Annals of Mathematical Sciences and Applications
Volume 1, Number 2, 341–377, 2016

Dimensionality reduction of parameter-dependent problems through proper orthogonal decomposition

Andrea Manzoni, Federico Negri, and Alfio Quarteroni

The numerical solution of partial differential equations (PDEs) depending on parametrized or random input data is computationally intensive. Reduced order modeling techniques, such as the reduced basis methods, have been developed to alleviate this computational burden, and are nowadays exploited to accelerate real-time analysis, as well as the solution of PDE-constrained optimization and inverse problems. These methods are built upon low-dimensional spaces obtained by selecting a set of snapshots from a parametrically induced manifold. However, for these techniques to be effective, both parameter-dependent and random input data must be expressed in a convenient form. To address the former case, the empirical interpolation method has been developed. In the latter case, a spectral approximation of stochastic fields is often generated by means of a Karhunen-Loève expansion. In all these cases, a low dimensional space to represent the function being approximated (PDE solution, parametrized data, stochastic field) can be obtained through proper orthogonal decomposition. Here, we review possible ways to exploit this methodology in these three contexts, we recall its optimality properties, and highlight the common mathematical structure beneath.

Keywords and phrases: Reduced order modeling, proper orthogonal decomposition, empirical interpolation, Karhunen-Loève expansion.

1. Introduction

The development of efficient numerical methods for differential problems is of paramount importance in scientific computing. Being able to solve a partial differential equation (PDE) using a handful of degrees of freedom – instead of a huge amount like in classical full-order techniques – is mandatory to speedup the repeated solution of the problem e.g. for a huge number of system configurations, operating conditions, or realizations when dealing with uncertain features. All these inputs can be related to physical coef-

ficients, source terms, boundary/initial conditions, or even to the spatial domain $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, wherein the PDE has to be solved.

In turn, these quantities can be *(i)* known parameter-dependent functions expressed in terms of *deterministic* input parameters, or *(ii)* modeled through *random* input parameters. In the former case, we denote by $\boldsymbol{\mu} \in \mathcal{P}_\mu \subset \mathbb{R}^P$ a vector of P deterministic input parameters; in the latter case, we denote by $\boldsymbol{\psi} \in \mathcal{P}_\psi \subset \mathbb{R}^Q$ a vector of Q random input parameters, by assuming that we can *parametrize* random fields by a finite number of random variables. Here $\mathcal{P}_\mu, \mathcal{P}_\psi$ denote parameter spaces – usually, closed, bounded subsets of $\mathbb{R}^P, \mathbb{R}^Q$, respectively. For instance, consider the parametrized PDE

$$\left\{ \begin{array}{rl} -\text{div}(\nu(\mathbf{x};\omega)\nabla u(\mathbf{x};\boldsymbol{\mu})) = & f(\mathbf{x};\boldsymbol{\mu}) \quad \text{in } \Omega \\ u(\mathbf{x};\boldsymbol{\mu}) = & h(\mathbf{x};\boldsymbol{\mu}) \quad \text{on } \partial\Omega, \end{array} \right. \tag{1}$$

where $\boldsymbol{\mu} \in \mathcal{P}$ is a vector of deterministic input parameters, $f(\mathbf{x};\boldsymbol{\mu})$ and $h(\mathbf{x};\boldsymbol{\mu})$ are $\boldsymbol{\mu}$-dependent functions, and $\nu = \nu(\mathbf{x};\omega)$ is a (positive) random field, that is, *(i)* for a fixed point $\mathbf{x} \in \Omega$, $\nu(\mathbf{x};\cdot)$ is a random variable over the outcome space Θ; *(ii)* for a fixed $\omega \in \Theta$, $\nu(\cdot;\omega)$ is a realization of the random field in Ω. In many contexts, problem (1) has to be *efficiently* solved for many input parameter values: this might be required, e.g, for the sake of sensitivity analysis, PDE-constrained optimization (where e.g. f and/or h are unknown functions to be determined in order to reach some prescribed target) and in any case when the PDE depends on some random input data – for instance, when using the stochastic collocation method, see e.g. [4]. At least three needs arise when dealing with (1).

First of all, even in the case where $\nu = \nu(\mathbf{x})$ is not a random quantity, the numerical approximation of problem (1) for any new $\boldsymbol{\mu}$ usually entails the solution of a linear system of large dimension $N_h \times N_h$. This latter stems from the discretization of the PDE according to standard techniques such as the Galerkin-finite element method, to which we refer to as *high-fidelity* technique (or *full-order* model), meaning that we assume that the problem can be solved up to any desired accuracy level. The *Reduced Basis (RB) method* [41] replaces the high-fidelity problem by a reduced problem of much smaller dimension, by seeking in a subspace of lower dimension $N \ll N_h$ an approximate solution expressed as a linear combination of suitable, problem-dependent, basis functions $\phi_1, \ldots, \phi_N$. The latter are generated from a given set of n_s solutions $\{u_h(\boldsymbol{\mu}^1), \ldots, u_h(\boldsymbol{\mu}^{n_s})\}$, $n_s \geq N$, of the high-fidelity problem, called *snapshots*, corresponding to a suitably chosen set of parameter values. Once a reduced basis has been built, the RB solution is expressed as

a linear combination of the basis functions,

$$u_N(\mathbf{x};\boldsymbol{\mu}) = \sum_{j=1}^{N} u_j(\boldsymbol{\mu})\phi_j(\mathbf{x}), \qquad \boldsymbol{\mu} \in \mathcal{P}_\mu.$$

The components of the RB solution $\mathbf{u}_N(\boldsymbol{\mu})$ are determined by requiring that a suitable (geometric) orthogonality criterion is fulfilled – this yields a linear system of N equations to be solved, the *RB system*.

At the same time, assembling for any new value of $\boldsymbol{\mu}$ the RB system efficiently – that is, purely relying on arrays of small dimension N – requires to represent the $\boldsymbol{\mu}$-dependent data in a suitable form, fulfilling the so-called *affine parametric dependence*. By this latter, we mean that we can approximate, e.g. $f(\mathbf{x};\boldsymbol{\mu})$, by

$$f_M(\mathbf{x};\boldsymbol{\mu}) = \sum_{j=1}^{M} \gamma_j(\boldsymbol{\mu})\rho_j(\mathbf{x}), \qquad \boldsymbol{\mu} \in \mathcal{P}_\mu,$$

where $\rho_1, \ldots, \rho_M$ are pre-selected, linearly independent basis functions. The coefficients $\gamma_1, \ldots, \gamma_M$ are obtained by requiring that the approximation is exact at a finite set of M (interpolation) points.

Finally, for the sake of computational convenience, we assume that the random field $\nu(\mathbf{x};\omega)$ can be expressed as an infinite sum of uncorrelated random variables $\{\psi_j(\omega)\}_{j=1}^{\infty}$, for a given set of real-valued functions $\{\nu_j(\mathbf{x})\}_{j=1}^{\infty}$. Since only few modes are *usually* relevant, a low-dimensional representation of a random field is obtained by truncating the infinite sum and retaining a (hopefully small) finite number Q of terms:

$$\nu_Q(\mathbf{x};\omega) = \sum_{j=1}^{Q} \psi_j(\omega)\nu_j(\mathbf{x}),$$

provided the variables $\psi_1, \psi_2, \ldots$ are uncorrelated. These variables can thus play the role of *random input parameters* $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_q) \in \mathcal{P}_\psi$.

Proper orthogonal decomposition (POD) can be exploited to obtain a low-dimensional representation of data in each of these contexts.

The strong development that reduced order modeling, low rank tensor approximations and problems involving PDEs depending on (possibly random) input parameters have undergone in these last two decades motivates the investigation of POD as *dimensional reduction technique* at least for two reasons: *(i)* a better understanding of the common mathematical structure beneath the three contexts above, as well as the optimality results they rely

on, can provide better insights also in view of developing efficient computational techniques; *(ii)* the need of reduction is nowadays shared by several applications involving large scale data and/or PDE problems, thus motivating a strong interplay among techniques historically developed to address one of the three goals. For instance, sampling techniques (such as sparse grids) mandatory to deal with PDEs depending on random input data prove to be extremely useful also when constructing a reduced basis space to better explore the solution manifold. POD can thus be exploited within:

1. the Reduced Basis method for the solution of a parametrized PDE;
2. the Empirical Interpolation method (EIM), and its discrete variant, for the affine approximation of a parameter-dependent function;
3. the Karhunen-Loève (KL) expansion[1] (or transformation) for the parametrization of a stochastic field.

For the sake of generality, we denote by $\boldsymbol{\tau}$ any kind of parameter vector, and by $g : \mathcal{P} \to V$ a generic function for which we seek a low-dimensional approximation, being V a Hilbert space; for any $\boldsymbol{\tau}$, $g = g(\,\cdot\,;\boldsymbol{\tau})$ is a spatial field over the domain Ω. The problem we focus on is thus related with the construction of low dimensional spaces for simultaneously approximating *all* the elements belonging to a parameter-dependent family

$$\mathcal{M} = \{g(\,\cdot\,;\boldsymbol{\tau})\ ,\ \boldsymbol{\tau} \in \mathcal{P}\} \subset V, \tag{2}$$

where $g = u$, $g = f$ or $g = \nu$ in cases 1, 2, and 3, respectively. Hence:

- first, we determine a basis $\{\zeta_i\}_{i=1}^N$ – fulfilling a suitable optimality criterion – to represent all the elements of $\mathcal{M}$. This stage may require the sampling of snapshots from $\mathcal{M}$ (e.g. in the cases 1 and 2);
- once the basis has been computed, we express each element of $\mathcal{M}$ as

$$g(\mathbf{x};\boldsymbol{\tau}) \approx g_N(\mathbf{x};\boldsymbol{\tau}) = \sum_{i=1}^{N} \zeta_i(\mathbf{x})\beta_i(\boldsymbol{\tau}). \tag{3}$$

  The vector of coefficients $\boldsymbol{\beta}(\boldsymbol{\tau}) \in \mathbb{R}^N$ is determined in different ways (e.g., by means of projection or interpolation) depending on the case.

[1] This is the way POD is called in the context of stochastic processes. In multivariate statistics, instead, POD goes under the name of principal component analysis, where the POD modes are called principal components.

Ideally, for a given tolerance $\varepsilon > 0$, the dimension N should be chosen such that

$$\left\| g(\mathbf{x};\boldsymbol{\tau}) - \sum_{i=1}^{N} \zeta_i(\mathbf{x})\beta_i(\boldsymbol{\tau}) \right\| \leq \varepsilon \qquad \forall \boldsymbol{\tau} \in \mathcal{P}.$$

In this paper, the basis $\{\zeta_i\}_{i=1}^N$ is systematically constructed through POD. This is not, of course, the only available option: greedy algorithms are for instance largely used in the construction of low-dimensional approximations, both in the EIM for approximating parameter-dependent functions [6] and in the RB method for parametrized PDEs [40, 39], relying in this latter case on (computable) a posteriori error estimations. Two additional techniques – indeed quite close to POD – for generating reduced spaces are the Centroidal Voronoi Tessellation [11, 12] and the Proper Generalized Decomposition (see for instance [37]). Another technique which can be used to parametrize random fields is polynomial chaos expansion, which exploits polynomial expansions in terms of independent random variables, see e.g. [51]. Both KL and polynomial chaos expansion have been widely used together with stochastic Galerkin approximations [22, 5] or collocation methods [4] for solving PDEs with random input data.

For the sake of notation, hereon we consider a *discretized* version of any quantity depending on spatial coordinates. In this respect, we introduce a discretized version of the spatial domain, made by N_h degrees of freedom – these latter might be related to a set of quadrature points defined over a triangulation Ω_h of Ω, to the vertices (or to the centers) of the elements whose union gives Ω_h. Hence, all the spatially dependent quantities are expressed by (parameter-dependent) vectors in $\mathbb{R}^{N_h}$. For instance, a random field $\nu(\mathbf{x};\omega)$ turns hereon into a random vector $\boldsymbol{\nu}(\omega)$, meaning that for any $\omega \in \Theta$ we deal with the evaluation of $\nu(\cdot;\omega)$ into N_h degrees of freedom, thus yielding a vector $\boldsymbol{\nu} \in \mathbb{R}^{N_h}$. We denote by V a generic functional space – remarkable examples are $V = L^2(\Omega)$ in the case of random fields, $V \subseteq H^1(\Omega)$ in the case of the solution of second-order elliptic PDEs, $V = C^0(\bar{\Omega})$ in the case of function interpolation. Moreover, $V_h \subset V$ denotes any finite-dimensional subspace of V, of dimension N_h, and $\{\varphi_j\}_{j=1}^{N_h}$ an orthonormal basis of V_h; a one-to-one map holds between the components of a vector $\mathbf{v} = (v_1, \ldots, v_{N_h}) \in \mathbb{R}^{N_h}$ and the corresponding function $v \in V_h$, that is, $v(\mathbf{x}) = \sum_{j=1}^{N_h} v_j \varphi_j(\mathbf{x})$.

The paper is organized as follows. We provide a continuous description of POD in Sect. 2, then we introduce the (more common) method of snapshots in Sect. 3. We show how to exploit this procedure to devise a RB method for solving parameter-dependent PDEs in Sect. 4 and a discrete version of

the empirical interpolation method for approximating parameter-dependent functions in Sect. 5, respectively. Then, we illustrate the Karhunen-Loève expansion for stochastic fields in Sect. 6. Finally, in Sect. 7 we present some numerical results dealing with the approximation of a parametrized elliptic PDE describing a stationary advection-diffusion problem.

2. POD on parametrized manifolds

Proper orthogonal decomposition is, in a very broad sense, a technique for reducing the dimensionality of given data by representing them onto an orthonormal basis which is optimal in a least-squares sense. The original variables are transformed into a new set of uncorrelated variables (called POD modes), the first few N modes ideally retaining most of the *energy* present in all of the original variables.

