An active-learning algorithm that combines sparse polynomial chaos expansions and bootstrap for structural reliability analysis

S. Marelli and B. Sudret

$Y = \sum_{\alpha \in A} y_\alpha \Psi_\alpha(X)$
$\int \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt$
$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp(-t^2/2)\, dt$
$\Phi\left(\frac{x-\mu}{\sigma}\right)$
$\beta_{HL} = \frac{\mu_R - \mu_S}{\sigma_R^2 + \sigma_S^2}$
ELEVATION

Chair of Risk, Safety and Uncertainty Quantification
Stefano-Franscini-Platz 5
CH-8093 Zürich

Data Sheet

Journal:	Structural Safety
Report Ref.:	RSUQ-2017-009
Arxiv Ref.:	`http://arxiv.org/abs/1709.01589` - [stat.CO]
DOI:	`https://doi.org/10.1016/j.strusafe.2018.06.003`
Date submitted:	August 24, 2017
Date accepted:	June 15, 2018

An active-learning algorithm that combines sparse polynomial chaos expansions and bootstrap for structural reliability analysis

S. Marelli and B. Sudret

August 9, 2018

Abstract

Polynomial chaos expansions (PCE) have seen widespread use in the context of uncertainty quantification. However, their application to structural reliability problems has been hindered by the limited performance of PCE in the tails of the model response and due to the lack of local metamodel error estimates. We propose a new method to provide local metamodel error estimates based on bootstrap resampling and sparse PCE. An initial experimental design is iteratively updated based on the current estimation of the limit-state surface in an active learning algorithm. The greedy algorithm uses the bootstrap-based local error estimates for the polynomial chaos predictor to identify the best candidate set of points to enrich the experimental design. We demonstrate the effectiveness of this approach on a well-known analytical benchmark representing a series system, on a truss structure and on a complex realistic frame structure problem.

Keywords: Polynomial Chaos Expansions, Adaptive Designs, Bootstrap, Structural Reliability, Active Learning

1 Introduction

Structural reliability analysis aims at computing the probability of failure of a system with respect to some performance criterion in the presence of uncertainty in its structural and operating parameters. Such uncertainty can be modelled by a random vector $\boldsymbol{X} \in \mathbb{R}^M$ with prescribed joint probability density function $f_{\boldsymbol{X}}$. The limit-state function g is defined over the support of $\boldsymbol{X}$ such that $\{\boldsymbol{x} : g(\boldsymbol{x}) \leq 0\}$ defines the failure domain, while $\{\boldsymbol{x} : g(\boldsymbol{x}) > 0\}$ defines the safe domain. The *limit state surface* implicitly defined by $g(\boldsymbol{x}) = 0$ lies at the boundary between the two domains. The probability of failure of such a system can be defined as (Melchers, 1999; Lemaire, 2009):

$$P_F = \int_{\{\boldsymbol{x}:g(\boldsymbol{x})\leq 0\}} f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x}. \tag{1}$$

A straightforward approach to compute the integral in Eq. (1) is to use of Monte Carlo Simulation (MCS). However, standard MCS approaches can often not be used in the presence of complex and computationally expensive engineering models, because of the large number of samples they require to estimate small probabilities (typically in the order of $\sim 10^{k+2}$ for $P_F \approx 10^{-k}$) with acceptable accuracy. Well-known methods based on local approximation of the limit-state function close to the failure domain (such as FORM (Hasofer and Lind, 1974) and SORM (Rackwitz and Fiessler, 1978)) can be more efficient, yet they are usually based on linearisation and tend to fail in real-case scenarios with highly non-linear structural models.

In contrast, methods based on surrogate modelling have gradually gained momentum in the last few years. Due to the nature of the problem of estimating low probabilities, most recent methods combine active-learning-based greedy algorithms with Gaussian process surrogate models (Kriging). Among the first works to propose this approach, the earliest applications in this context were the efficient global reliability analysis method (EGRA) by Bichon et al. (2008, 2011), and the active-learning reliability (AK-MCS) method based on Kriging by Echard et al. (2011). More recently, Kriging has been employed to devise quasi-optimal importance densities in Dubourg et al. (2013); Dubourg and Sudret (2014). Amongst other variations, polynomial-chaos-based Kriging has also been used as an alternative metamodelling technique (Schöbi et al., 2016) to overcome some of the limitations of pure Kriging-based methods. Additional works on the topic of Kriging and structural reliability can be found, including extensions of the original AK-MCS algorithm to more advanced sampling techniques (Echard et al., 2013; Balesdent et al., 2013), system reliability (Fauriat and Gayton, 2014) and for the exploration of multiple-failure regions (Cadini et al., 2014).

Polynomial chaos expansions (PCE) (Ghanem and Spanos, 1991) are a well-established tool in the context of uncertainty quantification, with applications in uncertainty propagation (Xiu and Karniadakis, 2002), sensitivity analysis (Le Gratiet et al., 2017) and, to a lesser degree, structural reliability (Sudret and Der Kiureghian, 2002). While often considered as an efficient surrogate modelling technique due to their global convergence behaviour, PCEs have been employed only seldom in reliability analysis (see, *e.g.* Notin et al. (2010)) due to their lack of accuracy in the tails of the model response distribution, which are essential in this field.

In addition, most active-learning approaches with surrogates require some form of local error estimate to adaptively enrich a small set of model evaluations close to the limit state surface. Kriging-based methods can rely on the Kriging variance for this task, but PCEs do not provide a natural equivalent.

In this paper, we leverage on the properties of regression-based sparse-PCE (Blatman and Sudret, 2011) to derive a local error estimator based on *bootstrap resampling.* We then use this estimator to construct an active-learning strategy that adaptively approximates the limit-state function with PCE by minimizing a *misclassification probability*-based learning function at every iteration. The method is then showcased on a standard benchmark functions representing a series system and on a realistic structural frame engineering example.

2 Methodology

2.1 Polynomial Chaos Expansions

Consider a finite variance model $Y = \mathcal{M}(\boldsymbol{X})$ representing the response of some quantity of interest (QoI) Y to the random input parameters $\boldsymbol{X} \in \mathbb{R}^M$, modelled by a joint probability distribution function (PDF) $f_{\boldsymbol{X}}$. Also consider the functional inner product defined by:

$$\langle g, h \rangle \equiv \int_{\boldsymbol{x} \in \Omega_{\boldsymbol{X}}} g(\boldsymbol{x}) h(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x} = \mathbb{E}\left[g(\boldsymbol{X}) h(\boldsymbol{X})\right] \tag{2}$$

where $\Omega_{\boldsymbol{X}}$ represents the input domain. Under the assumption of independence of the input variables, that is $f_{\boldsymbol{X}}(\boldsymbol{x}) = \prod\limits_{i=1}^{M} f_{X_i}(x_i)$, one can represent $\mathcal{M}(\boldsymbol{X})$ as the following *generalised polynomial chaos expansion* (see, *e.g.* Ghanem and Spanos (1991); Xiu and Karniadakis (2002)):

$$Y = \mathcal{M}(\boldsymbol{X}) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^M} y_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{X}), \tag{3}$$

where the $y_{\boldsymbol{\alpha}}$ are real coefficients and $\boldsymbol{\alpha}$ is a multi-index that identifies the degree of the multivariate polynomial $\Psi_{\boldsymbol{\alpha}}$ in each of the input variables X_i:

$$\Psi_{\boldsymbol{\alpha}} = \prod_{i=1}^{M} \phi_{\alpha_i}^{(i)}(X_i). \tag{4}$$

Here $\phi_{\alpha_i}^{(i)}$ is a polynomial of degree α_i that belongs to the family of orthogonal polynomials w.r.t. the marginal PDF f_{X_i}. For more details on the construction of such polynomials for both standard and arbitrary distributions, the reader is referred to Xiu and Karniadakis (2002).

