Global sensitivity analysis in high dimensions with PLS-PCE

Max Ehre, Iason Papaioannou, Daniel Straub

Engineering Risk Analysis Group, Technische Universität München, Arcisstr. 21, 80290 München, Germany

Abstract

Global sensitivity analysis is a central part of uncertainty quantification with engineering models. Variance-based sensitivity measures such as Sobol' and total-effect indices are amongst the most popular and commonly used tools for global sensitivity analysis. Mutiple sampling-based estimators of these measures are available, but they often come at considerable computational cost due to the large number of required model evaluations. If the computational model is expensive to evaluate, these approaches are quickly rendered infeasible. An alternative is the use of surrogate models, which reduce the computational cost per sample significantly. This contribution focuses on a recently introduced latent-variable-based polynomial chaos expansion (PCE) based on partial least squares (PLS) analysis, which is particularly suitable for high-dimensional problems. We develop an efficient way of computing variance-based sensitivities with the PLS-PCE surrogate. By back-transforming the surrogate model from its latent variable space-basis to the original input variable space-basis, we derive analytical expressions for the sought sensitivities. These expressions depend on the surrogate model coefficients exclusively. Thus, once the surrogate model is built, the variance-based sensitivities can be computed at negligible computational cost and no additional sampling is required. The accuracy of the method is demonstrated with two numerical experiments of an elastic truss and a thin steel plate.

Keywords: Uncertainty quantification, Global sensitivity analysis, Surrogate modelling, PLS-PCE, Dimensionality reduction, High dimensions,

1. Introduction

Surrogate models have received much attention due to their potential of alleviating computational cost significantly in applications requiring elaborate and expensive numerical models, see e.g. [1, 2, 3]. A common example is the propagation of uncertainties through computationally intensive numerical models. The general concept of surrogate modeling techniques is to establish an abstract, parametrized input-output-relation that has similar properties as the original model. The parameters of the surrogate model are determined based on a finite set of original model evaluations, to maximize similarity between the surrogate and the original model according to a suitable criterion. Subsequently, the surrogate model can be used to cheaply approximate the original model and, in the context of uncertainty quantification, compute statistics of the output or a quantity of interest derived thereof.

In many scenarios, a statistical characterization of the pure model output is less important than an analysis of its sensitivity with respect to changes and variability in the model inputs. Surrogate models have also proven useful in efficiently performing model sensitivity analysis - an otherwise computationally intensive task. Sensitivity analysis is a collection of measures and tools designed to determine how the random inputs and/or deterministic model parameters of a model influence its output or a quantity of interest derived from the output. Amongst these, one can discern local (derivative-based) [4, 5, 6, 7, 8] and global [9, 10, 11, 12, 13] sensitivity measures. Local sensitivities are suitable to determine the impact an input has on the output in the vicinity of a nominal value, by virtue of the model structure. However, they neglect the global significance of the input. Global sensitivity measures on the other hand take into account the entire input variable support as well as the variability of the inputs over their support. Regression-based sensitivity measures aim

at linearly regressing the output on its inputs to identify global sensitivity indices; this approach works well if the output depends approximately linearly on the inputs [11, 12]. A second category of global sensitivity measures is referred to as ANOVA (ANalysis Of VAriance) [14]. Generally speaking, these measures aim at quantifying an input variable's influence (or that of a combination of inputs) through identifying the fraction of output variance it causes. A recent, third category of global sensitivity measures can be summarized under the terms 'moment-independent' or 'distribution-based' [15, 16, 17, 18, 19, 20]. The underlying idea is to quantify the sensitivity of the output to a given input through the distance between the output density conditional on the given input from its unconditional counterpart. Our work focusses on variance-based sensitivity measures. The most commonly used variance-based measures are the Sobol' index [21] and the total-effect index [22], which can be computed using Monte Carlo methods [9, 23, 24, 25, 26], Fourier analysis [27, 28] or surrogate models as in [29, 30]. The works of [29, 30] have derived variance-based sensitivity measures directly from the model coefficients of conventional polynomial chaos expansions (PCE) [31] and polynomial-based low-rank approximations (LRA) [32], respectively.

Along these lines, we derive global, variance-based sensitivity measures for the model output from the coefficients of basis-adapted PCEs [33]. The basic idea of basis adaptation is to identify a low dimensional latent variable space and construct a PCE in this space. We focus on a recently introduced approach for identifying the latent variables and computing the corresponding PCE coefficients termed partial least squares-driven polynomial chaos expansion (PLS-PCE) [34]. PLS-PCE allows application of PCEs in very high dimensions. By back-transforming the PCE from the latent variable space to the original input variable space, we enable estimation of the sensitivity indices as with the standard PCE model [29].

The paper is structured as follows: In Section 2, we review the PLS-PCE surrogate model, its construction and some important properties. In Section 3, we give a brief introduction to variance-based sensitivity analysis and its application in the context of polynomial basis surrogate models. In Section 4, we develop the methodology to compute sensitivities based on the model coefficients. In Section 5, we demonstrate the new method based on two numerical examples and in Section 6 we provide some concluding remarks.

2. Partial least squares and polynomial chaos expansions

Let $\boldsymbol{X}$ be a random vector on the outcome space $\mathbb{R}^d$ with joint cumulative distribution function (CDF) $F_{\boldsymbol{X}}$ and $\mathcal{Y}(\boldsymbol{X}) = Y \in \mathbb{R}$. If $\mathcal{Y}$ is square-integrable, i.e. $\mathbb{E}_{\boldsymbol{X}}[\mathcal{Y}(\boldsymbol{X})^2] < \infty$, it belongs to a Hilbert space $\mathcal{H}$ with inner product of any two functions $g, h \in \mathcal{H}$

$$\langle g(\boldsymbol{X}), h(\boldsymbol{X})\rangle_{\mathcal{H}} = \mathbb{E}_{\boldsymbol{X}}[g(\boldsymbol{X})h(\boldsymbol{X})] \tag{1}$$

$$= \int_{\mathbb{R}^d} g(\boldsymbol{x})h(\boldsymbol{x})f_{\boldsymbol{X}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \tag{2}$$

where $f_{\boldsymbol{X}}(\boldsymbol{x})$ is the joint probability density function (PDF) of $\boldsymbol{X}$. g and h are orthogonal if

$$\langle g(\boldsymbol{x}), h(\boldsymbol{x})\rangle_{\mathcal{H}} = \mathbb{E}_{\boldsymbol{X}}[g(\boldsymbol{X})h(\boldsymbol{X})] = 0. \tag{3}$$

Note, that if g and h can be written as products of univariate functions g_i and h_i, $i = 1, ..., d$, of the components of $\boldsymbol{X}$ and these components are statistically independent, the following holds:

$$\langle g(\boldsymbol{x}), h(\boldsymbol{x})\rangle_{\mathcal{H}} = \prod_{i=1}^{d} \mathbb{E}_{X_i}[g_i(X_i)h_i(X_i)]. \tag{4}$$

2.1. Polynomial Chaos Expansion

Given a complete and orthonormal basis of $\mathcal{H}$, $\{h_i(\boldsymbol{X}), i \in \mathbb{N}\}$, Y may be expressed as a linear combination of the basis functions:

$$Y = \mathcal{Y}(\boldsymbol{X}) = \sum_{i=0}^{\infty} b_i h_i(\boldsymbol{X}). \tag{5}$$

Then, since $\mathcal{Y} \in \mathcal{H}$, the approximation

$$\widehat{Y}_n = \widehat{\mathcal{Y}}(\boldsymbol{X}) = \sum_{i=0}^{n} b_i h_i(\boldsymbol{X}) \tag{6}$$

asymptotically ($n \to \infty$) converges to Y in the mean-square sense. Henceforth, without loss of generality, we will consider the case $F_{\boldsymbol{X}} = \Phi_d$, where Φ_d denotes the d-variate independent standard-normal CDF. If the joint PDF of $\boldsymbol{X}$ is known, one can express $\boldsymbol{X}$ as a function of standard normal random variables through an iso-probabilistic transformation [35]. Then, one can construct an orthonormal polynomial basis of $\mathcal{H}$ using products of one-dimensional normalized Hermite polynomials

$$\Psi_{\boldsymbol{k}}(\boldsymbol{X}) = \prod_{i=1}^{d} \psi_{k_i}(X_i) \tag{7}$$

where $\{\psi_i(X), i \in \mathbb{N}\}$ are the normalized (probabilist) Hermite polynomials and $\boldsymbol{k} = (k_1, \ldots, k_d) \in \mathbb{N}^d$. The PCE of maximum total order p reads

$$\hat{Y}_p = \sum_{|\boldsymbol{k}| \leq p} b_{\boldsymbol{k}} \Psi_{\boldsymbol{k}}(\boldsymbol{X}). \tag{8}$$

The total number of basis functions in the PCE P is given combinatorially in terms of the dimensions d and the maximum total polynomial order p:

$$P = \binom{d+p}{p}. \tag{9}$$

The coefficients $\boldsymbol{b}$ are computed through a projection of $\mathcal{Y}$ onto the space spanned by $\{\Psi_{\boldsymbol{k}}, |\boldsymbol{k}| \leq p\}$, where the projection can be transformed into an equivalent ordinary least squares (OLS) problem [36]. Equation (9) indicates a fast growth of the associated regression problem with increasing dimension d, rendering PCEs intractable for high-dimensional problems. Sparse PCE methods have been proposed to relax this constraint by solving a modified, $\mathcal{L}_1$-regularized least-squares problem, which penalizes the number of terms in the expansion and thus reduces P [37]. This is also known under the term 'compressive sampling/sensing' [38, 39]. Nevertheless, the computation of a sparse PCE still requires computing the entirety of all possible basis elements, which can become a second (combinatorial) bottleneck in addition to the solution of the regression problem.

2.2. Basis adaptation

In order to address this problem, one may rotate the PCE representation onto a new basis defined by the new variables $\boldsymbol{Z} = \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{X}$, where $\boldsymbol{Q} \in \mathbb{R}^{d \times d}$ and $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q} = \boldsymbol{I}$, with $\boldsymbol{I}$ denoting the identity matrix. Then, an equivalent PCE representation is given by [33]

$$\hat{Y}_p^{\boldsymbol{Q}} = \sum_{|\boldsymbol{k}| \leq p} a_{\boldsymbol{k}} \Psi_{\boldsymbol{k}}(\boldsymbol{Z}) = \sum_{|\boldsymbol{k}| \leq p} a_{\boldsymbol{k}} \Psi_{\boldsymbol{k}}\left(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{X}\right). \tag{10}$$

The coordinate transformation allows for the construction of PCEs along important directions of the problem input space. These directions are defined by linear combinations of the original variable vector $\boldsymbol{X}$, the coefficients of which are stored in the rows of $\boldsymbol{Q}$. Then, by retaining only the $m < d$ most important directions in $\boldsymbol{Q}$, one obtains a matrix $\boldsymbol{Q}_m$ and the corresponding PCE reads

$$\hat{Y}_p^{\boldsymbol{Q}_m} = \sum_{|\tilde{\boldsymbol{k}}| \leq p} a_{\tilde{\boldsymbol{k}}} \Psi_{\tilde{\boldsymbol{k}}}\left(\boldsymbol{Q}_m^{\mathrm{T}}\boldsymbol{X}\right), \tag{11}$$

where $\tilde{\boldsymbol{k}} \in \mathbb{N}^m$. [33] compute the basis adaptation $\boldsymbol{Q}_m$ by evaluating first- or second-order PCE coefficients only with a sparse-grid numerical quadrature. [40] couple this approach with compressive sensing to simultaneously identify $\boldsymbol{Q}_m$ and the PCE coefficients in the latent space. In [34], we show that important directions can be identified efficiently based on a set of original function evaluations via partial least squares (PLS). The next section summarizes the approach.

2.3. Partial least squares-based PCE

PLS finds a relationship between variables $\boldsymbol{X}$ and Y based on N observations of both quantities [41, 42, 43]. $\mathcal{X} \in \mathbb{R}^{N\times d}$ stores observations from $\boldsymbol{X}$ and $\mathcal{Y} \in \mathbb{R}^{N\times 1}$ stores the corresponding responses. PLS sequentially identifies latent components $\boldsymbol{t}_i \in \mathbb{R}^{N\times 1}$ such that they have maximum covariance with $\mathcal{Y}$. After determining each $\boldsymbol{t}_i$, PLS assumes a linear relationship between $\boldsymbol{t}_i$ and $\mathcal{Y}$ and evaluates the corresponding coefficient a_i of $\boldsymbol{t}_i$ by OLS. After each sequence, the matrices $\mathcal{X}$ and $\mathcal{Y}$ are deflated by the contribution of the i-th PLS-component. Components are extracted until a certain error criterion is met, which can be formulated e.g. through the norm of the residual response vector or via cross-validation.

