

Extreme value oriented random field discretization based on an hybrid polynomial chaos expansion — Kriging approach

S. Dubreuil^a, N. Bartoli^a, C. Gogu^{b,*}, T. Lefebvre^a, J. Mas Colomer^a

^a *Onera-The French Aerospace Lab, Toulouse, F-31055, France*

^b *Université de Toulouse, CNRS, UPS, INSA, ISAE, Mines Albi, Institut Clément Ader (ICA), 3 rue Caroline Aigle, Toulouse F-31400, France*

Received 13 February 2017; received in revised form 18 December 2017; accepted 8 January 2018

Available online 11 January 2018

Highlights

- We seek random fields metamodels which are accurate around their extreme values.
- The metamodel can accurately represent the global extremum for any realization.
- To achieve this an adaptive discretization and enrichment strategy is proposed.
- Original random field representation combining polynomial chaos expansion and Kriging.
- Additional dimensional reduction achieved by Karhunen Loeve decomposition.
- Three example problems, including the aerodynamic design of an aircraft wing.

Abstract

This article addresses the characterization of extreme value statistics of continuous second order random field. More precisely, it focuses on the parametric study of engineering models under uncertainty. Hence, the quantity of interest of this model is defined on both a parametric space and a stochastic space. Moreover, we consider that the model is computationally expensive to evaluate. For this reason it is assumed that uncertainty propagation, at a single point of the parametric space, is achieved by polynomial chaos expansion. The main contribution of the present study is the development of an adaptive approach for the discretization of the random field modeling the quantity of interest. Objective of this new approach is to focus the computational budget over the areas of the parametric space where the minimum or the maximum of the field is likely to be for any realization of the stochastic parameters. To this purpose two original random field representations, based on polynomial chaos expansion and Kriging interpolation, are introduced. Moreover, an original adaptive enrichment scheme based on Kriging is proposed. Advantages of this approach with respect to accuracy and computational cost are demonstrated on several numerical examples. The proposed method is also illustrated on the parametric study of an aircraft wing under uncertainty.

© 2018 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Random field discretization; Adaptive design of experiments; Polynomial chaos expansion; Kriging

* Correspondence to: Université de Toulouse, F-31055, Toulouse, France.
E-mail address: christian.gogu@gmail.com (C. Gogu).

1. Introduction

This study addresses the characterization of extreme value statistics of continuous second order random fields. It is motivated by the analysis of parametric engineering problems under parametric uncertainty. Indeed, let $y = \mathcal{M}(\mathbf{x}, \mathbf{d})$ be an engineering model, where $\mathcal{M}$ is a deterministic solver, $\mathbf{d} \in \mathbb{R}^{n_d}$ is a vector of control variables of the model and $\mathbf{x}$ is a vector of n model parameters defined on a parametric space $S \subset \mathbb{R}^n$. The minimum (or maximum) value of $y(\mathbf{x}, \mathbf{d})$, as well as its position in the parametric space S , are generally of interest in the study of the model y (e.g. for optimization of the modeled system).

It is now assumed that parameters $\mathbf{x}$ and $\mathbf{d}$ are affected by uncertainty. In order to take into account this uncertainty, the probabilistic framework is used. Hence, a probability space (Ω, F, P) and a random vector $\xi : (\Omega, F) \rightarrow (\mathbb{R}^{n_s}, \mathcal{B}_{n_s})$, where $\mathcal{B}_{n_s}$ is the Borel σ algebra and n_s the stochastic dimension of the problem, are introduced. The parameters of the problem are expressed by the random variables $X(\mathbf{x}, \xi) = f(\mathbf{x}, \xi)$ and $D(\mathbf{d}, \xi) = g(\mathbf{d}, \xi)$, where $f : \mathbb{R}^n \times \mathbb{R}^{n_s} \rightarrow \mathbb{R}^n$ and $g : \mathbb{R}^{n_d} \times \mathbb{R}^{n_s} \rightarrow \mathbb{R}^{n_d}$ are the known mappings that describe how the uncertainty affects the parameters $\mathbf{x}$ and $\mathbf{d}$ respectively. The parametric random problem, $Y(X(\mathbf{x}, \xi), D(\mathbf{d}, \xi)) = \mathcal{M}(X(\mathbf{x}, \xi), D(\mathbf{d}, \xi))$ can thus be modeled as a scalar random field over the parametric space S . As the purpose of this article is the study of this random field with respect to the variables $\mathbf{x}$ and the random variables ξ , the problem can be rewritten, without any loss of generality, as $Y(X(\mathbf{x}, \xi), D(\mathbf{d}, \xi)) = Y(\mathbf{x}, \xi)$ in order to lighten the notations.

In the following we are interested in the random variable modeling the minimum (or the maximum) of the random field $Y(\mathbf{x}, \xi)$, denoted by

$$Y_{min}(\xi) = \min_{\mathbf{x} \in S} (Y(\mathbf{x}, \xi)), \quad (1)$$

as well as the random variable modeling the position where this minimum (or maximum) is reached,

$$X^*(\xi) = \arg \min_{\mathbf{x} \in S} (Y(\mathbf{x}, \xi)). \quad (2)$$

As they allow to characterize the dispersion of the value and the position of the minimum of the random parametric problem under study, the random variables $Y_{min}(\xi)$ and $X^*(\xi)$ are important information for many applications in the domains of optimization, robustness and reliability. Note that the approach we will propose can handle both cases (maximum or minimum) of extreme values of the considered random field (the passage from one to the other being done by taking the opposite of the random field). In order to simplify notations we will from now on only consider the minimization case.

The study of extreme values generally deals with a discrete collection of independent random variables. In this particular case theoretical results are available for modeling the probability distribution of the maxima (e.g. the generalized extreme value distribution, see [1]). Concerning the study of continuous random process and fields, readers are referred to [2] for theoretical basis. Once again, some theoretical results are available for some particular random process and fields (see [2] for details). Otherwise, one can rely on numerical approximations, for example, [3] proposes to use an approximation of the random field by Karhunen Loève (KL) decomposition and Gaussian mixture in order to compute the extreme value statistics by application of the Rice formula [4] on this approximation.

The approach proposed in this paper is complementary to these previous approaches, by addressing a typical industrial context in which the deterministic model $\mathcal{M}(\mathbf{x}, \xi)$ can only be sampled at a given space position $\mathbf{x}^{(0)} \in S$ and for a given random realization $\xi^{(0)} \in \mathbb{R}^{n_s}$ (this type of application is sometimes referred as *black box* problem). Moreover, it is assumed that the response of this deterministic model is the output of a computationally expensive numerical simulation making the use of brute force Monte Carlo approach intractable. In this context, uncertainty quantification (at a given space point $\mathbf{x}^{(0)}$) by polynomial chaos expansion (PCE) is now well established and will also be used in the present work. Indeed several studies have demonstrated the effectiveness of PCE to deal with uncertainty quantification, from the pioneering work by Ghanem and Spanos [5] on stochastic finite elements by Hermite PCE to the generalized PCE [6,7] and adaptive PCE [8,9]. Computation of PCE approximations also made great progress in the last decade, in particular with the development of non intrusive approaches [10–14] that allow to compute PCE only by sampling the deterministic solver $\mathcal{M}$. These developments make PCE easy to implement even on heavy computational solvers $\mathcal{M}$ (e.g. non-linear finite elements simulations).

The study of extreme values is obviously closely related to optimization. In the context of solving parametrized optimization problems, an increasingly popular approach is surrogate based optimization [15–17], where a surrogate model [18,19] is constructed in order to aid the optimization process. Kriging based optimization algorithms have

been extensively investigated and developed in the past in particular due to their ability to trade-off exploration and exploitation of the design space [20–22]. In our context the problem is also affected by uncertainties, similarly to robust optimization. Robust optimization consists in optimizing a given statistical measure of the random objective function. Many statistical measures have been studied (mean value, variance, quantile, probability of exceeding a threshold) and readers are referred to [23] for a review). Problems of optimization under uncertainty have also been treated by constructing Kriging metamodels in the *augmented* space $S \times \mathbb{R}^{n_s}$ (that respectively concerns the variables $\mathbf{x}$ and ξ), see [24,25] and [26]. While these approaches considered global approximations of the objective function, approaches have also been considered that sequentially construct local approximations in a trust region type framework [27,28]. Following this line, an approach integrating a trust region method for the optimization with an adaptive stochastic collocation method for uncertainty quantification is proposed in [29] and extended in [30]. In [31,26] the authors propose to use a trust region optimization framework, while constructing local Kriging surrogate models of the system output in the augmented space. Compared to these existing frameworks for optimization under uncertainty our approach differs by considering a somewhat different problem. Typical frameworks consider the optimization of the expected value or a certain quantile or any other probability measure of the objective function. On the other hand, the proposed approach seeks to approximate the objective function near its optimum for any realization of the objective function (i.e. for any instance of the uncertain parameters), allowing to construct good approximations of the probability distributions of the optimum and of its position. Our goal is thus to find an adaptive discretization of the design space, that yields an accurate approximation of the objective function near its optimum value for any possible realization, instead of an accurate approximation of just a metric on the objective function, such as the expected value or a quantile, as typically considered in previous works.

Moreover, we are particularly interested in applications in which the objective function involves multiple possible minima, implying that the probability distribution of the random variable modeling the minimum position of the objective function, i.e. $X^*(\xi)$, is multimodal. This means that the mapping from the random parameters ξ to $X^*(\xi)$ is expected to be highly non linear. Hence, direct approximation of $X^*(\xi)$ by non intrusive PCE is very difficult to achieve. This motivates the development of the proposed indirect approach devoted to the construction of an approximation of the random field modeling the objective function $Y(\mathbf{x}, \xi)$.

It is also worth mentioning that, in order to approximate the random variable modeling the extreme value (i.e. maximum or minimum) defined by Eq. (1) and by Eq. (2), it is necessary to approximate the random field $Y(\mathbf{x}, \xi)$ in any region since the extremum of the random field can *a priori* be located anywhere in the parametric space S . This is usually achieved by discretizing the random field (i.e. performing the uncertainty quantification at several points $\mathbf{x}^{(j)}$ of the parametric space S). Various discretization approaches have been proposed in the literature and the reader is referred to [32] for a review. Among the most popular approaches is the KL decomposition (see [33] for a review on numerical methods to compute this decomposition). A common point with all these approaches is that they generally try to reach the best possible approximation over the whole parametric space S . However, in the context of our study (determining the random variables in Eqs. (1) and (2)), one is only interested in having a good approximation of the random field in the areas of S where the minimum of the random field is likely to be.

In this respect, the aim of the present paper is to propose two different adaptive random field approximations, both based on PCE and Kriging. It should be noted that hybrid approaches between PCE and other approximation methods have already been proposed, for example in [34] in which PCE is combined with perturbation method and in [35] in which polynomials of the PCE are used as regression basis in universal Kriging metamodels. The idea of the two approaches proposed in this article is to handle the dependency with respect to the random variable ξ by PCE whereas the dependency with respect to the model parameter $\mathbf{x}$ is handled by Kriging. Indeed, in the context of our study (parametric uncertainties in parametrized engineering problems) it is assumed that the sensitivity of the model response with respect to the random parameters is linear or slightly non linear (the range of uncertainty is assumed small compared to the one of the design parameter, which is a reasonable hypothesis for the type of engineering problems we are interested in). Hence PCE is well suited to deal with approximation of the random variables $\forall \mathbf{x} \in S, Y(\mathbf{x}, \xi)$. Concerning the sensitivity of the model response with respect to the parameter $\mathbf{x}$, we assumed that it is more likely to be highly non linear as the range of variation of these parameters may be very large (exploration of a design space for example). Consequently, Kriging will be used for interpolation with respect to $\mathbf{x}$. Moreover, among the different interpolation approaches able to deal with non linear functions, Kriging is retained because the Gaussian process formulation is a keystone of the enrichment criterion proposed in this article. Indeed, the originality of the approaches we propose resides in their adaptivity, meaning that the approximations are iteratively constructed

to be accurate only in areas likely to highly contribute to the determination of the random variable of the extremum (Eq. (1)). Hence, the proposed approach can be seen as an *extreme value oriented* random field discretization method. Once the approximation of the random field is obtained, extreme value statistics of the random field are obtained, at a reasonable numerical cost, by Monte Carlo sampling on the approximation. It should be noted that, even if the article focuses on estimation of the extreme value statistics, the obtained approximation of the random field can also be used for various other problems, in which accuracy with respect to the extreme value is determinant, such as parametrized reliability problems. The next section details the two proposed random field approximations used for setting up the adaptive discretization algorithm. The first approximation approach is based on an interpolation of PCE coefficients, while the second approximation approach is based on an interpolation of a KL decomposition. Section 3 describes the adaptive strategy as well as the corresponding iterative algorithm and some details on its practical implementation. Section 4 is dedicated to numerical examples and comparison between the two proposed random field approximations. Finally, Section 5 draws conclusions and perspectives.

2. Discretization of random field by hybrid PCE–Kriging approaches

2.1. Uncertainty quantification by non intrusive PCE

In the following it is assumed that uncertainty quantification (UQ) at a point $\mathbf{x} \in S$ is carried out by PCE. This section recalls some basics about this approximation method.

The PCE of the second order random variable $Y(\mathbf{x}, \boldsymbol{\xi})$ is defined as,

$$Y(\mathbf{x}, \boldsymbol{\xi}) = \sum_{i=1}^{\infty} a_i(\mathbf{x}) \phi_i(\boldsymbol{\xi}), \quad \forall \mathbf{x} \in S \quad (3)$$

where convergence is in the mean square sense (see [7] and [36]). $\{\phi_i(\boldsymbol{\xi})\}_{i=1}^{\infty}$ is a Hilbertian basis of the Hilbert space containing the response $Y(\mathbf{x}, \boldsymbol{\xi})$ and $a_i(\mathbf{x})$ are the projection of $Y(\mathbf{x}, \boldsymbol{\xi})$ on the basis vector $\phi_i(\boldsymbol{\xi})$. In practice this expansion is truncated to P terms. The simplest truncation strategy consists in keeping all the polynomials with a degree less or equal to d . This choice leads to $P = \frac{(n_s+d)!}{n_s!d!}$ where n_s denoted the stochastic dimension of the problem.

Note that in the above decomposition the PCE can be seen as a variable separation approach since the PCE coefficients a_i only depend on the parameters $\mathbf{x}$ while the basis vectors ϕ_i only depend on the random variables $\boldsymbol{\xi}$. This separability is an important contributor to the efficiency of the approaches that we propose below and is in-line with multiple recent developments also based on variables separability.

For sake of simplicity, only the Hermite polynomials of standard normal random variables will be used in this study. Note however that this does not limit the applicability of the method to Gaussian inputs as one can always use iso probabilistic transforms (see [37]) to transform the random input into standard Gaussian variables. Hence, in the following it is assumed that $\boldsymbol{\xi}$ is a vector of n_s independent standard Gaussian random variables and $\{\phi_i(\boldsymbol{\xi})\}_{i=1}^{\infty}$ the n_s variate Hermite polynomials.

Concerning the computation of the unknown coefficients $a_i(\mathbf{x})$, the use of a non intrusive approach [10] is compulsory as it is assumed that the model $\mathcal{M}$ is a *black box* function (as opposed to the intrusive approach which needs a modification of the solver of $\mathcal{M}$). The two most used non intrusive approaches are the projection and regression approaches [10]. It is well known that both strategies suffer from the so called *curse of dimensionality* which makes them computationally intractable in high stochastic dimension (larger than 10 is a common threshold). This issue has been tackled (at least partially) for both strategies, by the use of sparse grid numerical integration in the case of the projection approach ([38,39] and [40] for example) and by sparse regression in the second case [11] and [13].

One can note that PCE offers an approximation of the random field $Y(\mathbf{x}, \boldsymbol{\xi})$ by truncating the series to only P terms such that,

$$Y(\mathbf{x}, \boldsymbol{\xi}) \approx \hat{Y}(\mathbf{x}, \boldsymbol{\xi}) = \sum_{i=1}^P a_i(\mathbf{x}) \phi_i(\boldsymbol{\xi}), \quad \forall \mathbf{x} \in S.$$

Hence, with respect to the problem of extreme value statistics, it is natural to introduce the random variables,

$$\hat{Y}_{\min}(\boldsymbol{\xi}) = \min_{\mathbf{x} \in S} (\hat{Y}(\mathbf{x}, \boldsymbol{\xi})) \quad (4)$$

and

$$\hat{X}^*(\xi) = \arg \min_{x \in S} (\hat{Y}(x, \xi)). \quad (5)$$

It is worth noting that solving directly this optimization problem implies, similarly to the original problems of Eqs. (1) and (2), that the PCE approximation is computed for all x in S which is again intractable in practice, as the UQ can only be performed point by point, by sampling the model. In order to solve this problem, the next section introduces two random field approximations based on PCE and Kriging interpolation.

2.2. Random field representations by a hybrid PCE–Kriging approach

2.2.1. Kriging interpolation

Kriging has been introduced by [41] in the geostatistics field. Literature on Kriging is wide and readers could refer, for example, to [42] or [43] for a detailed presentation. In the following only the background necessary to our application will be presented, in particular, only ordinary Kriging with constant trend is presented.