In the presence of a complex dependence structure between the input variables, it is always possible to construct isoprobabilistic transforms (*e.g.* Rosenblatt or Nataf transforms, see *e.g.* Lebrun and Dutfoy (2009)) to decorrelate the input variables prior to the expansion, even in the case of complex dependence modelled by vine copulas (Torre et al., 2017). For the sake of notational simplicity and without loss of generality, we will hereafter assume independent input variables.

In practical applications, the series expansion in Eq. (3) is traditionally truncated based on the maximal degree p of the expansion, thus yielding a set of basis elements identified by the multi-indices $\boldsymbol{\alpha} \in \mathcal{A} : \sum\limits_{i=1}^{M} \alpha_i \leq p$, with $\text{card}(\mathcal{A}) \equiv P = \binom{M+p}{p}$, or using more advanced truncation schemes that favour sparsity, *e.g.* hyperbolic truncation (Blatman and Sudret, 2010a). The corresponding expansion coefficients $\boldsymbol{y}_{\boldsymbol{\alpha}}$ can then be calculated efficiently via least-square analysis based on an existing sample of the input random vector

$\mathcal{X} = \left\{\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(N)}\right\}$, known as the *experimental design* (ED), and the corresponding model responses $\mathcal{Y} = \left\{y^{(1)}, \cdots, y^{(N)}\right\}$ as follows:

$$\boldsymbol{y}_{\boldsymbol{\alpha}} = \operatorname{argmin} \frac{1}{N} \sum_{i=1}^{N} \left[y^{(i)} - \sum_{\boldsymbol{\alpha} \in \mathcal{A}} y_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\boldsymbol{x}^{(i)}) \right]^2 . \tag{5}$$

When the number of unknown coefficients P is high (*e.g.* for high-dimensional inputs or high-degree expansions), regression strategies that favour sparsity are needed to avoid overfitting in the presence of a limited-size experimental design and to make the analysis at all feasible with a reasonable sample size N. Amongst them, *least angle regression* (LARS, Efron et al. (2004)), based on a regularized version of Eq. (5), has proven to be very effective in tackling realistic engineering problems even in relatively high dimensions (*i.e.* $M \sim 100$). In this paper, we adopt the full degree-adaptive, sparse PCE based on hybrid-LARS introduced in Blatman and Sudret (2011)), as implemented in the UQLab Matlab software ((Marelli and Sudret, 2014, 2017)).

2.2 Bootstrap-based local error estimation in PCE

2.2.1 Bootstrap in least-square regression

Adopting a least-square regression strategy to calculate the coefficients in Eq. (5) allows one to use the *bootstrap* resampling method (Efron, 1982) to obtain information on the variability in the estimated coefficients due to the finite size of the experimental design. Suppose that a set of estimators $\boldsymbol{\theta}$ is a function of a finite-size sample $\mathcal{X} = \left\{\boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(N)}\right\}$ drawn from the random vector $\boldsymbol{X}$. Then the *bootstrap* method consists in drawing B new sample sets $\left\{\mathcal{X}^{(1)}, \cdots, \mathcal{X}^{(B)}\right\}$ from the original $\mathcal{X}$ by *resampling with substitution.* This is achieved by randomly assembling $B-$times N realizations $\boldsymbol{x}^{(i)} \in \mathcal{X}$, possibly including repeatedly the same realization multiple times within each sample. The set of estimated quantities can then be re-calculated from each of the B samples, thus yielding a set of estimators $\boldsymbol{\Theta} = \left\{\boldsymbol{\theta}^{(1)}, \cdots, \theta^{(B)}\right\}$. This set of estimators can then be used to directly assess the variability of $\boldsymbol{\theta}$ due to the finite size of the experimental design $\mathcal{X}$, at no additional costs, *e.g.* by calculating statistics, or directly using each realization separately. Application of the bootstrap method combined with PCE to provide confidence bounds in the estimated P_F in structural reliability applications can be found in *e.g.* Notin et al. (2010); Picheny et al. (2010).

2.2.2 Bootstrap-PCE

We propose to use the bootstrap technique to provide local error estimates to the PCE predictions. The rationale is the following: the PCE coefficients $\boldsymbol{y}_{\boldsymbol{\alpha}}$ in Eq. (5) are estimated from the experimental design $\mathcal{X}$, therefore they can be resampled through bootstrap. This can be achieved by first generating a set of bootstrap-resampled experimental designs

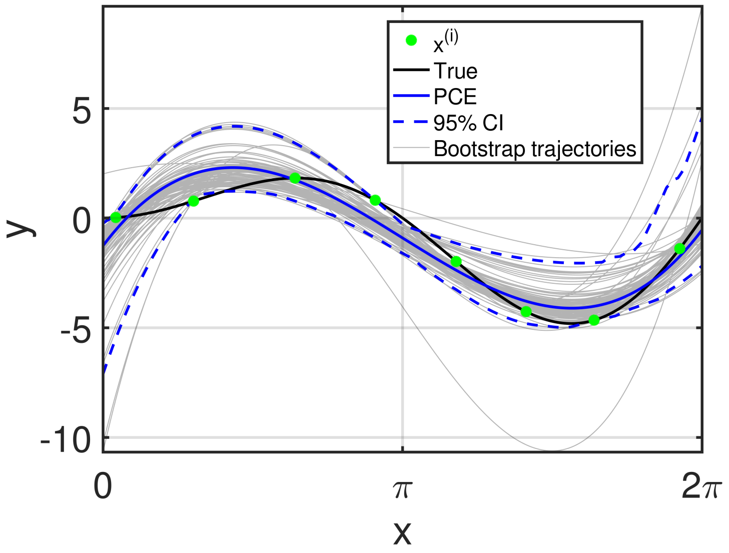


Figure 1: Bootstrap-resampled trajectories ($B = 100$) of the simple 1D analytical function in Eq. (7). The black line represents the true model, sampled at the 8 experimental design points $\boldsymbol{x}^{(i)}$ (green dots). The PCE surrogate is represented by the blue line, while the bootstrap trajectories are given by the gray lines. The corresponding 95% empirical inter quantile-range is given by the dashed blue lines.

$\left\{\mathcal{X}^{(b)}, \mathcal{Y}^{(b)}, b = 1, \cdots, B\right\}$. For each of the generated designs, one can calculate a corresponding set of coefficients $\boldsymbol{y}_{\boldsymbol{\alpha}}^{(b)}$, effectively resulting in a set of B different PCEs. Correspondingly, the response of each PCE can be evaluated at a point $\boldsymbol{x}$ as follows:

$$Y_{PC}^{(b)}(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} y_{\boldsymbol{\alpha}}^{(b)} \, \Psi_{\boldsymbol{\alpha}}(\boldsymbol{x}), \tag{6}$$

thus yielding a full response sample at each point $\left\{Y_{PC}^{(b)}(\boldsymbol{x}), b = 1, \cdots, B\right\}$. Therefore, empirical quantiles can be employed to provide local error bounds on the PCE prediction at each point, as well as to any derived quantity (*e.g.* P_F or sensitivity indices, see *e.g.* Picheny et al. (2010); Dubreuil et al. (2014)).