The nonlinear version of PLS in turn relaxes the assumption of a linear relationship between latent component and the response. A number of nonlinear PLS algorithms have been proposed [44]. Here we employ the approach of [45, 46], which introduces an additional loop into the algorithm for running a Newton-Raphson procedure iterating between the current latent component and the response. In the context of PCE, the nonlinear relationship between the $\{\boldsymbol{t}_i\}_{i=1,\dots,m}$ and the response is a one-dimensional Hermite polynomial expansion [34]. The coefficients of the PLS-driven PCE can be computed simultaneously with the latent variable structure as a byproduct of the PLS algorithm. Ultimately, the nonlinear PCE-driven

Algorithm 1 PCE-driven PLS algorithm [34]

Input Data matrix $\mathcal{X}$ and response matrix $\mathcal{Y}$
Center matrices: $\mathcal{X} \leftarrow \mathcal{X} - \bar{\mathcal{X}}$, $\mathcal{Y} \leftarrow \mathcal{Y} - \bar{\mathcal{Y}}$
Set $\mathcal{E} = \mathcal{X}$, $\mathcal{F} = \mathcal{Y}$, $i = 1$
repeat
  Compute weight: $\boldsymbol{w}_i^0 = \mathcal{E}^{\mathrm{T}}\mathcal{F}/\|\mathcal{E}^{\mathrm{T}}\mathcal{F}\|$
  for $q \leftarrow 1, p$ **do**
    Set $\boldsymbol{w}_i^q = \boldsymbol{w}_i^0$
    repeat
      Compute score: $\boldsymbol{t}_i^q = \mathcal{E}\boldsymbol{w}_i^q$
      Fit a 1D PCE of order q: $\hat{\boldsymbol{a}}_i^q \leftarrow \text{fit}\left[\mathcal{F} = (\boldsymbol{a}_i^q)^{\mathrm{T}}\boldsymbol{\psi}_q(\boldsymbol{t}_i^q) + \boldsymbol{\epsilon}\right]$
      Set $\hat{\mathcal{M}}_i^q(t) = (\hat{\boldsymbol{a}}_i^q)^{\mathrm{T}}\boldsymbol{\psi}_q(\boldsymbol{t}_i^q)(t)$
      Compute the error: $\hat{\mathcal{F}} = (\hat{\boldsymbol{a}}_i^q)^T\boldsymbol{\psi}_q(\boldsymbol{t}_i^q)$; $\boldsymbol{e} = \mathcal{F} - \hat{\mathcal{F}}$
      Compute: $\Delta\boldsymbol{w}_i^q = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{e}$ with $\boldsymbol{A} = \nabla_{\boldsymbol{w}}(\hat{\boldsymbol{a}}_i^q)^{\mathrm{T}}\boldsymbol{\psi}_q(\mathcal{E}\boldsymbol{w})$
      Set: $\boldsymbol{w}_i^q \leftarrow \boldsymbol{w}_i^q + \Delta\boldsymbol{w}_i^q$
      Normalize: $\boldsymbol{w}_i^q \leftarrow \boldsymbol{w}_i^q/\|\boldsymbol{w}_i^q\|$
    until $\|\Delta\boldsymbol{w}_i^q\|$ is smaller than ϵ_w
    Evaluate the relative leave-one-out error ϵ_{LOO}^q as in [37]
  end for
  Set $\{q_i, \hat{\boldsymbol{a}}_i^{q_i}, \boldsymbol{w}_i^{q_i}\}$ as the triple $\{q, \hat{\boldsymbol{a}}_i^q, \boldsymbol{w}_i^q\}$ with the smallest ϵ_{LOO}^q
  Compute score: $\boldsymbol{t}_i^{q_i} = \mathcal{E}\boldsymbol{w}_i^{q_i}$
  Compute load: $\boldsymbol{p}_i^{q_i} = \mathcal{E}^{\mathrm{T}}\boldsymbol{t}_i^{q_i}/((\boldsymbol{t}_i^{q_i})^{\mathrm{T}}\boldsymbol{t}_i^{q_i})$
  Deflate: $\mathcal{E} \leftarrow \mathcal{E} - \boldsymbol{t}_i^{q_i}(\boldsymbol{p}_i^{q_i})^{\mathrm{T}}$, $\mathcal{F} \leftarrow \mathcal{F} - (\hat{\boldsymbol{a}}_i^{q_i})^{\mathrm{T}}\boldsymbol{\psi}_{q_i}(\boldsymbol{t}_i^{q_i})$
  $i \leftarrow i + 1$
until change in $\|\mathcal{F}\|$ is smaller than ϵ_y
return $\{q_i, \hat{\boldsymbol{a}}_i^{q_i}, \boldsymbol{t}_i^{q_i}, \boldsymbol{w}_i^{q_i}, \boldsymbol{p}_i^{q_i}\}, i = 1\dots, m.$

PLS algorithm, which is developed in [34] and summarized in Algorithm 1, identifies m latent components. For each component, it returns the direction $\boldsymbol{r}_i$ and the 1-dimensional PCE along this direction, which is defined by its polynomial order q_i and the coefficient vector $\boldsymbol{a}_i$. The polynomial order is identified with leave-one-out cross validation. For each (i-th) latent component, the nonlinear PLS iteration is repeated for different polynomial orders and q_i is chosen as the order minimizing the leave-one-out error. The PLS-PCE reads

$$\hat{Y}_m^{\mathrm{PLS}} = a_0 + \sum_{i=1}^{m}(\boldsymbol{a}_i^{q_i})^T\boldsymbol{\psi}_{q_i}\left[(\boldsymbol{r}_i)^{\mathrm{T}}\tilde{\boldsymbol{X}}\right], \tag{12}$$

where $a_0 = \widehat{\mathbb{E}}[\mathcal{Y}]$, $\boldsymbol{\psi}_{q_i}(\boldsymbol{X})$ is a vector function assembling the evaluations of the one-dimensional Hermite polynomials up to order q_i and $\tilde{\boldsymbol{X}} = \boldsymbol{X} - \boldsymbol{\mu}_{\mathcal{X}}$, where $\boldsymbol{\mu}_{\mathcal{X}}$ is the columnwise sample mean of the training data $\mathcal{X}$. The PLS directions $\boldsymbol{r}_i$ can be evaluated in terms of the PLS weights $\boldsymbol{w}_i$ and loads $\boldsymbol{p}_i$ computed by Algorithm 1 through the following recursive relation [47]:

$$\begin{aligned} \boldsymbol{r}_1 &= \boldsymbol{w}_1 \\ \boldsymbol{r}_i &= \boldsymbol{w}_i - \boldsymbol{r}_{i-1}\left(\boldsymbol{p}_{i-1}^{\mathrm{T}}\boldsymbol{w}_i\right). \end{aligned} \tag{13}$$

Let $\boldsymbol{R} = [\boldsymbol{r}_1, \ldots, \boldsymbol{r}_m] \in \mathbb{R}^{d\times m}$ be the matrix that collects all PLS directions. The matrix $\boldsymbol{R}$ is not necessarilty orthogonal, i.e. in general $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R} \neq \boldsymbol{I}$. However, in [34] it is shown that $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R} \to \boldsymbol{I}$ as $N \to \infty$ and hence Eq. (12) is asymptotically equivalent to a PCE of the form of Eq. (11), where only the main effects in the latent components are considered.

3. Global sensitivity analysis

3.1. Variance-based sensitivity analysis

The idea behind variance-based sensitivity analysis for model outputs Y is to decompose $\mathbb{V}[Y]$ into partial variances that are attributable to variable combinations in the input $\boldsymbol{X}$. If $\boldsymbol{X}$ is jointly uniform on $[0,1]$ and its components are independent, this is accomplished by projecting Y onto a unique, orthogonal basis with respect to the uniform joint density. The representation of Y is then the Sobol'-Hoeffding decomposition [21], which reads:

$$f(\boldsymbol{X}) = f_0 + \sum_{i=1}^{d} f_i(X_i) + \sum_{i=1}^{d}\sum_{j=i+1}^{d} f_{ij}(X_i, X_j) + \cdots + f_{12\ldots d}(\boldsymbol{X}). \tag{14}$$

Each summand in equation (14) represents the influence of a distinct variable subset of $\boldsymbol{X}$, $\boldsymbol{X}_{\mathcal{A}}$, and due to the orthogonality property, the partial variance associated with $\mathcal{A}$ is given immediately by $\mathbb{V}[f_{\mathcal{A}}]$. The Sobol' index is then the ratio of the partial variance due to $f_{\mathcal{A}}$ and the total variance [21]:

$$S_{Y,\mathcal{A}} = \mathbb{V}[f_{\mathcal{A}}]/\mathbb{V}[Y]. \tag{15}$$

Alternatively, one can utilize the closed Sobol' index [48], which is based on the partial variance contributed by $\boldsymbol{X}_{\mathcal{A}}$ and any subset of $\boldsymbol{X}_{\mathcal{A}}$, i.e.,

$$S_{Y,\mathcal{A}}^{clo} = \sum_{\mathcal{B}\subseteq\mathcal{A}} \mathbb{V}[f_{\mathcal{B}}]/\mathbb{V}[Y]. \tag{16}$$

While the Sobol' and closed Sobol' indices are identical for single variables, i.e., $\mathrm{card}(\mathcal{A}) = 1$, the former represents the net interaction in between all elements of $\boldsymbol{X}_{\mathcal{A}}$ and the latter represents the total contribution of all elements of $\boldsymbol{X}_{\mathcal{A}}$. Finally, the total-effect index S^T [22, 48] is based on the partial variance contributed by all variable combinations containing any element from $\boldsymbol{X}_{\mathcal{A}}$, such that

$$S_{Y,\mathcal{A}}^{T} = \sum_{\mathcal{A}\cap\mathcal{B}\neq\emptyset} \mathbb{V}[f_{\mathcal{B}}]/\mathbb{V}[Y]. \tag{17}$$

The decomposition of equation (14) is generalizable to arbitrary joint distributions with independent components through an iso-probabilistic transformation.

3.2. PCE-based sensitivity analysis

A major benefit of representing a response $Y \in \mathcal{H}$ as $\hat{Y}$ with an orthogonal basis of $\mathcal{H}$ lies in the simplicity of finding statistical properties of $\hat{Y}$ and thus - if the model accurately represents Y - approximately of Y. Given the model representation (8) with P terms, e.g. the first two moments can be computed as

$$\mathbb{E}[\hat{Y}] = b_{\boldsymbol{0}}, \quad \mathbb{V}[\hat{Y}] = \sum_{0<|\boldsymbol{k}|\leq p} b_{\boldsymbol{k}}^2. \tag{18}$$

Moreover, [29] showed that the indices $S_{\hat{Y},\mathcal{A}}$ and $S^T_{\hat{Y},\mathcal{A}}$ of representation (8) can also be found merely by post-processing its coefficients $\boldsymbol{b}$. For a given subset of the input variables denoted by the index set $\mathcal{A}$, we define a boolean index vector $\mathcal{I}^{\mathcal{A}} \in \{0,1\}^{d\times 1}$ s.t. $\mathcal{I}^{\mathcal{A}}_i = 0$ if $i \notin \mathcal{A}$ and $\mathcal{I}^{\mathcal{A}}_i = 1$ if $i \in \mathcal{A}$. In the same way, we define such a boolean vector for the multi-index $\boldsymbol{k}$ s.t. $\mathcal{I}^{\boldsymbol{k}}_i = 0$ if $k_i = 0$ and $\mathcal{I}^{\boldsymbol{k}}_i = 1$ if $k_i > 0$. Then, the PCE-based sensitivity indices read

$$\hat{S}_{\hat{Y},\mathcal{A}} = \frac{1}{\mathbb{V}[\hat{Y}]} \sum_{\substack{\mathcal{I}^{\mathcal{A}}=\mathcal{I}^{\boldsymbol{k}},\\ 0<|\boldsymbol{k}|\leq p}} b^2_{\boldsymbol{k}}, \quad \hat{S}^{clo}_{\hat{Y},\mathcal{A}} = \frac{1}{\mathbb{V}[\hat{Y}]} \sum_{\substack{\mathcal{I}^{\mathcal{A}}-\mathcal{I}^{\boldsymbol{k}}\geq \boldsymbol{0},\\ 0<|\boldsymbol{k}|\leq p}} b^2_{\boldsymbol{k}}, \quad \hat{S}^T_{\hat{Y},\mathcal{A}} = \frac{1}{\mathbb{V}[\hat{Y}]} \sum_{\substack{(\mathcal{I}^{\mathcal{A}})^{\mathrm{T}}\mathcal{I}^{\boldsymbol{k}}\neq 0,\\ 0<|\boldsymbol{k}|\leq p}} b^2_{\boldsymbol{k}}. \tag{19}$$

