Proper orthogonal decomposition method for multiscale elliptic PDEs with random coefficients

Dingjiong Ma[a], Wai-ki Ching[a], Zhiwen Zhang[a,*]

[a] *Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong SAR.*

Abstract

In this paper we propose to use the proper orthogonal decomposition (POD) method to solve multiscale elliptic PDEs with random coefficients in the multi-query setting. Our method consists of offline and online stages. In the offline stage, a small number of reduced basis functions are constructed within each coarse grid block using the POD method. Moreover, local tensor spaces are defined to approximate the multiscale random solution space. In the online stage, a weak formulation is derived and discretized using the Garlerkin method to compute the solution. Since the reduced basis functions can efficiently approximate the high-dimensional solution space, our method is very efficient in solving multiscale elliptic PDEs with random coefficients. The convergence analysis of the proposed method is also presented. Finally, numerical results are presented to demonstrate the accuracy and efficiency of the proposed method for several multiscale stochastic problems with or without scale separation.

Keywords: Random partial differential equations (RPDEs); uncertainty quantification (UQ); proper orthogonal decomposition (POD); multiscale reduced basis; over-sampling method; high-contrast problem.

1. Introduction

Many physical and engineering applications involving uncertainty quantification (UQ) can be described by stochastic partial differential equations (SPDEs, i.e., PDEs driven by Brownian motion) or partial differential equations with random coefficients (RPDEs). In recent years, there has been an increased interest in the simulation of systems with uncertainties, and several numerical methods have been developed in the literature to solve SPDEs and RPDEs; see [18, 48, 8, 35, 23, 46, 7, 38, 37, 43, 45, 19] and references therein. In this work, we shall consider a challenging problem in the UQ, i.e., solving multiscale elliptic PDEs with random coefficients. Due to multiscale and random features in these solutions, it is extremely challenging to solve this type of problem. This motivates us to develop efficient numerical methods so that we can efficiently solve this type of problem.

Elliptic PDEs with multiscale coefficients have already become expensive to solve as they require tremendous computational resource to resolve the smallest-scale of the solution. When the multiscale coefficient has scale separation or periodic structure, one can apply homogenization theory to derive an effective equation, which allows us to solve the problem on relatively coarse mesh. However, in many applications, the multiscale coefficients usually do not satisfy the scale separation assumption or may not have periodic structures. In the past four decades, many efficient methods have been

*Corresponding author

Email addresses: martin35@hku.hk (Dingjiong Ma), wching@hku.hk (Wai-ki Ching), zhangzw@hku.hk (Zhiwen Zhang)

developed for the multiscale PDEs in the literature; see [2, 6, 20, 26, 16, 24, 28, 12, 30, 34, 15, 32, 41] and references therein.

Elliptic PDEs with multiscale and random coefficients (e.g. Eq.(1)) become more complicated since the appearance of the random dimension further increases the dimension of the solution space. Recently, Zabaras et al. proposed a stochastic variational multiscale method for diffusion in heterogeneous random media [5, 17], which combined the generalized polynomial chaos (gPC) method with variational multiscale method to perform model reduction. However, when the stochastic dimension direction is large, this method becomes expensive due to the fast growth of the number of the gPC basis elements. In [3], Ghanem et al. considered the probabilistic equivalence and stochastic model reduction in multiscale analysis. In [32], Kevrekidis et al. applied the equation-free idea to study the stochastic incompressible flows. The last author of this paper has made some progress in developing numerical methods for stochastic multiscale PDEs by exploring the low-dimensional structure of the solutions and constructing problem-dependent stochastic basis functions to solve these SPDEs. In [13, 50, 27], a data-driven stochastic method was proposed to solve stochastic partial differential equations with high-dimensional random input and/or multiscale coefficients.

In this paper, we shall use the proper orthogonal decomposition (POD) method to construct multiscale reduced basis functions, which can be used to solve multiscale PDEs with random coefficients in the multi-query setting. We consider the multiscale elliptic PDEs with random coefficients as follows,

$$-\nabla \cdot (a^{\varepsilon}(x,\omega)\nabla u^{\varepsilon}(x,\omega)) = f(x), \quad x \in D, \quad \omega \in \Omega, \tag{1}$$

where a homogeneous boundary condition is imposed. More details about the setting of Eq.(1) will be discussed in Section 3. Our method consists of offline and online stages. In the offline stage, we perform model reduction for the multiscale problem. Specifically, a small number of multiscale reduced basis functions are constructed within each coarse grid block using the POD method [9, 44, 47], which captures the multiscale features of the solution space. In addition, local tensor product spaces are constructed to approximate the solution space. It should be pointed out that the construction of the basis functions only depends on the PDE operator in Eq.(1) and does not depend on the forcing functions.

In the online stage, a weak formulation for Eq.(1) is derived in the tensor space and a global linear equation system can be obtained using the Galerkin method. Our method is efficient in solving (1) with a broad range of forcing functions. In addition, rigorous error analysis is provided to study the error between the numerical solution obtained from our method and the exact solution. Finally, numerical results are presented to demonstrate the accuracy and efficiency of the proposed method for several multiscale stochastic problems with or without scale separation.

The rest of the paper is organized as follows. To make this paper self-contained, we give a brief introduction of the POD method, the gPC basis functions, and the stochastic finite element method (SFEM) in Section 2. In section 3, we review the multiscale finite element method and present our derivations of the multiscale reduced basis functions. Then, the stochastic Galerkin discretization for the multiscale elliptic PDEs with random coefficients will be discussed in Section 4. In addition, the convergence analysis of our method will be presented. In Section 5, we discuss some numerical implementation issues and present several numerical results to demonstrate the accuracy and effectiveness of our method. Finally, some concluding remarks are given in Section 6.

2. Some preliminaries

2.1. Proper orthogonal decomposition method

The POD method, also known as Karhunen-Loève expansion (KLE) in stochastic process and signal analysis [31, 33], or the principal component analysis (PCA) in statistics [1], or singular value decomposition (SVD) in linear algebra, or the method of empirical orthogonal functions (EOF) in geophysical fluid dynamics [39, 21]. The POD method has firstly been introduced in solving the turbulence in fluid dynamics. It aims to generate optimally ordered orthonormal basis functions in the least squares sense for a given set of theoretical, experimental or computational data. Reduced-order models (ROMs) or surrogate models are then obtained by truncating this optimal basis functions, which provide considerable computational savings over the original high-dimensional problems. We refer the interested readers to [44, 9, 47, 29] and references therein for more details.

