. Numerical results are reasonnably good with the proposed approach. Nevertheless, we believe that counter examples could be found to put into jeopardy the strategy proposed in (23) and ours. As mentioned in (23): "we are not aware of theoretically justified methods of choosing these parameters; this is an important open problem". Of course the best approach would be to optimize with respect to B and the regularization parameters at the same time but then the cost could become prohibitive. An improvement could be to replace the line search step in the gradient descent (w.r.t B) by some BFGS or L-BFGS strategy that will require less iteration and further add a few gradient descent steps with respect to the kernels parameters. The latter improvement has not been implemented for the moment and will be the subject of further works.

On the other hand, since the regularization term ϵ_n as the same purpose of the two previous parameters it is fixed during the procedure to a default value of 10^{-4}.

Finally we can notice that the following introduction of the g-KDR approach is in particular motivated by the resolution of the same problem with a less expensive-to-evaluate cost function; enabling to perform cross-validation to tune the kernels parameters at a more reasonnable cost.

3.3. Gradient-based kernel dimension reduction (gKDR)

The previous kernel dimension reduction procedure is efficient and consistent under non-restrictive assumptions on the joint and marginal distributions of $\boldsymbol{X}$ and Y. However, the required non-convex optimization step makes it somewhat CPU intensive depending on the size of the training samples $(\boldsymbol{x}^i, y^i)$. Indeed, the KDR cost function involves the inversion of a large matrix when the number of data points increases involving large computational time when cross-validation is used for parameters selection. The gKDR method from Fukumizu and Leng (22) solves these issues with a much faster algorithm for dimension reduction subspace identification by only using an eigendecomposition after

Gram matrix manipulation.

The gKDR approach relies on the fact that, supposing (17) holds, the derivative $\partial\mathbb{E}(Y \mid \boldsymbol{X} = \boldsymbol{x})/\partial\boldsymbol{x}$ is contained in the SDR subspace and can be further used to estimate the projection matrix B. A non-parametric estimator of this derivative is obtained through covariance operators and the dimension reduction matrix B is identified as the solution of an eigenproblem with few matrices inversions. This fast procedure enables regularization and kernel parameter selection through cross-validation and can be applied to large training sets in high dimensions. For a detailed description of gKDR, we refer the reader to the original paper by Fukumizu and Leng (22).

Note that both KDR and gKDR require a priori knowledge of the SDR subspace dimension r which is seldom the case in practical applications. In the case of gKDR, we suggest a cross-validation procedure based on the mean-squared error of the standard regression estimate of Y w.r.t. $B_r^T\boldsymbol{X}$ where B_r is the dimension reduction matrix of rank r. This procedure can be made more efficient since gKDR defines the columns of the dimension reduction subspace matrix estimate $\hat{B}$ as the eigenvectors corresponding to the r largest eigenvalues of some symmetric $n \times n$ matrix (see Fukumizu and Leng (22)). This implies that if $r < s$, then the columns of $\hat{B}_r$ are the first r columns of $\hat{B}_s$, possibly up to a sign. It is therefore sufficient to only compute $\hat{B}$ for the largest candidate dimension for all cross-validation folds leading to a faster algorithm.

4. Importance sampling and metamodeling in the SDR subspace

In our context, as stated in 3.1, the aim of sufficient dimension reduction is to enable Kriging metamodelling in a space of much lower dimension, given by the dimension reduction subspace. We now consider a Kriging metamodel of the limit state function output Y by using a Gaussian process model $GP(\mathbf{f}^T\boldsymbol{\beta}, C)$ for the link function g_r of the dimension reduction model in (17). The ensuing Kriging predictor is denoted $\hat{G}_r$. For reasons that will be clarified in the sequel, we restrict ourselves to limit state functions that are defined for random inputs $\boldsymbol{X}$ with support in $\mathbb{R}^d$.

4.1. Quasi-optimal importance density with metamodelling in the dimension reduced space

Consider the natural extension of the original quasi-optimal importance density defined as

$$\tilde{q}_{r*}(\boldsymbol{x}) = \frac{\pi_r(B^T\boldsymbol{x})q(\boldsymbol{x})}{\int \pi_r(B^T\boldsymbol{x})q(\boldsymbol{x})\mathrm{d}\boldsymbol{x}} \tag{21}$$

where π_r is the probabilistic classification function of the Kriging metamodel in the SDR subspace

$$\pi_r(\boldsymbol{z}) = \Phi\left(-m_{\hat{G}_r}(\boldsymbol{z})/\sigma_{\hat{G}_r}(\boldsymbol{z})\right) \tag{22}$$

The potential error arising from the replacement of π by π_r is taken into account with the nugget effect simultaneously estimated with the others hyper-parameters of the GP model. Adding a nugget is almost equivalent to add a Gaussian noise to the model with the only difference that we still interpolate the simulated data. This "noisy" model is then use to construct the BLUP and therefore propagate to the probability π_r. This approach can be justified by a fisrt order approximation of the prior Gaussian process taking into account a (small) error on the projection matrix that would suggest to add a Gaussian error noise to G_r.

The ensuing augmented failure probability estimates and correction factor using the IS density $\tilde{q}_{r*}$ then read

$$P_{f,\epsilon} = \mathbb{E}_q(\pi_r(B^T \boldsymbol{X})) \tag{23}$$

$$\alpha_{\text{corr}} = \mathbb{E}_{\tilde{q}_{r*}}\left(\mathbb{1}_{g(\boldsymbol{X})\leq 0}/\pi_r(B^T \boldsymbol{X})\right) \tag{24}$$

4.2. Reduced Kriging metamodel refinements

Regarding the sampling based Kriging refinement strategy outlined in 2.3, a similar approach may readily be applied, that is we define the weighted margin probability density

$$h_r(\boldsymbol{x}) \propto w_r(B^T \boldsymbol{x}) q(\boldsymbol{x}) \tag{25}$$

where w_r is defined as w introduced in 2.3 but in the reduced space. The refinement strategy using a reduced metamodel may now proceed as follows: given a sample $\boldsymbol{x}^i$, $i = 1, \ldots, n$, drawn from a Markov chain targeting h_r, N_{add} cluster centers are determined via an Euclidean based k-means algorithm and then projected onto the dimension reduction subspace using matrix B. These projected points can then be added to the DoE. However this runs the risk that two cluster centers have very close projections which is inefficient for the construction of the metamodel of the link function in the projected space (17) and furthermore can lead to numerical instability of the Kriging surrogate.

