Structural reliability assessment through metamodel based importance sampling with dimension reduction

Achille Murangira, Miguel Munoz Zuniga, Timothée Perdrizet

Structural reliability assessment through metamodel based importance sampling with dimension reduction

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

[a]*IFP Energies Nouvelles, 1-4 Avenue du Bois Préau, 92852 Rueil-Malmaison*

Abstract

We present a method for reliability assessment through metamodel based importance sampling. The contribution is the use of the sufficient dimension reduction method which enables the construction of the limit state function Kriging surrogate in lower dimension. This metamodel is used to build an approximation of the optimal importance density as in the original MetaIS algorithm by Dubourg et al. (2011). 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 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 usually the limiting 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

1. Introduction

Structural reliability analysis 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 (l.s.f.) g which describes the safety level of the structure for a given input vector $X \in \mathbb{R}^d$. The failure domain F, corresponds to the set of inputs for which the performance function g is negative, i.e. $F = \{X \in \mathbb{R}^d \mid g(X) \leq 0\}$, and the failure probability is thus:

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

This probability may readily be estimated through standard Monte Carlo simulation, however since P_f is often less than 10^{-5} it takes roughly $10^7 - 10^8$ 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 an expensive computer model. Analytical approximations based on the most probable failure point have long been the practical alternative to Monte Carlo sampling. This has led to the popular FORM/SORM approximations (Ditlevsen and Madsen, 1996). However these approximations are often poor whenever there exists multiple failure modes or when the non-linearity are extreme. Moreover, the FORM/SORM methods do not provide any confidence bounds on the failure probability estimate. Importance sampling (IS) (Hammersley and Handscomb, 1964) 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 such as a mixture of standard distributions centred on the failure modes (Au et al., 1999) or a single design point (Melchers, 1989). Still, the number of evaluations of the limit state function required to reach a given accuracy may still be important, limiting the applicability of such a scheme. Subset simulation (Au and Beck, 2001) 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 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. AK-MCS (Echard et al., 2011), and AK-IS (Echard et al., 2013) are based on the active kriging principle, that is the performance function is replaced by a kriging metamodel which is iteratively refined so as to provide accurate kriging predictions in the vicinity of the limit state surface. While often superior to a simple FORM analysis, simple substitution of G by the metamodel gives generally non-consistent failure probability estimators. In the subset simulation framework, a similar approach has been proposed by Bourinet et al. (2011), where a SVM classifier is built to emulate the intermediate limit state surfaces.

To obtain consistent importance sampling estimators, the Meta-IS (Dubourg et al., 2011) method was introduced, where the

*Corresponding author

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 ($d \geq 20 - 30$), 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 dimension reduction techniques have been the object of much research in computational statistics in the last couple of decades. In this paper, we lay out a methodology for importance sampling based on the sufficient dimension reduction (SDR) framework (Chiaromonte and Cook, 2002) for regression. SDR is based on the assumption that the output statistical dependency on the input X can be described entirely by projecting X on a lower dimensional subspace. Based on this assumption, we propose to build a kriging metamodel in the dimension reduction 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 dimension reduction subspace. As for the correction factor which is based on a sample from an approximation to the optimal importance density, we demonstrate how sample generation can be achieved efficiently since MCMC sampling is only performed in the dimension reduction 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 dimension reduction tools, more specifically Kernel Dimension Reduction (KDR) (Fukumizu et al., 2009) and gradient-based KDR (Fukumizu and Leng, 2014). 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 and wave processes are are modelled as Gaussian processes 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 $X = (X_1, \ldots, X_d)^T$ with density q . As previously stated g denotes the limit state function.

Kriging models, also known as Gaussian process regression models, are flexible and efficient surrogates to complex computer codes. As most metamodels, they rely on an initial *design of experiments* (DoE) $y^* = \left\{x^1, \ldots, x^{N_D}\right\}$. 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 region of interest. For reliability purposes this region is in the vicinity of the limit state surface defined by $\{x \in \mathbb{R}^d \mid g(x) = 0\}$.

2.1. Gaussian process models

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

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

where $\mathbf{f}^T\boldsymbol{\beta}$ 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 constant vector in $\mathbb{R}^d$. Z is parametrized by its autocovariance function

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

where σ_G^2 is the GP variance and θ is a vector of parameters of the autocorrelation function R. Equation 2 is known as universal kriging (UK) as it allows for a non constant mean of the process G.
Prediction at a previously unobserved input x is based on the best linear unbiased prediction (BLUP) of $G(x)$ given the observations $\mathbf{y} = g(x^1, \ldots, x^{N_D})$ at the DoE y^*. The BLUP at x, denoted $\hat{G}(x)$, is a normal random variable $\mathcal{N}(m_{\hat{G}}, \sigma^2_{\hat{G}})$ where

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

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

are respectively the prediction mean and variance, $\mathbf{R}$ is the correlation matrix of the DoE defined by $\mathbf{R}_{ij} = R_\theta(x^i, x^j)$, $i, j = 1, \ldots, N_D$. $\mathbf{r} = \left[R_\theta(x - x^i)\right]_{i=1}^N$ is the cross-correlation vector between the prediction and the observations while $\mathbf{F}$ is the matrix defined by $\mathbf{F}_{ij} = f_j(x^i)$, $1 \leq i \leq N_D$, $1 \leq j \leq p$. The vector $\hat{\boldsymbol{\beta}}$ is the solution to a generalized least-squares problem

$$\hat{\boldsymbol{\beta}} = (\mathbf{F}^T\mathbf{R}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\mathbf{R}^{-1}\mathbf{y} \tag{6}$$

and $\mathbf{v}(x) = \mathbf{F}^T\mathbf{R}^{-1}\mathbf{r}(x) - \mathbf{f}(x)$.

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. (2011). The method is founded on the use of a quasi-optimal importance density which mimics the intractable optimal importance density.