This bootstrap-resampling strategy in Eq. (6) yields in fact a family of B surrogate models that can be interpreted as *trajectories*. Figure 1 showcases how such trajectories can be directly employed to assess confidence bounds on point-wise predictions on a simple 1D test function given by:

$$f(x) = x \sin(x), \qquad x \in [0, 2\pi], \tag{7}$$

where the single random variable is assumed to be uniformly distributed within the bounds $X \sim \mathcal{U}(0, 2\pi)$, and where $B = 100$ bootstrap samples have been used.

This process of bootstrap-based trajectory resampling to provide better estimates of point-wise confidence bounds has been recently explored in the Gaussian process modelling literature, see *e.g.*, den Hertog et al. (2006); van Beers and Kleijnen (2008).

We refer to this approach as to *bootstrap-PCE*, or bPCE in short.

2.2.3 Fast bPCE

Because the training of a PCE model with sparse least-square analysis may be time consuming, especially in high dimension and/or when an already large experimental design is available (*i.e.* $N \sim 10^3$), and because in this particular application we do not need very accurate estimates on the bounds of the derived quantities, we adopt a *fast bPCE* approach. In this approach, the sparse polynomial basis identified by the LARS algorithm during calibration is calculated only once from the available full experimental design $\mathcal{X}$, and bootstrapping is applied only to the final *hybrid* step, which consists in a classic ordinary least-square regression on the sparse basis (Blatman and Sudret, 2011).

In the presence of a very expensive model, however (*i.e.* requiring several hours for a single model run), we recommended to adopt full bootstrapping, including the estimation of the sparse PCE basis for each of the B bootstrapped experimental designs $\mathcal{X}^{(1,\cdots,B)}$.

2.3 Active bPCE-based reliability analysis

In this section we present an adaptation of the Adaptive PC-Kriging MCS algorithm in Schöbi et al. (2016) (based in turn on the original AK-MCS algorithm by Echard et al. (2011)), that makes use of the bPCE just introduced. Consistently with Echard et al. (2011); Schöbi et al. (2016), in the following we will refer to this algorithm as *active bootstrap-polynomial-chaos Monte-Carlo simulation* (A-bPCE). We follow the original idea of adaptively building a surrogate of the limit-state function starting from a small initial experimental design and subsequently refining it to optimize the surrogate performance for structural reliability. The ultimate goal of the adaptation is to retrieve an estimate of P_F that is comparable to that of a direct Monte Carlo simulation (MCS) using a large sample set with a much smaller experimental design. The algorithm is summarized as follows:

0. *Initialization*:
   (a) Generate an initial experimental design (*e.g.* through Latin hypercube sampling or uniform sampling of a ball (Dubourg, 2011)) and calculate the corresponding bPCE surrogate (see Section 2.3.1).
   (b) Generate a large reference MCS sample $\mathcal{X}_{MCS} = \left\{ \boldsymbol{x}^{(i)}_{MCS},\ i = 1, \cdots, N_{MCS} \right\}$ of size N_{MCS} (*e.g.* $N_{MCS} = 10^6 \gg N$). A discussion on the choice of a suitable MCS sample is given in Section 2.3.2).
1. Calculate a set of MCS estimators of the probability of failure: $\left\{ \widehat{P}_F^{(b)},\ b = 1, \cdots, B \right\}$ with the current bPCE surrogate.
2. Evaluate one or more suitable convergence criteria on $\widehat{P}_F^{(1,\cdots,B)}$ (see Section 2.3.3). If they are met, go to Step 5 (terminate the algorithm). Otherwise continue to the next

step.

3. Evaluate a suitable learning function on the MCS sample $\mathcal{X}_{MCS}$ (see Section 2.3.4). Choose one or more additional $\boldsymbol{x}_{MCS}^{(i)} \in \mathcal{X}_{MCS}$ and add them to the ED (see Section 2.3.5).
4. Update the bPCE surrogate on the new ED and return to Step 1
5. *Algorithm termination*: return the $\widehat{P}_F$ resulting from the PCE on the current ED, as well as the error bounds derived *e.g.* from the extremes or the empirical quantiles of the current $\widehat{P}_F^{(1,\cdots,B)}$ set.

A detailed description of each step of the algorithm is given in the following sections.

2.3.1 Initial experimental design

The initial experimental design is usually generated by space-filling sampling techniques of the random vector $\boldsymbol{X}$, such as Latin hypercube sampling (LHS) or pseudo-random sequences (*e.g.* Sobol' sequence). Alternative sampling techniques, such as the uniform sampling of a ball, have also proven effective in the context of structural reliability when low probabilities of failure are expected Dubourg (2011). Note that this initial set of model evaluations does not need to be a subset of the reference sample $\mathcal{X}_{MCS}$ used later to evaluate the P_F estimates during the iterations of the algorithm.

2.3.2 Inner MCS-based estimate of P_F

While the estimation of the P_F via MCS is trivial, as it simply entails counting the number of samples that belong to the failure domain, some discussion about the number of samples N_{MCS} in this step is needed. Throughout this paper, we opted to choose a single MCS sample $\mathcal{X}_{MCS}$ large enough to ensure a relatively small CoV for the P_F estimate at every iteration. This is by no means a requirement of this algorithm, but it simplifies significantly the notation (because $\mathcal{X}_{MCS}$ becomes independent on the current iteration) and in some cases (as noted in both Echard et al. (2011) and Schöbi et al. (2016)) it can result in stabler convergence, due to the lowered MCS noise in the estimation of P_F during each iteration. This technique is known as *common random numbers* in the context of repeated reliability analysis *.e.g.* in reliability-based design optimization (Taflanidis and Beck, 2008). It is entirely possible to redraw the $\mathcal{X}_{MCS}$ during every iteration, possibly each time with a different number of samples N_{MCS}.

The choice of $N_{MCS} = 10^6$ ensures that the CoV estimated probabilities of failure in the order of $P_F \geq 10^{-3}$ is always smaller than 5%, which we found suitable in our application examples. The choice of a single MCS sample drawn during the algorithm initialization also allows us to use the application examples to focus more on the convergence of the active learning part of A-bPCE, which is the focus of this paper.

In more general applications, the order of magnitude of P_F may be unknown. In this case, it is recommended instead to set a target desired CoV for the estimation of P_F at each iteration (as proposed in the original AK-MCS algorithm in Echard et al. (2011)), and gradually add samples to $\mathcal{X}_{MCS}$ until it is reached.

2.3.3 Convergence criteria

The proposed convergence criterion of choice is directly inspired by Schöbi et al. (2016); Notin et al. (2010) and it depends on the stability of the P_F estimate at the current iteration. Let us define:

$$\begin{aligned} \widehat{P}_F^+ &= \max_{b=1,\cdots,B}\left(\widehat{P}_F^{(b)}\right) \\ \widehat{P}_F^- &= \min_{b=1,\cdots,B}\left(\widehat{P}_F^{(b)}\right). \end{aligned} \tag{8}$$

Convergence is achieved when the following condition is satisfied for at least two consecutive iterations of the algorithm:

$$\frac{\widehat{P}_F^+ - \widehat{P}_F^-}{\widehat{P}_F} \leq \epsilon_{\widehat{P}_F}, \tag{9}$$

with $0.05 \leq \epsilon_{\widehat{P}_F} \leq 0.15$ in typical usage scenarios.

2.3.4 Learning function

A learning function is a function that allows one to rank a set of candidate points based on some utility criterion that depends on the desired application. In this case, we adopt the same heuristic approach proposed in Schöbi et al. (2016), by focusing on the probability of misclassification of the bPCE model on the candidate set given by $\mathcal{X}_{MCS}$.