Let X be a Hilbert space equipped with the inner product $(\cdot,\cdot)_X$ and $u(\cdot,t)\in X$, $t\in[0,T]$ be the solution of a dynamic system. In practice, we approximate the space X with a linear finite dimensional space V with $dimV = d$, where d represents the degree of freedom of the solution space. We should point out that d can be extremely large for a high-dimensional problem. Given a set of snapshot of solutions, a linear space V can be spanned, denoted as

$$V = \text{span}\{u(\cdot,t_1), u(\cdot,t_2), ..., u(\cdot,t_N)\}, \tag{2}$$

where $t_1,\cdots,t_N\in[0,T]$ are different time instances. The POD method aims to build a set of low-dimensional basis functions $\{\varphi_1(\cdot),\varphi_2(\cdot),...,\varphi_r(\cdot)\}$ with $r\ll\min(N,d)$ that optimally approximates the input solution snapshots. The optimality means that given any integer r and linear independent basis $\{\varphi_k(x)\}_{k=1}^r$, the POD basis functions minimize the following error

$$\frac{1}{N}\sum_{i=1}^{N}\left|\left|u(\cdot,t_i)-\sum_{k=1}^{r}(u(\cdot,t_i),\varphi_k(\cdot))_X\varphi_k(\cdot)\right|\right|_X^2, \tag{3}$$

subject to the constraints that $(\varphi_m(\cdot),\varphi_n(\cdot))_X=\delta_{mn}$, $1\le m,n\le r$, where $\delta_{mn}=1$ if $m=n$, otherwise $\delta_{mn}=0$.

Using the method of snapshot proposed by Sirovich [44], we know that the optimization problem (3) can be reduced to an eigenvalue problem

$$Kv=\lambda v, \tag{4}$$

where $K\in R^{N\times N}$ is the correlation matrix with (i,j)-element $K_{ij}=\frac{1}{N}(u(\cdot,t_i),u(\cdot,t_j))_X$. We sort the eigenvalues in a decreasing order as $\lambda_1\ge\lambda_2\ge...\ge\lambda_N>0$ and the corresponding eigenvectors are denoted by v_k, $k=1,...,N$. It can be shown that the POD basis functions are constructed by

$$\varphi_k(\cdot)=\frac{1}{\sqrt{\lambda_k}}\sum_{j=1}^{N}(v_k)_j u(\cdot,t_j),\quad 1\le k\le N, \tag{5}$$

where $(v_k)_j$ is the j-th component of the eigenvector v_k. The basis functions $\{\varphi_k\}_{k=1}^r$ minimizes the error (3). This result as well as the error formula were proved in [22].

Proposition 2.1 ([22]). *Let $\lambda_1\ge\lambda_2\ge...\ge\lambda_N>0$ denote the positive eigenvalues of K in (4). Then $\{\varphi_k\}_{k=1}^r$ constructed according to (5) is the set of POD basis functions of rank $r\le N$, and we have the following error formula:*

$$\frac{1}{N}\sum_{i=1}^{N}\left|\left|u(\cdot,t_i)-\sum_{k=1}^{r}(u(\cdot,t_i),\varphi_k(\cdot))_X\varphi_k(\cdot)\right|\right|_X^2=\sum_{k=r+1}^{N}\lambda_k.$$

In practice, we shall make use of the decay property of eigenvalues in λ_k and choose the first r dominant eigenvalues such that the ratio $\rho = \frac{\sum_{k=1}^{r}\lambda_k}{\sum_{k=1}^{N}\lambda_k}$ is big enough to achieve an expected accuracy, for instance $\rho = 99\%$. One would prefer the eigenvalues decays as fast as possible so that the fewer POD basis functions can ensure the higher accuracy.

2.2. The generalized Polynomial Chaos (gPC) method

The multiscale elliptic PDEs with random coefficients (1) are often used to model flows in heterogeneous porous media such as water aquifer and oil reservoirs, where $a^{\varepsilon}(x,\omega)$ is used to model permeability fields [49]. In practice, however, the data of the spatially varying permeability field is available only at limited locations in the physical domain D. Therefore, we assume $a(x,\omega)$ is a second-order random filed, i.e., $a(x,\omega) \in L^2(D\times\Omega)$, with $\mathbb{E}\left[a(x,\omega)\right] = \bar{a}(x)$ and covariance kernel $C(x,y)$, where we omit the superscript ε for notation simplicity. Using the KL expansion, the permeability field $a(x,\omega)$ reads

$$a(x,\omega) = \bar{a}(x) + \sum_{i=1}^{\infty}\sqrt{\lambda_i}\xi_i(\omega)\phi_i(x),$$

where $\{\lambda_i, \phi_i(x)\}$ are eigenvalues and eigenfunctions of the covariance kernel $C(x,y)$ and $\xi_i(\omega)$ are independent and identically distributed (i.i.d) random variables. In practice, we truncate the KLE into its first r terms and obtain the parametrization of the random coefficient as follows,

$$a(x,\omega) \approx \bar{a}(x) + \sum_{i=1}^{r}\sqrt{\lambda_i}\xi_i(\omega)\phi_i(x). \tag{6}$$

Since the random coefficient $a(x,\omega)$ in Eq.(1) is parameterized by r i.i.d. random variables, i.e., $\boldsymbol{\xi}(\omega) = (\xi_1(\omega), \cdots, \xi_r(\omega))$, the Doob-Dynkin's lemma [40] implies that the solution of Eq.(1) can also be represented by a functional of these random variables, i.e. $u(x,\omega) = u(x, \xi_1(\omega), ..., \xi_r(\omega))$.

