International Journal for Uncertainty Quantification, 4 (3): 243–271 (2014)

QUANTIFICATION OF UNCERTAINTY FROM HIGH-DIMENSIONAL SCATTERED DATA VIA POLYNOMIAL APPROXIMATION

Lionel Mathelin*

LIMSI-CNRS, BP 133, 91403 Orsay, France; Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

Original Manuscript Submitted: 06/24/2013; Final Draft Received: 01/27/2014

This paper discusses a methodology for determining a functional representation of a random process from a collection of scattered pointwise samples. The present work specifically focuses onto random quantities lying in a high-dimensional stochastic space in the context of limited amount of information. The proposed approach involves a procedure for the selection of an approximation basis and the evaluation of the associated coefficients. The selection of the approximation basis relies on the a priori choice of the high-dimensional model representation format combined with a modified least angle regression technique. The resulting basis then provides the structure for the actual approximation basis, possibly using different functions, more parsimonious and nonlinear in its coefficients. To evaluate the coefficients, both an alternate least squares and an alternate weighted total least squares methods are employed. Examples are provided for the approximation of a random variable in a high-dimensional space as well as the estimation of a random field. Stochastic dimensions up to 100 are considered, with an amount of information as low as about 3 samples per dimension, and robustness of the approximation is demonstrated with respect to noise in the dataset. The computational cost of the solution method is shown to scale only linearly with the cardinality of the a priori basis and exhibits a $(N_q)^s$, $2 \leq s \leq 3$, dependence with the number N_q of samples in the dataset. The provided numerical experiments illustrate the ability of the present approach to derive an accurate approximation from scarce scattered data even in the presence of noise.

KEY WORDS: *uncertainty quantification, least angle regression, high-dimensional model reduction, total least squares, alternate least squares, polynomial chaos*

1. INTRODUCTION

With the growing available computational power, and as more efficient numerical methods become available, domains as diverse as engineering, chemistry, psychometrics, medicine, finance, or social sciences now heavily rely on simulation for the prediction of more and more complex phenomena, often combining multimodels and high accuracy requirement. The prediction capability of modern simulations is often such that a new bottleneck for accuracy has emerged from the lack of relevant boundary and/or initial conditions (BICs) as well as parameters intrinsic to the model of the system at hand, e.g., diffusivity, viscosity, etc. These sources of uncertainty are hereafter simply referred to as BICs. They are often poorly known and have to be estimated or modeled. This introduces modeling errors which often constitute the main source of lack of accuracy in the simulation chain. This situation has triggered a renewed interest for stochastic modeling where it is explicitly accounted for uncertainty in the model. The BICs may sometimes be modeled from first principles but are often approximated in a functional form involving a set of influencing parameters and identified from experimental measurements. However, more often than not, only relatively few

*Correspond to Lionel Mathelin, E-mail: mathelin@limsi.fr

2152–5080/14/$35.00

measurements are available, in particular when a significant number of parameters is of influence so that representing the BICs takes the form of a high-dimensional approximation problem.

If the random process, which output is to be represented in closed form, is driven by known equations, efficient techniques may be used to determine its representation. In the specific case of high-dimensional quantities, tensor-based representations have proved to be effective when applicable. In particular, low-rank approximations based on an *a priori* chosen separated representation can be efficiently derived, see [1–4] in the context of uncertainty quantification (UQ). If a closed-form model description of the process at hand is not available, one is typically left with approximating it from a finite collection of instances, hereafter termed samples. When the process is known only from a closed numerical code used as a black-box or if measurements can be made arbitrarily (design of experiments), some properties of approximation theory can be exploited. For instance, measurements may be taken at some particular locations in the parameter space, possibly associating a weight to them, so that the random Quantity of Interest (QoI) can be represented in the retained approximation basis with good accuracy using (sparse) quadrature techniques, [5]; see also [6] for an application to UQ. Anisotropy in the QoI may be exploited by biasing the quadrature weights [7–9]. In [10], an alternate least squares (ALS) technique to estimate the coefficients has been considered with samples lying on a tensor-product grid. Another situation of design of experiment arises in importance sampling where the Markov-Chain Monte Carlo algorithm requires a new sample at a specific proposed location. This control over the samples usually brings efficiency and allows one to approximate a reasonably behaved QoI with accuracy.

A different situation occurs when the data are scattered, with no ability to choose the set of samples nor to add a measurement. This is a common situation, typically arising when samples come from a past experiment or are costly to acquire so that new samples cannot be taken. In this context, one has to resort to a regression-based approach and the coefficients of the approximation are then solution of an optimization problem. This type of approach was considered in [11–13].

In the present work, the focus is specifically put on deriving a closed-form approximation of a high-dimensional quantity of interest from a small, uncontrolled, collection of its samples. This requires one to determine an approximation basis finely tuned to the data at hand and an efficient way of evaluating the associated coefficients. To this aim, we rely on the fact that, as a counterpart of the curse of dimensionality associated with high-dimensional problems, real applications often reward with a *blessing* of dimensionality. Indeed, in many cases, the QoI can be well approximated in a low-dimensional subspace of the solution space, sometimes involving orders of magnitude fewer degrees-of-freedom. This typically occurs when the solution exhibits some degree of sparsity in the retained functional space. Efficient techniques have been proposed in the recent past to take advantage of this situation and essentially consist of matching the approximation with the observational data while promoting a sparse coefficient set. This class of methods works well in many different contexts and has been recently applied to the UQ framework [14, 15]. These techniques rely on the compressed sensing theory, e.g., [16, 17], and may seem well suited for the present problem as they promote a low cardinality approximation of the QoI. However, they require one to handle a potentially huge representation basis, or *dictionary*, and associated optimization problem, leading to severe memory and computation limitations in the present high-dimensional context.

In this paper, we present a solution method combining the strength of different techniques, taking advantage of the sparsity of the representation in a suitable basis and allowing an efficient approximation of a well-behaved multivariate function with a low number of degrees-of-freedom hence compatible with a small experimental dataset. The driving principle is first to consider a tight approximation basis based on *a priori* knowledge on the QoI at hand and to rely on the available data to further refine it. In a nutshell, an initial approximation basis is first considered in the high-dimensional model representation format (HDMR, [18, 19]), assuming it is suitable for representing the QoI. This initial basis is hereafter referred to as *a priori* basis. Next, available data are used to refine it by retaining only its most relevant basis functions through a constructive subset selection procedure based on a modification of the Least Angle Regression approach proposed in [20]. This *a posteriori* basis defines a skeleton from which a final basis is built and the associated coefficients are evaluated with an alternate least squares technique. The solution method allows us to approximate random variables as well as random fields and is shown here to outperform both sparse grids and tensored-based techniques.

The paper is organized as follows. The representation of a random quantity is central to the methodology discussed in this paper. Standard techniques for deriving a closed-form approximation of a random variable from a finite set of

samples are briefly recalled in Section 2. Similarly, different representation formats of functions in high-dimensional spaces are subsequently heavily used in the paper and a short discussion is given in Section 3. The proposed solution method is introduced and discussed in Section 4 and an algorithm is given. Scalability of the proposed approach together with its robustness with respect to noise in the data is also discussed. In Section 5, the present methodology is illustrated on a stochastic diffusion equation involving up to 100 dimensions and on the space-dependent solution of the shallow water equations with random parameters. Accuracy, robustness, and scalability of the proposed approach are shown. Concluding remarks close the paper in Section 6.

2. QUANTIFICATION OF UNCERTAINTY

Thanks to its pivotal role in the rest of the paper, the representation of a random quantity and standard ways of evaluating it in closed form from a discrete set of samples is now briefly discussed.

2.1 General Framework

Random quantities are defined on a probability space $(\Theta, \mathcal{B}_\Theta, \mu_\Theta)$ where Θ is the space of elementary events $\theta \in \Theta$, $\mathcal{B}_\Theta$ a σ-algebra defined on Θ, and μ_Θ a probability measure on $\mathcal{B}_\Theta$. To make the description of the problem amenable to a tractable representation, it is convenient to introduce a finite set of statistically independent random variables $\{\xi_i\}_{i=1}^d : \Theta \to \Xi_i \subseteq \mathbb{R}$, $\theta \mapsto \xi_i(\theta)$. The set of these d random variables is defined on a probability space $(\Xi, \mathcal{B}_\Xi, \mu_\Xi)$ with $\Xi = \times_{i=1}^d \Xi_i = \boldsymbol{\xi}(\Theta) \subseteq \mathbb{R}^d$, $\boldsymbol{\xi} := (\xi_1 \ \dots \ \xi_d)$, $\mathcal{B}_\Xi \subset 2^\Xi$ a σ-algebra on Ξ and $\mu_\Xi = \mu_\Theta \circ \boldsymbol{\xi}^{-1}$ the probability measure on $\mathcal{B}_\Xi$. Since the physical process at hand relies on random quantities belonging to $(\Theta, \mathcal{B}_\Theta, \mu_\Theta)$, a suitable description of its output, or its solution in case the physical process is described by a known mathematical model, may be determined in $(\Xi, \mathcal{B}_\Xi, \mu_\Xi)$ as justified by the Doob-Dynkin lemma.

In this work, we restrict ourselves to random variables of physical significance, i.e., real-valued second-order variables satisfying

$$\mathop{\mathscr{E}}_{\theta}\left[u(\theta)^2\right] := \int_\Theta u(\theta)^2 \, \mathrm{d}\mu_\Theta(\theta) = \int_\Xi u(\boldsymbol{\zeta})^2 \, \mathrm{d}\mu_\Xi(\boldsymbol{\zeta}) =: \mathop{\mathscr{E}}_{\boldsymbol{\xi}}\left[u(\boldsymbol{\xi})^2\right] < +\infty, \tag{1}$$

where $\mathscr{E}$ denotes the expectation operator and u is the quantity of interest (QoI). It is then natural to consider the space of square integrable functions $\mathcal{S}$ for describing real-valued functions of the random quantities:

$$\mathcal{S} := L^2(\Xi, \mu_\Xi) = \left\{ v : \Xi \to \mathbb{R}, \boldsymbol{\xi} \mapsto v(\boldsymbol{\xi}) \,;\, \mathop{\mathscr{E}}_{\boldsymbol{\xi}}\left[v(\boldsymbol{\xi})^2\right] < +\infty \right\}. \tag{2}$$

Upon introduction of a natural inner product of $\mathcal{S}$: $\langle v, w \rangle_{L^2(\Xi, \mu_\Xi)} := \displaystyle\int_\Xi v(\boldsymbol{\zeta}) \, w(\boldsymbol{\zeta}) \, \mathrm{d}\mu_\Xi(\boldsymbol{\zeta})$, $\forall\, v, w \in \mathcal{S}$, and the associated norm $\|v\|^2_{L^2(\Xi, \mu_\Xi)} := \langle v, v \rangle_{L^2(\Xi, \mu_\Xi)}$, $\mathcal{S}$ is a Hilbert space. Further, we define $\langle v \rangle_{L^2(\Xi, \mu_\Xi)} := \mathscr{E}_{\boldsymbol{\xi}}[v(\boldsymbol{\xi})]$. One can now rely on functional analysis results and take advantage of approximation theory techniques to characterize the output u. Introducing a Hilbertian basis $\{\psi_k\}_{k \in \mathbb{N}}$ of $\mathcal{S}$, the output can then be uniquely represented as $u(\boldsymbol{\xi}) = \displaystyle\sum_{\boldsymbol{\alpha}} c_{\boldsymbol{\alpha}} \, \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$.

The basis $\{\psi_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in \mathbb{N}}$ is typically chosen orthonormal with respect to the inner product $\langle v, w \rangle_{L^2(\Xi, \mu_\Xi)}$. Orthonormality of the basis leads to $\langle \psi_{\boldsymbol{\alpha}}, \psi_{\boldsymbol{\alpha}'} \rangle_{L^2(\Xi, \mu_\Xi)} = \delta_{\boldsymbol{\alpha}\boldsymbol{\alpha}'}$, $\forall\, \boldsymbol{\alpha}, \boldsymbol{\alpha}' \in \mathbb{N}$, with δ the Kronecker delta, and the decomposition coefficients $\{c_{\boldsymbol{\alpha}}\}$ then express as

$$c_{\boldsymbol{\alpha}} = \langle u, \psi_{\boldsymbol{\alpha}} \rangle_{L^2(\Xi, \mu_\Xi)} = \int_\Xi u(\boldsymbol{\zeta}) \, \psi_{\boldsymbol{\alpha}}(\boldsymbol{\zeta}) \, \mathrm{d}\mu_\Xi(\boldsymbol{\zeta}), \qquad \forall\, \boldsymbol{\alpha} \in \mathbb{N}. \tag{3}$$

For a given representation basis $\{\psi_{\boldsymbol{\alpha}}\}$ of $\mathcal{S}$, the output $u(\boldsymbol{\xi})$ is entirely characterized by the set of coefficients $\{c_{\boldsymbol{\alpha}}\}$. For computational purpose, the infinite dimensional representation is substituted with a finite dimensional approximation relying on a subset $\mathcal{J} \subset \mathbb{N}$ of the representation basis:

$$u(\boldsymbol{\xi}) \approx \sum_{\boldsymbol{\alpha} \in \mathcal{J}} c_{\boldsymbol{\alpha}} \, \psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}). \tag{4}$$

2.2 Computing a Data-Driven Approximation

As seen above, in many situations, a closed-form model of the QoI is not available or not reliable enough to be used and one can only rely on the sole available input-output information to approximate the output u. The solution method then consists of using a set of outputs given some inputs, i.e., samples of the process. One then looks for a functional form of the map between the set of random variables $\boldsymbol{\xi}^{(q)}$ and the output value $u\left(\boldsymbol{\xi}^{(q)}\right) =: u^{(q)}, \forall 1 \leq q \leq N_q$, where N_q is the size of the available experimental set. Approximating the output under the functional form of Eq. (4) results in evaluating the coefficients $\{c_\alpha\}$ from $\left\{\left(\boldsymbol{\xi}^{(q)}, u^{(q)}\right)\right\}_{q=1}^{N_q}$, $\boldsymbol{\xi}^{(q)} = \left(\xi_1^{(q)} \ \dots \ \xi_d^{(q)}\right)$.

2.2.1 Direct Evaluation

If the sampling can be controlled, in the sense that samples can be drawn arbitrarily, the popular Monte Carlo approach can be followed and the approximation coefficients are then estimated from

$$c_\alpha = \int_\Xi u\left(\boldsymbol{\zeta}\right) \psi_\alpha\left(\boldsymbol{\zeta}\right) \mathrm{d}\mu_\Xi\left(\boldsymbol{\zeta}\right) \approx \sum_q u\left(\boldsymbol{\xi}^{(q)}\right) \psi_\alpha\left(\boldsymbol{\xi}^{(q)}\right). \tag{5}$$

Monte Carlo-based estimation is very robust and easy to implement but suffers from a slow $\mathcal{O}\left({N_q}^{-1/2}\right)$ asymptotic convergence rate. However, since the convergence rate does not depend on the dimensionality of the integral, this is a wise choice for very high-dimensional problems where other methods fail. Alternatively, quasi-Monte Carlo methods generate a low-discrepancy sequence of samples improving the convergence rate of the evaluation for moderate- to high-dimensional problems.

For low to moderate dimensionality problems, the d-dimensional integral arising in Eq. (3) may be advantageously evaluated with a quadrature rule:

$$c_\alpha = \int_\Xi u\left(\boldsymbol{\zeta}\right) \psi_\alpha\left(\boldsymbol{\zeta}\right) \mathrm{d}\mu_\Xi\left(\boldsymbol{\zeta}\right) \approx \sum_q w^{(q)}\, u\left(\boldsymbol{\xi}^{(q)}\right) \psi_\alpha\left(\boldsymbol{\xi}^{(q)}\right), \tag{6}$$

where $\left\{w^{(q)}\right\}$ are the weights associated with the quadrature points $\left\{\boldsymbol{\xi}^{(q)}\right\}$ [21].

2.2.2 Regression

The above methods require some kind of control over the samples. If no experimental design can be exploited, a solution method is then to reformulate the evaluation of the coefficients as a minimization problem:

$$\boldsymbol{c} = \underset{\widetilde{\boldsymbol{c}} \in \mathbb{R}^{|\mathcal{J}|}}{\arg\min} \left\| \boldsymbol{u} - \Psi\, \widetilde{\boldsymbol{c}} \right\|_2, \tag{7}$$

with $\boldsymbol{c} = \left(c_1 \ \dots \ c_{|\mathcal{J}|}\right)^T$, $\boldsymbol{u} = \left(u^{(1)} \ \dots \ u^{(N_q)}\right)^T$, $\Psi \in \mathbb{R}^{N_q \times |\mathcal{J}|}$, $\Psi_{q\alpha} = \psi_\alpha\left(\boldsymbol{\xi}^{(q)}\right)$, and $|\mathcal{J}|$ the cardinality of the approximation basis $\{\psi_\alpha\}_{\alpha \in \mathcal{J}}$. For a full column rank Ψ, the solution is given by $\boldsymbol{c} = \Psi^+ \boldsymbol{u}$ which is typically evaluated using the Cholesky decomposition of the symmetric positive definite matrix $\Psi^T \Psi$ or the QR decomposition of Ψ. When the size of the dataset grows, this standard least squares (LS) problem may become computationally involved. The quasi-regression solution alleviates the computational burden and is given by

$$c_\alpha = \boldsymbol{\psi}_\alpha^T \boldsymbol{u} / \left\|\psi_\alpha\right\|_2^2, \qquad \boldsymbol{\psi}_\alpha = \left(\psi_\alpha\left(\boldsymbol{\xi}^{(1)}\right) \ \dots \ \psi_\alpha\left(\boldsymbol{\xi}^{(N_q)}\right)\right)^T, \qquad 1 \leq \alpha \leq |\mathcal{J}|. \tag{8}$$

Standard LS formulation as considered in Eq. (7) treats all predictors $\{\psi_\alpha\}_{\alpha=1}^{|\mathcal{J}|}$ the same way and uses the available data to estimate all the coefficients to produce an estimate with a low bias but often a large variance. As will

be discussed in Section 4.3.1, additional properties of the QoI may be exploited or imposed to the approximation coefficients. This class of approaches trades some increase in bias with a decrease in variance and often results in an improved accuracy. A suitable solution method then typically formulates as a penalized LS problem

$$\boldsymbol{c} = \underset{\widetilde{\boldsymbol{c}} \in \mathbb{R}^{|\mathcal{J}|}}{\arg\min} \left\| \boldsymbol{u} - \Psi\, \widetilde{\boldsymbol{c}} \right\|_2 + \mathscr{J}\left(\widetilde{\boldsymbol{c}}\right). \tag{9}$$

The properties of the penalized LS solution are driven by the choice of the function $\mathscr{J}$, the flexibility of which leads to a variety of solution techniques; see [22, 23]. Since we have no control over the sampling strategy, we will rely on regression to estimate the approximation coefficients. The discussion of an efficient LS formulation in the present context is postponed to Section 4.