Let us suppose that $\mathbf{g}_h(\boldsymbol{\tau}) \in L^2(\mathcal{P};\mathbb{R}^{N_h})$, i.e. $\int_{\mathcal{P}} \|\mathbf{g}_h(\boldsymbol{\tau})\|_2^2 \, d\boldsymbol{\tau} < \infty$ (that is, $\mathbf{g}_h(\boldsymbol{\tau})$ is a Hilbert-Schmidt kernel) and denote by $\mathcal{V}_N = \{\mathbb{W} \in \mathbb{R}^{N_h \times N} \colon \mathbb{W}^T\mathbb{W} = \mathbb{I}_N\}$ the set of all N-dimensional orthonormal bases. To find a POD basis of dimension N that approximates the parameter-dependent family $\mathcal{M} = \{\mathbf{g}_h(\boldsymbol{\tau}) : \boldsymbol{\tau} \in \mathcal{P}\} \subset \mathbb{R}^{N_h}$, we have to determine

$$\min_{\mathbb{W}\in\mathcal{V}_N} \int_{\mathcal{P}} \|\mathbf{g}_h(\boldsymbol{\tau}) - \mathbb{W}\mathbb{W}^T\mathbf{g}_h(\boldsymbol{\tau})\|_2^2 \, d\boldsymbol{\tau}. \tag{4}$$

The solution of this *nonconvex* optimization problem is due to E. Schmidt [46] (see also [48]), who first showed how to obtain an optimal low-rank approximation to a (compact) integral operator; this is indeed related to an infinite-dimensional analogue of the singular value decomposition (SVD) of a matrix.

Let us denote by $T : L^2(\mathcal{P}) \to \mathbb{R}^{N_h}$ the operator defined as

$$Tf = \int_{\mathcal{P}} \mathbf{g}_h(\boldsymbol{\tau}) f(\boldsymbol{\tau}) \, d\boldsymbol{\tau} \qquad \forall f \in L^2(\mathcal{P}); \tag{5}$$

similarly, its adjoint $T^* : \mathbb{R}^{N_h} \to L^2(\mathcal{P})$ is given by

$$(f, T^*\mathbf{w})_{L^2(\mathcal{P})} = (Tf, \mathbf{w})_2 \qquad \forall f \in L^2(\mathcal{P}), \mathbf{w} \in \mathbb{R}^{N_h},$$

that is,

$$T^*\mathbf{w} = (\mathbf{g}_h(\boldsymbol{\tau}), \mathbf{w})_2 \qquad \forall \mathbf{w} \in \mathbb{R}^{N_h}, \tag{6}$$

Note that both T and T^* depends on $\boldsymbol{\tau}$; however, we omit the dependence on $\boldsymbol{\tau}$ wherever it is clear from the context. Note also that $Tf \in \mathbb{R}^{N_h}$ and $\text{rank}(T) = r \leq N_h$, while $T^*\mathbf{w} = T^*\mathbf{w}(\boldsymbol{\tau}) \in L^2(\mathcal{P})$. Since $\mathbf{g}_h(\boldsymbol{\tau})$ is a Hilbert-Schmidt kernel, T is a Hilbert-Schmidt, and thus compact, operator [45, 18]. Moreover, the Hilbert-Schmidt norm[2] of T coincides with the norm of its kernel (see e.g. [44]), i.e.

$$\|T\|_{HS}^2 = \|\mathbf{g}_h\|_{L^2(\mathcal{P};\mathbb{R}^{N_h})}^2.$$

Since T is compact, $K = TT^* : \mathbb{R}^{N_h} \to \mathbb{R}^{N_h}$ and $C = T^*T : L^2(\mathcal{P}) \to L^2(\mathcal{P})$ are self-adjoint, non-negative, compact operators, given by

$$Cf = \int_{\mathcal{P}} (\mathbf{g}_h(\boldsymbol{\tau}), \mathbf{g}_h(\boldsymbol{\tau}'))_2 f(\boldsymbol{\tau}')\, d\boldsymbol{\tau}' \qquad \forall f \in L^2(\mathcal{P}), \tag{7}$$

$$K\mathbf{w} = \int_{\mathcal{P}} \mathbf{g}_h(\boldsymbol{\tau})(\mathbf{g}_h(\boldsymbol{\tau}), \mathbf{w})_2\, d\boldsymbol{\tau} \qquad \forall \mathbf{w} \in \mathbb{R}^{N_h},$$

respectively. Actually, since K is a linear map from $\mathbb{R}^{N_h}$ to $\mathbb{R}^{N_h}$, it is represented by the $N_h \times N_h$ symmetric positive definite matrix

$$K = \int_{\mathcal{P}} \mathbf{g}_h(\boldsymbol{\tau})\mathbf{g}_h^T(\boldsymbol{\tau})\, d\boldsymbol{\tau} \tag{8}$$

whose eigenvalues $\lambda_1 \geq \cdots \geq \lambda_r \geq 0$ and associated orthonormal eigenvectors $\boldsymbol{\zeta}_i \in \mathbb{R}^{N_h}$ satisfy

$$K\boldsymbol{\zeta}_i = \lambda_i \boldsymbol{\zeta}_i, \qquad i = 1, \ldots, r. \tag{9}$$

Moreover, also the operator C admits an eigendecomposition

$$C\psi_i = \gamma_i \psi_i, \qquad i = 1, \ldots, r$$

where the eigenvalues γ_i are indeed equal to the eigenvalues λ_i of K, whereas the eigenvectors $\psi_i \in L^2(\mathcal{P})$ are related to those of K as follows

$$\psi_i = \frac{1}{\sqrt{\lambda_i}} T^* \boldsymbol{\zeta}_i, \qquad i = 1, \ldots, r. \tag{10}$$

[2] If $\{e_n\}_{n=1}^{\infty}$ is an orthonormal basis for V and $L \in \mathcal{L}(V, W)$, the Hilbert-Schmidt (HS) norm of L is $\|L\|_{HS} = \left(\sum_{n=1}^{\infty} \|Le_n\|_W^2\right)^{1/2}$. The HS norm of a compact operator $L \in \mathcal{L}(V, W)$ is given by $\|L\|_{HS} = \left(\sum_{n=1}^{\infty} \sigma_n(L)^2\right)^{1/2}$, being $\{\sigma_n(L)\}_{n=1}^{\infty}$ the singular values of L, and is indeed similar to the Frobenius norm for matrices.

Indeed, thanks to (9)–(10)

(11)
$$C\psi_i = C\frac{1}{\sqrt{\lambda_i}}T^*\boldsymbol{\zeta}_i = T^*T\left(\frac{1}{\sqrt{\lambda_i}}T^*\boldsymbol{\zeta}_i\right) = \frac{1}{\sqrt{\lambda_i}}T^*K\boldsymbol{\zeta}_i = \sqrt{\lambda_i}T^*\boldsymbol{\zeta}_i = \lambda_i\psi_i.$$

Finally, since the $\boldsymbol{\zeta}_i$'s form an orthonormal basis for $\mathbb{R}^{N_h}$, $\mathbf{g}_h(\boldsymbol{\tau})$ admits the expansion

$$\mathbf{g}_h(\boldsymbol{\tau}) = \sum_{i=1}^{r} \boldsymbol{\zeta}_i(\mathbf{g}_h(\boldsymbol{\tau}), \boldsymbol{\zeta}_i)_2 = \sum_{i=1}^{r} \sqrt{\lambda_i}\boldsymbol{\zeta}_i\psi_i(\boldsymbol{\tau}), \tag{12}$$

and the following decomposition holds for T

$$T = \sum_{i=1}^{r} \sqrt{\lambda_i}\,\boldsymbol{\zeta}_i\,(\psi_i, \cdot)_{L^2(\mathcal{P})}. \tag{13}$$

Moreover, it can be easily proved that

$$\|T\|_{\mathcal{L}(L^2(\mathcal{P}),\mathbb{R}^{N_h})} = \sqrt{\lambda_1}, \qquad \|T\|_{HS} = \sqrt{\sum_{i=1}^{r} \lambda_i}; \tag{14}$$

furthermore, by truncating the sum (13) to the first N terms we obtain the best rank N approximation to the operator T, according to the following

Theorem 2.1 (Schmidt)**.** *The operator $T_N : L^2(\mathcal{P}) \to \mathbb{R}^{N_h}$ defined by*

$$T_N = \sum_{i=1}^{N} \sqrt{\lambda_i}\,\boldsymbol{\zeta}_i\,(\psi_i, \cdot)_{L^2(\mathcal{P})}, \qquad 0 \le N < r, \tag{15}$$

satisfies the following optimality property

$$\|T - T_N\|_{HS} = \min_{B\in\mathcal{B}_N} \|T - B\|_{HS} = \sqrt{\sum_{i=N+1}^{r} \lambda_i}, \tag{16}$$

where $\mathcal{B}_N = \{B \in \mathcal{L}(L^2(\mathcal{P}), \mathbb{R}^{N_h}) : \mathrm{rank}(B) \le N\}$.

See, e.g., [41, Theorem 6.2] for the proof. Theorem 2.1 hence provides a solution to the minimization problem (4), as stated in the following Proposition.

Proposition 2.1. *The POD basis* $\mathbb{V} = [\boldsymbol{\zeta}_1 \mid \ldots \mid \boldsymbol{\zeta}_N] \in \mathbb{R}^{N_h \times N}$ *is such that*

$$\int_{\mathcal{P}} \|\mathbf{g}_h(\boldsymbol{\tau}) - \mathbb{V}\mathbb{V}^T \mathbf{g}_h(\boldsymbol{\tau})\|_2^2 \, d\boldsymbol{\tau} = \min_{\mathbb{W} \in \mathcal{V}_N} \int_{\mathcal{P}} \|\mathbf{g}_h(\boldsymbol{\tau}) - \mathbb{W}\mathbb{W}^T \mathbf{g}_h(\boldsymbol{\tau})\|_2^2 \, d\boldsymbol{\tau}. \tag{17}$$

Moreover

$$\int_{\mathcal{P}} \|\mathbf{g}_h(\boldsymbol{\tau}) - \mathbb{V}\mathbb{V}^T \mathbf{g}_h(\boldsymbol{\tau})\|_2^2 \, d\boldsymbol{\tau} = \sum_{i=N+1}^{r} \lambda_i. \tag{18}$$

Proof. Since $\psi_i = \frac{1}{\sqrt{\lambda_i}} T^* \boldsymbol{\zeta}_i$, thanks to the definition of T^* we express T_N as

$$\begin{aligned} T_N f = \sum_{i=1}^{N} \sqrt{\lambda_i} \boldsymbol{\zeta}_i (\psi_i(\boldsymbol{\tau}), f(\boldsymbol{\tau}))_{L^2(\mathcal{P})} &= \int_{\mathcal{P}} (\mathbf{g}_h(\boldsymbol{\tau}), \boldsymbol{\zeta}_i)_2 \boldsymbol{\zeta}_i \, f(\boldsymbol{\tau}) \, d\boldsymbol{\tau} \\ &= \int_{\mathcal{P}} \mathbb{V}\mathbb{V}^T \mathbf{g}_h(\boldsymbol{\tau}) f(\boldsymbol{\tau}) \, d\boldsymbol{\tau} \qquad \forall f \in L^2(\mathcal{P}). \end{aligned}$$

Then

$$\|T - T_N\|_{HS}^2 = \|T - \mathbb{V}\mathbb{V}^T T\|_{HS}^2 = \int_{\mathcal{P}} \|\mathbf{g}_h(\boldsymbol{\tau}) - \mathbb{V}\mathbb{V}^T \mathbf{g}_h(\boldsymbol{\tau})\|_2^2 \, d\boldsymbol{\tau}.$$

Since $\|T - T_N\|_{HS}^2 = \min_{B \in \mathcal{B}_N} \|T - B\|_{HS}^2 \leq \min_{\mathbb{W} \in \mathcal{V}_N} \|T - \mathbb{W}\mathbb{W}^T T\|_{HS}^2$ and $T_N = \mathbb{V}\mathbb{V}^T T$, we conclude that

$$\|T - \mathbb{V}\mathbb{V}^T T\|_{HS}^2 = \min_{\mathbb{W} \in \mathcal{V}_N} \|T - \mathbb{W}\mathbb{W}^T T\|_{HS}^2. \qquad \square$$

3. POD by the method of snapshots

From a practical standpoint, generating the POD basis by solving problem (9) requires to form the operator K. This is clearly unfeasible since it would require the knowledge of $\mathbf{g}(\boldsymbol{\tau})$ for any $\boldsymbol{\tau} \in \mathcal{P}$ – as well as the possibility to evaluate the integral in (8); note that this operation is indeed made possible in the case of the KL expansion once the covariance function of the random field is known analytically.

The method of snapshots, introduced by Sirovich [47] is a way of determining the POD basis without explicitly calculating the operator K. To this end, consider a set $\Xi_s = \{\boldsymbol{\tau}^1, \ldots, \boldsymbol{\tau}^{n_s}\}$ of n_s parameter samples and the corresponding set of *snapshots* $\{\mathbf{g}_h(\boldsymbol{\tau}^1), \ldots, \mathbf{g}_h(\boldsymbol{\tau}^{n_s})\}$; we denote the snapshot matrix $\mathbb{S} \in \mathbb{R}^{N_h \times n_s}$ by

$$\mathbb{S} = [\mathbf{g}_1 \mid \ldots \mid \mathbf{g}_{n_s}],$$

where $\mathbf{g}_i = \mathbf{g}_h(\boldsymbol{\tau}^i) \in \mathbb{R}^{N_h}$, $1 \leq i \leq n_s$. We denote by $r \leq \min(N_h, n_s)$ the rank of $\mathbb{S}$, which is strictly smaller than n_s if the snapshot vectors are not all linearly independent. The matrix

$$\mathbb{K} = \sum_{i=1}^{n_s} \mathbf{g}_i \mathbf{g}_i^T = \mathbb{S}\mathbb{S}^T \in \mathbb{R}^{N_h \times N_h} \tag{19}$$

is the $\mathcal{P}$-discrete approximation of the one defined in (8) – note that in (19) the integral over $\mathcal{P}$ has been replaced by a sum over the set of snapshots – and its eigenvectors are such that

$$\mathbb{S}\mathbb{S}^T \boldsymbol{\zeta}_i = \lambda_i \boldsymbol{\zeta}_i, \qquad i = 1, \ldots, r.$$

Similarly, the matrix

$$\mathbb{C} = \mathbb{S}^T \mathbb{S} \in \mathbb{R}^{n_s \times n_s}$$

is the $\mathcal{P}$-discrete approximation of the operator C introduced in (7); the eigenpairs of $\mathbb{C}$

$$\mathbb{S}^T \mathbb{S} \boldsymbol{\psi}_i = \gamma_i \boldsymbol{\psi}_i$$

are related to those of $\mathbb{K}$ as follows: $\gamma_i = \lambda_i$ and

$$\boldsymbol{\zeta}_i = \frac{1}{\sqrt{\lambda_i}} \mathbb{S} \boldsymbol{\psi}_i, \qquad 1 \leq i \leq r. \tag{20}$$

In fact, similarly to (11), $\sqrt{\lambda_i}\boldsymbol{\psi}_i = \mathbb{S}^T \boldsymbol{\zeta}_i$, so that

$$\mathbb{C}\boldsymbol{\psi}_i = \mathbb{S}^T \mathbb{S} \boldsymbol{\psi}_i = \mathbb{S}^T \mathbb{S} \frac{1}{\sqrt{\lambda_i}} \mathbb{S}^T \boldsymbol{\zeta}_i = \sqrt{\lambda_i} \mathbb{S}^T \boldsymbol{\zeta}_i = \lambda_i \boldsymbol{\psi}_i.$$

Equation (20) is the $\mathcal{P}$-discrete counterpart of (10). Note also that $\lambda_i, i = 1, \ldots, r$ are the nonzero eigenvalues of both matrices $\mathbb{S}^T\mathbb{S}$ and $\mathbb{S}\mathbb{S}^T$, listed in nondecreasing order.

