

UNCERTAINTY QUANTIFICATION IN HIGH-DIMENSIONAL SPACES WITH LOW-RANK TENSOR APPROXIMATIONS

Katerina Konakli¹, and Bruno Sudret¹

¹Chair of Risk, Safety and Uncertainty Quantification, ETH Zurich
Stefano-Franscini-Platz 5, 8093 Zurich, Switzerland
e-mail: {konakli, sudret}@ibk.baug.ethz.ch

Keywords: uncertainty propagation, meta-models, polynomial chaos expansions, canonical decomposition, low-rank approximations.

Abstract. *Polynomial chaos expansions have proven powerful for emulating responses of computational models with random input in a wide range of applications. However, they suffer from the curse of dimensionality, meaning the exponential growth of the number of unknown coefficients with the input dimension. By exploiting the tensor product form of the polynomial basis, low-rank approximations drastically reduce the number of unknown coefficients, thus providing a promising tool for effectively dealing with high-dimensional problems. In this paper, first, we investigate the construction of low-rank approximations with greedy approaches, where the coefficients along each dimension are sequentially updated and the rank of the decomposition is progressively increased. Furthermore, we demonstrate the efficiency of the approach in different applications, also in comparison with state-of-art methods of polynomial chaos expansions.*

1 INTRODUCTION

Surrogate modeling is an effective approach for propagation of uncertainty from the input to a response quantity of interest through complex computational models. In this, a computational model is substituted by a *meta-model* that possesses similar statistical properties with the original model, but has a simple functional form and is therefore easy to evaluate.

Polynomial Chaos Expansions (PCE) constitute a popular meta-modeling technique, which has proven its efficiency in a wide range of applications. The key concept in PCE is to expand the response quantity of interest onto a basis of orthonormal multivariate polynomials obtained as tensor products of univariate polynomials in the random input parameters. By construction, the size of the multivariate basis and thus the number of unknown coefficients grow exponentially with the input dimension, rendering the classical PCE approach inefficient in high-dimensional problems. A remedy to this limitation is the use of sparse PCE, as demonstrated in [1, 2, 3]. Recently, a new class of models called Low-Rank Approximations (LRA) has been introduced ([4, 5, 6]) and is further investigated in this paper.

LRA exploit the tensor-product form of the polynomial basis to express the random response as a sum of a few rank-one functions. Such representations drastically decrease the number of unknown coefficients, which grows only linearly with the input dimension. Existing algorithms for building LRA are based on greedy approaches, where the polynomial coefficients along each dimension are sequentially updated and the rank of the approximation is progressively increased. The sizes of the associated error-minimization problems may be orders of magnitudes smaller than those in the classical PCE approach. However, such constructions involve open questions that call for further investigations.

In this paper, we shed light on the construction of LRA by investigating stopping criteria in the sequential updating of the polynomial coefficients and optimal rank selection based on cross-validation. Moreover, we confront LRA with state-of-art PCE methods and demonstrate the singular efficiency of the former in example applications involving the deflection of a beam, an analytical benchmark function and heat diffusion with spatially varying diffusion coefficient.

2 UNCERTAINTY PROPAGATION WITH META-MODELS

2.1 Mathematical setting

We consider a computational model, \mathcal{M} , with M -dimensional random input, \mathbf{X} . Due to the uncertainties in the input, the response quantity of interest, Y , becomes random. Therefore, the computational model represents the map

$$\mathbf{X} \in \mathcal{D}_{\mathbf{X}} \subset \mathbb{R}^M \mapsto Y = \mathcal{M}(\mathbf{X}) \in \mathbb{R}, \quad (1)$$

where $\mathcal{D}_{\mathbf{X}}$ is the support of \mathbf{X} . In general, the map $Y = \mathcal{M}(\mathbf{X})$ is not known analytically and may correspond to a complex process.

A meta-model $\widehat{\mathcal{M}}(\mathbf{X})$ mimics the behavior of $\mathcal{M}(\mathbf{X})$, while having a simple functional form. We herein focus on *non-intrusive* meta-modeling approaches, in which the original computational model is treated as a "black box". Building a meta-model in a non-intrusive manner requires an Experimental Design (ED) comprising a set of realizations of the input vector, $\mathcal{E} = \{\boldsymbol{\chi}^{(1)}, \dots, \boldsymbol{\chi}^{(N)}\}$, and the corresponding evaluations of the original model, $\mathcal{Y} = \{\mathcal{M}(\boldsymbol{\chi}^{(1)}), \dots, \mathcal{M}(\boldsymbol{\chi}^{(N)})\}$.

2.2 Error estimates

For a set of realizations of the input vector, $\mathcal{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \subset \mathcal{D}_{\mathbf{X}}$, and a function $\mathbf{x} \in \mathcal{D}_{\mathbf{X}} \mapsto a(\mathbf{x}) \in \mathbb{R}$, we introduce the discrete L_2 semi-norm

$$\|a\|_{\mathcal{X}} = \left(\frac{1}{n} \sum_{i=1}^n a^2(\mathbf{x}_i) \right)^{1/2}. \quad (2)$$

A good measure of accuracy of the meta-model response, \widehat{Y} , is the mean-square error of the difference $Y - \widehat{Y}$, called *generalization error*. Employing the notation in Eq. (2), the generalization error is estimated by

$$\widehat{Err}_G = \left\| \mathcal{M} - \widehat{\mathcal{M}} \right\|_{\mathcal{X}_{\text{val}}}^2, \quad (3)$$

where $\mathcal{X}_{\text{val}} = \{\mathbf{x}_1, \dots, \mathbf{x}_{n_{\text{val}}}\}$ is a sufficiently large set of realizations of the input vector, denoted *validation set*. The estimate of the relative generalization error, \widehat{err}_G , is obtained by normalizing \widehat{Err}_G with the empirical variance of $\mathcal{Y}_{\text{val}} = \{\mathcal{M}(\mathbf{x}_1), \dots, \mathcal{M}(\mathbf{x}_{n_{\text{val}}})\}$.