3. FUNCTIONAL REPRESENTATION OF RANDOM VARIABLES

3.1 Tensored Bases

As seen above, a random quantity is conveniently approximated in a Hilbertian basis $\{\psi_k\}$. If the random quantity is known, or expected, to exhibit a certain degree of smoothness along the stochastic space, a suitable and popular choice is to take advantage of this smoothness using a spectral-based approximation relying on polynomials. Early efforts toward this direction are the pioneering works of [24], who used univariate Hermite polynomials $\psi_\alpha\left(\xi_i\right)$ of zero-centered, unit variance, normal random variables $\xi_i \sim \mathcal{N}\left(0,1\right)$. These polynomials define an orthogonal basis of $L^2\left(\Xi_i,\, \mu_{\Xi i}\right)$, $\mu_{\Xi i} \propto e^{-(1/2)\xi_i^2}$. Tensorization of univariate Hermite polynomials ψ then leads to an orthogonal basis of $L^2\left(\Xi,\, \mu_\Xi\right)$:

$$\left\langle \psi_\alpha, \psi_{\alpha'} \right\rangle_{L^2(\Xi,\, \mu_\Xi)} \propto \int_\Xi \psi_\alpha\left(\boldsymbol{\zeta}\right) \psi_{\alpha'}\left(\boldsymbol{\zeta}\right) e^{-(1/2)(\boldsymbol{\zeta}^T \boldsymbol{\zeta})}\, \mathrm{d}\boldsymbol{\zeta} \propto \delta_{\alpha\alpha'}. \tag{10}$$

This can be extended to polynomials orthogonal with respect to different measures [25–27], and constitutes the so-called (generalized) polynomial chaos (PC) basis. A common practice is to consider an approximation space $\mathcal{S}_p$ spanned by polynomials of given maximum total degree p:

$$\mathcal{S}_p = \operatorname{span}\left(\left\{ \psi_{\boldsymbol{\alpha}}\left(\boldsymbol{\xi}\right) = \psi_{\alpha_1}\left(\xi_1\right) \ldots \psi_{\alpha_d}\left(\xi_d\right) \right\};\, \boldsymbol{\alpha} = \left(\alpha_1\, \ldots\, \alpha_d\right),\, \sum_{i=1}^{d} \alpha_i \leq p \right), \tag{11}$$

and the number of terms to be determined in the approximation (4) is then $|\mathcal{J}| = \begin{pmatrix} d+p \\ d \end{pmatrix}$. We adopt the convention $\psi_1 \equiv 1$. When the random quantity is not smooth enough for a low degree polynomial fit to be accurate, approximation schemes such as h/p-type refinement or multi-resolution analysis may be applied, see [28].

Some alternative representation formats specifically exploit the tensor-product structure of the Hilbert stochastic space $\mathcal{S}$ and approximates a d-variate function with a series of products of lower dimensional functions. Efficient algorithms allow us to determine the approximation coefficients of the representation by solving a series of low-dimensional problems while never considering the full-dimensional problem at once. A general presentation of tensor-structured numerical methods can be found in [29] while application to the approximation of a high-dimensional random quantity is considered in [3, 4, 10, 30]. For instance, a d-variate quantity may be approximated under a CANDECOMP-PARAFAC (CP) format [31, 32], with a sum of rank-1 terms, the simplest form of tensored-structure format:

$$u\left(\boldsymbol{\xi}\right) \approx \sum_{r=1}^{n_r} f_{1,r}\left(\xi_1\right)\, \ldots\, f_{d,r}\left(\xi_d\right), \tag{12}$$

with n_r the retained rank of the decomposition and $\{f_{i,r}\}_{i=1}^{d}$ univariate functions. Assuming pth-order polynomials for $\{f_{i,r}\}$, the resulting cardinality of the approximation is $d\, n_r\, p$. It thus exhibits a linear dependence with the number of dimensions, in contrast with the exponential dependence of the PC. Alternative decomposition techniques,

easier to evaluate and numerically more stable than decomposition (12), such as the Tucker or Tensor-Trains, can be considered, see [29]. A tensored-structure format then constitutes a method of choice for deriving memory- and CPU-efficient approximation of high-dimensional quantities. They also lead to a low-cardinality basis $|\mathcal{J}|$ so that the conditioning of the approximation method remains good, in the sense that $|\mathcal{J}| \leq N_q$, a crucial feature for deriving a good approximation from the scarce available data.

3.2 High-Dimensional Model Representation

An efficient alternative to these tensored-structure formats for representing high-dimensional quantities is discussed in [18, 19]. It consists of representing a quantity $u(\boldsymbol{\xi})$ with a sum of lower-dimensional terms accounting for increasing levels of interaction between the constitutive variables:

$$u(\boldsymbol{\xi}) = f_\emptyset + \sum_{i=1}^{d} f_i(\xi_i) + \sum_{\substack{i,j=1,\\ j>i}}^{d} f_{ij}(\xi_i,\xi_j) + \ldots + f_{12\ldots d}(\xi_1,\ldots,\xi_d) = \sum_{\boldsymbol{\gamma}\subseteq\{1,\ldots,d\}} f_{\boldsymbol{\gamma}}, \tag{13}$$

where $f_{\boldsymbol{\gamma}}$ are functions of $\mathcal{S}$ and depend only on a subset of variables $\boldsymbol{\xi}_{\boldsymbol{\gamma}} = \{\xi_i\}_{i\in\boldsymbol{\gamma}}$ and $\boldsymbol{\gamma}$ is a multi-index. This decomposition is exact, unique, and does not introduce any approximation. An important property is that the modes $\{f_{\boldsymbol{\gamma}}\}$ are mutually orthogonal: $\langle f_{\boldsymbol{\gamma}}, f_{\boldsymbol{\gamma}'}\rangle_{L^2(\Xi,\,\mu_\Xi)} = 0, \forall\,\boldsymbol{\gamma}\neq\boldsymbol{\gamma}' \subseteq \{1,\ldots,d\}$. The zeroth-order term $f_\emptyset$ accounts for the mean and is invariant across the entire domain Ξ, while the other modes are zero-mean:

$$f_\emptyset = \langle u\rangle_{L^2(\Xi,\,\mu_\Xi)}, \qquad \langle f_{\boldsymbol{\gamma}}\rangle_{L^2(\Xi,\,\mu_\Xi)} = 0, \quad \forall\,\boldsymbol{\gamma}\subseteq\{1,\ldots,d\}\setminus\emptyset. \tag{14}$$

The rationale behind the expected success of this so-called high-dimensional model representation (HDMR) is that many quantities of interest exhibit a significant dependence on low-dimensional groups of variables only, hence having negligible high order interaction decomposition terms. This leads to an efficient approximation of u with only a low N_l-order HDMR: $u(\boldsymbol{\xi}) \approx \sum_{\boldsymbol{\gamma}\subseteq\{1,\ldots,d\}} f_{\boldsymbol{\gamma}}(\boldsymbol{\xi}_{\boldsymbol{\gamma}})$, $|\boldsymbol{\gamma}| \leq N_l$. We denote $\mathcal{J}_f$ the set of retained modes, $\mathcal{J}_f := \{\boldsymbol{\gamma}\subseteq\{1,\ldots,d\}\,;|\boldsymbol{\gamma}|\leq N_l\}$.

Functions $\{f_{\boldsymbol{\gamma}}\}$ are evaluated with the application of a set of commuting projections $\{\mathcal{P}_i\}$ onto the output u. The projection $\mathcal{P}_i$ eliminates the effect of variable ξ_i while leaving the effect of the others unchanged. Letting $\mathcal{P}_\emptyset$ be the identity operator on $\mathcal{S}$, we define $\mathcal{P}_{\boldsymbol{\eta}} = \prod_{i\in\boldsymbol{\eta}} \mathcal{P}_i, \forall\,\boldsymbol{\eta}\subseteq\{1,\ldots,d\}$. Functions $\{f_{\boldsymbol{\gamma}}\}$ can then be written [33],

$$f_{\boldsymbol{\gamma}\subseteq\{1,\ldots,d\}\setminus\emptyset} = \mathcal{P}_{\{1,\ldots,d\}\setminus\boldsymbol{\gamma}}\,u - \sum_{\boldsymbol{\gamma}'\subsetneq\boldsymbol{\gamma}} f_{\boldsymbol{\gamma}'} = \sum_{\boldsymbol{\gamma}'\subseteq\boldsymbol{\gamma}} (-1)^{|\boldsymbol{\gamma}|-|\boldsymbol{\gamma}'|}\,\mathcal{P}_{\{1,\ldots,d\}\setminus\boldsymbol{\gamma}'}\,u, \qquad f_\emptyset = \mathcal{P}_{\{1,\ldots,d\}}\,u. \tag{15}$$

Defining projections as $\mathcal{P}_i\,u(\boldsymbol{\xi}) = \int_{\Xi_i} u(\xi_1,\ldots,\xi_{i-1},\zeta',\xi_{i+1},\ldots,\xi_d)\,\mathrm{d}\mu(\zeta')$, the measure μ determines the form of the projection. A popular choice consists of using $\mu = \mu_{\Xi i}$ so that the Analysis of Variance (ANOVA) decomposition is obtained. An example of application of the HDMR representation to the approximation of a random quantity is presented in [9].

Remark 1. *These different functional representations are not totally distinct. For instance, the PC basis defined in Eq.* (11) *can also be interpreted as a particular case of both HDMR and tensor-based expansion. For illustration, consider the following PC basis approximation space* $\mathcal{S}_p = \mathrm{span}\,(\{\psi_1(\equiv 1), \psi_2(\xi_1), \psi_2(\xi_2), \psi_3(\xi_1), \psi_2(\xi_1)\,\psi_2(\xi_2), \psi_3(\xi_2)\})$. *This corresponds to a HDMR representation with* $N_l = 2$ *and* $f_\emptyset \in \mathrm{span}\,(\psi_1)$, $f_1 \in \mathrm{span}\,(\psi_2(\xi_1), \psi_3(\xi_1))$, $f_2 \in \mathrm{span}\,(\psi_2(\xi_2), \psi_3(\xi_2))$, $f_{12} \in \mathrm{span}\,(\psi_2(\xi_1)\,\psi_2(\xi_2))$. *Further, this can also be reformatted in a* $n_r = 3$*-rank CP format, say with* $f_{1,1} \in \mathrm{span}\,(\psi_1)$, $f_{2,1} \in \mathrm{span}\,(\psi_1, \psi_2(\xi_2), \psi_3(\xi_2))$, $f_{1,2} \in \mathrm{span}\,(\psi_2(\xi_1))$, $f_{2,2} \in \mathrm{span}\,(\psi_1, \psi_2(\xi_2))$, $f_{1,3} \in \mathrm{span}\,(\psi_3(\xi_1))$ *and* $f_{2,3} \in \mathrm{span}\,(\psi_1)$.

4. QUANTIFYING UNCERTAINTY OF SCATTERED DATA

4.1 Setting Up the Stage

In the following, we will consider that the quantity of interest u is a scalar-valued random field, indexed by space and/or time $\boldsymbol{x} \in \mathbb{R}^{d_{\boldsymbol{x}}}$ and depending on a set of random variables $\boldsymbol{\xi} \in \mathbb{R}^d$. To approximate it, the only available piece of information is a collection of scattered samples $\left\{\boldsymbol{x}^{(q)}, \boldsymbol{\xi}^{(q)}, u^{(q)}\right\}_{q=1}^{N_q}$. In case these data come from an experimental context, the coordinates $\boldsymbol{\xi}^{(q)}$ are not directly measurable. They are then inferred from auxiliary observations and depend on the modelization.[1] Since the underlying random quantity u is only known through these samples, no governing equation for the QoI can be exploited and, say, Galerkin projection-based weak-formulation methods cannot be employed. Further, these samples are scattered and do not follow a deterministic rule so that no deterministic sampling strategy can be assumed. Quadrature-based techniques can then not be applied either and one has to resort to regression to estimate the coefficients of the approximation in the retained basis $\{\psi_\alpha\}$. Standard L^2-regression solves Eq. (7) which is only well-posed for a matrix Ψ such that $\Psi^T \Psi$ is invertible so that it requires the number of observations to be larger than the cardinality of the approximation basis, $N_q \geq |\mathcal{J}|$.

The choice of a good approximation basis in a general setting largely remains an open question. If one is given a dictionary of approximation functions, *a priori* selecting the best terms so that they can be evaluated from the data is a combinatorial optimization problem which algorithmic complexity quickly becomes intractable when the size of the dictionary grows. Dictionary-learning techniques require a training while availability of an independent training set cannot be assumed here.

The proposed approach is as follows. We separate the determination of an efficient representation format from the evaluation of the coefficients. We first choose an *a priori* general format for the approximation of u, Section 4.2. The selection of particular terms to be included in the approximation basis is left to a dedicated subset selection procedure which will further refine the approximation basis and make it as tight as possible, Section 4.3. A good *a priori* basis is motivated by results from compressed sensing which show that the number of samples necessary for accurately selecting the dominant basis functions of a K-sparse QoI (i.e., having K non-zero coefficients in the retained approximation basis) varies as $K \log(|\mathcal{J}|)$ [34], illustrating the fact that it becomes increasingly difficult to select the best terms when the size $|\mathcal{J}|$ of the *a priori* dictionary increases. The subset selection hence produces an *a posteriori* basis suitable for the data at hand. However, this basis is *linear* in its predictors as required by the selection method. To circumvent this limitation, the *a posteriori* basis is used as a skeleton only, of the best structure, and the final approximation of the QoI is evaluated with a different basis, of the same skeleton, but possibly nonlinear in its predictors, Section 4.4. A sketch of the solution method is shown in Fig. 1.

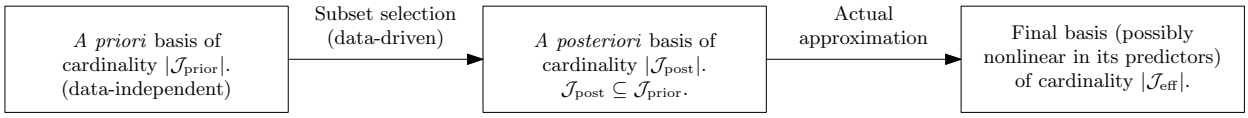

FIG. 1: Sketch of the solution method.

4.2 A Priori Choice of Representation of a Random Variable

We first focus on approximating a random variable and will discuss approximation of a more general random process in Section 4.8. The QoI is hence here a random variable $u(\boldsymbol{\xi})$.

In this work, we want to take advantage of the low order interactions of constitutive variables for many quantities of practical interest as mentioned in Section 3.2. Previous works have shown evidence of this low interaction configuration in various situations [9, 18, 19], and the QoI is hence chosen to be approximated under the HDMR form,

[1] For instance, in a fluid flow, the Reynolds number may be uncertain and modeled as a random variable parametrized by ξ_i. The value of ξ_i in each sample $\boldsymbol{\xi}^{(q)}$ is then auxiliary deduced from the measurement of the flow velocity V and the model $V(\xi_i)$.

Eq. (13). An example is considered in Appendix A and demonstrates that a general HDMR format approximation with a tensor-based description of the interaction modes $\{f_{\boldsymbol{\gamma}}\}$ involved in the HDMR may compare favorably with a full tensor-based approximation in terms of required number $|\mathcal{J}_{\rm prior}|$ of basis functions for a given reconstruction accuracy, even for reasonably large dimensional problems. This motivates our choice of an HDMR format for the *a priori*, data-independent, basis.

4.3 Subset Selection

We now build upon from the *a priori* basis and further improve it with an *a posteriori*, data-driven, procedure.

4.3.1 A Direct Approach

As discussed in Section 2.2, different techniques may be used to compute the coefficients of an approximation. In the case considered in this paper, the available data are scarce while the cardinality $|\mathcal{J}_{\rm prior}|$ of the *a priori* approximation basis may be large, in particular when the dimensionality d of the problem is large. It can then result in an ill-posed problem where one has to estimate $|\mathcal{J}_{\rm prior}|$ coefficients for each stochastic mode λ_n from $N_q \ll |\mathcal{J}_{\rm prior}|$ pieces of information. However, this situation often only reflects our lack of knowledge on the quantity at hand and how conservative this naive approximation method is. Indeed, high-dimensional problems are often intrinsically sparse and lower dimensional. In the present setting, it is likely that many dimensions actually hardly contribute to the approximation and that representing the dependence of the QoI along only a subset of the dimensions yields an acceptable accuracy. In our *a priori* HDMR representation, it means that many interaction modes $\{f_{\boldsymbol{\gamma}}\}$ can be discarded without significantly affecting the accuracy. The challenge for an efficient solution method is then to reveal and exploit the low-dimensional manifold onto which a good approximation of the solution lies. As an illustration, if $u(\boldsymbol{\xi}) = g(\xi_i)$ was depending only on one dimension i, $i \in \{1,\dots,d\}$, information theory allows us to show that one only requires $m+1+\lceil \log_2 d \rceil$ function evaluations to approximate a sufficiently smooth function $g \in C^s$, having s continuous derivatives, so that $\|u-\widehat{u}\|_{C(\Xi_i)} \leq a\, h^s$, $h := 1/m$, where $a \geq 0$ is related to a norm of g [35]. This number of samples actually is directly related to the number of information bits required to represent the integer $i \in [1,d]$.