Similarly to Proposition 2.1, the following result characterizes the POD basis obtained through the method of snapshots: starting from the snapshots collected in the matrix $\mathbb{S} \in \mathbb{R}^{N_h \times n_s}$, for any $N \leq n_s$ the POD basis $\mathbb{V} \in \mathbb{R}^{N_h \times N}$ of dimension N is defined as the set of the first N eigenvectors of $\mathbb{K}$.

Proposition 3.1. *Let* $\mathcal{V}_N = \{\mathbb{W} \in \mathbb{R}^{N_h \times N} : \mathbb{W}^T\mathbb{W} = \mathbb{I}_N\}$ *be the set of all* N*-dimensional orthonormal bases. The POD basis* $\mathbb{V} = [\boldsymbol{\zeta}_1 \mid \ldots \mid \boldsymbol{\zeta}_N] \in \mathbb{R}^{N_h \times N}$

is such that

$$\sum_{i=1}^{n_s} \|\mathbf{g}_i - \mathbb{V}\mathbb{V}^T\mathbf{g}_i\|_2^2 = \min_{\mathbb{W}\in\mathcal{V}_N} \sum_{i=1}^{n_s} \|\mathbf{g}_i - \mathbb{W}\mathbb{W}^T\mathbf{g}_i\|_2^2. \tag{21}$$

Moreover,

$$\sum_{i=1}^{n_s} \|\mathbf{g}_i - \mathbb{V}\mathbb{V}^T\mathbf{g}_i\|_2^2 = \sum_{i=N+1}^{r} \lambda_i. \tag{22}$$

By construction, the POD basis is orthonormal. Moreover, it minimizes, over all possible N-dimensional orthonormal bases $\mathbb{W} = [\mathbf{w}_1 \mid \ldots \mid \mathbf{w}_N] \in \mathbb{R}^{N_h\times N}$, the sum of the squares of the errors between each snapshot vector $\mathbf{u}_i$ and its projection onto the subspace spanned by $\mathbb{W}$.

From a practical standpoint, whenever the number n_s of snapshots is smaller than their dimension N_h it is worth computing the eigenvectors of the matrix $\mathbb{C}$, and then obtain the POD basis through the relation (20).

3.1. The relationship with singular value decomposition

The POD obtained by the method of snapshots is very closely related to the singular value decomposition (SVD) of the snapshots matrix. Since $\mathbb{S}\mathbb{S}^T$ and $\mathbb{S}^T\mathbb{S}$ are symmetric matrices, the left (resp. right) singular vectors of $\mathbb{S}$ turn out to be the eigenvectors of $\mathbb{S}\mathbb{S}^T$ (resp. $\mathbb{S}^T\mathbb{S}$). The singular values of $\mathbb{S}$ and the eigenvalues of $\mathbb{S}^T\mathbb{S}$ and $\mathbb{S}\mathbb{S}^T$ are such that

$$\sigma_i(\mathbb{S}) = \sqrt{\lambda_i(\mathbb{S}^T\mathbb{S})}, \quad i = 1, \ldots, r$$

so that if $\mathbb{S} = \mathbb{U}\Sigma\mathbb{Z}^T$, being $\mathbb{U} = [\boldsymbol{\zeta}_1 \mid \ldots \mid \boldsymbol{\zeta}_{N_h}] \in \mathbb{R}^{N_h\times N_h}$ and $\mathbb{Z} = [\boldsymbol{\psi}_1 \mid \ldots \mid \boldsymbol{\psi}_{n_s}] \in \mathbb{R}^{n_s\times n_s}$ orthogonal matrices, and $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_r) \in \mathbb{R}^{N_h\times n_s}$ with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$, then

$$\mathbb{S}\boldsymbol{\psi}_i = \sigma_i\boldsymbol{\zeta}_i \quad \text{and} \quad \mathbb{S}^T\boldsymbol{\zeta}_i = \sigma_i\boldsymbol{\psi}_i, \qquad i = 1, \ldots, r. \tag{23}$$

Hence, the POD basis $\mathbb{V}$ of dimension N, $N \leq n_S$, is realized by picking the first N columns of the matrix $\mathbb{U}$.

Proposition 3.1 also provides a practical criterion to select the minimal POD dimension $N \leq r$ such that the projection error is smaller than a desired tolerance $\varepsilon_{\mathrm{POD}} > 0$. In fact, from (22), it follows that the error in the POD basis is equal to the sum of the squares of the singular values

corresponding to the neglected POD modes. Indeed, it is sufficient to choose N as the smallest integer such that

$$I(N) = \frac{\sum_{i=1}^{N} \sigma_i^2}{\sum_{i=1}^{r} \sigma_i^2} \geq 1 - \varepsilon_{\text{POD}}^2, \tag{24}$$

that is the energy retained by the last $r - N$ modes is equal or smaller than $\varepsilon_{\text{POD}}^2$. $I(N)$ represents the percentage of energy of the snapshots captured by the first N POD modes, and it is referred to as the *relative information content* of the POD basis.

Remark 3.1. *Computing the POD basis by solving an eigenvalue problem for the matrix $\mathbb{C} = \mathbb{S}^T\mathbb{S}$ yields inaccurate results for the modes associated to small singular values. This is due to the roundoff errors introduced while constructing $\mathbb{C}$ and the fact that $\kappa(\mathbb{C}) = (\kappa(\mathbb{S}))^2$. In such cases it is recommended to construct the POD basis by means of stable algorithms for the computation of the so-called thin SVD, see e.g. [23].*

3.2. Snapshots selection: how to sample the parameter space?

The POD basis is, among all basis of dimension $N < n_s$, the one which best approximates the set of solution snapshots $\{\mathbf{g}_h(\boldsymbol{\tau}^1), \ldots, \mathbf{g}_h(\boldsymbol{\tau}^{n_s})\}$. We still have to discuss how to select the parameter samples $\Xi_s = \{\boldsymbol{\tau}^1, \ldots, \boldsymbol{\tau}^{n_s}\}$ so that the corresponding snapshots set is sufficiently representative of the solutions set $\mathcal{M}$. The key is to regard the discrete minimization problem (21) as an approximation of the continuous minimization problem (4). To this end, it is useful to introduce a suitable quadrature formula

$$\int_{\mathcal{P}} f(\boldsymbol{\tau})\, d\boldsymbol{\tau} \approx \sum_{i=1}^{n_s} w_i f(\boldsymbol{\tau}^i), \tag{25}$$

to approximate the integrals over $\mathcal{P}$, for any continuous function $f : \mathcal{P} \to \mathbb{R}$, where $w_i > 0$ and $\boldsymbol{\tau}^i$ are suitable quadrature weights and quadrature points, respectively. Then,

$$\int_{\mathcal{P}} \|\mathbf{g}_h(\boldsymbol{\tau}) - \mathbb{W}\mathbb{W}^T\mathbf{g}_h(\boldsymbol{\tau})\|_2^2\, d\boldsymbol{\tau} \approx \sum_{i=1}^{n_s} w_i \|\mathbf{g}_h(\boldsymbol{\tau}^i) - \mathbb{W}\mathbb{W}^T\mathbf{g}_h(\boldsymbol{\tau}^i)\|_2^2. \tag{26}$$

More in general, the parameter samples location are defined by a suitable quadrature formula in the parameter space. Indeed, the right-hand side of

(26) differs from (21) only by the weighting coefficients w_i. However, by defining $\mathbb{D} = \text{diag}(w_1, \dots, w_{n_s}) \in \mathbb{R}^{n_s \times n_s}$,

$$\sum_{i=1}^{n_s} w_i \, \|\mathbf{g}_h(\boldsymbol{\tau}^i) - \mathbb{W}\mathbb{W}^T \mathbf{g}_h(\boldsymbol{\tau}^i)\|_2^2 = \|\mathbb{S}\mathbb{D}^{1/2} - \mathbb{W}\mathbb{W}^T \mathbb{S}\mathbb{D}^{1/2}\|_F^2. \tag{27}$$

The POD basis associated to the snapshots set $\mathcal{M}_h^s = \{\mathbf{g}_h(\boldsymbol{\tau}^1), \dots, \mathbf{g}_h(\boldsymbol{\tau}^{n_s})\}$ is given by $\mathbb{V} = [\boldsymbol{\zeta}_1 \mid \dots \mid \boldsymbol{\zeta}_N]$, with $\boldsymbol{\zeta}_i = \frac{1}{\sigma_i^s}\mathbb{S}\mathbb{D}^{1/2}\widetilde{\boldsymbol{\psi}}_i$ and

$$\widetilde{\mathbb{S}}^T \widetilde{\mathbb{S}} \widetilde{\boldsymbol{\psi}}_i = \sigma_i^s \widetilde{\boldsymbol{\psi}}_i, \quad \text{with} \quad \widetilde{\mathbb{S}} = \mathbb{S}\mathbb{D}^{1/2}.$$

Remark 3.2. *For the sake of simplicity, we have addressed the method of snapshots by considering constant weights $w_i = 1$ when performing approximation* (26) *– see equation* (21)*. This indeed corresponds to choose $\mathbb{D} = \mathbb{I}$.*

Let us now denote by $\mathbb{V}^\infty$ the POD basis solution of the *continuous* minimization problem (17) and by $\mathbb{V}^{n_s}$ the POD basis minimizing (27); moreover we define

$$\mathcal{E}(\mathbb{W}) = \int_{\mathcal{P}} \|\mathbf{g}_h(\boldsymbol{\tau}) - \mathbb{W}\mathbb{W}^T \mathbf{g}_h(\boldsymbol{\tau})\|_2^2 \, d\boldsymbol{\tau},$$

$$\mathcal{E}^s(\mathbb{W}) = \sum_{i=1}^{n_s} w_i \|\mathbf{g}_h(\boldsymbol{\tau}^i) - \mathbb{W}\mathbb{W}^T \mathbf{g}_h(\boldsymbol{\tau}^i)\|_2^2.$$

Thanks to the optimality result (17) it follows that $\mathcal{E}(\mathbb{V}^\infty) \leq \mathcal{E}(\mathbb{V}^{n_s})$. Furthermore,

$$\mathcal{E}(\mathbb{V}^{n_s}) \leq |\mathcal{E}(\mathbb{V}^{n_s}) - \mathcal{E}^s(\mathbb{V}^{n_s})| + \mathcal{E}^s(\mathbb{V}^{n_s}) \leq E_{\text{s}} + \sum_{i=N+1}^{r} (\sigma_i^s)^2, \tag{28}$$

where E_s denotes the quadrature (or sampling) error. We remark that:

- the retained energy criterion (24) can serve as a reliable estimate of the projection error $\mathcal{E}(\mathbb{V}^{n_s})$, provided an appropriate sampling of the parameter space is performed. In fact, if E_s is smaller than the truncation error (i.e. the second term) in (28), then

$$\mathcal{E}(\mathbb{V}^{n_s}) \lesssim \sum_{i=N+1}^{r} (\sigma_i^s)^2$$

and therefore

$$\frac{\mathcal{E}(\mathbb{V}^{n_s})}{\|\mathbf{g}_h\|^2_{L^2(\mathcal{P};\mathbb{R}^{N_h})}} \lesssim \varepsilon^2_{\text{POD}};$$

- the quadrature error E_s depends on: (i) the quadrature formula chosen, (ii) the number of quadrature points, (iii) the smoothness of the integrand, (iv) the dimension P of the parameter space. Moreover, suitable estimates of E_s are available depending on the quadrature formula used;
- different sampling strategies can be employed depending on the dimension of the parameter space. While tensorial (also called full factorial) sampling is suitable for low-dimensional problems (typically for $P \leq 3$), statistical methods like random (Monte Carlo) or latin hypercube (LHS) sampling (see, e.g., [16, 32]) and sparse grids (see, e.g., [21, 10]) are recommended as soon as the dimension of the parameter space gets large.

Remark 3.3. *Note that in the case of time-dependent PDEs, we might be interested in using POD to compress the trajectory* $\mathcal{M} = \{\mathbf{u}_h(t),\ t \in [t_0, t_f]\}$*, for which* $\boldsymbol{\tau} = t$*. For this (indeed, very relevant) case, the snapshots are* naturally *selected among the solutions* $\mathbf{u}_h(t^n)$ *computed at each time step* $t^n = t_0 + n\Delta t$*.*

4. Reduced basis approximation of parametrized PDEs

Proper orthogonal decomposition provides a well-known tool for constructing a reduced order model for a PDE problem. Here we focus on the case of parametrized PDEs, although in the context of reduced-order modeling POD has been mostly used – in conjunction with (Petrov-)Galerkin projection methods – to build reduced-order models of time-dependent problems [29, 30, 35, 43]. Only recently POD has been exploited in the context of parametrized systems [15, 8, 9, 27, 49, 3]; see, e.g., [38, 50] for further details. We also highlight that former applications of POD in scientific computing were concerned with the simulation of turbulent flows [33], and the extraction of (both spatial and temporal) coherent structures appearing in turbulent flows [1, 2, 19]; further details can be found, e.g., in [7, 26, 28].