Let us define a function $g : S \rightarrow \mathbb{R}$. It is assumed that this function has been sampled at n_{doe} points $\mathbf{x}^{(j)} \in S$ called a design of experiments (doe) and denoted by $S_{n_{doe}} = [\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(j)}, \dots, \mathbf{x}^{(n_{doe})}]^t$. The corresponding vector of responses is denoted by $\mathbf{g} = [g(\mathbf{x}^{(1)}), \dots, g(\mathbf{x}^{(j)}), \dots, g(\mathbf{x}^{(n_{doe})})]^t$. Kriging assumes that the function to be interpolated (g in this example) is a realization of a Gaussian process. Hence, at a point $\mathbf{x}^{(0)}$ that does not belong to the doe, Kriging prediction $\tilde{g}(\mathbf{x}^{(0)})$ of $g(\mathbf{x}^{(0)})$ is a Gaussian random variable of mean μ and standard deviation σ such that

$$\tilde{g}(\mathbf{x}^{(0)}, \eta) = \mu(\mathbf{x}^{(0)}) + \sigma(\mathbf{x}^{(0)})\eta \quad (6)$$

where η is a standard normal random variable. Expressions of μ and σ are given by

$$\mu(\mathbf{x}^{(0)}) = \beta + \mathbf{r}(\mathbf{x}^{(0)})^t \mathbf{R}^{-1}(\mathbf{g} - \mathbf{1}\beta) \quad (7)$$

and

$$\sigma^2(\mathbf{x}^{(0)}) = s^2 (1 + \mathbf{r}(\mathbf{x}^{(0)})^t \mathbf{R}^{-1} \mathbf{r}(\mathbf{x}^{(0)})) \quad (8)$$

where $\mathbf{r}(\mathbf{x}^{(0)}) = [r(\mathbf{x}^{(0)} - \mathbf{x}^{(1)}), \dots, r(\mathbf{x}^{(0)} - \mathbf{x}^{(j)}), \dots, r(\mathbf{x}^{(0)} - \mathbf{x}^{(n_{doe})})]^t$, $\mathbf{R}$ is the covariance matrix of the doe i.e. $\mathbf{R}_{ij} = r(\mathbf{x}^{(i)} - \mathbf{x}^{(j)})$, $i = 1, \dots, n_{doe}$, $j = 1, \dots, n_{doe}$ where $r(\mathbf{x}; \boldsymbol{\theta})$ is a covariance function with parameters $\boldsymbol{\theta} = [\theta_1, \dots, \theta_i, \dots, \theta_n]$ (correlation length in each direction), $\mathbf{1} = [1, \dots, 1]^t$, β and s are scalar parameters. It should be noted that a large literature is devoted to the optimal construction of Kriging models by carefully selecting the covariance function and optimizing its parameters. Readers interested in this topic are referred to [42] or [43] and references within. In the following, the covariance function is set to squared exponential

$$r(\mathbf{x}; \boldsymbol{\theta}) = s^2 \prod_{i=1}^n \exp\left(-\frac{1}{\theta_i} x_i^2\right)$$

and parameters $\boldsymbol{\theta}$, β and s are estimated by maximization of the likelihood function (classical method in the literature). These choices are quite common for Kriging, though relatively arbitrary. A full investigation of their impact was not carried out in order to lighten the article and focus it on the proposed methodology for the discretization of second order random field dedicated to the study of its extreme value statistics. Note that using some advanced methods for the selection of the covariance function and the optimization of its parameters rather than the classical one used in the following could potentially lead to a further improvement of the proposed approach.

2.2.2. Interpolation of the PCE coefficients

For the purpose of explanation it is assumed that UQ by PCE has been carried out at n_{doe} points $\mathbf{x}^{(j)}$, $j = 1, \dots, n_{doe}$ forming an initial doe. Hence, the random variables $Y(\mathbf{x}^{(j)}, \xi)$ are approximated by their truncated PCE denoted by,

$$\hat{Y}(\mathbf{x}^{(j)}, \xi) = \sum_{i=1}^P a_i(\mathbf{x}^{(j)}) \phi_i(\xi), \quad j = 1, \dots, n_{doe}. \quad (9)$$

Then, assuming that the same basis $\{\phi_i(\xi)\}_{i=1}^P$ is used at each point $\mathbf{x}^{(j)}$, $j = 1, \dots, n_{doe}$, a way to approximate the random field $\hat{Y}(\mathbf{x}, \xi)$ is to interpolate the value of the PCE coefficients $a_i(\mathbf{x})$, $i = 1, \dots, P$ over the design space $\mathcal{S}$ based on the initial doe. It should be noted that each coefficient $a_i(\mathbf{x})$ is a deterministic function from $\mathcal{S}$ to $\mathbb{R}$ and that fitting the Kriging models for each $a_i(\mathbf{x})$ presents no particular difficulty. According to the description of Kriging interpolation in Section 2.2.1, at a point $\mathbf{x}^{(0)}$ that does not belong to the doe, Kriging prediction $\tilde{a}_i(\mathbf{x}^{(0)})$ of $a_i(\mathbf{x}^{(0)})$ is a Gaussian random variable of mean $\mu_{\tilde{a}_i}(\mathbf{x}^{(0)})$ and standard deviation $\sigma_{\tilde{a}_i}(\mathbf{x}^{(0)})$ such that $\tilde{a}_i(\mathbf{x}^{(0)}, \eta_i) = \mu_{\tilde{a}_i}(\mathbf{x}^{(0)}) + \sigma_{\tilde{a}_i}(\mathbf{x}^{(0)})\eta_i$ where η_i is a standard normal random variable. Expressions of $\mu_{\tilde{a}_i}(\mathbf{x}^{(0)})$ and $\sigma_{\tilde{a}_i}(\mathbf{x}^{(0)})$ are obtained similarly to Eq. (7) and Eq. (8) respectively.

Hence, the approximation of the PCE of the random field $Y(\mathbf{x}, \xi)$ reads,

$$\tilde{Y}(\mathbf{x}, \xi, \boldsymbol{\eta}) = \sum_{i=1}^P \tilde{a}_i(\mathbf{x}, \eta_i) \phi_i(\xi), \quad \forall \mathbf{x} \in \mathcal{S} \quad (10)$$

where $\boldsymbol{\eta} = [\eta_1, \dots, \eta_i, \dots, \eta_P]^T$ is a random vector of P independent standard normal random variables (cf. Eq. (6)). These random variables represent the Kriging uncertainty, i.e. the uncertainty related to the fact that for each coefficient $a_i(\mathbf{x})$ a Kriging approximation is one possible realization of a conditioned random field. In the following this representation will be denoted by **PCE-KG**.

At this point, some remarks can be done.

- It is assumed that the same basis vectors $\{\phi_i(\xi)\}_{i=1}^P$ are used at each point $\mathbf{x}^{(j)}$, $j = 1, \dots, n_{doe}$, in order to have a meaningful approximation of the PCE coefficients by Kriging models. This assumption reduces the adaptability of the method to enhance PCE approaches such as adaptive basis [8] or sparse basis [11] and [13]. This point, and some possible solutions, will be discussed in Section 5.
- One can note that independence of the PCE coefficients is assumed for the Kriging interpolation. Indeed, P independent Kriging models are built (one per PCE coefficients). A more rigorous way to achieve the interpolation could be to consider the function from $\mathcal{S}$ into $\mathbb{R}^n$ and to build a n -values Kriging model (this allows to use the information of dependence between the PCE coefficients). Nevertheless, building such a Kriging model increases significantly the computational efforts as the whole correlation structure must be determined. Hence, the independence simplification is made in order to keep the computational cost low (moreover, in this case the Kriging interpolations can be built in parallel). Note that this assumption does not have any detrimental effect on the quality of the approximations.
- One can note that the conditional random variable,

$$\tilde{Y}(\mathbf{x}, \xi, \boldsymbol{\eta})|_{\xi=\xi^{(0)}} = \tilde{Y}(\mathbf{x}, \xi^{(0)}, \boldsymbol{\eta}) = \sum_{i=1}^P \tilde{a}_i(\mathbf{x}, \eta_i) \phi_i(\xi^{(0)}), \quad \forall \mathbf{x} \in \mathcal{S}$$

is a linear combination of independent Gaussian random variables and thus is itself a Gaussian random variable of mean

$$\mu_{\tilde{Y}}(\mathbf{x}, \xi^{(0)}) = \sum_{i=1}^P \mu_{\tilde{a}_i}(\mathbf{x}) \phi_i(\xi^{(0)}), \quad \forall \mathbf{x} \in \mathcal{S}$$

and variance

$$\sigma_{\tilde{Y}}^2(\mathbf{x}, \xi^{(0)}) = \sum_{i=1}^P \sigma_{\tilde{a}_i}^2(\mathbf{x}) \phi_i^2(\xi^{(0)}), \quad \forall \mathbf{x} \in \mathcal{S}.$$

This last point will be particularly important for the practical implementation of the adaptive discretization method proposed in Section 3. Finally, one can note that, in this direct interpolation approach, the number of Kriging interpolations to build is equal to the size of the polynomial basis P used in the PCE approximation. As this number grows rapidly with the polynomial degree d and the stochastic dimension of the problem n_s , we introduced in the next subsection a second random field representation allowing, in some situations, a reduction of the number of Kriging interpolations to build.

2.2.3. Karhunen Loève decomposition, interpolation of the mean and eigenvectors

In the following we are interested in the KL decomposition of $Y(\mathbf{x}, \xi)$ which is assumed to be a real valued continuous second order random field. We denote by $\mu_Y(\mathbf{x}) : S \rightarrow \mathbb{R}$ its mean function, and by $k(\mathbf{x}, \mathbf{x}') : S \times S \rightarrow \mathbb{R}$ its covariance function. Then, the KL decomposition stands that $\forall \mathbf{x} \in S$,

$$Y(\mathbf{x}, \xi) = \mu_Y(\mathbf{x}) + \sum_{i=1}^{\infty} \gamma_i(\xi) \sqrt{\lambda_i} \phi_i(\mathbf{x}) \quad (11)$$

where convergence is in the mean square sense. γ_i are zero mean, unit variance, mutually uncorrelated random variables equal to,

$$\gamma_i(\xi) = \frac{1}{\sqrt{\lambda_i}} \int_S (Y(\mathbf{x}, \xi) - \mu_Y(\mathbf{x})) \phi_i(\mathbf{x}) d\mathbf{x}. \quad (12)$$

λ_i and ϕ_i are respectively the eigenvalue and the eigenvectors of the covariance function $k(\mathbf{x}, \mathbf{x}')$ solving the second kind Fredholm equation,

$$\int_S k(\mathbf{x}, \mathbf{x}') \phi_i(\mathbf{x}') d\mathbf{x}' = \lambda_i \phi_i(\mathbf{x}). \quad (13)$$

In practice, Eq. (13) is solved by numerical integration methods (see [33] for a review of these methods). Concerning the random variables γ_i , one could note that, in the particular case of Gaussian random fields, these random variables are standard Gaussian independent random variables. In non Gaussian cases, the distribution of the random variables γ_i is numerically approximated thanks to Eq. (12) (for example [3] uses Gaussian kernels).

In the particular context of UQ by PCE, the continuous second order random field $Y(\mathbf{x}, \xi)$ is approximated by $\hat{Y}(\mathbf{x}, \xi) = \sum_{i=1}^P a_i(\mathbf{x}) \phi_i(\xi)$, $\forall \mathbf{x} \in S$. This random field admits the following KL decomposition,

$$\hat{Y}(\mathbf{x}, \xi) = \mu_{\hat{Y}}(\mathbf{x}) + \sum_{j=1}^{\infty} \hat{\gamma}_j(\xi) \sqrt{\hat{\lambda}_j} \hat{\phi}_j(\mathbf{x}), \quad \forall \mathbf{x} \in S.$$

In [44] it is shown that PCE provides a natural way to achieve a KL decomposition. The authors use this approach to reduce the dimension of the stochastic problem for the resolution of non linear coupled problems. Below we recall the feature of the method proposed in [44] that is relevant in the present work. As in the previous Section 2.2.2 it is assumed that the UQ has been performed by PCE at an initial doe of size n_{doe} and we denote by $\mathbf{a}_i = [a_i(\mathbf{x}^{(1)}), \dots, a_i(\mathbf{x}^{(j)}), \dots, a_i(\mathbf{x}^{(n_{doe})})]^t$, $i = 1, \dots, P$ the vectors whose components are the PCE coefficients a_i computed at the doe. In the following we are interested in the KL decomposition of the random vector $\hat{Y}_{n_{doe}}(\xi) = [\hat{Y}(\mathbf{x}^{(1)}, \xi), \dots, \hat{Y}(\mathbf{x}^{(j)}, \xi), \dots, \hat{Y}(\mathbf{x}^{(n_{doe})}, \xi)]^t$, which reads (see [44] and Appendix A.1 for calculation details),

$$\hat{Y}_{n_{doe}}(\xi) = \mu_{\hat{Y}} + \sum_{j=1}^{n_{doe}} \left(\sum_{i=2}^P \mathbf{a}_i^t \hat{\phi}_j \phi_i(\xi) \right) \hat{\phi}_j \quad (14)$$

where $\mu_{\hat{Y}} = \mathbf{a}_1$ and $\hat{\phi}_j$, $j = 1, \dots, n_{doe}$ are the eigenvalues of the covariance matrix $\mathbf{K}_{\hat{Y}} = \sum_{i=2}^P \mathbf{a}_i \mathbf{a}_i^t$.

From Eq. (14) we propose to approximate the random field $Y(\mathbf{x}, \xi)$ by Kriging interpolation of the mean value and of the eigenvectors based on the vectors $\mu_{\hat{Y}}$ and $\hat{\phi}_j$ respectively. This leads to the following representation of the random field,

$$\tilde{Y}(\mathbf{x}, \xi, \eta) = \tilde{\mu}_{\hat{Y}}(\mathbf{x}, \eta_0) + \sum_{j=1}^{n_{doe}} \left(\sum_{i=2}^P \mathbf{a}_i^t \hat{\phi}_j \phi_i(\xi) \right) \tilde{\varphi}_j(\mathbf{x}, \eta_j), \quad \forall \mathbf{x} \in S \quad (15)$$

where $\tilde{\mu}_{\hat{Y}}(\mathbf{x}, \eta_0)$ and $\tilde{\varphi}_j(\mathbf{x}, \eta_j)$ are respectively the Kriging interpolation of the mean function $\mu_{\hat{Y}}(\mathbf{x})$ and of the eigenvectors $\hat{\phi}_j(\mathbf{x})$. Posing $\eta = [\eta_0, \dots, \eta_j, \dots, \eta_{n_{doe}}]^t$ a random vector of $n_{doe} + 1$ independent standard normal random variables, we express, based on Eq. (6), the Kriging approximations $\tilde{\mu}_{\hat{Y}}(\mathbf{x}, \eta_0) = \mu_{\tilde{\mu}_{\hat{Y}}}(\mathbf{x}) + \sigma_{\tilde{\mu}_{\hat{Y}}}(\mathbf{x}) \eta_0$ and $\tilde{\varphi}_j(\mathbf{x}, \eta_j) = \mu_{\tilde{\varphi}_j}(\mathbf{x}) + \sigma_{\tilde{\varphi}_j}(\mathbf{x}) \eta_j$. In the following this representation will be denoted by **full PCE-KL-KG**.

Some remarks can be made.

- Similarly to the method involving direct interpolation of the PCE coefficients (Section 2.2.2), it is assumed that the functions to interpolate are independent and $n_{doe} + 1$ Kriging models are build. In the case of the

eigenvectors this hypothesis is obviously abusive as we know that the eigenvectors are orthogonal (which involves a dependence between them). Once more, however, constructing independent approximations is not detrimental to the quality of the approximations.

- Usually, the traditional numerical integration schemes that have been proposed to solve Eq. (13) converge for all points in S . Objective of the adaptive numerical scheme proposed in the following is not to achieve convergence for all points in S but, only in the areas of S where the minimum of the random field is likely to be, in order to decrease computational costs.
- One can note that in Eq. (14) the n_{doe} eigenvectors are used in order to have the exact KL decomposition of the random vector $\hat{Y}_{n_{doe}}(\xi)$. However, in some cases the KL decomposition can be truncated by keeping only the M highest eigenvectors (which is the idea of classical dimensional reduction approaches (e.g. [44])). In practice the eigenvectors to keep in the truncated decomposition are selected by choosing the M vectors with the highest eigenvalue such that

$$\frac{\sum_{i=1}^M \hat{\lambda}_k^{n_{doe}}}{\sum_{i=1}^{n_{doe}} \hat{\lambda}_k^{n_{doe}}} = 1 - \epsilon_{KL} \quad (16)$$

where $\hat{\lambda}_1^{n_{doe}} \geq \hat{\lambda}_k^{n_{doe}} \geq \hat{\lambda}_{n_{doe}}^{n_{doe}}$. In practice this remark is very important in our methodology as it allows to decrease the number of Kriging models to build from $n_{doe} + 1$ in the complete case to $M + 1$ in the truncated case. Numerical experiments in Section 4 will present examples that emphasize the practical interest of this truncation. This representation will be hence denoted by **truncated PCE-KL-KG**.

- Finally, as for the case of interpolation of the PCE coefficients (Section 2.2.2), one can note that the conditional random variable,

$$\tilde{Y}(\mathbf{x}, \xi, \eta)|_{\xi=\xi^{(0)}} = \tilde{Y}(\mathbf{x}, \xi^{(0)}, \eta) = \tilde{\mu}_{\tilde{Y}}(\mathbf{x}, \eta_0) + \sum_{j=1}^{n_{doe}} \left(\sum_{i=1}^P \mathbf{a}_i^t \hat{\phi}_j \phi_i(\xi^{(0)}) \right) \tilde{\varphi}_j(\mathbf{x}, \eta_j), \quad \forall \mathbf{x} \in S$$

is a Gaussian random variable of mean,

$$\mu_{\tilde{Y}}(\mathbf{x}, \xi^{(0)}) = \mu_{\tilde{\mu}_{\tilde{Y}}}(\mathbf{x}) + \sum_{j=1}^{n_{doe}} \left(\sum_{i=1}^P \mathbf{a}_i^t \hat{\phi}_j \phi_i(\xi^{(0)}) \right) \mu_{\tilde{\varphi}_j}(\xi^{(0)}), \quad \forall \mathbf{x} \in S$$

and variance,

$$\sigma_{\tilde{Y}}^2(\mathbf{x}, \xi^{(0)}) = \sigma_{\tilde{\mu}_{\tilde{Y}}}^2(\mathbf{x}) + \sum_{j=1}^{n_{doe}} \left(\sum_{i=1}^P \mathbf{a}_i^t \hat{\phi}_j \phi_i(\xi^{(0)}) \right)^2 \sigma_{\tilde{\varphi}_j}^2(\xi^{(0)}), \quad \forall \mathbf{x} \in S.$$

This is again of great practical interest for efficient computations within the adaptive approach we will propose in Section 3.

2.2.4. Discussion on the proposed random field discretizations

It should be noted that both random field discretizations $\tilde{Y}(\mathbf{x}, \xi, \eta)$, defined by Eq. (10) and Eq. (15), depend on the *primary uncertainty* of the problem, modeled by the random vector ξ , and on the *Kriging uncertainty* due to the use of Kriging interpolations and modeled by the random vector η .

However, it is interesting to note a major difference between the two proposed random field representations that concerns the number of Kriging models to build and thus, the numerical cost of the method. Table 1 recalls the number of Kriging metamodels to be build for each random representation introduced. In the case of *PCE – KG* representation (Eq. (10)), the number of Kriging interpolations to build is equal to the size P of the polynomial basis and we recall that $P = \frac{(n_s+d)!}{n_s!d!}$. Hence, it is independent of the dimension of the space S and of the size n_{doe} of the doe. In the case of *PCE – KL – KG* representation (Eq. (15)) the number of Kriging metamodels to build is equal to $n_{doe} + 1$ in its full version or to $M + 1$ in its truncated version. Hence, it depends on the correlation structure of the random field under study, thus it implicitly also depends on the size of the space S . It is clear that for a given problem,

Table 1

Number of Kriging metamodels to build according to the random field representation used. n_s is the stochastic dimension of the problem, d is the largest polynomial degree used in the PCE, n_{doe} is the size of the doe and M is the number of modes kept according to Eq. (16).