However, meta-models are typically used in cases when a large number of model evaluations is not affordable. It is thus desirable to estimate the generalization error using only the information contained in the ED. One such error estimate is the empirical error, \widehat{Err}_E , given by

$$\widehat{Err}_E = \left\| \mathcal{M} - \widehat{\mathcal{M}} \right\|_{\mathcal{E}}^2, \quad (4)$$

where the index \mathcal{E} emphasizes that the average is carried out over the ED. The relative empirical error, \widehat{err}_E , is obtained by normalizing \widehat{Err}_E with the empirical variance of $\mathcal{Y} = \{\mathcal{M}(\boldsymbol{\chi}^{(1)}), \dots, \mathcal{M}(\boldsymbol{\chi}^{(N)})\}$. Unfortunately, the empirical error tends to underestimate the generalization error, which might be severe in cases of overfitting.

A good compromise between accurate error estimation and affordable computational cost is the use of Cross-Validation (CV) techniques. The CV-based error measures can provide fair approximations of the generalization error by relying only on the ED. In the general case of k -fold cross-validation, first, the ED is randomly partitioned into k sets of approximately equal size. Then, a meta-model is built considering all but one of the partitions and the excluded set is used to evaluate the generalization error; by alternating through the k sets, k meta-models are obtained. The average generalization error of the k meta-models provides an estimate of the generalization error of the meta-model built with the full ED.

3 POLYNOMIAL CHAOS EXPANSIONS

We assume that the components of $\mathbf{X} = \{X_1, \dots, X_M\}$ are independent with joint Probability Density Function (PDF) $f_{\mathbf{X}}(\mathbf{x})$ and marginal PDFs $f_{X_i}(x_i)$, $i = 1, \dots, M$. A PCE approximation of $Y = \mathcal{M}(\mathbf{X})$ in Eq. (1) has the form ([7])

$$\widehat{Y} = \mathcal{M}^{\text{PCE}}(\mathbf{X}) = \sum_{\boldsymbol{\alpha} \in \mathcal{A}} y_{\boldsymbol{\alpha}} \Psi_{\boldsymbol{\alpha}}(\mathbf{X}), \quad (5)$$

where $\{\Psi_{\boldsymbol{\alpha}}, \boldsymbol{\alpha} \in \mathcal{A}\}$ is a set of multivariate polynomials that are orthonormal with respect to $f_{\mathbf{X}}$, $\{y_{\boldsymbol{\alpha}}, \boldsymbol{\alpha} \in \mathcal{A}\}$ is the set of corresponding polynomial coefficients and $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_M)$ are multi-indices.

The multivariate polynomials that comprise the PCE basis are obtained by tensorization of appropriate univariate polynomials, *i.e.*

$$\Psi_{\alpha}(\mathbf{X}) = \prod_{i=1}^M P_{\alpha_i}^{(i)}(X_i), \quad (6)$$

where $P_{\alpha_i}^{(i)}$ is a polynomial of degree α_i in the i -th input variable belonging to a family of polynomials that are orthonormal with respect to f_{X_i} . For instance, a uniform variable with support $[-1, 1]$ is associated with the family of Legendre polynomials, whereas a standard normal variable is associated with the family of Hermite polynomials. Other cases can be treated through an isoprobabilistic transformation of \mathbf{X} to a basic random vector. Cases with mutually dependent input variables can also be treated through an isoprobabilistic transformation (*e.g.* Nataf transformation) to a vector of independent standard variables.

The set of multi-indices \mathcal{A} in Eq. (5) is determined by an appropriate truncation scheme. A common scheme, also employed in the subsequent example applications (Section 5), consists in selecting multivariate polynomials up to a total degree p^t , *i.e.* $\{\Psi_{\alpha}, \alpha \in \mathbb{N}^M : \sum_{i=1}^M \alpha_i \leq p^t\}$. The corresponding number of terms in the truncated series is

$$\text{card}\mathcal{A} = \binom{M + p^t}{p^t} = \frac{(M + p^t)!}{M!p^t!}. \quad (7)$$

For other advanced truncation schemes, the reader is referred to [2]. Note in Eq. (7) the exponential growth of the number of terms with M , which is known as the *curse of dimensionality*.