4. Global sensitivity analysis with PLS-PCE

Here we derive expressions for $S_{\hat{Y}}$ and $S^T_{\hat{Y}}$ for $\hat{Y}$ of the form (12). Note, that if the columns of $\boldsymbol{R}$ form an orthonormal basis, i.e., if $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R} = \boldsymbol{I}$, the sensitivity indices of any latent variable component $Z_i = \boldsymbol{r}_i^{\mathrm{T}}\tilde{\boldsymbol{X}}$ can be obtained immediately as

$$S_{\hat{Y}^{\mathrm{PLS}}_m,Z_i} = S^{clo}_{\hat{Y}^{\mathrm{PLS}}_m,Z_i} = S^T_{\hat{Y}^{\mathrm{PLS}}_m,Z_i} = \sum_{j=1}^{q_i}(a^{q_i}_{ij})^2 \Bigg/ \sum_{i=1}^{m}\sum_{j=1}^{q_i}(a^{q_i}_{ij})^2. \tag{20}$$

However, as discussed in Section 2.3, the condition on $\boldsymbol{R}$ only holds asymptotically as $N \to \infty$. In practice, the sensitivity indices associated with the latent variables are of less interest than those of the original inputs. That is, one is interested in computing sensitivities of $\hat{Y}^{\mathrm{PLS}}_m$ to the original input vector $\boldsymbol{X}$ rather than $\boldsymbol{Z}$. For convenience, we restate the format of $\hat{Y}^{\mathrm{PLS}}_m$:

$$\hat{Y}^{\mathrm{PLS}}_m = a_0 + \sum_{i=1}^{m}(\boldsymbol{a}^{q_i}_i)^{\mathrm{T}}\boldsymbol{\psi}_{q_i}\left[(\boldsymbol{r}_i)^{\mathrm{T}}(\boldsymbol{X}-\boldsymbol{\mu}_{\mathcal{X}})\right]. \tag{21}$$

In the following two subsections we derive expressions and state corresponding algorithms for computing the Sobol' and total-effect indices of the PLS-PCE model response with respect to $\boldsymbol{X}$ for both large and small sample sizes.

4.1. Computation in the asymptotic limit $N \to \infty$

Asymptotically, i.e. for $N \to \infty$, we have

$$\lim_{N\to\infty} \boldsymbol{\mu}_{\mathcal{X}} = \boldsymbol{0}$$

and [34] proves that

$$\lim_{N\to\infty} \|\boldsymbol{r}_i\| = 1 \quad i = 2,\ldots,m,$$

while the first PLS-direction $\|\boldsymbol{r}_1\|$ always has length 1. [49] provides a multinomial theorem for non-normalized probabilist's Hermite polynomial of order k, which is restated in the following for the normalized poylnomials.

Theorem 1. *Let $j \in \mathbb{N}_0$ be the polynomial order of the normalized probabilist's Hermite polynomial ψ_j, $d \in \mathbb{N}$ and $\boldsymbol{X} \in \mathbb{R}^{d\times 1}$. Let further $\boldsymbol{k} \in \mathbb{N}^d_0$ be an index set and $\boldsymbol{s} \in \mathbb{R}^{d\times 1}$ such that $\sum_{i=1}^d s_i^2 = 1$. Then,*

$$\begin{aligned}\psi_j(\boldsymbol{s}^{\mathrm{T}}\boldsymbol{X}) &= \sqrt{j!}\sum_{|\boldsymbol{k}|=j}\prod_{l=1}^{d}\frac{s_l^{k_l}}{\sqrt{k_l!}}\psi_{k_l}(X_l)\\ &= \sqrt{j!}\sum_{|\boldsymbol{k}|=j}\frac{s_1^{k_1}\cdot s_2^{k_2}\cdot\ldots\cdot s_d^{k_d}}{\sqrt{k_1!\cdot k_2!\cdot\ldots\cdot k_d!}}\Psi_{\boldsymbol{k}}(\boldsymbol{X}).\end{aligned} \tag{22}$$

Therefore, in the asymptotic limit we can use Eq. (22) to write

$$\hat{Y}_m^{\text{PLS}} = a_0 + \sum_{i=1}^{m} \sum_{|\boldsymbol{k}|\leq q_i} a_{i|\boldsymbol{k}|}^{q_i} \sqrt{|\boldsymbol{k}|!} \frac{r_{i1}^{k_1} \cdot r_{i2}^{k_2} \dots r_{id}^{k_d}}{\sqrt{k_1! \cdot k_2! \dots k_d!}} \Psi_{\boldsymbol{k}}(\boldsymbol{X}). \tag{23}$$

In practice, the sample mean decays towards 0 relatively fast, such that the approximation error introduced by neglecting the variable centering in equation (23) is typically orders of magnitude smaller than the leading error introduced by the surrogate model itself. The error due to $\|\boldsymbol{r}_i\| \neq 1$ grows with the number of included components m (with $m = 1$, the representation is exact since $\|\boldsymbol{r}_1\| = 1$ always). Alternatively, it is possible to derive exact expressions with respect to both non-zero sample mean and non-unit-length component directions and we will do so in Section 4.2.

Equation (23) is merely a linear combination of m standard PCEs, each representing a latent component in standard PCE format, so that we can write

$$\hat{Y}_m^{\text{PLS}} = a_0 + \sum_{|\boldsymbol{k}|\leq q_{max}} c_{\boldsymbol{k}} \Psi_{\boldsymbol{k}}(\boldsymbol{X}), \tag{24}$$

where

$$q_{max} = \max_{i\in\{1,\dots,m\}} (q_i). \tag{25}$$

The equivalent PCE coefficients $\boldsymbol{c}$ read

$$c_{\boldsymbol{k}} = \sum_{i=1}^{m} a_{i|\boldsymbol{k}|}^{q_i} \sqrt{|\boldsymbol{k}|!} \prod_{l=1}^{d} \frac{r_{il}^{k_l}}{\sqrt{k_l!}}, \tag{26}$$

where $\left\{a_{i|\boldsymbol{k}|}^{q_i} : q_i < |\boldsymbol{k}|\right\} = 0$. Thus, one can apply the standard post-processing defined by equations (19) to format (24) in order to obtain variance-based sensitivity indices based on [29]. The corresponding subroutine is referred to as PCE_sensitivites.

Algorithm 2 efficiently determines the Sobol' and total-effect indices of a PLS-PCE model. The algorithm requires to compute the set of multi-indices $\boldsymbol{k} = (k_1, ..., k_d) \in \mathbb{N}^d$ that satisfy $|k| \leq q_{max}$. This multi-index set is only computed once at the beginning using a routine termed multi_index. Various methodologies such as the ball-box-algorithm [50] have been proposed for computing the set of multi-indices of a PCE. We observe that the index set $\boldsymbol{k}$ required for the PLS-based sensitivity indices is equivalent to that of a full PCE formulation of maximum polynomial order q_{max}. Fortunately, the additional degrees of freedom emerging from the latent variable formulation (i.e. the $\boldsymbol{r}_i$) lead to significantly smaller required polynomial degrees in PLS-PCE compared to sparse and classical PCE models. That is, the computational bottleneck of computing $\boldsymbol{k}$ is not critical in most applications.

Algorithm 2 PLS-based sensitivities - asymptotic case

Input PLS components $q_i, \boldsymbol{a}^{q_i}, \boldsymbol{r}_i \ \forall \ i = 1, ..., m$
Initialize $\boldsymbol{c} = \boldsymbol{0}$
$q_{max} = \max_{i \in \{1,...,m\}} (q_i)$
Compute PCE multi-index set, e.g.: $\boldsymbol{\alpha} \leftarrow \text{multi_index}(d, q_{max}) \in \mathbb{N}^{P \times d}$
for $i \leftarrow 1, m$ **do**
    $j = 1$
    while $|\boldsymbol{\alpha}_j| \leq q_i$ **do**
        Set current multi-index $\boldsymbol{k} \leftarrow \boldsymbol{\alpha}_j \in \mathbb{N}^{1 \times d}$
        Compute coefficients $c_{\boldsymbol{k}} \leftarrow a^{q_i}_{i|\boldsymbol{k}|} \sqrt{|\boldsymbol{k}|!} \prod_{l=1}^{d} \frac{r_{il}^{k_l}}{\sqrt{k_l!}}$
        Augment: $c_j \leftarrow c_j + c_{\boldsymbol{k}}$
        $j \leftarrow j + 1$
    end while
end for
Compute sensitivities: $S, S^{clo}, S^T \leftarrow \text{PCE_sensitivities}(\boldsymbol{\alpha}, \boldsymbol{c})$
Return S, S^{clo}, S^T

4.2. Corrections for small samples sizes

The presented methodology to extract sensitivities from a PLS-PCE model is asymptotically exact. However, PLS-PCE is a surrogate modeling technique, which is particularly suitable when the number of samples is small compared to the problem dimension. In this case, the exact sensitivities accounting for non-zero sample mean $\boldsymbol{\mu}_{\mathcal{X}}$ and non-unit-length PLS-directions $\boldsymbol{r}_i$ can still be derived. Rewriting equation (21), we have

$$\hat{Y}_m^{\text{PLS}} = a_0 + \sum_{i=1}^{m} (\boldsymbol{a}_i^{q_i})^{\text{T}} \boldsymbol{\psi}_{q_i} \left(\|\boldsymbol{r}_i\| (\tilde{\boldsymbol{r}}_i)^{\text{T}} (\boldsymbol{X} - \boldsymbol{\mu}_{\mathcal{X}}) \right), \tag{27}$$

where $\tilde{\boldsymbol{r}}_i = \boldsymbol{r}_i / \|\boldsymbol{r}_i\|$. We can view the argument of each polynomial ψ_j in Eq. (27) as an affine transformation of the PLS-component $z = \tilde{\boldsymbol{r}}_i^{\text{T}} \boldsymbol{X}$. Then, expressing each $\psi_j(\beta z + \gamma)$ in Eq. (27) in terms of $\psi_j(z)$, where $\beta = \|\boldsymbol{r}_i\|$ and $\gamma = -(\boldsymbol{r}_i)^{\text{T}} \boldsymbol{\mu}_{\mathcal{X}}$, provides a representation of the PLS-PCE model that can be exactly transformed in the form of Eq. (23) even for small sample sizes (but in the asymptotic limit as well). We start with the following well-known product theorem [51, 52]:

$$\psi_j(\beta z) = \sqrt{j!} \sum_{l=0}^{\lfloor j/2 \rfloor} \frac{\beta^{j-2l} (\beta^2 - 1)^l}{\sqrt{(j-2l)!} 2^l l!} \psi_{j-2l}(z). \tag{28}$$

Moreover, expanding $\psi_j(z + \gamma)$ in a Taylor series around z yields

$$\psi_j(z + \gamma) = \sum_{t=0}^{j} \binom{j}{t} \sqrt{\frac{t!}{j!}} \gamma^{j-t} \psi_t(z). \tag{29}$$

Consequently,

$$\begin{aligned} \boldsymbol{\psi}_j \left(\|\boldsymbol{r}_i\| (\tilde{\boldsymbol{r}}_i)^{\text{T}} (\boldsymbol{X} - \boldsymbol{\mu}_{\mathcal{X}}) \right) &\overset{(28)}{=} \sqrt{j!} \sum_{l=0}^{\lfloor j/2 \rfloor} \zeta_1(j,l) \psi_{j-2l}((\tilde{\boldsymbol{r}}_i)^{\text{T}} (\boldsymbol{X} - \boldsymbol{\mu}_{\mathcal{X}})) \\ &\overset{(29)}{=} \sqrt{j!} \sum_{l=0}^{\lfloor j/2 \rfloor} \zeta_1(j,l) \sum_{t=0}^{j-2l} \zeta_2(j,l,t) \psi_t((\tilde{\boldsymbol{r}}_i)^{\text{T}} \boldsymbol{X}) \\ &\overset{(22)}{=} j! \sum_{l=0}^{\lfloor j/2 \rfloor} \zeta_1(j,l) \sum_{t=0}^{j-2l} \zeta_2(j,l,t) \sum_{|\boldsymbol{k}|=t} \zeta_3(\boldsymbol{k}) \Psi_{\boldsymbol{k}}(\boldsymbol{X}), \end{aligned} \tag{30}$$

where

$$\zeta_1(j,l) = \frac{\|\boldsymbol{r}_i\|^{j-2l}(\|\boldsymbol{r}_i\|^2-1)^l}{\sqrt{(j-2l)!2^l l!}} \tag{31}$$

$$\zeta_2(j,l,t) = \binom{j-2l}{t}\sqrt{\frac{t!}{(j-2l)!}}(-(\boldsymbol{r}_i)^{\mathrm{T}}\boldsymbol{\mu}_{\mathcal{X}})^{j-2l-t} \tag{32}$$

$$\zeta_3(\boldsymbol{k}) = \frac{\tilde{r}_{i1}^{k_1}\cdot\tilde{r}_2^{k_2}\dots\tilde{r}_d^{k_d}}{\sqrt{k_1!\cdot k_2!\dots k_d!}}. \tag{33}$$

In the same way as for the asymptotic formulation, this yields cumulative coefficients for a given multivariate basis function associated with $\boldsymbol{k}$. An efficient way to compute the equivalent PCE-coefficients, which ensures the index set $\boldsymbol{\alpha}$ only has to be computed once, is presented by Algorithm 3.