To avoid this, we suggest another strategy based on a clustering in the projected space. We first generate a space-filling design, of size n, in the full space: a sphere of "maximal radius" (to be presented later) in $\mathbb{R}^d$. Then the following three step are performed:

1. let $\boldsymbol{z}_i = B^T \boldsymbol{x}^i$, $i = 1, \ldots, n$ be the projected samples,
2. find N_{add} cluster centers $\{\boldsymbol{c}_j\}_{j=1}^{N_{add}}$ of the $\boldsymbol{z}_i$ dataset,
3. let i_j be the index of the closest projections z^i to c_j: add $(\boldsymbol{z}^{i_j}, y^j = g_r(\boldsymbol{z}^{i_j}))$ by evaluating $y^j = g(\boldsymbol{x}^{i_j})$.

The last step is a tweak that avoid the issue we would face if we directly back transform the cluster centers in $\mathbb{R}^r$ to $\mathbb{R}^d$ by using a generalized inverse of the projecting matrix: $B^{\#}$. Indeed, in the case where the input domain $\mathbb{X}$ is not

$\mathbb{R}^d$, there is no guarantee that $\boldsymbol{x} = (B^T)^{\#}\boldsymbol{z} \in \mathbb{X}$.

4.3. The case of Gaussian inputs

The augmented failure probability and correction factor (23) and (24) normally require ordinary Monte Carlo or MCMC sampling in a d-dimensional space. Admittedly, (23) does not resort to expensive limit state function evaluations but a small value of the target failure probability may require a substantial amounts of sampling in $\mathbb{R}^d$ and calls to the Kriging predictor. More importantly, the estimation of the correction factor requires sampling from $\tilde{q}_{r*}$ which is achieved through an MCMC algorithm such as the Metropolis-Hastings (MH) algorithm. Tuning the MH algorithm in high dimensions can prove cumbersome but approaches such as the modified Metropolis-Hastings algorithm (Au and Beck (2)) may prove successful.

Let us assume that the input $\boldsymbol{X}$ is a standard multivariate random vector with Gaussian components. The estimation of (23) and (24) is made more efficient as shown below.

Lemma 1. (i) *The augmented failure probability $P_{f,\epsilon}$ can be expressed as $P_{f,\epsilon} = \mathbb{E}(\pi_r(\boldsymbol{Z}))$ where $\boldsymbol{Z}$ is an r dimensional standard normal variable.*

(ii) *Let $B_a = [B, B_\perp]$ where the columns of $B_\perp$ form an orthonormal basis of* $\mathrm{span}(B)^\perp$. *Let $\boldsymbol{W}_2 \sim \mathcal{N}(\boldsymbol{0}_{d-r\times 1}, I_{d-r})$, $\boldsymbol{W}_1 \sim p_{\boldsymbol{W}_1}(\boldsymbol{w}_1) = \pi_r(\boldsymbol{w}_1)\varphi_r(\boldsymbol{w}_1)/P_{f,\epsilon}$ where φ_r is the standard r-dimensional multinormal pdf and $\boldsymbol{W} = [\boldsymbol{W}_1^T, \boldsymbol{W}_2^T]^T$. Then $\tilde{\boldsymbol{X}} = B_a^{-T}\boldsymbol{W}$ is distributed according to $\tilde{q}_{r*}(\boldsymbol{x})$.*

The first result regarding the augmented failure probability computation is a straightforward consequence of the fact that $\boldsymbol{Z} = B^T\boldsymbol{X} \sim \mathcal{N}(0, I_r)$ since $B^T B = I_r$. The proof of (ii) is deferred to the supplementary material.

The practical implications of these results are that:

- the estimation of $P_{f,\epsilon}$ requires sampling of a standard distribution in $\mathbb{R}^r$ instead of $\mathbb{R}^d$.
- the estimation of the correction factor boils down to MCMC sampling in a space of expected much lower dimension r which voids some of the limitations of high dimensional MCMC. The second part of the sampling procedure is a straightforward standard Gaussian generation which makes the whole procedure faster than performing MCMC sampling in $\mathbb{R}^d$ directly where $d \gg r$ potentially.
- note that lemma 1 can also be used to sample the weighted margin probability density: the method is strictly the same if one replaces the probabilistic classification function with the margin probability of the reduced Kriging model. Also, given that we only need the projection of cluster centers of the samples onto the dimension reduction subspace, this implies that we only need to simulate a chain with stationary un-normalized distribution $w_r(\boldsymbol{w}_1)\varphi_r(\boldsymbol{w}_1)$.

4.4. The general case

In the general case, one might simply consider the limit state function in the standard Gaussian space $\mathbb{U}$, which is denoted g_0. The $\mathbb{U}$-space formulation is classically achieved through an iso-probabilistic transform $T : \boldsymbol{X} \in \mathbb{X} \mapsto T(\boldsymbol{X}) = \boldsymbol{U} \in \mathbb{U}$ so that $g_0(\boldsymbol{U}) = g\left(T^{-1}(\boldsymbol{U})\right)$. In this paper, we consider the Nataf transform (see Liu and Kiureghian (38), Dutfoy and Lebrun (18, 19) for details and the implications of this kind of transformation), defined by $T = T_2 \circ T_1$, where

$$T_1 : \boldsymbol{X} \mapsto \boldsymbol{Z} = T_1(\boldsymbol{X}) = \left(\Phi^{-1}(F_1(X_1)) \dots \Phi^{-1}(F_d(X_d))\right)^T \tag{26}$$

$T_2 : \boldsymbol{Z} \mapsto T_2(\boldsymbol{Z}) = L_0^{-1}\boldsymbol{Z}$ where $L_0 L_0^T = R_0$ is the linear correlation matrix of $\boldsymbol{Z}$, Φ is the standard normal cdf and $F_1, \dots, F_d$ are the marginal CDFs of $\boldsymbol{X}$.

One could simply try dimension reduction on the response model $Y = g_0(\boldsymbol{U})$, but, because of the non-linear nature of the Nataf transform, even if there exist a projection matrix B that reduces the dimension in the $\mathbb{X}$-space, with respect to the function g, there is no guarantee that it is also the case in the $\mathbb{U}$-space with respect to the function $g \circ T^{-1}$. If T was linear then it would be the case. The $\mathbb{U}$-space formulation being convenient for practical reliability analysis it might be preferable to work in it after achieving dimension reduction in the $\mathbb{X}$-space. A quasi-optimal importance density in the $\mathbb{U}$-space is then obtained as

$$\tilde{q}_{r*}(\boldsymbol{u}) = \pi_r(B^T T^{-1}(\boldsymbol{u}))\varphi_d(\boldsymbol{u})/P_{f,\epsilon} \tag{27}$$

with $P_{f,\epsilon} = \int \pi_r(B^T T^{-1}(\boldsymbol{u}))\varphi_d(\boldsymbol{u})\mathrm{d}\boldsymbol{u}$. The correction factor reads

$$\alpha_{\text{corr}} = \mathbb{E}_{\tilde{q}_{r*}}\left(\frac{\mathbb{1}_{g_0(\boldsymbol{U})\leq 0}}{\pi_r(B^T T^{-1}(\boldsymbol{U}))}\right) \tag{28}$$

Now, because of the non-linear transformation, drawing samples from (27) is not as simple as in the case of Gaussian inputs $\boldsymbol{X}$. A chain targeting $\tilde{q}_{r*}$ has to be simulated in $\mathbb{R}^d$ via a relevant MCMC method. For this purpose, we suggest using a modified Metropolis-Hastings algorithm (Au and Beck (2)).