Due to the availability of the bootstrap response samples $\mathcal{Y}_{MCS}^{(1,\cdots,B)}$, it is straightforward to define a measure of the misclassification probability U_{FBR} (where the subscript FBR stands for *failed bootstrap replicates*) at each point $\boldsymbol{x}_{MCS}^{(i)} \in \mathcal{X}_{MCS}$ as follows:

$$U_{FBR}(\boldsymbol{x}_{MCS}^{(i)}) = \left|\frac{B_S(\boldsymbol{x}_{MCS}^{(i)}) - B_F(\boldsymbol{x}_{MCS}^{(i)})}{B}\right| \tag{10}$$

where $B_S(\boldsymbol{x}_{MCS}^{(i)})$ and $B_F(\boldsymbol{x}_{MCS}^{(i)})$ are the number of safe (resp. failed) bPCE replicate predictions at point $\boldsymbol{x}_{MCS}^{(i)}$ (with $B_S(\boldsymbol{x}_{MCS}^{(i)}) + B_F(\boldsymbol{x}_{MCS}^{(i)}) = B$). When all the B replicates consistently classify $\boldsymbol{x}_{MCS}^{(i)}$ in the safe or in the failure domain, $U_{FBR} = 1$ (minimum misclassification probability). In contrast, $U_{FBR} = 0$ corresponds to the case when the replicates are equally distributed between the two domains. In the latter case, 50% of the B bootstrap PCEs predict that $\boldsymbol{x}_{MCS}^{(i)}$ is in the safe domain, while the other 50% predicts that $\boldsymbol{x}_{MCS}^{(i)}$ belongs to the failure domain. Therefore, maximum epistemic uncertainty on the classification of a point $\boldsymbol{x}_{MCS}^{(i)}$ is attained when U_{FBR} is minimum.

2.3.5 Enrichment of the experimental design

The aim of the iterative algorithm described in Section 2.3 is to obtain a surrogate model that minimizes the misclassification probability. As a consequence, the learning function in Eq. (10) can be directly used to obtain a single-point enrichment criterion. The next best candidate point for the ED $\boldsymbol{x}^* \in \mathcal{X}_{MCS}$ is given by:

$$\boldsymbol{x}^* = \underset{\boldsymbol{x}^{(i)} \in \mathcal{X}_{MCS}}{\operatorname{argmin}} \left(U_{FBR}(\boldsymbol{x}^{(i)}) \right). \tag{11}$$

Due to the global character of regression-based PCE, it can be beneficial to add multiple points in each iteration to sample several interesting regions of the parameter space simultaneously. The criterion in Eq. (11) can be extended to include K distinct points simultaneously by following the approach in Schöbi et al. (2016). A *limit state margin* region is first defined as the set of points such that $U_{FBR} < 1$ (*i.e.* those point with non-zero misclassification probability at the current iteration). Subsequently, k-means clustering techniques (see, *e.g.*, Zaki and Meira (2014)) can be used at each iteration to identify K disjoint regions $\left\{ \mathcal{X}_{MCS}^{(1,\cdots,K)} \right\}$ in the limit-state margin. Then, Eq. (11) can be directly applied to each of the subregions to obtain K different enrichment points:

$$\boldsymbol{x}_k^* = \underset{\boldsymbol{x}^{(i)} \in \mathcal{X}_{MCS}^{(k)}}{\operatorname{argmin}} \left(U_{FBR}^{(k)}(\boldsymbol{x}_k^{(i)}) \right), \quad k = 1, \cdots, K \tag{12}$$

where $\boldsymbol{x}_k^* \in \mathcal{X}_{MCS}^{(k)}$ is the $k-$th enrichment sample and $U_{FBR}^{(k)}$ is the learning function evaluated on the $k-$th region of the parameter space.

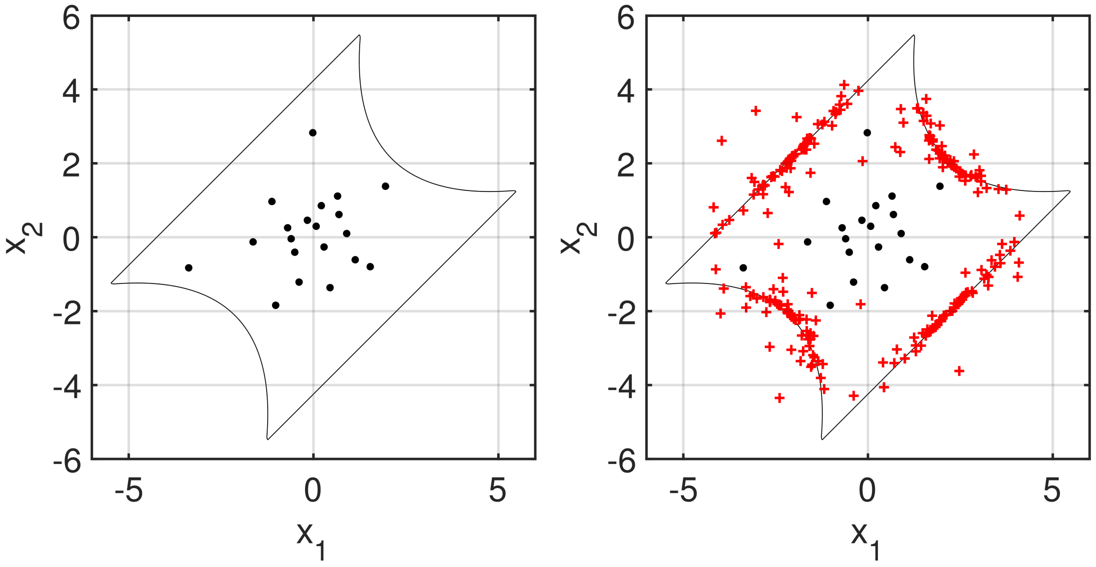
Note that this approach is also convenient when parallel computing facilities are available and in the presence of computationally expensive objective functions, as the evaluation of the K enrichment points can be carried out simultaneously.

3 Results on benchmark applications

All the algorithm development and the final calculations presented in this section were performed with the polynomial chaos expansions and reliability analysis modules of the UQLab software for uncertainty quantification (Marelli and Sudret, 2014, 2017; Marelli et al., 2017).

3.1 Series system

A common benchmark for reliability analysis functions is given by the four-branch function, originally proposed in Waarts (2000), that represents a series system comprising four components with different failure criteria. Although it is a simple analytical function, it shows multiple failure regions and a composite limit-state surface. Its two-dimensional limit state



a. *Initial experimental design*

b. *Final experimental design*

Figure 2: Four branch function: limit state surface (black line) and experimental design before (black circles) and after (red crosses) enrichment

function reads:

$$g(\boldsymbol{x}) = \min \begin{Bmatrix} 3 + 0.1(x_1 + x_2)^2 - \frac{x_1+x_2}{\sqrt{2}} \\ 3 + 0.1(x_1 + x_2)^2 + \frac{x_1+x_2}{\sqrt{2}} \\ (x_1 - x_2) + \frac{6}{\sqrt{2}} \\ (x_2 - x_1) + \frac{6}{\sqrt{2}} \end{Bmatrix} \tag{13}$$

where the two random input variables $X_1 \sim \mathcal{N}(0, 1)$ and $X_2 \sim \mathcal{N}(0, 1)$ are modelled as independent standard normals. Failure occurs when $g(\boldsymbol{x}) \leq 0$.

Due to the multi-failure shape of the limit-state surface (represented as a solid black line in Figure 2), classic methods like FORM/SORM and importance sampling tend to fail with this benchmark problem. The reference failure probability of $P_F = 4.460 \cdot 10^{-3}$ is obtained through an extremely large MCS ($N_{MCS} = 10^8$).