	PCE – KG	Full PCE – KL – KG	Truncated PCE – KL – KG
Nb of Kriging metamodels	$P = \frac{(n_s+d)!}{n_s!d!}$	$n_{doe} + 1$	$M + 1$

depending on its stochastic dimension n_s , the degree d of the polynomial chaos used, the dimension of the space $\mathcal{S}$ and the correlation structure of the random field, the use of one of these two approaches would be more advantageous over the other. The applications section will present some examples in order to study the numerical cost of both approaches and will try to define some good practices in order to choose *a priori* the best strategy with respect to accuracy and numerical cost.

As both representation methods are based on Kriging interpolations, the next section presents an original development that takes advantage of the uncertainty due to Kriging interpolation in order to adaptively improve the doe until the accuracy on the random variable of interest $\hat{Y}_{min}(\xi)$ reaches a given level.

3. Iterative discretization of the random field

3.1. Discretization of $\hat{Y}_{min}(\xi)$

In this section it is still assumed that the UQ has been carried out at n_{doe} points, forming an initial doe denoted by $S_{n_{doe}} \in \mathcal{S}$. We propose an adaptive doe enrichment strategy aimed at rapidly improving the characterization of the random variable of interest $Y_{min}(\xi)$.

Let us first introduce the random variable $Y_{min}|_{x(S_{n_{doe}})}(\xi)$ which is the random variable modeling the minimum value of the random vector

$$[Y(\mathbf{x}^{(1)}, \xi), \dots, Y(\mathbf{x}^{(j)}, \xi), \dots, Y(\mathbf{x}^{(n_{doe})}, \xi)]^t, \text{ i.e.,}$$

$$Y_{min}(\xi)|_{x(S_{n_{doe}})} = \min_{x \in S_{n_{doe}}} (Y(\mathbf{x}, \xi)).$$

According to the context of our study (UQ by PCE), we define the random variable $\hat{Y}_{min}|_{x(S_{n_{doe}})}(\xi)$ which is the random variable modeling the minimum value of the random vector $[\hat{Y}(\mathbf{x}^{(1)}, \xi), \dots, \hat{Y}(\mathbf{x}^{(j)}, \xi), \dots, \hat{Y}(\mathbf{x}^{(n_{doe})}, \xi)]^t$, i.e.,

$$\hat{Y}_{min}(\xi)|_{x(S_{n_{doe}})} = \min_{x \in S_{n_{doe}}} (\hat{Y}(\mathbf{x}, \xi)). \quad (17)$$

This random variable can be interpreted as an approximation of $\hat{Y}_{min}(\xi)$ based on a finite number of sampling location $\mathbf{x}^{(j)}$, $j = 1, \dots, n_{doe}$ where the UQ has been carried out by PCE. Based on the definition given by Eq. (17), the idea of the method proposed in this article is to find the next sampling location $\mathbf{x}^{(new)}$ where the UQ should be performed in order to increase our knowledge of the random variable $\hat{Y}_{min}(\xi)$ and thus of $Y_{min}(\xi)$. To this purpose next part is devoted to the definition of an enrichment criterion.

3.2. Enrichment criterion

In the literature devoted to global optimization of deterministic *black box* functions based on Kriging metamodels initiated by Jones et al. [20], an enrichment criterion refers to a criterion that allows to choose the next point where the *black box* function should be evaluated in order to bring the most relevant information about the optimum of the deterministic function. An efficient enrichment criterion must be able to find a compromise between *exploration* (i.e. increasing the accuracy of the Kriging metamodel over the whole parametric space so that no localized minimum can stay invisible to the metamodel) and *exploitation* (i.e. increasing the accuracy of the metamodel only where the global minimum is susceptible to be). Several enrichment criteria have been proposed and compared in the literature (see [45]).

Note that none of the existing criteria is applicable in the context of our study since we are not interested here in determining the optimum of a deterministic function but we are rather interested in determining the distribution of the optimum of a random field (in the sense of Eq. (2)). In our context, we thus propose a new criterion inspired from the

pioneer criterion defined in [20] and called Expected Improvement (*EI*). In order to distinguish it from the classical *EI* that deals with deterministic optimization we denote our criterion by EI_{RF} where the subscript *RF* stands for Random Field.

According to the approximations defined in the previous sections, we define the EI_{RF} by, $\forall \mathbf{x} \in \mathcal{S}$

$$EI_{RF}(\mathbf{x}) = \mathbb{E} \left[\left(\hat{Y}_{min}(\xi) |_{\mathbf{x}(S_{ndoe})} - \tilde{Y}(\mathbf{x}, \xi, \eta) \right) \mathbb{1}_{\tilde{Y}(\mathbf{x}, \xi, \eta) \leq \hat{Y}_{min}(\xi) |_{\mathbf{x}(S_{ndoe})}} \right] \quad (18)$$

where

$$\mathbb{1}_{\tilde{Y}(\mathbf{x}, \xi, \eta) \leq \hat{Y}_{min}(\xi) |_{\mathbf{x}(S_{ndoe})}} = 0 \text{ if } \tilde{Y}(\mathbf{x}, \xi, \eta) > \hat{Y}_{min}(\xi) |_{\mathbf{x}(S_{ndoe})} \text{ and}$$

$$\mathbb{1}_{\tilde{Y}(\mathbf{x}, \xi, \eta) \leq \hat{Y}_{min}(\xi) |_{\mathbf{x}(S_{ndoe})}} = 1 \text{ if } \tilde{Y}(\mathbf{x}, \xi, \eta) \leq \hat{Y}_{min}(\xi) |_{\mathbf{x}(S_{ndoe})}.$$

One can note that $EI_{RF}(\mathbf{x})$ is positive for $\mathbf{x} \notin S_{ndoe}$ and that $EI_{RF}(\mathbf{x}) = 0$ if $\mathbf{x} \in S_{ndoe}$.

The point $\mathbf{x}^{(new)}$ where the UQ should be performed is thus solution of the optimization problem,

$$\mathbf{x}^{(new)} = \arg \max_{\mathbf{x} \in \mathcal{S}} (EI_{RF}(\mathbf{x})). \quad (19)$$

It can be noticed that this criterion still realizes a compromise between *exploitation* and *exploration*. Indeed, at a given point $\mathbf{x}^{(0)} \in \mathcal{S}$, a positive value of the Expected Improvement (as defined by Eq. (18)) could have two reasons:

- The uncertainty due to the Kriging interpolation (modeled by the random vector η) increases the dispersion of the random variable $\tilde{Y}(\mathbf{x}^{(0)}, \xi, \eta)$, thus the random event $(\hat{Y}_{min}(\xi) |_{\mathbf{x}(S_{ndoe})} - \tilde{Y}(\mathbf{x}^{(0)}, \xi, \eta))$ may have a non negligible probability. This helps the *exploration* of the parametric space.
- If the uncertainty associated to the Kriging interpolation is low (compared to the one due to ξ), then a positive value of $EI_{RF}(\mathbf{x}^{(0)})$ means that the random variable $\tilde{Y}(\mathbf{x}^{(0)}, \xi, \eta) \approx \hat{Y}(\mathbf{x}^{(0)}, \xi)$ contributes significantly to the random variable $Y_{min}(\xi)$. This corresponds to the *exploitation* of the model (defined by Eq. (10) or Eq. (15)).

By considering the expectation over the whole random space (ξ and η) the criterion defined by Eq. (18) automatically realizes a compromise between *exploration* and *exploitation*. Obviously, this criterion is heuristic and its efficiency will be investigated on several numerical examples in Section 4.

One can note that the Expected Improvement criterion defined by Eq. (22) could be adapted for *deterministic* interpolation such as Radial Basis Function [46]. In that case the approximation of the random field $\tilde{Y}$ will no longer depends on η but only on ξ , nevertheless the criterion EI_{RF} can still be computed (the expectation is, in that simple case, only with respect to ξ). However, by taking into account interpolation uncertainty (modeled by the random variables η), Kriging will help the exploration of the parametric domain $\mathcal{S}$ and increases the chance of finding the various areas of the parametric domain $\mathcal{S}$ where the minimum is likely to be. On the contrary, a *deterministic* interpolation approach will put too much confidence in the exploitation of the approximation $\tilde{Y}$ which can lead to a poor approximation in some important areas of the parametric domain $\mathcal{S}$. This behavior will be illustrated on numerical examples in Section 4.

The next section discusses the practical computation of the EI_{RF} defined by Eq. (18) and describes the proposed extreme value oriented random field discretization algorithm.

3.3. Implementation and proposed algorithm

The aim of the proposed algorithm is to add points iteratively to an initial doe until a given convergence criterion is reached. Hence, from an initial doe of size n_{doe} , the criterion defined by Eq. (18) is computed using one of the two random field representations presented in Sections 2.2.2 and 2.2.3. Maximization of this criterion gives the next point $\mathbf{x}^{(new)}$ where the UQ should be performed in order to improve our knowledge on the random variable $\hat{Y}_{min}$. Then, $\mathbf{x}^{(new)}$ is added to the doe and the UQ by PCE is performed in order to compute $\hat{Y}(\mathbf{x}^{(new)}, \xi)$. Adding this new information, the random field representation is updated and the previous steps are repeated until convergence.

The keystone of the proposed algorithm is the optimization of the EI_{RF} defined by Eq. (18). In the literature devoted to optimization of deterministic black box function this criterion can be evaluated analytically which makes its optimization relatively easy to solve (see [20]). In our context, the computation of the expected value Eq. (18) could not be derived analytically. Hence, a direct approach will be to estimate it by Monte Carlo sampling. It should be noted that Eq. (18) only involves polynomial functions with a very low computational cost which makes such an approach

tractable; in particular no call to the expensive simulator $\mathcal{M}$ is required for the computation of the EI_{RF} of Eq. (18). Furthermore, by using the fact that the conditional random variables $\tilde{Y}(\mathbf{x}, \boldsymbol{\xi}, \boldsymbol{\eta})|_{\boldsymbol{\xi}=\boldsymbol{\xi}^{(0)}}$, $\forall \mathbf{x} \in \mathcal{S}$ are Gaussian random variables (see remarks at the end of Section 2.2.2 and Section 2.2.3) one can simplify the estimation of Eq. (18). Indeed, we recall that, for both decompositions, the random vectors $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are independent. Then, we denote by $E_{\boldsymbol{\xi}}$ the expected value with respect to the random variable $\boldsymbol{\xi}$, and by $E_{\boldsymbol{\eta}}$ the expected value with respect to the random variable $\boldsymbol{\eta}$. Using these notations, Eq. (18) can be rewritten, $\forall \mathbf{x} \in \mathcal{S}$

$$EI_{RF}(\mathbf{x}) = E_{\boldsymbol{\xi}} \left[E_{\boldsymbol{\eta}} \left[\left(\hat{Y}_{min}(\boldsymbol{\xi})|_{\mathbf{x}(S_{ndoe})} - \tilde{Y}(\mathbf{x}, \boldsymbol{\xi}, \boldsymbol{\eta}) \right) \mathbb{1}_{\tilde{Y}(\mathbf{x}, \boldsymbol{\xi}, \boldsymbol{\eta}) \leq \hat{Y}_{min}(\boldsymbol{\xi})|_{\mathbf{x}(S_{ndoe})}} \right] \right]. \quad (20)$$

By estimating the expected value $E_{\boldsymbol{\xi}}$ by the Monte Carlo method with n_{sim} simulations in Eq. (20), one gets, $\forall \mathbf{x} \in \mathcal{S}$

$$EI_{RF}(\mathbf{x}) \approx \frac{1}{n_{sim}} \sum_{k=1}^{n_{sim}} E_{\boldsymbol{\eta}} \left[\left(\hat{Y}_{min}(\boldsymbol{\xi}^{(k)})|_{\mathbf{x}(S_{ndoe})} - \tilde{Y}(\mathbf{x}, \boldsymbol{\xi}^{(k)}, \boldsymbol{\eta}) \right) \mathbb{1}_{\tilde{Y}(\mathbf{x}, \boldsymbol{\xi}^{(k)}, \boldsymbol{\eta}) \leq \hat{Y}_{min}(\boldsymbol{\xi}^{(k)})|_{\mathbf{x}(S_{ndoe})}} \right] \quad (21)$$

in which $\hat{Y}_{min}(\boldsymbol{\xi}^{(k)})|_{\mathbf{x}(S_{ndoe})}$ is a deterministic value obtained by evaluating Eq. (17) and $\tilde{Y}(\mathbf{x}, \boldsymbol{\xi}^{(k)}, \boldsymbol{\eta}) = \tilde{Y}(\mathbf{x}, \boldsymbol{\xi}, \boldsymbol{\eta})|_{\boldsymbol{\xi}=\boldsymbol{\xi}^{(k)}}$ is a Gaussian random variable of mean $\mu_{\tilde{Y}}(\mathbf{x}, \boldsymbol{\xi}^{(k)})$ and variance $\sigma_{\tilde{Y}}^2(\mathbf{x}, \boldsymbol{\xi}^{(k)})$ (expressions of $\mu_{\tilde{Y}}(\mathbf{x}, \boldsymbol{\xi}^{(k)})$ and $\sigma_{\tilde{Y}}^2(\mathbf{x}, \boldsymbol{\xi}^{(k)})$ depend on the random field representation and are given in Section 2.2.2 for the *PCE – KG* representation and in Section 2.2.3 for the *PCE – KL – KG* representation). Finally, one can note that the expected value with respect to $\boldsymbol{\eta}$ in Eq. (21) is very close to the expression of the *EI* in the context of deterministic optimization and it can be shown that,

$$\begin{aligned} E_{\boldsymbol{\eta}} \left[\left(\hat{Y}_{min}(\boldsymbol{\xi})|_{\mathbf{x}(S_{ndoe})} - \tilde{Y}(\mathbf{x}, \boldsymbol{\xi}, \boldsymbol{\eta}) \right) \mathbb{1}_{\tilde{Y}(\mathbf{x}, \boldsymbol{\xi}, \boldsymbol{\eta}) \leq \hat{Y}_{min}(\boldsymbol{\xi})|_{\mathbf{x}(S_{ndoe})}} \right] = \\ (\hat{Y}_{min}(\boldsymbol{\xi}^{(k)})|_{\mathbf{x}(S_{ndoe})} - \mu_{\tilde{Y}}(\mathbf{x}, \boldsymbol{\xi}^{(k)})) \mathcal{F} \left(\frac{\hat{Y}_{min}(\boldsymbol{\xi}^{(k)})|_{\mathbf{x}(S_{ndoe})} - \mu_{\tilde{Y}}(\mathbf{x}, \boldsymbol{\xi}^{(k)})}{\sigma_{\tilde{Y}}(\mathbf{x}, \boldsymbol{\xi}^{(k)})} \right) \\ + \sigma_{\tilde{Y}}(\mathbf{x}, \boldsymbol{\xi}^{(k)}) f \left(\frac{\hat{Y}_{min}(\boldsymbol{\xi}^{(k)})|_{\mathbf{x}(S_{ndoe})} - \mu_{\tilde{Y}}(\mathbf{x}, \boldsymbol{\xi}^{(k)})}{\sigma_{\tilde{Y}}(\mathbf{x}, \boldsymbol{\xi}^{(k)})} \right) \end{aligned} \quad (22)$$

where f and $\mathcal{F}$ are respectively the probability density function and the cumulative distribution function of the standard normal distribution.

By using Eq. (22) into Eq. (21) one gets the implementation of the Monte Carlo estimator of the EI_{RF} that is used in the following. In practice, we use a Python implementation and in order to benefit from the efficient vectorial calculation, the Monte Carlo method is performed by batch of 10 000 simulations until the coefficient of variation of estimator of the EI_{RF} is less than 5%.

Concerning the resolution of the optimization problem of Eq. (19) a global optimizer is best suited due to the property of the EI_{RF} (null at the doe point and positive elsewhere), which typically leads to many local maxima. In the applications section we rely on the Constrained Optimization BY Linear Approximation (COBYLA) algorithm [47] to solve Eq. (19). We use a multi-start approach in order to maximize the chance of finding the global maximum.

We finish this description of the proposed algorithm by discussing the possible convergence criteria. First of all, it is assumed that the user can only perform a limited number of UQ by PCE. This constraint sets the maximum number of points that can be added to the doe and thus the maximum number of iterations. Nevertheless, this criterion does not give information about the accuracy of the random field representation obtained after this maximum number of iterations. To address this aspect we propose to use the change in the maximum value of the EI_{RF} as convergence criterion. Convergence is thus defined as,

$$\frac{\max_{\mathbf{x} \in \mathcal{S}}(EI_{RF}^{(k)}(\mathbf{x}))}{\max_{\mathbf{x} \in \mathcal{S}}(EI_{RF}^{(0)}(\mathbf{x}))} \leq \epsilon_{EI_{RF}} \quad (23)$$

where $\max_{\mathbf{x} \in \mathcal{S}}(EI_{RF}^{(0)}(\mathbf{x}))$ is the maximum value of the EI_{RF} computed on the initial doe, $\max_{\mathbf{x} \in \mathcal{S}}(EI_{RF}^{(k)}(\mathbf{x}))$ is the maximum value of the EI_{RF} computed at iteration k and $\epsilon_{EI_{RF}}$ is the coefficient of reduction of the EI_{RF} used to define convergence.

Algorithm 1 summarizes the implementation of the proposed method for extreme value oriented adaptive discretization of random fields.