As far as the sampling based refinement criterion is concerned, we redefine the weighted margin probability in a similar fashion, that is

$$h_r(\boldsymbol{u}) = w_r(B^T T^{-1}(\boldsymbol{u}))\varphi_d(\boldsymbol{u}) \tag{29}$$

Since this density may have modes that are far apart, slice sampling is recommended as an alternative to standard Metropolis Hastings which typically struggles in this setting, as noted by Dubourg (14). Once samples are obtained, we back-transform them to the physical input domain $\mathbb{X}$ and proceed as outlined in 2.3.

4.5. The Meta-ISDR algorithm

We hereafter sketch the Meta-ISDR algorithm, assuming $\boldsymbol{X}$ to be a standard Gaussian vector. The general case may be obtained by considering the modifications stated in 4.4. Note that we assume a preliminary single/multiple failure points search but this is optional. The main parameters of the algorithm, along with sensible default parameters values (in brackets) are:

- $\boldsymbol{x}_1^*, \ldots, \boldsymbol{x}_m^*$: $m \geq 1$ most probable failure points (optional)
- N_{SDR}: number of samples for the SDR basis estimation
- r_{max}: maximum SDR subspace dimension
- K_{cv} (5): number of CV folds for the reduced dimension estimation
- $k_{\mathcal{X}}$, $k_{\mathcal{Y}}$: Gaussian kernels used in SDR.
- N_{max}: maximum number of limit state function evaluations
- R_{max}: radius of the hypersphere enclosing the design of experiments
- N_D^{init} $(2d)$: initial size of DoE
- N_{fill} (10^5) number of samples generated for the space-filling design
- $N_D^{\min}$, $N_D^{\max}$ minimum/maximum number of points in final design of experiments
- N_{add}: number of points added to DoE during refinement step
- $\alpha_{\text{LOO}}^{\min}$ (0.1), $\alpha_{\text{LOO}}^{\max}$ (10): min/max of leave-one-out criteria for the metamodel quality assessment
- N_{aug}: number of samples used for the estimation of the augmented failure probability $\hat{P}_{f,\epsilon}$
- δ_ϵ, δ_α: target coefficient of variation (c.o.v.) for $\hat{P}_{f,\epsilon}$ and α_{corr}

Although the random input vector is not bounded, the radius parameter R_{max}, which defines the domain of the Kriging metamodel, can be set by choosing R_{max} such that $\mathbb{P}(\|\boldsymbol{X}\|^2 > R_{max}^2)$ is much smaller than P_f. For instance if $P_f >> 10^{-b}$ for some positive integer b then solving $\mathbb{P}(\|\boldsymbol{X}\|^2 > R_{max}^2) = 10^{-b}$ yields $R_{max} = \sqrt{q_{\chi_d^2, 1-10^{-b}}}$ where $q_{\chi_d^2, 1-10^{-b}}$ is the quantile of order $1 - 10^{-b}$ of the chi-squared distribution with d degrees of freedom. The complete algorithm is given in algorithm 1.

5. Numerical illustrations

We now investigate the performance of the metamodel-based importance sampling procedure with dimension reduction on two analytical test cases and a realistic high dimensional industrial application consisting in estimating the

Algorithm 1 Meta-ISDR algorithm for standard Gaussian inputs

SDR Basis estimation:

0. Research of design points. (optional)

1. Sample N_{SDR} SDR learning points. **If** step 0 is skipped, sample $\boldsymbol{x}^i_{SDR}$ from a uniform distribution in the hypersphere of radius R_{max} **Else** combine the latter distribution in a mixture with (local) uniform distributions within spheres of given radius centered at the discovered DP. Compute $y^i_{SDR} = g(\boldsymbol{x}^i_{SDR})$.

2. Using gKDR algorithm find a dimension reduction subspace estimate $\hat{B}, \hat{r} \leftarrow \texttt{gKDR}(\boldsymbol{x}_{SDR}, \boldsymbol{y}_{SDR}, k_X, k_Y, K_{cv}, r_{max})$

Initial DoE:

1. Add design points $\boldsymbol{x}^*_1, \ldots, \boldsymbol{x}^*_m$ to DoE (optional)

2. Space-filling design: sample N_{fill} points in the hypersphere of radius R_{max} and, following the clustering strategy presented at the end of section 4.2, add $N_D^{init} - m$ initial DoE points in the reduced subspace by evaluating $y_j^{init} = g(\boldsymbol{x}_j^{init})$, $j = 1, \ldots, N_D^{init} - m$

3. Initial metamodel: fit a Kriging model to initial DoE and compute $\hat{\alpha}_{\text{LOO}}$

4. Metamodel refinement:

While $N_D < N_D^{\min}$ **or** $\hat{\alpha}_{\text{LOO}} \notin [\alpha^{\min}_{\text{LOO}}, \alpha^{\max}_{\text{LOO}}]$

sample $n_1 \in [10^4, 5 \times 10^4]$ points z^i, $i = 1, \ldots, n_1$ from h_r (defined in (25)) using procedure from lemma 1.

add N_{add} points to DoE according to the procedure outlined in 4.2.

$N_D \leftarrow N_D + N_{add}$.

If $N_D > N_D^{\max}$ **break**

end If

end While

Failure probability estimation:

1. Augmented probability estimation: compute $\hat{P}_{f,\epsilon}, \hat{\delta}_\epsilon$ using standard Monte Carlo or the splitting introduced in section 2.2

2. Correction factor estimation: run an MCMC of length $N_{corr} = b + N_{max} - N_{SDR} - N_D$ targeting $\tilde{q}_{r*}$ as described in lemma 1. Compute $\hat{\alpha}_{\text{corr}}$ and $\hat{\delta}_\alpha = \hat{\sigma}_{\text{corr}}/\hat{\alpha}_{\text{corr}}$ according to (13) and (14)

3. Failure probability estimate: $\hat{P}_f = \hat{\alpha}_{\text{corr}} \hat{P}_{f,\epsilon}$, $\hat{\delta} = \sqrt{\hat{\delta}^2_\alpha + \hat{\delta}^2_\epsilon + \hat{\delta}^2_\alpha \hat{\delta}^2_\epsilon}$

failure probability of a wind turbine under stationary wind loads.