The initial experimental design for the A-bPCE algorithm was obtained with a space-filling LHS sample consisting of $N_{ini} = 20$ points drawn from the input distributions (black dots in Figure 2). Three points at a time were added to the experimental design during the enrichment phase of the algorithm. The number of replications for the A-bPCE algorithm is set to $B = 100$. After extensive testing, the algorithm was found to be very weakly dependent on the number of bootstrap replications, provided a minimum of $B \geq 20$ was provided. Indeed, the boostrap samples are used to identify areas of relatively large prediction variability, but an accurate estimate of such variability is never really needed by the algorithm. Degree adaptive sparse PCE (with maximum degree in the range $p \in [2, 10]$) based on LARS (Blatman and Sudret, 2011) was used to calibrate the PCE metamodel at each iteration. For validation and comparison purposes, a similar analysis was performed on the same initial

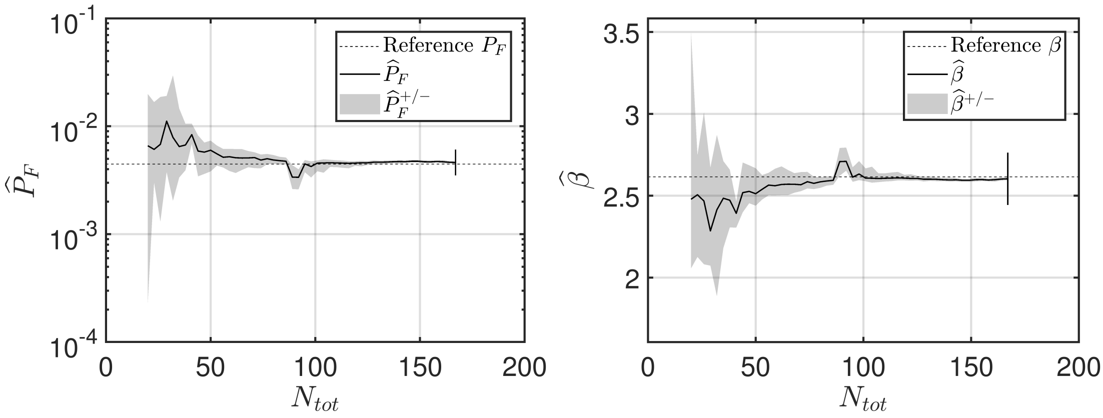


Figure 3: Convergence curves of the four-branch limit-state function. The reference P_F and β are given by dotted lines.

ED with the AK-MCS module of UQLab, with an anisotropic Matérn 5/2 ellipsoidal multivariate Kriging correlation function (Rasmussen and Williams, 2006; Marelli et al., 2017; Lataniotis et al., 2017). The convergence criterion in Eq. (9) was set to $\epsilon_{\widehat{P}_F} = 0.05$ for both the AK-MCS and A-bPCE algorithms.

Convergence was achieved after 49 iterations, resulting in a total cost (including the initial experimental design) of $N_{tot} = 185$ model evaluations. The experimental design points added during the iterations are marked by red crosses on panel (b) of Figure 2. As expected, the adaptive algorithm tends to enrich the experimental design close to the limit state surface as it is adaptively learned during the iterations. A graphical representation of the convergence of the algorithm is shown in Figure 3, where the estimated $\widehat{P}_F$ is plotted against the total number of model evaluations N_{tot}. The shaded area represents the 95% confidence bounds based on the empirical quantiles as estimated from the bootstrap sample.

The final results of the analysis are summarized in Table 1, where the generalised reliability index $\beta = -\Phi^{-1}(P_F)$ is also given for reference. For comparison, the reference MCS probability as well as an estimate from AK-MCS are also given. The latter converged to a comparably accurate estimate of P_F, at the cost of a slightly higher number of model evaluations. Note that for engineering purposes, the algorithm could have been stopped earlier, *i.e.* when a 5% accuracy on the generalized reliability index is attained. In this case, the algorithm would have converged to a comparable result ($\widehat{\beta} = 2.56$) with only 50 runs of the model. The final sparse PCE model after enrichment contained a total of $P = 12$ basis elements of degree up to $p = 5$.

3.2 Two-dimensional truss structure

To test the algorithm on a more realistic engineering benchmark, consider the two-dimensional truss structure sketched in Figure 4. This structure has been previously analysed in several works, see *e.g.* (Blatman and Sudret, 2011, 2010b; Schöbi et al., 2016). The truss comprises

Table 1: Comparison of different reliability analysis methods for the four branch function

Algorithm	$\widehat{P}_F$	$[\widehat{P}_F^-, \widehat{P}_F^+]$	$\widehat{\beta}$	$[\widehat{\beta}^-, \widehat{\beta}^+]$	N_{tot}
MCS (ref.)	4.46×10^{-3}	-	2.62	-	10^8
AK-MCS	4.52×10^{-3}	$[4.38, 4.65] \times 10^{-3}$	2.61	$[2.60, 2.62]$	200
A-bPCE	4.62×10^{-3}	$[4.5, 4.7] \times 10^{-3}$	2.60	$[2.59, 2.61]$	167

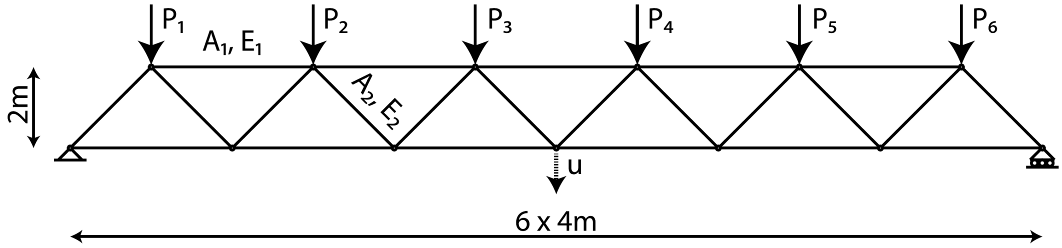


Figure 4: Two-dimensional truss structure with uncertain parameters. Probability distributions of the geometrical parameters $\{A_i, E_i\}$ and of the loads $\{P_1, \cdots, P_6\}$ are given in Table 2

23 bars and 13 nodes, with deterministic geometry yet with uncertain material properties and random loads. The components of the input random vector $\boldsymbol{X} = [A_{1,2}, E_{1,2}, P_{1,\dots,6}]^{\mathsf{T}}$ include the cross-section and the Young's modulus $\{A_1, E_1\}$ of the horizontal bars, the cross-section and the Young's modulus $\{A_2, E_2\}$ of the diagonal bars and the six random loads $\{P_1, \cdots, P_6\}$. They are considered mutually independent and their distributions are given in Table 2. An in-house developed Matlab-based finite-element solver is used to calculate the displacement at midspan $u(\boldsymbol{X})$, counted positively downwards.

This structure operates in the nominal range as long as the midspan displacement is smaller than a critical threshold $\tau = 12$ cm, which can be cast as the following limit-state function:

$$g(\boldsymbol{x}) = \tau - u(\boldsymbol{x}) \tag{14}$$

where $g(\boldsymbol{x}) \leq 0$ if the system is in a failure state.

Because the FEM computational model is relatively cheap to evaluate, we could run a direct MCS-analysis with $N = 10^6$ samples to provide the reference $P_F = 1.52 \cdot 10^{-3}$ for

Table 2: Two-dimensional truss structure: definition of the probabilistic model of the input variables (Schöbi et al., 2016)

Variable	Distribution	Mean	Standard Deviation
E_1, E_2 (Pa)	Lognormal	2.1×10^{11}	2.1×10^{10}
A_1 (m^2)	Lognormal	2.0×10^{-3}	2.0×10^{-4}
A_2 (m^2)	Lognormal	1.0×10^{-3}	1.0×10^{-4}
$P_1, \cdots, P_6$ (N)	Gumbel	5.0×10^4	7.5×10^3

Table 3: Comparison of the estimation of P_F with several algorithms for the truss structure example.