Once the basis has been specified, the set of coefficients $\mathbf{y} = \{y_{\alpha}, \alpha \in \mathcal{A}\}$ may be computed by minimizing the mean-square error of the approximation over the ED, *i.e.*

$$\mathbf{y} = \arg \min_{\mathbf{v} \in \mathbb{R}^{\text{card}\mathcal{A}}} \left\| \mathcal{M} - \sum_{\alpha \in \mathcal{A}} v_{\alpha} \Psi_{\alpha} \right\|_{\mathcal{E}}^2. \quad (8)$$

Eq. (8) represents an Ordinary Least-Squares (OLS) minimization problem that can be solved using well-known techniques. However, the required size of ED may be prohibitively large in cases with large M . More efficient solutions schemes can be devised by considering respective regularized problems. For example, the Least Angle Regression (LAR) method ([8]) disregards insignificant terms from the set of predictors, thus yielding *sparse* meta-models. A variation is the hybrid LAR method ([2]), which employs the LAR algorithm to select the best set of predictors and subsequently, estimates the coefficients using OLS.

A good measure of the accuracy of a PCE meta-model is the Leave-One Out (LOO) error, obtained by CV with $k = N$ (see Section 2.2). According to the description of the CV approach, this would require building N different meta-models, by sequentially setting apart each point of the ED. However, algebraic manipulations allow evaluation of the LOO error from a *single* PCE based on the full ED: Let us denote by h_i the i -th diagonal term of matrix $\Psi(\Psi^T\Psi)^{-1}\Psi^T$, where $\Psi = \{\Psi_{ij} = \Psi_j(\mathbf{x}^{(i)}), i = 1, \dots, N; j = 1, \dots, \text{card}\mathcal{A}\}$. Then, the LOO error can be computed as

$$\widehat{Err}_{LOO} = \frac{1}{N} \sum_{i=1}^N \left(\frac{\mathcal{M}(\mathbf{x}^{(i)}) - \mathcal{M}^{\text{PCE}}(\mathbf{x}^{(i)})}{1 - h_i} \right)^2. \quad (9)$$

The relative LOO error, \widehat{err}_{LOO} , is obtained by normalizing \widehat{Err}_{LOO} with the empirical variance of $\mathcal{Y} = \{\mathcal{M}(\mathbf{x}^{(1)}), \dots, \mathcal{M}(\mathbf{x}^{(N)})\}$. Because \widehat{err}_{LOO} may be too optimistic, one may use

instead a corrected estimate, given by ([9])

$$\widehat{err}_{LOO}^* = \widehat{err}_{LOO} \left(1 - \frac{\text{card}\mathcal{A}}{N} \right)^{-1} (1 + \text{tr}((\mathbf{\Psi}^T \mathbf{\Psi})^{-1})). \quad (10)$$

4 LOW-RANK APPROXIMATIONS

A canonical decomposition of $Y = \mathcal{M}(\mathbf{X})$ in Eq. (1) has the form

$$\widehat{Y} = \sum_{l=1}^R b_l \left(\prod_{i=1}^M v_l^{(i)}(X_i) \right), \quad (11)$$

where $v_l^{(i)}$ denotes a univariate function of X_i and b_l is the normalizing coefficient of the rank-1 function in the parenthesis. The number of terms R in Eq. (11) defines the rank of the decomposition, leading to the name Low-Rank Approximations (LRA) in cases when R is small.

Of interest herein are LRA with $v_l^{(i)}(X_i)$ expanded onto a polynomial basis that is orthonormal with respect to the marginal f_{X_i} (see Section 3), *i.e.*

$$\widehat{Y} = \sum_{l=1}^R b_l \left(\prod_{i=1}^M \left(\sum_{k=0}^{p_i} z_{k,l}^{(i)} P_k^{(i)}(X_i) \right) \right), \quad (12)$$

where $P_k^{(i)}$ is the k -th degree univariate polynomial in the i -th input variable, of maximum degree p_i , and $z_{k,l}^{(i)}$ is the coefficient of $P_k^{(i)}$ in the l -th rank-1 term. A representation of $Y = \mathcal{M}(\mathbf{X})$ in the form of Eq. (12) drastically reduces the number of unknowns compared to Eq. (5). For example, when $p_i = p$ for $i = 1, \dots, M$, the number of unknowns is $P = ((p+1)M+1)R$, which grows only linearly with M .

LRA can be efficiently built with greedy approaches that involve progressive increase of the rank of the decomposition by adding rank-1 terms and successive updating of the polynomial coefficients along separate dimensions (see *e.g.* [4, 6]). We herein adopt the skeleton of the algorithm proposed by [4], comprising a sequence of pairs of a *correction step*, where a rank-1 tensor is built, and an *updating step*, where the normalizing coefficients are determined/updated. Details are given in the sequel.

Let \widehat{Y}_r denote the rank- r approximation of $Y = \mathcal{M}(\mathbf{X})$, *i.e.*

$$\widehat{Y}_r = \sum_{l=1}^r b_l w_l, \quad (13)$$

with

$$w_l = \prod_{i=1}^M \left(\sum_{k=0}^{p_i} z_{k,l}^{(i)} P_k^{(i)}(X_i) \right), \quad (14)$$

and let $R_r = Y - \widehat{Y}_r$ denote the corresponding residual.