5.1. Academic examples

We consider two analytical limit state functions whose dimensionality can be varied to illustrate the impact of an increasing number of inputs. The Meta-ISDR strategy presented here is confronted with standard reliability methods, namely FORM, standard Monte Carlo simulation, importance sampling with a proposal centered on the design point(s) (IS-FORM) and the standard Meta-IS extrated from the R package MISTRAL. Note that, with the exception of FORM, we only consider methods that yield consistent probability estimation. The design point in FORM was computed using a sequential quadratic programming algorithm as described in Liu and Der Kiureghian (37). In all numerical tests, when $d > 2$, the maximum projection space dimension: r_{max}, was set to 10 and 2 when $d = 2$.

First test function. The first performance function extracted from Bourinet (6) reads $g_1(\boldsymbol{X}) = d + a\sigma\sqrt{d} - b^T\boldsymbol{X}$ where the X_i are d i.i.d. lognormal variables with mean 1 and standard deviation $\sigma = 0.2$, $a = 3$ and $b = \left(1 \cdots 1\right)^T$. We consider 3 cases with increasing dimension: $d = 2, 50, 100$. It is obvious that a sufficient dimension reduction subspace is spanned by b so we can expect sizeable performance improvements in high dimensions for the Meta-ISDR method. As can be expected for $d = 2$, the dimension reduction might not worth it since it comes at an additional cost of limit state function evaluations. In table 1, $d = 2$, we notice that the total number of calls to g_1 for Meta-ISDR is 5 times that of standard MetaIS because of the SDR extra cost. With 300 simulations assign to the dimension reduction step, the method selects two projection vector which correspond to a rotation of the initial canonical dimensions. The first unit vector is very close to a normalized b. The method balance the latter approximation by adding a second orthogonal dimension. For higher dimensions however, it appears that the metamodelling, in a reduced space, limits the number of limit state function evaluations: this reduction occurs at the Kriging model construction but is even more noticeable for the correction factor estimation. This is probably because a more accurate Kriging model is obtained in the reduced dimensional SDR subspace leading to an importance sampling density closer to the optimal one hence achieving lower estimation variance. The selected reduced dimension was 5 with one of the unit vector close to a normalized b. Again, to balance for the approximation, a few additional dimension were added by the method.

Second test function. The second test function is a 2-dimensional non-linear function: $g_2(\boldsymbol{X}) = 1.2 - X_2 + 0.2(X_1 + X_2)^4$. The purpose of this function is to illustrate, in a non-linear context, how the dimension reduction technique proposed can retrieve a basis well suited for following reliability analysis. A similar example is also studied in dimension $d = 50$ such that only the fifth first variables are influential: $g_2(\boldsymbol{X}) = 1.2 - X_5 + 0.2(c_1X_1 + c_2X_2 + c_3X_3 + c_4X_4)^4$. with $c = (0.1, 1, 0.5, 1)$. The purpose of this latter extension is to demonstrate the efficiency of the proposed strategy to retrieve an appropriate basis when faced with a non-linear model and unequal importance weights on the variables.

In both instances of function g_2, all the variables follow a standard normal distribution.

In table 2, on the $d = 2$ version of the function, with only an hundred simulations the SDR method find a projection basis which corresponds to a rotation of the canonical one where the first unit vector point in the direction of the design point as illustrated in figure 1. We then achieve a precise probability estimation with 200 additional function evaluations. Although, a more suitable rotation can be achieved by adding more model evaluations to the SDR step. Indeed, with 300 simulations for the SDR step, the first unit vector of the orthogonal basis point in the direction of greater variation of the limit state function around the design point as illustrated in figure 1. Again 200 additional evaluations are necessary to achieve the same coefficient of variation as previously. In table 2, facing the high dimensional non-linear function version of g_2, a large number of samples (3500) were requiered to learn precisely the expected basis of projection. Indeed, the estimated first basis correspond to a close approximation of a normalized c and the second basis correspond to the standard fifth dimension basis. However, to compensate for the two first basis approximations, one additional basis was proposed by the SDR method. On this specific example, we see that an IS strategy centered on the DP is a good compromise because of the unimodal and localized main area of failure. Nevertheless the SDR method enables to recover relevant informations about the influential part of the feature space. The following MetaIS analysis requires a non-negligible number of evaluations mainly related to the α_{corr} estimation. The MetaIS parameters were set to default values and a reduction of the number of evaluations can be expected with a finer adjustement of these parameters. Further numerical tests, not presented here, suggest two possible improvement of the Meta-ISDR method: first to remove the orthogonal constraint on the projection basis and second to replace the linear model used in the tuning (by cross validation) of the projection dimension (r) by a non-parametric model such as a Kriging model (potentially scaled for a fast model estimation). We can also emphasize the fact that the Meta-IS algorithm (independently of the SDR) requires a fine tuning when facing high dimensional and non-linear models.

5.2. Reliability assessment of a wind turbine in stationary conditions

We now turn to an industrial case study where the goal is the reliability analysis of an onshore wind turbine under wind loading. We use a 5MW wind turbine mode developed by the National Renewable Energy Laboratory (NREL). The turbine has a 90m hub height and implements a blade and generator control strategy. The mechanical response to the wind inflow is obtained using the FAST software which provides extreme and fatigue loads for a wide array of turbines. The random wind field is described in terms of its (u, v, w) coordinate where u is a vector pointing in the same directions as the mean wind flow, (v, w) completes the orthogonal basis and thus mean wind speeds in the v and w directions are zero. For the purpose of our analysis, we neglect the turbulent wind flow components in the v and w directions and consider the wind speed along u as the random input load.

Method	**FORM**	**Monte Carlo**	**IS-FORM**	**Meta-IS**	**Meta-ISDR**
d = 2					
N	18	5×10^5	8818	121	519(300)
$\hat{P}_f$	3.81×10^{-3}	4.93×10^{-3}	4.90×10^{-3}	4.88×10^{-3}	4.86×10^{-3}
c.o.v.	-	2%	2%	1.3%	1.3%
d = 50					
N	255	1.4×10^6	14255	3255	1281(500)
$\hat{P}_f$	1.54×10^{-3}	1.89×10^{-3}	1.95×10^{-3}	1.87×10^{-3}	1.85×10^{-3}
c.o.v.	-	1.9%	1.9%	2%	1.9%
d = 100					
N	505	1.5×10^6	18505	5505	2288(1000)
$\hat{P}_f$	3.74×10^{-5}	1.73×10^{-3}	1.76×10^{-3}	1.76×10^{-3}	1.71×10^{-3}
c.o.v.	-	1.8%	1.9%	2%	1.9%

Table 1. Failure probability estimation: analytical limit state function $g_1(\boldsymbol{X}) = d + a\sigma\sqrt{d} - b^T\boldsymbol{X}$. N is the overall total number of function evaluations. For Meta-ISDR, N_{SDR} is specified in parentheses.