2.2.1. Standard importance sampling

Importance sampling is a well known variance reduction method where the quantity of interest is an expectation of a integrable function. In the context of reliability methods, the

expectation of interest is $P_f = \mathbb{E}_q(\mathbb{1}_{g(X)\leq 0}) = \int_{\mathbb{R}^d} \mathbb{1}_{g(x)\leq 0} q(x)\mathrm{d}x$. Let $\tilde{q}$ be a density such that:

$$\mathbb{1}_{g(x)\leq 0} q(x) \neq 0 \implies \tilde{q}(x) \neq 0 \tag{7}$$

$$\mathbb{E}_q\left(\mathbb{1}_{g(X)\leq 0}\frac{q(X)}{\tilde{q}(X)}\right) < \infty \tag{8}$$

IS stems from the equality $P_f = \mathbb{E}_q(\mathbb{1}_{g(X)\leq 0}) = \mathbb{E}_{\tilde{q}}\left(\mathbb{1}_{g(X)\leq 0}\frac{q(X)}{\tilde{q}(X)}\right)$. Given an i.i.d. sample $x^{(i)}$, $i = 1, \ldots, n$ from $\tilde{q}$, the importance sampling estimator of P_f reads

$$\hat{P}_f^{IS} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{g(x^{(i)})\leq 0}\frac{q(x^{(i)})}{\tilde{q}(x^{(i)})} \tag{9}$$

The above estimator is unbiased and its variance is given by

$$\mathrm{Var}_{\tilde{q}}(\hat{P}_f^{IS}) = \frac{1}{n}\mathbb{E}_{\tilde{q}}\left(\mathbb{1}_{g(X)\leq 0}\frac{q^2(X)}{\tilde{q}^2(X)}\right) - \frac{P_f^2}{n} \tag{10}$$

which can readily be estimated by

$$\hat{\sigma}_{IS}^2 = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{g(x^{(i)})\leq 0}\frac{q^2(x^{(i)})}{\tilde{q}^2(x^{(i)})} - \frac{(\hat{P}_f^{IS})^2}{n} \tag{11}$$

An important result is that $\mathrm{Var}_{\tilde{q}}(\hat{P}_f^{IS}) = 0$ (Rubinstein and Kroese, 2008), for the following optimal proposal density

$$\tilde{q}_{opt}(x) = \frac{\mathbb{1}_{g(x)\leq 0} q(x)}{P_f} \tag{12}$$

which is generally difficult to sample from because it depends on the complex limit state function g. In Meta-IS samples are drawn from a quasi-optimal density obtained by replacing the indicator function with a a continuous probabilistic classification function linked to the kriging predictor.

2.2.2. Probabilistic classification function and quasi-optimal importance density

Since the Kriging prediction $\hat{G}(x)$ at an unknown location x is a normal random variable, it's straightforward to compute the following probabilistic classification function

$$\pi(x) = \mathbb{P}_0(\hat{G}(x) \leq 0) \tag{13}$$

where $\mathbb{P}_0$ is the probability measure associated with the Gaussian process G. It turns out that

$$\pi(x) = \begin{cases} \Phi(-\frac{m_{\hat{G}}(x)}{\sigma_{\hat{G}}(x)}) & \text{if } x \notin y^* \\ \mathbb{1}_{G(x)\leq 0} & \text{if } x \in y^* \end{cases} \tag{14}$$

where Φ is the standard normal cumulative distribution function (cdf). Now by swapping $\pi(x)$ and $\mathbb{1}_{G(x)\leq 0}$ in (12), one obtains the following quasi-optimal importance density (Dubourg et al., 2011):

$$\tilde{q}_*(x) = \frac{\pi(x)q(x)}{P_{f,\epsilon}} \tag{15}$$

where $P_{f,\epsilon} = \int \pi(x)q(x)\mathrm{d}x$ is the *augmented* failure probability. One would be tempted to use the simpler density $\tilde{q}_{\mathrm{meta}}(x) = \frac{\mathbb{1}_{m_{\hat{G}}(x)\leq 0} q(x)}{P_f}$ but this choice does not guarantee that the proposal density dominates the unnormalized pdf $\mathbb{1}_{g(x)\leq 0}q(x)$ as required by condition (7). Note that this is condition is always satisfied for $\tilde{q}_*$ no matter the Kriging model.

2.2.3. The Meta-IS failure probability estimator

Having defined the quasi-optimal instrumental density, the failure probability estimate may now be broken down as follows Dubourg et al. (2011)

$$P_f^{IS} = \int \mathbb{1}_{g(x)\leq 0}\frac{q(x)}{\tilde{q}_*(x)}\tilde{q}_*(x)\mathrm{d}x = P_{f,\epsilon}\int \frac{\mathbb{1}_{g(x)\leq 0}}{\pi(x)}\tilde{q}_*(x)\mathrm{d}x \tag{16}$$

$$= P_{f,\epsilon}\alpha_{\mathrm{corr}} \tag{17}$$

where $\alpha_{\mathrm{corr}} = \mathbb{E}_{\tilde{q}_*}(\frac{\mathbb{1}_{g(X)\leq 0}}{\pi(X)})$ is a factor that corrects for the bias in the augmented probability estimator which relies solely on the metamodel. The value α_{corr} actually depends on how accurate the Kriging predictor is in the failure domain F. It's clear that whenever the metamodel is highly accurate in the failure domain, the ratio $\frac{\mathbb{1}_{g(X)\leq 0}}{\pi(X)}$ is close to 1 and so is α_{corr}. In case of lower accuracy, there are regions of F with high uncertainty on the sign of the Kriging prediction, i.e. $\pi(x)$ is far from 1. This yields a correction factor also different from 1.
Let $x^{(i)}$, $i = 1, \ldots, N_{\mathrm{meta}}$ be an i.i.d. sample from the prior density q, then

$$\hat{P}_{f,\epsilon} = \frac{1}{N_{\mathrm{meta}}}\sum_{i=1}^{N_{\mathrm{meta}}} \pi(x^{(i)}) \tag{18}$$

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

$$\hat{\sigma}_\epsilon^2 = \frac{1}{N_{\mathrm{meta}}(N_{\mathrm{meta}}-1)}\sum_{i=1}^{N_{\mathrm{meta}}}\left(\pi(x^{(i)}) - \bar{\pi}_{N_{\mathrm{meta}}}\right)^2 \tag{19}$$

where $\bar{\pi}_{N_{\mathrm{meta}}}$, $\bar{\gamma}_{N_{\mathrm{corr}}}$ is the respective sample average. From a practical standpoint, sampling from q is usually straightforward whereas obtaining draws from $\tilde{q}_*$ requires adequate algorithms such as Monte Carlo Markov Chain simulation. Let $t^{(j)}$, $j = 1, \ldots$ be a chain with stationary distribution $\tilde{q}_*$, then from the ergodic theorem $\frac{1}{N}\sum_{j=1}^{N}\frac{\mathbb{1}_{g(t^{(j)})\leq 0}}{\pi(t^{(j)})}$ converges a.s. to α_{corr} as $N \to \infty$. A biased finite sample estimate of α_{corr} is then

$$\hat{\alpha_{\mathrm{corr}}} = \frac{1}{N_{\mathrm{corr}}}\sum_{j=b+1}^{N_{\mathrm{corr}}+b}\frac{\mathbb{1}_{g(t^{(j)})\leq 0}}{\pi(t^{(j)})} \tag{20}$$

where b is a burnin parameter. The standard squared error can be assessed as

$$\hat{\sigma}_{\mathrm{corr}}^2 = \frac{1}{N_{\mathrm{corr}}}\left(\frac{1}{N_{\mathrm{corr}}}\sum_{j=b+1}^{N_{\mathrm{corr}}+b} h(t^j) - \hat{\alpha_{\mathrm{corr}}}^2\right)(1+\hat{\gamma}_{\mathrm{corr}}) \tag{21}$$

where $h(t^j) = \frac{\mathbb{1}_{g(t^{(j)})\leq 0}}{\pi^2(t^{(j)})}$ and γ_{corr} is an estimator of $2\sum_{k=0}^{\infty}\mathrm{Corr}(h(t^{(0)}), h(t^{(k)}))$, which can be estimated as detailed

in Dubourg (2011). In the original paper (Dubourg et al., 2011), samples from the quasi-optimal density are obtained through a modified Metropolis-Hastings scheme (Au and Beck, 2001). 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.