Let $\rho(\cdot)$ denote the distribution function of $\xi_i(\omega)$ and $\{\mathrm{H}_i(\xi)\}_{i=1}^{\infty}$ denote the one-dimensional polynomials that are orthogonal to each other with respect to the distribution $\rho(\xi)$, i.e.,

$$\int_{\Omega}\mathrm{H}_i(\xi)\mathrm{H}_j(\xi)\rho(\xi)d\xi = \delta_{ij}.$$

For some commonly used distributions, such as the Gaussian distribution and the uniform distribution, such orthogonal polynomial sets are Hermite polynomials and Legendre polynomials, respectively. For general distributions, such polynomial set can be obtained numerically. Furthermore, by a tensor product representation, we can use the one-dimensional polynomial $\mathrm{H}_i(\xi)$ to construct a sufficient orthonormal basis $\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$'s of $\mathbb{L}^2(\Omega)$ as follows:

$$\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \prod_{i=1}^{r}\mathrm{H}_{\alpha_i}(\xi_i), \quad \boldsymbol{\alpha} \in \mathfrak{J}_r^{\infty}, \tag{7}$$

where $\boldsymbol{\alpha}$ is a multi-index and $\mathfrak{J}_r^{\infty}$ is a multi-index set of countable cardinality,

$$\mathfrak{J}_r^{\infty} = \left\{\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_r) \,|\, \alpha_i \geq 0, \alpha_i \in \mathbb{N}\right\}.$$

The zero multi-index corresponding to $\mathbf{H}_{\mathbf{0}}(\boldsymbol{\xi}) = 1$, which is used to represent the mean of the solution. Clearly, the cardinality of $\mathfrak{J}_r^{\infty}$ is infinite. For the purpose of numerical computations, we prefer a

finite set of polynomials. There are many choices of truncations. One possible choice is the set of polynomials whose total orders are at most p, i.e.,

$$\mathfrak{J}_r^p = \left\{ \boldsymbol{\alpha} \,|\, \boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_r), \alpha_i \geq 0, \alpha_i \in \mathbb{N}, |\alpha| = \sum_{i=1}^{r} \alpha_i \leq p \right\}. \tag{8}$$

The cardinality of $\mathfrak{J}_r^p$ in (8) or the number of basis functions, denoted by $N_p = |\mathfrak{J}_r^p|$, is equal to $\frac{(p+r)!}{p!r!}$. We may simply write such a truncated set as $\mathfrak{J}$ when there is no ambiguity arises. The orthonormal basis $\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$ is the generalized Polynomial Chaos (gPC) basis, see [18, 48, 23] for more details.

2.3. Stochastic finite element method (SFEM)

The weak solution $u(x,\omega)$ to Eq.(10) is defined in a Hilbert space that has a tensor structure $V = H_0^1(D) \otimes L^2(\Omega)$. A straightforward way to solve Eq.(10) is the stochastic finite element method (SFEM) [18], where one seeks the numerical solution in a finite dimensional solution space with a tensor product form, i.e., $V_{h,\mathfrak{J}} = X_h \otimes \Xi_{\mathfrak{J}}$. Here X_h is the finite element space spanned by the fine-scale nodal basis functions and $\Xi_{\mathfrak{J}}$ is the random space spanned by the gPC basis functions $\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$, $\boldsymbol{\alpha} \in \mathfrak{J}$.

In the SFEM, we first represent the solution $u(x,\omega)$ using the gPC basis functions $\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$ as $u(x,\omega) = \sum_{\boldsymbol{\alpha}\in\mathfrak{J}} u_{\boldsymbol{\alpha}}(x)\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$. Then, we represent the gPC expansion coefficients $u_{\boldsymbol{\alpha}}(x)$ using the fine-scale finite element nodal basis functions $\phi_i(x)$ as $u_{\boldsymbol{\alpha}}(x) = \sum_{i=1}^{N_x} u_{\boldsymbol{\alpha}i}\phi_i(x)$, where $\phi_i(x)$ are defined on a fine mesh with size $h < \varepsilon$ and N_h is dimension of the finite element space. This results in an equivalent representation of the solution $u(x,\omega)$ in the tensor space $V_{h,\mathfrak{J}}$ as

$$u_{h,p}(x,\omega) = \sum_{\boldsymbol{\alpha}\in\mathfrak{J}} \sum_{i=1}^{N_x} u_{\boldsymbol{\alpha}i}\phi_i(x)\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}). \tag{9}$$

Finally, we apply the Galerkin method (with test space $V_{h,\mathfrak{J}}$) to derive a linear equation system for $u_{\boldsymbol{\alpha}i}$. When the dimension in random space (e.g. r in Eq.(6)) is large and/or the coefficient contains multiscale features, the size of the coupled linear equation system becomes extremely huge. Thus, the SFEM requires tremendous computational resources to solve Eq.(1).

3. Model reduction for multiscale basis functions

We shall develop the model reduction method to solve Eq.(1). Our method consists of offline and online stages. In the offline stage, we construct multiscale reduced basis functions using the POD method. To this end, we consider the following multiscale elliptic PDEs with random coefficients,

$$-\nabla \cdot (a^\varepsilon(x,\omega)\nabla u^\varepsilon(x,\omega)) = f(x), \quad x \in D, \omega \in \Omega, \tag{10}$$

$$u^\varepsilon(x,\omega) = 0, \qquad x \in \partial D, \tag{11}$$

where $D \in R^d$ is a bounded spatial domain, Ω is a sample space, and the smallest-scale information is parameterized by ε. The coefficient $a^\varepsilon(x,\omega) = a^\varepsilon(x,\xi_1(\omega),...,\xi_r(\omega))$ is parameterized by r independent random variables and is uniformly coercive almost surely, i.e., there exist $a_{\min}, a_{\max} > 0$, such that

$$P(\omega \in \Omega : a^\varepsilon(x,\omega) \in [a_{\min}, a_{\max}], \forall x \in D) = 1. \tag{12}$$

The force function $f(x) \in L^2(D)$ (not just $H^{-1}(D)$) is assumed to be resolved on a coarse grid, which is necessary for the compactness of the solution space.