While determining which interaction modes are dominant is an NP-hard problem in general, recent results have shown that a good estimation of the best subset can be obtained as the solution of a convex optimization problem. In particular, the LASSO formulation [36] has been proven effective. One of its formulations, referred to as *basis pursuit denoising*, writes

$$\boldsymbol{c} = \underset{\widetilde{\boldsymbol{c}}\in\mathbb{R}^{|\mathcal{J}_{\rm prior}|}}{\arg\min} \|\widetilde{\boldsymbol{c}}\|_1 \quad \text{s.t.} \quad \|\boldsymbol{u} - \Psi\,\widetilde{\boldsymbol{c}}\|_2 \leq \epsilon, \tag{16}$$

with Ψ the matrix of evaluations of the approximation basis and ϵ the approximation residual. Efforts from the signal processing community, where the theory supporting these results is termed *compressed sensing*, have demonstrated its good recovery properties in the case where $N_q < |\mathcal{J}_{\rm prior}|$, e.g., [17, 37, 38]. In particular, this formulation achieves provable and robust recovery bounds.[2]

The compressed sensing technique was proved very effective and is now being applied in many areas, including uncertainty quantification, [14, 15]. However, standard implementations of the algorithm require the sensing matrix Ψ to be available. This bears an intrinsic limitation when it comes to high-dimensional problems as it requires the use of the whole dictionary at once from which to select the basis functions associated with the dominant coefficients. While effective, this approach is not deemed tractable for high-dimensional problems, neither in terms of storage requirement nor CPU burden.

[2]For a sufficiently incoherent set of approximation and test functions, a K-sparse solution $\boldsymbol{c}$ to Eq. (16) satisfies [39] $\|\boldsymbol{c}^\star - \boldsymbol{c}\|_2 \lessapprox h\left(\epsilon + \|\boldsymbol{c}^\star - \boldsymbol{c}_K^\star\|_1/\sqrt{K}\right)$, where $h>0$ is a constant depending on the set of approximation and test functions and $\boldsymbol{c}_K^\star$ is the K-term approximation of $\boldsymbol{c}^\star$ given by an oracle, i.e., it is the best K-term approximation of $\boldsymbol{c}^\star$ if one was given full knowledge of it.

4.3.2 A Progressive Selection

To circumvent the issues identified above, we here use a bottom-to-top approach which achieves a forward stagewise regression by progressively revealing important basis functions. Introduced by [20, 23], the least angle regression selection (LARS) technique relies on analytical solutions to speed up computations and essentially follows the piecewise linear regularization path of the LASSO.[3] One advantage of LARS over other techniques is that the potential dictionary is never stored nor used as a whole. A LARS approach in the UQ framework was also considered in [40].

We consider the following polynomial approximation $\widetilde{f}_{\boldsymbol{\gamma}}$ of $f_{\boldsymbol{\gamma}}$:

$$f_{\boldsymbol{\gamma}}\left(\{\xi_i\}_{i\in\boldsymbol{\gamma}}\right) \approx \widetilde{f}_{\boldsymbol{\gamma}}\left(\{\xi_i\}_{i\in\boldsymbol{\gamma}}\right) := \sum_{\boldsymbol{\alpha},\,|\boldsymbol{\alpha}|\le\widetilde{p}} c_{\boldsymbol{\gamma},\boldsymbol{\alpha}}\,\psi_{\boldsymbol{\alpha}}\left(\{\xi_i\}_{i\in\boldsymbol{\gamma}}\right), \qquad \psi_{\boldsymbol{\alpha}} = \prod_{i\in\boldsymbol{\gamma}} \psi_{\alpha_i}\left(\xi_i\right), \tag{17}$$

with $\boldsymbol{\alpha} = (\alpha_i, i \in \boldsymbol{\gamma})$, $\alpha_i \in \{1,\ldots,\widetilde{p}\}$. Interaction modes $\{f_{\boldsymbol{\gamma}}\}$ are then approximated in $\mathbb{P}_{\widetilde{p}}$, the space of polynomials with maximum total degree $\widetilde{p}$, by modes $\left\{\widetilde{f}_{\boldsymbol{\gamma}}\right\}$ linear in their coefficients.

In the present framework, the HDMR approximation format naturally leads to *groups* of predictors whose importance in describing the QoI u follows a similar trend. These groups are defined by the subsets $\{\mathcal{J}_{\boldsymbol{\gamma}}\}$ of predictors which belong to a given interaction mode $f_{\boldsymbol{\gamma}}$, $\mathcal{J}_{\boldsymbol{\gamma}} = \left\{\psi_{\boldsymbol{\alpha}}\left(\{\xi_i\}_{i\in\boldsymbol{\gamma}}\right)\right\}$, and are likely to be strongly correlated. For instance, if the QoI exhibits a strong dependence on a given dimension ξ_j, one then wants to incorporate the whole set of predictors $\left\{\psi_{\boldsymbol{\alpha}}\left(\{\xi_i\}_{i\in\boldsymbol{\gamma}}\right)\right\}$, $\boldsymbol{\gamma} : j \in \boldsymbol{\gamma}$ without evaluating their relevance individually. One then looks for an approximation which is sparse at the level of groups of functions. Note that grouping predictors significantly alleviates the computational cost associated with the subset selection as further discussed in Section 4.7.

It is important to recall that this approximation format is made only for the subset selection step and is independent of the format the QoI will finally be approximated in. The selection of groups reduces to selection of interaction modes $f_{\boldsymbol{\gamma}}$ and leaves the possibility for using different formats between the subset selection step and the coefficients evaluation step: an interaction mode found to be dominant is incorporated to the active dictionary $\mathcal{J}_{f,\mathrm{post}}$ independently of the way its contribution to the approximation of u is actually determined in the end. Indeed, since the LARS technique only applies to predictors *linear* in their coefficients, an approximation $\widetilde{f}_{\boldsymbol{\gamma}}$ of the form (17) is suitable for the selection of the dominant groups. However, the final approximation $\widehat{f}_{\boldsymbol{\gamma}}$ of the retained $f_{\boldsymbol{\gamma}}$ may rely on predictors *nonlinear* in their coefficients: the subset selection step only serves to determine which interaction modes will be considered in the *a posteriori* approximation basis, the "skeleton" $\{f_{\boldsymbol{\gamma}} : \boldsymbol{\gamma} \in \mathcal{J}_{f,\mathrm{post}}\}$.

The selection is made using a modified LARS approach and the following optimization problem is solved:

$$\boldsymbol{c} = \underset{\widetilde{\boldsymbol{c}}\in\mathbb{R}^{|\mathcal{J}|}}{\arg\min} \|\boldsymbol{u} - \Psi\,\widetilde{\boldsymbol{c}}\|_2^2 + \tau \sum_{\boldsymbol{\gamma}\in\mathcal{J}_f} \|\widetilde{\boldsymbol{c}}_{\boldsymbol{\gamma}}\|_{K_{\boldsymbol{\gamma}}}, \tag{18}$$

with $\tau > 0$ the regularization parameter and $\|\cdot\|_{K_{\boldsymbol{\gamma}}}$ a norm induced by a positive definite matrix $K_{\boldsymbol{\gamma}}$. All predictors within a group $\boldsymbol{\gamma}$ are here weighted similarly so that we use a scaled identity matrix $K_{\boldsymbol{\gamma}} = \mathrm{I}_{|\mathcal{J}_{\boldsymbol{\gamma}}|}/|\mathcal{J}_{\boldsymbol{\gamma}}|, \forall\,\boldsymbol{\gamma} \in \mathcal{J}_f$. The regularization term is a combination of L^2- and L^1-norms and penalizes the L^1-norm of the "group" vector to promote a collective behavior: either a group is basically active (nonzero $K_{\boldsymbol{\gamma}}$-norm) or inactive, essentially disregarding the detailed behavior within the group. This group LARS (gLARS) strategy was first proposed in [41] and the algorithm presented in [42] was modified to solve the optimization problem (18).

[3] In a nutshell, it consists of selecting, from the *a priori* set $\mathcal{J}_{\mathrm{prior}}$, the predictor (approximation function) which is most correlated with the current residual, move this predictor to the active set $\mathcal{J}_{\mathrm{post}}$, compute the increment solution vector by minimizing the residual L^2-norm, and follow the descent direction along the increment vector until a predictor from the inactive set becomes as correlated with the residual as those from the active set. The whole process is then repeated and allows to sequentially build the optimal subset of approximation functions by exploring the Pareto front defined by the competition between the two terms of the unconstrained formulation of the optimization problem of Eq. (16).

The set of dominant modes $\{f_{\boldsymbol{\gamma}}\}$ is first determined by the gLARS approach with a low approximation order $\widetilde{p}$ and the basis is subsequently further refined by a LARS step, using L^1-regularization, onto these selected modes only now approximated with a higher $\widetilde{p}$ for improved accuracy.

4.4 Functional Spaces for the Final Approximation Basis

We now discuss the general methodology for approximating a random variable $u(\boldsymbol{\xi})$, from a finite set of its realizations. An *a priori* choice of representation format was first made, Section 4.2, and was adjusted based on the data through the subset selection procedure, the *a posteriori* step, Section 4.3. This has selected a set of groups, or interaction modes, $\{f_{\boldsymbol{\gamma}}\}_{\boldsymbol{\gamma}\in\mathcal{J}_{f,\text{post}}}$ deemed to most contribute to the HDMR representation of the QoI u. The actual approximation of u will rely on these selected groups but does not bear restriction on the linearity with respect to the coefficients so that different suitable formats, possibly nonlinear, can then be considered.

Many possibilities exist to determine an approximation of $\{f_{\boldsymbol{\gamma}}, \boldsymbol{\gamma}\in\mathcal{J}_{f,\text{post}}\}$ in a polynomial space, e.g., maximum partial degree, maximum total degree, hyperbolic cross, etc. For sake of simplicity, the space $\mathbb{P}_p$ of polynomials with maximum total degree p is retained as a reasonable compromise between cardinality $|\mathcal{J}_{\boldsymbol{\gamma}}|$ and expected accuracy of the approximation $\widehat{f}_{\boldsymbol{\gamma}}$:

$$f_{\boldsymbol{\gamma}}\left(\{\xi_i\}_{i\in\boldsymbol{\gamma}}\right) \approx \widehat{f}_{\boldsymbol{\gamma}}\left(\{\xi_i\}_{i\in\boldsymbol{\gamma}}\right) = \sum_{\boldsymbol{\alpha},|\boldsymbol{\alpha}|\le p} c_{\boldsymbol{\gamma},\boldsymbol{\alpha}}\,\psi_{\boldsymbol{\alpha}}\left(\{\xi_i\}_{i\in\boldsymbol{\gamma}}\right), \qquad \psi_{\boldsymbol{\alpha}}\left(\{\xi_i\}_{i\in\boldsymbol{\gamma}}\right) = \prod_{i\in\boldsymbol{\gamma}} \psi_{\alpha_i}(\xi_i),$$

$$\boldsymbol{\alpha} = (\alpha_i, i\in\boldsymbol{\gamma}), \qquad \alpha_i\in\{1,\dots,p\}, \qquad 1\le|\boldsymbol{\gamma}|\le N_l^{(\text{PC})}\le\min(N_l,p). \tag{19}$$

The cardinality associated with this approximation of $f_{\boldsymbol{\gamma}}$ at a given iteration level $l=|\boldsymbol{\gamma}|$ is $|\mathcal{J}_{\boldsymbol{\gamma}}| = p!/(l!\ (p-l)!)$ and usually provides an accurate approximation with a low number of coefficients for low dimensions $|\boldsymbol{\gamma}|$.

When the dimension $|\boldsymbol{\gamma}|$ increases, the number of terms in $\widehat{f}_{\boldsymbol{\gamma}}$ decreases and eventually degenerates for $|\boldsymbol{\gamma}|>p$. For modes of interaction order highe-than a prescribed threshold $N_l^{(\text{PC})}$, a low-rank canonical decomposition is instead considered:

$$f_{\boldsymbol{\gamma}}\left(\{\xi_i\}_{i\in\boldsymbol{\gamma}}\right) \approx \widehat{f}_{\boldsymbol{\gamma}}\left(\{\xi_i\}_{i\in\boldsymbol{\gamma}}\right) = \sum_{r=1}^{n_r}\prod_{i\in\boldsymbol{\gamma}}\sum_{\alpha=1}^{p} c_{\boldsymbol{\gamma},\alpha}^{r,i}\,\psi_\alpha(\xi_i), \qquad N_l^{(\text{PC})}<|\boldsymbol{\gamma}|\le N_l\le d. \tag{20}$$

The maximum number of modes at a given interaction level l is $d!/((d-l)!\,l!)$. Relying on an approximation in $\mathbb{P}_p$ for interaction modes of order $|\boldsymbol{\gamma}|\le N_l^{(\text{PC})}$ and on low-rank approximation for higher interaction order modes, with maximum rank n_r, the total cardinality of this approximation format is bounded from above by

$$|\mathcal{J}_{\text{eff}}| \le \sum_{l=0}^{N_l^{(\text{PC})}} \frac{d!\,p!}{(d-l)!\ (l!)^2\ (p-l)!} + \sum_{l=N_l^{(\text{PC})}+1}^{N_l} \frac{d!}{(d-l)!\,l!}\,n_r\,l\,p. \tag{21}$$

4.5 Algorithm for Approximating a Random Variable

We will denote $\mathcal{J}_{f,\text{eff}}$ the set of modes $\left\{\widehat{f}_{\boldsymbol{\gamma}}\right\}_{\boldsymbol{\gamma}\subseteq\mathcal{J}_{f,\text{post}}}$ finally considered for the approximation of u and $\mathcal{J}_{\text{eff}}$ the set of associated predictors $\{\psi_{\boldsymbol{\alpha}}\}$. The interaction modes are estimated sequentially. Once a new mode is evaluated, the whole approximation may be updated by reevaluating the coefficients of the predictors $\{\psi_{\boldsymbol{\alpha}}\}$ already evaluated of the current evaluation set $\mathcal{J}_{f,\text{eff}}\in\mathcal{J}_{f,\text{post}}$. Let $\boldsymbol{z}=\left(z^{(1)}\ \dots\ z^{(N_q)}\right)^T$ be the residual vector after basis functions $\widehat{f}_{\boldsymbol{\gamma}}, \boldsymbol{\gamma}\in\mathcal{J}_{f,\text{eff}}$ have been evaluated. The coefficients involved in the next mode $\widehat{f}_{\boldsymbol{\gamma}}, \boldsymbol{\gamma}\in\mathcal{J}_{f,\text{post}}\backslash\mathcal{J}_{f,\text{eff}}$ to be evaluated are then determined. If $\boldsymbol{\gamma}$ is such that $|\boldsymbol{\gamma}|\le N_l^{(\text{PC})}$, they are computed from the following system of equations[4]:

[4]While not found necessary here, the solution of the LS problem may be regularized by adding a generic term of the form $\beta\,\|L\,\widetilde{\boldsymbol{c}}\|_2$. A typical choice is $L=I_{|\mathcal{J}_{\boldsymbol{\gamma}}|}$ but one may also want to consider nondiagonal matrices L.

$$\left\{ \begin{array}{l} \boldsymbol{c}_{\boldsymbol{\gamma},\cdot} = \underset{\widetilde{\boldsymbol{c}}\in\mathbb{R}^{|\mathcal{J}_{\boldsymbol{\gamma}}|}}{\arg\min} \left\| \boldsymbol{z} - \Psi\, \widetilde{\boldsymbol{c}} \right\|_2, \qquad \forall\, i \in \boldsymbol{\gamma}, \quad \boldsymbol{\gamma} \subseteq \mathcal{J}_{f,\mathrm{post}}, \quad |\boldsymbol{\gamma}| \leq N_l^{(\mathrm{PC})}, \end{array} \right. \tag{22}$$

with $\boldsymbol{c}_{\boldsymbol{\gamma},\cdot} = \left(\boldsymbol{c}_{\boldsymbol{\gamma},\alpha_i}, i \in \boldsymbol{\gamma}\right)^T$ and

$$\begin{aligned} z^{(q)} &= u^{(q)} - \sum_{\boldsymbol{\gamma}'\subseteq\mathcal{J}_{f,\mathrm{eff}}\setminus\boldsymbol{\gamma}} \widehat{f}_{\boldsymbol{\gamma}'}\left(\left\{\xi_i^{(q)}\right\}_{i\in\boldsymbol{\gamma}'}\right), \qquad \boldsymbol{z} = \left(z^{(1)} \,\ldots\, z^{(N_q)}\right)^T, \\ \Psi_{q\boldsymbol{\alpha}} &= \psi_{\boldsymbol{\alpha}}\left(\left\{\xi_i^{(q)}\right\}_{i\in\boldsymbol{\gamma}}\right), \qquad \Psi = \left[\Psi_{q\boldsymbol{\alpha}}\right]. \end{aligned} \tag{23}$$