To better highlight the most relevant features of the reduced basis (RB) method, we restrict ourselves to the case of steady, linear PDE problems, although the described techniques have been generalized in many ways to several nonlinear (possibly unsteady) PDEs. Further details can be found,

e.g., in [41]. Let us denote by $\boldsymbol{\tau} = \boldsymbol{\mu} \in \mathcal{P} \subset \mathbb{R}^P$ a vector of parameters and by $\mathbf{g}_h(\boldsymbol{\tau}) = \mathbf{u}_h(\boldsymbol{\mu})$ the solution of the following high-fidelity system

$$\mathbb{A}_h(\boldsymbol{\mu})\mathbf{u}_h(\boldsymbol{\mu}) = \mathbf{f}_h(\boldsymbol{\mu}), \tag{29}$$

where $\mathbb{A}_h(\boldsymbol{\mu}) \in \mathbb{R}^{N_h \times N_h}$, $\mathbf{f}_h(\boldsymbol{\mu}) \in \mathbb{R}^{N_h}$ are the $\boldsymbol{\mu}$-dependent *stiffness* matrix and the right-hand side vector, respectively. Such a system can be obtained by discretizing problem (1) with (for instance) the finite element method, see e.g. [42] for further details.

The idea of the RB method is to generate an approximate solution to problem (29) under the form $\mathbb{V}\mathbf{u}_N(\boldsymbol{\mu})$, i.e. as a linear combination of the RB functions represented by the columns of the matrix $\mathbb{V} \in \mathbb{R}^{N_h \times N}$. To build this latter, POD can be performed starting from a set of n_s snapshots $\{\mathbf{u}_h(\boldsymbol{\mu}^1), \ldots, \mathbf{u}_h(\boldsymbol{\mu}^{n_s})\}$, corresponding to a set of n_s selected parameters $\Xi_s = \{\boldsymbol{\mu}^1, \ldots, \boldsymbol{\mu}^{n_s}\} \subset \mathcal{P}$.

Once a N-dimensional basis has been built, the coefficients $\mathbf{u}_N(\boldsymbol{\mu})$ are determined through a projection process leading to the solution of a linear system of dimension N. To begin with, let us define the residual of the high-fidelity problem

$$\mathbf{r}_h(\mathbf{v}_h; \boldsymbol{\mu}) = \mathbf{f}_h(\boldsymbol{\mu}) - \mathbb{A}_h(\boldsymbol{\mu})\mathbf{v}_h \qquad \forall \mathbf{v}_h \in \mathbb{R}^{N_h},$$

so that the vector

$$\mathbf{r}_h^N = \mathbf{r}_h(\mathbb{V}\mathbf{u}_N; \boldsymbol{\mu}) \tag{30}$$

can be regarded as the residual of the high-fidelity problem computed on the RB solution. A classical criterion to obtain the RB solution is to force the residual (30) to be orthogonal to the subspace $\mathbf{V}_N$ generated by the columns of $\mathbb{V}$

$$\mathbb{V}^T(\mathbf{f}_h(\boldsymbol{\mu}) - \mathbb{A}_h(\boldsymbol{\mu})\mathbb{V}\mathbf{u}_N(\boldsymbol{\mu})) = \mathbf{0}, \tag{31}$$

that is, to require that the *orthogonal projection* of (30) onto $\mathbf{V}_N$ is zero. This is the reason why RB methods can be regarded as projection-based methods. For any given $\boldsymbol{\mu} \in \mathcal{P}$, (29) is thus replaced by the *RB problem*

$$\mathbb{A}_N(\boldsymbol{\mu})\mathbf{u}_N(\boldsymbol{\mu}) = \mathbf{f}_N(\boldsymbol{\mu}), \tag{32}$$

where $\mathbf{u}_N(\boldsymbol{\mu}) \in \mathbb{R}^N$ is the *reduced* vector of degrees of freedom. $\mathbb{A}_N(\boldsymbol{\mu}) \in \mathbb{R}^{N \times N}$ and $\mathbf{f}_N(\boldsymbol{\mu}) \in \mathbb{R}^{N \times N}$ are related to the arrays of the high-fidelity

problem (29) through the following identities

$$\mathbb{A}_N(\boldsymbol{\mu}) = \mathbb{V}^T \mathbb{A}_h(\boldsymbol{\mu}) \mathbb{V}, \qquad \mathbf{f}_N(\boldsymbol{\mu}) = \mathbb{V}^T \mathbf{f}_h(\boldsymbol{\mu}). \tag{33}$$

More general strategies (like Petrov-Galerkin methods) require

$$\mathbb{W}^T(\mathbf{f}_h(\boldsymbol{\mu}) - \mathbb{A}_h(\boldsymbol{\mu})\mathbb{V}\mathbf{u}_N(\boldsymbol{\mu})) = \mathbf{0}$$

instead of (31), yielding

$$\mathbb{A}_N(\boldsymbol{\mu}) = \mathbb{W}^T \mathbb{A}_h(\boldsymbol{\mu}) \mathbb{V}, \qquad \mathbf{f}_N(\boldsymbol{\mu}) = \mathbb{W}^T \mathbf{f}_h(\boldsymbol{\mu})$$

instead of (33), and are obtained by imposing that the residual (30) is orthogonal to the subspace $\mathbf{W}_N$ generated by the columns of a matrix $\mathbb{W} \neq \mathbb{V}$.

The 2-norm of the error between the high-fidelity solution $\mathbf{u}_h(\boldsymbol{\mu})$ and the RB approximation $\mathbb{V}\mathbf{u}_N(\boldsymbol{\mu})$ can be bounded in terms of the residual (30) as follows

$$\|\mathbf{u}_h(\boldsymbol{\mu}) - \mathbb{V}\mathbf{u}_N(\boldsymbol{\mu})\|_2 \leq \frac{1}{\sigma_{\min}(\mathbb{A}_h(\boldsymbol{\mu}))} \|\mathbf{r}_h(\mathbb{V}\mathbf{u}_N; \boldsymbol{\mu})\|_2,$$

where $\sigma_{\min}(\mathbb{A}_h(\boldsymbol{\mu}))$ denotes the smallest singular value of $\mathbb{A}_h(\boldsymbol{\mu})$; a very similar result indeed holds for the V-norm of the error, see, e.g. [41].

For the sake of computational efficiency, both the assembling and the solution of the RB system must be performed independently of the high-fidelity problem dimension N_h. This is achieved provided the high-fidelity arrays in (29) are expressed under the *separable* form

$$\mathbb{A}_h(\boldsymbol{\mu}) = \sum_{q=1}^{Q_a} \theta_a^q(\boldsymbol{\mu}) \mathbb{A}_h^q, \qquad \mathbf{f}_h(\boldsymbol{\mu}) = \sum_{k=1}^{Q_f} \theta_f^k(\boldsymbol{\mu}) \mathbf{f}_h^k \tag{34}$$

where $\theta_a^q, \theta_f^k : \mathcal{P} \subset \mathbb{R}^p \to \mathbb{R}$ are Q_a, Q_f scalar functions and $\{\mathbb{A}_h^q\}_{q=1}^{Q_a}$, $\{\mathbf{f}_h^k\}_{k=1}^{Q_f}$ are $\boldsymbol{\mu}$-independent matrices (respectively, vectors). By virtue of this *affine parametric dependence* of $\mathbb{A}_h(\boldsymbol{\mu})$ and $\mathbf{f}_h(\boldsymbol{\mu})$, by inserting (34) in (33) we obtain

$$\mathbb{A}_N(\boldsymbol{\mu}) = \sum_{q=1}^{Q_a} \theta_a^q(\boldsymbol{\mu}) \mathbb{A}_N^q, \qquad \mathbf{f}_N(\boldsymbol{\mu}) = \sum_{k=1}^{Q_f} \theta_f^k(\boldsymbol{\mu}) \mathbf{f}_N^k.$$

The arrays $\mathbb{A}_N^q = \mathbb{V}^T \mathbb{A}_h^q \mathbb{V} \in \mathbb{R}^{N \times N}$, $\mathbf{f}_N^k = \mathbb{V}^T \mathbf{f}_h^k \in \mathbb{R}^N$ can be computed and stored once (and for all) during a possibly expensive *offline* stage, thus enabling a N_h-independent assembling of the system (33) during the *online*

stage, for any given $\boldsymbol{\mu} \in \mathcal{P}$. In this respect, EIM (or its discrete version) is essential to recover a separable expansion like (34) in all those cases where such a property is not *built-in*. A recent matrix version of the discrete EIM bypassing the interpolation of functions and acting directly on matrices has been recently proposed in [36].

5. (Discrete) EIM for parameter-dependent functions

We now turn to the approximation of a family of parameter-dependent functions $\mathcal{M} \subset C^0(\bar{\Omega})$ through the EIM, [6, 34]; see also [24]. The key idea is to determine, for any $\boldsymbol{\tau}$, an interpolant of $\mathbf{g}_h(\boldsymbol{\tau})$ on a set of basis functions built by sampling $\mathbf{g}_h$ at a suitably selected set of points in $\mathcal{P}$, instead than using predefined basis functions; and to choose interpolation points adaptively, through an iterative procedure, again depending on the function being interpolated and the domain Ω where the problem is set. The key idea is to construct an interpolant of $\mathbf{g}_h(\boldsymbol{\tau})$ built upon (i) a set of basis functions obtained by sampling $\mathbf{g}_h$ at a suitably selected set of points in $\mathcal{P}$, (rather than a predefined basis functions); (ii) a set of interpolation points adaptively chosen through an iterative procedure.

EIM indeed allows to recover a very fast rate of convergence of the approximation with respect to N, relying on *ad hoc* selection of basis functions and interpolation points. These two operations can be performed at the same time – performing a *greedy algorithm*, as in what is usually referred to EIM [6, 34] – or in two subsequent stages. In this latter case, the construction of the basis functions is performed before the selection of the interpolation points, for instance by operating a proper orthogonal decomposition on a set of n_s snapshots $\{\mathbf{g}_h(\boldsymbol{\mu}^1), \ldots, \mathbf{g}_h(\boldsymbol{\mu}^{n_s})\}$ obtained by evaluating $\mathbf{g}_h$ for a suitably chosen set of parameter values. We refer to this approach as to *discrete EIM* (DEIM, [13]). Similarly to EIM, DEIM thus approximates a nonlinear function $\mathbf{g}_h \colon \boldsymbol{\mu} \in \mathcal{P} \subset \mathbb{R}^P \to \mathbf{g}_h(\boldsymbol{\mu}) \in \mathbb{R}^{N_h}$ by projection onto a low-dimensional subspace spanned by a basis $\mathbb{Q}$,

$$\mathbf{g}_h(\boldsymbol{\mu}) \approx \mathbb{Q}\boldsymbol{\gamma}_M(\boldsymbol{\mu}), \tag{35}$$

where $\mathbb{Q} = [\boldsymbol{\rho}_1, \ldots, \boldsymbol{\rho}_M] \in \mathbb{R}^{N_h \times M}$ and $\boldsymbol{\gamma}_M(\boldsymbol{\mu}) \in \mathbb{R}^M$ is the corresponding vector of coefficients, with $M \ll N_h$. The difference lies in the construction of the basis $\mathbb{Q}$, that is obtained operating a POD on a set of snapshots

$$\mathbb{S} = [\mathbf{g}_h(\boldsymbol{\mu}^1_{DEIM}) \mid \ldots \mid \mathbf{g}_h(\boldsymbol{\mu}^{n_s}_{DEIM})],$$

instead of being embedded in the EIM greedy algorithm. Note that for both EIM and DEIM the interpolation points are iteratively selected with the same greedy algorithm. DEIM thus requires to:

(i) construct a set of snapshots obtained by sampling $\mathbf{g}_h(\boldsymbol{\mu})$ at values $\boldsymbol{\mu}^i_{DEIM}$, $i = 1, \ldots, n_s$ and apply POD to extract a M-dimensional basis $\boldsymbol{\rho}_1, \ldots, \boldsymbol{\rho}_M$, $M \leq n_s$;

(ii) select iteratively M indices $\mathfrak{I} \subset \{1, \cdots, N_h\}$, $|\mathfrak{I}| = M$ from the basis $\mathbb{Q}$ using a greedy procedure, which minimizes at each step the interpolation error over the snapshots set measured in the maximum norm. This operation is indeed the same as for the selection of the EIM magic points, see [6, 34];

(iii) given a new $\boldsymbol{\mu}$, in order to compute the coefficients vector $\boldsymbol{\gamma}(\boldsymbol{\mu})$, interpolation constraints are imposed at the M points corresponding to the selected indices, thus requiring the solution of the linear system

$$\mathbb{Q}_{\mathfrak{I}}\, \boldsymbol{\gamma}_M(\boldsymbol{\mu}) = \mathbf{g}_{\mathfrak{I}}(\boldsymbol{\mu}), \tag{36}$$

where $\mathbb{Q}_{\mathfrak{I}} \in \mathbb{R}^{M \times M}$ is formed by the $\mathfrak{I}$ rows of $\mathbb{Q}$ and $\mathbf{g}_{\mathfrak{I}}(\boldsymbol{\mu}) \in \mathbb{R}^M$ is obtained by evaluating $\mathbf{g}_h(\boldsymbol{\mu})$ in the same selected indices.

We point out that the 2-norm of the error between $\mathbf{g}_h$ and its DEIM approximation $\mathbf{g}_M$ can be bounded as

$$\|\mathbf{g}_h(\boldsymbol{\mu}) - \mathbb{Q}\boldsymbol{\gamma}_M(\boldsymbol{\mu})\|_2 \leq \|\mathbb{Q}_{\mathfrak{I}}^{-1}\|_2 \, \|(\mathbb{I} - \mathbb{Q}\mathbb{Q}^T)\mathbf{g}_h(\boldsymbol{\mu})\|_2, \tag{37}$$

with

$$\|(\mathbb{I} - \mathbb{Q}\mathbb{Q}^T)\mathbf{g}_h(\boldsymbol{\mu})\|_2 \approx \sigma_{M+1}, \tag{38}$$

being σ_{M+1} the first discarded singular value of the matrix $\mathbb{S}$ when selecting M basis through the POD procedure. This approximation holds for any $\boldsymbol{\mu} \in \mathcal{P}$ provided a suitable sampling in the parameter space has been carried out to build the snapshot matrix $\mathbb{S}$; see e.g. [14] for further details.