Initialization:

maximum number of iterations n_{max} , current iteration $n_{iter} = 0$,

convergence criterion $\epsilon_{EI_{RF}}$,

initial doe $\mathcal{S}_{n_{doe}}$ of n_{doe} points,

initial vector of PCE random responses $[\hat{Y}(\mathbf{x}^{(1)}, \boldsymbol{\xi}), \dots, \hat{Y}(\mathbf{x}^{(j)}, \boldsymbol{\xi}), \dots, \hat{Y}(\mathbf{x}^{(n_{doe})}, \boldsymbol{\xi})]^t$,

compute the new random field representation based on Eq. (10) or Eq. (15) $\rightarrow$ compute Kriging interpolations

find $\mathbf{x}^{(new)} = \arg \max_{\mathbf{x} \in \mathcal{S}} (EI_{RF}^{(0)}(\mathbf{x}))$

while convergence not achieved (Eq. (23)) $\Rightarrow$ False and $n_{iter} \leq n_{max}$ **do**

perform the UQ at $\mathbf{x}^{(new)}$

add $\mathbf{x}^{(new)}$ to the doe

add $\hat{Y}(\mathbf{x}^{(new)}, \boldsymbol{\xi})$ to the vector of PCE random responses

compute the new random field representation based on Eq. (10) or Eq. (15) $\rightarrow$ compute Kriging interpolations

$n_{iter} = n_{iter} + 1$

find $\mathbf{x}^{(new)} = \arg \max_{\mathbf{x} \in \mathcal{S}} (EI_{RF}^{(n_{iter})}(\mathbf{x}))$

end

Algorithm 1: Extreme value oriented adaptive discretization of continuous random field by an hybrid PCE–Kriging approach

The next section presents the post processing step that allows to approximate the extreme value statistics of the random field under study.

3.4. Approximation of the extreme value statistics

For the purpose of explanation it is assumed that Algorithm 1 has converged satisfying Eq. (23). Hence, the random field has been discretized at n_{doe} points $\mathbf{x}^{(j)} \in \mathcal{S}$, $j = 1, \dots, n_{doe}$. Satisfaction of the convergence criterion defined by the Eq. (23) expresses the fact that the possible improvement on our knowledge of the random variable $\hat{Y}_{min}(\boldsymbol{\xi})$ by adding a new point is negligible. This translates the fact that the random field representation $\tilde{Y}(\mathbf{x}, \boldsymbol{\xi}, \boldsymbol{\eta})$ reaches an acceptable accuracy in the areas of $\mathcal{S}$ where the extreme value is likely to be and consequently that the uncertainty due to the Kriging interpolation is negligible in these areas. Hence, we propose to simplify the random fields representation given by Eq. (10) and Eq. (15) using only the mean value of the Kriging interpolations. This leads to the following representations,

$$\tilde{Y}(\mathbf{x}, \boldsymbol{\xi}) = \sum_{i=1}^P \mu_{\tilde{a}_i}(\mathbf{x}) \phi_i(\boldsymbol{\xi}), \quad \forall \mathbf{x} \in \mathcal{S} \quad (24)$$

in the case of the *PCE – KG* representation, and to

$$\tilde{Y}(\mathbf{x}, \boldsymbol{\xi}) = \mu_{\tilde{\mu}_Y}(\mathbf{x}) + \sum_{j=1}^{n_{doe}} \left(\sum_{i=2}^P \mathbf{a}_i^t \hat{\boldsymbol{\phi}}_j \phi_i(\boldsymbol{\xi}) \right) \mu_{\tilde{\phi}_j}(\mathbf{x}), \quad \forall \mathbf{x} \in \mathcal{S} \quad (25)$$

in the case of *PCE – KL – KG* representation.

At this step, Eq. (24) (or equivalently Eq. (25)) can be considered as an emulator or a metamodel of the random field $Y(\mathbf{x}, \boldsymbol{\xi})$ optimized to be accurate in the areas where the random field is likely to reach its minimum value. Hence we propose to use it to approximate the probability distributions of $\hat{Y}_{min}(\boldsymbol{\xi})$ and $\hat{X}^*(\boldsymbol{\xi})$ (Eq. (4) and Eq. (5)) by,

$$\hat{Y}_{min}(\boldsymbol{\xi}) = \min_{\mathbf{x} \in \mathcal{S}} (\tilde{Y}(\mathbf{x}, \boldsymbol{\xi})) \quad (26)$$

and

$$\hat{X}^*(\boldsymbol{\xi}) = \arg \min_{\mathbf{x} \in \mathcal{S}} (\tilde{Y}(\mathbf{x}, \boldsymbol{\xi})). \quad (27)$$

It should be noted that random field representation given by Eq. (24) (or equivalently by Eq. (25)) only involves polynomial functions and thus makes the optimization problem Eq. (26) easy to solve by the Monte Carlo method (*i.e.* solving the minimization problem of Eq. (26) for a large number of samples $\boldsymbol{\xi}^{(j)}$, $j = 1, \dots, n_{sim}$). Moreover,

we would like to add that, for both random field representations, the partial derivatives with respect to $\mathbf{x}$ only involves the partial derivatives of the mean value of the Kriging models which makes them analytically computable. For these reasons, the optimization problem of Eq. (26) for a given sample $\xi^{(j)}$ is efficiently solved by the Sequential Least Square Programming (SLSQP) algorithm by [48] with multistart to increase the chance of finding the global optimum. Note that other global optimization algorithms could be used instead.

Next section is devoted to numerical examples studying the convergence of $\tilde{Y}_{min}(\xi)$ and $\tilde{X}^*(\xi)$ to $Y_{min}(\xi)$ and $X^*(\xi)$ respectively.

4. Numerical applications

Three numerical examples are presented in this section. The first two, presented in Section 4.1, deal with Gaussian random fields of spatial dimension 1 and 2. Gaussian random fields are first studied as they can be represented exactly on the Hermite PCE which allows to study the proposed method without any error due to the UQ by PCE. Then, a non Gaussian example involving the computation of aerodynamic loads on an aircraft wing by a numerical model is presented in Section. 4.2.

4.1. Gaussian case

4.1.1. Exponentially correlated Gaussian process

We consider the one dimensional Gaussian process defined on $S = [-2, 2]$, with mean $\mu(x) = \sin(x\pi)$ and covariance function $cov(x^{(1)}, x^{(2)}) = \sigma \exp(-\frac{|x^{(1)} - x^{(2)}|}{l})$, where l is the correlation length. It should be noted that, for this particular choice of covariance function, the KL decomposition can be computed analytically (see [5] for the solution of the Fredholm equation in this case). Thus, in the following we consider the random process defined by, $\forall x \in S$,

$$Y(x, \xi) = \mu(x) + \sum_{i=1}^{n_s} \xi_i \sqrt{\lambda_i} \varphi_i(x) \quad (28)$$

where expression of eigenvalues λ_i and eigenvectors φ_i can be found in [5] (note that $\lambda_1 > \lambda_i > \lambda_{n_s}$). The truncation n_s is chosen according to the following empiric rule $\frac{\lambda_{n_s+1}}{\sum_{i=1}^{n_s} \lambda_i} < 0.01$.

Concerning the PCE approximation, it is clear that the field defined by Eq. (28) can be represented exactly on the Hermite polynomial basis $\{\phi_i(\xi)\}_{i=1}^{n_s+1}$ made of a constant term and the n_s first order terms ξ_i , leading to,

$$\hat{Y}(x, \xi) = \sum_{i=1}^P a_i(x) \phi_i(\xi)$$

where $P = n_s + 1$, $a_1(x) = \mu(x)$ and $a_i(x) = \sqrt{\lambda_i} \varphi_i(x)$, $i = 2, \dots, P$. Hence, at a given point $x^{(j)} \in S$ the random variable $Y(x^{(j)}, \xi)$ is perfectly represented by its PCE i.e $Y(x^{(j)}, \xi) = \hat{Y}(x^{(j)}, \xi)$. Motivation for studying this particular Gaussian case is that it allows to investigate the characteristics of the proposed method without any error due to the PCE approximation.

The proposed method is investigated with respect to a correlation length equals to $l = 0.2$. According to the representation given by Eq. (28) and the previously described truncation rule, this correlation length leads to stochastic dimensions $n_s = 23$. Fig. 1 presents 5 random realizations of this field. One can note that the sinusoidal mean shape is strongly perturbed and the random realizations may have more than two local minima.

The proposed method is now applied starting from an initial doe of size $n_{doe} = 2$ such that $x^{(1)} \approx -0.23$ and $x^{(2)} \approx 0.27$ (sampled by Latin Hypercube Sampling, LHS). The two random fields representations introduced in Section 2.2.2 by Eq. (10) and in Section 2.2.3 by Eq. (15) are compared. Note that the $PCE - KL - KG$ representation is used in its full version for this first application. The stopping criterion defined by Eq. (23) is set to $\epsilon_{ELRF} = 0.01$. Fig. 2 presents the final doe (points on the x -axis) as well as a reference random realization and its approximations by the two random field representations.

In Fig. 2 it is notable that both random field representations lead to very close final doe with respect to the number of added points and their positions. Moreover, one can note that the adaptive strategy performed well in this example as Algorithm 1 focuses the exploration around the two regions of S where the minimum of the process is likely to

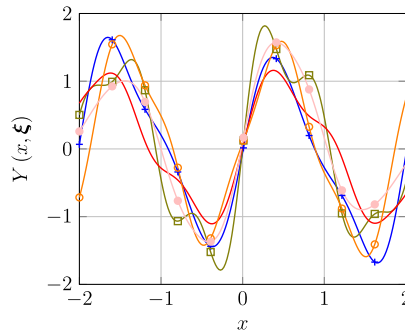


Fig. 1. Random realizations of the exponentially correlated random process defined by Eq. (28) for correlation length $l = 0.2$.

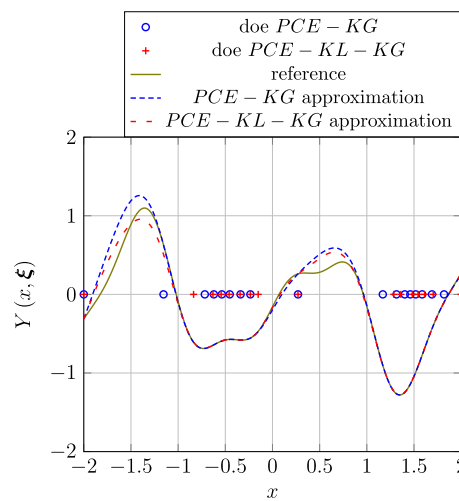


Fig. 2. Final doe obtained with the proposed adaptive approach using both random field representation ($PCE - KG$ and full $PCE - KL - KG$) with $\epsilon_{EIRF} = 0.01$. A comparison with a single reference random realization of the process is also presented in order to illustrate the interest of the adaptive doe.

Table 2

Size of the final doe obtained by the proposed adaptive approach with $\epsilon_{EI} = 10^{-2}$, with respect to the random field representation used.

	$PCE - KG$	Full $PCE - KL - KG$
Final n_{doe}	17	16

be (neighborhood of $x = -0.5$ and $x = 1.5$). Table 2 gives the size of the final doe with respect to the random field representation used. First, one can note that the choice of the random field representation has almost no influence on the final doe size. One can also note in Fig. 2 that the random trajectory is poorly approximated in the regions of S where it is likely to reach high values but it is perfectly approximated in the two regions where it reaches its minimum values. This case illustrates the philosophy of our approach which is to focus the computational effort in the region of S where the minimum is likely to be.

Fig. 2 presents an illustration on a single random trajectory. However, to be efficient in the approximation of the probability distributions of $Y_{min}(\xi)$ and $X^*(\xi)$, the proposed approach must be efficient for every random realization (over the whole stochastic space). In order to address the global accuracy of the proposed method two mean relative

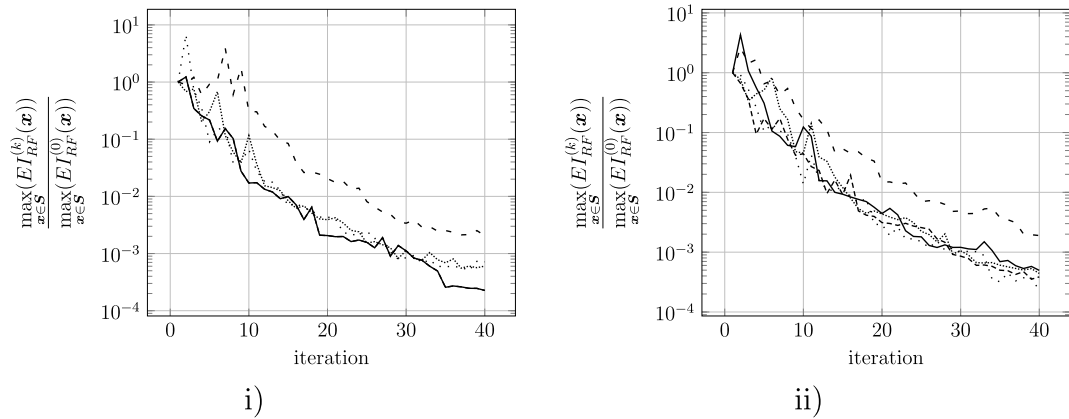


Fig. 3. Evolution of the convergence criterion defined by Eq. (23) on 40 iterations. 5 replications are presented. (i) *PCE – KG* representation, (ii) *PCE – KL – KG* representation.

errors are introduced,

$$Err_{Y_{min}} = E \left[\left| \frac{Y_{min}(\xi) - \tilde{Y}_{min}(\xi)}{Y_{min}(\xi)} \right| \right] \quad (29)$$

and

$$Err_{X^*} = E \left[\left| \frac{X^*(\xi) - \tilde{X}^*(\xi)}{X^*(\xi)} \right| \right]. \quad (30)$$

where $Y_{min}(\xi)$ and $X^*(\xi)$ are the reference values and $\tilde{Y}_{min}(\xi)$ and $\tilde{X}^*(\xi)$ are the approximations obtained by optimization of the proposed random field approximations (see Eq. (26) and Eq. (27)). Estimations of these two quantities will be evaluated by Monte Carlo sampling of 1000 simulations.

First of all, Fig. 3 presents the evolution of the convergence criterion defined by Eq. (23) on 40 iterations. As this criterion depends on the initial doe, 5 replications are presented in Fig. 3.

Several remarks can be made on Fig. 3. First, one can note that, the convergence of the proposed algorithm is relatively independent of the method retained for the random field representation. Indeed, results obtained by the *PCE – KG* representation (Fig. 3(i)) and by the *PCE – KL – KG* representation (Fig. 3(ii)) show that the convergence criterion decreases with the same tendency for both random field representations. Second, even if the convergence criterion is globally decreasing when the doe size increases, this convergence is not monotonic. This behavior is expected and it is due to the exploration of the parametric domain S by the proposed Algorithm 1. Finally, concerning the dispersion with respect to the initial doe, for both random field representations, the 5 replicates show the same global behavior which illustrates the robustness of the approach on this example. However, one can note that, one replicate is clearly shifted compared to the 4 others, this shows that in a case of *bad* initial doe, the proposed algorithm may need some more iterations but still converges.

In order to access the global accuracy of the proposed method over the whole random space, Figs. 4 and 5 present the estimation of the relative errors defined by Eq. (29) and by Eq. (30) respectively. Results are given for the two random field representations and for 3 values of the convergence criterion $\epsilon_{EIRF} = 10^{-1}$, 10^{-2} , 10^{-3} . Concerning the initial doe, the same 5 replicates as in Fig. 3 are used.

First of all, one can note in Figs. 4 and 5 that decreasing the convergence criterion allows to reduce the dispersion between the five replicates and globally increase the accuracy of the method (clearly visible between $\epsilon_{EIRF} = 10^{-1}$ and $\epsilon_{EIRF} = 10^{-2}$ and only a slight improvement by decreasing it to $\epsilon_{EIRF} = 10^{-3}$). From a quantitative point of view one can note that with $\epsilon_{EIRF} = 10^{-2}$ both methods lead to very low relative errors (less than 0.5% on the value of the minimum and around 2.5% on its position). In order to illustrate the link between those results and some statistical results such as the probability density function (PDF), the Fig. 6 presents the approximations of the PDF of $Y_{min}(\xi)$ and $X^*(\xi)$ compared to a reference PDF computed by Monte Carlo method (all the PDF curves are obtained by kernel

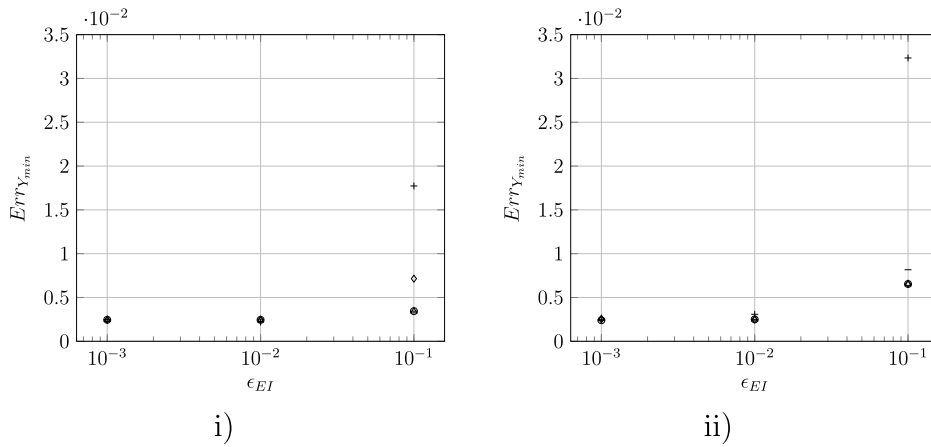


Fig. 4. Relative error on the value of the minimum defined by the Eq. (29) for 3 different values of the convergence criterion $\epsilon_{EI} = 10^{-3}$, $\epsilon_{EI} = 10^{-2}$, $\epsilon_{EI} = 10^{-1}$. 5 replications are presented, each one with a different symbol. (i) PCE-KG (ii) PCE-KL-KG.

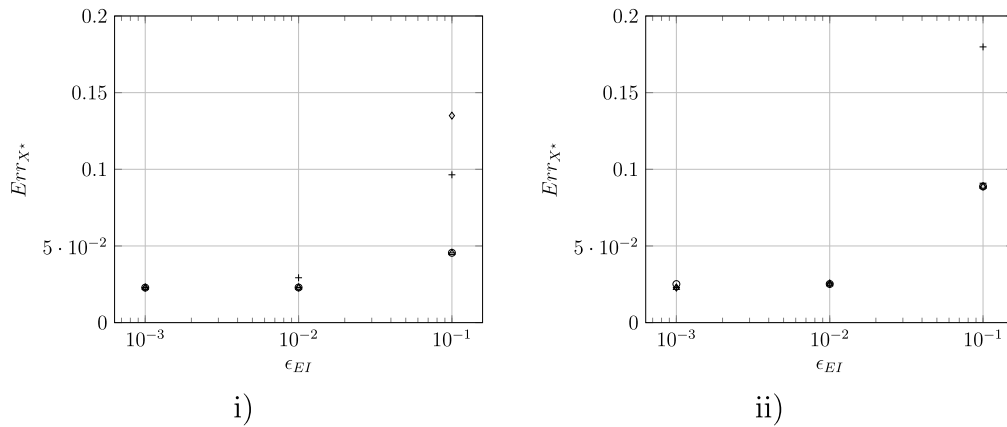


Fig. 5. Relative error on the position of the minimum defined by the Eq. (30) for 3 different values of the convergence criterion $\epsilon_{EI} = 10^{-3}$, $\epsilon_{EI} = 10^{-2}$, $\epsilon_{EI} = 10^{-1}$. 5 replications are presented, each one with a different symbol. (i) PCE-KG (ii) PCE-KL-KG.

smoothing, with standard normal kernel, over 1000 simulations). For both approaches, Fig. 6 presents results obtained with $\epsilon_{EIRF} = 10^{-2}$.