2.2.4. *Efficient splitting estimator for the augmented failure probability*

While the estimation of the augmented failure probability $P_{f,\epsilon}$ only resorts to the Kriging predictor, its computation in a high dimensional through standard Monte Carlo simulation can still be improved. One idea due to Sudret et al. (2012) is to use a splitting estimator of $P_{f,\epsilon}$, much like in subset simulation which is base on the following equality:

$$P_{f,\epsilon} = P_{\epsilon,1} \prod_{k=2}^{s} P_{\epsilon,k|k-1} \tag{22}$$

where

$$P_{\epsilon,1} = \mathbb{E}_q(\pi_1(X)) \tag{23}$$

$$P_{\epsilon,k|k-1} = \mathbb{E}_{\tilde{q}_{k*}}\left[\frac{\pi_{(k)}(X)}{\pi_{(k-1)}(X)}\right] \tag{24}$$

where $\pi_{(k)}(x) = \Phi(-\frac{l_k - m_{\hat{G}}(x)}{\sigma_{\hat{G}}(x)})$ and $\tilde{q}_{k*}(x) \propto \pi_{(k-1)}(x)q(x)$ is the quasi-optimal density for a failure level l_{k-1}. The thresholds $l_1 > \ldots > l_s = 0$ should be such that the intermediate probabilities $P_{\epsilon,k|k-1}$ are easy to estimate (around 0.1-0.3 for instance). They can be fixed a priori or by running a subset simulation algorithms on the Kriging mean of the limit state function and using the corresponding intermediate levels as the l_k thresholds. Notice that $P_{\epsilon,1}$ can evaluated by standard Monte Carlo efficiently since it is a relatively large probability and q. Estimation of $P_{\epsilon,k|k-1}$ and its standard error $\sigma_{\epsilon,k|k-1}$ can be achieved by running a Monte Carlo Markov Chain targeting $\tilde{q}_{k*}$. Assuming $\hat{\delta}_{\epsilon,k} = \frac{\hat{\sigma}_{\epsilon,k|k-1}}{\hat{P}_{\epsilon,k|k-1}}$ is the coefficient of variation for the estimation of $P_{\epsilon,k|k-1}$, a conservative estimate of the overall c.o.v. of $P_{f,\epsilon}$ is

$$\hat{\delta}_\epsilon = \sum_{i=1}^{s} \delta_{\epsilon,k}^2 \tag{25}$$

2.2.5. *Kriging metamodel refinement*

Before computing both quantities $\hat{P}_{f,\epsilon}$ and α_{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 has to 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 (2009), Picheny et al. (2010) or Bect et al. (2012). The sign of the performance function G which delineates the failure domain is the most uncertain in this target region. It is usually determined by optimizing some *in-fill* criterion (Bect et al., 2012). However, the *in-fill* criterion might feature multiple optima or cause the optimizer to be stuck on a local plateau. Therefore, unless one uses a multi-start scheme (which wouldn't 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. (2011) 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, which acts as a surrogate to the function G. The idea is to cluster points generated from the density of the inputs given that they belong to the margin. In practice after generating a sufficiently large number of samples (through accept-reject for instance), k-means clustering could be used to obtain any given number of points. The Kriging metamodel enhancement proceeds in a similar fashion by drawing a large number of samples from the weighted margin probability density defined as

$$h(x) \propto w(x)q(x) \tag{26}$$

where

$$w(x) = \mathbb{P}_0(\hat{G}(x) \in [-t_{1-\alpha}\sigma_{\hat{G}}(x), t_{1-\alpha}\sigma_{\hat{G}}(x)]) \tag{27}$$

$$= \Phi\left(\frac{t_{1-\alpha}\sigma_{\hat{G}}(x) - m_{\hat{G}}(x)}{\sigma_{\hat{G}}(x)}\right) - \Phi\left(\frac{-t_{1-\alpha}\sigma_{\hat{G}}(x) - m_{\hat{G}}(x)}{\sigma_{\hat{G}}(x)}\right) \tag{28}$$

is the margin probability (Dubourg, 2011) and $t_{1-\alpha}$ is the $1-\alpha$ quantile of the standard normal pdf. 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) can provide a sufficiently large sample with stationary distribution h. The k-means algorithm then provides N_{add} clusters from these samples, which 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

$$\alpha_{\text{LOO}} = \frac{1}{N_D} \sum_{i=1}^{N_D} \frac{\mathbb{1}_{G(x_i)\leq 0}}{\pi_{-i}(x_i)} \tag{29}$$

where $\pi_{-i}(x) = \Phi\left(-\frac{\hat{m}_{G_{\mathcal{D}\setminus x_i}}(x_i)}{\hat{\sigma}_{G_{\mathcal{D}\setminus x_i}}(x_i)}\right)$ is the classification function obtained by removing observation x_i from the DoE $\mathcal{D}$. As stated in section 2.2.3, a factor 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's 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 amounts of l.s.f. 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 metamodel 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 Hessian matrix singularity during marginal likelihood optimization, especially when the kernel is parametrized by a scalar parameter for each input variable (e.g. the length-scale parameters for the squared-exponential (SE) kernel).
Attempts have been made recently to propose adequate representations for high dimensional Kriging. For instance, additive Kriging (Durrande et al., 2012) can be used by assuming an additive model for the emulator, i.e. that model is a sum of univariate metamodels. The resulting kernel function is simply the sum of the kernel submodels. 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. (67).