6. Karhunen-Loève expansion for stochastic fields

In the last decade, the development of numerical techniques for the solution of PDEs depending on random input data has undergone an incredible expansion. This has gone with the widespread of applications in the huge realm of *uncertainty quantification* (UQ), including e.g. parameter estimation, uncertainty propagation, model calibration, and many other aspects.

See, e.g., [31, 25] for further details. In all these cases, *probabilistic methods* involve a statistical characterization of uncertain inputs. In the simplest case, the mathematical model depends on a set of (say, physical) uncertain parameters $\boldsymbol{\tau} = \boldsymbol{\psi} \in \mathbb{R}^Q$ that may be represented as a random vector $\boldsymbol{\psi}(\omega) = (\psi_1(\omega), \ldots, \psi_Q(\omega)) \in \mathbb{R}^Q$ for any outcome $\omega \in \Theta$, with a given joint probability distribution; here Θ denotes the outcome set and $\boldsymbol{\psi} : \Theta \to \mathbb{R}^Q$ denotes a generic Q-dimensional random vector.

In more complex situations, input data might be spatially-dependent (or time-dependent) random quantities, meaning that they may vary randomly from one point of the physical domain Ω to another (or from one time instant to another), so that input uncertainty has to be described in terms of random fields. Focusing on the spatially-dependent, scalar case, a random field $\nu = \nu(\mathbf{x}; \omega)$ is a stochastic process taking values in $\mathbb{R}$ for any $\mathbf{x} \in \Omega$, for any outcome $\omega \in \Theta$, meaning that *(i)* for a fixed point $\mathbf{x} \in \Omega$, $\nu(\mathbf{x}; \cdot)$ is a random variable over Θ; *(ii)* for a fixed $\omega \in \Theta$, $\nu(\cdot; \omega)$ is a realization of the random field in Ω. A more convenient finite dimensional representation of a random field in terms of Q (possibly uncorrelated) random variables that indeed play the role of parameters can be obtained by means of a Karhunen-Loève (KL) expansion.

Let us denote by (Θ, Σ, P) a complete probability space where Θ is the outcome set, Σ is a σ-algebra of events and $P : \Sigma \to [0, 1]$ is a probability measure. Let us denote by $\boldsymbol{\nu} : \Theta \to \mathbb{R}^{N_h}$ a random field on the probability space (Θ, Σ, P) with finite second order moments, that is, $\boldsymbol{\nu} \in L^2_P(\Theta; \mathbb{R}^{N_h})$; alternatively, $\boldsymbol{\nu}$ is said to be a second-order stochastic process; here

$$L^2_P(\Theta; \mathbb{R}^{N_h}) = \left\{ \mathbf{v} : \Theta \to \mathbb{R}^{N_h} : \mathbf{v} \text{ measurable}, \int_\Theta \|\mathbf{v}(\omega)\|_2^2 \, dP(\omega) < +\infty \right\}.$$

The subscript $_P$ indicates that the integrals are weighted with respect to the probability measure P. The expectation of the random field $\boldsymbol{\nu}(\omega)$ is

$$\overline{\boldsymbol{\nu}} = \mathbb{E}[\boldsymbol{\nu}(\cdot)] = \int_\Theta \boldsymbol{\nu}(\omega) \, dP(\omega), \tag{39}$$

while its covariance matrix $\mathbb{K} \in \mathbb{R}^{N_h \times N_h}$ is given by

$$\begin{aligned} (\mathbb{K})_{ij} = &\ \mathbb{E}\left[(\boldsymbol{\nu}_i(\cdot) - \overline{\boldsymbol{\nu}}_i(\cdot))(\boldsymbol{\nu}_j(\cdot) - \overline{\boldsymbol{\nu}}_j(\cdot)) \right] \\ = &\ \mathbb{E}\left[(\nu(\cdot, \mathbf{x}_i) - \overline{\nu}(\cdot, \mathbf{x}_i))(\nu(\cdot, \mathbf{x}_j) - \overline{\nu}(\cdot, \mathbf{x}_j)) \right]. \end{aligned}$$

Equivalently,

$$\mathbb{K} = \int_\Theta (\boldsymbol{\nu}(\omega) - \overline{\boldsymbol{\nu}})(\boldsymbol{\nu}(\omega) - \overline{\boldsymbol{\nu}})^T \, dP(\omega).$$

The covariance is a two-point statistics that models the spatial correlation of the stochastic field between two points $\mathbf{x}_i, \mathbf{x}_j \in \Omega_h$. Moreover, the covariance matrix is symmetric and semi-positive definite; indeed,

$$\mathbf{x}^T \mathbb{K}\, \mathbf{x} = \int_\Theta \left(\mathbf{x}^T(\boldsymbol{\nu}(\omega) - \overline{\boldsymbol{\nu}})\right)^2 dP(\omega) \geq 0 \qquad \forall \mathbf{x} \in \mathbb{R}^{N_h}.$$

Therefore, $\mathbb{K}$ has N_h non-negative real eigenvalues $\lambda_1 \geq \cdots \geq \lambda_{N_h} \geq 0$ and corresponding eigenvectors $\boldsymbol{\zeta}_i \in \mathbb{R}^{N_h}$ such that

$$\mathbb{K}\, \boldsymbol{\zeta}_i = \lambda_i \boldsymbol{\zeta}_i, \qquad i = 1, \ldots, N_h.$$

Since the $\boldsymbol{\zeta}_i$'s form an orthonormal basis for $\mathbb{R}^{N_h}$, $\boldsymbol{\nu}(\omega)$ admits the following expansion

$$\boldsymbol{\nu}(\omega) = \overline{\boldsymbol{\nu}} + \sum_{i=1}^{N_h} \sqrt{\lambda_i} \boldsymbol{\zeta}_i \psi_i(\omega), \tag{40}$$

where the random variables $\psi_i(\omega)$ are defined as

$$\psi_i(\omega) = \frac{1}{\sqrt{\lambda_i}} (\boldsymbol{\nu}(\omega) - \overline{\boldsymbol{\nu}}, \boldsymbol{\zeta}_i)_2.$$

These latter are uncorrelated random variables with zero mean and unit variance:

$$\mathbb{E}[\psi_i(\cdot)] = \frac{1}{\sqrt{\lambda_i}} \int_\Theta (\boldsymbol{\nu}(\omega) - \overline{\boldsymbol{\nu}}, \boldsymbol{\zeta}_i)_2 \, dP(\omega) = \frac{1}{\sqrt{\lambda_i}} (\overline{\boldsymbol{\nu}} - \overline{\boldsymbol{\nu}}, \boldsymbol{\zeta}_i)_2 = 0,$$

$$\begin{aligned} \mathbb{E}[\psi_i(\cdot)\psi_j(\cdot)] &= \frac{1}{\sqrt{\lambda_i \lambda_j}} \int_\Theta \boldsymbol{\zeta}_i^T (\boldsymbol{\nu} - \overline{\boldsymbol{\nu}})(\boldsymbol{\nu} - \overline{\boldsymbol{\nu}})^T \boldsymbol{\zeta}_j \, dP(\omega) = \frac{1}{\sqrt{\lambda_i \lambda_j}} \boldsymbol{\zeta}_i^T \mathbb{K}\, \boldsymbol{\zeta}_j \\ &= \frac{1}{\sqrt{\lambda_i \lambda_j}} \boldsymbol{\zeta}_i^T \lambda_j\, \boldsymbol{\zeta}_j = \frac{\lambda_j}{\sqrt{\lambda_i \lambda_j}} \delta_{ij}. \end{aligned}$$

The expression (40) is referred to as the Karhunen-Loève (KL) expansion of the random field $\boldsymbol{\nu}(\omega)$. It is a series expansion involving deterministic functions (the eigenvectors $\boldsymbol{\zeta}_i$) and random variables (the ψ_i). The deterministic functions are fixed by the form of the covariance of the random field, whereas the joint probability law of the ψ_i's remains unknown in the absence of information other than the second-order properties of the process. A convenient (and indeed very common) situation arises when the random field

is Gaussian, leading to significant simplifications. Indeed, the KL expansion of a Gaussian field involves random variables $\psi_i(\omega)$ which are not only uncorrelated, but also independent. The latter is a useful property in practical applications, since it makes the KL expansions very useful for generating samples of Gaussian stochastic processes; for instance, is is extensively used in the stochastic Galerkin and collocation methods for SPDEs.

The most remarkable property of the KL decomposition is its optimality in the mean square sense. Indeed, when considering the truncated KL expansion $\boldsymbol{\nu}_Q(\omega)$ consisting of the first Q modes,

$$\boldsymbol{\nu}_Q(\omega) = \overline{\boldsymbol{\nu}} + \sum_{i=1}^{Q} \sqrt{\lambda_i}\, \boldsymbol{\zeta}_i\, \psi_i(\omega) \tag{41}$$

the following result holds (see Proposition 2.1):

$$\min_{\hat{\boldsymbol{\nu}}} \mathbb{E}\left[\|\boldsymbol{\nu}(\cdot) - \hat{\boldsymbol{\nu}}(\cdot)\|_2\right] = \mathbb{E}\left[\|\boldsymbol{\nu}(\cdot) - \boldsymbol{\nu}_Q(\cdot)\|_2\right],$$

where $\hat{\boldsymbol{\nu}}$ denotes any approximation of $\boldsymbol{\nu}(\cdot)$ in the form of a series expansion involving a finite number of random variables. Moreover, the mean square truncation error decreases monotonically with Q, at a rate that depends on the decay of the spectrum of the matrix $\mathbb{K}$.

7. A numerical example

In this section we solve a parametrized elliptic PDE describing a stationary advection-diffusion problem, in which *(i)* the diffusion coefficient is a stochastic field, and *(ii)* the parametrized source term features a nonaffine parametric dependence. We show how POD can be exploited to obtain a low-dimensional representation of these two quantities, as well as of the solution of the problem. In spite of its simplicity, this example allows us to highlight those (indeed, far from trivial) aspects involved in the construction of a reduced-order model in the case of a moderately large number of parameters and a complex parametric dependence.

For the sake of illustration, we consider a simple mass transfer problem describing e.g. the behavior of pollutant emissions released by an industrial chimney into the atmosphere. The evolution of the pollutant concentration can be modeled by the following advection-diffusion-reaction equation on the (parameter-independent) domain $\Omega = (0,1) \times (0,0.5)$,

$$\begin{cases} -\mathrm{div}(0.25\nu(\mathbf{x};\omega)\nabla u + \boldsymbol{b}(\mu_3)\cdot\nabla u + a_0 u = f(\mathbf{x};\boldsymbol{\mu}) & \text{in } \Omega \\ \qquad\qquad \nu(\mathbf{x};\omega)\nabla u \cdot \boldsymbol{n} = 0 & \text{on } \Gamma_N = \partial\Omega, \end{cases} \tag{42}$$

where $u = u(\boldsymbol{\mu})$ is the unknown pollutant concentration, $a_0 > 0$ represents the intensity of reaction processes,

$$f(\mathbf{x}; \boldsymbol{\mu}) = \exp\left(-\frac{(x_1 - \mu_1)^2 + (x_2 - \mu_2)^2}{0.2^2}\right) \tag{43}$$

describes the pollutant emission, characterized in terms of its position (μ_1, μ_2), and $\boldsymbol{b}(\mu_3)$ is a (constant in space) advection field,

$$\boldsymbol{b}(\mu_3) = [\cos(\mu_3) \ \ \sin(\mu_3)]^T \tag{44}$$

representing e.g. the wind speed (see, e.g., [20] for a similar application). The molecular diffusivity of the chemical species is modeled as a random field, in order to account for its spatial heterogeneity and lack of precise knowledge. In particular, we assume that:

- $\nu(\mathbf{x}, \omega) : \overline{\Omega} \times \Theta \to \mathbb{R}$ is a Gaussian field, given in our case by $\nu(\mathbf{x}, \cdot) \sim \mathcal{N}(m, \sigma^2)$ $\forall \mathbf{x} \in \Omega$, where $\mathcal{N}(m, \sigma^2)$ denotes a Gaussian probability distribution with expected value m and variance σ^2. For the case at hand, we set $\sigma^2 = 1$ and $m = 4$, to ensure that the diffusivity takes positive values almost everywhere[3];
- $\nu(\mathbf{x}, \omega) : \overline{\Omega} \times \Theta \to \mathbb{R}$ is such that for $\mathbf{x}, \mathbf{x}' \in \Omega$ the covariance function $C_\nu(\mathbf{x}, \mathbf{x}') = Cov[\nu(\mathbf{x}, \cdot)\nu(\mathbf{x}', \cdot)]$ only depends on the distance $\|\mathbf{x} - \mathbf{x}'\|$ – that is, ν is a second-order, isotropic stationary field – and $C_\nu(\mathbf{x}, \mathbf{x}') = C_\nu(\|\mathbf{x} - \mathbf{x}'\|)$ is a Lipschitz continuous, positive definite function. Here we consider a simple Gaussian covariance function,

$$C_\nu(\|\mathbf{x} - \mathbf{x}'\|) = \sigma^2 \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{L_c}\right)$$

  where $L_c > 0$ denotes the correlation length; for the case at hand, $\sigma^2 = 1$ and $L_c^2 = 0.05$. Other choices for the correlation function are possible, depending on the application.

For the sake of simplicity, here we assume that randomness only affects the diffusion coefficient and not the forcing term f. However, this is the

[3] Clearly, this is not rigorous, since the support of the normal distribution is $\mathbb{R}$. However, by taking $m = 4$, we have that $S(\mathbf{x}) = \{\omega \in \Theta : \nu(\mathbf{x}, \omega) > 0\}$ is such that $\mathbb{P}(S(\mathbf{x})) = 0.9999683$ for any $\mathbf{x} \in \Omega$. A check to avoid realizations of the Gaussian field yielding to negative pointwise values is performed. To ensure the positivity of the diffusion coefficient, a better option would be to define $\tilde{\nu}(\mathbf{x}, \omega) = e^{\nu(\mathbf{x}, \omega)}$, with $\nu(\mathbf{x}, \cdot)$ denoting a Gaussian random variable.

most interesting case, since the solution u depends nonlinearly on ν and linearly on f and the boundary data. Other similar situations are related to groundwater flows modeling in hydrology, where the diffusion coefficient describes the permeability of a porous medium, the subsoil.