Algorithm 3 PLS-based sensitivities - non-asymptotic case

Input PLS components $\boldsymbol{a}^i, q_i, \boldsymbol{r}_i \ \forall\ i = 1, ..., m$
Initialize $\boldsymbol{c} = \boldsymbol{0}$
$q_{max} = \max\limits_{i\in\{1,\dots,m\}}(q_i)$
Compute PCE multi-index set, e.g.: $\boldsymbol{\alpha} \leftarrow$ multi_index$(d, q_{max}) \in \mathbb{N}^{P\times d}$
for $i \leftarrow 1, m$ **do**
  for $j \leftarrow 0, q_i$ **do**
    for $l \leftarrow 0, \lfloor j/2 \rfloor$ **do**
      Compute $\zeta_1(j,l)$ from equation (31)
      for $t \leftarrow 0, j-2l$ **do**
        Compute $\zeta_2(j,l,t)$ from equation (32)
        Get the multi-index subset of length t: $\boldsymbol{A} = \{\boldsymbol{\alpha}_p : |\boldsymbol{\alpha}_p| = t\}$
        for $p \leftarrow 1, len(\boldsymbol{A})$ **do**
          Get current multi-index $\boldsymbol{k} \leftarrow \boldsymbol{\alpha}_p \in \mathbb{N}^{1\times d}$
          Compute $\zeta_3(\boldsymbol{k})$ from equation (33)
          Augment: $c_p \leftarrow c_p + j! a_{it}^{q_i}\zeta_1\zeta_2\zeta_3$
          $p \leftarrow p+1$
        end for
      end for
    end for
  end for
end for
Compute sensitivities: $S, S^{clo}, S^T \leftarrow$ PCE_sensitivities$(\boldsymbol{\alpha}, \boldsymbol{c})$
Return S, S^{clo}, S^T

4.3. A comment on variance-based sensitivity analysis with the standard basis adaptation format

In the previous two subsections, we have derived expressions for the Sobol' and total-effect index of a basis-adapted PCE of the format (12), which differs from the standard basis adaptation format (11). While the former is a summation of m univariate PCEs of m latent variables, the latter represents an m-variate PCE of the latent variables. In the following we state an expression for the standard basis-adapted PCE format in terms of Hermite polynomials of the original inputs only. We write $\tilde{\boldsymbol{k}} \in \mathbb{N}^m$ and $\boldsymbol{k} \in \mathbb{N}^d$ to express

multi-indices in the latent and original input spaces, respectively. Then,

$$\begin{aligned}
\hat{Y}_p^{\boldsymbol{Q}_m} &\overset{(11)}{=} \sum_{|\tilde{\boldsymbol{k}}|\leq p} a_{\tilde{\boldsymbol{k}}} \Psi_{\tilde{\boldsymbol{k}}} \left(\boldsymbol{Q}_m^{\mathrm{T}} \boldsymbol{X}\right) \\
&\overset{(22)}{=} \sum_{|\tilde{\boldsymbol{k}}|\leq p} a_{\tilde{\boldsymbol{k}}} \prod_{j=1}^{m} \sum_{|\boldsymbol{k}|=\tilde{k}_j} \sqrt{|\boldsymbol{k}|!} \prod_{\ell=1}^{d} \frac{Q_{m,\ell j}^{k_\ell}}{\sqrt{k_\ell!}} \psi_\ell \left(X_\ell\right) \\
&= \sum_{|\tilde{\boldsymbol{k}}|\leq p} a_{\tilde{\boldsymbol{k}}} \prod_{j=1}^{m} \sum_{|\boldsymbol{k}|=\tilde{k}_j} \zeta_{\boldsymbol{k}} \Psi_{\boldsymbol{k}} \left(\boldsymbol{X}\right), \quad \zeta_{\boldsymbol{k}} = \sqrt{|\boldsymbol{k}|!} \prod_{\ell=1}^{d} \frac{Q_{m,\ell j}^{k_\ell}}{\sqrt{k_\ell!}}.
\end{aligned} \tag{34}$$

Back-transforming format (34) to a standard PCE is non-trivial due to the inner summation, which induces multiple occurences of the $\boldsymbol{k}$-th multivariate basis polynomial $\Psi_{\boldsymbol{k}}$ in the full expansion. Instead of back-transforming Eq. (34) to a standard PCE, one may plug it directly in the definitions of the variance-based sensitivity indices in Eqs. (15) & (17) and collect the partial variance contributions associated with any $\Psi_{\boldsymbol{k}}$. Such an approach is chosen in [30] for deriving variance-based sensitivities of canonical LRAs. We leave this task for a future work. We remark that for the case where the basis-adapted format of Eq. (11) does not consider mixed effects, i.e. it is $|\tilde{\boldsymbol{k}}| = \max(\tilde{k}_i)$, the coefficients of the equivalent standard PCE representation will be identical to the ones defined in Eq. (26) with r_{ij} set to $Q_{m,ij}$. In such case, Algorithm 2 is directly applicable for computing the sensitivity indices.

5. Numerical experiments

In this section, we evaluate the proposed methods with one low-dimensional and one high-dimensional numerical experiment. We examine the performance of both the asymptotic approximation proposed in Algorithm 2 and the exact computation performed with Algorithm 3. The results are compared to those obtained with other surrogate modelling techiques, namely sparse PCE based on least-angle regression [37] and LRAs in the canonical polyadics format [30]. For all polynomial bases (sparse PCE, LRA, PLS-PCE) we prescribe a maximum degree of $p = 10$, for the LRA a maximum rank of $R = 10$ and for the PLS-PCE the same maximum number of components, i.e. $m = 10$. When computing LRAs, we use an adaptive scheme for the rank selection while considering every polynomial order up to p within the selected ranks [32]. All experimental designs are generated via latin hypercube sampling. In both examples, we draw 100 random experimental designs to quantify the relative error mean and standard deviation, where the error for a quantity Q with respect to to its reference solution Q_{ref} is defined as their difference:

$$\epsilon_Q = Q - Q_{ref}.$$

In our studies, Q refers to the Sobol' and the Total-effect indices and the respective reference solutions Q_{ref} are obtained with double-loop Monte Carlo and the estimators stated in [23, 26] depending on the example. The latter estimators are based on drawing a single set of n independent input samples of dimensions d and splitting them in two equally sized subsets. Permuting columns between the subsets yields a dependent sample of size $n(d+2)/2$ based on which all Sobol' and total-effect indices can be estimated. We refer to these as permutation matrix estimators.

5.1. Elastic truss

We consider an elastic truss that consists of 23 rods as depicted in Figure 1 [53]. Horizontal and diagonal rods have cross-sections A_1, A_2 and Young's moduli E_1, E_2, respectively. The truss sustains 6 vertical point loads P_1 - P_6. The input variable definitions are provided in Table (1). We compute Sobol' and total-effect indices for the maximum truss deflection u_{max} using Algorithm 2. Reference solutions (direct Monte Carlo - DMC) are obtained based on $n = 10^6$ independent samples (Figure 2) with the permutation matrix estimators.

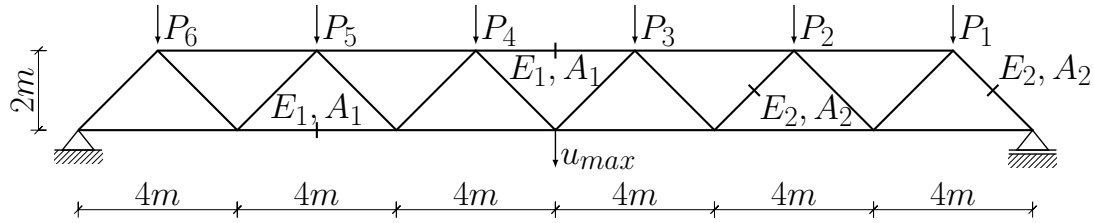

Figure 1: 2-D truss example.

Table 1: Input variable definitions of the truss example.

Random Variable	Distribution	Mean	Standard deviation
A_1 $[m^2]$	Log-Normal	$2 \cdot 10^{-3}$	$2 \cdot 10^{-4}$
A_2 $[m^2]$	Log-Normal	$1 \cdot 10^{-3}$	$1 \cdot 10^{-4}$
E_1, E_2 $[Pa]$	Log-Normal	$2.1 \cdot 10^{11}$	$2.1 \cdot 10^{10}$
P_1 - P_6 $[N]$	Gumbel	$5.0 \cdot 10^4$	$7.5 \cdot 10^3$

Figure 2 indicates good agreement of the PLS-PCE-based sensitivities with the reference solution. Figure 3 shows that all three surrogate-based sensitivity indices are estimated with similar mean relative error and convergence rate as N increases. Figure 4 indicates the same for the relative error variance. Figure 5 shows a performance comparison of Algorithms 2 and 3 for the truss. Here, the asymptotic approximation introduces only negligible error into the sensitivity estimate. For variables of little significance, which yield small absolute values for the corresponding sensitivity indices (e.g. E_2, A_2), the surrogate modelling error is much larger than the error due to the asymptotic approximation of the sensitivity indices so that there is no visible difference left in the error plots.

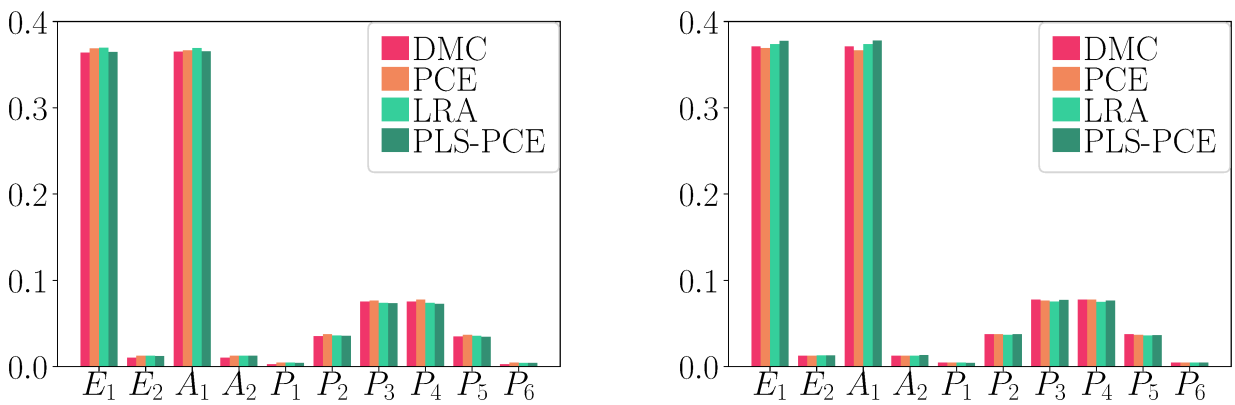

Figure 2: Sobol' (left) and total-effect (right) indices of u_{max} obtained with $N = 100$.