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, 2002). The idea of SDR is to find a low rank matrix B such that regression of the output Y on X can be performed by replacing X by $B^T X$. Without requirements on the rank r of matrix B, B always exists since $B = I_d$ does the trick. However, for many regression models, it is frequent for the output to be only a function of a linear combination of inputs as well as an optional noise independent of the inputs. For instance, if Y depends only on a subset of covariates, B is a sparse matrix with zero rows on the irrelevant predictors. Indeed, assuming without loss of generality that $X^T = (X_{\mathcal{A}}^T, X_{-\mathcal{A}}^T)$ where $X_{\mathcal{A}}$ (resp. $X_{-\mathcal{A}}$) is the subset of relevant (resp. irrelevant) components, then $B = \begin{pmatrix} I_{|\mathcal{A}|} \\ O_{d-|\mathcal{A}|\times|\mathcal{A}|} \end{pmatrix}$ satisfies the SDR requirements. More formally, $S(B) = \text{span}(B)$ is dimension reduction subspace if

$$Y \perp\!\!\!\perp X \mid B^T X \tag{30}$$

where $B \in \mathbb{R}^{d\times r}$ and $\perp\!\!\!\perp$ stands for independence. Equivalently, (30) could be rewritten as

$$Y = g_r(B^T X, \varepsilon) \tag{31}$$

where ε is a noise term independent of X and g_r an unknown function. An object of interest is the central subspace (CS) which is defined as the intersection of all dimension reduction subspace and therefore has minimal dimension: it exists under mild conditions (see (Cook, 1996). 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 $r \leq 10$ usual GP metamodels can then be applied to model the response $Y = g(X)$ 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

$$Y = g_r(B^T X) \tag{32}$$

In this case, the surrogate model is fitted to the link function g_r. To deal with the fact that model (32) does not take into account the noise in the original dimension reduction model (7), we suggest to use a nugget term in the corresponding metamodel. Let us stress that this is equivalent to Kriging with noisy observations when considering data points outside the DoE. However, for points in the experimental design, the Kriging metamodel with nugget effect interpolates the observations which apparently contradicts the original model (31). This shouldn't however be a major problem since the whole objective is to design an emulator that is close enough to the original limit state function. 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 (32) shouldn't 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.

3.2. SDR estimation methods

Let us get back at dimension reduction subspace estimation. The first arguably popular SDR method is sliced inverse regression (Li, 1991) which finds a basis of the central subspace under restrictive conditions on the marginal distributions of X. Although very efficient computationally, it is difficult to use in the context of metamodelling for reliability purposes. Indeed, we are interested in reducing the dimension for the regression of the output Y in a region that is large enough to part of the failure domain. Therefore, it is necessary to obtain a training set $\{x_{tr}^i\}_{i=1}^{N_{tr}}$ consisting in a random samples that fill the input space but also populate the failure domain. Because SIR constrains the training set to satisfy restrictive conditions on the marginal distributions of the training sample, it seems difficult to find such a sampling distribution that also explores efficiently both the safety and failure domains. Also, SIR tends to miss some SDR directions in some special cases (Cook, 2000). Other methods such as Sliced Average Variance Estimation (SAVE) (Cook, 2000) and contour regression (Li et al., 2005) have similar limitations.
More recently, kernel methods have lead to the *kernel dimension reduction*(KDR) algorithm (Fukumizu et al., 2009)

and its gradient based counterpart (gKDR, Fukumizu and Leng (2014)): these algorithms 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. These are therefore preferable methods to learn a dimension reduction subspace in very general settings.

3.3. Standard and gradient based kernel dimension reduction

3.3.1. Kernel dimension reduction

The basis of KDR is to express the conditional independence criterion (30) as an optimization problem that reaches its minimum at the 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 (X, Y) be a random variable on $\mathcal{X}\times\mathcal{Y}$ with distribution probability P_{XY} where $(\mathcal{X}, \mathcal{B}_\mathcal{X})$ and $(\mathcal{Y}, \mathcal{B}_\mathcal{Y})$ are measure 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 $\mathbb{E}(k_\mathcal{X}(X, X)) < \infty$ and $\mathbb{E}(k_\mathcal{Y}(Y, Y)) < \infty$. The conditional independence criterion of (30) in KDR relies on the cross-covariance operator $\Sigma_{YX} : \mathcal{H}_\mathcal{X} \mapsto \mathcal{H}_\mathcal{Y}$ defined as (Fukumizu et al., 2009)

$$\langle g, \Sigma_{YX} f\rangle_{\mathcal{H}_\mathcal{Y}} = \mathbb{E}\left(f(X) - \bar{f})(g(Y) - \bar{g})\right) \tag{33}$$

for all $f \in \mathcal{H}_\mathcal{X}$ and $g \in \mathcal{H}_\mathcal{Y}$. The covariance operator Σ_{XX} is defined similarly. The conditional covariance operator $\Sigma_{YY|X}$ is defined as

$$\Sigma_{YY|X} = \Sigma_{YY} - \Sigma_{YY}^{1/2} V_{YX} V_{XY} \Sigma_{YY}^{1/2} \tag{34}$$

where V_{YX}, V_{XY} are bounded operators such that $\Sigma_{YX} = \Sigma_{XX}^{1/2} V_{YX} \Sigma_{YY}^{1/2}$ and $\Sigma_{XY} = \Sigma_{YY}^{1/2} V_{YX} \Sigma_{XX}^{1/2}$ and the conditional cross-covariance is sometimes abusively expressed as $\Sigma_{YY|X} = \Sigma_{YY} - \Sigma_{YX}^{1/2}\Sigma_{XX}^{-1}\Sigma_{XY}^{1/2}$. Let $B \in \mathbb{R}^{d\times r}$ a matrix such that $\text{span}(B)$ is a dimension reduction subspace and $B^T B = I_r$. Letting k_d a positive kernel and $k_\mathcal{X}^B$ a positive definite kernel on $\mathbb{R}^r$ defined by $k_\mathcal{X}^B(x, x') = k_d(B^T x, B^T x')$, define the (cross-)covariance operators w.r.t. kernel $k_\mathcal{X}^B$ as Σ_{YX}^B, Σ_{XY}^B, Σ_{YY}^B and Σ_{XX}^B. The conditional cross covariance operator $\Sigma_{YY|X}^B$ associated to $k_\mathcal{X}^B$ is then defined as in (34). 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 conditional cross-covariance operators (Fukumizu et al., 2009).

$$\Sigma_{YY|X} = \Sigma_{YY|X}^B \Leftrightarrow Y \perp\!\!\!\perp X \mid B^T X \tag{35}$$

In addition to (35), we have $\Sigma_{YY|X}^B \geq \Sigma_{YY|X}$ (with the order of self-adjoint operators). Thus, in order to find the SDR matrix B, the KDR algorithm minimizes $\text{Tr}(\Sigma_{YY|X}^B)$ subject to $B^T B = I_r$. The practical algorithm takes as inputs n i.i.d. samples (x^i, y^i) from P_{XY} and solves

$$\min_{B^T B = I_r} \text{Tr}\left[G_Y (G_X^B + n\varepsilon_n I_n)^{-1}\right] \tag{36}$$

where G_Y and G_X are the centred Gram matrices defined by $G_X = HK_XH$, $G_Y = HK_YH$, $(K_X)_{ij} = k_\mathcal{X}(x^i, x^j)$, $(K_Y)_{ij} = k_\mathcal{X}(y^i, y^j)$, $1 \leq i, j \leq n$ 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. (2009) for more details).