In the r -th correction step, the rank-1 tensor w_r is built, by solving the minimization problem

$$w_r = \arg \min_{\omega \in \mathcal{W}} \|R_{r-1} - \omega\|_{\mathcal{E}}^2, \quad (15)$$

where \mathcal{W} denotes the space of rank-1 tensors. In the first correction step, one has $R_0 = Y$. Eq. (15) is solved by first, assigning initial arbitrary values to the polynomial coefficients and then, successively updating the coefficients in each dimension $i = 1, \dots, M$, while "freezing" the coefficients in all other dimensions at their current values. Thus, the coefficients in dimension j , $\mathbf{z}_r^{(j)} = \{z_{1,r}^{(j)} \dots z_{p_j,r}^{(j)}\}$, are obtained as the solution of

$$\mathbf{z}_r^{(j)} = \arg \min_{\zeta \in \mathbb{R}^{p_j}} \left\| R_{r-1} - \left(\prod_{i \neq j} \sum_{k=0}^{p_i} z_{k,r}^{(i)} P_k^{(i)} \right) \left(\sum_{k=0}^{p_j} \zeta_k P_k^{(j)} \right) \right\|_{\mathcal{E}}^2. \quad (16)$$

The above minimization problem is of size $p_j + 1$, where typically $p_j < 20$, and can be easily solved using OLS.

Note that a correction step may involve several iterations over the set of dimensions; the optimal number of iterations remains an open question. We herein propose a stopping criterion that combines the number of iterations over the set $\{1, \dots, M\}$, denoted I_r , and the decrease in the relative empirical error between two successive iterations, denoted $\Delta \widehat{err}_r$. The relative empirical error, \widehat{err}_r , is obtained by normalizing

$$\widehat{err}_r = \|R_{r-1} - w_r\|_{\mathcal{E}}^2 \quad (17)$$

with the empirical variance of $\mathcal{Y} = \{\mathcal{M}(\boldsymbol{\chi}^{(1)}), \dots, \mathcal{M}(\boldsymbol{\chi}^{(N)})\}$. According to the proposed criterion, the algorithm exits the r -th correction step if either I_r reaches a maximum allowable value, I_{max} , or $\Delta \widehat{err}_r$ becomes smaller than a threshold, $\Delta \widehat{err}_{min}$.

After the completion of a correction step, the algorithm moves to an updating step, in which the coefficients $\mathbf{b} = \{b_1, \dots, b_r\}$ are obtained as the solution of

$$\mathbf{b} = \arg \min_{\beta \in \mathbb{R}^r} \left\| \mathcal{M} - \sum_{l=1}^r \beta_l w_l \right\|_{\mathcal{E}}^2. \quad (18)$$

In the r -th updating step, the value of the element b_r is determined for the first time, whereas the values of the elements $\{b_1, \dots, b_{r-1}\}$ are updated from their previous values. The minimization problem in Eq. (18) is of size r (recall that small ranks are of interest in LRA) and can be easily solved using OLS.

In the above algorithm, the progressive adding of rank-1 terms results in a set of LRA of rank $\{1, \dots, r\}$ at the r -th step. [5] propose selection of the optimal rank, $R \in \{1, \dots, r_{max}\}$, with 3-fold CV (see Section 2.2). By using 3-fold CV, one obtains three meta-models for each rank $\{1, \dots, r_{max}\}$ as well as the respective error estimates. The rank R corresponding to the smallest average error over the three meta-models is identified as optimal. Then, a new meta-model of rank R is built using the full ED. The average error for the selected rank provides an estimate of the generalization error of the final LRA.

5 EXAMPLE APPLICATIONS

5.1 Beam deflection

In this example, we use LRA to represent the mid-span deflection of a simply supported beam under a concentrated load acting at the midpoint of the span. Table 1 shows the distributions of the $M = 5$ input random variables of the model, *i.e.* the width, b , and height, h , of the rectangular cross section, the length, L , the Young's modulus, E , and the load, P . All input

Table 1: Distributions of input random variables.

Variable	Distribution	mean	CoV
b (m)	Lognormal	0.15	0.05
h (m)	Lognormal	0.3	0.05
L (m)	Lognormal	5	0.01
E (MN/m ²)	Lognormal	3e4	0.15
P (MN)	Lognormal	0.01	0.20

random variables are independent. The mid-span deflection as a function of the input variables reads $Y = PL^3/4Ebh^3$.

In the following analysis, we consider ED of varying sizes drawn with Sobol sampling and a validation set of size $n_{\text{val}} = 10^6$ drawn with Monte Carlo Simulation (MCS). We note that a validation set is typically not available in a real-case scenario, but is used herein to obtain reliable estimates of the generalization errors and thus, assess the accuracy of our approaches. Employing an isoprobabilistic transformation of the input variables to standard normal variables, Hermite polynomials are used to build the basis functions.

We first investigate selection of optimal rank among a set of candidate values $\{1, \dots, 20\}$. After preliminary investigations, the common polynomial degree is set to $p_1 = \dots = p_5 = 5$ and the stopping criterion in the correction step is defined by $I_{\text{max}} = 50$ and $\Delta \widehat{err}_{\text{min}} = 10^{-8}$. For different sizes of the ED, the left graph of Figure 1 compares the rank identified as optimal using the 3-fold CV approach with the actual optimal rank leading to the minimum generalization error; the corresponding relative generalization errors are shown in the right graph of the same figure. We observe that although the two ranks do not coincide in all cases, the differences in the corresponding generalization errors are negligible. We highlight the accuracy of LRA, even when the size of the ED is as small as $N = 50$.