Concerning the approximation of Y_{min} (Fig. 6(i)) one can note that both approaches perfectly approximated the reference results (curves are almost superposed). Only, slight differences are visible at the upper tail of the PDF. It is also interesting to note that the shape of the PDF is non Gaussian which shows the capability of the proposed method to handle non Gaussian distributions. Fig. 6(ii) presents the approximations of the PDF of X^* compared to the reference PDF obtained by Monte Carlo method. This bimodal PDF clearly exposes two modes, respectively centered around the deterministic minima at $x = -0.5$ and $x = 1.5$. Approximations of this PDF by the two proposed approaches are almost perfect, only slight differences appear at the lower tails (around $x = -1$ and $x = 1$).

Finally, we investigate the possibility of dimensional reduction by using the truncated version of the *PCE – KL – KG* representation. First, we propose to study the final doe of size 18, used for the results presented by Fig. 6 and to investigate the spectrum of the covariance matrix defined by Eq. (35) (see Section 2.2.3). Fig. 7 presents the value of $\frac{\sum_{i=1}^M \hat{\lambda}_k^{n_{doe}}}{\sum_{i=1}^{n_{doe}} \hat{\lambda}_k^{n_{doe}}}$ where n_{doe} is the size of the doe, $\hat{\lambda}_k^{n_{doe}}$ are the eigenvalues of the covariance matrix computed on this doe and M is the number of modes kept in the truncation.

One can note in Fig. 7 that the 10 first eigenvalues represent the majority of the covariance spectrum. This is very interesting in comparison to the 23 modes kept in the KL expansion of the whole process. It seems that, the minimum

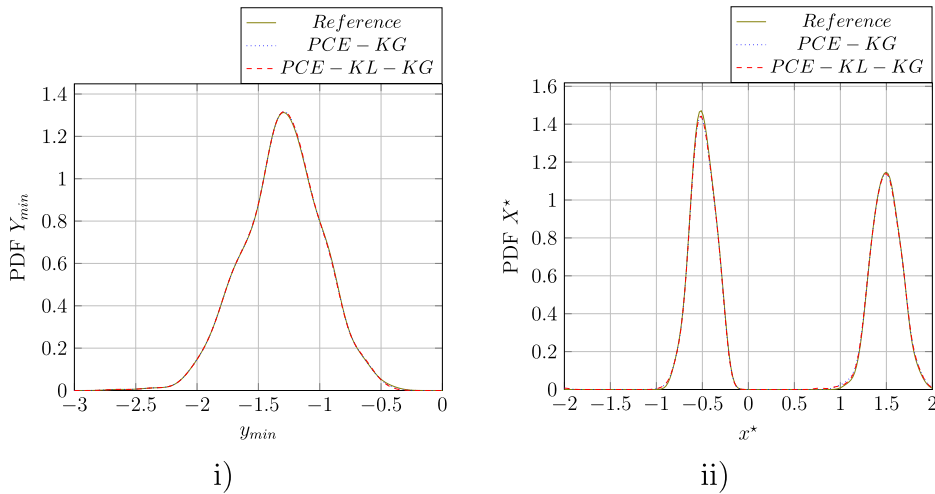


Fig. 6. Comparison of the PDF obtained by both random field approximations with the reference PDF obtained by Monte Carlo method ($\epsilon_{EIRF} = 10^{-2}$). (i) PDF of Y_{min} , (ii) PDF of X^* .

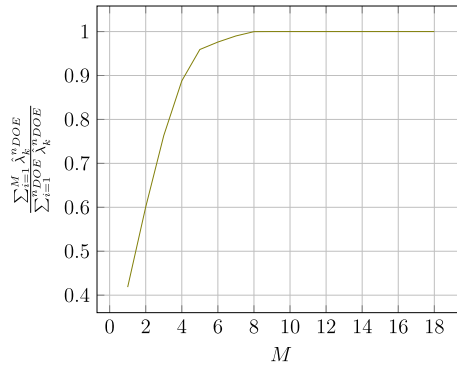


Fig. 7. Illustration of the covariance matrix computed on a doe of size 18. Truncation criterion as a function of the number of modes M kept for the truncation.

value of the random field can be fairly approximated with less modes than the whole random field. This remark allows for an important numerical time reduction in the proposed approach.

In order to study the accuracy of the truncated version of the $PCE - KL - KG$ representation, the same example as before is tested but, at each iteration of Algorithm 1, only the M modes that satisfy the truncation criterion defined by Eq. (16) are kept in the random field representation. It should be noted that the convergence criterion is still set to $\epsilon_{EIRF} = 10^{-2}$. Fig. 8 presents the evolution of the number of modes kept in the random field representation with respect to the iteration of Algorithm 1. Two values of the convergence criterion ϵ_{KL} (see Eq. (16)) are tested, $\epsilon_{KL} = 10^{-3}$ and $\epsilon_{KL} = 10^{-6}$. In order to evaluate the gain compared to the full method, Fig. 8 also presents the evolution of the doe size (we recall that, in its full version, the number of modes kept in the random field representation is equal to the doe size).

Several remarks can be done in Fig. 8. First, one can note that the truncated strategy converges after approximately the same number of iterations as the full strategy (17 iterations for the full strategy compared to 14 for the truncated strategy with $\epsilon_{KL} = 10^{-3}$ and to 15 for the truncated strategy with $\epsilon_{KL} = 10^{-6}$). Second, the truncation strategy is efficient for dimension reduction. Indeed, one can note that with $\epsilon_{KL} = 10^{-3}$ the number of modes kept in the random field representation evolves from 2 to 8 which leads to an important dimension reduction after iteration 7. The same remark can be done with $\epsilon_{KL} = 10^{-6}$ adding that a supplementary mode is kept at iteration 9. Finally, concerning

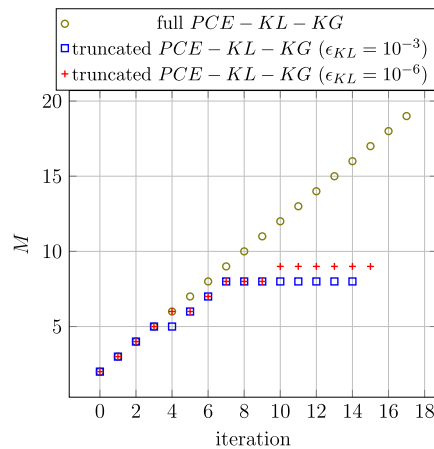


Fig. 8. Evolution of the number of modes M retained with respect to the iteration of Algorithm 1. Comparison between the full $PCE - KL - KG$ representation and the truncated $PCE - KL - KG$ representation with $\epsilon_{KL} = 10^{-3}$ or $\epsilon_{KL} = 10^{-6}$ (see Eq. (16)).

Table 3

Comparison of the estimators of the mean relative errors defined by Eq. (29) and Eq. (30) obtained by full and the truncated $PCE - KL - KG$ representations ($\epsilon_{KL} = 10^{-3}$ or $\epsilon_{KL} = 10^{-6}$).

	$Err_{Y_{min}}$	Err_{X^*}
Full	0.25%	2.52%
$\epsilon_{KL} = 10^{-3}$	0.72%	6.09%
$\epsilon_{KL} = 10^{-6}$	0.33%	2.75%

the global accuracy, Table 3 compares the estimations of the mean relative errors defined by Eq. (29) and Eq. (30) obtained by full and the truncated $PCE - KL - KG$ representations.

Table 3 reveals that the truncated strategy with $\epsilon_{KL} = 10^{-3}$ slightly deteriorates the accuracy whereas it has a comparable accuracy when ϵ_{KL} is set to 10^{-6} .

To conclude on this first example, one can say that, the 3 tested strategies for the representation of the random field ($PCE - KG$, full $PCE - KL - KG$ and truncated $PCE - KL - KG$) lead to comparable accuracy with respect to the minimum value of the field and its position. However, their numerical costs, are different. For the $PCE - KG$ strategy one has to build a Kriging interpolation of each coefficient which leads to 24 interpolations. For the full $PCE - KL - KG$ representation, the number of interpolations to build is equal to $n_{doe} + 1$. This seems much more efficient, however the numerical cost corresponding to the construction and resolution of the eigenvalue problem is not negligible. Finally, the truncated $PCE - KL - KG$ allows a significant dimension reduction leading to smaller number of Kriging interpolations to be built. This strategy is the most efficient on this example. However, the difficulty is then to choose the threshold ϵ_{KL} (previous example shows that its influence can be non negligible on the global accuracy of the method). Moreover we would like to add that the dimension reduction is case dependent and cannot be guaranteed *a priori*. For these reasons, it seems reasonable to try the truncated $PCE - KL - KG$ representation with a low threshold ϵ_{KL} (10^{-6}). If this leads to a number of interpolations to be built higher than the number of PCE coefficients, then one should switch to the $PCE - KG$ representation.

Remarks:

- This example illustrates the type of problem we are interested in. Indeed, PCE approximation of $Y(x, \xi)$ is very efficient as $Y(x, \xi)$ is Gaussian, nevertheless an accurate direct PCE approximation of $X^*(\xi)$ is very difficult to reach. In order to detail the computational cost involved by the direct and the proposed indirect approach, we use a sample of size $2P$ for the computation of the PCE coefficients. The number of evaluations of the objective function with the proposed indirect approach is equal to: the size of the final doe obtained with the proposed algorithm times $2P$. As $Y(x, \xi)$ is Gaussian, Hermite polynomials of degree $d = 1$ are retained for

Table 4

Number of objective function calls and estimators of the mean relative error on $X^*(\xi)$ in case of a direct PCE approximation.

d	nb evaluations	Err_X^*
1	1632	116%
2	20694	108%
3	177467	95%

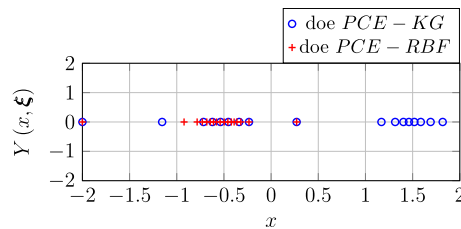


Fig. 9. Comparison of the final doe obtained by applying Algorithm 1 using $PCE - KG$ representation or $PCE - RBF$ representation.

the PCE basis leading to $P = n_s + 1 = 24$, which finally leads to 816 calls to the objective function (we assume the $PCE - KG$ representation leading to a final doe of size 17, see Table 2). We now try the direct PCE approximation of $X^*(\xi)$. Polynomial degree from $d = 1$ to $d = 3$ are tested. The corresponding numbers of objective function calls used to create the sample necessary to construct the PCE approximation are given in Table 4 (deterministic optimization problems used to construct the sample are solved with SLSQP algorithm and 2 starting points) as well as the mean relative error estimated on 1000 simulations. Results presented by Table 4 clearly show that direct PCE approximation is inefficient in the case where $X^*(\xi)$ is multimodal.

- As explained at the end of Section 3.2, Kriging interpolation offers the advantage of quantifying the interpolation uncertainty which allows to help the *exploration* of the parametric domain S . On this example this advantage allows the proposed algorithm to explore the two areas of S where the minimum is likely to be. In order to highlight this advantage, the same example is studied using RBF interpolation (with a Gaussian radial basis function) instead of Kriging (see end of Section 3.2 for the computation of the enrichment criterion in that case). Using the same initial doe ($x^{(1)} \approx -0.23$ and $x^{(2)} \approx 0.27$), the direct interpolation of PCE coefficients and the same number of iterations as the one obtained with the Kriging approach (15 iterations, see Table 2), the proposed algorithm added 15 points only around the possible minimum located around $x = -0.5$ and totally missed the one around $x = 1.5$ as presented by Fig. 9. In order to capture the two promising areas we have to increase the size of initial doe from 2 to 4 (which allows a better initial exploration). This example illustrates the fact that, by taking into account interpolation uncertainty, the proposed Kriging approach will be more efficient and robust to the initial doe than a RBF approach.

4.1.2. 2 dimensional case

We now consider a 2 dimensional Gaussian random field defined on $S = [-5, 5] \times [-5, 5]$, with mean $\mu(\mathbf{x}) = \sum_{i=1}^2 \frac{x_i^2}{4000} - \prod_{i=1}^2 \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$. The random field is then created by adding 3 linear random terms such that,

$$Y(\mathbf{x}, \xi) = \mu(\mathbf{x}) + \xi_1 + 0.1(\xi_2 x_1 + \xi_3 x_2) \quad (31)$$

where ξ is a standard Gaussian random vector of independent components (ξ_1 , ξ_2 and ξ_3). Fig. 10 presents the mean value of the field (3D visualization Fig. 10(i) and contour plot Fig. 10(ii)). It should be noted that $\mu(\mathbf{x})$ counts 5 local minima and that the global minimum is reached at $\mathbf{x} = (0, 0)$.

Fig. 11 presents 4 realizations of the random field, for each one the position of the minimum is highlighted. This figure illustrates that the random variable modeling the position of the minimum (X^*) exhibits 4 modes located around the four local minima at the corners of the domain S . As for the previous example, studying a Gaussian random field

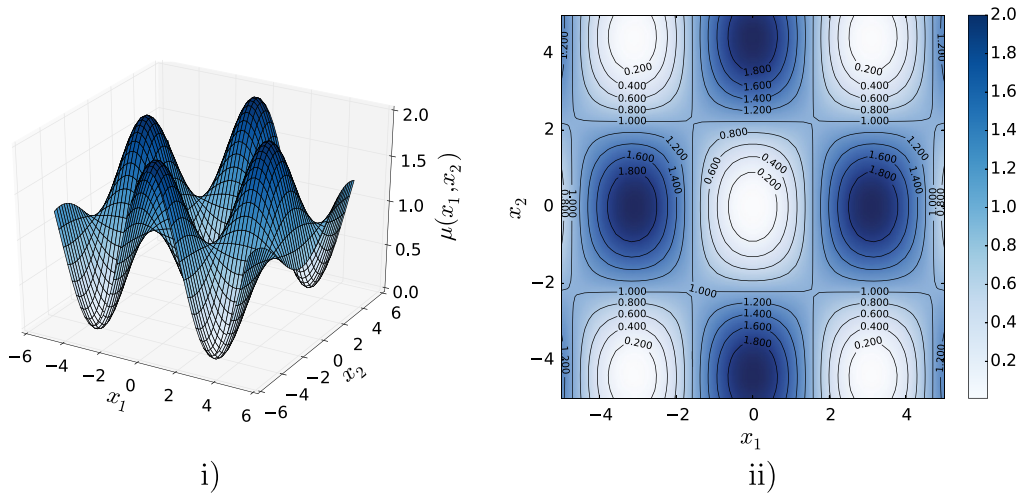


Fig. 10. Mean value of the 2D Gaussian random field counting 5 local minima, (i) 3D visualization, (ii) contour plot.

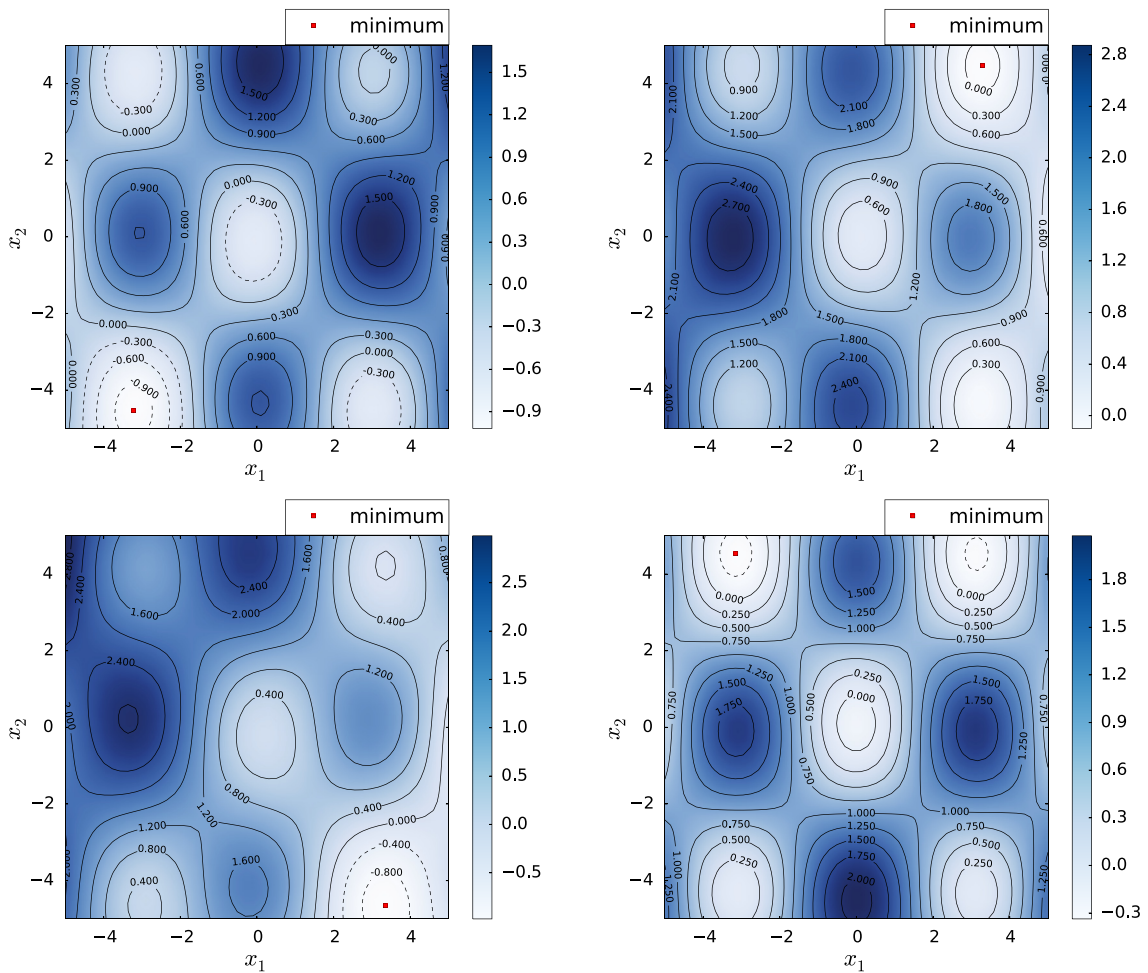


Fig. 11. Random realizations of the 2D Gaussian random field defined by Eq. (31). For each realization the position of the global minimum is given.