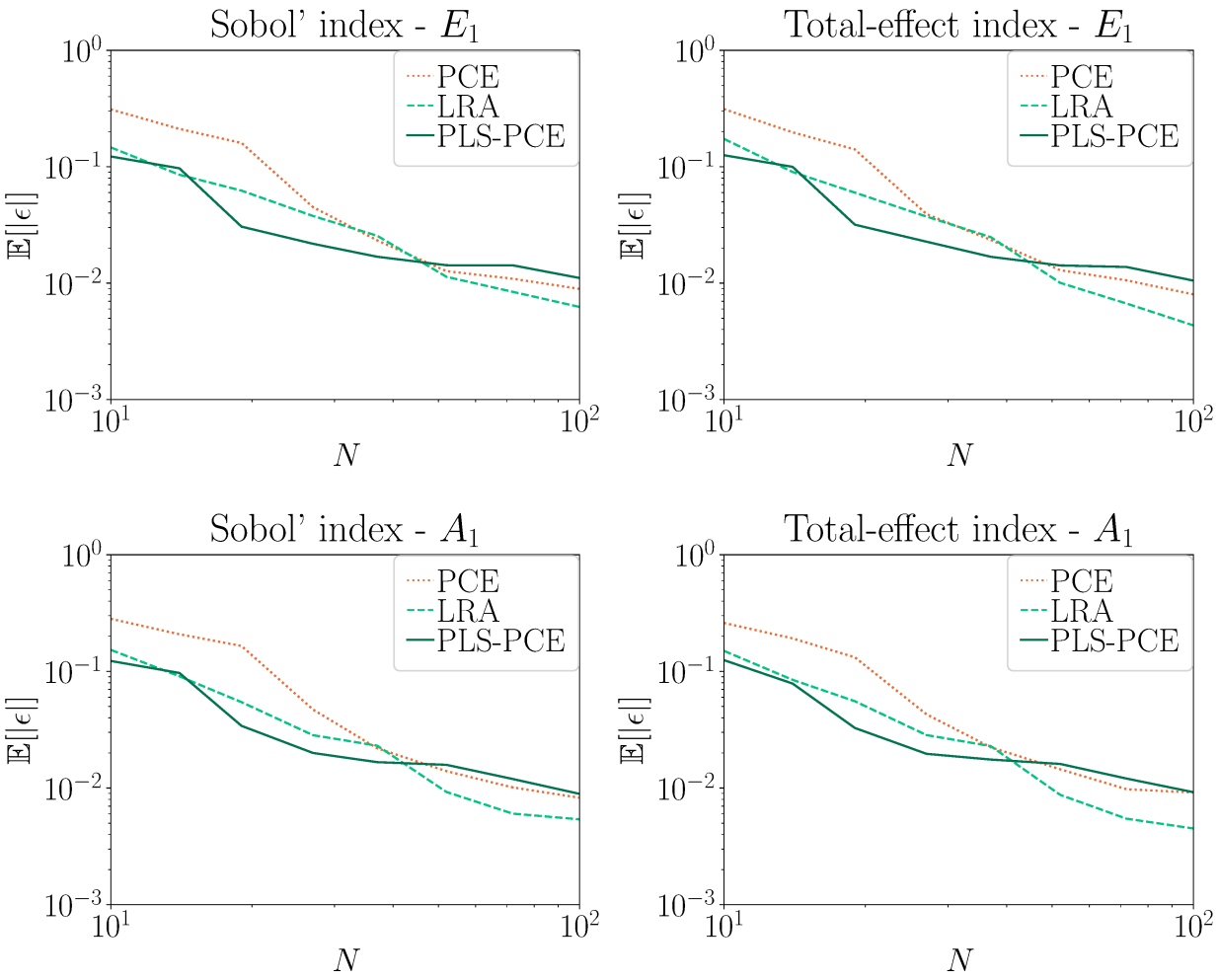

Figure 3: Mean relative errors for the two most influential inputs E_1 and A_1, computed with PCE, LRA and PLS-PCE.

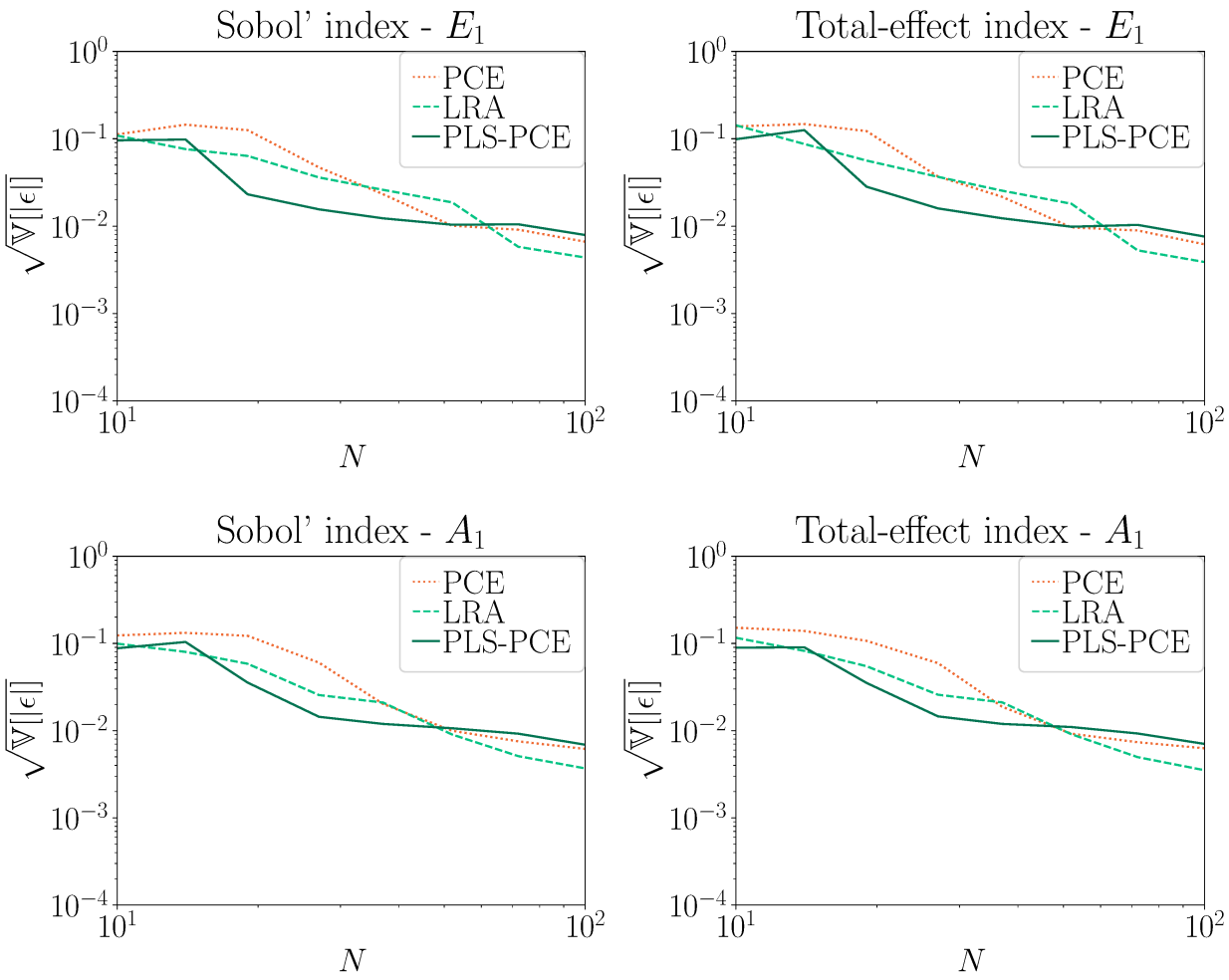

Figure 4: Relative error standard deviation for the two most influential inputs E_1 and A_1, computed with sparse PCE, LRA and PLS-PCE.

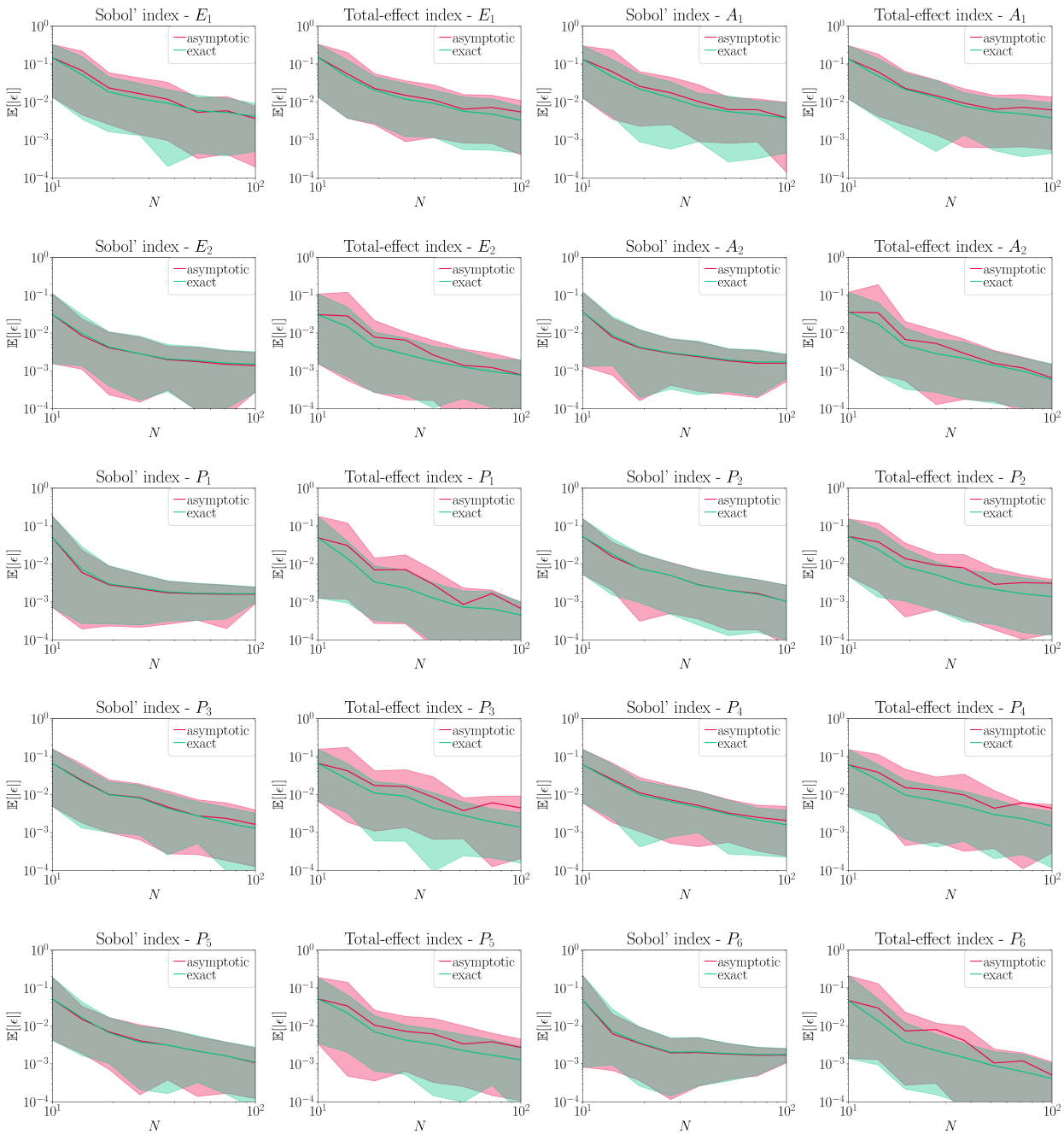

Figure 5: Mean relative errors with 90% CI, computed with asymptotic approximation (Algorithm 2) and exactly (Algorithm 3).