To solve for the coefficients associated with predictors nonlinear in their coefficients, an alternate least squares (ALS) approach is used, reformulating the nonlinear problem into a set of coupled linear equations:

$$\left\{ \begin{array}{l} \boldsymbol{c}_{\boldsymbol{\gamma},\cdot}^{r,i} = \underset{\widetilde{\boldsymbol{c}}\in\mathbb{R}^{p}}{\arg\min} \left\| \boldsymbol{z}_i - \Psi\, \widetilde{\boldsymbol{c}} \right\|_2, \qquad \forall\, i \in \boldsymbol{\gamma} \subseteq \mathcal{J}_{f,\mathrm{post}}, \quad N_l^{(\mathrm{PC})} < |\boldsymbol{\gamma}| \leq N_l, \end{array} \right. \tag{24}$$

with $\boldsymbol{c}_{\boldsymbol{\gamma},\cdot}^{r,i} = \left(\boldsymbol{c}_{\boldsymbol{\gamma},1}^{r,i} \,\ldots\, \boldsymbol{c}_{\boldsymbol{\gamma},p}^{r,i}\right)^T$ and

$$\begin{aligned} z_i^{(q)} &= u^{(q)} - \sum_{\boldsymbol{\gamma}'\subseteq\mathcal{J}_{f,\mathrm{eff}}} \widehat{f}_{\boldsymbol{\gamma}'}\left(\left\{\xi_i^{(q)}\right\}_{i\in\boldsymbol{\gamma}'}\right) - \sum_{r'=1}^{r-1} \prod_{i'\in\boldsymbol{\gamma}} \sum_{\alpha'=1}^{p} c_{\boldsymbol{\gamma},\alpha'}^{r',i'}\, \psi_{\alpha'}\left(\xi_{i'}^{(q)}\right), \\ \Psi_{q\alpha} &= \psi_{\alpha}\left(\xi_i^{(q)}\right) \prod_{i'\in\boldsymbol{\gamma},i'\neq i} \sum_{\alpha'=1}^{p} c_{\boldsymbol{\gamma},\alpha'}^{r,i'}\, \psi_{\alpha'}\left(\xi_{i'}^{(q)}\right), \qquad \Psi = \left[\Psi_{q\boldsymbol{\alpha}}\right]. \end{aligned} \tag{25}$$

This whole step is embedded in a loop over the modes $f_{\boldsymbol{\gamma}}, \boldsymbol{\gamma} \in \mathcal{J}_{f,\mathrm{post}}$ retained by the subset selection procedure. The cross-validation error (CVε) is estimated from $\widehat{N_q}$ validation samples $\left\{\boldsymbol{\xi}^{(\widehat{q})}, u^{(\widehat{q})}\right\}_{\hat{q}=1}^{\widehat{N_q}}$ independent from the N_q samples of the training set.[5] If the cross-validation error has increased over the last two loops, the approximation basis is likely to have become too large with respect to the available data and iterations are stopped. The retained basis is then the one that has led to the lowest CVε. If CVε keeps decreasing, the next interaction mode as selected by the subset selection step is considered and added to the current active set $\mathcal{J}_{f,\mathrm{eff}}$ and the whole iteration is carried out. Once the approximation is determined, the coefficients are updated with the same sequential technique using both the training and the validation points, $N_q + \widehat{N_q}$. The approximation accuracy is estimated by the relative L^2-norm ε of the approximation error estimation evaluated from a $\widetilde{N_q}$-point test set $\left\{\left(x^{(\widetilde{q})}, \boldsymbol{\xi}^{(\widetilde{q})}, u^{(\widetilde{q})}\right)\right\}_{\tilde{q}=1}^{\widetilde{N_q}}$, independent from the training set:

$$\varepsilon^2 := \left\| \boldsymbol{u} - \widehat{\boldsymbol{u}} \right\|_2^2 / \left\| \boldsymbol{u} \right\|_2^2, \qquad \boldsymbol{u} = \left(u^{(1)} \,\ldots\, u^{(\widetilde{N_q})}\right), \qquad \widehat{\boldsymbol{u}} = \left(\widehat{u}^{(1)} \,\ldots\, \widehat{u}^{(\widetilde{N_q})}\right). \tag{26}$$

The global methodology is summarized in Algorithm 1. Statistical moments can be readily evaluated from the present HDMR of the QoI, see Appendix B.

[5] A ratio $\widehat{N_q}/N_q \simeq 1/2$ is typically accepted as a reasonable splitting of the set of samples. We here use the simplest cross-validation method but more sophisticated techniques (k-fold, Leave-One-Out, etc.) are available, see for instance [23]. While more accurate, they are significantly more computationally expensive.

Algorithm 1: Sketch of the solution method for approximating a random variable $u(\boldsymbol{\xi})$

Select an *a priori* basis in HDMR format. Choose p, N_l, $N_l^{(\mathrm{PC})}$, n_r and $\widetilde{p}$. Initialize $\boldsymbol{z} = \left(u^{(1)} \dots u^{(N_q)}\right)^T$.
Subset selection step. Solve the LASSO optimization problem with the gLARS algorithm $\longrightarrow$ sequence of *a posteriori* approximation bases indexed by s with ordered groups $\mathcal{J}_{f,\mathrm{post}} = \left\{\boldsymbol{\gamma}^{(s)}\right\}$. Initialize s and $\mathcal{J}_{f,\mathrm{eff}}$: $s \leftarrow 0$, $\mathcal{J}_{f,\mathrm{eff}} \leftarrow \emptyset$.
Solve the approximation problem:
repeat
  $s \leftarrow s+1$.
  Consider the next mode $\widehat{f}_{\boldsymbol{\gamma}^{(s)}}$ from the set $\mathcal{J}_{f,\mathrm{post}}$ selected in (2): $\mathcal{J}_{f,\mathrm{eff}} \leftarrow \mathcal{J}_{f,\mathrm{eff}} \bigcup \boldsymbol{\gamma}^{(s)}$.
  Solve for the approximation coefficients $\{\boldsymbol{c}_{\boldsymbol{\gamma}}\}_{\boldsymbol{\gamma}\in\mathcal{J}_{f,\mathrm{eff}}}$ by alternately solving for the coefficients of modes $\left\{\widehat{f}_{\boldsymbol{\gamma}}\right\}_{\boldsymbol{\gamma}\in\mathcal{J}_{f,\mathrm{eff}}}$, Eqs. (22, 24). [**Update step**]
  Estimate the cross-validation error CVε and evaluate the current approximation $\widehat{\boldsymbol{u}} = \left(\widehat{u}^{(1)} \dots \widehat{u}^{(N_q)}\right)^T$.
  Update the residual $\boldsymbol{z} \leftarrow \boldsymbol{u} - \widehat{\boldsymbol{u}}$.
until *CVε has increased over the last two passes s and $s-1$.*
$\mathcal{J}_{f,\mathrm{eff}} \leftarrow \mathcal{J}_{f,\mathrm{eff}} \backslash \left\{\boldsymbol{\gamma}^{(s)}, \boldsymbol{\gamma}^{(s-1)}\right\}$.
Update the coefficients $\{\boldsymbol{c}_{\boldsymbol{\gamma}}\}_{\boldsymbol{\gamma}\in\mathcal{J}_{f,\mathrm{eff}}}$ of the retained modes with the extended set of data $\left\{\boldsymbol{\xi}^{(q)}, u^{(q)}\right\}_{q=1}^{N_q+\widehat{N_q}}$. It finally yields $\widehat{u}(\boldsymbol{\xi})$ expressed in the basis $\left\{\widehat{f}_{\boldsymbol{\gamma}}\right\}_{\boldsymbol{\gamma}\in\mathcal{J}_{f,\mathrm{eff}}}$.

4.6 Robust Estimation

An important concern when deriving a methodology is the robustness with respect to noise and a more robust alternative to the methodology discussed so far is now presented.

To evaluate the approximation coefficients once an approximation basis is determined from the subset selection step, a standard approach is to minimize a norm between target observations and reconstructed approximation as done in the previous section, Eqs. (22) and (24): the approximation coefficients of a given mode $\widehat{f}_{\boldsymbol{\gamma}}$ are basically given by $\boldsymbol{c}_{\boldsymbol{\gamma}} = \arg\min_{\widetilde{\boldsymbol{c}}\in\mathbb{R}^{|\mathcal{J}_{\boldsymbol{\gamma}}|}} \|\boldsymbol{z} - \Psi\,\widetilde{\boldsymbol{c}}\|_2$, with $\Psi \in \mathbb{R}^{N_q\times|\mathcal{J}_{\boldsymbol{\gamma}}|}$ the matrix of the $\boldsymbol{\gamma}$-group predictors evaluated in $\left\{\boldsymbol{\xi}^{(q)}\right\}$ and $\boldsymbol{z}$ the target residual vector. The solution to this least squares problem is equivalently obtained from

$$\{\boldsymbol{c}, \Delta\boldsymbol{z}\} = \underset{\widetilde{\boldsymbol{c}}\in\mathbb{R}^{|\mathcal{J}_{\boldsymbol{\gamma}}|}}{\arg\min} \left\|\widetilde{\Delta\boldsymbol{z}}\right\|_F \quad \text{s.t.} \quad \boldsymbol{z} + \widetilde{\Delta\boldsymbol{z}} = \Psi\,\widetilde{\boldsymbol{c}}, \tag{27}$$

which minimizes the Frobenius norm of the residual vector. This implicitly assumes no error in the coordinates $\left\{\boldsymbol{\xi}^{(q)}\right\}$ at which the target is evaluated. For instance, these coordinates may be known as the solution of auxiliary inference problems. This brings errors so that the actual coordinates vector is only estimated with an error $\Delta\boldsymbol{\xi}^{(q)}$. Since Ψ depends on $\boldsymbol{\xi}$, an error predictor matrix $\Delta\Psi(\boldsymbol{\xi}, \Delta\boldsymbol{\xi}) := \Psi(\boldsymbol{\xi}+\Delta\boldsymbol{\xi}) - \Psi(\boldsymbol{\xi})$ arises and the estimation problem (27) then rewrites as a total least squares problem [43]:

$$\{\boldsymbol{c}, \Delta\Psi, \Delta\boldsymbol{z}\} = \underset{\widetilde{\boldsymbol{c}}\in\mathbb{R}^{|\mathcal{J}_{\boldsymbol{\gamma}}|}}{\arg\min} \left\|\widetilde{\Delta\Psi}\,\widetilde{\Delta\boldsymbol{z}}\right\|_F \quad \text{s.t.} \quad \boldsymbol{z} + \widetilde{\Delta\boldsymbol{z}} = \left(\Psi + \widetilde{\Delta\Psi}\right)\,\widetilde{\boldsymbol{c}}. \tag{28}$$

The realizations of the error in the data $\left\{\Delta\boldsymbol{\xi}^{(q)}(\theta), \Delta z^{(q)}(\theta)\right\}$ are modeled to follow the distribution of zero-mean *iid* variables. Further, predictors may be correlated:

$$\underset{\theta}{\mathscr{E}}\left[\left(\Delta\psi_{\boldsymbol{\alpha}} - \underset{\theta}{\mathscr{E}}\left[\Delta\psi_{\boldsymbol{\alpha}}\right]\right)\left(\Delta\psi_{\boldsymbol{\alpha}'} - \underset{\theta}{\mathscr{E}}\left[\Delta\psi_{\boldsymbol{\alpha}'}\right]\right)\right] \neq 0, \tag{29}$$

with $\Delta\psi_{\boldsymbol{\alpha}} = \Delta\psi_{\boldsymbol{\alpha}}\left(\boldsymbol{\xi}^{(q)}, \Delta\boldsymbol{\xi}^{(q)}\right)$. A general approach to solve the weighted total least squares (wTLS) problem of Eq. (28) consists of the minimization of the usual weighted residual sum of squares ρ^2 [44]:

$$\rho^2 := \mathrm{vec}\,(\Delta X)^T\ \Lambda^{-1}\ \mathrm{vec}\,(\Delta X)\,, \qquad \Delta X := (\Delta\Psi\ \Delta\boldsymbol{z})^T\,, \qquad X := (\Psi\ \boldsymbol{z})^T\,, \tag{30}$$

where the 'vec' operator unfolds a generic $m \times n$ matrix into a mn vector and X is the data matrix. The covariance matrix for X, $\Lambda := \left\langle \mathrm{vec}\left(X - \langle X\rangle_{N_q}\right)\ \mathrm{vec}\left(X - \langle X\rangle_{N_q}\right)^T \right\rangle_{N_q}$, is evaluated and the minimization problem (28) is solved using the ALS-based algorithm proposed in [45].

As will be shown in the numerical experiments examples, Section 5.1.4, the present total LS formulation allows one to improve the approximation quality from noisy data.

Remark 2. *When a large amount N_q of experimental information is available, the data matrix $X \in \mathbb{R}^{(|\mathcal{J}_\Upsilon|+1)\times N_q}$ can be large. The resulting correlation matrix Λ then has potentially very large dimensions. However, since the noise is assumed independent from one sample to another, Λ has a block diagonal structure. Further, it is a symmetric definite positive matrix, allowing for additional reduction of the storage requirement. The structure of Λ is then exploited in solving the weighted total least squares problem above through sparse storage and operations.*

4.7 Asymptotic Numerical Complexity

While the primary motivation for this work is to determine an accurate representation of a random quantity from a small set of its realizations, it is desirable that the solution method remains computationally tractable. As seen above, the algorithm for approximating a random variable is essentially two fold.

The selection process essentially consists in sequentially building a subset, Section 4.3. Each step of the sequence involves solving a LS problem of growing size and finding the basis function, or group of functions, within the *a priori* set $\mathcal{J}_{\mathrm{prior}}$ most correlated with the current residual. The matrix of the LS problem is $\Psi \in \mathbb{R}^{N_q\times|\mathcal{J}_{\mathrm{post}}|}$, with $|\mathcal{J}_{\mathrm{post}}|$ the cardinality of the current set of selected basis functions. The LS problem is solved via a QR decomposition of Ψ in $\mathcal{O}\left(N_q\,|\mathcal{J}_{\mathrm{post}}|^2\right)$ operations. The iterative selection process is carried out with a growing active set $\mathcal{J}_{\mathrm{post}}$ until the problem becomes ill-posed, i.e., until $|\mathcal{J}_{\mathrm{post}}|$ is about N_q. We use grouped LARS and denote $\overline{|\mathcal{J}_{\boldsymbol{\gamma},\mathrm{post}}|}$ the average cardinality of the retained group predictors, i.e., the average number of basis functions in the group added to the active set. The subset selection process retains n_f groups of variables so that the total cost associated with the LS step of the subset selection is

$$\mathscr{J}_{\mathrm{LS}} = \mathcal{O}\left(\sum_{s=1}^{n_f} N_q\ \left(\overline{|\mathcal{J}_{\boldsymbol{\gamma},\mathrm{post}}|}\,s\right)^2\right). \tag{31}$$

As groups of predictors are moved to the active set, the size of the remaining *a priori* set decreases, $|\mathcal{J}_{\mathrm{prior}}|^{(\mathrm{current})} \simeq |\mathcal{J}_{\mathrm{prior}}| - s\,\overline{|\mathcal{J}_{\boldsymbol{\gamma},\mathrm{post}}|}$. The cost associated with the evaluation of the correlation for each predictor in the inactive set is then

$$\mathscr{J}_{\mathrm{correl}} = \mathcal{O}\left(\sum_{s=1}^{n_f} N_q\ \left(|\mathcal{J}_{\mathrm{prior}}| - s\,\overline{|\mathcal{J}_{\boldsymbol{\gamma},\mathrm{post}}|}\right)\right) \propto N_q. \tag{32}$$

In practice, the cost associated with the evaluation of the correlation of the predictors in the inactive set with the current residual dominates so that the whole cost of the subset selection finally approximates as

$$\mathscr{J}_{\mathrm{subsel}} = \mathscr{J}_{\mathrm{LS}} + \mathscr{J}_{\mathrm{correl}} \simeq \mathcal{O}\left(N_q\,|\mathcal{J}_{\mathrm{prior}}|\,n_f - N_q\,\overline{|\mathcal{J}_{\boldsymbol{\gamma},\mathrm{post}}|}\,\frac{n_f\ (n_f+1)}{2}\right). \tag{33}$$

The second step of the solution method deals with the evaluation of the approximation coefficients, Sections 4.4–4.5. The cost associated with evaluating the coefficients of an lth interaction order mode, $1 \le l \le N_l^{(\mathrm{PC})}$, encompasses the matrix Ψ assembly cost $\mathcal{O}\left(N_q\,p!/\left(l!\ (p-l)!\right)\right)$ and the LS solution $\mathcal{O}\left(N_q\ \left(p!/\left(l!\ (p-l)!\right)\right)^2\right)$. Since modes

$\left\{\widehat{f}_{\boldsymbol{\gamma}}\right\}_{\boldsymbol{\gamma}\in\mathcal{J}_{f,\mathrm{eff}}}$ already evaluated may be updated once an additional one from the selected set is considered, the total cost is the sum of an arithmetic sequence. Its exact formulation depends on the selected set and is difficult to derive in closed-form. As a simple example, updating all coefficients for each new mode $\widehat{f}_{\boldsymbol{\gamma}}$ considered, neglecting the cost associated with first-order interaction modes and assuming only second-order interaction modes are retained in the *a posteriori* set, an upper bound for the cost writes

$$\mathscr{J}_{\mathrm{coef}} \leq \sum_{s=1}^{|\mathcal{J}_{f,\mathrm{eff}}|} \left[\mathcal{O}\left(s\, N_q\, p^l\right) + \mathcal{O}\left(s\, N_q\, p^{2l}\right)\right], \qquad \text{with } l=2, \tag{34}$$

where $|\mathcal{J}_{f,\mathrm{eff}}|$ is the number of groups finally retained for the approximation by the CV test, see Algorithm 1. A quantitative discussion of the numerical cost is given in Section 5.1.5 with an illustrative example.