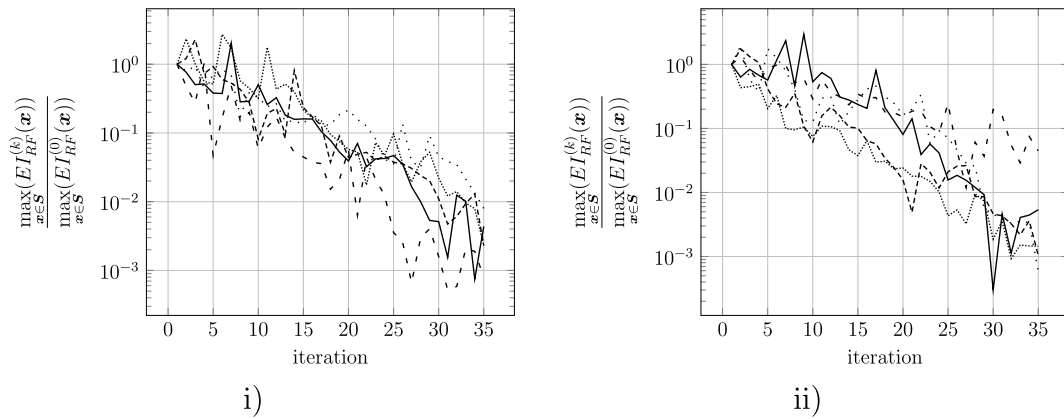


Fig. 12. Evolution of the convergence criterion defined by Eq. (23) on 35 iterations (5 replicates). (i) *PCE* – *KG* representation, (ii) *PCE* – *KL* – *KG* representation.

allows to represent it exactly by the Hermite polynomial chaos and then to focus on the proposed adaptive approach. Note that as the random field is made by 3 independent standard Gaussian random variables, its *PCE* representation counts 4 terms. This means that the *PCE* – *KG* representation of the field needs the construction of only 4 Kriging interpolations. The proposed method is now applied using the two random field representations *PCE* – *KG* and truncated *PCE* – *KL* – *KG* with $\epsilon_{KL} = 10^{-6}$. Note that, based on the conclusion of the first example, the full *PCE* – *KL* – *KG* will not be studied anymore, and thus the notation *PCE* – *KL* – *KG* implicitly refers to the truncated *PCE* – *KL* – *KG* representation.

Fig. 12 presents the evolution of the convergence criterion defined by Eq. (23) with respect to the number of iterations of Algorithm 1 for the two random field representations. Initial doe is sampled by LHS and counts 4 points.

One can observe the same behavior as for the first example. The convergence criterion is globally decreasing while the number of iterations increasing. However, this convergence is not monotonic. Compared to the 1D case (Fig. 3), one can note that the dispersion between the five replicates is larger and that the results obtained with the *PCE* – *KL* – *KG* representation (Fig. 12(ii)) seems to be more dispersed than the one obtained with the *PCE* – *KG* representation (Fig. 12(i)). Nevertheless, as for the 1D example, both representations lead to convergence of the Algorithm 1 and only small differences in the number of iterations necessary to reach a given value of the convergence criterion are notable. Moreover, we add that for the *PCE* – *KL* – *KG* representation the number of modes kept according to the defined criterion (we recall that $\epsilon_{KL} = 10^{-6}$) remains constant all over the iterations and is equal to 3. This demonstrates the efficiency of the empiric KL representation for finding an appropriate subspace to represent the random field from a limited number of sampling locations.

In the following, the convergence criterion $\epsilon_{EI_{RF}}$ is set to $\epsilon_{EI_{RF}} = 0.01$. Fig. 13(i) presents the initial and final doe obtained by running Algorithm 1 on this example. It should be noted that both representations lead to a convergence of Algorithm 1 after 28 iterations, leading to a final doe of 32 points. Moreover, the two representations lead to different but very close point locations. These final doe are clearly located at the four corners of the domain S , where the minimum is likely to be. In order to illustrate the interest of our adaptive strategy we propose to compare the results obtained using these adaptive doe to the one obtained using a doe constructed by LHS sampling of 32 points and presented by Fig. 13(ii).

In order to access the accuracy of the proposed approach, Figs. 14 and 15 present 3 random realizations of the random field (one per column). For these realizations, Fig. 14 and Fig. 15 also present the approximations obtained by the *PCE* – *KG* representation and the *PCE* – *KL* – *KG* representation respectively. Moreover, we also present results obtained by constructing these two representations on doe obtained by LHS with 32 points (Fig. 13(ii)). These results are respectively denoted by *LHS* – *PCE* – *KG* and *LHS* – *PCE* – *KL* – *KG*. In every case the position of the global minimum is highlighted.

This example illustrates the benefit of the proposed adaptive approach for the estimation of extreme value statistics. Indeed, one can note that the approximations of the 3 realizations obtained with the adaptive approaches (Fig. 14 second row and Fig. 15 second row) are globally inaccurate on the whole domain, but perfectly represent the

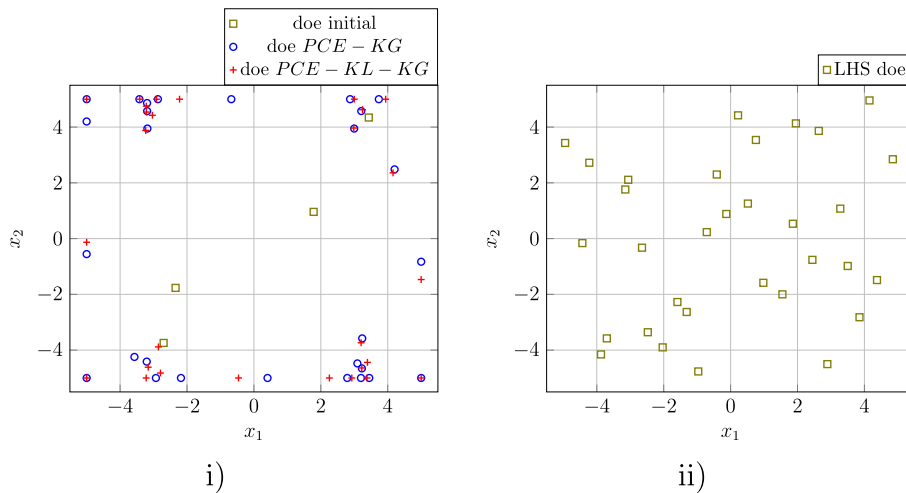


Fig. 13. (i) Comparison of the final doe obtained by applying Algorithm 1 using $PCE - KG$ representation and $PCE - KL - KG$ representation. (ii) doe of 32 points constructed by LHS and used for comparison.

Table 5

Comparisons of the mean relative errors defined by Eq. (29) and Eq. (30) obtained by the various approximations approaches on the example defined by Eq. (31).

	$Err_{Y_{min}}$	Err_{X^*}
$PCE - KG$	0.69%	1.68%
$PCE - KL - KG$	0.16%	0.37%
$LHS - PCE - KG$	3.60%	31.61%
$LHS - PCE - KL - KG$	4.09%	45.32%

realizations in the four corner of the domain where the minimum value is likely to be. This allows for a perfect approximation of the minimum position (as shown by the comparison with the reference results). In comparison, the approximations obtained on a LHS doe containing the same number of points (Fig. 14 third row and Fig. 15 third row) allow a better representation of the realizations over the whole domain but are clearly inaccurate for the determination of the minimum position. Especially, one can note that for the third realization (third columns of Fig. 14 and Fig. 15), approximations of the minimum position by LHS doe are in the wrong corner of the domain.

In order to assess the accuracy over the whole stochastic space, estimations of the two mean relative errors defined by Eq. (29) and Eq. (30) are now performed. As for the previous example, these mean relative errors are estimated over 1000 samples. Table 5 compares these results between the different approaches.

Table 5 shows that the $PCE - KL - KG$ representation leads to significantly better results on this example. However, the results obtained with the $PCE - KG$ representation are also very satisfying. In addition to this comment, Fig. 16 presents the approximations of the PDF of Y_{min} obtained with the two proposed representations (kernel smoothing over 1000 samples). One can note that both approximations perfectly represent the reference results and that the 3 curves are not distinguishable. Finally, the mean relative errors are also estimated on the approaches using the LHS doe with the same number of points. As presented by Table 5 the results obtained by these approaches are strongly inaccurate which highlight the advantage of the proposed adaptive approach compared to classical LHS sampling.

It is also notable that the results on this example are better than the one obtained on the first example. It should be noted that, even if it deals with a 2 dimensional parametric space, it only involves a stochastic space of dimension 3 which makes it easy to solve by PCE. Interest of this example was to illustrate the advantages of the proposed adaptive strategy on a test case where the minimum position can be located at various, relatively distant areas of the parametric domain. In this respect, results presented by Figs. 14, 15 and Table 5 allow to be confident in the capability of the proposed approach.

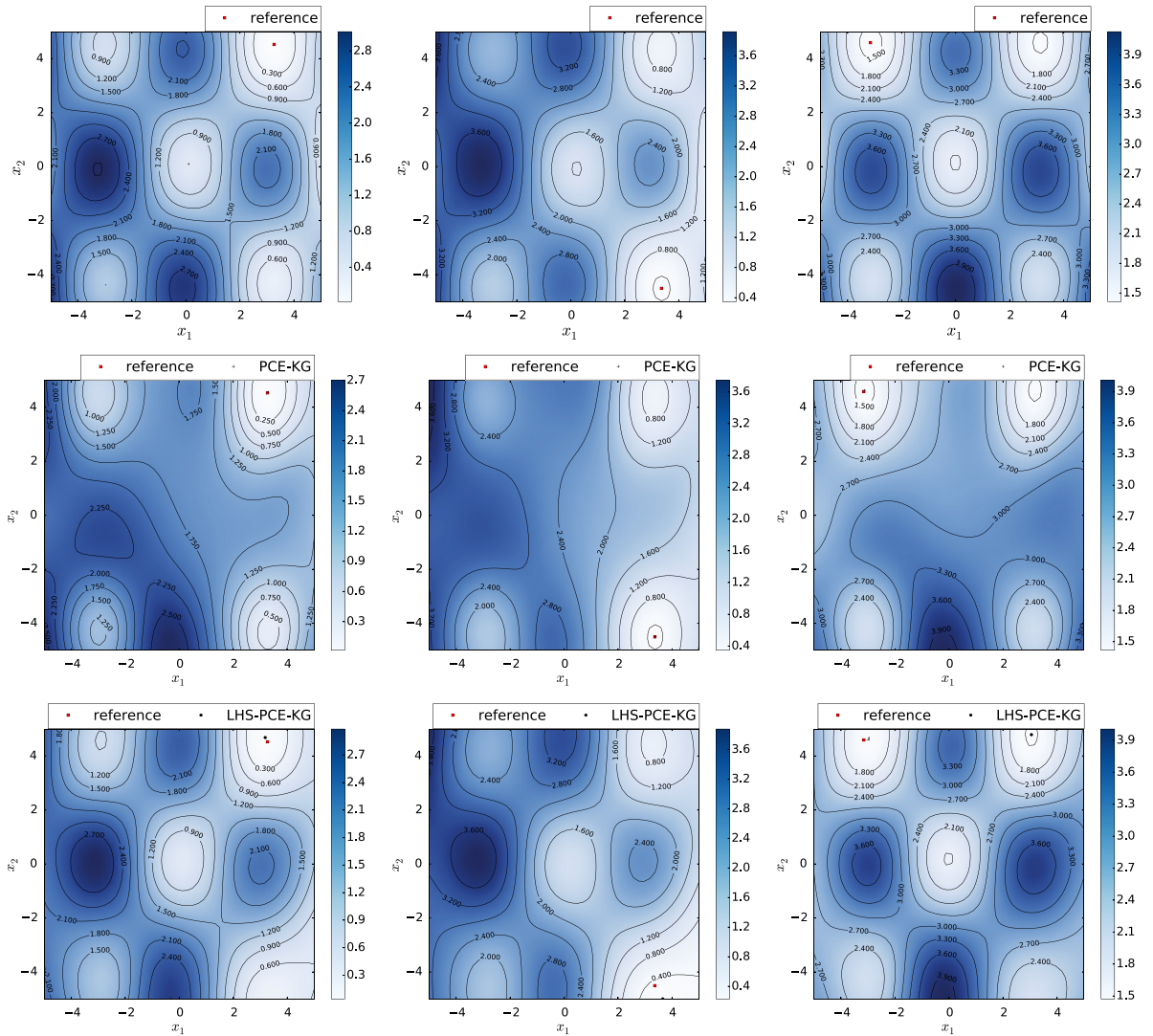


Fig. 14. Random realizations of the 2D Gaussian random field defined by Eq. (31). First row: reference results, second row: approximation using adaptive doe constructed by Algorithm 1 and $PCE - KG$ representation, third row: approximation using LHS doe and $PCE - KG$ representation. Position of the global minimum is given for the reference and the approximations.

Remark: As for the first example, deterministic interpolation by RBF instead of Kriging is also tested on this example. The comparison is made for the same initial doe (see Fig. 13) and for the direct interpolation of the PCE coefficients. As for the example in one dimension, the doe obtained after 28 iterations (number of iterations necessary for our approach to converge) is not as efficient as the one we obtain with our approach. Indeed, even if all the possible minimum locations are explored, only a single point is sampled in the top left corner of the domain (see Fig. 17). This behavior highlights the importance of the exploration of the domain by taking into account the interpolation uncertainty in the Kriging methodology.

4.2. Non Gaussian case: a conceptual aircraft wing design

4.2.1. Presentation

This example presents the parametric study of an aircraft wing under uncertainty. The shape of the wing is defined by 4 variables, namely the root chord c_{root} , the tip chord c_{tip} , the span l and the sweep angle ϕ . We are interested in

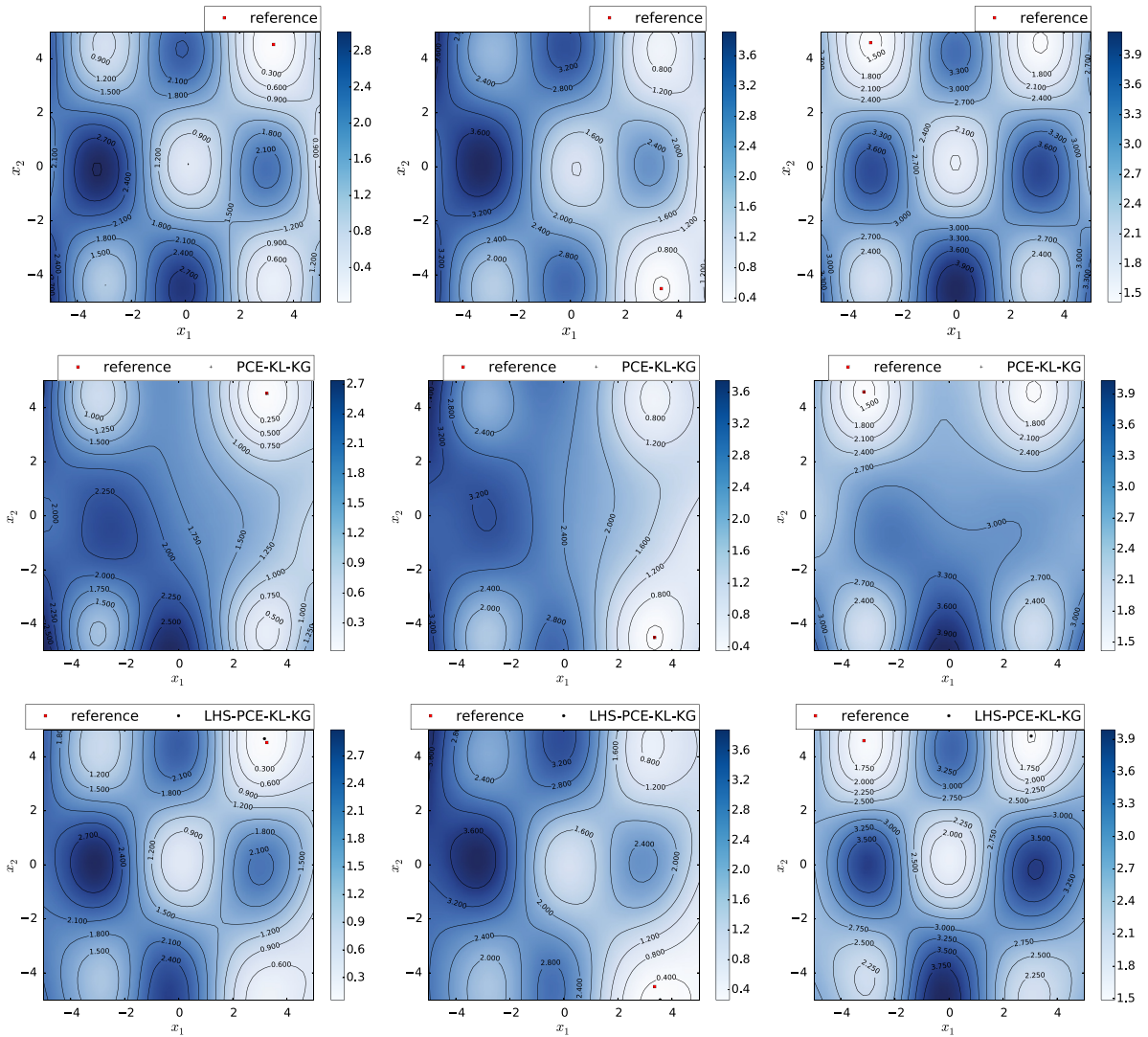


Fig. 15. Random realizations of the 2D Gaussian random field defined by Eq. (31). First row: reference results, second row: approximation using adaptive doe constructed by Algorithm 1 and $PCE - KL - KG$ representation, third row: approximation using LHS doe and $PCE - KL - KG$ representation. Position of the global minimum is given for the reference and the approximations.

the aerodynamic behavior of the wing, hence we also defined the angle of attack α . A symmetric NACA0010 airfoil is used. The objective of this conceptual design phase is to roughly evaluate the aerodynamic loads. Hence, a 3D potential flow panel code (PANAIR/A502) [49] is used to evaluate these loads. This code uses a discretization of the wing geometry by panels and returns the pressure coefficients computed at the center of each panel (see [50]). The aerodynamic coefficients of the wing are obtained by numerically integrating the pressure distribution over the wing surface. Fig. 18(i) presents the wing jig shape as well as the mesh (Fig. 18(ii)) used for the computation of the aerodynamic loads.