3.3.2. 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 X and Y. However, the required non-convex optimization step makes it somewhat CPU intensive depending on the size of the training samples (x^i, y^i). Moreover, kernel parameter selection and the choice of the regularization term is needed and cross-validation for parameter selection is not possible because of the computing cost of the optimization procedure. The gKDR method (Fukumizu and Leng, 2014) solves these issues with a much faster algorithm for dimension reduction subspace identification. It relies on the fact that the derivative $\partial\mathbb{E}(Y \mid X = x)/\partial x$ is contained in the SDR subspace. 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 (2014).

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 in 1 based on the mean-squared error of the regression estimate of Y w.r.t. $B_r^T X$ where B_r is the dimension reduction matrix of rank r.

Algorithm 1 CV procedure for dimension selection

Require: samples (x^i, y^i), $i = 1, \ldots, n$, K_{cv}: number of CV folds, r_{max}: maximum SDR subspace dimension, kernels $k_\mathcal{X}$,$k_\mathcal{Y}$

Split dataset into K_{cv} folds of similar size $C_1, \ldots, C_{K_{cv}}$

for $r = 1$ to r_{max} **do**

  for $j = 1$ to K_{cv} **do**

  – Estimate $\hat{B}_r$ using (x^i, y^i), $i \notin C_j$

  – Estimate the mean-squared error $e_{r,j}^2$ of k-NN regression of Y on $\hat{B}_r^T X$ using validation fold (x^i, y^i), $i \in C_j$

  end for

  Estimate the regression error of Y on $B_r^T X$ by $e_r^2 = \frac{1}{K_{cv}} \sum_{j=1}^{K_{cv}} e_{r,j}^2$

end for

$\hat{r} \leftarrow \arg\min_{1\leq r\leq r_{max}} e_r^2$ is the SDR dimension estimate

Compute $\hat{B} = \hat{B}_{\hat{r}}$ using the whole dataset

As a matter of fact, algorithm 1 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 (2014)). 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 the faster algorithm 2.

Algorithm 2 Fast CV procedure for dimension selection

Require: samples (x^i, y^i), $i = 1, \ldots, n$, K_{cv}: number of CV folds, r_{max}: maximum SDR subspace dimension, kernels $k_{\mathcal{X}}$,$k_{\mathcal{Y}}$

Split dataset into K_{cv} folds of similar size $C_1, \ldots, C_{K_{cv}}$

for $j = 1$ to K_{cv} **do**

Compute $\hat{B}^{(j)} = \hat{B}^{(j)}_{r_{max}} \in \mathbb{R}^{d \times r_{max}}$ using (x^i, y^i), $i \notin C_j$

end for

for $r = 1$ to r_{max} **do**

for $j = 1$ to K_{cv} **do**

$\hat{B}^{(j)}_r \leftarrow \hat{B}^{(j)}_{:,1:r}$ matrix of first r columns

Estimate the mean-squared error $e^2_{r,j}$ of k-NN regression of Y on $\left(\hat{B}^{(j)}_r\right)^T X$ using validation fold (x^i, y^i), $i \in C_j$

end for

Estimate the regression error of Y on $B_r^T X$ by $e_r^2 = \frac{1}{K_{cv}} \sum_{j=1}^{K_{cv}} e^2_{r,j}$

end for

$\hat{r} \leftarrow \arg\min_{1 \leq r \leq r_{max}} e_r^2$ is the SDR dimension estimate

Compute $\hat{B} = \hat{B}_{\hat{r}}$ using the whole dataset

4. Importance sampling with dimension reduction subspace metamodel

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 32. The ensuing Kriging predictor is denoted $\hat{G}_r$. For reasons that will be clear in the sequel, we restrict ourselves to limit state functions that are defined for random inputs $X \in \mathbb{R}^d$.

4.1. Quasi-optimal importance density with reduced metamodel

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

$$\tilde{q}_{r*}(x) = \frac{\pi_r(B^T x) q(x)}{\int \pi_r(B^T x) q(x) \mathrm{d}x} \tag{37}$$

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

$$\pi_r(z) = \Phi\left(-\frac{m_{\hat{G}_r}(z)}{\sigma_{\hat{G}_r}(z)}\right) \tag{38}$$

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

$$P_{f,\epsilon} = \mathbb{E}_q(\pi_r(B^T X)) \tag{39}$$

$$\alpha_{\text{corr}} = \mathbb{E}_{\tilde{q}_{r*}}\left(\frac{\mathbb{1}_{g(X)\leq 0}}{\pi_r(B^T X)}\right) \tag{40}$$

4.2. Reduced Kriging metamodel refinements

Regarding the sampling based Kriging refinement strategy outlined in 2.2.5, a similar approach may readily be applied, that is we define the weighted margin probability density

$$h_r(x) \propto w_r(B^T x) q(x) \tag{41}$$

where w_r is defined as in (28). The refinement strategy using a reduced metamodel may now proceed as follows: given a sample x^i, $i = 1, \ldots, n$ drawn from a Markov chain targeting h_r, N_{add} cluster centers are determined via a 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 since the metamodel actually emulates the link function in the simplified regression model (32). The same remark can be made in the initial space filling design.

To avoid this, we suggest another strategy. The DoE for the reduced Kriging model is a set of input-evaluation pairs in $\mathbb{R}^r \times \mathbb{R}$ in which the evaluated function is the link function g_r of the SDR model. Because g_r is unknown, we should use the true limit state function g. To do so, a mapping from $\mathbb{R}^r$ to $\mathbb{R}^d$ is necessary. The simplified model states that $Y = g_r(B^T X)$. On the other hand $Y = g(X)$. Hence if $(z^*, y^* = g_r(z^*))$ is a point to be added to the DoE, letting $x^* = (B^T)^{\#} z^*$ where $B^T(B^T)^{\#} = I_d$ ensures that $y^* = g(x^*) = g_r(B^T(B^T)^{\#} z^*) = g_r(z^*)$. Therefore, (z^*, y^*) can be added to the design of experiments by evaluating $y^* = g((B^T)^{\#} z^*)$. In the case where the input domain $\mathbb{X}$ is not $\mathbb{R}^d$, there is no guarantee that $x^* = (B^T)^{\#} z^* \in \mathbb{X}$, therefore the following tweak can be adopted:

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

4.3. The case of Gaussian inputs

The augmented failure probability and correction factor (39) and (40) normally require ordinary Monte Carlo or MCMC sampling in a d-dimensional space. Admittedly, (39) doesn't 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 though approaches such as the modified Metropolis-Hastings algorithm (Au and Beck, 2001) may prove

successful.
Let us assume the input X to be a standard multivariate Gaussian variable. The estimation of (39) and (40) is made more efficient as shown below.