We are interested in evaluating suitable statistics of the PDE solution, such as its average or moments, as well as moments of other quantities of interest expressed by proper functionals of the solution. We will not further elaborate on numerical methods to deal with these issues; we only mention that the simplest approach would consist in employing a Monte Carlo method combined with a deterministic solver of the PDE, thus entailing: *(i)* the generation of many samples of the input random field; *(ii)* correspondingly, the approximation of the PDE solution for each sample and, finally, *(iii)* the evaluation of the expected value of the quantity of interest by sample average. Since a huge number of samples is required when employing Monte Carlo methods, replacing a high-fidelity solver with a reduced-order approximation such as the one based on POD would yield remarkable overall computational speedups.

For the case at hand, this requires to *parametrize* the random field by expressing it under the form (41), and to express the parametric dependence of the source term so that the resulting right-hand side vector of the linear system features a separable form. It is worth noting that a similar *separable* form is then obtained, by applying the KL expansion, on the diffusion coefficient as well[4] and that the number Q of selected modes (see (41)) yields the same amount of additional parameters $\mu_4, \ldots, \mu_{3+Q}$ on the problem (42).

Here we report the numerical results obtained by applying the three main techniques discussed in the paper – the RB method, the discrete EIM and the KL expansion – on problem (42). As high-fidelity approximation we consider the finite element method with linear finite elements over a triangular mesh yielding $N_h = 5305$ degrees of freedom. We rely on the `redbKIT` package [41] for performing EIM and POD, whereas we take advantage of the `random field simulation` package [17] for performing the KL expansion.

First, we perform the KL expansion of the random field $\nu(\mathbf{x}, \omega)$; see Sect. 6. Note that for the case at hand the spatial correlation $\mathbb{K}$ is a prescribed matrix of dimension $N_h \times N_h$, so that we can directly compute its eigendecomposition; the first $Q = 9$ modes $\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_9$ of the KL expansion of the random field are reported in Fig. 1, whereas the eigenvalues of the

[4]Note that if the diffusion coefficient is modeled through a lognormal (instead than a Gaussian) field $\tilde{\nu} = e^{\nu}$, the affine parametric dependence of the stiffness matrix has to be restored by further applying DEIM on the function $\tilde{\nu}$.

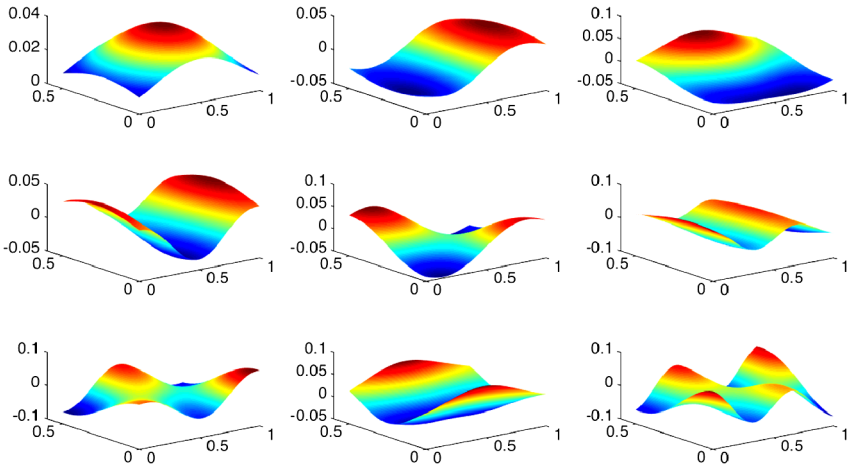

Figure 1: KL expansion of the diffusivity field: first $Q = 9$ modes $\boldsymbol{\zeta}_1, \ldots, \boldsymbol{\zeta}_9$.

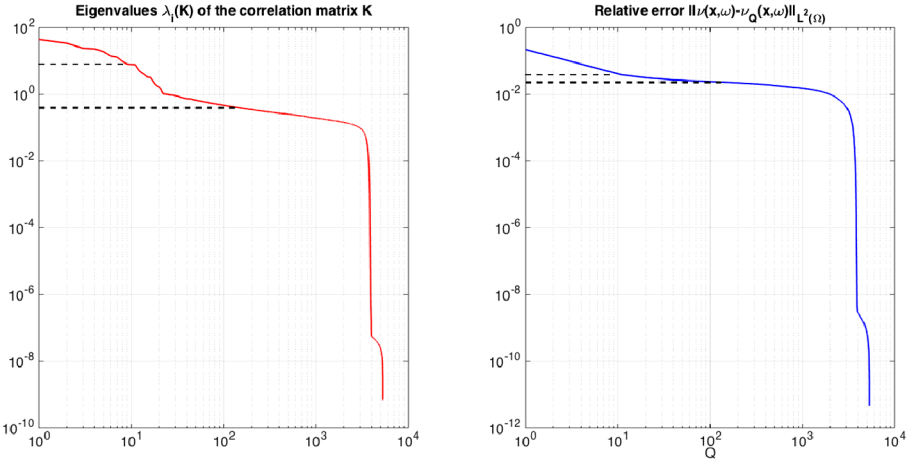

Figure 2: KL expansion of the diffusivity field. Left: eigenvalues of the spatial correlation matrix $\mathbb{K}$. Right: relative $L^2(\Omega)$ errors $\|\nu(\mathbf{x},\omega) - \nu_Q(\mathbf{x},\omega)\|_{L^2(\Omega)}/\|\nu(\mathbf{x},\omega)\|_{L^2(\Omega)}$ for increasing values of Q, averaged over a sample of 50 randomly selected realizations. The levels of accuracy corresponding to $Q = 9$ and $Q = 150$ are reported in both plots.

matrix $\mathbb{K}$ are reported in Fig. 2, left. Moreover, in our case the random variables $\psi_1, \ldots, \psi_Q$ appearing in (41) are independent, following a standard Gaussian distribution $\mathcal{N}(0, 1)$.

Table 1: KL expansion of the diffusivity field. Relative $L^2(\Omega)$ errors $\|\nu(\mathbf{x},\omega) - \nu_Q(\mathbf{x},\omega)\|_{L^2(\Omega)}/\|\nu(\mathbf{x},\omega)\|_{L^2(\Omega)}$ for $Q = 3, 6, 9, 150$, averaged over a sample of 50 random realizations

$Q = 3$	$Q = 6$	$Q = 9$	$Q = 150$
0.1334	0.0856	0.0524	0.0251

As a matter of fact, the additional parameters $\mu_4, \ldots, \mu_{3+Q}$ are sampled from standard Gaussian distributions; nevertheless, for the sake of the subsequent sampling of the (augmented) parameter space $\mathcal{P} \subset \mathbb{R}^{Q+3}$, in view of the construction of a reduced model for the parametrized PDE on a closed and compact parameter space, we restrict ourselves to retain only those samples such that $-4 \leq \mu_i \leq 4$, $i = 4, \ldots, Q+3$. Hereon, we consider the cases $Q = 0, 3, 6, 9$, in order to deal with *(i)* a constant diffusivity field, and *(ii)* random fields obtained with KL expansions of increasing accuracy, yet keeping the dimension of the parameter space relatively small; this latter ranges from $P = 3$ to $P = 12$ in the cases $Q = 0$, $Q = 9$, respectively.

The accuracy of the KL expansion for increasing values of Q has been evaluated over a sample of 50 realizations; the average relative $L^2(\Omega)$ errors $\|\nu(\mathbf{x},\omega) - \nu_Q(\mathbf{x},\omega)\|_{L^2(\Omega)}/\|\nu(\mathbf{x},\omega)\|_{L^2(\Omega)}$ between the random field and its KL expansion are then reported in Fig. 2, right. Obviously, the error and the computed eigenvalues of $\mathbb{K}$ decays at a similar rate; for $Q = 9$ the relative error is about 0.05, whereas it falls below 0.025 for $Q = 150$; see also Table 1. Some realizations of the random field ν and of its KL expansions obtained for $Q = 3, 6, 9, 150$ are reported in Fig. 3.

We now turn to the approximation of the function $f(\mathbf{x};\boldsymbol{\mu})$ defined in (43) representing the distributed source for the advection-diffusion-reaction problem (42). The presence of the exponential function yields a nonaffine dependence of $f(\mathbf{x};\boldsymbol{\mu})$ on the parameters μ_1, μ_2, thus requiring to rely on DEIM to approximate the family of parameter-dependent functions

$$\mathcal{M} = \{f(\mathbf{x};\mu_1,\mu_2)\,,\ (\mu_1,\mu_2) \in [0.3, 0.7] \times [0.15, 0.35]\};$$

we recall that $\mu_3, \mu_4, \ldots, \mu_{Q+3}$ do not affect $f(\mathbf{x};\boldsymbol{\mu})$.

Following the procedure detailed in Sect. 5, we select $n_s = 1000$ parameter points obtained by latin hypercube sampling, get the corresponding snapshots by evaluating f for the selected parameter points, perform POD on this set of snapshots and select a basis of dimension $M = 62$ by imposing a tolerance $\varepsilon_{POD} = 10^{-4}$ in the criterion (24). Afterwards, we select a set $\mathfrak{I} \subset \{1, \cdots, N_h\}$ of indices, $|\mathfrak{I}| = 62$, by running a greedy algorithm to determine the interpolation points, obtaining the distribution of points

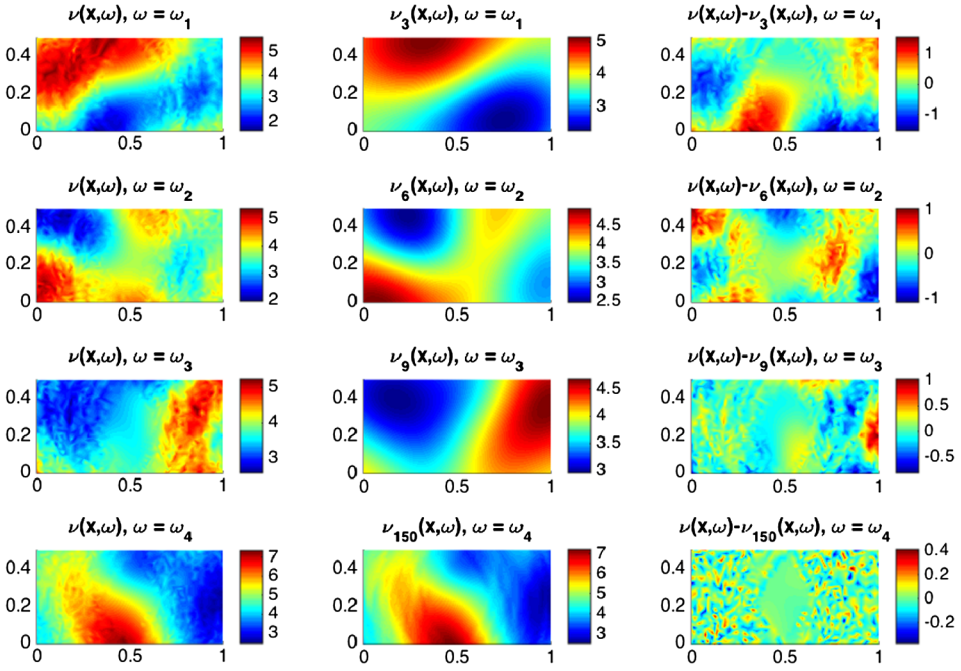

Figure 3: KL expansion of the diffusivity field. Different realizations of $\nu(\mathbf{x}, \omega)$, $\omega = \omega_i$, $i = 1, \ldots, 4$ randomly selected, with increasing levels $Q = 3, 6, 9, 150$ (from top to bottom). Left: realization of the random field; middle: KL expansion with Q terms; right: discrepancy $\nu(\mathbf{x}, \omega_i) - \nu_Q(\mathbf{x}, \omega_i)$.

reported in Fig. 5. Note that the construction of an affine approximation $\mathbf{f}_M(\boldsymbol{\mu})$ is independent on the already performed KL expansion on the random diffusivity field. We report the decay of both the singular values of the snapshot matrix $\mathbb{S}$ and the errors $\|\mathbf{f}(\boldsymbol{\mu}) - \mathbf{f}_M(\boldsymbol{\mu})\|_2$ evaluated on a test sample of 50 randomly selected couples (μ_1, μ_2) in Fig. 4; some instances of affine approximations f_M for increasing values of $M = 10, 20, 60$, and the corresponding errors $\mathbf{f}(\boldsymbol{\mu}) - \mathbf{f}_M(\boldsymbol{\mu})$ are then shown in Fig. 6.

The decay of both the singular values of $\mathbb{S}$ and of the approximation errors $\|\mathbf{f}(\boldsymbol{\mu}) - \mathbf{f}_M(\boldsymbol{\mu})\|_2$, is fast, even if a moderately large number of terms M is necessary to keep the approximation errors under a threshold of about 10^{-4}. This is due to the fact that the dependence on the parameters μ_1, μ_2 of the function f localizes the peak of the function on the spatial domain, thus making the approximation of any $f(\mathbf{x}; \boldsymbol{\mu})$ by a linear combination of elements extracted from $\mathcal{M}$ intrinsically difficult.