In the following we study the influence of the parameter $x = \phi$ (sweep angle in radian), defined on the space $S = [-\frac{\pi}{3}, \frac{\pi}{3}]$, on the response defined by

$$Y(x, \xi) = (c_l(x, \xi) - 0.5)^2 + 0.1c_y(x, \xi) \quad (32)$$

where c_l and c_y are respectively the lift coefficient and the pitching moment coefficient (calculated at the root leading edge). This quantity of interest is made by mixing an objective on the lift of the wing (term $(c_l(x, \xi) - 0.5)^2$) and a

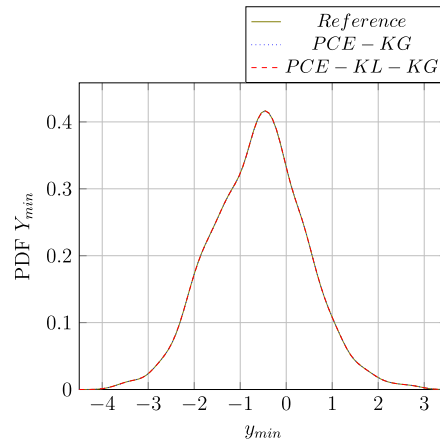


Fig. 16. PDF of Y_{min} , approximations obtained by the two proposed approaches.

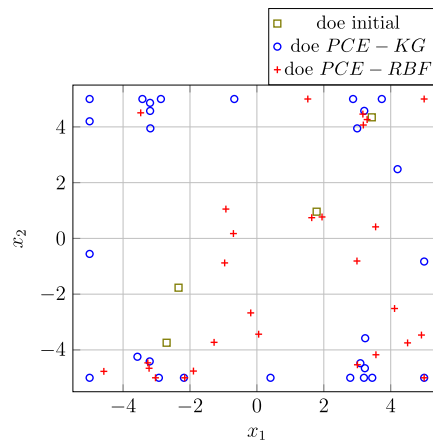


Fig. 17. Comparison of the final doe obtained by applying Algorithm 1 using $PCE - KG$ representation or $PCE - RBF$ representation.

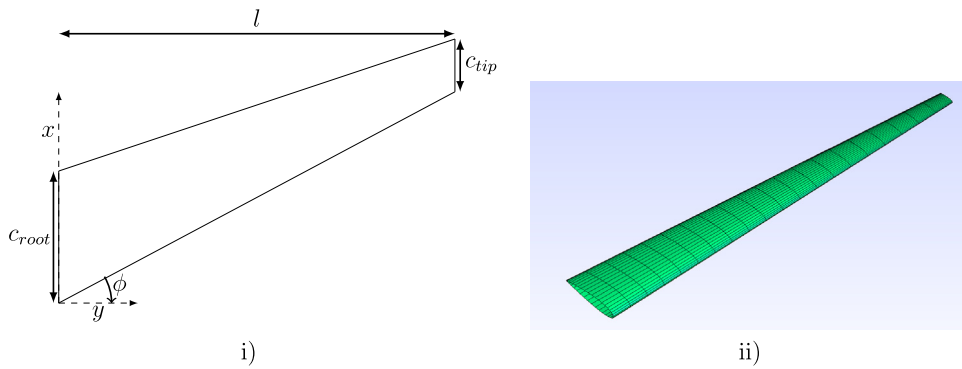


Fig. 18. (i) Shape of the wing under study. (ii) Mesh of the wing used for the computation of the aerodynamic loads.

Table 6

Means and coefficients of variation of the Gaussian random variables affecting the problem of conceptual aircraft wing design.

Variables	Mean	Coefficient of variation
c_{root}	5.67 m	4%
l	17.04 m	4%
α	2.5 deg	6%

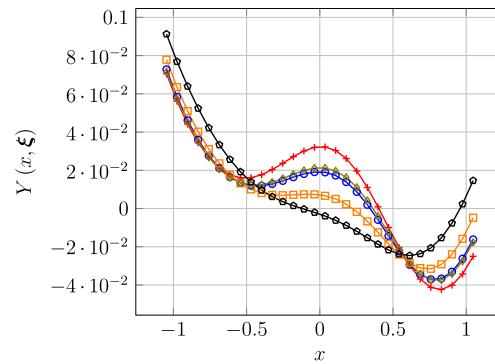


Fig. 19. Examples of realizations of the random process defined by Eq. (32), x is the sweep angle.

constraint term on the pitching moment (term $0.1c_y(x, \xi)$). Moreover, ξ is a vector of 3 independent normal random variables modeling the uncertainties that affect the angle of attack, the span and the tip chord. Numerical values for the mean and coefficient of variation of these random variables are given in Table 6.

Finally, the problem under study has a spatial dimension equal to one and a stochastic dimension equal to 3. It should be noted that the computation of the aerodynamic coefficients c_l and c_y involves non linear transformations of the random vector ξ . Hence, the random process $Y(x, \xi)$ is non Gaussian.

Fig. 19 presents some samples of the random process under study. For this figure, the parametric space S is discretized with 30 linearly spaced values. One could note that, as the parameter under study and the random variables affect the geometry of the wing, it is necessary to create a new geometry and a new mesh at each evaluation of the model $Y(x, \xi)$. With the computer we use for this example (CPU core 2 DUO @ 3 GHz, memory 4 GB), the whole computation of the aerodynamic loads (*i.e.* creation of the geometry, meshing and resolution of the potential fluid equation by the panel method) lasts around 8.6 s. Hence, as an example, the computational time necessary to compute the values used in Fig. 19 is $8.6 \times 30 \times 5 \approx 1290 \text{ s} \approx 21.5 \text{ min}$.

Fig. 19 shows that the realizations of the process defined by Eq. (32) count two local minima and that the global minimum is reached approximately for $\phi \in [0.5, 1.0]$. Physically this can be explained by the fact that the lift objective (term $(c_l(x, \xi) - 0.5)^2$) can reach 0 for a positive or a negative value of the sweep angle (which respectively corresponds to conventional or forward-swept wings). However, adding the constraint term on the pitching moment coefficient (term $0.1c_y(x, \xi)$) forces the minimum to be uniquely defined at the positive sweep angle (conventional solution).

4.2.2. Results

The proposed method is now applied to this example. Contrary to the Gaussian examples presented in Section 4.1 the PCE of the random process cannot be computed exactly and analytically. Hence, we rely on the non intrusive PCE with a computation of the coefficients based on regression (see [10]). This method involves the sampling of the model over the stochastic space. After a convergence study we decide to set the maximal degree of the PCE to $d = 3$. Accordingly the size of the Hermite polynomial basis is equal to $P = 20$. Hence, the sample size for the computation of the coefficients is set to 40 (twice the number of coefficients to determine is commonly used in the regression approach). One uncertainty quantification step, at a given point $x^{(0)} \in S$, thus last approximately $8.6 \times 40 = 344 \text{ s} \approx 5.75 \text{ min}$.

Table 7

Conceptual aircraft wing design example, comparison of the computational time between Monte Carlo method and the proposed approach to solve 1000 optimization problems.

	Monte Carlo reference	$PCE - KL - KG$
Calls to the model	19811	800
Resolution of 1000 optimization problems	41.73 h	0.41 h
Adaptive construction of the $PCE - KL - KG$ representation	–	1.91 h
Total time	41.73 h	2.32 h

Then, the proposed Algorithm 1 is applied using the truncated $PCE - KL - KG$ method with $\epsilon_{KL} = 10^{-6}$. This leads to a number of modes kept equal to 4 (reached after 2 iterations). The convergence criterion is set to $\epsilon_{ELRF} = 10^{-2}$. The initial doe counts 2 points and the Algorithm 1 converges after 18 iterations. Hence, the final doe counts 20 points. This implies $20 \times 40 = 800$ evaluations of the model $Y(x, \xi)$ for the UQ step. This step thus last $8.6 \times 800 = 6880 \text{ s} \approx 1.91 \text{ h}$ and it should be noted that the computational time of the other steps of Algorithm 1 is negligible compared to the UQ step. Fig. 20 presents the final doe as well as the approximations of five realizations of the random process.

One can note that Algorithm 1 performs well on this non Gaussian example as the final doe is very dense in the range $[0.5, 1.0]$, where the minimum is likely to be. Moreover, the minimum position and its value are also well approximated for the 5 random realizations presented by Fig. 20. Once again one can note that, on the five random realizations presented by Fig. 20, the $PCE - KL - KG$ representations are very accurate in the range $[0.5, 1.0]$ whereas they are less accurate elsewhere (especially in the range $[-1, -0.5]$).

We now study the accuracy of the method by estimating the mean relative errors defined by Eq. (29) and Eq. (30) with a sample of size 1000. This leads to $Err_{Y_{min}} \approx 0.025\%$ and $Err_{X^*} \approx 0.44\%$. These very low relative errors illustrate the efficiency of the proposed approach on a industrial based non Gaussian example. Fig. 21 presents the PDF of Y_{min} and X^* estimated by kernel smoothing (standard normal kernel) over the sample of size 1000. As expected, the approximations of the PDF of Y_{min} and X^* by the proposed approach are very accurate. More precisely, for the PDF of Y_{min} (Fig. 21(i)) the approximation and the reference curves are perfectly superposed and for the PDF of X^* (Fig. 21(ii)) only slight differences appear.

Finally, we would like to give some insight about the numerical cost of our approach compared to brute force Monte Carlo used for reference results. These reference results are obtained by the SLSQP algorithm with evaluation of the derivative by finite differences. In order to reduce the computational time we choose only one starting point at $x = 0.5$. Table 7 presents the number of function calls to solve the 1000 optimization problems as well as the computational time it involves. Thus, reference results are obtained in approximately 41.73 h.

In comparison, the resolution of the optimization problem on the truncated $PCE - KL - KG$ approximation lasts, on these 1000 realizations, 0.41 h. We note that, on the $PCE - KL - KG$ approximation, the optimization problem is solved by multi-start SLSQP algorithm; that we provided the analytic expressions of the partial derivatives (which are available here due to the use of the Kriging metamodels) and that we use the final doe as starting points. For a fair comparison, the time spend in the adaptive construction of the $PCE - KL - KG$ approximation must be added to the time spend for the optimization. This leads to a total computational time for the proposed approach equal to $1.91 \text{ h} + 0.41 \text{ h} = 2.32 \text{ h}$, whereas the computational time for the reference results is 41.73 h. Hence, we obtained a computational saving by a factor of 18 although the evaluation of the model only lasts around 8.6 s (which can be considered as very fast for a numerical model). It should be noted that the computational savings obtained with the proposed approach compared to the Monte Carlo method is going to increase as the computational time of a single evaluation of the model increases and as the number of optimization problems to solve increases (size of the sample for the Monte Carlo method).

5. Conclusion

This paper presents an original adaptive random field discretization dedicated to the approximation of extreme value statistics. The proposed approach is based on a representation of the random field by Polynomial Chaos Expansion (PCE) and Kriging interpolation. Taking advantages of some of the Kriging properties, we define a criterion that allows to construct iteratively the design of experiments (doe) where the uncertainty quantification by PCE must

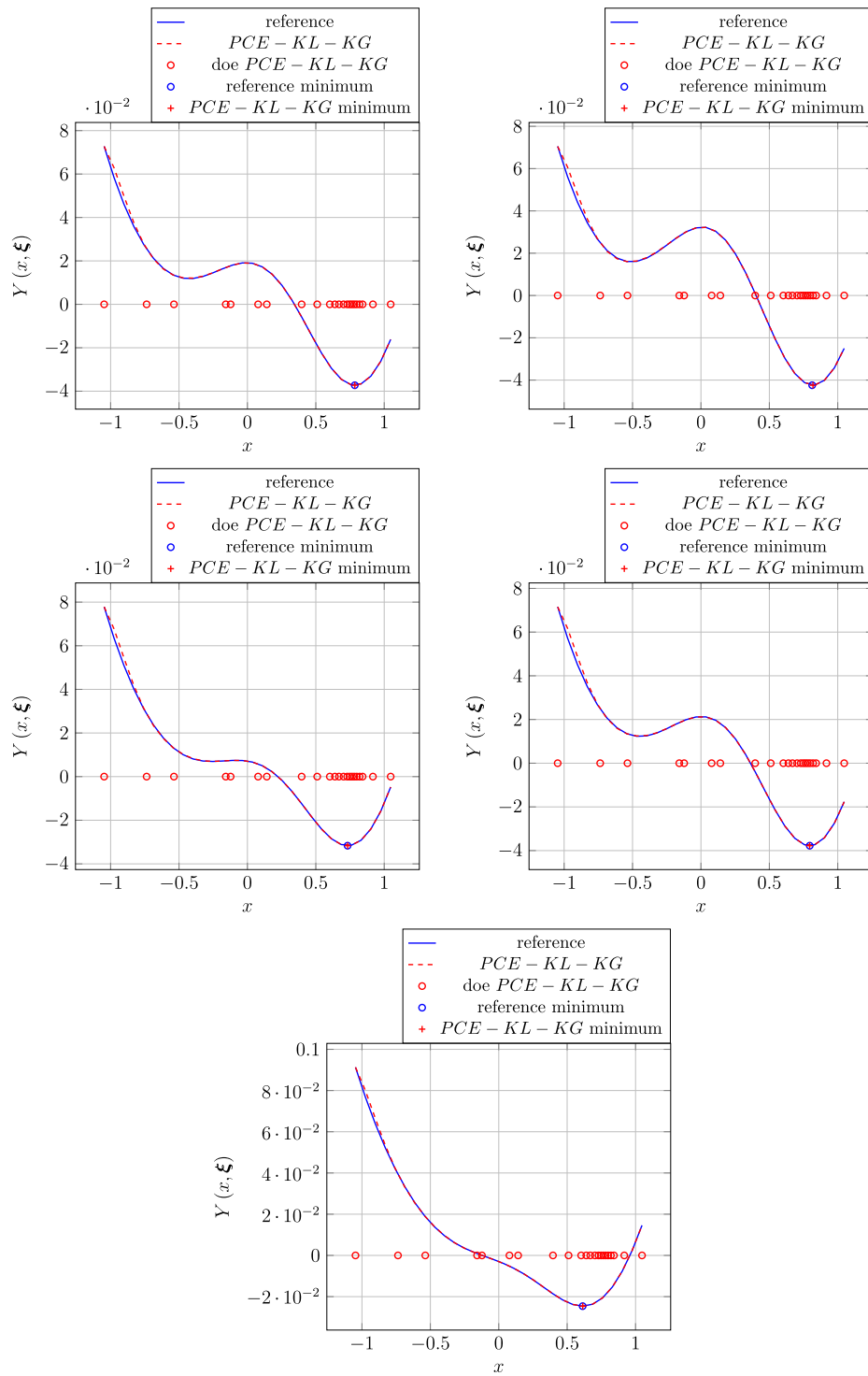


Fig. 20. Illustration of approximation of the random process defined by Eq. (32) by the proposed approach using the truncated $PCE - KL - KG$ random field representation. 5 random realizations are presented.

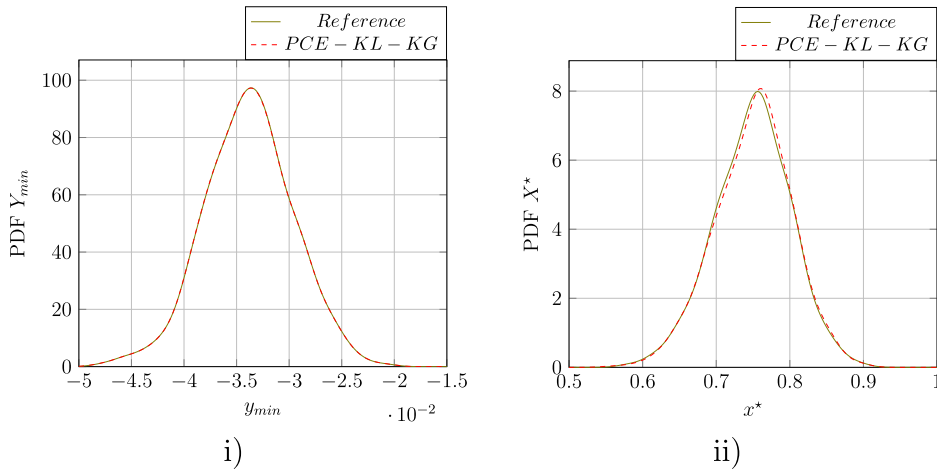


Fig. 21. Conceptual aircraft wing design example, PDF obtained by the truncated $PCE - KL - KG$ representation compared to reference results, (i) PDF of Y_{min} , (ii) PDF of X^* .

be performed, in order to bring relevant information with respect to the extreme value statistics of the random field. An efficient and accurate numerical estimation of this criterion is also proposed. Finally, Algorithm 1 presents the global strategy leading to the construction of an hybrid PCE-Kriging approximation of the random field (note that two random field representations are proposed and compared). Closed form expressions of these representations are given by Eq. (10) and by Eq. (15). Moreover, partial derivatives with respect to the parametric variable x can be derived analytically for both representations which make them very efficient for the estimation of extreme value statistics by direct Monte Carlo method. The efficiency of the proposed approach is illustrated on 2 Gaussian examples in spatial dimension 1 and 2 and stochastic dimension equal to 23 and 3 respectively. A non Gaussian example, of spatial dimension 1 and stochastic dimension 3, dealing with the parametric study of an aircraft wing and involving the computation of aerodynamics loads by a numerical code is also presented.

Concerning the perspective of this research, one can note that a major drawback of the proposed approach is that the complexity of the representations increases rapidly with respect to both the spatial and the stochastic dimensions (well known phenomenon of *curse of dimensionality*). To address this issue we would first like to emphasize that for both random field representations proposed ($PCE - KG$ Eq. (10) and $PCE - KL - KG$ Eq. (15)), the spatial and the stochastic dimension are separated. Indeed, the spatial dimension is handled by the Kriging metamodels whereas the stochastic dimension is handled by the PCE. Hence, we must be able to take benefit of some of the recent developments of these two methods to deal with larger dimension with respect to both the spatial and the stochastic space. For example, concerning the Kriging interpolations, the issue of dealing with large spatial dimension has been tackled by [51] in which Partial Least Square is used for dimension reduction. Concerning the stochastic dimension, it could be of great interest to adapt the sparse PCE proposed by [11] that uses Least Angle Regression. However, the method we proposed assumes that the same PCE basis is used at each point x of the parametric space S . Hence, adaptation of the method proposed by [11] imposes to construct a sparse PCE basis that is able to represent the random responses over the whole parametric space S . A first solution might be to construct this sparse basis in an iterative way along with the discretization of the random field. However, there is no guarantee that a sparse basis might be sufficient to represent the random response over the whole parametric space. This behavior is case dependent and will be studied in future developments.