3.1. Multiscale finite element basis functions

To make this paper self-contained, we first give a brief review of the multiscale finite element method (MsFEM) [24]. In the MsFEM, we first partition the domain D into coarse grid blocks $\{D^k, 1 \leq k \leq K\}$, which can be triangles or quadrilaterals. The union of all triangles or quadrilaterals covers the closure of D, and the intersection of different triangles or quadrilaterals is either empty, a common node, or a common edge. The mesh size of coarse grid is H, which is $H \gg \varepsilon$. For each $\omega \in \Omega$, we solve the following cell problems to obtain the multiscale finite element basis functions $\phi^{kl}(x,\omega)$ on $D^k, 1 \leq k \leq K$,

$$-\nabla \cdot (a^\varepsilon(x,\omega)\nabla\phi^{kl}(x,\omega)) = 0, \quad x \in D^k, \quad \omega \in \Omega, \tag{13}$$

$$\phi^{kl}(x,\omega) = \theta^{kl}(x), \quad x \in \partial D^k, \quad l = 1, ..., d, \tag{14}$$

where d is the number of vertexes on D^k (d=3 for triangle elements and d=4 for quadrilateral elements) and $\theta^{kl}(x)$ are Dirichlet boundary conditions. In practice, we choose $\theta^{kl}(x)$ to be the bilinear or linear basis functions. By solving the Eq.(13), the multiscale information of the Eq.(10) can be captured through $\phi^{kl}(x,\omega)$. Let $\phi^{kl}(x,\omega) \in H^1(D^k)$ be solutions of Eqns.(13) and (14). We have

$$\phi^{kl}(x,\omega) = \phi_0^{kl}(x,\omega) + \theta^{kl}(x). \tag{15}$$

The unknown homogeneous part $\phi_0^{kl}(x,\omega) \in H_0^1(D^k)$ satisfies the following variational formulation

$$b^\varepsilon(\phi_0^{kl}(x,\omega), v; \omega) = f^\varepsilon(v;\omega), \quad \forall v \in H_0^1(D^k), \tag{16}$$

where

$$b^\varepsilon(w, v; \omega) = \int_{D^k} a^\varepsilon(x,\omega)\nabla w \cdot \nabla v dx, \tag{17}$$

$$f^\varepsilon(v;\omega) = -\int_{D^k} a^\varepsilon(x,\omega)\nabla\theta^{kl}(x) \cdot \nabla v dx. \tag{18}$$

For each $\omega \in \Omega$, Eq.(13) is a deterministic problem. The convergence of the MsFEM for deterministic problems has been proved in [25]. When the Eq.(13) has random coefficient, we have

Theorem 3.1. *For each $\omega \in \Omega$, let $a^\varepsilon(x,\omega) = a(\frac{x}{\varepsilon},\omega)$ with $a(\frac{x}{\varepsilon},\omega)$ be periodic in $\frac{x}{\varepsilon}$ and smooth. Let $u^\varepsilon(x,\omega)$ be the solution of Eq.(10) and $u_H^\varepsilon(x,\omega)$ be the solution obtained from the space spanned by the multiscale basis functions with linear boundary conditions. Then, we have*

$$||u^\varepsilon - u_H^\varepsilon||_{L^2(D)} \leq C_1 H^2 ||f||_{L^2(D)} + C_2 \frac{\varepsilon}{H}, \tag{19}$$

where $\varepsilon < H$ and C_1, C_2 are generic constants that do not depend on ε and H.

The MsFEM gives the correct homogenized result as ε tends to zero. However, when $H \sim \varepsilon$, the multiscale solution has a large error, which is due to the resonance effect between the grid scale H and the small-scale parameter ε of the problem. To reduce the resonance effect, Hou et al. proposed an over-sampling technique in [25, 16]. The main observation is that the boundary layer in the boundary is strongly localized within a width of order $O(\varepsilon)$. If we solve Eqns.(13) and (14) in a domain with size larger than $H + \varepsilon$ and only use the interior information to construct the basis, we can significantly reduce the effect of the boundary layer on the basis functions.

Specifically, for each $\omega \in \Omega$, let ψ^{kl} be the basis functions satisfying the homogeneous elliptic equation in a larger domain $S^k \supset D^k$ (with $dist(S^k, \partial D^k) \geq \varepsilon$), $1 \leq k \leq K$,

$$-\nabla \cdot (a^{\varepsilon}(x,\omega)\nabla\psi^{kl}(x,\omega)) = 0, \qquad x \in S^k, \tag{20}$$

$$\psi^{kl}(x,\omega) = \theta^{kl}(x), \quad x \in \partial S^k, \quad l = 1, ..., d, \tag{21}$$

where d is the number of nodes on S^k and $\theta^{kl}(x)$ are linear functions defined on the boundary of S^k. The MsFEM basis functions $\phi^{ki}(x,\omega)$ are obtained by taking linear combination of $\psi^{kl}(x,\omega)$, i.e.,

$$\phi^{ki}(x,\omega) = \sum_{l=1}^{d} c_{il}\psi^{kl}(x,\omega)|_{D^k}, \quad i,l = 1, ..., d, \quad 1 \leq k \leq K, \tag{22}$$

where the coefficients c_{il} are determined by the condition $\phi^{ki}(x_l,\omega) = \delta_{il}$ and x_l are the nodes of domain D^k. The over-sampling technique results in a non-conforming MsFEM method though, it significantly improves the convergence rate of the MsFEM [16]. In this paper, we shall let $\phi^{ki}(x,\omega)$ denote the multiscale finite element basis functions obtained by using the over-sampling technique in [25, 16]. In practice, we partition the coarse grids into fine grids with mesh size $h \ll \varepsilon$ and numerically solve the cell problem (13) or (20) to obtain $\phi^{ki}(x,\omega)$.

3.2. Construction of multiscale reduced basis functions

When Eq.(13) has random coefficient, the MsFEM becomes prohibitively expensive since we need to compute $\phi^{ki}(x,\omega)$ for each sample of the permeability field $a^{\varepsilon}(x,\omega)$. To address this issue, we use the POD method to build a small number of reduced basis functions that enable us to obtain the approximated multiscale finite element basis functions within each coarse grids in a cheaper way.