4.8 Approximation of a Random Process by a Separated Representation

The approximation of a random process, say, a space-dependent uncertain quantity $u(\boldsymbol{x},\boldsymbol{\xi})$ is now considered in the form of separation of variables:

$$u(\boldsymbol{x},\boldsymbol{\xi}) \approx w_0(\boldsymbol{x}) + \sum_{n=1}^{N} w_n(\boldsymbol{x})\,\lambda_n(\boldsymbol{\xi}) \equiv \sum_{n=0}^{N} w_n(\boldsymbol{x})\,\lambda_n(\boldsymbol{\xi}), \qquad \lambda_0 \equiv 1. \tag{35}$$

The 'spatial' modes are associated with all physical dimensions the random process may be indexed upon (space, time, ...) so that $\boldsymbol{x} = (x_1\, x_2 \ldots t \ldots) \subseteq \mathbb{R}^{d_{\boldsymbol{x}}}$. They are defined as: $w_n(\boldsymbol{x}) = \sum_{l=1}^{|\mathcal{J}_x|} c_{l,n}^{(w)}\,\phi_l(\boldsymbol{x})$ with $\{\phi_l\}$ a chosen truncated basis of cardinality $|\mathcal{J}_x|$. The functional form of "stochastic" modes $\{\lambda_n\}$ and their evaluation was discussed in Sections 4.2–4.5.

The spatial and stochastic modes of the approximation (35) are sequentially determined in turn. Let $\|v\|_{N_q} := \langle v,v\rangle_{N_q}$ be the norm induced by the data-driven inner product: $\langle\cdot,\cdot\rangle_{N_q} : \mathbb{R}\times\mathbb{R}\to\mathbb{R}, (v,w)\mapsto \langle v,w\rangle_{N_q} := \sum_{q=1}^{N_q} v^{(q)}\,w^{(q)}$. Assuming $\{\lambda_n\}$ known and projecting Eq. (35) onto the space spanned by $\{\phi_l\}$, the coefficients $\left\{c_{l,n}^{(w)}\right\}_l$ of the deterministic mode w_n are the solution of the following problem:

$$\begin{aligned} &\langle u, \phi_k\,\lambda_n\rangle_{N_q} = \left\langle \sum_{n'=0}^{n-1} w_{n'}\,\lambda_{n'} + \sum_{l=1}^{|\mathcal{J}_x|} c_{l,n}^{(w)}\,\phi_l\,\lambda_n, \phi_k\,\lambda_n \right\rangle_{N_q}, \qquad \forall\, 1\leq k\leq |\mathcal{J}_x|, \\ \Leftrightarrow \quad &\boldsymbol{c}_{\cdot,n}^{(w)} = \underset{\widetilde{\boldsymbol{c}}\in\mathbb{R}^{|\mathcal{J}_x|}}{\arg\min} \left\| \boldsymbol{u} - \sum_{n'=0}^{n-1} \boldsymbol{w}_{n'}\odot\boldsymbol{\lambda}_{n'} - (\Phi\,\widetilde{\boldsymbol{c}})\odot\boldsymbol{\lambda}_n \right\|_2, \end{aligned} \tag{36}$$

where $\Phi\in\mathbb{R}^{N_q\times|\mathcal{J}_x|}$, $\Phi_{ql} = \phi_l\left(\boldsymbol{x}^{(q)}\right)$, $\boldsymbol{w}_n = \left(w_n\left(x^{(1)}\right)\ \ldots\ w_n\left(x^{(N_q)}\right)\right)^T$, $\boldsymbol{\lambda}_n = \left(\lambda_n\left(\boldsymbol{\xi}^{(1)}\right)\ \ldots\ \lambda_n\left(\boldsymbol{\xi}^{(N_q)}\right)\right)^T$ and $\odot$ is the Hadamard product. Similarly, the stochastic mode λ_n is evaluated by determining the set of coefficients $\left\{c_n^{(\lambda)}\right\}$ minimizing $\left\|\boldsymbol{u} - \sum_{n'=0}^{n-1}\boldsymbol{w}_{n'}\odot\boldsymbol{\lambda}_{n'} - \boldsymbol{w}_n\odot\boldsymbol{\lambda}_n\left(\left\{c_n^{(\lambda)}\right\}\right)\right\|_2$ using Algorithm 1 presented in Section 4.5. The spatial mode w_n is then evaluated from Eq. (36) given all the other information and the whole iteration is repeated until convergence of the pair $\{w_n,\lambda_n\}$. The next pair can then be determined with the same methodology with $n\leftarrow n+1$. The algorithm is summarized in Algorithm 2.

Remark 3. *If the separated approximation grows beyond a few modes, it is beneficial to update the coefficients of, say, the spatial modes for improved accuracy: solve for $\{w_0,\ldots,w_n\}$ given $\{\boldsymbol{u},\lambda_0,\ldots,\lambda_n\}$.*

Algorithm 2: Sketch of the solution method for approximating a random process

Choose $|\mathcal{J}_x|$. Initialize $\boldsymbol{z} = \boldsymbol{u}$ and $\boldsymbol{\lambda}_0 = \mathbf{1}$ and set $n \leftarrow 0$.
repeat
  while $\|\lambda_n\|_{N_q}$ *not converged* **do**
    Solve for coefficients $\left\{c_{l,n}^{(w)}\right\}_{l=1}^{|\mathcal{J}_x|}$ **of the deterministic mode** given $\boldsymbol{\lambda}_n$ and $\boldsymbol{z}$ using Eq. (36) and normalize $w_n(\boldsymbol{x}) = \sum_l c_{l,n}^{(w)}\, \phi_l(\boldsymbol{x})$, $\|w_n\|_{N_q} = 1$.
    Solve for the stochastic mode $\lambda_n(\boldsymbol{\xi})$ using Algorithm 1 given $\boldsymbol{w}_n$ and $\boldsymbol{z}$. If $n = 0$, $\boldsymbol{\lambda}_0 \leftarrow \mathbf{1}$.
  Set $\boldsymbol{z} \leftarrow \boldsymbol{z} - \boldsymbol{w}_n \odot \boldsymbol{\lambda}_n$, and $n \leftarrow n + 1$.
until *a termination criterion is met (for instance,* $\|\lambda_n\|_{N_q}$ *below a threshold or maximum rank* $n > N$ *reached).*

5. NUMERICAL EXPERIMENTS

The methodology developed in the previous sections is now demonstrated on a set of examples. Different aspects of the global solution method are illustrated on a 1D stochastic diffusion equation. A more computationally involved example is next considered with a shallow water problem with multiple sources of uncertainty.

5.1 Stochastic Diffusion Equation

We consider a steady-state stochastic diffusion equation on $\Omega \times \Xi$, $\Omega = [x_-, x_+] \subset \mathbb{R}$, with deterministic Dirichlet boundary conditions:

$$\nabla_x \left(\nu\left(x, \boldsymbol{\xi}'\right) \nabla_x u\left(x, \boldsymbol{\xi}\right)\right) = F\left(x, \boldsymbol{\xi}''\right), \qquad u\left(x_-, \boldsymbol{\xi}\right) = u_-, \quad u\left(x_+, \boldsymbol{\xi}\right) = u_+. \tag{37}$$

The right-hand side F is a random source field and ν is a space-dependent random diffusion coefficient modeled as

$$\begin{aligned} \nu\left(x, \boldsymbol{\xi}'\right) &= \nu_0(x) + \sum_{k=1}^{d_\nu} \sqrt{\sigma_{\nu,k}}\, \omega_{\nu,k}(x)\, \xi_k', \qquad \boldsymbol{\xi}' = \left(\xi_1' \ \ldots \ \xi_{d_\nu}'\right), \\ F\left(x, \boldsymbol{\xi}''\right) &= F_0(x) + \sum_{k=1}^{d_F} \sqrt{\sigma_{F,k}}\, \omega_{F,k}(x)\, \xi_k'', \qquad \boldsymbol{\xi}'' = \left(\xi_1'' \ \ldots \ \xi_{d_F}''\right), \end{aligned} \tag{38}$$

with $\nu_0(x) = 1$ and $F_0(x) = -1$ the respective mean values. The random variables $\left\{\xi_1', \ldots, \xi_{d_\nu}', \xi_1'', \ldots, \xi_{d_F}''\right\}$ are chosen mutually independent and uniformly distributed on $[0, 1]$. The spatial modes $\omega_{\nu,k}(x)$ and $\omega_{F,k}(x)$, and their associated amplitude $\sqrt{\sigma_{\nu,k}}$ and $\sqrt{\sigma_{F,k}}$, are the first dominant eigenfunctions of the following eigenproblems:

$$\begin{aligned} \int_\Omega K_\nu(x, x')\, \omega_{\nu,k}(x')\, \mathrm{d}x' &= \sigma_{\nu,k}\, \omega_{\nu,k}(x), \qquad K_\nu(x, x') = \sigma_\nu^2\, e^{-[(x-x')^2/2\, L_{c,\nu}^2]}, \\ \int_\Omega K_F(x, x')\, \omega_{F,k}(x')\, \mathrm{d}x' &= \sigma_{F,k}\, \omega_{F,k}(x), \qquad K_F(x, x') = \sigma_F^2\, e^{-[(x-x')^2/2\, L_{c,F}^2]}, \end{aligned} \tag{39}$$

with K_ν and K_F the correlation kernels. The random fields properties are chosen as $\sigma_\nu = 0.7$, $\sigma_F = 0.7$, $L_{c,\nu} = 0.3$, $L_{F,\nu} = 0.3$. Note that $F\left(\cdot, \boldsymbol{\xi}''\right) < 0$ a.e. and $\nu\left(\cdot, \boldsymbol{\xi}'\right) > 0$ a.e. so that the problem remains coercive. The spectra of the operators associated with these eigenproblems are here the same and decay quickly thanks to the high correlation length as can be appreciated from Table 1 where the dominant eigenvalues are given. The resulting problem is then anisotropic in Ξ in the sense that the degree of dependence of the input random parameters along the dimensions $\{\xi_1, \ldots, \xi_8\}$ strongly varies.

TABLE 1: Upper part of the spectrum of both eigenproblems (39)

σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7	σ_8
0.1815	0.1396	0.0906	0.0450	0.0236	0.0097	0.0035	0.0011

Denoting $\boldsymbol{\xi} = \left(\boldsymbol{\xi}' \, \boldsymbol{\xi}''\right) \in \mathbb{R}^d$, $d = d_\nu + d_F$, the solution u is approximated in a rank-N separated form: $u\left(x, \boldsymbol{\xi}\right) \approx \hat{u}\left(x, \boldsymbol{\xi}\right) = w_0\left(x\right) + \sum_{n=1}^{N} w_n\left(x\right) \lambda_n\left(\boldsymbol{\xi}\right)$. The stochastic approximation basis relies on an HDMR format with a maximum interaction order $N_l = 3$ and 1D Legendre polynomials $\{\psi_\alpha\}_{\alpha=1}^{p}$ of maximum degree $p = 8$.

In this section, the focus is on approximating a purely random quantity, i.e., disregarding its spatial dependence. We then rely on samples of the solution $u\left(x, \boldsymbol{\xi}\right)$ taken at a given spatial location $x^\star$: $\left\{u^{(q)} := u\left(x^\star, \boldsymbol{\xi}^{(q)}\right)\right\}_{q=1}^{N_q}$.

5.1.1 Influence of the Number of Samples

We first focus on the achieved accuracy in the approximation with a given budget $N_q + \widehat{N_q}$ samples. The number of test points $\widetilde{N_q}$ to estimate the approximation error ε, Eq. (26), is chosen sufficiently large so that ε is well estimated, $\widetilde{N_q} = 10,000$. In Fig. 2, the performance of the present gLARS-ALS methodology is compared with both a plain HDMR approximation, i.e., with no subset selection hence considering the whole *a priori* approximation basis, and a PC approximation with a sparse grid technique. The Smolyak scheme associated with a Gauss-Patterson quadrature rule is used as the sparse grid, with varying number of points in the 1D quadrature rule and varying levels. The dimensionality of the stochastic space is $d = 8$.

The sparse grid is seen to require a large number of samples to reach a given approximation accuracy.[6] The HDMR-format approximation, with various interaction orders N_l, provides a better performance than PC/Smolyak but

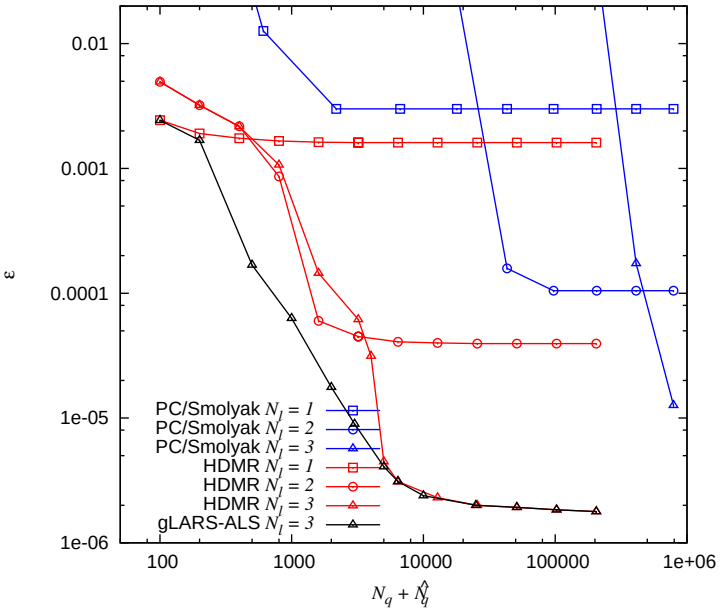

FIG. 2: Convergence of the approximation with the number of samples $N_q + \widehat{N_q}$. Different approximation methods are compared: plain HDMR, PC/Smolyak scheme sparse grid spectral decomposition, and the present gLARS-ALS. The convergence is plotted in terms of ε. $d = 8$, $p = 8$, $N_l = 3$, $N_l^{(\mathrm{PC})} = 3$.

[6]Note that the plain Smolyak scheme is used here, which does not exploit anisotropy in the response surface. More sophisticated Smolyak-based approximations have been developed, see [7], and are expected to provide better results.

still requires more points to reach a given accuracy than the present gLARS-ALS method which performs significantly better in approximating the QoI from a given dataset. The gLARS-ALS approximation error is also seen to be smooth and monotonic when the amount of information varies. When N_q is large enough, the subset selection step becomes useless as all $|\mathcal{J}_{\text{prior}}|$ terms of the *a priori* basis can be evaluated from the large amount of information and the gLARS-ALS performance is then similar to that of the HDMR. Note that the benefit of a subset selection step in terms of accuracy improvement increases with the dimension d as the size $|\mathcal{J}_{\text{prior}}|$ of the potential dictionary then grows.

5.1.2 Influence of the Stochastic Dimension

The approximation accuracy of the present method is now studied when the dimension of the stochastic space varies. The same problem as above is considered but with various truncation orders of the source F and the diffusion coefficient ν definitions, see Eq. (38). The solution of the diffusion problem (37) is of dimension $d = d_F + d_\nu$ and the dimensions d_F and d_ν are varied together, $d_F = d_\nu$. The resulting approximation error is plotted in Fig. 3 for different d when the number of available samples varies. From $d = 8$ to $d = 40$, the required number of points for a given accuracy is seen to increase significantly, between a 2- and a 10-fold factor. However this is much milder than the increase in the potential approximation basis cardinality, i.e., if not subset selection was done, as $|\mathcal{J}_{\text{prior}}|$ shifts from $10,565$ $(d = 8)$ to 1.7×10^6 $(d = 40)$, demonstrating the efficiency of the subset selection step which activates only a small fraction of the dictionary. When d further increases from 40 to 100 for a given N_q, the performance remains essentially the same with hardly any loss of accuracy: T he solution method is able to capture the low-dimensional manifold onto which the solution essentially lies and an increase in the size of the solution space hardly affects the number of samples it requires. This capability is a crucial feature when available data are scarce and the solution space is very large. As an illustration, when $d = 100$, and with the parameters retained, the potential cardinality of the approximation basis is about 27×10^6 while the number of available samples is $\mathcal{O}(100 - 10,000)$. It clearly illustrates the pivotal importance of the subset selection step. Note that if one substitutes a PC approximation to the present HDMR format, about 352×10^9 terms need be evaluated with the present settings, a clearly daunting task.

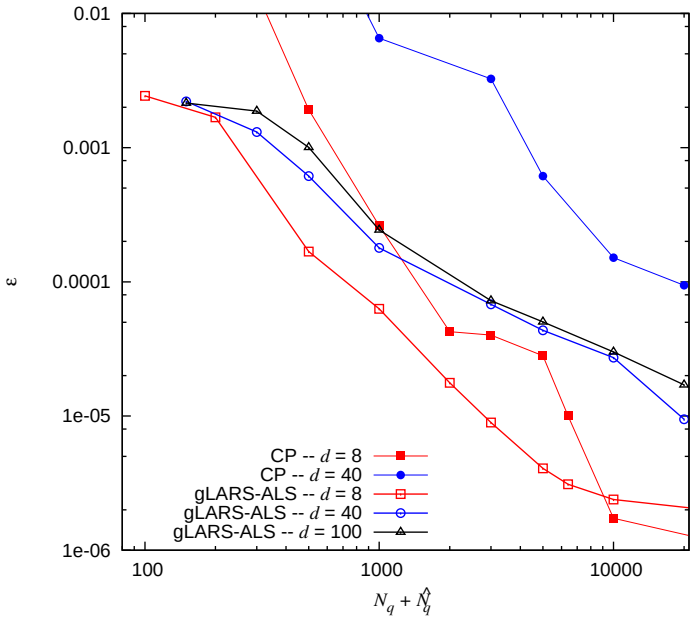

FIG. 3: Convergence of the approximation with the number of samples $N_q + \widehat{N_q}$ and for different dimensionality of the QoI. The present gLARS-ALS approach is compared with a CANDECOMP-PARAFAC-type technique (labeled "CP").

For sake of completeness, the approximation given by a CP format, Eq. (12), is also considered. The univariate functions $\{f_{i,r}\}$ are approximated with the same polynomial approximation as in the present gLARS-ALS approach and a Tikhonov-based regularized ALS technique is used to determine each $f_{i,r}$ in turn given the others. Upon convergence, the next set of modes $\{f_{1,r+1}, \ldots, f_{d,r+1}\}$ is evaluated until a maximum rank n_r set by cross validation. At each rank r, the best approximation, as estimated by cross validation, is retained from a set of initial conditions and regularization parameter values. As can be appreciated from Fig. 3, the number of samples required for a given approximation error is significantly larger than with the present gLARS-HDMR method.