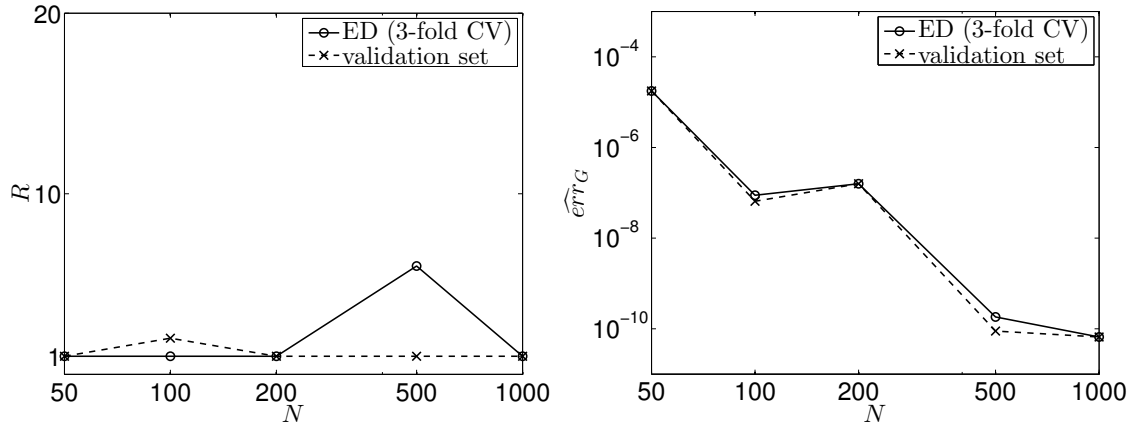


Figure 1: Beam deflection: Selected rank (left) and corresponding relative generalization error (right) for varying sizes of the experimental design.

Next, we examine optimal values of the error threshold in the correction step, while other parameters are fixed to their values above. For three different sizes of the ED and $\Delta \widehat{err}_{\text{min}}$ varying from 10^{-9} to 10^{-4} , the left graph of Figure 2 shows the relative generalization errors for ranks selected with 3-fold CV. The right graph of the figure shows the corresponding total number of iterations in the 20 steps, $\sum I_r$. We observe that for the smaller ED, setting a smaller threshold $\Delta \widehat{err}_{\text{min}}$ significantly improves the LRA accuracy, but increases the required number

of iterations; the value of $\Delta\widehat{err}_{min}$ has a lesser effect for the larger ED.

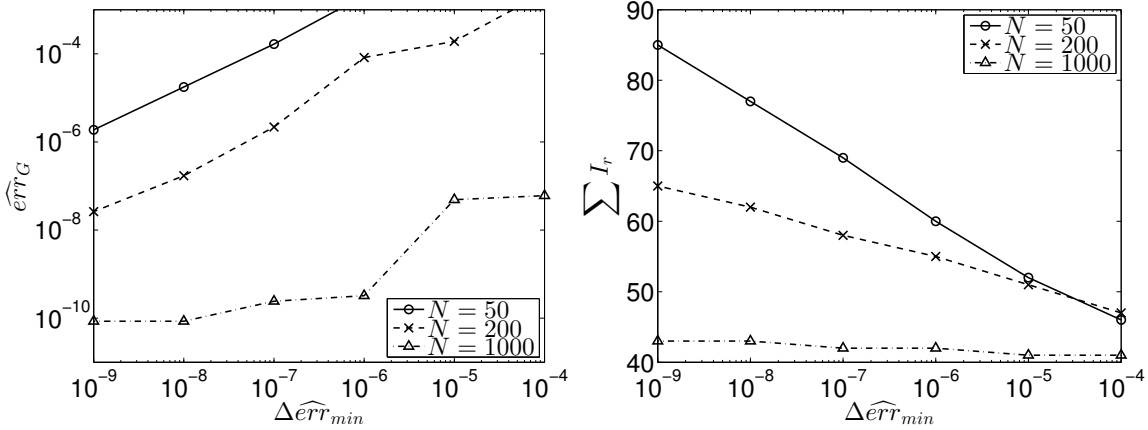


Figure 2: Beam deflection: Relative generalization error (left) and corresponding total number of iterations (right) for varying error thresholds in the stopping criterion.

We now confront LRA with PCE, using the same polynomial family to build the bases and computing the PCE coefficients with the hybrid LAR method (see Section 3). For both approaches, we consider meta-models of optimal polynomial degrees: in LRA, the optimal common maximum degree, p , is selected with 3-fold CV, whereas in PCE, the optimal maximum total degree, p^t , is identified by means of the corrected LOO error (see Eq. (10)). Other parameters in LRA are the same as in Figure 1. For N varying in 20 – 500, the left and right graphs of Figure 3 respectively show the selected polynomial degrees and the corresponding relative generalization errors. For all considered N , LRA outperform PCE, yielding meta-models that are 2 to 3 orders of magnitude more accurate. It is remarkable that only $N = 20$ points suffice to build LRA with an accuracy of the order of 10^{-4} .