5.2. Steel plate

For the second example, we consider a modified version of the example given in [54], which consists of a low-carbon steel plate of length 0.32 m, width 0.32 m, thickness $t = 0.01$ m, and a hole of radius 0.02 m located at the center. The Poisson ratio is set to $\nu = 0.29$ and the density of the plate is $\rho = 7850$ kg/m^3. The horizontal and vertical displacements are constrained at the left edge. Both the surface load q, which acts on the right plate side, and the plate's Young's modulus $E(x, y)$ are considered uncertain and spatially variable. Both are described by homogeneous random fields in two and one spatial dimension, respectively. E has log-normal marginal distribution, mean value $\mu_E = 2 \times 10^5$ MPa and standard deviation $\sigma_E = 3 \times 10^4$ MPa. The autocorrelation function of the underlying Gaussian field $\ln E$ is modeled by the isotropic exponential model

$$\rho_{\ln E}(\Delta x, \Delta y) = \exp(-\sqrt{\Delta x^2 + \Delta y^2}/l_E) \tag{35}$$

with correlation length $l_E = 0.08$m. The random field $\ln E$ is discretized by a Karhunen-Loève-expansion (KLE), i.e.

$$E(x, y) = \exp\left\{\mu_{\ln E} + \sigma_{\ln E} \sum_{i=1}^{d_E} \sqrt{\lambda_i^E} \varphi_i^E(x, y) \xi_i^E\right\}, \tag{36}$$

where $\mu_{\ln E}$ and $\sigma_{\ln E}$ are the parameters of the log-normal marginal distribution of E, $\{\lambda_i^q, \varphi_i^E\}$ are the eigenpairs of the correlation kernel (35) and $\boldsymbol{\xi}^E \in \mathbb{R}^{d_E \times 1}$ is a standard-normal random vector. The number of terms in the expansion d_E is chosen such that 90% of the random field variance is represented by the discretization (36), which yields $d_E = 169$. Selected eigenfunctions of $\rho_{\ln E}$ are shown in Figure 7. q also has log-normal marginal distribution with mean value $\mu_q = 60$ MPa and standard deviation $\sigma_q = 12$ MPa. The autocorrelation function of the underlying Gaussian field $\ln q$ is also modeled by an isotropic exponential model,

$$\rho_{\ln q}(\Delta y) = \exp(-|\Delta y|/l_q) \tag{37}$$

with correlation length $l_q = 0.02$m. The random field $\ln q$ is also discretized by a KLE, s.t.

$$q(y) = \exp\left\{\mu_{\ln q} + \sigma_{\ln q} \sum_{i=1}^{d_q} \sqrt{\lambda_i^q} \varphi_i^q(y) \xi_i^q\right\}, \tag{38}$$

where $\mu_{\ln q}$ and $\sigma_{\ln q}$ are the parameters of the log-normal marginal distribution of q, $\{\lambda_i^q, \varphi_i^q\}$ are the eigenpairs of the load correlation kernel (37) and $\boldsymbol{\xi}^q \in \mathbb{R}^{d_q \times 1}$ is a standard-normal random vector constituting the model input of the plate together with $\boldsymbol{\xi}^E$, i.e. $\boldsymbol{X} = [\boldsymbol{\xi}^q, \boldsymbol{\xi}^E]^T$. With the same accuracy criterion on the represented random field variance as for E ($> 90\%$) , one obtains $d_q = 32$. The first 4 eigenfunctions of $\rho_{\ln q}$ are shown in Figure 8.
The stress field

$$\boldsymbol{\sigma}(x, y) = [\sigma_x(x, y), \sigma_y(x, y), \tau_{xy}(x, y)]^T,$$

strain field

$$\boldsymbol{\epsilon}(x, y) = [\epsilon_x(x, y), \epsilon_y(x, y), \gamma_{xy}(x, y)]^T$$

and displacement field

$$\boldsymbol{u}(x, y) = [u_x(x, y), u_y(x, y)]^T$$

of the plate are given through elasticity theory, namely the Cauchy-Navier equations [55]. Given the configuration of the plate, the model can be simplified under the plane stress hypothesis, which yields

$$G(x, y)\nabla^2 \boldsymbol{u}(x, y) + \frac{E(x, y)}{2(1 - \nu)} \nabla(\nabla \cdot \boldsymbol{u}(x, y)) + \boldsymbol{b} = 0. \tag{39}$$

Therein, $G(x, y) = E(x, y)/(2(1 + \nu))$ is the shear modulus, and $\boldsymbol{b} = [b_x, b_y]^T$ is the vector of body forces acting on the plate. Eq. (39) is discretized with a finite-element method. That is, the spatial domain of

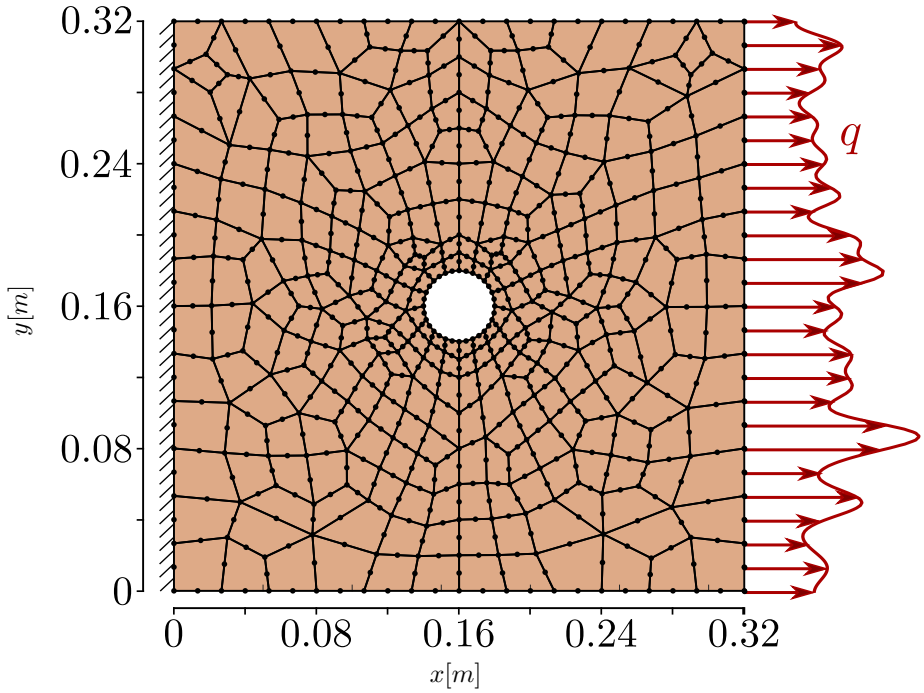

Figure 6: FE-mesh of 2D-plate model with uncertain Young's modulus E subject to random surface load q.

the plate is discretized into 282 eight-noded quadrilateral elements, as shown in Figure 6. The scalar model output is the maximally occurring first principal plane stress

$$\sigma_1 = 0.5(\sigma_x + \sigma_y) + \sqrt{[0.5(\sigma_x + \sigma_y)]^2 + \tau_{xy}^2}.$$

The FE-model of the plate with random inputs is illustrated in Figure 6. We compute sensitivity indices for both the random variables characterizing the uncertainty associated with the single modes of the KL-decompositions as well as for the two random fields as such. The random field sensitivity analysis can be understood as interpreting each random field as a single input. Thus the variance decomposition of the model output with respect to its inputs is computed once with respect to each single random variable input and once with respect to two subgroups of these random variables each characterizing one of the random fields in the problem description. Random field-oriented Sobol' indices are always of closed form (see Equation (16)) as the classical Sobol' index would conceal most of the contributed variance. The PLS-PCE-based sensitivity indices are computed with Algorithm 2 and compared against LRA-based sensitivities and Monte Carlo reference solutions. The latter (DMC) are obtained with $4 \cdot 10^6$ samples using the double-loop ($2 \cdot 10^3$ samples per stage) to compute the random field sensitivities and with $n = 2 \times 10^4$ independent samples using the permutation matrix estimators to compute the random variable sensitivities. For this application, sparse PCE surrogates are difficult to obtain beyond relatively low total polynomial degrees ($p \leq 3$) due to the large input dimension d. PCE-based sensitivity indices are thus not considered in this study.

The random field-oriented sensitivity index means are plotted in Figure 9 and attribute a larger influence to the random load field $q(y)$ compared to the material parameter field $E(x, y)$. The PLS-PCE-based indices are consistently closer to the reference than the LRA-based indices while both surrogate-based indices approximate the reference well. According to Figure 10, both surrogate indices converge in standard deviation with increasing N while the LRA-based mean relative error does not decrease further beyond a certain sample size. The PLS-PCEs are superior in this respect as the corresponding mean relative error decreases further as $N \to 10^3$. The means of the sensitivity indices of the random variables correspond-

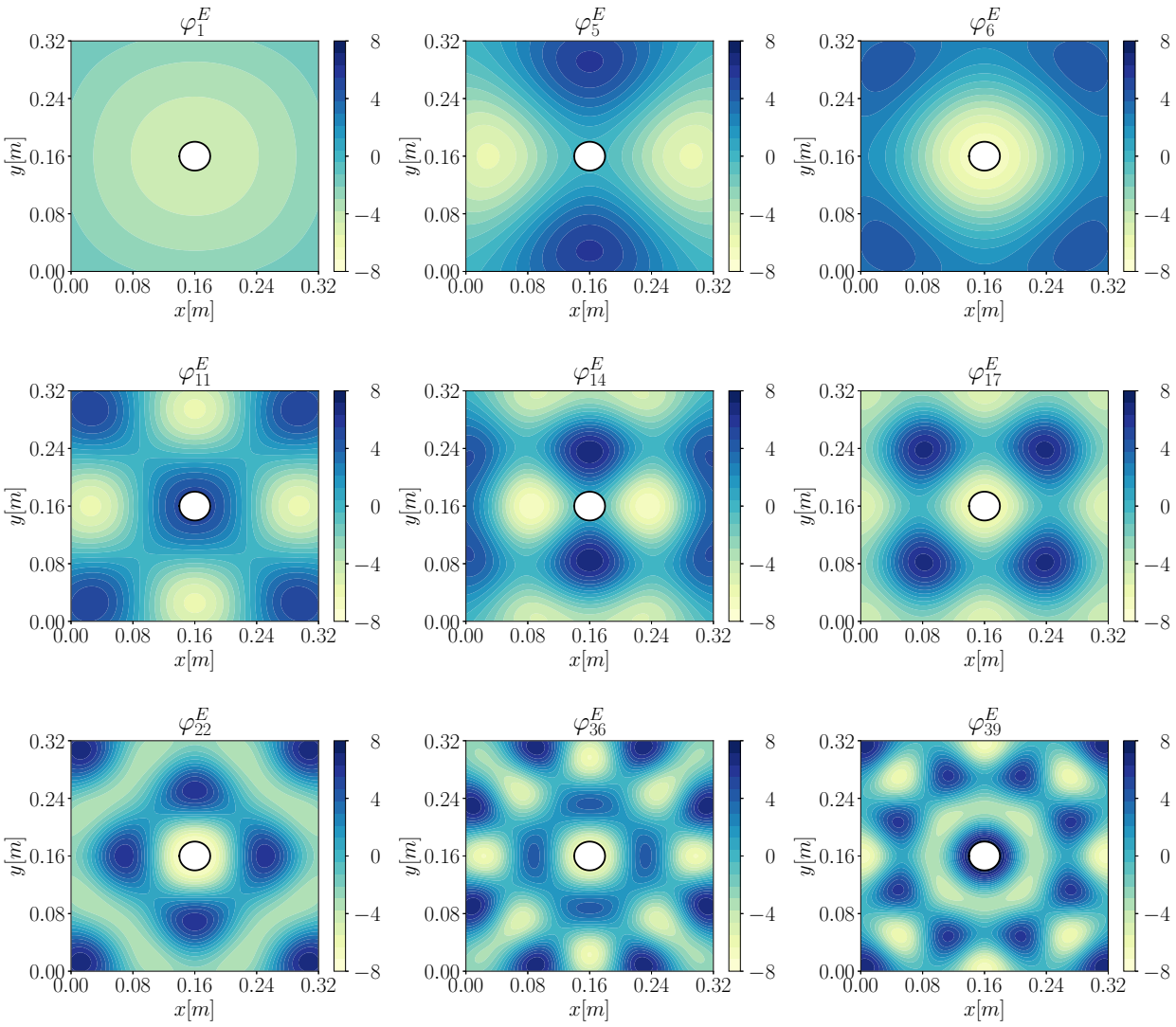

Figure 7: The nine most important eigenfunctions of the exponential, isotropic Young's modulus correlation kernel $\rho_{\ln E}$.

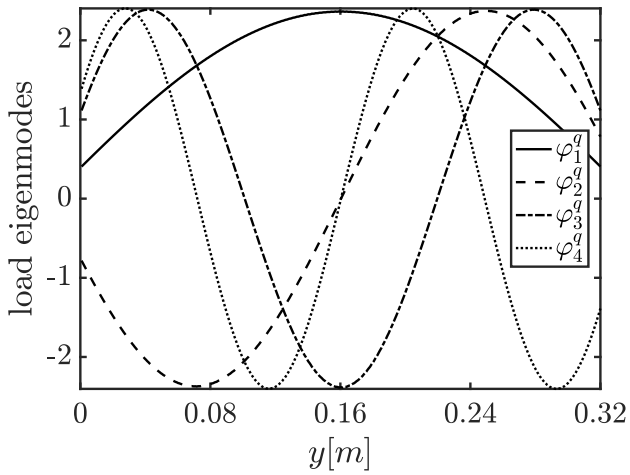

Figure 8: First four eigenfunctions of exponential load correlation kernel $\rho_{\ln q}$, $\varphi_1^q(y)$ - $\varphi_4^q(y)$.