5.1.3 Subset Selection

To further illustrate the subset selection step, the set of second- and third-order interaction retained modes $\{f_{\boldsymbol{\gamma}}\}_{\boldsymbol{\gamma}\in\mathcal{J}_{f,\mathrm{post}}}$ are plotted in Fig. 4 in the $d = 40$ case. Each bullet represents one of the d stochastic dimensions and each line connects two (2nd order, left plot) or three (3rd order, right plot) dimensions, denoting a retained mode. The first $d_F = 20$ of the 40 dimensions are associated with the source term F in the stochastic equation and are represented as the solid bullets of the first two quadrants, $d \in [1, 20]$. The other $d_\nu = 20$ dimensions are associated with the uncertain diffusion coefficient ν and are plotted as open bullets in the third and fourth quadrants, $d \in [21, 40]$. The dimensions introduced by these two quantities are sorted with the associated magnitude of the eigenvalues σ_F and σ_ν of their kernel, see Eqs. (39), which decreases along the counterclockwise direction. Hence, the norm of the eigenvalues of the kernel associated with F decreases when one goes counterclockwise from the first to the second quadrant. Likewise, the norm of the eigenvalues associated with dimensions introduced by ν decreases from the third to the fourth quadrant. Dominant dimensions of the stochastic space for the output u approximation are thus expected to lie at the beginning of the first and/or third quadrant.

From the plot of second-order modes (left), the subset selection process is seen to retain interaction modes mainly associated with dominant eigenvalues of both F and ν: they mainly link bullets from the first (dominant) dimensions associated with F to the first (dominant) dimensions associated with ν, as one might expect. Further, modes associated with two dimensions both introduced by ν are seen to be selected while two dimensions both associated with F are

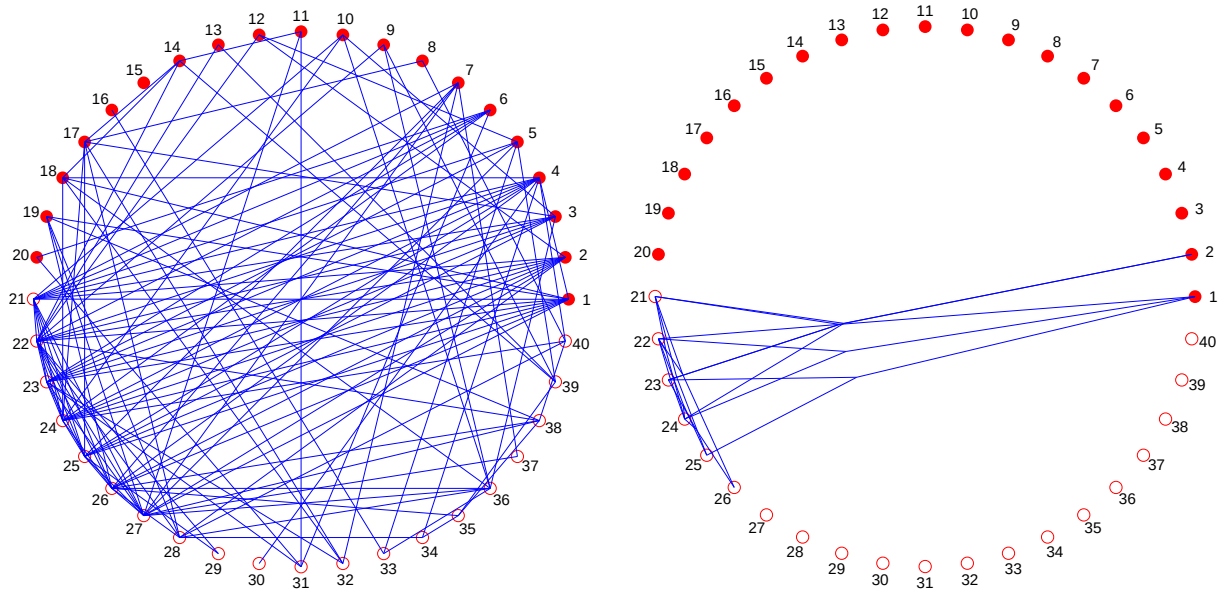

FIG. 4: Graphical representation of the interaction modes retained by the subset selection procedure. Left: second-order modes are plotted as a line linking two dimensions (bullets). Right: third-order modes are represented as 3-branch stars and connect three dimensions.

rarely connected: the subset selection procedure is able to capture the nonlinearity associated with ν in the QoI and retains corresponding interaction modes. Indeed, note from Eq. (37) that the source term F interacts linearly with the solution u while the diffusion coefficient is nonlinearly coupled with u and hence, interaction modes between two dimensions introduced by F do not contribute to the approximation. The third-order modes (right plot) also illustrate the nonlinearity associated with ν: The retained modes either connect dimensions associated with ν only or with one F-related and two ν-related dimensions. Again, no two dimensions of F are connected, consistently with the linear dependence of u with F. These results illustrate the effectiveness of the procedure to unveil the dominant dependence structure and to discard unnecessary approximation basis functions.

5.1.4 Robustness

The robustness of the approximation against measurement noise is now investigated. The dataset is corrupted with noise. Denoting the nominal value with a star as superscript, noise in the coordinates is modeled as

$$\boldsymbol{\xi}^{(q)} = \boldsymbol{\xi}^{(q)^\star} + s\,\boldsymbol{\zeta}^{(q)}, \qquad \forall\, 1 \le q \le N_q, \qquad s > 0. \tag{40}$$

The noise is modeled as an additive d-dimensional, zero-centered, unit variance, Gaussian random vector $\boldsymbol{\zeta}$ biased so that $\boldsymbol{\xi}^{(q)} \in [-1,1]^d, \forall\, q$. It is independent from one sample q to another. Without loss of generality, measurements are here modeled as being corrupted with a multiplicative noise: $u^{(q)} = u^{(q)^\star}\left(1 + s_u\,\zeta_u^{(q)}\right)$, with $s_u = 0.2$ and $\zeta_u \sim \mathcal{N}(0,1)$.

The evolution of the approximation accuracy when the noise intensity s in the coordinates varies is plotted in Fig. 5 in terms of error estimation ε. We compare gLARS-ALS using standard least squares (LS) with its "robust" counterpart relying on weighted total least squares (wTLS) as discussed in Section 4.6.

When the noise intensity increases, the error exponentially increases, quickly deteriorating the quality of the approximation with a noise standard deviation here as low as $s = 3 \times 10^{-5}$. When the noise is strong (low signal-to-noiseratio), both the LS and the wTLS methods achieve poor accuracy. However, if the dataset is only mildly

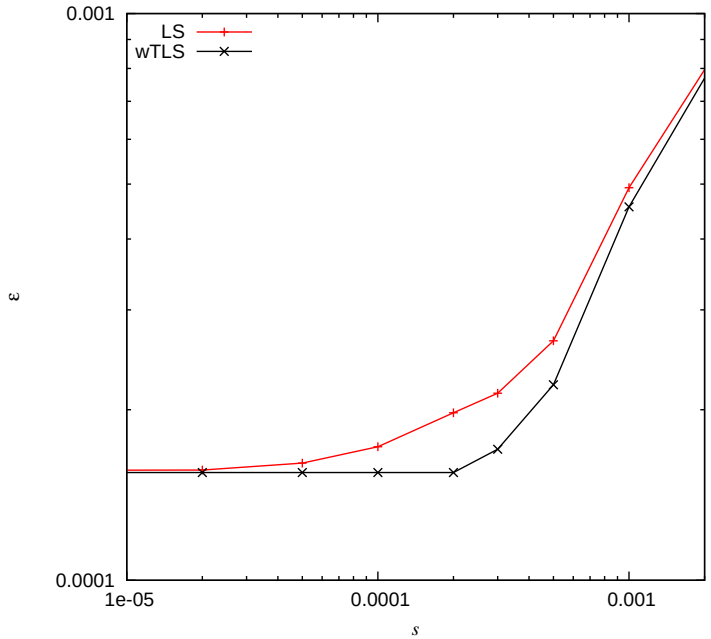

FIG. 5: Robustness of the approximation with respect to noise in the data: approximation error ε from the standard least squares (LS) and weighted total least squares (wTLS). $d = 5$, $N_q = 500$.

corrupted with noise, the wTLS approach is seen to achieve a significantly better accuracy than the standard LS, while the solution process is significantly slower than that using the standard LS. The present paper is based on the assumption that the critical part of the whole solution chain of determining a good approximation of the QoI is the data acquisition and that the cost of the post-processing part is not the main issue. However, while it is useful only on a range of SNR and somehow computationally costly, this feature is deemed important for a successful solution method in an experimental context where noise is naturally present.

If the noisy dataset is unbiased, possible further improvements upon the wTLS approach include lowering the complexity of the approximation model. Indeed, the well-known bias-variance tradeoff indicates that a more robust, while less accurate, approximation can be obtained when the complexity of the retained model decreases. To improve the robustness of our present approach, a natural way is hence to trade some accuracy for some additional robustness. For instance, a predictor selection within each retained groups $\left\{f_{\boldsymbol{\gamma}\in\mathcal{J}_{f,\mathrm{post}}}\right\}$ can be considered, further lowering the final number of coefficients involved in the approximation and likely improving its robustness with respect to noise in the data. This could be achieved by estimating the approximation coefficients via a *penalized* (total) LS problem as mentioned in Section 2.2.2.

5.1.5 Scaling of the Solution Algorithm

In this section, the numerical complexity associated with the different steps of the solution method is illustrated in terms of computational time. Numerical experiments are carried out with varying number of samples N_q and solution space dimensions d. When one is varying, the other remains constant. The nominal parameters are $d = 40$ (dimension of the stochastic space), $p = 6$ (maximum total order of the Legendre polynomials), $N_q = 1000$ (number of samples), $N_l = 3$ (maximum interaction order of the truncated HDMR approximation), $\widetilde{p} = 5$ (maximum total polynomial order in the subset selection step).

Numerical results are gathered in Fig. 6. The asymptotic behavior of the number n_f of required subset selection iterations as introduced in Section 4.7 might be different according to which limit is considered. For the present stochastic diffusion problem, first and second interaction order modes tend to be selected first. Assuming the active set $\mathcal{J}_{f,\mathrm{post}}$ is dominated by first and second interaction order modes, it can easily be shown that the number of retained groups then satisfies

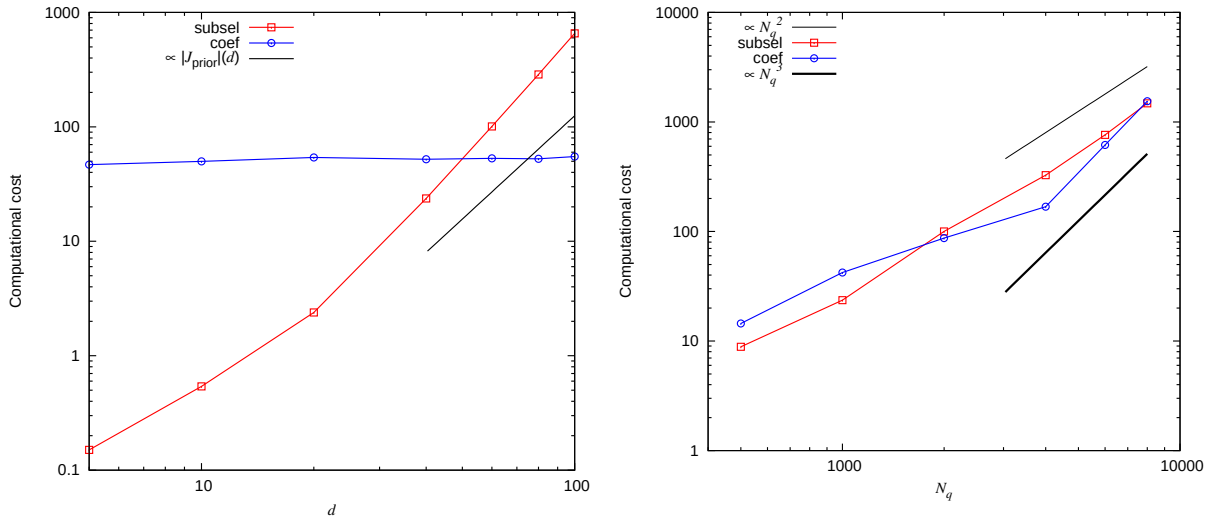

FIG. 6: Numerical cost of the subset selection and coefficients evaluation steps as a function of the stochastic dimension d and size of the dataset N_q. Approximation coefficients are fully updated for each new mode. Nominal parameters are $d = 40$, $p = 6$, $N_q = 1000$, $N_l = 3$, $\widetilde{p} = 5$, $N_l^{(\mathrm{PC})} = 3$.

$$n_f \leq 1 + n_{f_1} + \min\left[\frac{d\,(d-1)}{2}, 2\,\frac{N_q - n_{f_1}\,\widetilde{p}}{\widetilde{p}\,(\widetilde{p}+1)}\right], \qquad n_{f_1} \leq \min\left[d, \frac{N_q}{\widetilde{p}}\right]. \tag{41}$$

In the present example, second-order interaction groups dominate the retained set so that the number of retained groups tends to scale as $n_f \propto N_q/\widetilde{p}^2$. From Eq. (33) and for the present nominal parameters, it results in the following limit behavior for the subset selection step:

$$\begin{aligned}
\lim_{d\to+\infty} \mathscr{J}_{\text{subsel}} &\propto {N_q}^2\,|\mathcal{J}_{\text{prior}}|/\widetilde{p}^2 \qquad \longrightarrow \text{here}: \quad \propto {N_q}^2\,d^{N_l}\,\widetilde{p}^{N_l-2},\\
\lim_{N_q\to+\infty} \mathscr{J}_{\text{subsel}} &\propto {N_q}^2\,|\mathcal{J}_{\text{prior}}|/\widetilde{p}^2 \qquad \longrightarrow \text{here}: \quad \propto {N_q}^2\,d^{N_l}\,\widetilde{p}^{N_l-2}.
\end{aligned} \tag{42}$$

Similarly, the cost associated with the coefficients evaluation is considered. The number of interaction modes $\mathcal{J}_{f,\text{eff}}$ effectively varies between 1 and $\mathcal{O}\left(N_q/p^2\right)$ along the solution procedure, and, since the cost associated with solving the LS problem dominates that of the matrix assembly, the cost of their evaluation finally simplifies in $\mathscr{J}_{\text{coef}} \propto \mathcal{O}\left({N_q}^2\,p\right)$ or $\mathscr{J}_{\text{coef}} \propto \mathcal{O}\left({N_q}^3/p\right)$ depending on whether the coefficients are updated whenever an additional group is considered or not, see Section 4.5 and step (7) in Algorithm 1. In the present regime, the cost is found not to depend on d.

These asymptotic behaviors are consistent with the numerical experiments as can be appreciated from Fig. 6. The coefficients are here updated whenever a new mode from the selected set is considered, hence $\mathscr{J}_{\text{coef}} \propto \mathcal{O}\left({N_q}^3/p\right)$. It is seen that the subset selection step scales less favorably than the coefficients evaluation step with the dimensionality of the random variable. This stresses the benefit of a carefully chosen *a priori* approximation basis to reduce as much as possible the cardinality $|\mathcal{J}_{\text{prior}}|$.

5.2 Approximation of the Solution Random Field

We now consider the approximation of the space-dependent random solution $u\,(x, \boldsymbol{\xi})$ under the form (35) using Algorithm 2. The approximation obtained from different number of samples $\left\{x^{(q)}, \boldsymbol{\xi}^{(q)}, u^{(q)}\right\}$ is compared with the Karhunen-Loève modes, computed from a full knowledge of the QoI, hereafter referred to as the reference solution.[7]

The simulation relies on the following parameters: $|\mathcal{J}_x| = 32$, $p = 10$, $d = 6$, $N_l = 3$, $N_l^{(\text{PC})} = 3$. The potential approximation basis cardinality is about $|\mathcal{J}_x|\,|\mathcal{J}_{\text{prior}}| \simeq 10^5$. Fig. 7 shows the first and second spatial modes, $w_1\,(x)$ and $w_2\,(x)$ for different sizes of the dataset, $N_q = 1000$, 3000, 9000, and $26,000$. The mean mode $w_0\,(x)$ is virtually indistinguishable from the reference solution mean mode for any of the dataset sizes and is not plotted. On the left plot $[w_1\,(x)]$, it is seen that the approximation is decent, even with as low as $N_q = 1000$ samples. For $N_q = 3000$, the approximation is good. This $(1 + d) = 7$-dimensional case corresponds to ${N_q}^{1/(1+6)} \simeq 3.1$ samples per solution space dimension only and about $N_q/\left(|\mathcal{J}_x|\,|\mathcal{J}_{\text{prior}}|\right) \simeq 3\%$ of the potentially required information.

For approximating the second spatial mode (Fig. 7, right plot), more points are needed to reach a good accuracy but $N_q = 26,000$ is seen to already deliver a good performance. Quantitative approximation error results are gathered in Table 2 for various separation ranks N and number of samples N_q.

The satisfactory performance of the present method can be understood from the upper part of the Karhunen-Loève approximation (normalized) spectrum plotted in Table 3. The norm of the eigenvalues decays quickly so that the first two modes contribute more than 90% of the QoI L^2-norm, showing that this problem efficiently lends itself to the present separation of variables-based methodology.