Algorithm	$\widehat{P}_F$	$[\widehat{P}_F^-, \widehat{P}_F^+]$	$\widehat{\beta}$	$[\widehat{\beta}^-, \widehat{\beta}^+]$	N_{tot}
MCS (ref.)	1.52×10^{-3}	-	2.96	-	10^6
FORM	0.76×10^{-3}	-	3.17	-	160
SORM	1.63×10^{-3}	-	2.94	-	372
AK-MCS	1.52×10^{-3}	$[1.44, 1.59] \times 10^{-3}$	2.96	$[2.90, 3.01]$	300
A-bPCE	1.48×10^{-3}	$[1.43, 1.54] \times 10^{-3}$	2.97	$[2.96, 2.98]$	129

validation purposes. Additionally, standard FORM and SORM analyses were run to estimate the non-linearity of the limit-state surface. FORM underestimated the failure probability of a factor of almost 2 and a cost of $N_{FORM} = 160$ model runs, while SORM achieved a good accuracy at a cost of $N_{SORM} = 372$ model runs, which suggests that the underlying problem is non-linear. Neither of the two methods, however, provides confidence interval on their estimates.

The A-bPCE algorithm was initialized with an experimental design consisting in a uniform sampling of a ball (for details, see *e.g.* Dubourg (2011)) of size $N_{ini} = 30$, while the sparse adaptive PCE was given a polynomial degree range $1 \leq p \leq 10$, hyperbolic truncation with q-norm $q = 0.75$ (Blatman and Sudret, 2011) and maximum allowed interaction $r = 2$ (Marelli and Sudret, 2017). The internal MCS sample size was $N_{MCS} = 10^6$ and the algorithm was set to add $K = 3$ new samples per iteration. The stopping criterion in Eq. (9) was set to $\epsilon_{\widehat{P}_F} = 0.10$. For comparison purposes, we also ran a standard AK-MCS analysis with the same initial experimental design and convergence criterion. The covariance family of choice for the underlying Kriging model was chosen as Gaussian.

Table 3 presents a comparison of the estimated $\widehat{P}_F$ with the aforementioned analyses. Both AK-MCS and A-bPCE estimates of P_F include the reference value within the confidence bounds set by the convergence criterion. However, for this particular example and choice of convergence criterion, A-bPCE achieved convergence significantly faster than AK-MCS, with a total cost of 129 model evaluations, as compared to the 300 required by AK-MCS, resulting in a final PCE of degree $p = 3$ with $P = 43$ basis elements.

Overall, A-bPCE provides a stable estimate of the failure probability and confidence intervals at a cost that is lower than FORM for this example.

3.3 Top-floor displacement of a structural frame

Figure 5 shows a well known, high dimensional benchmark in structural reliability applications (Liu and Der Kiureghian, 1991; Blatman and Sudret, 2010a). It consists on a three-span, five story frame structure that is subject to horizontal loads. Both the loads and the properties of the elements of the frame (see Table 4) are uncertain. Of interest is the top-floor

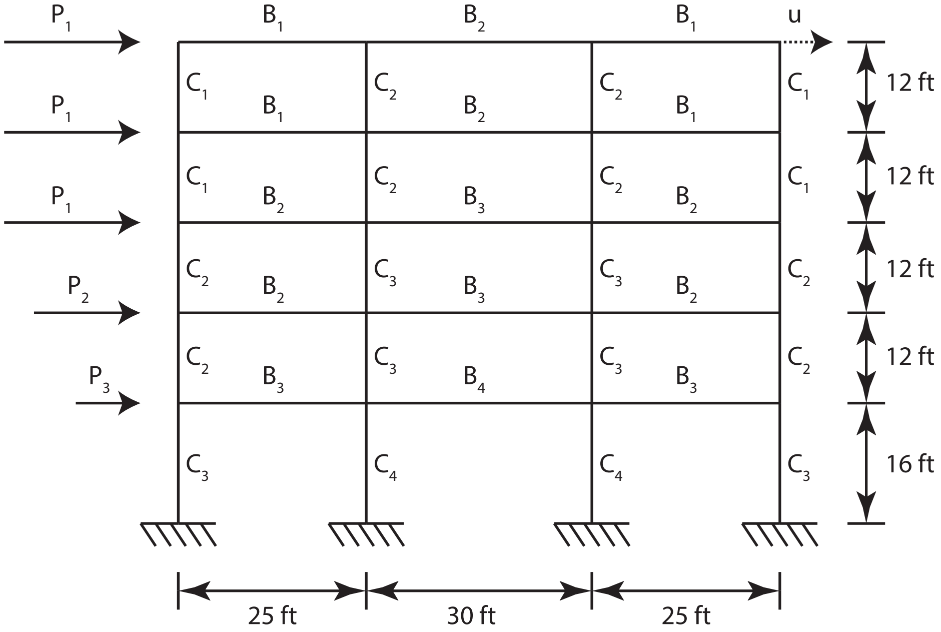


Figure 5: 20-dimensional structural frame in Section 3.3. The distributions of the input variables are reported in Table 5

.

horizontal displacement at the top right corner u.

The uncertainties on the applied loads $P_1, ..., P_3$, the Young's moduli E_1 and E_2, the moments of inertia $I_6, ..., I_{13}$ and the cross sections $A_{14}, ..., A_{21}$ are modelled by a 21-dimensional joint random vector $\boldsymbol{Z} = \{P_1, ..., A_{21}\}$ with marginal distributions given in Table 5.

Additionally, a Gaussian copula (Lebrun and Dutfoy, 2009) is used to model dependence between the variables. The elements of the Gaussian copula correlation matrix $\boldsymbol{R}$ are given as:

- $R_{E_1,E_2} = 0.9$ – the two Young's moduli are highly correlated;
- $R_{A_i,I_i} = 0.95$ – each element's cross-sectional area is highly correlated to the corresponding moment of inertia;
- $R_{A_i,I_j} = R_{I_i,I_j} = R_{A_i,A_j} = 0.13$ – the correlation between the properties of different elements is much lower;
- All the remaining elements of $\boldsymbol{R}$ are set to 0.

A critical displacement of $\tau = 5$cm is identified as the maximum admissible threshold for the displacement u, hence resulting in the limit-state function:

$$g(\boldsymbol{z}) = \tau - u(\boldsymbol{z}) \tag{15}$$

where $u(\boldsymbol{z})$ is the displacement on the top right corner calculated with an in-house FEM code. Due to the associated computational costs, the maximum available budget for the calculation

Table 4: Frame structure: properties of the elements shown in Figure 5

Element	Young's modulus	Moment of inertia	Cross-sectional area
B_1	E_4	I_{10}	A_{18}
B_2	E_4	I_{11}	A_{19}
B_3	E_4	I_{12}	A_{20}
B_4	E_4	I_{13}	A_{21}
C_1	E_5	I_6	A_{14}
C_2	E_5	I_7	A_{15}
C_3	E_5	I_8	A_{16}
C_4	E_5	I_9	A_{17}

of a reference solution is in this case limited to $N_{REF} = 4 \cdot 10^4$. Therefore, the reference solution is calculated with standard importance sampling (IS) (Melchers, 1999) instead of direct MCS. In addition to Importance sampling, we also ran FORM and SORM. Due to the non-linearity of the problem, FORM significantly underestimated P_F, while SORM provided an accurate estimate. However, due to the high dimensionality of the input space, the associated cost in terms of model evaluation was relatively high, with $N_{SORM} = 1146$ model runs, since all the gradients of the limit-state function are computed using finite-differences.