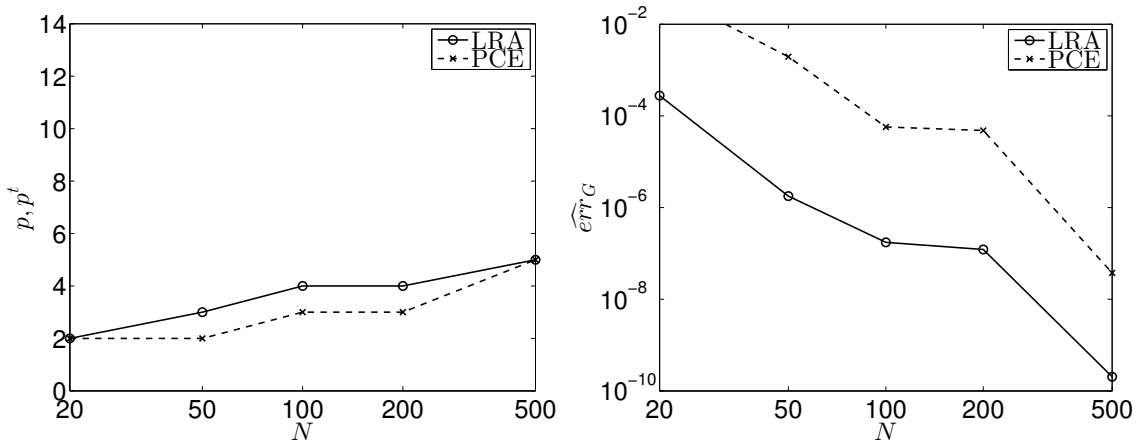


Figure 3: Beam deflection: Optimal polynomial degrees (left) and corresponding relative generalization errors (right) of LRA and PCE meta-models for varying sizes of the experimental design.

5.2 Sobol function

In this example, we use LRA to approximate the Sobol function, given by

$$Y = \prod_{i=1}^M \frac{|4X_i - 2| + c_i}{1 + c_i}, \quad (19)$$

where $\mathbf{X} = \{X_1, \dots, X_M\}$ are independent random variables uniformly distributed over $[0, 1]$ and $\mathbf{c} = \{c_1, \dots, c_M\}$ are non-negative constants. We herein consider the case with $M = 8$ and $\mathbf{c} = \{1, 2, 5, 10, 20, 50, 100, 500\}^T$. As in the previous example, we employ Sobol sampling to draw ED of varying sizes and MCS to draw a validation set of size $n_{\text{val}} = 10^6$.

We confront LRA with PCE, using Legendre polynomials for the bases and computing the PCE coefficients with the hybrid LAR method. As in the previous example, the optimal degrees are identified by means of 3-fold CV and the corrected LOO error in LRA and PCE, respectively. After preliminary investigations, the parameters in the stopping criterion in LRA are set to $I_{\text{max}} = 50$ and $\Delta \widehat{e}_{\text{TR}}^{\text{min}} = 10^{-4}$ (in this case, use of a smaller $\Delta \widehat{e}_{\text{TR}}^{\text{min}}$ was found to have a minor effect on the accuracy of the meta-model, while significantly increasing the number of iterations). For N varying in $50 - 500$, the left graph of Figure 4 shows the selected polynomial degrees, whereas the right graph shows the corresponding relative generalization errors. For all considered N , LRA clearly outperform PCE. Note that the small regression problems involved in the construction of LRA allow use of higher-degree polynomials that are more appropriate for representing the Sobol function (recall that p is the polynomial degree of each univariate function in LRA, whereas p^t is the total degree of the multivariate basis of the PCE).

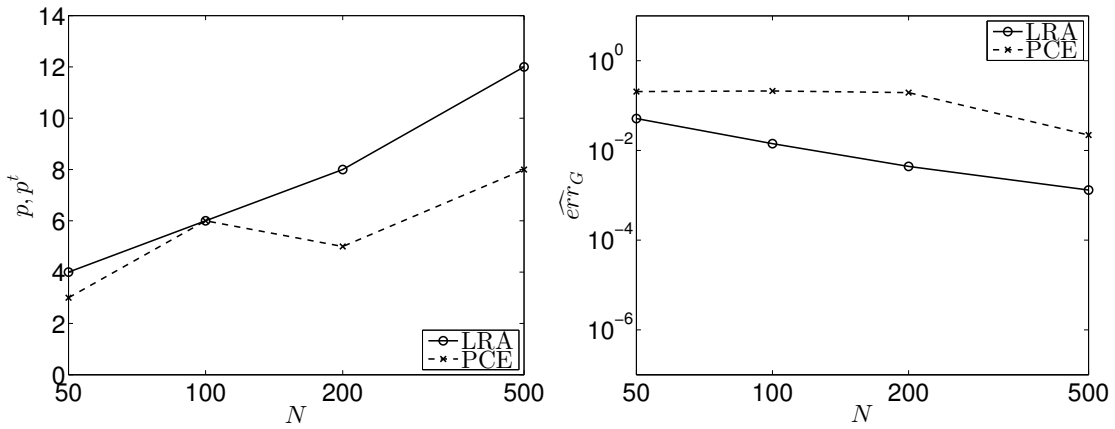


Figure 4: Sobol function: Optimal polynomial degrees (left) and corresponding relative generalization errors (right) of LRA and PCE meta-models for varying sizes of the experimental design.