[7] The spatial $\{w_n\}$ and stochastic modes $\{\boldsymbol{\lambda}_n\}$ are sequentially determined from (36) via an ALS approach. Since the decomposition is two-dimensional, $u\,(\boldsymbol{x}, \boldsymbol{\xi}) \approx \sum_{n=0}^{N} w_n\,(\boldsymbol{x})\,\boldsymbol{\lambda}_n\,(\boldsymbol{\xi})$, the approximation problem is convex, see for instance [46], and the ALS approach converges to the best rank-1 approximation of the matricized $\boldsymbol{u}$ in the Frobenius sense. If the data-driven inner product $\langle\cdot,\cdot\rangle_{N_q}$ was inducing a cross norm (it only induces a semi-norm), then $\langle w\,\boldsymbol{\lambda}, w\,\boldsymbol{\lambda}\rangle_{N_q} = \|\boldsymbol{w}\|_2^2\,\|\boldsymbol{\lambda}\|_2^2$ and the pair $(\boldsymbol{w}, \boldsymbol{\lambda})$ would be the dominant rank-1 approximation of the matricized $\boldsymbol{u}$. The Karhunen-Loève decomposition of u is thus the reference solution one should obtain in the particular case where the empirical inner product induces a cross norm and $N_q \to \infty$.

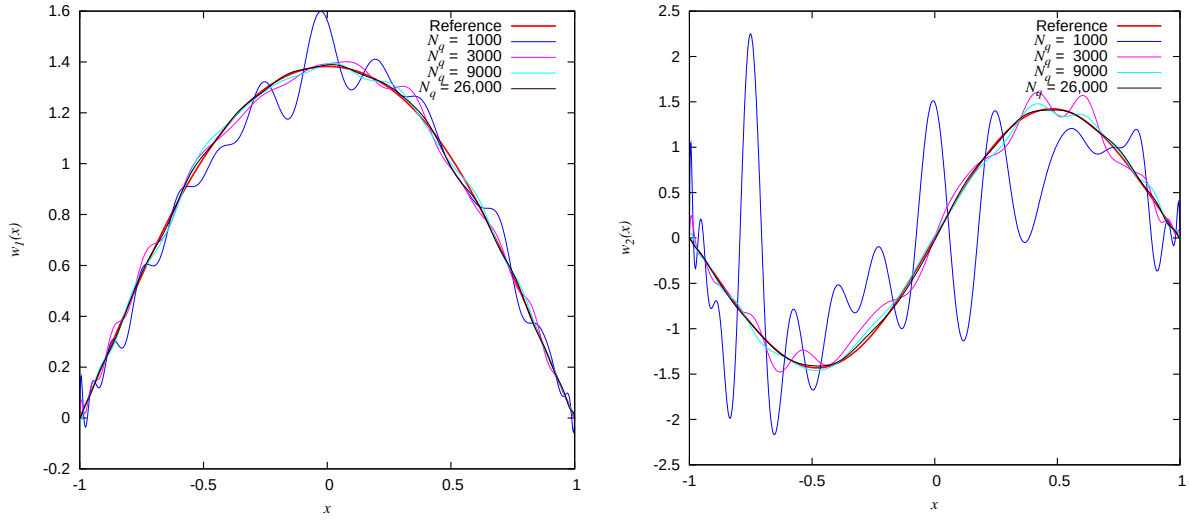

FIG. 7: First $[w_1(x)$, left] and second $[w_2(x)$, right] spatial approximation modes of the stochastic diffusion solution. The reference (Karhunen-Loève) solution is plotted for comparison (thick line).

TABLE 2: Evolution of the approximation error ε, as defined in Eq. (26), with the decomposition rank N and the number of samples N_q

$N_q \setminus N$	0	1	2
1000	5.5×10^{-3}	7.4×10^{-4}	7.4×10^{-4}
3000	5.5×10^{-3}	4.2×10^{-4}	2.7×10^{-4}
9000	5.5×10^{-3}	3.1×10^{-4}	1.0×10^{-4}
26,000	5.4×10^{-3}	2.8×10^{-4}	6.2×10^{-5}

TABLE 3: Normalized upper spectrum of the Karhunen-Loève approximation

i	1	2	3	4	5	6	7	8	9	10
σ_i	144	30.2	13.6	2.64	1.37	0.250	0.0817	0.0167	0.00891	0.00232

5.3 A Shallow Water Flow Example

The methodology is now applied to the approximation of the stochastic solution of a shallow water flow simulation with multiple sources of uncertainty. It is a simple model for the simulation of wave propagation on the ocean surface. Waves are here produced by the sudden displacement of the sea bottom at a given magnitude in time, extension, and location, all uncertain.

5.3.1 Model

The problem is governed by the following set of equations:

$$\frac{D\,v_1}{D\,t} = f_C\,v_2 - g\,\frac{\partial h}{\partial x_1} - b\,v_1 + S_{v_1}, \tag{43}$$

$$\frac{D\,v_2}{D\,t} = -f_C\,v_1 - g\,\frac{\partial h}{\partial x_2} - b\,v_2 + S_{v_2}, \tag{44}$$

$$\frac{\partial h}{\partial t} = -\frac{\partial \left(v_1 \ (H+h)\right)}{\partial x_1} - \frac{\partial \left(v_2 \ (H+h)\right)}{\partial x_2} + S_h, \tag{45}$$

where $(v_1(\boldsymbol{x}, \boldsymbol{\xi}, t) \ v_2(\boldsymbol{x}, \boldsymbol{\xi}, t))$ is the velocity vector at the surface, $\boldsymbol{x} = (x_1 \, x_2) \in \Omega \subset \mathbb{R}^2$, $h(\boldsymbol{x}, \boldsymbol{\xi}, t)$ the elevation of the surface from its position at rest, $H(\boldsymbol{x})$ the sea depth, f_C models the Coriolis force, b is the viscous drag coefficient, g the gravity constant and $S_{v_1}(\boldsymbol{x}, \boldsymbol{\xi}, t)$, $S_{v_2}(\boldsymbol{x}, \boldsymbol{\xi}, t)$, $S_h(\boldsymbol{x}, \boldsymbol{\xi}, t)$ are the source fields. Without loss of generality, the drag b and the Coriolis force f_C are neglected. No slip boundary conditions apply for the velocity. The sources are modeled as acting on h only, $S_{v_1} \equiv 0$ and $S_{v_2} \equiv 0$. S_h models the source term acting on h due to, say, an underwater seismic event. The fluid density and the free surface pressure are implicitly assumed constant. Full details on the numerical implementation of a similar problem are given in [47].

5.3.2 Sources of Uncertainty

Let $\boldsymbol{\xi} = \left(\boldsymbol{\xi}' \, \boldsymbol{\xi}''\right)$. The source S_h is uncertain and is modeled as a time-dependent, spatially distributed quantity:

$$S_h(\boldsymbol{x}, \boldsymbol{\xi}, t) = a_t\left(\boldsymbol{\xi}', t\right) \ a_\xi(\xi_1'') \ \exp\left(-\frac{\left(\boldsymbol{x} - \boldsymbol{x}_{S_h}(\xi_3'')\right)^T \left(\boldsymbol{x} - \boldsymbol{x}_{S_h}(\xi_3'')\right)}{\sigma_{S_h}(\xi_2'')^2}\right), \tag{46}$$

where $a_t\left(\boldsymbol{\xi}', t\right)$ is a given time envelope, $a_\xi(\xi_1'')$ the uncertain source magnitude, $\sigma_{S_h}(\xi_2'')$ drives the uncertain source spatial extension, and $\boldsymbol{x}_{S_h}(\xi_3'')$ is the uncertain spatial location. The time envelope $a_t\left(\boldsymbol{\xi}', t\right)$ is described with an N_a-term expansion:

$$a_t\left(\boldsymbol{\xi}', t\right) = \overline{a_t}(t) + \sum_{i=1}^{N_a} \sqrt{\lambda_i} \ \xi_i'(\theta) \ \varphi_i^{a_t}(t), \tag{47}$$

with $\boldsymbol{\xi}' = \left(\xi_1' \ \ldots \ \xi_{N_a}'\right)$ the stochastic germ associated with the uncertainty in a_t. Random variables $\{\xi_i'\}_{i=1}^{N_a}$ are *iid*, uniformly distributed. The solution of the shallow water problem then lies in a $(d = N_a + 3)$-dimensional stochastic space.

5.3.3 Approximation from an Available Database

As an illustration of the methodology, we aim at approximating the sea surface field at a fixed amount of time $t^\star$ after a seismic event. The QoI is then a random field $u(\boldsymbol{x}, \boldsymbol{\xi}) = h(\boldsymbol{x}, \boldsymbol{\xi}, t^\star)$. An accurate description of this field is of importance for emergency plans in case of a seaquake. Sea level measurements of the surface at various spatial locations from past events constitute the dataset $\left\{\boldsymbol{x}^{(q)}, \boldsymbol{\xi}^{(q)}, h\left(\boldsymbol{x}^{(q)}, \boldsymbol{\xi}^{(q)}, t^\star\right)\right\}_{q=1}^{N_q}$ used to derive an approximation of u under a separated form: $u(\boldsymbol{x}, \boldsymbol{\xi}) \approx \langle u \rangle_{N_q}(\boldsymbol{x}) + \sum_{n=1}^{N} w_n(\boldsymbol{x}) \ \lambda_n(\boldsymbol{\xi})$.

The solution method here relies on a $N_q = 37,000$-sample dataset complemented with $\widehat{N_q} = 5000$ cross-validation samples and a $\widetilde{N_q} = 5000$ set for error estimation. We consider a $N_a = 5$ expansion for the time envelope, leading to a stochastic dimension of $d = 5 + 3 = 8$. The effective number of samples per dimension is then about ${N_q}^{1/(d_{\boldsymbol{x}}+d)} \simeq 2.9$. The approximation is determined based on a $|\mathcal{J}_x| = 484$ spatial discretization DOFs (spectral elements) at the deterministic level and $p = 6$th order Legendre polynomials $\{\psi_\alpha\}$, $N_l = 3$, $N_l^{(\mathrm{PC})} = 3$, for the stochastic modes. The cardinality of this *a priori* basis is then $|\mathcal{J}_x| \, |\mathcal{J}_{\mathrm{prior}}| \simeq 770 \times 10^3 \gg N_q$, again relying on an efficient subset selection step to make the approximation problem well-posed.

The approximation error when the rank N varies is shown in Table 4. It is seen that estimating the mean spatial mode w_0 leads to a relative error of about 0.12 while adding the first (w_1, λ_1) and second (w_2, λ_2) pair drops it to about 0.05. Further adding pairs does not lower the approximation error with this dataset and more samples are needed to accurately estimate them. Spatial modes w_0 and w_1 of the separated approximation are plotted in Fig. 8 for illustration.

TABLE 4: Relative approximation error ε evolution with the decomposition rank N. $N_q = 37{,}000$

N	0	1	2	3
ε	0.117	0.056	0.046	0.044

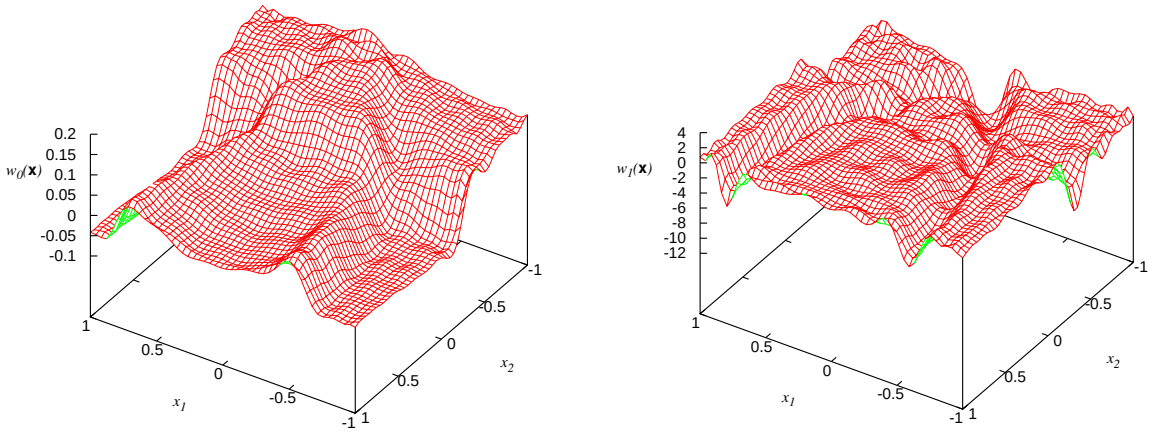

FIG. 8: Mean $[w_0(\boldsymbol{x}) \equiv \langle u \rangle_{N_q}(\boldsymbol{x})$, left] and first $[w_1(\boldsymbol{x})$, right] spatial modes.

6. CONCLUSION

In this paper, a methodology was proposed for deriving a functional representation of a random process only known through a collection of its pointwise evaluations. The proposed method essentially relies on an efficient determination of an approximation basis consistent with the available information. This involves the choice of an *a priori* canonical HDMR format combined with tuning the basis via a data-driven subset selection step. This subset selection is carried out in a bottom-to-top manner, as opposed to a top-to-bottom manner as done in the compressed sensing standard framework. It essentially sorts the HDMR modes (groups of predictors) by their contribution in approximating the quantity of interest. The final approximation can rely on a different functional description of the modes, typically of higher order and/or nonlinear in the coefficients.

The method is progressive, data-driven, and was shown to here outperform current approximation techniques in terms of accuracy for a given number of samples. Its efficiency was demonstrated in two examples which have shown its ability to achieve a good approximation accuracy from a small dataset, as long as the quantity at hand is essentially lying on a low-dimensional manifold. In particular, the dominant dimensions are naturally revealed so that all the available information can be dedicated to approximate relevant dependences only. Through a total least squares approach, it was also shown that some robustness can be achieved, an important feature if the dataset comes from experiments. Using a robust approximation was shown to bring up to a twofold improvement upon the approximation error using standard least squares, but at the price of a computational overhead. The global solution method scales reasonably well, exhibiting a linear dependence with the cardinality of the *a priori* basis dictionary and a quadratic or cubic dependence with the number of samples, depending on the coefficients update strategy.

The present work was focused on a general methodology, disregarding fine-tuning aspects. Among other things, a natural improvement would be to carry out a predictor selection within each retained groups $\left\{f_{\boldsymbol{\gamma}\in\mathcal{J}_{f,\text{post}}}\right\}$, further lowering the number of coefficients involved in the approximation. Moreover, the tensor structure of the Hilbert stochastic space can be exploited and developments towards a data-driven multilinear algebra effective tool for high-dimensional uncertainty quantification are currently carried out.

ACKNOWLEDGMENTS

The author gratefully acknowledges Tarek El Moselhy and Faidra Stavropoulou for stimulating discussions and useful comments. This work is part of the TYCHE project (ANR-2010-BLAN-0904) supported by the French Research National Agency (ANR).

REFERENCES

1. Nouy, A., A generalized spectral decomposition technique to solve a class of linear stochastic partial differential equations, *Comput. Methods Appl. Mech. Eng.*, 196(45-48):4521–4537, 2007.
2. Nouy, A., A priori model reduction through proper generalized decomposition for solving time-dependent partial differential equations, *Comput. Methods Appl. Mech. Eng.*, 199:1603–1626, 2010.
3. Nouy, A., Proper generalized decompositions and separated representations for the numerical solution of high dimensional stochastic problems, *Arch. Comput. Methods Eng.*, 17:403–434, 2010.
4. Matthies, H. and Zander, E., Solving stochastic systems with low-rank tensor compression, *Linear Algebra Appl.*, 436(10):3819–3838, 2012.
5. Novak, E. and Ritter, K., Simple cubature formulas with high polynomial exactness, *Construct. Approx.*, 15:499–522, 1999.
6. Xiu, D. and Hesthaven, J., High-order collocation methods for differential equations with random inputs, *SIAM J. Sci. Comput.*, 27:1118–1139, 2005.
7. Nobile, F., Tempone, R., and Webster, C., An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data, *SIAM J. Numer. Anal.*, 46(5):2411–2442, 2007.
8. Ganapathysubramanian, B. and Zabaras, N., Sparse grid collocation methods for stochastic natural convection problems, *J. Comput. Phys.*, 225:652–685, 2007.
9. Ma, X. and Zabaras, N., An adaptive high-dimensional stochastic model representation technique for the solution of stochastic partial differential equations, *J. Comput. Phys.*, 10:3884–3915, 2010.
10. Doostan, A. and Iaccarino, G., A least-squares approximation of partial differential equations with high-dimensional random inputs, *J. Comput. Phys.*, 228:4332–4345, 2009.
11. Choi, S.-K., Grandhi, R., Canfield, R., and Pettit, C., Polynomial chaos expansion with latin hypercube sampling for estimating response variability, *AIAA J.*, 42(6):1191–1198, 2004.
12. Berveiller, M., Sudret, B., and Lemaire, M., Stochastic finite element: A non intrusive approach by regression, *J. Eur. Meca. Num.*, 15(1-2-3):81–92, 2006.
13. Beylkin, G., Garcke, J., and Mohlenkamp, M., Multivariate regression and machine learning with sums of separable functions, *SIAM J. Sci. Comput.*, 31(3):1840–1857, 2009.
14. Doostan, A. and Owhadi, H., A non-adapted sparse approximation of PDEs with stochastic inputs, *J. Comput. Phys.*, 230(8):3015–3034, 2011.
15. Mathelin, L. and Gallivan, K., A compressed sensing approach for partial differential equations with random input data, *Commun. Comput. Phys.*, 12:919–954, 2012.
16. Candès, E. and Tao, T., Decoding by linear programming, *IEEE Trans. Inform. Theory*, 51:4203–4215, 2004.
17. Donoho, D., Compressed sensing, *IEEE Trans. Infor. Theo.*, 52(4):1289–1306, 2006.
18. Rabitz, H. and Alış, O., General foundations of high-dimensional model representations, *J. Math. Chem.*, 25:197–233, 1999.
19. Alış, O. and Rabitz, H., Efficient implementation of high-dimensional model representations, *J. Math. Chem.*, 29(2):127–142, 2001.
20. Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R., Least angle regression, *Ann. Stat.*, 32:407–499, 2004.
21. Abramowitz, M. and Stegun, I., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 10th edition, Dover Publications, New York, 1972.
22. Hesterberg, T., Choi, N. H., Meier, L., and Fraley, C., Least angle and ℓ_1 penalized regression: A review, *Stat. Surveys*, 2:61–93, 2008.