Lemma 1. (i) *The augmented failure probability $P_{f,\epsilon}$ can be expressed as $P_{f,\epsilon} = \mathbb{E}(\pi_r(Z))$ where 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 $W_2 \sim \mathcal{N}(0_{d-r\times 1}, I_{d-r})$, $W_1 \sim p_{W_1}(w_1) = \frac{\pi_r(w_1)\varphi_r(w_1)}{P_{f,\epsilon}}$ where φ_r is the standard r-dimensional multinormal pdf and $W = [W_1^T, W_2^T]^T$. Then $\tilde{X} = B_a^{-T} W$ is distributed according to $\tilde{q}_{r*}(x)$.*

The first result regarding the augmented failure probability computation is a straightforward consequence of the fact that $Z = B^T X \sim \mathcal{N}(0, I_r)$ since $B^T B = I_r$. The proof of (ii) is deferred to the appendix.
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 unnormalized distribution $w_r(w_1)\varphi_r(w_1)$.

4.4. The general case

Assuming Gaussian inputs yields a simple importance sampling scheme after linear dimension reduction. 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 isoprobabilistic transform $T : X \in \mathbb{X} \mapsto T(X) = U \in \mathbb{U}$ so that $g_0(U) = g\left(T^{-1}(U)\right)$. In this paper, T we consider the Nataf transform defined as $T = T_2 \circ T_1$, where

$$T_1 : X \mapsto Z = T_1(X) = \begin{pmatrix} \Phi^{-1}(F_1(X_1)) \\ \vdots \\ \Phi^{-1}(F_d(X_d)) \end{pmatrix} \tag{42}$$

$T_2 : Z \mapsto T_2(Z) = L_0^{-1} Z$ where $L_0 L_0^T = R_0$ is the linear correlation matrix of Z, Φ is the standard normal cdf and $F_1, \ldots, F_d$ are the marginal CDFs of X.
One could simply try dimension reduction on the response model $Y = g_0(U)$, but because of the non-linear nature of the Nataf transform, there is no guarantee that an SDR type hypothesis of the form $Y \perp\!\!\!\perp U \mid \tilde{B}^T U$ holds for some $\tilde{B}$ even if $Y \perp\!\!\!\perp X \mid B^T X$. Nevertheless, the $\mathbb{U}$ space formulation being convenient for practical reliability analysis, to achieve dimension reduction, we suggest to simply backtransform to the $\mathbb{X}$ space and perform linear dimension reduction. A quasi-optimal importance density in the $\mathbb{U}$ space is then obtained as

$$\tilde{q}_{r*}(u) = \frac{\pi_r(B^T T^{-1}(u))\varphi_d(u)}{P_{f,\epsilon}} \tag{43}$$

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

$$\alpha_{\mathrm{corr}} = \mathbb{E}_{\tilde{q}_{r*}}\left(\frac{\mathbb{1}_{g_0(U)\leq 0}}{\pi_r(B^T T^{-1}(U))}\right) \tag{44}$$

Now, because of the non-linear transformation, drawing samples from (43) is not as simple as in the case of Gaussian inputs X. A chain targeting $\tilde{q}_{r*}$ has to be simulated in $\mathbb{R}^d$ either via relevant MCMC method. For this purpose, we suggest using a modified Metropolis-Hastings algorithm (Au and Beck, 2001). As far as the sampling based refinement criterion is concerned, we redefine the weighted margin probability in a similar fashion, that is

$$h_r(u) = w_r(B^T T^{-1}(u))\varphi_d(u) \tag{45}$$

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 (2011). Once samples are obtained, we backtransform them to the physical input domain $\mathbb{X}$ and proceed as outlined in 2.2.5.

4.5. Alternative probabilistic classification function

A common challenge in estimating the correction factor is that the quasi-optimal density $\tilde{q}_{r*}$ may be multimodal in which case a standard Metropolis-Hastings scheme may be trapped in a mode and fail to provide an accurate sample from the target density. One way to deal with this is to use an ad-hoc MCMC scheme such as population MCMC (Liu, 2001). Instead of changing the sampling scheme, we suggest modifying the quasi-optimal density by altering the probabilistic classification function π_r. The modification consists in replacing the standard normal cdf with a logistic function L in the definition of π_r:

$$\pi_r(z) = L\left(-\frac{m_{\hat{G}_r}(z)}{\sigma_{\hat{G}_r}(z)}\right) \tag{46}$$

where $L(x) = \frac{1}{1+\exp(-x)}$. With this definition, π_r no longer corresponds to the probability that the Kriging predictor falls into the failure domain, however it is a reasonable approximation. The main advantage is that L decreases less slowly to zero as $x \to \infty$, hence the resulting importance density $\tilde{q}_{r*}$ has modes that are separated by valleys where the density achieves larger values. For this choice of IS density, the standard MH scheme has better mixing properties and thus provides more reliable estimates of the correction factor in multimodal settings.

4.6. The Meta-ISDR algorithm

We hereafter sketch the Meta-ISDR algorithm, assuming 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 in brackets are:

- $x_1^*, \ldots, x_m^* : m \geq 1$ 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
- $\alpha_{\mathrm{LOO}}^{\min}(0.1)$, $\alpha_{\mathrm{LOO}}^{\max}(10)$: 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: target coefficient of variation (c.o.v.) for $\hat{P}_{f,\epsilon}$
- δ_α: target coefficient of variation (c.o.v.) for 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 $\mathbb{P}(\|X\|^2 > R_{max}^2)$ is much smaller than P_f. For instance if $P_f >> 10^{-b}$ for some positive integer b then solving $\mathbb{P}(\|X\|^2 > R_{max}^2) = 10^{-b}$ 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 table 3.

5. Numerical illustrations

We now investigate the performance of the metamodel-based importance sampling procedure with dimension reduction on a simple analytical setting and a realistic high dimensional industrial application consisting in estimating the failure probability of a floating wind turbine under stationary wind and wave loads.

5.1. Academic example

We consider an analytical limit state function whose dimensionality can be varied to illustrate the impact of increasing number of inputs. The performance function reads

$$g(X) = d + a\sigma\sqrt{d} - \sum_{i=1}^{d} X_i \tag{47}$$

where the X_i are d i.i.d. lognormal variables with mean 1 and standard deviation $\sigma = 0.2$ and $a = 1$. We consider 3 cases with increasing dimension, $d = 2, 50, 100$. It is obvious that a sufficient dimension reduction subspace is spanned by $b = \left(1 \cdots 1\right)^T$ since $g(X) = d + a\sigma\sqrt{d} - b^T X$ so we can expect sizeable performance improvements for high dimensions for the Meta-ISDR method. The 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 (IS-FORM) and the standard Meta-IS. 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 (1991).

Method	**FORM**	**Monte Carlo**	**IS-FORM**	**MetaIS**	**MetaIS-DR**
d = 2					
N	18	5×10^5	8818	103	518
$\hat{P}_f$	3.81×10^{-3}	4.93×10^{-3}	4.90×10^{-3}	4.88×10^{-3}	4.94×10^{-3}
c.o.v.	-	2%	2%	1.3%	2.1%
d = 50					
N	255	1.4×10^6	14000	3255	1026
$\hat{P}_f$	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	18000	5505	1783
$\hat{P}_f$	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 (47)

As can be expected for $d = 2$, the dimension reduction is not worth it since it comes at an additional cost of limit state function evaluations: from table 5.1 the total calls to g for the MetaIS-SDR is 5 times that of standard MetaIS. 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 one-dimensional SDR subspace leading to an importance sampling density closer to the optimal one hence achieving lower estimation variance.