In our method, we partition the physical domain D into a number of sub-domains D^k, $k = 1, ..., K$ and construct the partition of unity functions $\chi_k(x)$ [36], which allow us to restrict the global solution on each local coarse grid. They satisfy

$$D = \bigcup_{k=1}^{K} D^k, \quad \text{supp}\big(\chi_k(x)\big) \subset D^k, \quad \chi_k(x) \geq 0, \quad \sum_{k=1}^{K} \chi_k(x) = 1, \quad \forall x \in D. \tag{23}$$

Assumption 3.2. *We assume that in the partition of unity the number of sub-domains D^k that intersects D^l is bounded for each l and*

$$K \leq \frac{C}{H^d}, \quad diam(D^k) = O(H), \quad |\nabla\chi_k(x)| \leq \frac{C}{H}, \quad \forall x \in D. \tag{24}$$

Assumption 3.2 can be easily satisfied if we use uniform coarse grids; see Fig.1 for a partition of D. In Fig.1, we show a typical sub-domain D^k, which is centered at x_k and consists of four quadrilaterals D_1^k, D_2^k, D_3^k, and D_4^k, i.e., $D^k = \cup D_j^k, j = 1, 2, 3, 4$. On each D_j^k, we define a bilinear quadrilateral nodal basis function that equals one at x_k and equals zero at other nodes. Let $\chi_k(x)$ denote the combination of these four functions, which is our partition of unity function associated with x_k. Equipped with the partition of unity functions, we shall compute multiscale reduced basis functions within each D^k.

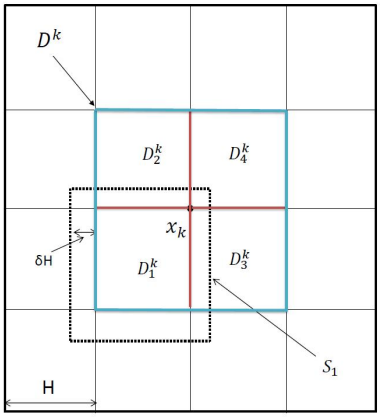

Figure 1: Illustrations of the uniform coarse meshes and a typical sub-domain D^k.

First, we compute samples of multiscale finite element basis functions within each D^k. Let $\{a^\varepsilon(x,\omega_q)\}_{q=1}^Q$ be a set of realization of the coefficients that are obtained using Monte Carlo method or stochastic collocation method [7], where Q is the number of realizations. Recall that D^k consists of four quadrilaterals D_1^k, D_2^k, D_3^k, and D_4^k in Fig.1. For each $a^\varepsilon(x,\omega_q)$ and within each D_j^k, $j=1,...,4$, we use the oversampling techniques (see Eqns.(20)-(22)) to compute the multiscale finite element basis function $\phi_j^k(x,\omega_q)$, where δH is the oversamping mesh size. We glue the four basis functions together into one basis function $\phi^k(x,\omega_q)$, which is defined on the domain D^k. We remark that the basis functions $\phi^k(x,\omega_q)$ are non-conforming and the interested reader is referred to [16] for the proof of the convergence rate of the non-conforming MsFEM.

Let $\mathcal{N}_H$ denote the set of interior vertices of coarse grids. For each $x_k \in \mathcal{N}_H$, let $\{\phi^k(x,\omega_q)\}_{q=1}^Q$ denote the MsFEM nodal basis functions associated with x_k, which satisfy $\phi^k(x_l,\omega_q)=\delta_{kl}$, $\forall x_l \in \mathcal{N}_H$, $\forall \omega_q \in \Omega$. We aim to obtain a set of multiscale reduced basis from $\{\phi^k(x,\omega_q)\}_{q=1}^Q$.

Let $\zeta_0^k(x)$ denote the sample mean of the basis functions, which is obtained using the Monte Carlo method or the stochastic collocation method. We substract the mean $\zeta_0^k(x)$ from the samples of the MsFEM basis and obtain the fluctuation of the MsFEM basis $\varphi^k(x,\omega_q)=\phi^k(x,\omega_q)-\zeta_0^k(x)$, $q=1,...,Q$. We define the snapshot of the fluctuations as

$$V=\text{span}\{\varphi^k(x,\omega_1),...,\varphi^k(x,\omega_Q)\}. \tag{25}$$

We apply the POD method to the snapshot of the fluctuations V and build a set of basis functions $\{\zeta_1^k(x),\zeta_2^k(x),...,\zeta_{m_k}^k(x)\}$ with $m_k \ll Q$ that optimally approximates the input solution snapshots. More details can be found in Section 2.1. Moreover, we have the following approximate property

Proposition 3.3. *Let $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_{m_k} \geq \lambda_{m_k+1} \geq ... > 0$ denote the positive eigenvalues of the covariance kernel associated with the snapshot of the fluctuations V and the corresponding eigenfunctions are $\zeta_1^k(x)$, ..., $\zeta_{m_k}^k(x)$,.... Then, $\{\zeta_j^k(x)\}_{j=1}^{m_k}$ will be the POD basis set and we have the following error formula hold:*

$$\frac{1}{Q}\sum_{j=1}^{Q}\left|\left|\varphi^k(x,\omega_j)-\sum_{i=1}^{m_k}\left(\varphi^k(x,\omega_j),\zeta_i^k(x)\right)_X\zeta_i^k(x)\right|\right|_X^2=\sum_{s=m_k+1}^{Q}\lambda_s, \tag{26}$$

where $X=H^1(D)$ and the number m_k will be determined according to the ratio $r=\frac{\sum_{j=1}^{m_k}\lambda_j}{\sum_{j=1}^{Q}\lambda_j}$.

Notice that $\zeta_0^k(x)$ and $\zeta_i^k(x)$, $i = 1, ..., m_k$ approximately capture the mean profile and fluctuation of the MsFEM basis associated with x_k, respectively. Thus, we expand the solution to Eqns.(20)-(22) as follows,

$$\phi^k(x,\omega) \approx \zeta_0^k(x) + \sum_{i=1}^{m_k} c_i(\omega)\zeta_i^k(x), \tag{27}$$

and use the Galerkin method to determine the coefficients $c_i(\omega)$.

Remark 3.1. To compute the multiscale reduced basis functions, we need to partition the coarse grids D^k into fine-scale quadrilateral elements with mesh $h \ll \varepsilon$, which requires a certain amount of computational cost in the offline stage. However, the pre-computed reduced basis functions can be repeatedly used in the online stage for different force functions $f(x)$, which results in considerable savings if we need to solve Eq.(10) with many different force functions.

Remark 3.2. Alternatively, we can apply a greedy-type algorithm to compute the multiscale reduced basis $\{\zeta_i^k\}_{i=0}^{m_k}$ [42]. However, the focus of this paper is to develop efficient methods for multiscale problems with random coefficients so we simply choose the POD method.