5.3 Heat diffusion with spatially varying diffusion coefficient

We consider a two-dimensional stationary heat diffusion problem defined on the square domain $D = (-0.5, 0.5) \times (-0.5, 0.5)$ shown in Figure 5, where the temperature field $T(\mathbf{z})$, $\mathbf{z} \in D$, is described by the partial differential equation

$$-\nabla(\kappa(\mathbf{z}) \nabla T(\mathbf{z})) = 500I_A(\mathbf{z}), \quad (20)$$

with boundary conditions $T = 0$ on the top boundary and $\nabla T \mathbf{n} = 0$ on the left, right and bottom boundaries, where \mathbf{n} denotes the vector normal to the boundary. In Eq. (20), $A =$

$(0.2, 0.3) \times (0.2, 0.3)$ is a square domain within D (see Figure 5) and I_A is the indicator function equal to 1 if $\mathbf{z} \in A$ and 0 otherwise. The diffusion coefficient, $\kappa(\mathbf{z})$, is a lognormal random field defined by

$$\kappa(\mathbf{z}) = \exp[a_\kappa + b_\kappa g(\mathbf{z})], \quad (21)$$

where $g(\mathbf{z})$ is a standard Gaussian random field and the parameters a_κ and b_κ are such that the mean and standard deviation of κ are $\mu_\kappa = 1$ and $\sigma_\kappa = 0.3$, respectively. The random field $g(\mathbf{z})$ is characterized by an autocorrelation function $\rho(\mathbf{z}, \mathbf{z}') = \exp(-\|\mathbf{z} - \mathbf{z}'\|^2/0.2^2)$. The quantity of interest, Y , is the average temperature in the square domain $B = (-0.3, -0.2) \times (-0.3, -0.2)$ within D (see Figure 5).

To facilitate solution of the problem, the random field $g(\mathbf{z})$ is represented using the Expansion Optimal Linear Estimation (EOLE) method ([10]). By truncating the EOLE series after the first M terms, $g(\mathbf{z})$ is approximated by

$$\hat{g}(\mathbf{z}) = \sum_{i=1}^M \frac{\xi_i}{\sqrt{l_i}} \phi_i^\top \mathbf{C}_{z\zeta}. \quad (22)$$

In the above equation, $\{\xi_1, \dots, \xi_M\}$ are independent standard normal variables; $\mathbf{C}_{z\zeta}$ is a vector with elements $\mathbf{C}_{z\zeta}^{(k)} = \rho(\mathbf{z}, \zeta_k)$, where $\{\zeta_1, \dots, \zeta_n\}$ are the points of an appropriately defined mesh in D ; and (l_i, ϕ_i) are the eigenvalues and eigenvectors of the correlation matrix $\mathbf{C}_{\zeta\zeta}$ with elements $\mathbf{C}_{\zeta\zeta}^{(k,l)} = \rho(\zeta_k, \zeta_l)$, where $k, l = 1, \dots, n$. We select $M = 53$ in order to satisfy

$$\sum_{i=1}^M l_i / \sum_{i=1}^n l_i \geq 0.99. \quad (23)$$

The underlying deterministic problem is solved with an in-house finite-element analysis code developed in Matlab environment. The employed finite-element discretization with triangular Delaunay elements is shown in Figure 5.

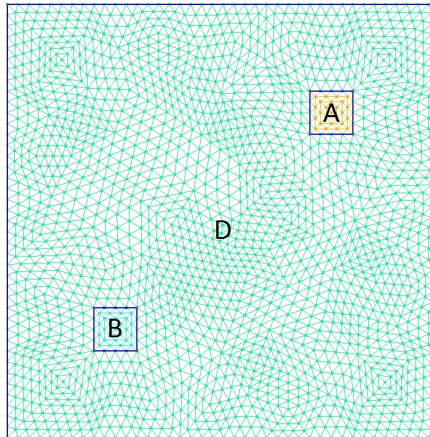


Figure 5: Domain and finite element mesh in heat diffusion problem.

In the following, we confront LRA with PCE (in conjunction with hybrid LAR), in which the basis functions are made of Hermite polynomials. We use an ED of size $N = 200$ drawn with Sobol sampling to build the meta-models and a MCS-based validation set of size $n_{\text{val}} = 10^4$ to assess their accuracy. Figure 6 shows the temperature fields corresponding to two samples of

the ED. The polynomial degree and rank of the optimal LRA meta-model (based on 3-fold CV) are $p = 1$ and $R = 1$, respectively; the total polynomial degree of the optimal PCE meta-model is $p^t = 1$ (based on the LOO error). In Figure 7, we compare the kernel smoothing estimates of the PDF of Y obtained with the LRA and PCE meta-models to the reference PDF based on the validation set. In the left graph, the PDFs are shown in the normal scale, whereas in the right graph, they are shown in a logarithmic scale to highlight the behavior at the tails. The PDF estimate obtained with the LRA meta-model is obviously superior, providing a fairly good approximation of the reference PDF over the whole range of the response values.