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 analysis to the wind

Algorithm 3 Meta-ISDR algorithm for standard Gaussian inputs

SDR Basis estimation:

Sample $x^i_{SDR} \sim \mathcal{N}(0, I_d)$, $i = 1, \dots, N_{SDR}$ and compute $y^i_{SDR} = g(x^i_{SDR})$

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

Initial DoE:

Add design points $x^*_1, \dots, x^*_m$ to DoE (optional)

Space-filling design: sample N_{fill} points in the hypersphere of radius R_{max} and add $N^{init}_D - m$ corresponding cluster centers x^{init}_j to initial DoE by evaluating $y^{init}_j = g(x^{init}_j)$, $j = 1, \dots, N^{init}_D - m$

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

Metamodel refinement:

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

  sample $n_1 \in [10^4, 5 \times 10^4]$ points z^i, $i = 1, \dots, n_1$ from h_r (defined in (41)) using procedure from lemma 1.

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

  if $N_D > N^{\max}_D$ **then break**

  end if

end while

Failure probability estimation

Augmented probability estimation: compute $\hat{P}_{f,\epsilon}, \hat{\delta}_\epsilon$ using standard Monte Carlo or the splitting algorithm in 2.2.4

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 $\hat{\delta}_\alpha = \frac{\hat{\sigma}_{\text{corr}}}{\hat{\alpha_{\text{corr}}}}$ according to (20) and (21)

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

inflow is obtained using the FAST software which provides extreme and fatigue loads for a 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 towards 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.

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 (Shinozuka and Deodatis, 1991)

$$X(t) = U_{10} + \sum_{i=1}^{n} (u_i \sigma_i \cos(\omega_i t) - \bar{u}_i \sigma_i \sin(\omega_i t)) \tag{48}$$

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 $\sigma^2_i = S(\omega)\mathrm{d}\omega_i$ where $S(\omega)$ 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, we extracted 10-minute wind speed time series corresponding to a turbine rated wind speed of $U_{10} = 11.5$ m/s and turbulence intensity $I = 6\%$. These measurements where broken down into 5 minutes segments. A p.s.d. estimate was then obtained by averaging the periodograms of each segment. The resulting spectrum is shown in figure 1. For simulation purposes, the spectrum is discretized into $n = 50$ harmonics so that the random vector $U = [u_1, \dots, u_n, \bar{u}_1, \dots, \bar{u}_n]$ describing X is 100-dimensional.

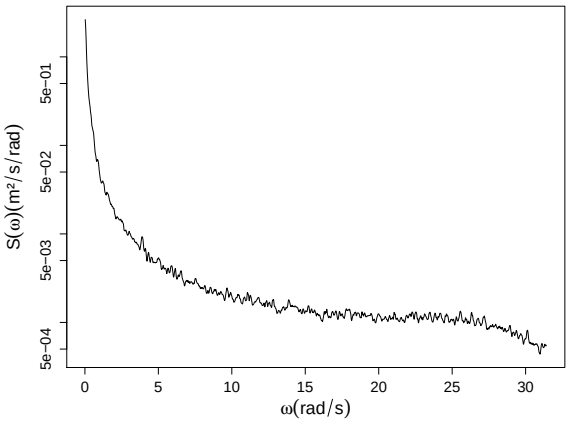

Figure 1: One-sided power spectral density of wind speed process

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 $\{Y(t), t \geq 0\}$ the mechanical response, we seek the following quantity

$$P_f(T) = \mathbb{P}(\exists t \in [0, T] \mid Y(t) \geq s) \tag{49}$$

where $T = 10$ minutes. Equivalently, we may introduce a time-varying limit state function $g(t, U) = s - Y(t)$ where 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 that a related quantity is the expected number of downcrossings $\mathbb{E}(N^+(0, T))$ (of level 0 by the limit state g) which is such that

$$P_f(T) \leq \mathbb{P}(g(0, U) \leq 0) + \mathbb{E}(N^+(0, T)) \tag{50}$$

Furthermore, $\mathbb{E}(N^+(0, T)) = \int \nu^+(t)\mathrm{d}t$, where

$$\nu^+(t) = \lim_{\Delta t \to 0} \frac{\mathbb{P}\left(g(t, U) > 0, g(t + \Delta t, U) \leq 0\right)}{\Delta t} \tag{51}$$

is the outcrossing rate. This 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. (2011). 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

$$\hat{\nu}^+ = \frac{\mathbb{P}\left(g(t_0, U) > 0, g(t_0 + \Delta t, U) \leq 0\right)}{\Delta t} \tag{52}$$

In our settings, we consider a threshold on the tower-top displacement $s = 0.4$ m and $\Delta t = 0.01$. The simulation length is one minute.

5.2.3. Instantaneous failure probability

To illustrate the performance of the MetaIS-DR algorithm with respect to other simulation-based reliability algorithms, we first look at the instantaneous failure probability at time $t_0 = 60$ s. 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, 2001):

- We considered a multiple failure point (MFP) search based on the approach of Der Kiureghian and Dakessian (1998): 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, 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.

- 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

- The subset simulation algorithm was ran using the original algorithm by Au and Beck (2001). 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:

- a maximum dimension of the SDR subspace $r_{max} = 6$

- 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 4.6).

- a Kriging model with constant mean β and squared exponential covariance function

- initial DoE of size $N_D^{init} = 10r + M$: Latin Hypercube Sampling of size $10 \times r + m$ "design points", where r is the estimated reduced dimension

- minimum/maximum size of the DoE $N_D^{\min} = 60$, $N_D^{\max} = 300$

- number of points added to DoE during refinement $N_{add} = 2r$

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.