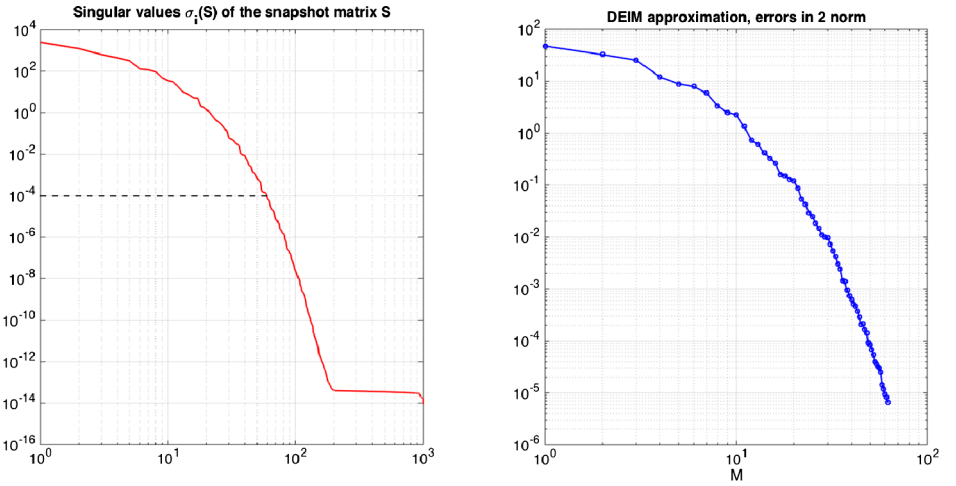

Figure 4: DEIM approximation of source term. Left: singular values of the snapshot matrix $\mathbb{S}$ obtained by selecting $n_s = 1000$ points $(\mu_1, \mu_2) \in [0.3, 0.7] \times [0.15, 0.35]$. Right: errors $\|\mathbf{f}(\boldsymbol{\mu}) - \mathbf{f}_M(\boldsymbol{\mu})\|_2$ for increasing values of $M = 1, \ldots, 62$, averaged over a sample of 50 randomly chosen parameter points.

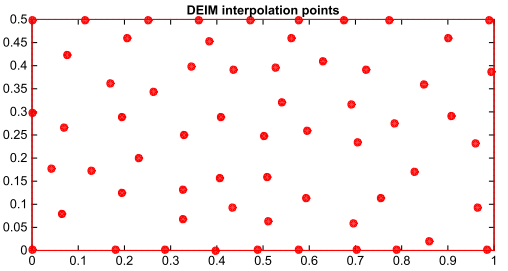

Figure 5: DEIM approximation of the source term: interpolation points.

Finally, we build a reduced-order model for efficiently evaluating the solution of problem (42) relying on the Galerkin-RB method, thus resulting in a KL-DEIM-G-RB approximation. We employ POD to construct the RB space, once the diffusivity field and the distributed source have been approximated by using KL and DEIM, respectively, in order to recover an affine parametric dependence. We perform the construction of a RB space in four different cases, taking into account a constant diffusivity field (case $Q = 0$ KL modes, yielding $P = 3$ parameters) or random diffusivity fields (cases $Q = 3, 6, 9$ KL modes, yielding $P = 6, 9, 12$ parameters, respectively).

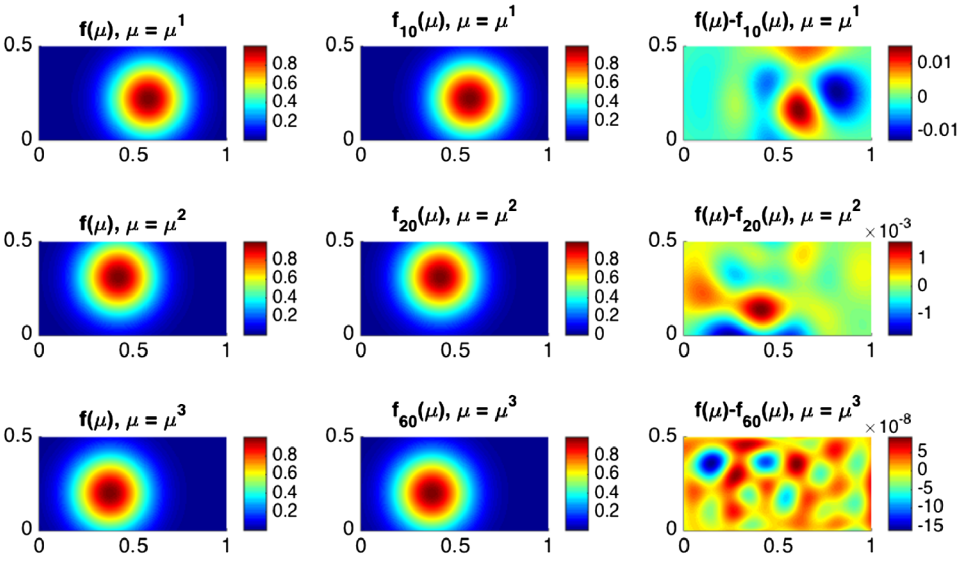

Figure 6: DEIM approximation of the source term. Evaluation of $\mathbf{f}(\boldsymbol{\mu}^i)$ and $\mathbf{f}_M(\boldsymbol{\mu}^i)$ for $\boldsymbol{\mu} = \boldsymbol{\mu}_i$, $i = 1, \ldots, 3$ randomly selected, with increasing DEIM dimension $M = 10, 20, 60$ (from top to bottom). Left: evaluation of $\mathbf{f}$; middle: DEIM approximation $\mathbf{f}_M$ with M terms; right: error $\mathbf{f}(\boldsymbol{\mu}^i) - \mathbf{f}_M(\boldsymbol{\mu}^i)$.

In all these cases we rely on the procedure detailed in Sect. 4: starting from n_s snapshots of the affine approximation of the high-fidelity problem, we then select N modes by imposing a tolerance $\varepsilon_{POD} = 10^{-6}$ on the POD algorithm. The number of snapshots is increasing with the dimension of Q, ranging from $n_s = 1,000$ to $n_s = 10,000$ when passing from $Q = 0$ to $Q = 9$, to take into account the increasing parameter dimension $P = Q+3$. Because of the variability induced by the increasing number of KL modes, we obtain RB spaces of increasing dimension N by keeping ε_{POD} constant, ranging from $N = 42$ to $N = 1357$ when passing from $Q = 0$ to $Q = 9$.

The singular values of the snapshot matrices obtained in the four cases $Q = 0, 3, 6, 9$, as well as the decay of the error $\|\mathbf{u}_h(\boldsymbol{\mu}) - \mathbb{V}\mathbf{u}_N(\boldsymbol{\mu})\|_V$ with respect to the dimension N of the RB space, are reported in Fig. 7. The decay rate of both these quantities indicates that the problem becomes more hardly reducible when the dimension $P = Q + 3$ of the parameter space increases.

We report the RB approximation $u_N(\boldsymbol{\mu})$ of the problem (42), as well as the error $u_h(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})$ between the FE and the RB approximation, for different values of $\boldsymbol{\mu}$ and Q, keeping N and M fixed, in Fig. 8. Note that by

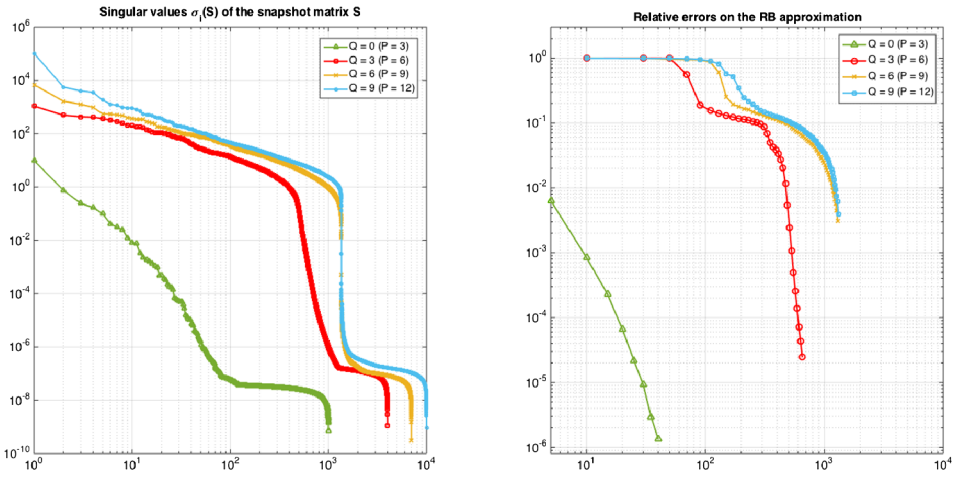

Figure 7: RB approximation of the PDE solution. Left: singular values of the snapshot matrix $\mathbb{S}$ obtained by selecting $n_s = 1000, 4000, 7000, 10000$ snapshots for $Q = 0, 3, 6, 9$ KL modes. Right: errors $\|\mathbf{u}_h(\boldsymbol{\mu}) - \mathbb{V}\mathbf{u}_N(\boldsymbol{\mu})\|_V$ for increasing values of N, averaged over a sample of 50 randomly chosen parameter points. In this case $M = 62$ DEIM terms have been selected.

taking $a_0 = 1$ and the advection field (44) – with $\mu_3 \in [0, 180°]$ – advection or reaction are not dominating the diffusion, thus yielding a non-negligible role to the presence of a random diffusivity field. This can also be seen by the decay of the singular values of the correlation matrices $\mathbb{S}$ built over the solution snapshots, showing a slower decay for larger values of Q. On the other hand, limiting ourselves to (relatively) few KL modes is crucial to limit the size of the parameters space; in turn, this yields less fluctuating random fields showing a higher spatial regularity.

Similarly, we report in Fig. 9 the RB approximations obtained for different RB space dimensions N, keeping both Q and M fixed, for randomly selected parameter values. Obviously, the accuracy of the RB approximation improves when the the dimension of the RB space increases, as predicted by the decay of the singular values of the snapshot matrix (see Fig. 7).

Finally, we investigate the role of the DEIM approximation on the accuracy of the RB approximation. We report in Fig. 10 the errors between the KL-DEIM-G-RB solution $u_N(\boldsymbol{\mu})$ and the FE approximation $u_h(\boldsymbol{\mu})$ involving the same KL expansion of the random field, evaluated for different RB space dimensions N and number of DEIM terms M, averaged over a sample of 50 parameter values randomly selected. From this error analysis it follows that a sufficiently large number of both DEIM terms and RB functions

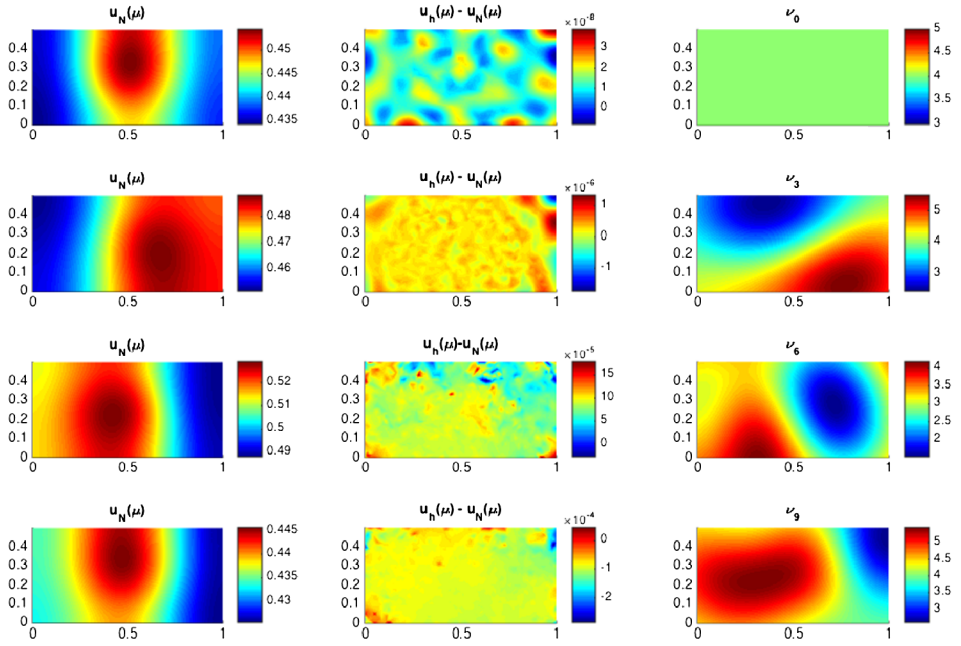

Figure 8: RB approximation of the PDE solution. Left: RB approximation $u_N(\boldsymbol{\mu})$; center: error $u_h(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})$ between the FE and the RB approximation; right: KL expansion of the diffusivity field. From top to bottom $Q = 0, 3, 6, 9$ KL modes have been considered. Different parameter values have been selected for each case.

has to be considered to achieve a good accuracy with respect to the high-fidelity approximation. Moreover, the DEIM accuracy must be higher than the RB accuracy to avoid that the DEIM approximation errors dominate – as, e.g., in the case it happens, for instance, when only $M = 10, 20$ terms are selected.

We conclude this section by highlighting some remarkable computational facts; see Tab. 2 for further details. Since the size of the high-fidelity problem is quite small (only $N_h = 5375$ degrees of freedom), the CPU time[5] required to evaluate the KL expansion of the random field and the DEIM approximation of the source term is of 28.79 s and 160.12 s, respectively. These operations have to be performed only once, as well as the assembling of the

[5]All the computations have been performed on a single core of a 2,8 GHz Intel Core i5 processor, with 16 GB RAM. All the linear systems (for both FE and RB problems) are solved using the direct solver provided by Matlab.