Method	**FORM**	**Monte Carlo**	**IS-FORM**	**Meta-IS**	**Meta-ISDR**
d = 2					
N	64	6×10^4	9064	537	265(100)
$\hat{P}_f$	8.6×10^{-2}	3.3×10^{-2}	3.3×10^{-2}	3.7×10^{-2}	3.4×10^{-2}
c.o.v.	-	2.1%	2.3%	4%	2%
d = 50					
N	306	10^4	3306	NA	6306(3500)
$\hat{P}_f$	8.2×10^{-3}	3.6×10^{-3}	3.6×10^{-3}	NA	3.4×10^{-3}
c.o.v.	NA	5.2%	3.9%	NA	4.9%

Table 2. Failure probability estimation: analytical limit state functions $g_2(\boldsymbol{X})$. N is the overall total number of function evaluations. For Meta-ISDR, N_{SDR} is specified in parentheses.

5.2.1. Wind speed model

The wind speed process $\{X(t), t \geq 0\}$ was modelled as a stationary Gaussian process according to a spectral expansion model, as in Shinozuka and Deodatis (46),

$$X(t) = U_{10} + \sum_{i=1}^{n} (u_i \sigma_i \cos(\omega_i t) - \bar{u}_i \sigma_i \sin(\omega_i t)) \tag{30}$$

u_i, $\bar{u}_i$ are standard independent normal variables, ω_i are the frequencies with increment $\mathrm{d}\omega_i = w_{i+1} - w_i$ and $\sigma_i^2 = S(\omega_i)\mathrm{d}\omega_i$ where S is the power spectrum density (p.s.d.) of $\{X(t), t \geq 0\}$. U_{10} is the 10-minute mean wind speed. The spectrum was estimated from real measurements on the Danish coastal site Hornsrev. To achieve near-stationarity,

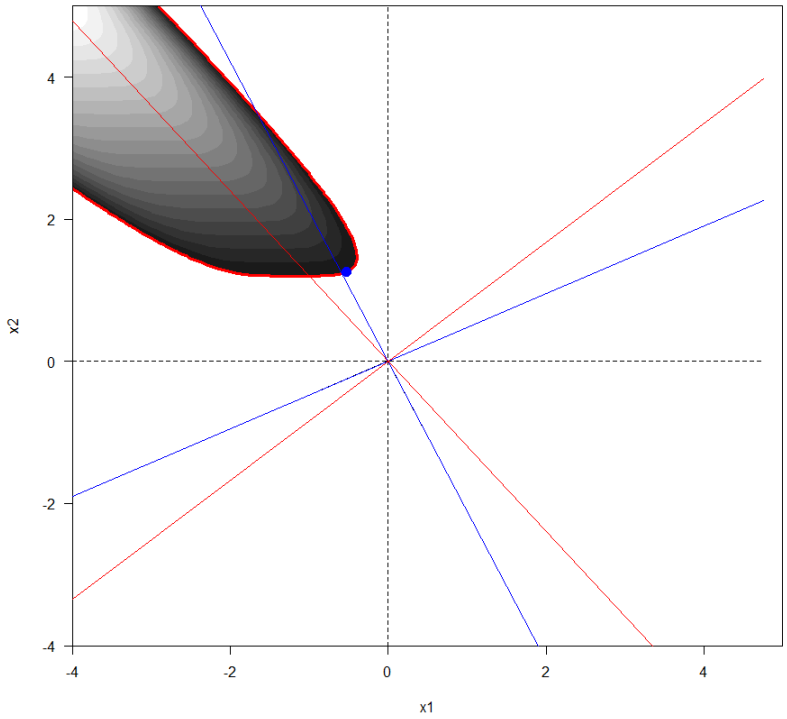


Figure 1. Illustration of the limit state function g_2 with $d = 2$ (bold red curve), failure area (shade of grey) and design point (blue dot). New axes defined by the obtained 2d projection basis with respectively 100 (blue lines) and 300 (red lines) sufficient dimension learning samples.

we extracted 10-minute wind speed time series corresponding to a turbine rated mean speed of $U_{10} = 11.5$ m/s and turbulence intensity $I = 6\%$. These measurements where broken down into 5 minutes segments. A power spectrum density estimate was then obtained by averaging the periodograms of each segment. For simulation purposes, the spectrum is discretized into $n = 50$ harmonics so that the random vector $\boldsymbol{U} = [u_1, \ldots, u_n, \bar{u}_1, \ldots, \bar{u}_n]$ describing the random process X is 100-dimensional.

5.2.2. *Time variant reliability analysis*

The objective of this reliability analysis problem is to evaluate the 10-minute failure probability of the turbine. Failure is hereafter characterized as the exceedance of a distance threshold s by the tower-top displacement due to fore-aft forces. Letting $\{Y(t), t \geq 0\}$ be the mechanical response, we seek the following quantity

$$P_f(T) = \mathbb{P}(\exists t \in [0, T] \mid Y(t) \geq s) \tag{31}$$

where $T = 10$ minutes. Equivalently, we may introduce a time-varying limit state function $g(t, \boldsymbol{U}) = s - Y(t)$ where $\boldsymbol{U}$ are the coefficients of the spectral expansion model, which yields $P_f(T) = \mathbb{P}(\exists t \in [0, T] \mid g(t, U) \leq 0)$.

It is well known (Andrieu-Renaud et al. (1)) that a related quantity is the expected number of downcrossings (of level 0 by the limit state g) which bound the seek failure probability and which is given by the integral over time of the outcrossing rate defined as

$$\nu^+(t) = \lim_{\Delta t \to 0} \mathbb{P}\left(g(t, \boldsymbol{U}) > 0, g(t + \Delta t, \boldsymbol{U}) \leq 0\right) \Delta t^{-1} \tag{32}$$

The outcrossing rate is the actual quantity of interest in the usual time-varying reliability framework. In most random vibration models, the response Y is modelled thanks to a linearized stochastic differential equation. The non linearity and non Gaussian nature of Y is then due to the loading force term (see Jensen et al. (25). Assuming wind speed stationarity, one does not automatically obtain output stationarity because of the coupling that goes on between the loading force and the response. However, simulations show that for stationary wind conditions, the response can generally be considered stationary after a transitory period which corresponds to the memory of the system. This means ν^+ is independent of the time t for sufficiently large t. Therefore, its computation can be obtained by considering the limit state function at a fixed time t_0. In practice, the simulation length must be sufficiently high so as to discard the initial transient part of the response signal. We however acknowledge that the stationarity assumption of the response assuming stationary inputs must be assessed more carefully but this is left out for a future study.