3.3. New discretization method using the multiscale reduced basis functions

Using the partition of unity functions and our multiscale reduced basis functions, we can approximate the solution $u^\varepsilon(x,\omega)$ to Eq.(10) as

$$u^\varepsilon(x,\omega) = \sum_{\boldsymbol{\alpha}\in\mathfrak{J}} u_{\boldsymbol{\alpha}}^\varepsilon(x)\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \sum_{\boldsymbol{\alpha}\in\mathfrak{J}} \Big(\sum_{k=1}^{K} \chi_k(x) u_{\boldsymbol{\alpha}}^\varepsilon(x)\Big)\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \sum_{\boldsymbol{\alpha}\in\mathfrak{J}} \Big(\sum_{k=1}^{K} u_{\boldsymbol{\alpha}}^{\varepsilon,k}(x)\Big)\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}), \tag{28}$$

where $\chi_k(x)$ is the partition of unity function associated with sub-domain D_k (see Eq.(23)) and $u_{\boldsymbol{\alpha}}^{\varepsilon,k}(x) = \chi_k(x)u_{\boldsymbol{\alpha}}^\varepsilon(x)$. Instead of discretizing the multiscale solution $u_{\boldsymbol{\alpha}}^{\varepsilon,k}(x)$ on a fine-scale partition of the sub-domain D_k, i.e.,

$$u_{\boldsymbol{\alpha}}^{\varepsilon,k}(x) = \sum_{i=1}^{N_h^k} u_{\boldsymbol{\alpha}i}^{\varepsilon,k}\phi_i(x), \tag{29}$$

where $\phi_i(x)$ are FE nodal basis functions defined on D_k with mesh size $h < \varepsilon$ and N_h^k is the number of FE basis functions, we approximate $u_{\boldsymbol{\alpha}}^{\varepsilon,k}(x)$ using our multiscale reduced basis functions $\{\zeta_i^k(x)\}_{i=0}^{m_k}$

$$u_{\boldsymbol{\alpha}}^{\varepsilon,k}(x) = \sum_{i=0}^{m_k} u_{\boldsymbol{\alpha}i}^{\varepsilon,k}\zeta_i^k(x). \tag{30}$$

This leads to the new discretization of the solution $u^\varepsilon(x,\omega)$ to Eq.(10) as

$$u_{h,p}^\varepsilon(x,\omega) = \sum_{\boldsymbol{\alpha}\in\mathfrak{J}}\sum_{k=1}^{K}\sum_{i=0}^{m_k} u_{\boldsymbol{\alpha}i}^{\varepsilon,k}\zeta_i^k(x)\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}). \tag{31}$$

The number of unknowns in Eq.(31) is far less than that of SFEM discretization (9) because multiscale reduced basis functions have already captured the low-dimensional structures in the solution space. Finally, we substitute the new discretization form (31) into Eq.(10) and apply the Galerkin method in the tensor space to determine the expansion coefficient $u_{\boldsymbol{\alpha}i}^{\varepsilon,k}$. The stiffness matrix obtained from our new discretization method can be used to solve Eq.(10) with many different force functions $f(x)$.

Remark 3.3. We focus on constructing multiscale reduced basis functions in physical space to reduce the computational cost in this paper. Therefore, we simply choose the gPC basis functions $\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$ in the random space. One can adopt the data-driven stochastic method developed in [13] to construct more efficient stochastic basis functions. However, we do not want to complicate the presentation by pursuing this avenue.

4. Error analysis

We shall analyze the error between the numerical solution obtained using our method, denoted by $u_{POD}(x,\omega)$ and the exact solution $u(x,\omega)$. We assume the fine grids with mesh size $h \ll \varepsilon$ so that the error between the numerical solution $u_h(x,\omega)$ obtained on the fine grids and $u(x,\omega)$ is accurate enough and we simply use $u_h(x,\omega)$ as a reference solution. To demonstrate the error, we need to consider errors from the physical space and the random space approximation. Before proceeding to the analysis, we first introduce some notations and assumptions.

4.1. Some notations and assumptions

We define the norm in the tensor space $L^2(D)\otimes L^2(\Omega)$ as

$$||u||_{L^2(D)\otimes L^2(\Omega)} = \Big(\int_\Omega \big(\int_D u(x,\omega)^2 dx\big) dP(\omega)\Big)^{\frac{1}{2}}, \tag{32}$$

where $P(\omega)$ is the probability distribution function of random variable $\xi(\omega)$. In addition, we need higher regularity in the random space when we estimate the convergence rate of our method in the random space. Let $D^\nu_\xi u(x,\cdot)$ denote the ν-th order mixed derivatives of $u(x,\cdot)$ with respect to the variable $\xi = (\xi_1, ..., \xi_r)$ in the random space, where $\nu = (\nu_1, ..., \nu_r)$ and ν_i are nonnegative integers. Then, we define the norm and seminorm in the random space as follows:

$$||u(x,\cdot)||^2_{H^p(\Omega)} = \int_\Omega \sum_{|\nu|\le p} |D^\nu_\xi u(x,\cdot)|^2 dP(\omega), \quad |u(x,\cdot)|^2_{H^p(\Omega)} = \int_\Omega \sum_{|\nu|= p} |D^\nu_\xi u(x,\cdot)|^2 dP(\omega). \tag{33}$$

In addition, we need the following assumption for the stability of the solution with respect to the random dimension [11, 19], which is satisfied if $a(x,\xi(\omega))$ satisfies certain regularity conditions. We refer interested reader to [14] for more details.

Assumption 4.1. *If $u(x,\omega)$ is the solution to Eqns. (10) and (11) and $u(x,\omega) \in H^p(\Omega), \forall x \in D$. Then we have the stability estimate as follows*

$$||u(x,\cdot)||_{H^p(\Omega)} \le C_1 ||f(x)||_{L^2(D)}, \quad \forall x \in D, \tag{34}$$

where the constant C_1 depends on the value of a_{min} and a_{max}.

4.2. Error Analysis

After introducing the necessary notations and assumption, we are in the position to proceed with the error analysis. Applying the triangle inequality, we divide the error into three parts

$$||u_h - u^{POD}|| \le ||u_h - u_h^{gPC}|| + ||u_h^{gPC} - u_H^{gPC}|| + ||u_H^{gPC} - u^{POD}||, \tag{35}$$

where u_h refers to the reference solution obtained using FEM with fine mesh h and no discretization in random space, u_h^{gPC} is computed using FEM with fine mesh h in physical space and gPC basis in

random space, u_H^{gPC} is computed using MsFEM basis with coarse mesh H in physical space and gPC basis in random space, and u^{POD} is obtained using the multiscale reduced basis in physical space and gPC basis in random space. We have assumed that the error between u_h and the exact solution u to Eqns. (10) and (11) is negligible.