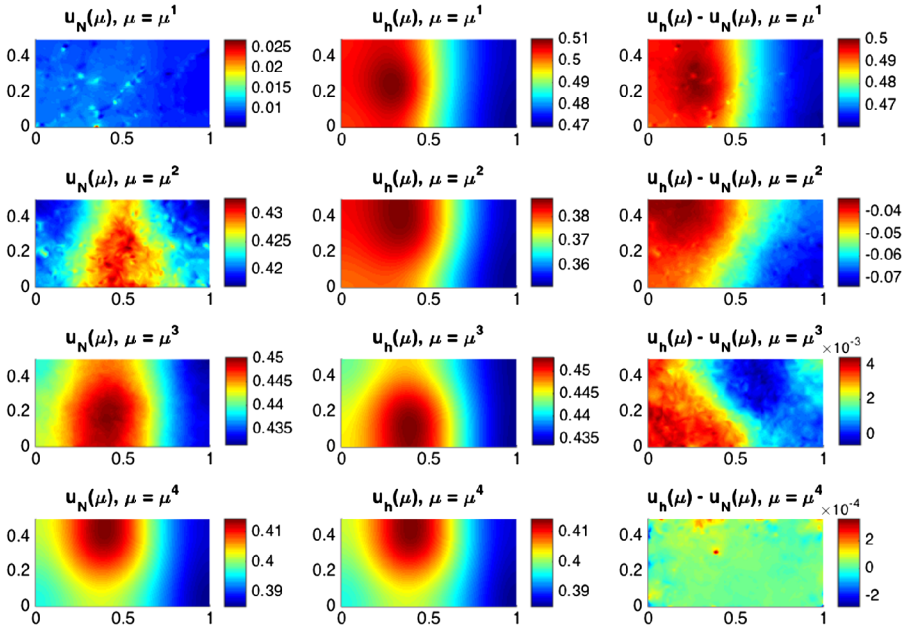

Figure 9: RB approximation of the PDE solution. Left: RB approximation $u_N(\boldsymbol{\mu})$; center: FE approximation $u_h(\boldsymbol{\mu})$; right: error $u_h(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})$ between the FE and the RB approximation. The same number $Q = 9$ of KL modes has been selected. A RB space of dimension $N = 50, 250, 750, 1357$ (from top to bottom) has been considered.

parameter-independent FE arrays $\mathbb{A}_h^q$, $q = 1, \ldots, Q_a$ and $\mathbf{f}_h^k$, $k = 1, \ldots, Q_f$, which in our case requires 13.34 s. Once this pre-processing stage is performed, we construct the RB space; the CPU time $t_{FE}^{offline}$ required by the construction of the RB space (and related arrays) increases for growing P, because of the larger amount of snapshots to be computed and the larger SVD decomposition (23) to be determined in order to find the POD basis functions. The CPU time t_{RB}^{online} entailed by the solution of the RB problem (32) also increases for larger RB dimensions N, thus yielding smaller and smaller computational speedups. Nevertheless, in any case the RB problem requires between 0.0025 and 0.607 seconds to be solved, whereas the corresponding FE problem takes about 1 second to be solved. When moving to problems on a larger scale regarding the dimension of the high-fidelity approximation, the computational speedup shall become even larger, thus making the use of reduced order models essential to perform many-query problems.

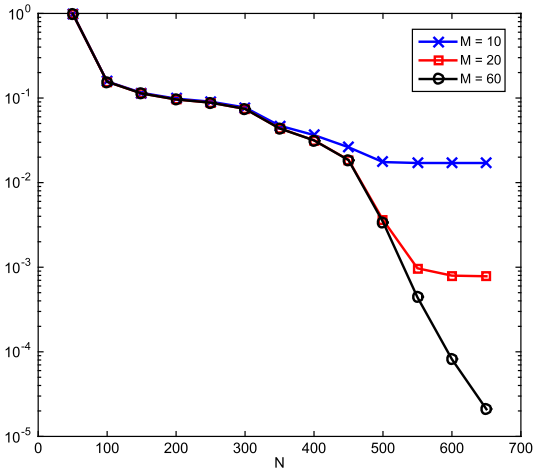

Figure 10: RB approximation of the PDE solution. Online convergence (with respect to N) of the error between the KL-DEIM-G-RB and the FE approximation for different values of M in the case $Q = 3$, averaged over a sample of 50 randomly chosen parameter points.

Table 2: Computational details. RB spaces have been built using POD with a tolerance $\varepsilon_{POD} = 10^{-6}$, starting from a high-fidelity FE approximation of dimension $N_h = 5375$. In addition to $t_{FE}^{offline}$, also the time to evaluate the KL expansion of the random field (28.79 s), to perform DEIM approximation of the source field (160.12 s) and to assemble the high-fidelity arrays (13.34 s) must taken into account. The number Q_f of affine terms at the right-hand side refers to the case where all the $M = 62$ DEIM terms are considered

	$Q = 0$	$Q = 3$	$Q = 6$	$Q = 9$
# parameters P	3	6	9	12
# affine terms lhs Q_a	4	7	10	13
# affine terms rhs Q_f	62	62	62	62
# POD snapshots n_s	1000	4000	7000	10000
RB space dimension N	42	658	1296	1314
RB evaluation t_{RB}^{online} (s)	$2.50 \cdot 10^{-3}$	$8.28 \cdot 10^{-2}$	$4.21 \cdot 10^{-1}$	$6.07 \cdot 10^{-1}$
RB construction $t_{FE}^{offline}$ (s)	19.58	145.27	441.97	1105.25
Dofs ratio N_h/N	128	8	4	4

References

[1] N. Aubry. On the hidden beauty of the proper orthogonal decomposition. *Theor. Comp. Fluid. Dyn.*, 2:339–352, 1991.

[2] N. Aubry, W. Y. Lian, and E. S. Titi. Preserving symmetries in the proper orthogonal decomposition. *SIAM J. Sci. Comput.*, 14(2):483–505, 1993. MR1204243

[3] C. Audouze, F. De Vuyst, and P. B. Nair. Reduced-order modeling of parameterized PDEs using time-space parameter principal component analysis. *Int. J. Numer. Methods Engrg.*, 80(10):1025–1057, 2009. MR2589527

[4] I. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM J. Numer. Anal.*, 45(3):1005–1034, 2007. MR2318799

[5] I. Babuška, R. Tempone, and G. E. Zouraris. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM J. Numer. Anal.*, 42(2):800–825, 2004. MR2084236

[6] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera. An 'empirical interpolation' method: application to efficient reduced-basis discretization of partial differential equations. *C. R. Math. Acad. Sci. Paris*, 339(9):667–672, 2004. MR2103208

[7] G. Berkooz, P. Holmes, and J. L. Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *Annu. Rev. Fluid Mech.*, 25(1):539–575, 1993. MR1204279

[8] T. Bui-Thanh, M. Damodaran, and K. Willcox. Proper orthogonal decomposition extensions for parametric applications in transonic aerodynamics (AIAA Paper 2003-4213). In *Proceedings of the 15th AIAA Computational Fluid Dynamics Conference*, 2003.

[9] T. Bui-Thanh, K. Willcox, and O. Ghattas. Parametric reduced-order models for probabilistic analysis of unsteady aerodynamics applications. *AIAA Journal*, 46(10), 2008.

[10] H. J. Bungartz and M. Griebel. Sparse grids. *Acta Numerica*, 13:147–269, 2004. MR2249147

[11] J. Burkardt, M. D. Gunzburger, and H. C. Lee. Centroidal voronoi tessellation-based reduced-order modeling of complex systems. *SIAM J. Sci. Comput.*, 28(2):459–484, 2006. MR2231716

[12] J. Burkardt, M. D. Gunzburger, and H. C. Lee. POD and CVT-based reduced-order modeling of Navier-Stokes flows. *Comput. Meth. Appl. Mech. Engrg.*, 196(1-3):337–355, 2006. MR2270135

[13] S. Chaturantabut and D. C. Sorensen. Discrete empirical interpolation for nonlinear model reduction. Technical Report TR09-05, Department of Computational and Applied Mathematics, Rice University, 2009.

[14] S. Chaturantabut and D. C. Sorensen. Nonlinear model reduction via discrete empirical interpolation. *SIAM J. Sci. Comput.*, 32(5):2737–2764, 2010. MR2684735

[15] E. A. Christensen, M. Brøns, and J. N. Sørensen. Evaluation of proper orthogonal decomposition–based decomposition techniques applied to parameter-dependent nonturbulent flows. *SIAM J. Sci. Comput.*, 21:1419–1434, 1999. MR1742325

[16] W. G. Cochran. *Sampling techniques.* John Wiley & Sons, Chichester, 2007.

[17] P. Constantine. Random Field Simulation, `matlab` package. URL: http://www.mathworks.com/matlabcentral/fileexchange/27613-random-field-simulation, 2010–2012.

[18] J. B. Conway. *A Course in Operator Theory.* American Mathematical Soc., Providence, 2000. MR1721402

[19] A. E. Deane, I. G. Kevrekidis, G. E. Karniadakis, and S. A. Orszag. Low-dimensional models for complex geometry flows: Application to grooved channels and circular cylinders. *Phys. Fluids A*, 3:2337, 1991.

[20] L. Dedè and A. Quarteroni. Optimal control and numerical adaptivity for advection–diffusion equations. *ESAIM Math. Modelling Numer. Anal.*, 39(5):1019–1040, 2005. MR2178571

[21] T. Gerstner and M. Griebel. Numerical integration using sparse grids. *Numer. algorithms*, 18(3-4):209–232, 1998. MR1669959

[22] R. G. Ghanem and P. D. Spanos. *Stochastic Finite Elements: A Spectral Approach.* Springer, 1991. MR1083354

[23] G. H. Golub and C. F. Van Loan. *Matrix Computations.* The John Hopkins University Press, Baltimore, fourth edition, 2013. MR3024913

[24] M. A. Grepl, Y. Maday, N. C. Nguyen, and A. T. Patera. Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations. *ESAIM Math. Modelling Numer. Anal.*, 41(3):575–605, 2007. MR2355712

[25] M D. Gunzburger, C. G. Webster, and G. Zhang. Stochastic finite element methods for partial differential equations with random input data. *Acta Numerica*, 23:521–650, 2014. MR3202242

[26] P. J. Holmes, J. L. Lumley, and G. Berkooz. *Turbulence, coherent structures, dynamical systems and symmetry.* Cambridge University Press, Cambridge, 1998. MR1422658

[27] M. Kahlbacher and S. Volkwein. Galerkin proper orthogonal decomposition methods for parameter dependent elliptic systems. *Discuss. Math., Differ. Incl. Control Optim.*, 27(1):95–117, 2007. MR2413807

[28] G. Kerschen, J. C. Golinval, A. F. Vakakis, and L. A. Bergman. The method of proper orthogonal decomposition for dynamical characterization and order reduction of mechanical systems: an overview. *Nonlin. Dyn.*, 41:147–169, 2005. MR2157178

[29] K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for parabolic problems. *Num. Math.*, 90:117–148, 2001. MR1868765

[30] K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. *SIAM J. Numer. Anal.*, 40(2):492–515, 2003. MR1921667

[31] O.P. Le Maître and O. M. Knio. *Spectral Methods for Uncertainty Quantification With Applications to Computational Fluid Dynamics.* Computational Science and Engineering. Springer Science+Business Media B.V, 2010. MR2605529

[32] S. L. Lohr. *Sampling: Design and Analysis.* Cengage Learning, Boston, second edition, 2010. MR3057878

[33] J. Lumley. The structure of inhomogeneous turbulent flows. In A. M. Yaglom and V. I. Takarski, editors, *Atmospheric Turbulence and Radio Wave Propagation*, pages 166–178. Nauka, Moscow, 1967.

[34] Y. Maday, N. C. Nguyen, A. T. Patera, and G. S. H. Pau. A general multipurpose interpolation procedure: the magic points. *Commun. Pure Appl. Anal.*, 8(1):383–404, 2009. MR2449115

[35] M. Meyer and H. G. Matthies. Efficient model reduction in non-linear dynamics using the Karhunen-Loève expansion and dual-weighted-residual methods. *Comput. Mech.*, 31(1-2):179–191, 2003.

[36] F. Negri, A. Manzoni, and D. Amsallem. Efficient model reduction of parametrized systems by matrix discrete empirical interpolation. *J. Comp. Phys.*, 303:431–454, 2015. MR3422722

[37] A. Nouy. Proper generalized decompositions and separated representations for the numerical solution of high dimensional stochastic problems. *Arch. Comput. Methods Engrg.*, 17:403–434, 2010. MR2739946

[38] R. Pinnau. Model reduction via proper orthogonal decomposition. In W. H. A. Schilders, H. A. van der Vorst, and J. Rommes, editors, *Model Order Reduction: Theory, Research Aspects and Applications*, volume 13 of *Mathematics in Industry*, pages 96–109. Springer, Berlin Heidelberg, 2008. MR2497746

[39] C. Prud'homme, D. V. Rovas, K. Veroy, L. Machiels, Y. Maday, A. T. Patera, and G. Turinici. Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods. *J. Fluid Eng.*, 124(1):70–80, 2002.

[40] C. Prud'homme, D. V. Rovas, K. Veroy, and A. T. Patera. A mathematical and computational framework for reliable real-time solution of parametrized partial differential equations. *ESAIM Math. Modelling Numer. Anal.*, 36(5):747–771, 2002. MR1955536

[41] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced Basis Methods for Partial Differential Equations. An Introduction*, volume 92 of *Unitext*. Springer, 2016. MR3379913

[42] A. Quarteroni and A. Valli. *Numerical Approximation of Partial Differential Equations*. Springer-Verlag, Berlin-Heidelberg, 1994. MR1299729

[43] M. Rathinam and L. R. Petzold. A new look at proper orthogonal decomposition. *SIAM J. Numer. Anal.*, 41(5):1893–1925, 2003. MR2035011

[44] M. Renardy and R. C. Rogers. *An Introduction to Partial Differential Equations*, volume 13. Springer-Verlag, New York, second edition, 2004. MR2028503

[45] J. R. Retherford. *Hilbert space: Compact Operators and the Trace Theorem*. Cambridge University Press, Cambridge, 1993. MR1237405

[46] E. Schmidt. Zur theorie der linearen und nichtlinearen integralgleichungen. *Math. Ann.*, 63(4):433–476, 1907. MR1511415

[47] L. Sirovich. Turbulence and the dynamics of coherent structures, part i: Coherent structures. *Quart. Appl. Math.*, 45(3):561–571, 1987. MR0910462

[48] G. W. Stewart. On the early history of the singular value decomposition. *SIAM review*, 35(4):551–566, 1993. MR1247916

[49] T. Tonn, K. Urban, and S. Volkwein. Optimal control of parameter-dependent convection-diffusion problems around rigid bodies. *SIAM J. Sci. Comput.*, 32(3):1237–1260, 2010. MR2639237

[50] S. Volkwein. Proper orthogonal decomposition: Theory and reduced-order modelling, 2013. Lecture Notes, University of Konstanz.

[51] D. Xiu and G. E. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24:619–644, 2002. MR1951058

Andrea Manzoni
Ecole Polytechnique Fédérale de Lausanne
CH-1015 Lausanne
Switzerland
E-mail address: andrea.manzoni@epfl.ch

Federico Negri
Ecole Polytechnique Fédérale de Lausanne
CH-1015 Lausanne
Switzerland
E-mail address: federico.negri@epfl.ch

Alfio Quarteroni
Ecole Polytechnique Fédérale de Lausanne
CH-1015 Lausanne
Switzerland
E-mail address: alfio.quarteroni@epfl.ch

Received December 23, 2015