An approximation of the outcrossing rate may be obtained by considering an integration time-step $\Delta t << 1$ so that

$$\hat{\nu}^+ = \mathbb{P}\left(H_{t_0, t_0 + \Delta t}(\boldsymbol{U}) \leq 0\right) \Delta t^{-1} \tag{33}$$

where $H_{t_0, t_0 + \Delta t}(\boldsymbol{U}) = \min\left(-g(t_0, \boldsymbol{U}), g(t_0 + \Delta t, \boldsymbol{U})\right)$. In our settings, we consider a threshold on the tower-top displacement $s = 0.4$ m and $\Delta t = 0.01$ s. The simulation length t_0 is one minute.

5.2.3. *Instantaneous failure probability*

To illustrate the performance of the Meta-ISDR algorithm with respect to other simulation-based reliability algorithms, we first look at the instantaneous failure probability at time $t_0 = 60$ s denoted $P_{f,t_0} := \mathbb{P}\left(g(t_0, \boldsymbol{U}) < 0\right)$. As for the academic examples, we also implement for comparison purposes a design point based method, an importance sampler, and subset simulation (Au and Beck (2)).

- We considered a multiple failure point (MFP) search based on the approach of Der Kiureghian and Dakessian (12): this method has the ability to find several failure points which contribute significantly to the failure probability. This algorithm essentially modifies the limit state function in the vicinity of a previously found

design point thus redefining a new constrained optimization problem geared towards finding a failure point far enough from previous solutions. The constrained optimization method used in our simulations was the sequential quadratic approximation (SQA) algorithm from Sinoquet and Langoüet (47), a gradient free method which proved efficient in terms of number of limit state function evaluations. The MFP search was done with a limit of 5 distinct design points since no more dominant failure zone are expected from the failure function.

- The importance sampler used a mixture proposal where each component was a unit covariance Gaussian distribution centered on a failure points identified by the MFP approach and has the same weight in the mixture.
- The subset simulation algorithm was ran using the original algorithm by Au and Beck (2) with an added thinning step and with a different length for the MCMC chain. The conditional probabilities where set to $p_0 = 0.1$ and a modified Metropolis-Hastings chain of length $N_{subset} = 1000$ was generated at each subset level, using a thinning interval of length 3.

The Meta-ISDR algorithm was implemented with the following setup.

- The maximum dimension of the SDR subspace is fixed to $r_{max} = 6$. This choice is case/user dependent. In our case, we selected the value $rmax = 6$ such that the overall calculus time was reasonnable with regard to the computing time for one evaluation of the used simulator.
- The dimension reduction subspace was estimated with the KDR algorithm using $N_{KDR} = 1000$ training samples. The training set was constructed by drawing 50% samples from a uniform distribution in the hypersphere in $\mathbb{R}^d$ of radius $R_{max} = \sqrt{q_{\chi^2,1-10^{-8}}}$ and 50% samples from a Gaussian mixture distribution centered on the design points . Note that this R_{max} parameter corresponds to a lower bound on the instantaneous failure probability of 10^{-8} (see section 4.5). As discussed in section 3.1, this latter training DoE enables the detection of the projection space which explains the output not only globaly but also around the limit state,
- A kriging model with constant trend β and an anisotropic Matern-5/2 covariance function is considered. The trend, the kernel correlation lengths and the nugget were estimated by maximizing the concentrated log-likelihood.
- The initial DoE used for the kriging learning stage is of size $N_D^{init} = 10r + m$: Latin Hypercube Sampling of size $10 \times r$ and the additional m "design points", where r is the estimated reduced dimension.
- The minimum/maximum size of the DoE for the kriging model refinement are set to $N_D^{\min} = 60$, $N_D^{\max} = 300$.
- The number of points added to the DoE during one step of refinement is $N_{add} = 2r$.

For all methods, the target coefficient of variation was 7%. For the Meta-ISDR algorithm, a maximum number of $N_{max} = 11000$ limit state function evaluations was set, including the number of calls due to the preliminary multiple failure points.

	$\hat{P}_{f,t_0}$	c.o.v.	# G calls	$P_{f,\epsilon}$	α_{corr}
Multi-FORM	3.78×10^{-6}	NA	4657	NA	NA
IS	2.68×10^{-5}	7%	84657	NA	NA
MetaIS	2.41×10^{-5}	5.3%	14657	5.49×10^{-5}	0.44
Meta-ISDR	2.34×10^{-5}	5.8%	10645	3.25×10^{-5}	0.7198
SS	2.73×10^{-5}	27%	57000	NA	NA

Table 3. Instantaneous Failure probability estimation: method comparison

The multiple failure points algorithm found two significant design points. A FORM approximation based on the lowest reliability index was then performed, while both design points were used in the importance sampler and Meta-ISDR algorithm. The results in table 5.2.3 indicate the good performance of the Meta-ISDR algorithm as it achieves both the target coefficient of variation while also requiring the least amount of limit state function evaluations. The estimated reduced dimension was $\hat{r} = 2$ and the first projection basis matched the direction of the main design point. The standard MetaIS requires 37% more limit state evaluations. The performance of the subset simulation algorithm in terms of estimation error is due to the limited length of Markov chains simulated at each level, which is not enough to mitigate the chain's autocorrelation. A better error can be achieved by increasing the thinning interval for instance but this results in running a longer chain and more evaluations of the limit state.

5.2.4. Outcrossing rate estimation

As previously stated, the object of interest for time-variant reliability analysis in stationary conditions is the outcrossing rate ν^+. The formulation (33) lends itself to standard reliability analysis algorithms and is used to compute numerical outcrossing rate estimates. The FORM outcrossing rate was computed using an approximation due to Koo et al. (29).

Surprisingly, the FORM approximation for this case study seems more accurate than the corresponding instantaneous failure probability computed in the previous subsection and is comparable to the estimation of Meta-ISDR. We however expect the FORM outcrossing rate estimation to be inaccurate for more pronounced non-linearities. Again the estimated reduced dimension was $\hat{r} = 2$ and the first projection basis matched the direction of the main design point. Since the Meta-ISDR cost includes the MFP search, we can see the difference of number of calls to the failure function

	$\hat{\nu}^+$	c.o.v.	# G calls	$\nu^+_{f,\epsilon}(Q)$	$\alpha_{\text{corr}}(Q)$
Multi-FORM	2.50×10^{-5}	NA	4657	NA	NA
Meta-ISDR	1.99×10^{-5}	0.16	10657	2.07×10^{-3}	9.63×10^{-3}

Table 4. Short-term outcrossing rate estimation

between the two approaches as the price to pay, for this specific case study, to get a certified result on the outcrossing rate. In this example, a posteriori, this price to pay is not necessary but still should always be recommanded.

6. Conclusion

A concern of Monte Carlo based methods for structural reliability assessment such as standard importance sampling or subset simultation is their high computational cost as the dimension exceeds a few tenths and/or the failure probability gets below 10^{-4}.