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 MetaIS-DR algorithm. The results in table 5.2.3 indicate the good performance of the MetaIS-DR algorithm as it achieves both the target coefficient of variation while also requiring the least amount of

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

Table 2: Instantaneous Failure probability estimation: method comparison

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

Table 3: Short-term outcrossing rate estimation

limit state function evaluations. 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 isn't 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^+. To evaluate this quantity, let us rewrite (52) in a more convenient fashion

$$\begin{aligned}\hat{\nu}^+ &= \frac{\mathbb{P}\left(g(t_0,U)>0, g(t_0+\Delta t,U)\leq 0\right)}{\Delta t}\\ &= \frac{\mathbb{P}\left[\min\left(-g(t_0,U), g(t_0+\Delta t,U)\right)\leq 0\right]}{\Delta t}\\ &- \mathbb{P}\left[g(t_0,U)=0, g(t_0+\Delta t,U)\leq 0\right]\\ &= \frac{\mathbb{P}\left[\min\left(-g(t_0,U), g(t_0+\Delta t,U)\right)\leq 0\right]}{\Delta t}\\ &= \frac{\mathbb{P}\left(H_{t_0,t_0+\Delta t}(U)\leq 0\right)}{\Delta t}\end{aligned} \tag{53}$$

where $H_{t_0,t_0+\Delta t}(U) = \min\left(-g(t_0,U), g(t_0+\Delta t,U)\right)$ In (53), we implicitly assume that the response process $\{Y(t), t \geq 0\}$ and therefore the limit state $g(t_0,U)$, admit a density leading to $\mathbb{P}\left(g(t_0,U)=0, g(t_0+\Delta t,U)\leq 0\right) = 0$. The formulation (53) 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. (2005) for i.i.d. standard Gaussian inputs u_i:

$$\nu^+(t_0) = \frac{1}{2\pi\beta} e^{-\frac{1}{2}\beta^2} \sqrt{\sum_{i=1}^{n}\left(u_i^{*2}(t_0) + \bar{u}_i^{*2}(t_0)\right)\omega_i^2} \tag{54}$$

where $\beta = \|u^*(t_0)\|$ is the reliability index and $u^*(t_0)$ is the design point of the limit state function $g_(t_0,\cdot)$.

Surprisingly, the FORM approximation for this case study is more accurate than the corresponding instantaneous failure probability computed in the previous subsection and is comparable to the estimation of MetaIS-DR. We however expect the FORM outcrossing rate estimation to be inaccurate for more pronounced non-linearities.

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 demands 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 well known academic example which illustrates the impact on 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 (subset simulation, standard importance sampling and standard MetaIS).

Appendix A. Quasi-optimal density sampling in Meta-ISDR

We provide the proof of lemma 1. Let $B \in \mathbb{R}^{d\times r}$ a matrix such that $B^TB = I_r$ and span(B) is a dimension reduction subspace for the regression of Y on X . Let $B_a = [B, B_\perp] \in \mathbb{R}^{d\times d}$ where the columns of $B_\perp$ form an orthonormal basis of span$(B)^\perp$. Now assume $X \sim \mathcal{N}(0, I_d)$.

X has density $q(x) = \varphi_d(x) = (2\pi)^{-\frac{d}{2}} \exp(-\frac{\|x\|^2}{2})$ where φ_d is the d dimensional multinormal pdf. If x_B and $x_{B_\perp}$ are the orthogonal projections of x on span(B) and span(B) respectively, then $x = x_B + x_{B_\perp} = B(B^Tx) + B_\perp(B_\perp^Tx)$ and $\|x\|^2 = \|x_B\|^2 + \|x_{B_\perp}\|^2 = \|B^Tx\|^2 + \|B_\perp^Tx\|^2$. Hence,

$$\begin{aligned}q(x) &= (2\pi)^{-\frac{d}{2}} \exp(-\frac{\|x_B\|^2 + \|x_{B_\perp}\|^2}{2})\\ &= (2\pi)^{-\frac{r}{2}} \exp(-\frac{\|B^Tx\|^2}{2})(2\pi)^{-\frac{d-r}{2}} \exp(-\frac{\|B_\perp^Tx\|^2}{2})\\ &= \varphi_r(B^Tx)\varphi_{d-r}(B_\perp^Tx)\end{aligned}$$

and the quasi-optimal density is equal to

$$\tilde{q}_{r*}(x) = \frac{\pi_r(B^T x)\varphi_r(B^T x)\varphi_{d-r}(B_\perp^T x)}{P_{f,\epsilon}}$$

Now let $\tilde{X} \sim \tilde{q}_{r*}$ and consider the mapping $\tilde{X} \mapsto W = B_a^T \tilde{X}$. By a change of variable, for any continuous bounded function ψ from $\mathbb{R}^d$ to $\mathbb{R}^d$

$$\begin{aligned}\mathbb{E}(\psi(W)) &= \int_{\mathbb{R}^d} \psi(B_a^T \tilde{x})\tilde{q}_{r*}(\tilde{x})\mathrm{d}\tilde{x} \\ &= \underbrace{\frac{1}{|\det(B_a^T)|}}_{=1} \int_{\mathbb{R}^d} \psi(w)\tilde{q}_{r*}(B_a^{-T} w)\mathrm{d}w\end{aligned}$$

which implies that the density of W is

$$p_W(w) = \tilde{q}_{r*}(B_a^{-T} w) \tag{A.1}$$

$$= \frac{1}{P_{f,\epsilon}}\pi_r(B^T B_a^{-T} w)\varphi_r(B^T B_a^{-T} w)\varphi_{d-r}(B_\perp^T B_a^{-T} w) \tag{A.2}$$

$$= \frac{1}{P_{f,\epsilon}}\pi_r(w_1)\varphi_r(w_1)\varphi_{d-r}(w_2) \tag{A.3}$$

where $w^T = (w_1^T, w_2^T)$. By the same argument, if W has density $p_W(w) = \frac{1}{P_{f,\epsilon}}\pi_r(w_1)\varphi_r(w_1)\varphi_{d-r}(w_2)$ then $\tilde{X} = B_a^{-T} W \sim \tilde{q}_{r*}$. It is clear from (A.3) that to sample $W = (W_1^T, W_2^T)^T$ from p_W, suffice it to sample $W_2 \sim \mathcal{N}(0_{d-r\times 1}, I_{d-r})$, and $W_1 \sim p_{W_1}(w_1) = \frac{\pi_r(w_1)\varphi_r(w_1)}{P_{f,\epsilon}}$.

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

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

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

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

Chiaromonte, F., Cook, R., 2002. Sufficient dimension reduction and graphics in regression. Annals of the Institute of Statistical Mathematics 54 (4), 768–795.

Cook, R. D., 1996. Regression Graphics. Wiley.

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

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

Ditlevsen, O., Madsen, H., 1996. Structural Reliability Methods. John Wiley & Sons Inc.

Dubourg, V., 2011. Adaptive surrogate models for reliability analysis and reliability-based design optimization. Ph.D. thesis, Blaise Pascal University - Clermont II.

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

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

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

Echard, B., Gayton, N., Lemaire, M., Relun, N., 2013. 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.

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

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

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

Jensen, J., Olsen, A., Mansour, A., 2011. Extreme wave and wind predictions. Ocean Engineering 38, 2244–2253.

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

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

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

Liu, J., 2001. Monte Carlo Strategies for Scientific Computing. Springer.

Liu, P.-L., Der Kiureghian, A., 1991. Optimization algorithms for structural reliability. Structural Safety 9 (3), 161 – 177.

Melchers, R., 1989. Importance sampling in structural systems. Structural Safety 6, 3–10.

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

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

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

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

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

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