Finally, note that recent researches show that Low Rank Approximations (LRA) could improve the performance of PCE (and sparse PCE), especially for uncertainty quantification in large stochastic dimension (see [52]). Hence, it is an interesting perspective to adapt the LRA framework to the method we proposed here. Note that this adaptation also raises the issue of constructing a basis able to represent the random quantity of interest over the whole parametric space S (in the framework of LRA similar problems have been studied in [53]).

Acknowledgments

The research presented in this paper has been performed in the framework of the AGILE project (Aircraft third Generation MDO for Innovative Collaboration of Heterogeneous Teams of Experts) and has received funding from the European Union Horizon 2020 Programme (H2020-MG-2014-2015) under Grant Agreement No. 636202. This work has been also partially supported by the French National Research Agency (ANR) through the ReBReD project under grant ANR-16-CE10-0002. These supports are gratefully acknowledged.

Appendix

A.1. Detail on the KL decomposition of a PCE random vector

Here are some details about the calculation steps that allow to express the KL decomposition of $\hat{\mathbf{Y}}_{n_{doe}}(\boldsymbol{\xi})$. First of all, we recall the expression of $\hat{\mathbf{Y}}_{n_{doe}}(\boldsymbol{\xi})$ with respect to $\mathbf{a}_i$,

$$\hat{\mathbf{Y}}_{n_{doe}}(\boldsymbol{\xi}) = \sum_{i=1}^P \mathbf{a}_i \phi_i(\boldsymbol{\xi}). \quad (33)$$

where $\{\phi_i(\boldsymbol{\xi})\}_{i=1}^P$ still denotes the basis formed by the n_s variate Hermite polynomials. Then, taking advantage of the PCE properties, the mean of the vector $\hat{\mathbf{Y}}_{n_{doe}}(\boldsymbol{\xi})$, denoted by $\boldsymbol{\mu}_{\hat{\mathbf{Y}}}$, and its covariance matrix, denoted by $\mathbf{K}_{\hat{\mathbf{Y}}}$ are expressed thanks to PCE coefficients by,

$$\boldsymbol{\mu}_{\hat{\mathbf{Y}}} = \mathbf{a}_1 \quad (34)$$

and

$$\mathbf{K}_{\hat{\mathbf{Y}}} = \sum_{i=2}^P \mathbf{a}_i \mathbf{a}_i^t. \quad (35)$$

Hence, solving the eigenproblem

$$\mathbf{K}_{\hat{\mathbf{Y}}} \hat{\boldsymbol{\phi}}_i = \hat{\lambda}_i^{n_{doe}} \hat{\boldsymbol{\phi}}_i$$

where $\hat{\lambda}_i^{n_{doe}}$ and $\hat{\boldsymbol{\phi}}_i \in \mathbb{R}^{n_{doe}}$ are the n_{doe} eigenvalues and eigenvectors of $\mathbf{K}_{\hat{\mathbf{Y}}}$, allows to express the random vector $\hat{\mathbf{Y}}_{n_{doe}}$ by,

$$\hat{\mathbf{Y}}_{n_{doe}}(\boldsymbol{\xi}) = \boldsymbol{\mu}_{\hat{\mathbf{Y}}} + \sum_{j=1}^{n_{doe}} \hat{\gamma}_j^{n_{doe}}(\boldsymbol{\xi}) \sqrt{\hat{\lambda}_j^{n_{doe}}} \hat{\boldsymbol{\phi}}_j$$

where $\hat{\gamma}_j^{n_{doe}}(\boldsymbol{\xi})$ are zero mean and unit variance random variables such that,

$$\hat{\gamma}_j^{n_{doe}} = \frac{1}{\sqrt{\hat{\lambda}_j^{n_{doe}}}} (\hat{\mathbf{Y}}_{n_{doe}}(\boldsymbol{\xi}) - \boldsymbol{\mu}_{\hat{\mathbf{Y}}})^t \hat{\boldsymbol{\phi}}_j. \quad (36)$$

Then, by substituting $\hat{\mathbf{Y}}_{n_{doe}}$ by its expression (Eq. (33)) in Eq. (36), one gets the PCE of the random variable $\hat{\gamma}_j^{n_{doe}}$,

$$\hat{\gamma}_j^{n_{doe}} = \sum_{i=2}^P \hat{\gamma}_{j,i}^{n_{doe}} \phi_i(\boldsymbol{\xi}), \quad \text{where } \hat{\gamma}_{j,i}^{n_{doe}} = \frac{1}{\sqrt{\hat{\lambda}_j^{n_{doe}}}} \mathbf{a}_i^t \hat{\boldsymbol{\phi}}_j.$$

Finally, expression of the random vector $\hat{\mathbf{Y}}_{n_{doe}}(\boldsymbol{\xi})$ reads,

$$\hat{\mathbf{Y}}_{n_{doe}}(\boldsymbol{\xi}) = \boldsymbol{\mu}_{\hat{\mathbf{Y}}} + \sum_{j=1}^{n_{doe}} \left(\sum_{i=2}^P \mathbf{a}_i^t \hat{\boldsymbol{\phi}}_j \phi_i(\boldsymbol{\xi}) \right) \hat{\boldsymbol{\phi}}_j. \quad (37)$$

References

- [1] E. Castillo, *Extreme Value Theory in Engineering*, Vol. 1, Elsevier Inc., 1988.
- [2] J. Azaïs, M. Wschebor, *Level Sets and Extrema of Random Processes and Fields*, Vol. 1, John Wiley & Sons, Inc., 2008.
- [3] F. Poirion, Karhunen Loève expansion and distribution of non-Gaussian process maximum, *Probab. Eng. Mech.* 43 (2016) 85–90. <http://dx.doi.org/10.1016/j.probengmech.2015.12.005>. URL <http://www.sciencedirect.com/science/article/pii/S0266892015300667>.
- [4] S.O. Rice, Mathematical analysis of random noise, *Bell Syst. Tech. J.* 23 (3) (1944) 282–332. <http://dx.doi.org/10.1002/j.1538-7305.1944.tb00874.x>.
- [5] R.G. Ghanem, P.D. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Springer, New York, NY, 1991.
- [6] D. Xiu, G.E. Karniadakis, The Wiener-Askey polynomial chaos for stochastic differential equations, *SIAM J. Sci. Comput.* 24 (2) (2002) 619–644. <http://dx.doi.org/10.1137/S1064827501387826>.
- [7] C. Soize, R. Ghanem, Physical systems with random uncertainties: Chaos representations with arbitrary probability measure, *SIAM J. Sci. Comput.* 26 (2) (2004) 395–410.
- [8] G. Poëtte, D. Lucor, Non intrusive iterative stochastic spectral representation with application to compressible gas dynamics, *J. Comput. Phys.* 231 (9) (2012) 3587–3609. <http://dx.doi.org/10.1016/j.jcp.2011.12.038>. URL <http://www.sciencedirect.com/science/article/pii/S0021999112000058>.
- [9] X. Wan, G.E. Karniadakis, An adaptive multi-element generalized polynomial chaos method for stochastic differential equations, *J. Comput. Phys.* 209 (2) (2005) 617–642. <http://dx.doi.org/10.1016/j.jcp.2005.03.023>. URL <http://www.sciencedirect.com/science/article/pii/S0021999105001919>.
- [10] M. Berveiller, B. Sudret, M. Lemaire, Stochastic finite elements: a non-intrusive approach by regression, *Eur. J. Comput. Mech.* 15 (1–3) (2006) 81–92.
- [11] G. Blatman, B. Sudret, Adaptive sparse polynomial chaos expansion based on least angle regression, *J. Comput. Phys.* 230 (6) (2011) 2345–2367. <http://dx.doi.org/10.1016/j.jcp.2010.12.021>. URL <http://www.sciencedirect.com/science/article/pii/S0021999110006856>.
- [12] D. Xiu, J.S. Hesthaven, High-order collocation methods for differential equations with random inputs, *SIAM J. Sci. Comput.* 27 (3) (2005) 1118–1139. <http://dx.doi.org/10.1137/040615201>.
- [13] S. Abraham, M. Raisee, G. Ghorbaniasl, F. Contino, C. Lacor, A robust and efficient stepwise regression method for building sparse polynomial chaos expansions, *J. Comput. Phys.* 332 (2017) 461–474. <http://dx.doi.org/10.1016/j.jcp.2016.12.015>. URL <http://www.sciencedirect.com/science/article/pii/S0021999116306684>.
- [14] J. Hampton, A. Doostan, Coherence motivated sampling and convergence analysis of least squares polynomial chaos regression, *Comput. Methods Appl. Mech. Engrg.* 290 (2015) 73–97. <http://dx.doi.org/10.1016/j.cma.2015.02.006>. URL <http://www.sciencedirect.com/science/article/pii/S004578251500047X>.
- [15] N.V. Queipo, R.T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, P.K. Tucker, Surrogate-based analysis and optimization, *Progr. Aerosp. Sci.* 41 (1) (2005) 1–28.
- [16] A.I. Forrester, A.J. Keane, Recent advances in surrogate-based optimization, *Prog. Aerosp. Sci.* 45 (1) (2009) 50–79.
- [17] T.W. Simpson, T.M. Mauery, J.J. Korte, F. Mistree, Kriging models for global approximation in simulation-based multidisciplinary design optimization, *AIAA J.* 39 (12) (2001) 2233–2241.
- [18] J.P. Kleijnen, *Design and Analysis of Simulation Experiments*, Vol. 230, Springer, 2015.
- [19] F.A. Viana, C. Gogu, R.T. Haftka, Making the most out of surrogate models: tricks of the trade, in: *Proceedings of the 2010 ASME IDETC and CIE*, Montreal, Canada, Paper No. DETC2010-28813, 2010.
- [20] D.R. Jones, M. Schonlau, W.J. Welch, Efficient global optimization of expensive black-box functions, *J. Global Optim.* 13 (4) (1998) 455–492. <http://dx.doi.org/10.1023/A:1008306431147>.
- [21] A. Sobester, S.J. Leary, A.J. Keane, On the design of optimization strategies based on global response surface approximation models, *J. Global Optim.* 33 (1) (2005) 31–59.
- [22] S. Shan, G.G. Wang, Survey of modeling and optimization strategies to solve high-dimensional design problems with computationally-expensive black-box functions, *Struct. Multidiscip. Optim.* 41 (2) (2010) 219–241.
- [23] S.M. Göhler, T. Eifler, T.J. Howard, Robustness metrics: Consolidating the multiple approaches to quantify robustness, *ASME. J. Mech. Des.* 138 (11) (2016) 111407–111407–12.
- [24] B.J. Williams, *Sequential Design of Computer Experiments to Minimize Integrated Response Functions* (Ph.D. thesis), The Ohio State University, 2000 AAI9983009.
- [25] J. Janusevskis, R. Le Riche, Simultaneous kriging-based estimation and optimization of mean response, *J. Global Optim.* 55 (2) (2013) 313–336. <http://dx.doi.org/10.1007/s10898-011-9836-5>.
- [26] J. Zhang, A. Taflanidis, J. Medina, Sequential approximate optimization for design under uncertainty problems utilizing kriging metamodeling in augmented input space, *Comput. Methods Appl. Mech. Engrg.* 315 (2017) 369–395. <http://dx.doi.org/10.1016/j.cma.2016.10.042>. URL <http://www.sciencedirect.com/science/article/pii/S0045782516314529>.
- [27] N.M. Alexandrov, J.E. Dennis, R.M. Lewis, V. Torczon, A trust-region framework for managing the use of approximation models in optimization, *Struct. Multidiscip. Optim.* 15 (1) (1998) 16–23.
- [28] A.R. Conn, N.I. Gould, P.L. Toint, *Trust Region Methods*, Society for Industrial and Applied Mathematics, 2000.
- [29] D.P. Kouri, M. Heinkenschloss, D. Ridzal, B.G. van Bloemen Waanders, A trust-region algorithm with adaptive stochastic collocation for PDE optimization under uncertainty, *SIAM J. Sci. Comput.* 35 (4) (2013) A1847–A1879.
- [30] D.P. Kouri, M. Heinkenschloss, D. Ridzal, B.G. van Bloemen Waanders, Inexact objective function evaluations in a trust-region algorithm for PDE-constrained optimization under uncertainty, *SIAM J. Sci. Comput.* 36 (6) (2014) A3011–A3029.

- [31] A.A. Taflanidis, J.C. Medina, Simulation-based optimization in design-under-uncertainty problems through iterative development of metamodels in augmented design/random variable space, *Simul. Model. Methodol. Technol. Appl.* (2015) 251–273.
- [32] B. Sudret, *Uncertainty Propagation and Sensitivity Analysis in Mechanical Models Contributions to Structural Reliability and Stochastic Spectral Methods*, Tech. Rep., Université Blaise Pascal - Clermont II, 2007.
- [33] W. Betz, I. Papaioannou, D. Straub, Numerical methods for the discretization of random fields by means of the karhunen loève expansion, *Comput. Methods Appl. Mech. Engrg.* 271 (2014) 109–129. <http://dx.doi.org/10.1016/j.cma.2013.12.010>. URL <http://www.sciencedirect.com/science/article/pii/S0045782513003502>.
- [34] B. Pascual, S. Adhikari, Hybrid perturbation-polynomial chaos approaches to the random algebraic eigenvalue problem, *Comput. Methods Appl. Mech. Engrg.* 217–220 (2012) 153–167. <http://dx.doi.org/10.1016/j.cma.2012.01.009>. URL <http://www.sciencedirect.com/science/article/pii/S0045782512000205>.
- [35] P. Kersaudy, B. Sudret, N. Varsier, O. Picon, J. Wiart, A new surrogate modeling technique combining kriging and polynomial chaos expansions application to uncertainty analysis in computational dosimetry, *J. Comput. Phys.* 286 (2015) 103–117. <http://dx.doi.org/10.1016/j.jcp.2015.01.034>. URL <http://www.sciencedirect.com/science/article/pii/S0021999115000388>.
- [36] R.H. Cameron, W.T. Martin, The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals, *Ann. of Math.* 48 (2) (1947) 385–392 URL <http://www.jstor.org/stable/1969178>.
- [37] R. Lebrun, A. Dutfoy, A generalization of the nataf transformation to distributions with elliptical copula, *Probab. Eng. Mech.* 24 (2) (2009) 172–178. <http://dx.doi.org/10.1016/j.proengmech.2008.05.001>. URL <http://www.sciencedirect.com/science/article/pii/S0266892008000507>.
- [38] V. Yadav, S. Rahman, Adaptive-sparse polynomial dimensional decomposition methods for high-dimensional stochastic computing, *Comput. Methods Appl. Mech. Engrg.* 274 (2014) 56–83. <http://dx.doi.org/10.1016/j.cma.2014.01.027>. URL <http://www.sciencedirect.com/science/article/pii/S0045782514000504>.
- [39] Z. Perko, L. Gilli, D. Lathouwers, J. Kloosterman, Grid and basis adaptive polynomial chaos techniques for sensitivity and uncertainty analysis, *J. Comput. Phys.* 260 (2014) 54–84. <http://dx.doi.org/10.1016/j.jcp.2013.12.025>. URL <http://www.sciencedirect.com/science/article/pii/S0021999113008322>.
- [40] J.G. Winokur, *Adaptive Sparse Grid Approaches to Polynomial Chaos Expansions for Uncertainty Quantification* (Ph.D. thesis), Duke University, 2015. <http://dukespace.lib.duke.edu/dspace/handle/10161/9845>.
- [41] D.G. Krige, *A Statistical Approach to Some Mine Evaluations and Allied Problems at the Witwatersrand* (Master's thesis), University of Witwatersrand, 1951.
- [42] C.E. Rasmussen, C.K.I. Williams, *Gaussian Processes for Machine Learning*, Adaptive Computation and Machine Learning, MIT Press, Cambridge, Mass, 2006.
- [43] A. Forrester, A. Sobester, A. Keane, *Engineering Design Via Surrogate Modelling: A Practical Guide*, John Wiley & Sons, Hoboken, NJ, 2008.
- [44] M. Arnst, R. Ghanem, E. Phipps, J. Red-Horse, Dimension reduction in stochastic modeling of coupled problems, *Internat. J. Numer. Methods Engrg.* 92 (11) (2012) 940–968. <http://dx.doi.org/10.1002/nme.4364>.
- [45] M.J. Sasena, *Flexibility and Efficiency Enhancements for Constrained Global Design Optimization with Kriging Approximations* (Ph.D. thesis), University of Michigan, 2002.
- [46] M. Buhmann, *Radial Basis Functions: Theory and Implementations*, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, 2003.
- [47] M.J.D. Powell, Direct search algorithms for optimization calculations, *Acta Numer.* 7 (1998) 287–336. <http://dx.doi.org/10.1017/S0962492900002841>.
- [48] D. Kraft, *A Software Package for Sequential Quadratic Programming*, Tech. Rep. DFVLR-FB-88-28, DLR German Aerospace Center Institute for Flight Mechanics, Koln, Germany, 1988.
- [49] T. Derbyshire, K. Sidwell, PAN AIR Summary Document, Tech. Rep. NASA Contractor Report 3250, NASA, 1982.
- [50] J. Katz, A. Plotkin, *Low-Speed Aerodynamics*, second ed., 2001. <http://dx.doi.org/10.1017/CBO9780511810329>.
- [51] M.A. Bouhlel, N. Bartoli, A. Otsmane, J. Morlier, Improving kriging surrogates of high-dimensional design models by partial least squares dimension reduction, *Struct. Multidiscip. Optim.* 53 (5) (2016) 935–952. <http://dx.doi.org/10.1007/s00158-015-1395-9>.
- [52] K. Konakli, B. Sudret, Polynomial meta-models with canonical low-rank approximations: Numerical insights and comparison to sparse polynomial chaos expansions, *J. Comput. Phys.* 321 (2016) 1144–1169. <http://dx.doi.org/10.1016/j.jcp.2016.06.005>. URL <http://www.sciencedirect.com/science/article/pii/S0021999116302303>.
- [53] M. Chevreuil, A. Nouy, Model order reduction based on proper generalized decomposition for the propagation of uncertainties in structural dynamics, *Internat. J. Numer. Methods Engrg.* 89 (2) (2012) 241–268. <http://dx.doi.org/10.1002/nme.3249>.