23. Hastie, T., Tibshirani, R., and Friedman, J., *The Elements of Statistical Learning*, 2nd ed., Springer Verlag, Berlin, 2009.
24. Wiener, N., The homogeneous chaos, *Am. J. Math.*, 60(4):897–936, 1938.
25. Ghanem, R. and Spanos, P., *Stochastic Finite Elements. A Spectral Approach*, rev. edition, Springer Verlag, Berlin, 2003.
26. Xiu, D. and Karniadakis, G., The Wiener-Askey polynomial chaos for stochastic differential equations, *SIAM J. Sci. Comput.*, 24(2):619–644, 2002.
27. Soize, C. and Ghanem, R., Physical systems with random uncertainties: Chaos representations with arbitrary probability measure, *SIAM J. Sci. Comput.*, 26(2):395–410, 2004.
28. Le Maître, O. and Knio, O., *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*, Springer, Berlin, 2010.
29. Khoromskij, B., Tensors-structured numerical methods in scientific computing: Survey on recents advances, *Chemometrics Intell. Lab. Syst.*, 110(1):1–19, 2012.
30. Khoromskij, B. and Schwab, C., Tensor-structured Galerkin approximation of parametric and stochastic elliptic pdes, *SIAM J. Sci. Comput.*, 33(1):1–25, 2011.
31. Harshman, R., Foundations of the parafac procedure: Models and conditions for an "explanatory" multimodal factor analysis, *UCLA Working Papers in Phonetics*, 16:1–84, 1970.
32. Carroll, J. and Chang, J.-J., Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart-Young" decomposition, *Psychometrika*, 35:283–319, 1970.
33. Kuo, F., Sloan, I., Wasilkowski, G., and Woźniakoski, H., On decompositions of multivariate functions, *Math. Comput.*, 79(270):953–966, 2009.
34. Candès, E. and Romberg, J., Sparsity and incoherence in compressive sampling, *Inv. Problems*, 23:969–985, 2006.
35. DeVore, R., Petrova, G., and Wojtaszczyk, P., Approximation of functions of few variables in high dimensions, *Constr. Approx.*, 33:125–143, 2011.
36. Tibshirani, R., Regression shrinkage and selection via the lasso, *J. R. Statist. Soc. Ser. B*, 58:267–288, 1996.
37. Chen, S., Donoho, D., and Saunders, M., Atomic decomposition by basis pursuit, *SIAM J. Sci. Comput.*, 20:33–61, 1999.
38. Candès, E. and Tao, T., Near-optimal signal recovery from random projections: Universal encoding strategies, *IEEE Trans. Inform. Theory*, 52:5406–5425, 2004.
39. Cai, T., Wang, L., and Xu, G., New bounds for restricted isometry constants, *IEEE Trans. Inform. Theory*, 56(9):4388–4394, 2010.
40. Blatman, G. and Sudret, B., Adaptive sparse polynomial chaos expansion based on least angle regression, *J. Comput. Phys.*, 230(6):2345–2367, 2011.
41. Yuan, M. and Lin, Y., Model selection and estimation in regression with grouped variables, *J. R. Statist. Soc. Ser. B*, 68:49–67, 2006.
42. Xie, J. and Zeng, L., Group variable selection methods and their applications in analysis of genomic data, in Feng, J., Fu, W., and Sun, F. (Eds.), *Frontiers in Computational and Systems Biology, Computational Biology*, Vol. 15, pp. 231–248, Springer-Verlag, Berlin, 2010.
43. Golub, G. and van Loan, C., *Matrix Computations*, 4th edition, JHU Press, Baltimore, MD, USA, 2012.
44. Markovsky, I. and van Huffel, S., Overview of total least squares methods, *Signal Processing*, 87(10):2283–2302, 2007.
45. Wentzell, P., Andrews, D., Hamilton, D., Faber, K., and Kowalski, B., Maximum likelihood principal component analysis, *J. Chemometrics*, 11:339–366, 1997.
46. Grasedyck, L., Hierarchical singular value decomposition of tensors, *SIAM J. Matrix Anal. Appl.*, 31(4):2029–2054, 2010.
47. Mathelin, L., Desceliers, C., and Hussaini, M., Stochastic data assimilation with a polynomial chaos parametric estimation, *Comput. Mech.*, 47(6):603–616, 2011.
48. Sobol, I., Sensitivity estimates for nonlinear mathematical models, *Math. Modell. Comput. Exper.*, 1(4):407–414, 1993.
49. Homma, T. and Saltelli, A., Importance measures in global sensitivity analysis of nonlinear models, *Rel. Eng. Syst. Safety*, 52(1):1–17, 1996.

APPENDIX A. A MOTIVATING EXAMPLE

To assess the choice of our *a priori* functional form for approximating a random variable, and while choosing a good basis is problem-dependent, let us consider a simple motivating example in the form of the 1D stochastic diffusion equation presented in Section 5.1, briefly recalled here for sake of convenience:

$$\nabla_x \left(\nu\left(x,\boldsymbol{\xi}\right) \nabla_x u\left(x,\boldsymbol{\xi}\right)\right) = F\left(x,\boldsymbol{\xi}\right), \qquad u\left(x_-,\boldsymbol{\xi}\right) = u_-,\ u\left(x_+,\boldsymbol{\xi}\right) = u_+. \tag{A.1}$$

The solution u is approximated under a separated format $u\left(x,\boldsymbol{\xi}\right) \approx \sum_{n=0}^{N} w_n\left(x\right) \lambda_n\left(\boldsymbol{\xi}\right)$. The approximation space for the spatial modes $\{w_n\left(x\right)\}$ is given and we here focus on the accuracy of the approximation with different representations for the stochastic modes $\{\lambda_n\left(\boldsymbol{\xi}\right)\}$. Each stochastic mode is determined either in a CP-like format, Eq. (12), or as an HDMR decomposition Eq. (13). In the latter case, interaction modes $\{f_{\boldsymbol{\gamma}}\}$ are approximated with a low-rank canonical decomposition on tensorized, unit-normed, univariate polynomials of maximum degree p: $f_{\boldsymbol{\gamma}}\left(\{\xi_i\}_{i\in\boldsymbol{\gamma}}\right) \approx \sum_{r=1}^{n_r} \prod_{i\in\boldsymbol{\gamma}} \sum_{\alpha=2}^{p} c_{\boldsymbol{\gamma},\alpha}^{r,i}\, \psi_\alpha\left(\xi_i\right)$. This approximation is hereafter referred to as a CP-HDMR decomposition. Similarly, univariate functions $\{f_{i,r}\}_{i=1}^{d}$ involved in the CP decomposition Eq. (12) are approximated with the same polynomials: $f_{i,r}\left(\xi_i\right) \approx \sum_{\alpha=1}^{p} c_{\alpha,i,r}\, \psi_\alpha\left(\xi_i\right)$.

The representation of the stochastic modes here relies on $p = 8$th-order univariate Legendre polynomials $\{\psi_\alpha\}$. The dimension of the problem is chosen to be $d_\nu = d_F = 5$ so that $d = 10$.

The CP-HDMR expansion is here built sequentially, starting with 0th and first-order interaction modes only. From this first approximation of the output, the set of dominant dimensions is estimated from the L^2-norm of univariate interaction modes $\{f_{\boldsymbol{\gamma}}\}_{|\boldsymbol{\gamma}|=1}$. Only second-order interaction modes $\{f_{\boldsymbol{\gamma}}\}_{|\boldsymbol{\gamma}|=2}$ in these dominant dimensions are next estimated and the set of dominant dimensions is then further refined based on both first and second-order interaction modes via the sensitivity Sobol indices, see Appendix B. Third-order modes are then computed for this new set of dominant dimensions only and the procedure is repeated until some stopping criterion is met, for instance a maximum interaction order N_l or a maximum basis cardinality $|\mathcal{J}|$. The number of samples N_q is here chosen sufficiently large so that full knowledge of u can be assumed. The approximation error then only comes from the choice of the approximation basis format, allowing a comparison. This section is loose on details, focusing on the main conclusions and leaving more in-depth discussion for main text sections.

First, the accuracy of the CP-HDMR approximation as a function of the decomposition rank N is studied in terms of ε, Eq. (26). Plotted in Fig. A.1, the approximation error estimation ε decreases when the maximum interaction order N_l increases from 1 to 3 and as the decomposition rank N increases.

The approximation is seen to improve exponentially fast as the number of modes N in the separated representation increases until it reaches a plateau. Increasing the interaction order leads to an improved approximation: increasing from first to second order brings more than a one-order-of-magnitude improvement in the approximation error estimation and an additional order of magnitude from $N_l = 2$ to $N_l = 3$. In this $d = 10$ example, the approximation hence exhibits a high convergence rate with N_l, supporting our assumption that low-order interactions dominate the HDMR decomposition.

This CP-HDMR approximation of the stochastic modes is now compared with a CP-like approximation in the form of Eq. (12). To evaluate the CP decomposition, we use an algorithm similar to that in [3]. Both decompositions rely on the same approximation basis for the deterministic modes $\{w_n\}$. We focus on the accuracy of the reconstruction as a function of the cardinality of the whole approximation basis both for $d = 10$ and $d = 40$ when the maximum decomposition rank N varies, see Fig. A.2. The total cardinality increases as more terms are considered in the decomposition series.

The accuracy of the representation is seen to improve as more terms are considered in the series expansion and both the CP and the CP-HDMR formats exhibit an exponential convergence with the total size $|\mathcal{J}|$ of the decomposition. The first-order CP-HDMR quickly reduces the error but plateaus as the functional space is small. The third order CP-HDMR decomposition is more costly in number of coefficients to evaluate to reach a given error and, in the present settings, should only be considered if high accuracy is needed.

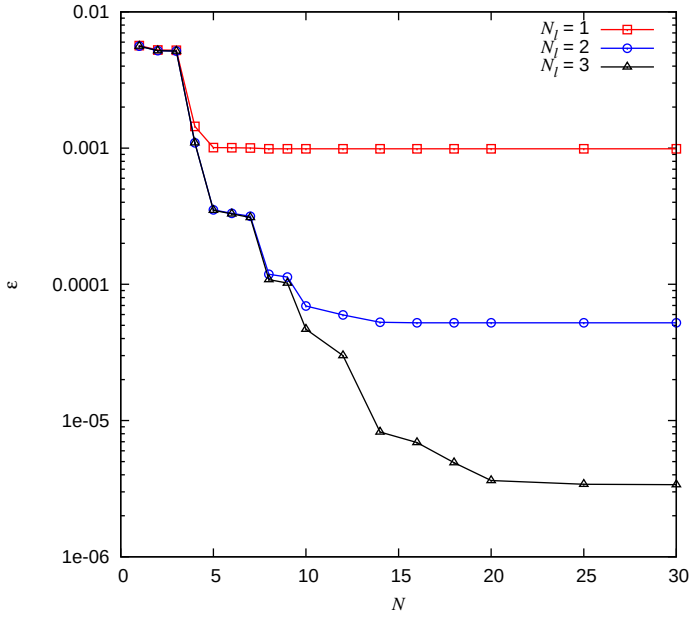

FIG. A.1: Convergence of the error estimation of the approximation with the order N of the separated representation and the maximum interaction order N_l of the HDMR expansion.

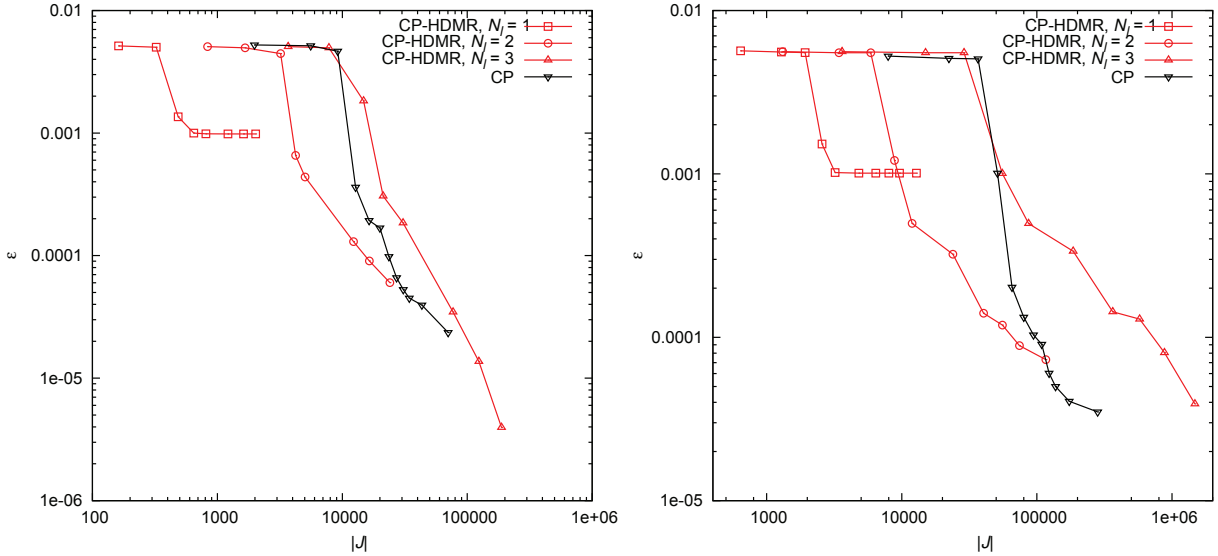

FIG. A.2: Convergence of the approximation error estimation ε with the total cardinality $|\mathcal{J}|$ of the representation basis. Approximations of the stochastic modes with the CP-HDMR and CP-like format are compared. $d = 10$ (left) and $d = 40$ (right).

Unless the targeted accuracy is really high, this motivating example tends to indicate that, for a reasonable required accuracy, a CP-HDMR format involves fewer unknowns than a CP-like decomposition, both for a low $d = 10$- and a moderate $d = 40$-dimensional problem. This is an important point since the number of coefficients which can be evaluated with a reasonable accuracy from experimental data is directly related to the size of the available dataset. Finally, a CP-HDMR format allows a great flexibility in representing interaction modes $\{f_{\boldsymbol{\gamma}}\}$. In particular,

a more parsimonious representation is used in the main text and achieves a similar accuracy with a lower number of terms.

APPENDIX B. STATISTICS AND SENSITIVITY ANALYSIS

Once an approximation of a random variable $u(\boldsymbol{\xi})$ is obtained, it is easy to estimate its first statistical moments. From the HDMR format properties, the estimated mean is simply given by the first term of the decomposition: $\langle u \rangle_{L^2(\Xi,\,\mu_\Xi)} \simeq f_\emptyset$.

Thanks to the orthogonality property of the modes $\{f_{\boldsymbol{\gamma}}\}$, the variance $\mathrm{Var}(u) := \left\langle \left(u - \langle u \rangle_{L^2(\Xi,\,\mu_\Xi)} \right)^2 \right\rangle_{L^2(\Xi,\,\mu_\Xi)}$ approximates as the sum of the variance of the individual interaction modes:

$$\begin{aligned}\mathrm{Var}(u) &\simeq \sum_{\boldsymbol{\gamma}\in\mathcal{J}_{f,\mathrm{eff}}\setminus\emptyset} \mathrm{Var}\left(\widehat{f}_{\boldsymbol{\gamma}}\right), \\ &= \sum_{\substack{\boldsymbol{\gamma}\in\mathcal{J}_{f,\mathrm{eff}}\setminus\emptyset \\ |\boldsymbol{\gamma}|\le N_l^{(\mathrm{PC})}}} \sum_{\boldsymbol{\alpha},|\boldsymbol{\alpha}|\le p} c_{\boldsymbol{\gamma},\boldsymbol{\alpha}}^2 \, \|\psi_{\boldsymbol{\alpha}}\|^2_{L^2(\Xi,\,\mu_\Xi)} + \sum_{\substack{\boldsymbol{\gamma}\in\mathcal{J}_{f,\mathrm{eff}}\setminus\emptyset \\ N_l^{(\mathrm{PC})}<|\boldsymbol{\gamma}|\le N_l}} \sum_{r,r'=1}^{n_r(\boldsymbol{\gamma})} \prod_{i\in\boldsymbol{\gamma}} \sum_{\alpha=1}^{p} c_{\boldsymbol{\gamma},\alpha}^{r,i}\, c_{\boldsymbol{\gamma},\alpha}^{r',i} \, \|\psi_\alpha\|^2_{L^2(\Xi,\,\mu_\Xi)}, \end{aligned} \tag{B.1}$$

where use was made of the orthogonality of the Hilbertian basis $\{\psi_\alpha\}$.

Other standard statistical quantities are the sensitivity indices $\{S_{\boldsymbol{\gamma}}\} := \mathrm{Var}\left(\widehat{f}_{\boldsymbol{\gamma}}\right) / \mathrm{Var}(\widehat{u})$ which essentially represent the relative part of the variance of the QoI due to the interaction of a given set of input random variables only, [48, 49]. From Eq. (B.1), it immediately follows that $\sum_{\boldsymbol{\gamma}\subseteq\{1,\ldots,d\}\setminus\emptyset} S_{\boldsymbol{\gamma}} = 1$ and the explicit expression of the sensitivity indices is straightforward to derive from the HDMR format. In practice, it is often more useful to assess the influence of a given input onto the variance of the QoI with the total sensitivity indices $\{S_{T,i}\}_{i=1}^d$:

$$S_{T,i} := \frac{\sum_{\boldsymbol{\gamma}\subseteq\{1,\ldots,d\}\setminus\emptyset:\, i\in\boldsymbol{\gamma}} \mathrm{Var}\left(\widehat{f}_{\boldsymbol{\gamma}}\right)}{\mathrm{Var}(\widehat{u})}, \qquad 1 \le i \le d. \tag{B.2}$$

Again, using Eq. (B.1), this quantity is straightforward to estimate once the approximation of $u(\boldsymbol{\xi})$ is available.