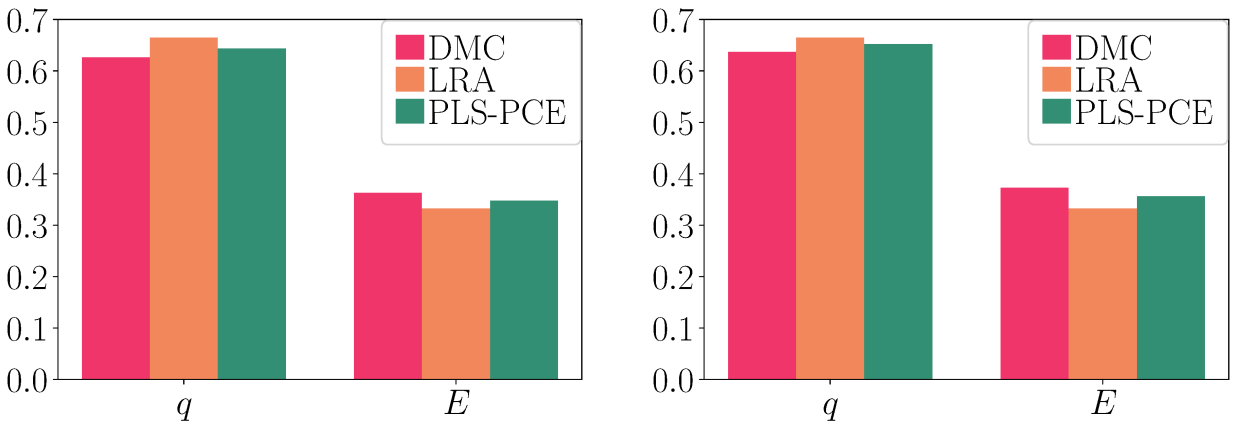

Figure 9: Closed Sobol' (left) and total-effect (right) indices for the two random fields q and E obtained with $N = 10^3$.

ing to the KL eigenmodes (the 10 most important ones) are plotted in Figure 11. The random variable corresponding to the first mode of the load random field is by far the most important input. The next 9 random variable inputs in the ranking all correspond to modes of the material parameter field $E(x, y)$ and their associated eigenfunctions are plotted in Figure 7. For the largest mode-oriented sensitivity indices, the PLS-PCE yields slightly better approximations of the reference than the LRA while both approximations are fairly accurate. Figures 12 & 13 show the convergence study of the mode-oriented Sobol' and total-effect indices as N increases. The dominant random variable ξ_1^q reproduces the convergence behaviour of the random field-oriented sensitivity study. Namely, the mean relative error of the LRA-based sensitivity reaches a plateau and increases again after reaching a certain experimental design size ($N \approx 300$). The remainder of the 10 most important inputs (all E-modes) show consistent convergence in the relative error mean and standard deviation as N increases.

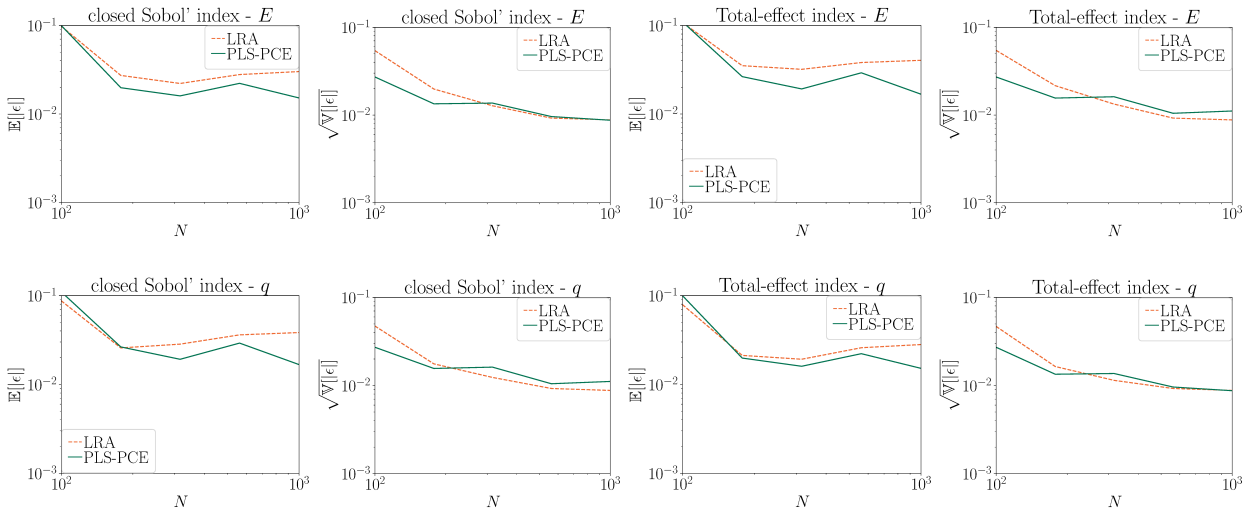

Figure 10: Relative error mean and standard deviation for sensitivity estimates of the maximum first main stress to q and E.

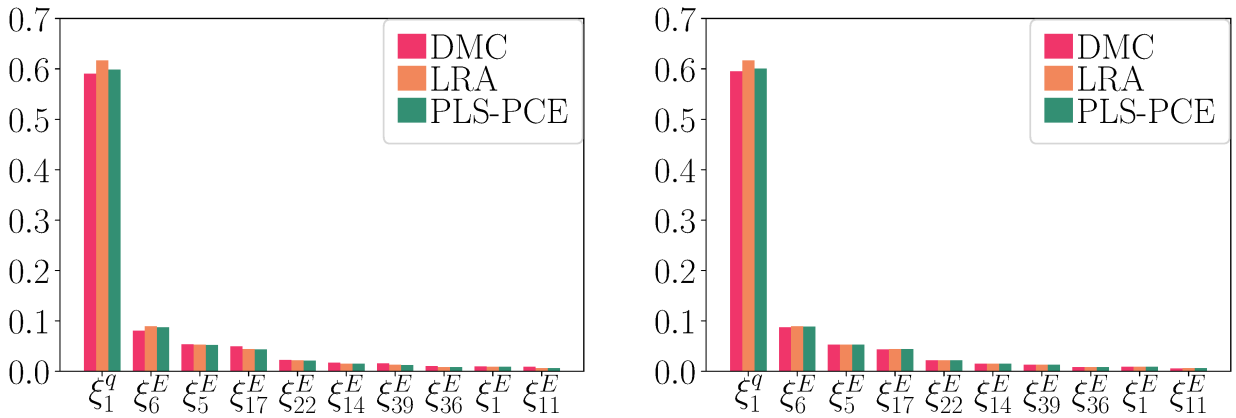

Figure 11: Sobol' (left) and total-effect (right) indices for the ten most important random field modes obtained with $N = 10^3$.

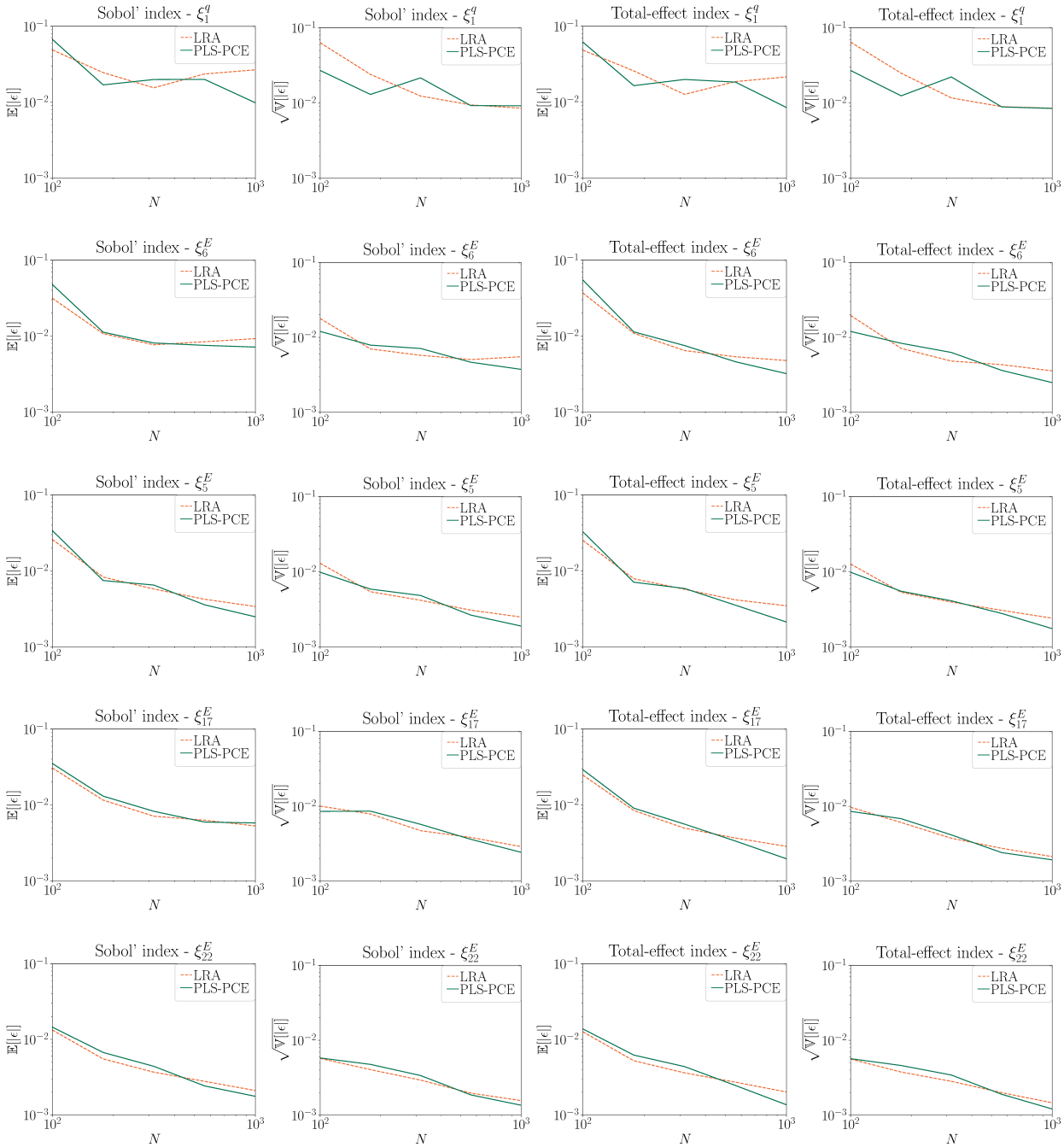

Figure 12: Relative error mean and standard deviation for the Sobol' and total-effect indices of the ten (1.-5.) most influential model inputs computed with LRA and PLS-PCE over varying experimental design size N.

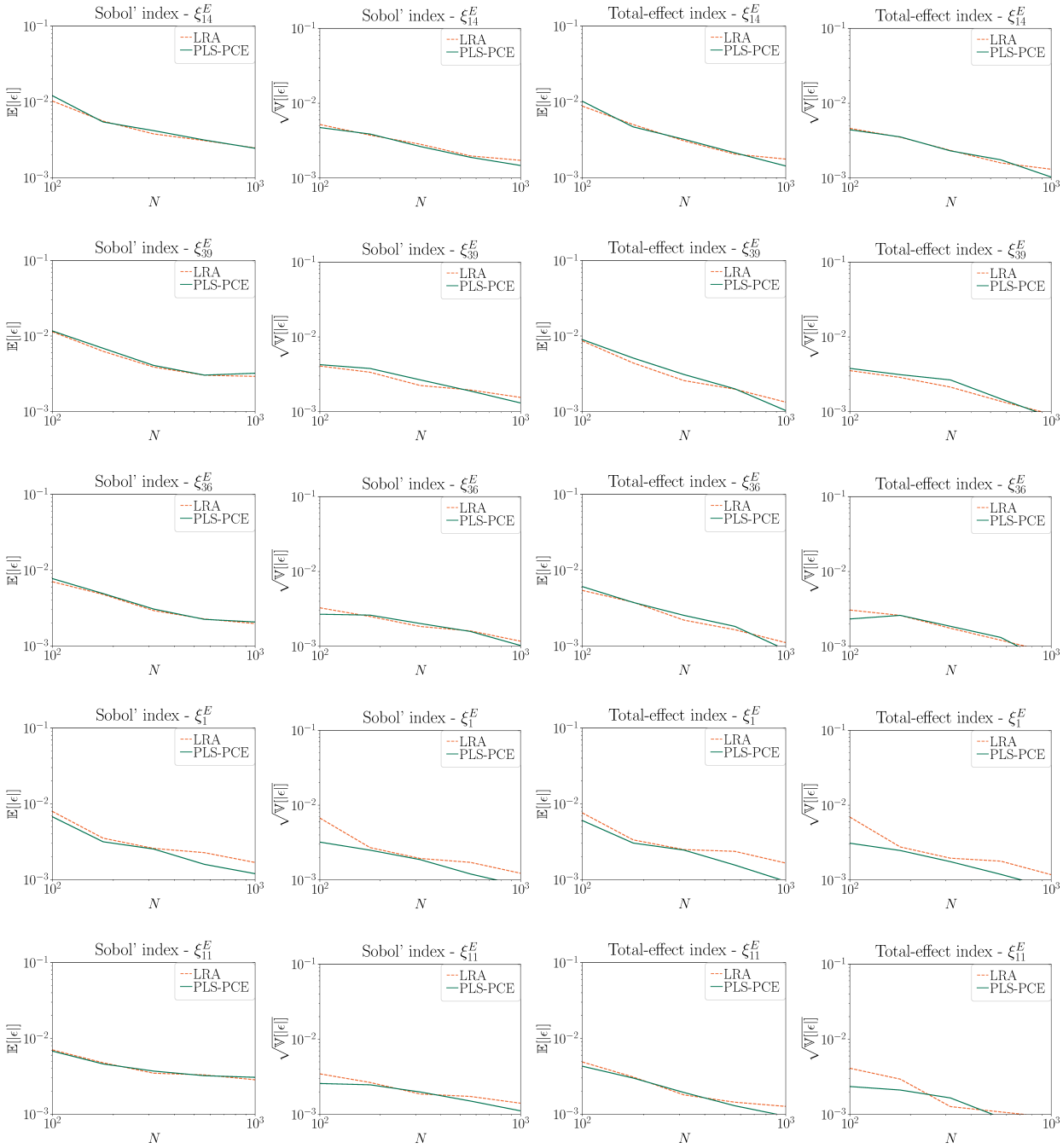

Figure 13: Relative error mean and standard deviation for the Sobol' and total-effect indices of the ten (6.-10.) most influential model inputs computed with LRA and PLS-PCE over varying experimental design size N.