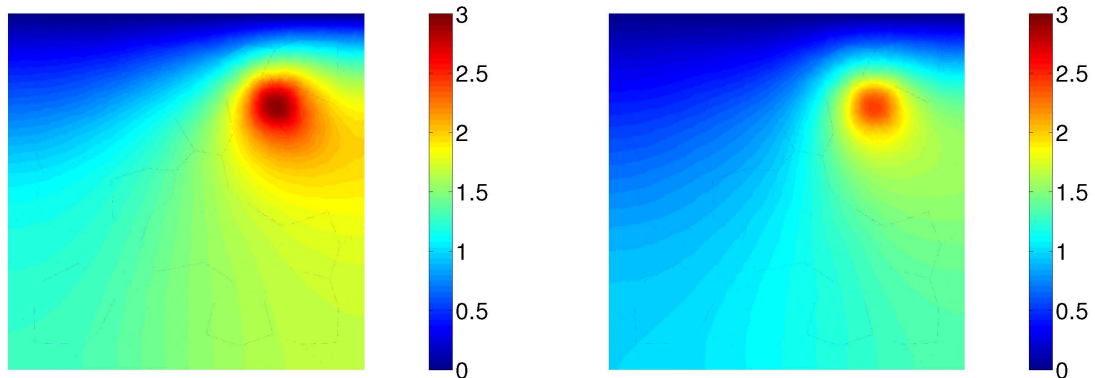


Figure 6: Heat diffusion problem: Example realizations of the temperature field.

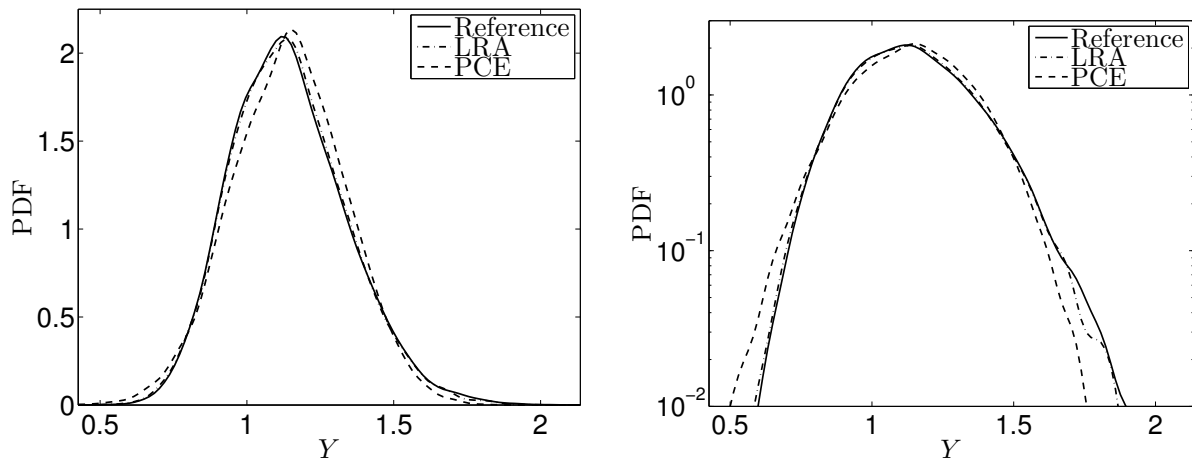


Figure 7: Heat diffusion problem : Comparison of the PDF estimates based on LRA and PCE (using an experimental design of size $N = 200$) to the reference PDF (left: normal scale, right: logarithmic scale).

6 CONCLUSIONS

Low-Rank Approximations (LRA) comprise an effective tool for uncertainty propagation in problems with high-dimensional input. In this paper, we considered LRA built with polynomial functions and shed light on their construction with greedy approaches. Furthermore, we demonstrated the accuracy of LRA in example applications involving the deflection of a simply-supported beam, the Sobol function and a two-dimensional stationary heat-diffusion problem

where the diffusion coefficient is a random field. In the same applications, we showed that LRA outperform the popular meta-modeling technique of polynomial chaos expansions, by using the same polynomial families to build the basis functions.

REFERENCES

- [1] G. Blatman, B. Sudret, An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. *Probabilistic Engineering Mechanics*, **25**, 183–197, 2010.
- [2] G. Blatman, B. Sudret, Adaptive sparse polynomial chaos expansion based on least angle regression. *Journal of Computational Physics*, **230**, 2345–2367, 2011.
- [3] A. Doostan, H. Owhadi, A non-adapted sparse approximation of PDEs with stochastic inputs. *Journal of Computational Physics*, **230**, 3015–3034, 2011.
- [4] M. Chevreuil, O. Rai, A. Nouy, Sampling based tensor approximation method for uncertainty propagation. *Proceedings of the 11th International Conference on Structural Safety and Reliability*, New York, June 2013.
- [5] M. Chevreuil, R. Lebrun, A. Nouy, P. Rai, A least-squares method for sparse low rank approximation of multivariate functions. *arXiv preprint arXiv:1305.0030*, 2013.
- [6] A. Doostan, A. Validi, G. Iaccarino, Non-intrusive low-rank separated approximation of high-dimensional stochastic models. *Comput. Methods Appl. Mech. Eng.*, **263**, 42–55, 2013.
- [7] D. Xiu, G.E. Karniadakis, The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, **24**, 619–644, 2002.
- [8] B. Efron, T. Hastie, I. Johnstone, B. Tibshirani, Least angle regression. *Annals of Statistics*, **32**, 407–499, 2004.
- [9] O. Chapelle, V. Vapnik, Y. Bengio, Model selection for small sample regression. *Machine Learning*, **48**, 9–23, 2002.
- [10] C. C. Li, A. Der Kiureghian, Optimal discretization of random fields. *Journal of Engineering Mechanics*, **119**, 1136–1154, 1993.