The A-bPCE algorithm was initialized with an experimental design consisting of an LHS sampling of the input random vector of size $N_{ini} = 40$. Sparse PCE was carried out with a q-norm truncation with $q = 0.75$ and maximum allowed interaction $r = 2$. Note that the initialization is essentially the same as for the truss structure in the previous application. The internal MCS sample size was $N_{MCS} = 10^6$, with single point enrichment per iteration. The stopping criterion in Eq. (9) was in this case set to $\epsilon_{\widehat{P}_F} = 0.15$. For comparison purposes, an AK-MCS analysis was also run on the same initial design, with similar settings and a Gaussian covariance family.

A comparison of the results is gathered in Table 6. Due to the different estimation method between the reference probability (importance sampling) and the active learning-based methods (which rely on an inner MCS), no direct comparison of the results is possible as in the previous cases. Indeed, even fixing the same random seeds would result in different estimates due to the different methodologies. Therefore, confidence bounds are given for all the three methods: 95% confidence bounds for IS (Melchers, 1999), and $P_F^{\pm}$ for both AK-MCS and A-bPCE. The three methods give comparable results, albeit with significant differences in the convergence behaviour. In particular, both AK- and A-bPCE resulted in a slight underestimation of the probability of failure w.r.t. the reference solution by IS, which in turn is slightly overestimated with respect to the reference result quoted in the literature (Blatman and Sudret, 2010a). However, AK-MCS did not converge in the allotted

Table 5: Frame structure: definition of the probabilistic model of the input variables (Blatman and Sudret, 2010a)

Variable	Distribution	Mean	Standard Deviation
P_1 (kN)	Lognormal	133.454	40.04
P_2 (kN)	Lognormal	88.97	35.59
P_3 (kN)	Lognormal	71.175	28.47
E_4 (kN/m^2)	Truncated Gaussian*	2.1738×10^7	1.9152×10^6
E_5 (kN/m^2)	Truncated Gaussian*	2.3796×10^7	1.9152×10^6
I_6 (m^4)	Truncated Gaussian*	8.1344×10^{-3}	1.0834×10^{-3}
I_7 (m^4)	Truncated Gaussian*	1.1509×10^{-2}	1.2980×10^{-3}
I_8 (m^4)	Truncated Gaussian*	2.1375×10^{-2}	2.5961×10^{-3}
I_9 (m^4)	Truncated Gaussian*	2.5961×10^{-2}	3.0288×10^{-3}
I_{10} (m^4)	Truncated Gaussian*	1.0812×10^{-2}	2.5961×10^{-3}
I_{11} (m^4)	Truncated Gaussian*	1.4105×10^{-2}	3.4615×10^{-3}
I_{12} (m^4)	Truncated Gaussian*	2.3279×10^{-2}	5.6249×10^{-3}
I_{13} (m^4)	Truncated Gaussian*	2.5961×10^{-2}	6.4902×10^{-3}
A_{14} (m^2)	Truncated Gaussian*	3.1256×10^{-1}	5.5815×10^{-2}
A_{15} (m^2)	Truncated Gaussian*	3.7210×10^{-1}	7.4420×10^{-2}
A_{16} (m^2)	Truncated Gaussian*	5.0606×10^{-1}	9.3025×10^{-2}
A_{17} (m^2)	Truncated Gaussian*	5.5815×10^{-1}	1.1163×10^{-1}
A_{18} (m^2)	Truncated Gaussian*	2.5302×10^{-1}	9.3025×10^{-2}
A_{19} (m^2)	Truncated Gaussian*	2.9117×10^{-1}	1.0232×10^{-1}
A_{20} (m^2)	Truncated Gaussian*	3.7303×10^{-1}	1.2093×10^{-1}
A_{21} (m^2)	Truncated Gaussian*	4.1860×10^{-1}	1.9537×10^{-1}

* Truncated in the domain $[0, +\infty]$. The quoted moments refer to the full, untruncated Gaussian distributions.

Table 6: Comparison of the estimation of P_F with several algorithms for the top-floor displacement of a structural frame example

Algorithm	$\widehat{P}_F$	$[\widehat{P}_F^-, \widehat{P}_F^+]$	$\widehat{\beta}$	$[\widehat{\beta}^-, \widehat{\beta}^+]$	N_{tot}
IS (ref.)	1.54×10^{-3}	$[1.51, 1.56] \times 10^{-3}$	2.96	$[2.95, 2.97]$	40241
FORM	1.01×10^{-3}	-	3.09	–	241
SORM	1.52×10^{-3}	-	2.96	–	1146
AK-MCS	1.48×10^{-3}	$[0.8, 4.96] \times 10^{-3}$	2.97	$[2.57, 3.12]$	300
A-bPCE	1.49×10^{-3}	$[1.42, 1.62] \times 10^{-3}$	2.97	$[2.94, 2.98]$	235

maximum number of model evaluations, and its confidence bounds remained remarkably large with respect to A-bPCE. A-bPCE converged instead at a total cost of approximately 200 model evaluations to the target $\epsilon_{\widehat{P}_F}$, with a final sparse PCE of degree 2, counting 30 non-zero coefficients. For both active-learning-based methods, the reference solution lies within the given confidence bounds. Moreover, the confidence bounds on the reliability index $\widehat{\beta}$ show that the results are stable to within 2% of the calculated values.

Finally, it is interesting to mention that for this example the costs of FORM and A-bPCE were comparable, but the latter provides a much less biased estimate, and includes confidence bounds.

4 Conclusions and outlook

A novel approach to solving reliability problems with polynomial chaos expansions has been proposed. The combination of the bootstrap method and sparse regression enabled us to introduce local error estimation in the standard PCE predictor. In turn, this allows one to construct active learning algorithms similar to AK-MCS to greedily enrich a relatively small initial experimental design so as to efficiently estimate the probability of failure of complex systems.

This approach has shown comparable performance *w.r.t.* to the well established AK-MCS method on both a simple analytical benchmark function and in two high-dimensional engineering applications of increasing complexity.

Extensions of this approach can be envisioned in two main directions:

- the simulation-based reliability analysis method can be extended beyond simple MCS (*e.g.* by using importance sampling (Dubourg et al., 2013), line sampling (Pradlwarter et al., 2007) or subset simulation (Dubourg et al., 2011)) to achieve better $\widehat{P}_F$ estimates at each iteration, especially for very low probabilities of failure;
- remote parallel computing facilities may be used during the enrichment phase of the

algorithms with expensive computational models when adding more than one point at a time;

- the use of bootstrap to enable local error estimation in an active learning context can be used also with different regression-based surrogate modelling techniques, including *e.g.* low rank tensor approximations (Konakli and Sudret, 2016).

Additionally, the bPCE approach itself introduced in this work can be used also outside of a pure reliability analysis context, as it provides an effective local error estimate for PCE. It has been used, *e.g.* in the context of reliability-based design optimization in Moustapha and Sudret (2017). Indeed the lack of this feature (as opposed to Kriging) has somewhat hindered its usage in more advanced active-learning applications.

References

References

Balesdent, M., J. Morio, and J. Marzat (2013). Kriging-based adaptive importance sampling algorithms for rare event estimation. *Structural Safety 44*, 1–10.