6. Conclusion

This paper derives analytical expressions for variance-based sensitivity indices of PLS-PCE surrogate model outputs and formulates two algorithms for their efficient computation. The expressions for the sensitivities involve only the surrogate model coefficients. Thus, these sensitivities can be computed with negligible additional computational effort once the surrogate model is identified by a mere post-processing of the model coefficients. For the first algorithm, a multinomial theorem for Hermite polynomials is applied to derive expressions for the sensitivity measures based on the model coefficients, which is asymptotically exact. That is, with the number of samples $N \to \infty$, these estimators exactly match the theoretical Sobol' and total-effect indices of the surrogate model. For the second algorithm, corrected expressions are derived, which are exact also when N is small. These can be used when the experimental design is small. The sensitivity estimates obtained with both algorithms are compared to Monte-Carlo-based reference solutions as well as estimates based on sparse PCEs and LRAs for two different numerical examples: A low-dimensional ($d = 10$) truss example as well as a high-dimensional plate example ($d = 201$). For both examples, the PLS-PCE-based sensitivity estimates approximate the reference solution well and perform at least as good as the two alternative surrogate-based estimators. Finally, we comment on the applicability of our approach to general basis adaptation formats: We recast the standard basis adapted format in terms of the original input but find that a back-transformation to a standard PCE format based on the stated expression is non-trivial. Instead, sampling-free variance-based sensitivity indices may be computed for standard basis-adapted PCEs by analysing the stated expression term by term as done for canonical LRAs in [30]. We leave this to future research. The presented procedure can be extended to multivariate output in combination with the PLS2 algorithm [56].

7. Acknowledgment

This project was supported by the German Research Foundation (DFG) through Grant STR 1140/6-1 under SPP 1886.

References

[1] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning, Springer Series in Statistics, Springer New York Inc., 2001.
[2] A. I. J. Forrester, A. Sobester, A. J. Keane, Engineering Design via Surrogate Modelling - A Practical Guide., Wiley, 2008.
[3] B. Sudret, Meta-models for structural reliability and uncertainty quantification, Proc. 5th Asian-Pacific Symp. Struct. Reliab. (APSSRA'2012), Singapore (2012).
[4] J. Iott, R. T. Haftka, H. M. Adelman, Selecting step sizes in sensitivity analysis by finite differences, NASA Technical Memorandum 86382 (1985).
[5] F. M. Alam, K. R. McNaught, T. J. Ringrose, Using morris' randomized oat design as a factor screening method for developing simulation metamodels, in: Simulation Conference, 2004. Proceedings of the 2004 Winter, volume 1, IEEE.
[6] P.-A. Ekström, R. Broed, Sensitivity analysis methods and a biosphere test case implemented in eikos, Posiva Working Report 31 (2006) 84.
[7] I. Papaioannou, K. Breitung, D. Straub, Reliability sensitivity analysis with Monte Carlo methods, in: Safety, Reliability, Risk and Life-Cycle Performance of Structures and Infrastructures, CRC Press, 2014, pp. 5335–5342.
[8] I. Papaioannou, K. Breitung, D. Straub, Reliability sensitivity estimation with sequential importance sampling, Structural Safety 75 (2018) 24 – 34.
[9] A. Saltelli, K. Chan, E. Scott, Sensitivity Analysis, John Wiley & Sons, Inc., 2000.
[10] J. C. Helton, J. D. Johnson, C. J. Sallaberry, C. B. Storlie, Survey of sampling-based methods for uncertainty and sensitivity analysis, Reliability Engineering & System Safety 91 (2006) 1175–1209.
[11] B. Iooss, P. Lemaître, A review on global sensitivity analysis methods, in: C. Meloni, G. Dellino (Eds.), Uncertainty management in Simulation-Optimization of Complex Systems: Algorithms and Applications, Springer, 2015.
[12] P. Wei, Z. Lu, J. Song, Variable importance analysis: a comprehensive review, Reliability Engineering & System Safety 142 (2015) 399–432.
[13] E. Borgonovo, E. Plischke, Sensitivity analysis: a review of recent advances, European Journal of Operational Research 248 (2016) 869–887.
[14] B. Efron, C. Stein, The jackknife estimate of variance, The Annals of Statistics (1981) 586–596.
[15] E. Borgonovo, A new uncertainty importance measure, Reliability Engineering & System Safety 92 (2007) 771 – 784.
[16] Q. Liu, T. Homma, A new importance measure for sensitivity analysis, Journal of Nuclear Science and Technology 47 (2010) 53–61.

[17] L. Li, Z. Lu, J. Feng, B. Wang, Moment-independent importance measure of basic variable and its state dependent parameter solution, Structural Safety 38 (2012) 40 – 47.
[18] E. Borgonovo, S. Tarantola, E. Plischke, M. Morris, Transformations and invariance in the sensitivity analysis of computer experiments, Journal of the Royal Statistical Society: Series B (Statistical Methodology) 76 (2014) 925–947.
[19] C. Zhou, Z. Lu, W. Li, Sparse grid integration based solutions for moment-independent importance measures, Probabilistic Engineering Mechanics 39 (2015) 46–55.
[20] G. Greegar, C. Manohar, Global response sensitivity analysis of uncertain structures, Structural Safety 58 (2016) 94–104.
[21] I. Sobol', Sensitivity estimates for nonlinear mathematical models, Math. Modeling & Comp. Exp 1 (1993) 407–414.
[22] T. Homma, A. Saltelli, Importance measures in global sensitivity analysis of nonlinear models, Reliability Engineering & System Safety 52 (1996) 1 – 17.
[23] M. J. Jansen, Analysis of variance designs for model output, Computer Physics Communications 117 (1999) 35 – 43.
[24] I. M. Sobol, Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates, Mathematics and computers in simulation 55 (2001) 271–280.
[25] S. Kucherenko, et al., Global sensitivity indices for nonlinear mathematical models, review, Wilmott Mag 1 (2005) 56–61.
[26] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, S. Tarantola, Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index, Computer Physics Communications 181 (2010) 259–270.
[27] R. Cukier, C. Fortuin, K. E. Shuler, A. Petschek, J. Schaibly, Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. i theory, The Journal of chemical physics 59 (1973) 3873–3878.
[28] G. Archer, A. Saltelli, I. Sobol, Sensitivity measures, anova-like techniques and the use of bootstrap, Journal of Statistical Computation and Simulation 58 (1997) 99–120.
[29] B. Sudret, Global sensitivity analysis using polynomial chaos expansions, Reliability Engineering & System Safety 93 (2008) 964 – 979.
[30] K. Konakli, B. Sudret, Global sensitivity analysis using low-rank tensor approximations, Reliability Engineering & System Safety 156 (2016) 64 – 83.
[31] D. Xiu, G. E. Karniadakis, The Wiener–Askey polynomial chaos for stochastic differential equations, SIAM Journal on Scientific Computing 24 (2002) 619–644.
[32] M. Chevreuil, R. Lebrun, A. Nouy, P. Rai, A least-squares method for sparse low rank approximation of multivariate functions, SIAM/ASA Journal on Uncertainty Quantification 3 (2015) 897–921.
[33] R. Tipireddy, R. Ghanem, Basis adaptation in homogeneous chaos spaces, Journal of Computational Physics 259 (2014) 304 – 317.
[34] I. Papaioannou, M. Ehre, D. Straub, PLS-based adaptation for efficient pce representation in high dimensions, Journal of Computational Physics 387 (2019) 186 – 204.
[35] M. Rosenblatt, Remarks on a multivariate transformation, The annals of mathematical statistics 23 (1952) 470–472.
[36] M. Berveiller, B. Sudret, M. Lemaire, Stochastic finite element: a non intrusive approach by regression, European Journal of Computational Mechanics 15 (2006) 81–92.
[37] G. Blatman, B. Sudret, Adaptive sparse polynomial chaos expansion based on least-angle regression, Journal of Computational Physics 230 (2011) 2345 – 2367.
[38] A. Doostan, H. Owhadi, A non-adapted sparse approximation of PDEs with stochastic inputs, Journal of Computational Physics 230 (2011) 3015–3034.
[39] L. Yan, L. Guo, D. Xiu, Stochastic collocation algorithms using ℓ_1-minimization, International Journal for Uncertainty Quantification 2 (2012) 279–293.
[40] P. Tsilifis, X. Huan, C. Safta, K. Sargsyan, G. Lacaze, J. C. Oefelein, H. N. Najm, R. G. Ghanem, Compressive sensing adaptation for polynomial chaos expansions, Journal of Computational Physics 380 (2019) 29 – 47.
[41] S. Wold, A. Ruhe, H. Wold, W. Dunn, III, The collinearity problem in linear regression. the partial least squares (PLS) approach to generalized inverses, SIAM Journal on Scientific and Statistical Computing 5 (1984) 735–743.
[42] A. Höskuldsson, PLS regression methods, Journal of Chemometrics 2 (1988) 211–228.
[43] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemometrics and Intelligent Laboratory Systems 58 (2001) 109–130.
[44] R. Rosipal, Nonlinear partial least squares: An overview, Chemoinformatics and Advanced Machine Learning Perspectives: Complex Computational Methods and Collaborative Techniques (2010).
[45] S. Wold, N. Kettaneh-Wold, B. Skagerberg, Nonlinear pls modeling, Chemometrics and Intelligent Laboratory Systems 7 (1989) 53 – 65. Proceedings of the First Scandinavian Symposium on Chemometrics.
[46] G. Baffi, E. Martin, A. Morris, Non-linear projection to latent structures revisited (the neural network PLS algorithm), Computers & Chemical Engineering 23 (1999) 1293–1307.
[47] A. Höskuldsson, Pls regression methods, Journal of Chemometrics 2 (1988) 211–228.
[48] C. Prieur, S. Tarantola, Variance-based sensitivity analysis: Theory and estimation algorithms, in: R. Ghanem, D. Higdon, H. Owhadi (Eds.), Handbook of Uncertainty Quantification, Springer International Publishing, 2017, pp. 1217–1239.
[49] F. Buet-Golfouse, A multinomial theorem for hermite polynomials and financial applications, Applied Mathematics 06 (2015) 1017 – 1030.
[50] B. Sudret, A. Der Kiureghian, Stochastic finite element methods and reliability: a state-of-the-art report, Department of Civil and Environmental Engineering, University of California, 2000.
[51] J. P. L. S. Takeyuki Hida, Hui-Hsiung Kuo, White noise : an infinite dimensional calculus, Mathematics and Its Applications 253, Dordrecht: Springer, 1993.
[52] E. Schlögl, Option pricing where the underlying assets follow a gram/charlier density of arbitrary order, Journal of Economic Dynamics and Control 37 (2013) 611 – 632.

[53] S. H. Lee, B. M. Kwak, Response surface augmented moment method for efficient reliability analysis, Structural Safety 28 (2006) 261 – 272.
[54] P.-L. Liu, K.-G. Liu, Selection of random field mesh in finite element reliability analysis, Journal of Engineering Mechanics 119 (1993) 667–680.
[55] C. Johnson, Numerical solution of partial differential equations by the finite element method, Dover Publications, 2009.
[56] L. Li, I. Papaioannou, D. Straub, Partial least squares-based polynomial chaos expansion for global sensitivity analysis of dynamic models in high dimensions, Manuscript, 2019.