Recently, the MetaIS algorithm, which combines a Kriging metamodel of the failure surface and an importance sampling procedure, enables more efficient estimation by sampling a so-called quasi-optimal density which acts as a surrogate to the optimal IS density. In high dimensional settings however, the Kriging metamodel construction requires a non-negligible number of limit state function evaluations so as to achieve sufficient accuracy in the vicinity of the failure region. This number influences directly the achievable variance reduction by the IS scheme.

Considering that in some reliability problems, the performance function depends on a projection of the input variables on a lower-dimension subspace, we have suggested to build a surrogate to the limit state function in this reduced subspace. The proposed approach leverages recent sufficient dimension reduction techniques to find this subspace. A cross-validation type procedure is suggested in order to infer the dimension of the reduced subspace. The MetaIS algorithm is then cast into this framework and yields particularly efficient MCMC sampling for Gaussian distributed inputs. Its applicability in dimensions up to 100 is demonstrated on a known academic example which illustrates the impact, of constructing a metamodel in the reduced subspace, on the efficacy of the probability estimator for a given confidence level. Finally, an industrial case study focused on the extreme response prediction of a wind turbine shows a notable reduction in the computational cost, compared to existing approaches which yield consistent estimators such as: subset simulation, standard importance sampling and standard MetaIS.

We believe that the Meta-ISDR approach can be further improved by removing the orthogonal constraint on the projection basis and by replacing the linear model, used in the tuning of the projection dimension parameter, by a non-parametric model such as a Kriging model (potentially scaled for a fast model estimation) or some cheap-to-evaluate non-linear regressor. Choosing the relevant number of evaluations to allocate to the SDR estimation is also

an important question since a bad SDR approximation can lead to difficulties in the following reliability analysis specially in high dimension.

References

[1] C. Andrieu-Renaud, B. Sudret, and M. Lemaire. The phi2 method: a way to compute time-variant reliability. Reliability engineering & system safety, 84:75–86, 2004.

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

[3] S-K. Au, C. Papadimitriou, and J.L. Beck. Reliability of uncertain dynamical systems with multiple design points. Structural Safety, 21: 113–133, 1999.

[4] S.K. Au. Augmenting approximate solutions for consistent reliability analysis. Probabilistic Engineering Mechanics, 22(1):77 – 87, 2007. ISSN 0266-8920. doi: https://doi.org/10.1016/j.probengmech.2006.08.004. URL http://www.sciencedirect.com/science/article/pii/S026689200600035X.

[5] J. Bect, D. Ginsbourger, L. Li, and E. Vasquez. Sequential design of computer experiments for the estimation of a probability of failure. Statistics and Computing, 22:773–793, 2012.

[6] J.-M. Bourinet. Rare-event probability estimation with adaptive support vector regression surrogates. Reliability Engineering & System Safety, 150:210 – 221, 2016.

[7] J.-M. Bourinet, F. Deheeger, and M. Lemaire. Assessing small failure probabilities by combined subset simulation and Support Vector Machines. Structural Safety, 33(6):343 – 353, 2011. ISSN 0167-4730.

[8] V. Caron, A. Guyader, M. Munoz Zuniga, and B. Tuffin. Some recent results in rare event estimation. In ESAIM PROCEEDINGS, january 2014, Vol. 44, p. 239-259, 2014.

[9] F. Chiaromonte and R.D. Cook. Sufficient dimension reduction and graphics in regression. Annals of the Institute of Statistical Mathematics, 54(4):768–795, 2002. ISSN 0020-3157.

[10] P.G. Constantine. Active subspaces - emerging ideas for dimension reduction in parameter studies. In SIAM spotlights, 2015.

[11] R. D. Cook. SAVE: A method for dimension reduction and graphics in regression. Communications in Statistics - Theory and Methods, 29: 2109–2121, 2000.

[12] A. Der Kiureghian and T. Dakessian. Multiple design points in first and second order reliability. Structural Safety, 20:37–49, 1998.

[13] O. D. Ditlevsen and H. O. Madsen. Structural Reliability Methods. John Wiley & Sons Inc, 1996. ISBN 0471960861.

[14] V. Dubourg. Adaptive surrogate models for reliability analysis and reliability-based design optimization. PhD thesis, Blaise Pascal University - Clermont II, 2011.

[15] V. Dubourg, F. Deheeger, and B. Sudret. Metamodel-based importance sampling for structural reliability analysis. Probabilistic Engineering Mechanics, 33:47–57, 2011.

[16] V. Dubourg, B. Sudret, and F. Deheeger. Metamodel-based importance sampling for structural reliability analysis. Probabilistic Engineering Mechanics, 33:47 – 57, 2013. ISSN 0266-8920. doi: https://doi.org/10.1016/j.probengmech.2013.02.002. URL http://www.sciencedirect.com/science/article/pii/S0266892013000222.

[17] N. Durrande, D. Ginsbourger, and O. Roustant. Additive covariance kernels for high-dimensional gaussian process modeling. Annales de la Faculté de Sciences de Toulouse, 21:481–499, 2012.

[18] A. Dutfoy and R. Lebrun. A generalization of the nataf transformation to distributions with elliptical copula. Probabilistic Engineering Mechanics, 24:172–178, 2009.

[19] A. Dutfoy and R. Lebrun. Do rosenblatt and nataf isoprobabilistic transformations really differ ? Probabilistic Engineering Mechanics, pages 577–584, 2009.

[20] B. Echard, N. Gayton, and M. Lemaire. AK-MCS: An active learning reliability method combining kriging and monte carlo simulation. Structural Safety, 33(2):145 – 154, 2011. ISSN 0167-4730.

[21] B. Echard, N. Gayton, M. Lemaire, and N. Relun. A combined importance sampling and kriging reliability method for small failure probabilities with time-demanding numerical models. Reliability Engineering & System Safety, 111(0):232 – 240, 2013. ISSN 0951-8320.

[22] K. Fukumizu and C. Leng. Gradient-based kernel dimension reduction for regression. Journal of the American Statistical Association, 109 (505):359–370, 2014.

[23] K. Fukumizu, F. R. Bach, and M. I. Jordan. Kernel dimension reduction in regression. The Annals of Statistics, 37:1871–1905, 2009.

[24] J. M. Hammersley and D. C. Handscomb. Monte Carlo Methods. Methuen and Co., 1964.

[25] J.J. Jensen, A.S. Olsen, and A. Mansour. Extreme wave and wind predictions. Ocean Engineering, 38:2244–2253, 2011.

[26] Z. Jiang and J. Li. High dimensional structural reliability with dimension reduction. Structural Safety, 69:35 – 46, 2017. ISSN 0167-4730. doi: https://doi.org/10.1016/j.strusafe.2017.07.007. URL http://www.sciencedirect.com/science/article/pii/S0167473016302077.