Bichon, B., M. Eldred, L. Swiler, S. Mahadevan, and J. McFarland (2008). Efficient global reliability analysis for nonlinear implicit performance functions. *AIAA Journal 46*(10), 2459–2468.

Bichon, B., J. McFarland, and S. Mahadevan (2011). Efficient surrogate models for reliability analysis of systems with multiple failure modes. *Reliab. Eng. Sys. Safety 96*(10), 1386–1395.

Blatman, G. and B. Sudret (2010a). An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. *Prob. Eng. Mech. 25*, 183–197.

Blatman, G. and B. Sudret (2010b). Efficient computation of global sensitivity indices using sparse polynomial chaos expansions. *Reliab. Eng. Sys. Safety 95*, 1216–1229.

Blatman, G. and B. Sudret (2011). Adaptive sparse polynomial chaos expansion based on Least Angle Regression. *J. Comput. Phys 230*, 2345–2367.

Cadini, F., F. Santos, and E. Zio (2014). An improved adaptive Kriging-based importance technique for sampling multiple failure regions of low probability. *Reliab. Eng. Sys. Safety 131*, 109–117.

den Hertog, D., J. P. C. Kleijnen, and A. Y. D. Siem (2006). The correct Kriging variance estimated by bootstrapping. *Journal of the Operational Research Society 57*(4), 400–409.

Dubourg, V. (2011). *Adaptive surrogate models for reliability analysis and reliability-based design optimization.* Ph. D. thesis, Université Blaise Pascal, Clermont-Ferrand, France.

Dubourg, V. and B. Sudret (2014). Metamodel-based importance sampling for reliability sensitivity analysis. *Structural Safety 49*, 27–36.

Dubourg, V., B. Sudret, and J.-M. Bourinet (2011). Reliability-based design optimization using Kriging and subset simulation. *Struct. Multidisc. Optim. 44*(5), 673–690.

Dubourg, V., B. Sudret, and F. Deheeger (2013). Metamodel-based importance sampling for structural reliability analysis. *Prob. Eng. Mech. 33*, 47–57.

Dubreuil, S., M. Berveiller, F. Petitjean, and M. Salaün (2014). Construction of bootstrap confidence intervals on sensitivity indices computed by polynomial chaos expansion. *Reliab. Eng. Sys. Safety 121*, 263–275.

Echard, B., N. Gayton, and M. Lemaire (2011). AK-MCS: an active learning reliability method combining Kriging and Monte Carlo simulation. *Structural Safety 33*(2), 145–154.

Echard, B., N. Gayton, M. Lemaire, and N. Relun (2013). A combined importance sampling and Kriging reliability method for small failure probabilities with time-demanding numerical models. *Reliab. Eng. Syst. Safety 111*, 232–240.

Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans*, Volume 38. SIAM.

Efron, B., T. Hastie, I. Johnstone, and R. Tibshirani (2004). Least angle regression. *Annals of Statistics 32*, 407–499.

Fauriat, W. and N. Gayton (2014). AK-SYS: An adaptation of the AK-MCS method for system reliability. *Reliab. Eng. and Sys. Safety*, 137–144.

Ghanem, R. and P. Spanos (1991). *Stochastic finite elements – A spectral approach.* Springer Verlag, New York. (Reedited by Dover Publications, Mineola, 2003).

Hasofer, A.-M. and N.-C. Lind (1974). Exact and invariant second moment code format. *J. Eng. Mech. 100*(1), 111–121.

Konakli, K. and B. Sudret (2016). Reliability analysis of high-dimensional models using low-rank tensor approximations. *Prob. Eng. Mech. 46*, 18–36.

Lataniotis, C., S. Marelli, and B. Sudret (2017). UQLab user manual – Kriging. Technical report, Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich. Report UQLab-V1.0-105.

Le Gratiet, L., S. Marelli, and B. Sudret (2017). *Metamodel-based sensitivity analysis: polynomial chaos expansions and Gaussian processes*, Chapter 8. Handbook on Uncertainty Quantification, Springer.

Lebrun, R. and A. Dutfoy (2009). Do Rosenblatt and Nataf isoprobabilistic transformations really differ? *Prob. Eng. Mech. 24*(4), 577–584.

Lemaire, M. (2009). *Structural reliability.* Wiley.

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

Marelli, S., R. Schöbi, and B. Sudret (2017). UQLab user manual – Reliability analysis. Technical report, Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich. Report UQLab-V1.0-107.

Marelli, S. and B. Sudret (2014). UQLab: A framework for uncertainty quantification in Matlab. In *Vulnerability, Uncertainty, and Risk (Proc. 2nd Int. Conf. on Vulnerability, Risk Analysis and Management (ICVRAM2014), Liverpool, United Kingdom)*, pp. 2554–2563.

Marelli, S. and B. Sudret (2017). UQLab user manual – Polynomial chaos expansions. Technical report, Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich. Report UQLab-V1.0-104.

Melchers, R.-E. (1999). *Structural reliability analysis and prediction.* John Wiley & Sons.

Moustapha, M. and B. Sudret (2017). Quantile-based optimization under uncertainties using bootstrap polynomial chaos expansions. In *Proc. 12th Internatinoal Conference on Structural Safety and Reliability (ICOSSAR), August 6-10, 2017, Vienna, Austria.*

Notin, A., N. Gayton, J. L. Dulong, M. Lemaire, P. Villon, and H. Jaffal (2010). RPCM: a strategy to perform reliability analysis using polynomial chaos and resampling. *European Journal of Computational Mechanics 19*(8), 795–830.

Picheny, V., N. Kim, and R. Haftka (2010). Application of bootstrap method in conservative estimation of reliability with limited samples. *Struct. Multidisc. Optim. 41*(2), 205–217.

Pradlwarter, H., G. Schuller, P. Koutsourelakis, and D. Charmpis (2007). Application of line sampling simulation method to reliability benchmark problems. *Structural Safety 29*(3), 208 – 221.

Rackwitz, R. and B. Fiessler (1978). Structural reliability under combined load sequences. *Computers & Structures 9*, 489–494.

Rasmussen, C. and C. Williams (2006). *Gaussian processes for machine learning* (Internet ed.). Adaptive computation and machine learning. Cambridge, Massachusetts: MIT Press.

Schöbi, R., B. Sudret, and S. Marelli (2016). Rare event estimation using Polynomial-Chaos-Kriging. *ASCE-ASME J. Risk Uncertainty Eng. Syst., Part A: Civ. Eng.*. D4016002.

Sudret, B. and A. Der Kiureghian (2002). Comparison of finite element reliability methods. *Prob. Eng. Mech. 17*, 337–348.

Taflanidis, A. A. and J. L. Beck (2008). Stochastic subset optimization for optimal reliability problems. *Prob. Eng. Mech 23*, 324–338.

Torre, E., S. Marelli, P. Embrechts, and B. Sudret (2017). A general framework for data-driven uncertainty quantification under complex input dependencies using vine copulas. *arXiv 1709.08626.* Under revision in Probabilistic Engineering Mechanics.

van Beers, W. C. and J. P. Kleijnen (2008). Customized sequential designs for random simulation experiments: Kriging metamodeling and bootstrapping. *European Journal of Operational Research 186*(3), 1099 – 1113.

Waarts, P.-H. (2000). *Structural reliability using finite element methods: an appraisal of DARS: Directional Adaptive Response Surface Sampling.* Ph. D. thesis, Technical University of Delft, The Netherlands.

Xiu, D. and G. Karniadakis (2002). The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput. 24*(2), 619–644.

Zaki, M. and W. Meira (2014). *Data Mining and Analysis: fundamental Concepts and Algorithms.* Cambridge University Press.