We first consider the error between u_h and u_h^{gPC} and get the estimate result. To illustrate the main idea, we assume the coefficient in Eq.(10) is parameterized by one-dimensional random variable $\xi(\omega) = \xi_1(\omega)$ that follows uniform distribution $U[-1, 1]$ and the basis functions in gPC method are Legendre polynomials. But we emphasize that the convergence estimate (36) holds for general gPC methods if we tensorize the orthogonal polynomials and use the multi-index.

Lemma 4.2. *Let $u_h(x, \omega)$ be the reference solution obtained using FEM with fine mesh h and $u_h^{gPC}(x, \omega)$ be computed using FEM with same mesh in physical space and gPC basis in random space. Then we get the following error estimate*

$$||u_h - u_h^{gPC}||_{L^2(D)\otimes L^2(\Omega)} \leq C_2 N^{-p} ||f(x)||_{L^2(D)}, \tag{36}$$

where N is the highest order of polynomial basis in the gPC method, p is an integer that quantifies the regularity of $u_h(x, \omega)$ in the random space (see Assumption 4.1), and C_2 is independent of N but depends on a_{min} and a_{max}.

Proof. Let $L_k(\xi(\omega))$ be the Legendre polynomial of order k and S_N be the space spanned by Legendre polynomials of degree at most N, i.e., $S_N = span\{L_k(\xi(\omega))\}_{k=0}^N$. Let P_N denote the projection operator on S_N. Specifically, we have the projection of $u_h(\omega)$ onto S_N defined as $P_N u_h(\omega) = \sum_{k=0}^{N} u_k L_k(\xi(\omega))$, where the coefficients $u_k = \frac{(u_h, L_k)}{(L_k, L_k)}$ and the inner product of two functions are defined as $(v, w) \equiv \int_\Omega v(\omega) w(\omega) dP(\omega)$. To estimate the decay rate in the projection coefficients, we use the property that Legendre polynomials satisfy the Sturm-Liouville eigenvalue problem as follows,

$$\mathscr{L} L_k(\xi(\omega)) = \frac{d}{d\xi}\big((1 - (\xi(\omega))^2)\frac{d}{d\xi}\big) L_k(\xi(\omega)) = -k(k+1) L_k(\xi(\omega)). \tag{37}$$

Some simple calculations imply that

$$(u_h, L_k) = -\frac{1}{2k(k+1)} \int_{-1}^{1} u_h \mathscr{L} L_k dP(\omega) = -\frac{1}{k(k+1)} (\mathscr{L} u_h, L_k). \tag{38}$$

Then, we repeat the above derivation and get $(u, L_k) = (-\frac{1}{k(k+1)})^l (\mathscr{L}^l u, L_k)$, where $l \geq 1$ is an integer. Finally, we obtain the error estimate of the projection approximation as follows:

$$\begin{aligned} ||u_h(\omega) - P_N u_h(\omega)||^2_{L^2(\Omega)} &= \sum_{k=N+1}^{\infty} \frac{(u_h, L_k)^2_{L^2(\Omega)}}{(L_k, L_k)_{L^2(\Omega)}} = \sum_{k=N+1}^{\infty} \frac{1}{(k(k+1))^{2l} ||L_k||^2_{L^2(\Omega)}} (\mathscr{L}^l u_h, L_k)^2_{L^2(\Omega)} \\ &\leq N^{-4l} ||\mathscr{L}^l u_h||^2_{L^2(\Omega)} \leq C_2 N^{-4l} ||u_h||^2_{H^{2l}(\Omega)}, \end{aligned} \tag{39}$$

where the Parseval's identity is used in the first equation. In the gPC method, we know that $u_h^{gPC}(x, \omega) = P_N u_h(x, \omega)$. If we take the regularity index $p = 2l$ and use the stability assumptions (34) and integrate over the physical space, we prove the lemma. □

When the coefficient in Eq.(10) is parameterized by r i.i.d. random variables $\xi = (\xi_1, ..., \xi_r)$, we use the multi-index $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_r), 0 \leq \alpha_i \leq N_i, \alpha_i \in \mathbb{N}$ to label the gPC basis and the multi-index

$\nu = (\nu_1, ..., \nu_r)$, $\nu_i \geq 0$, $\nu_i \in \mathbb{N}$ to label the order of the mixed derivatives. We define the tensorized Legendre polynomials by $L_{\boldsymbol{\alpha}}(\xi(\omega)) = \prod_{i=1}^{r} L_{\alpha_i}(\xi_i(\omega))$ and $S_N = span\{L_{\boldsymbol{\alpha}}(\xi(\omega))\}$. Let P_N denote the projection operator on S_N, i.e., $P_N u_h(\omega) = u_{\boldsymbol{\alpha}} L_{\boldsymbol{\alpha}}(\xi(\omega))$, where the Einstein summation convention is used and $N = \prod_{i=1}^{r} N_i$.

Corollary 4.3. *Let $u_h(x,\omega)$ be the reference solution and $u_h^{gPC}(x,\omega) = P_N u_h(x,\omega)$ be the gPC solution obtained using the same mesh. Then we get the convergence estimate as follows,*

$$||u_h(x,\omega) - P_N u_h(x,\omega)||^2_{L^2(\Omega)} \leq (N_1^{-4\nu_1} N_2^{-4\nu_2} ... N_r^{-4\nu_r}) ||\mathscr{L}^{\nu} u||^2_{L^2(\Omega)}. \tag{40}$$

If the highest order of the polynomials are same in each random variable and let $|\nu| = \sum_{i=1}^{r} \nu_i$, we get

$$||u_h(x,\omega) - P_N u_h(x,\omega)||^2_{L^2(\Omega)} \leq N_1^{-4|\nu|} ||\mathscr{L}^{\nu} u||^2_{L^2(\Omega)} \leq C_2 N_1^{-4|\nu|} ||u||^2_{H^{2|\nu|}(\Omega)}. \tag{41}$$

Remark 4.1. One may choose the best N-term Galerkin approximations in the gPC method [14], which reduces the total number of basis and maintain an optimal convergence rate.