[27] L.S. Katafygiotis and K.M. Zuev. Geometric insight into the challenges of solving high-dimensional reliability problems. Probabilistic Engineering Mechanics, 23(2):208 – 218, 2008. ISSN 0266-8920. doi: https://doi.org/10.1016/j.probengmech.2007.12.026. URL http://www.sciencedirect.com/science/article/pii/S0266892007000707. 5th International Conference on Computational Stochastic Mechanics.

[28] A. Der Kiureghian and T. Dakessian. Multiple design points in first and second-order reliability. Structural Safety, 20(1):37 – 49, 1998. ISSN 0167-4730. doi: https://doi.org/10.1016/S0167-4730(97)00026-X. URL http://www.sciencedirect.com/science/article/pii/S016747309700026X.

[29] H. Koo, A. Der Kiureghian, and K. Fujimura. Design-point excitation for non-linear random vibrations. Probabilistic Engineering Mechanics, 20:136–147, 2005.

[30] C. Lataniotis, S. Marelli, and B. Sudret. Extending classical surrogate modelling to ultrahigh dimensional problems through supervised dimensionality reduction: a data-driven approach, 2018.

[31] N. Lelievre. Développement des méthodes AK pour l'analyse de fiabilité. Focus sur les évènements rares et la grande dimension. PhD thesis, 2018. URL http://www.theses.fr/2018CLFAC045. Thèse de doctorat dirigée par Gayton, Nicolas Génie mécanique Clermont Auvergne 2018.

[32] N. Lelièvre, P. Beaurepaire, C. Mattrand, and N. Gayton. Ak-mcsi: A kriging-based method to deal with small failure probabilities and time-consuming models. Structural Safety, 73:1 – 11, 2018. ISSN 0167-4730. doi: https://doi.org/10.1016/j.strusafe.2018.01.002. URL http://www.sciencedirect.com/science/article/pii/S0167473017301558.

[33] B. Li, H. Zha, and F. Chiaromonte. Contour regression: a general approach to dimension reduction. The Annals of Statistics, 33:1580–1616, 2005.

[34] J. Li and D. Xiu. Evaluation of failure probability via surrogate models. Journal of Computational Physics, 229(23):8966 – 8980, 2010. ISSN 0021-9991. doi: https://doi.org/10.1016/j.jcp.2010.08.022. URL http://www.sciencedirect.com/science/article/pii/S0021999110004699.

[35] K.-C. Li. Sliced inverse regression for dimension reduction. Journal of the American Statistical Association, 86:316–342, 1991.

[36] L. Li, J. Bect, and E. Vasquez. Bayesian subset simulation: a kriging-based subset simulation algorithm for the estimation of small probabilities of failure. In In Proceedings of PSAM 11 and ESREL 2012. Helsinki, Finland. June 25-29, 2012, 2012.

[37] P.-L. Liu and A. Der Kiureghian. Optimization algorithms for structural reliability. Structural Safety, 9(3):161 – 177, 1991. ISSN 0167-4730. doi: =.

[38] P.-L. Liu and A. Der Kiureghian. Multivariate distribution models with prescribed marginals and covariances. Probabilistic Engineering Mechanics, 1(2):105 – 112, 1986. ISSN 0266-8920. doi: https://doi.org/10.1016/0266-8920(86)90033-0. URL http://www.sciencedirect.com/science/article/pii/0266892086900330.

[39] R.E. Melchers. Importance sampling in structural systems. Structural Safety, 6:3–10, 1989.

[40] J. Morio and M. Balesdent. Estimation of Rare Event Probabilities in Complex Aerospace and Other Systems. A Practical Approach. Woodhead Publishing, 2015.

[41] R.N. Neal. Slice sampling. Ann. of Statist., 31:705–767, 2003.

[42] V. Picheny, D. Ginsbourger, O. Roustant, R.T. Haftka, and N-H. Kim. Adaptive designs of experiments for accurate approximation of a target region. Journal of Mechanical Design, 132:1–9, 2010.

[43] Carl Edward Rasmussen. Gaussian processes for machine learning. MIT Press, 2006.

[44] R. Rubinstein and D. Kroese. Simulation and the Monte Carlo method. Wiley, 2008.

[45] J. Sacks, W.J. Welch, T.J. Mitchell, and H.P. Wynn. Design and analysis of computer experiments. Statistical Science, 4(4):409 – 435, 1989.

[46] M. Shinozuka and G. Deodatis. Simulation of stochastic processes by spectral representation. Applied Mechanics Review, 44:191–204, 1991.

[47] D. Sinoquet and H. Langoüet. Sqa: a generic trust region derivative free optimization method for black box industrial applications. In ICCOPT international conference on continuous optimization, 2013.

[48] B. Sudret, V. Dubourg, and J. Bourinet. Enhancing meta-model-based importance sampling by subset simulation. In 16th IFIP Working Conference on Reliability and Optimization of Structural Systems, 2012.

[49] Z. Sun, J. Wang, R. Li, and C. Tong. Lif: A new kriging based learning function and its application to structural reliability analysis. Reliability Engineering & System Safety, 157:152 – 165, 2017. ISSN 0951-8320. doi: https://doi.org/10.1016/j.ress.2016.09.003. URL http://www.sciencedirect.com/science/article/pii/S0951832016304847.

[50] E. Vasquez and J. Bect. A sequential bayesian algorithm to estimate a probability of failure. In 15th IFAC Symposium on System Identification, 2009.

[51] Jun Xu and Ding Wang. Structural reliability analysis based on polynomial chaos, voronoi cells and dimension reduction technique. Reliability Engineering & System Safety, 185:329 – 340, 2019. ISSN 0951-8320. doi: https://doi.org/10.1016/j.ress.2019.01.001. URL http://www.sciencedirect.com/science/article/pii/S0951832018307749.

[52] W. Yao, G. Tang, N. Wang, and X. Chen. An improved reliability analysis approach based on combined form and beta-spherical importance sampling in critical region. Structural and Multidisciplinary Optimization, 60(1):35–58, Jul 2019. ISSN 1615-1488. doi: 10.1007/s00158-019-02193-y. URL https://doi.org/10.1007/s00158-019-02193-y.

[53] G. Yi, J. Q. Shi, and J. Q. Choi. Penalized GPR and classification. Biometrics, 2011:1285–1294, 67.

[54] W. Yun, Z. Lu, and X. Jiang. A modified importance sampling method for structural reliability and its global reliability sensitivity analysis. Structural and Multidisciplinary Optimization, 57(4):1625–1641, Apr 2018. ISSN 1615-1488. doi: 10.1007/s00158-017-1832-z. URL https://doi.org/10.1007/s00158-017-1832-z.