Remark 4.2. The orthogonal polynomials arise from a differential equation of the form

$$Q(x) f''(x) + L(x) f'(x) = \lambda f(x), \tag{42}$$

where $Q(x)$ and $L(x)$ are polynomials and the function $f(x)$ and the constant λ are obtained by solving this Sturm-Liouville type eigenvalue problem. The solutions of (42) have singularities unless λ takes on specific values. Let $\{\lambda_k, f_k(x)\}$, $k = 0, 1, ...$ denote the corresponding eigenvalues and eigenfunctions. Then, $\{f_k(x)\}$ forms a set of orthogonal polynomials and the eigenvalues satisfy $\lambda_k = \frac{k(k-1)}{2} Q''(x) + kL'(x)$, see [4, 10]. Notice that $\lambda_k \sim O(k^2)$, we have the error estimate (36) holds for other orthogonal polynomials.

We then analyse the error between $u_h^{gPC}(x,\omega)$ and $u_H^{gPC}(x,\omega)$ and obtain the following lemma.

Lemma 4.4. *Let $u_h^{gPC}(x,\omega)$ and $u_H^{gPC}(x,\omega)$ be defined as above. We get the following error estimate*

$$||u_h^{gPC} - u_H^{gPC}||_{L^2(D)\otimes L^2(\Omega)} \leq C_3 H^2 ||f||_{L^2(D)} + C_4 \frac{\varepsilon}{H} \tag{43}$$

where $\varepsilon < H$ and C_3, C_4 are constants that do not depend on ε and H.

Proof. Since $u_h^{gPC}(x,\omega)$ and $u_H^{gPC}(x,\omega)$ use the same gPC basis in the random space, we only need to analyze the error in physical space. For each realization ω_i, the problem (10) and (11) are reduced to a deterministic problem and $u_h^{gPC}(x,\omega_i)$ is a FEM reference solution and $u_H^{gPC}(x,\omega_i)$ is the solution obtained using the MsFEM with over-sampling technique. Then, Theorem 3.1 implies

$$||u_h^{gPC}(x,\omega_i) - u_H^{gPC}(x,\omega_i)||_{L^2(D)} \leq C_3 H^2 ||f(x)||_{L^2(D)} + C_4 \frac{\varepsilon}{H}. \tag{44}$$

Integrate over the random space, we prove Lemma 4.4. □

Finally, we estimate the error between $u_H^{gPC}(x,\omega)$ and $u^{POD}(x,\omega)$.

Lemma 4.5. *Let u_H^{gPC} be the solution obtained using the MsFEM basis on coarse mesh with size H in physical space and gPC basis in random space, and u^{POD} be the solution obtained using our method. We have the error estimate,*

$$||u_H^{gPC} - u^{POD}||_{L^2(D)\otimes L^2(\Omega)} \leq C_5 \sum_{i=m^*+1}^{Q} \lambda_i, \tag{45}$$

where C_5 is a generic constant, m^* is an integer that will be defined in the proof, Q is the sample number in the basis snapshots, and λ_i are eigenvalues of the covariance kernel of the snapshot.

Proof. Since $u_H^{gPC}(x,\omega)$ and $u^{POD}(x,\omega)$ are computed using the same gPC basis functions in the random space, we only need to analyze the error in physical space. For each realize ω_s, we have $u_H^{gPC}(x,\omega_s) = \sum_{k=1}^K U_k\phi^k(x,\omega_s)$ and $u^{POD}(x,\omega_s) = \sum_{k=1}^K \tilde{U}_k\tilde{\phi}^k(x,\omega_s)$, where $\phi^k(x,\omega_s)$ denote the MsFEM basis associated with the sub-domain D_k and $\tilde{\phi}^k(x,\omega_s)$ denote the approximated basis obtained using our multiscale reduced basis functions. We shall use the Galerkin method to compute the numerical solutions U_k and $\tilde{U}_k$, $k = 1,...,K$ and estimate the error between $u_H^{gPC}(x,\omega_s)$ and $u^{POD}(x,\omega_s)$.

Let $A = (a_{ij})$ and $\tilde{A} = (\tilde{a}_{ij})$ denote the stiffness matrices associated with the MsFEM and our method, where the entries in the matrices are defined as

$$a_{ij} = \int \nabla\phi^i(x,\omega_s)a^\varepsilon(x,\omega_s)\nabla\phi^j(x,\omega_s)dx, \tag{46}$$

$$\tilde{a}_{ij} = \int \nabla\tilde{\phi}^i(x,\omega_s)a^\varepsilon(x,\omega_s)\nabla\tilde{\phi}^j(x,\omega_s)dx. \tag{47}$$

We estimate the difference between the entries a_{ij} and $\tilde{a}_{ij}$ by

$$\begin{aligned}
|a_{ij} - \tilde{a}_{ij}| &= \Big| \int \nabla\phi^i(x,\omega_s)a^\varepsilon(x,\omega_s)\nabla\phi^j(x,\omega_s)dx - \int \nabla\tilde{\phi}^i(x,\omega_s)a^\varepsilon(x,\omega_s)\nabla\tilde{\phi}^j(x,\omega_s)dx \Big| \\
&\le \Big| \int \nabla\phi^i a^\varepsilon\nabla\phi^j dx - \int \nabla\tilde{\phi}^i a^\varepsilon\nabla\phi^j dx \Big| + \Big| \int \nabla\tilde{\phi}^i a^\varepsilon\nabla\phi^j dx - \int \nabla\tilde{\phi}^i a^\varepsilon\nabla\tilde{\phi}^j dx \Big| \\
&\le ||a^\varepsilon||_{L^\infty(D)} \Big(||\phi^j||_{H^1(D)} ||\nabla\phi^i - \nabla\tilde{\phi}^i||_{L^2(D)} + ||\tilde{\phi}^i||_{H^1(D)} ||\nabla\phi^j - \nabla\tilde{\phi}^j||_{L^2(D)} \Big).
\end{aligned} \tag{48}$$

Recall that $a^\varepsilon(x,\omega_s)$ is bounded almost surely, $\phi^i(x,\omega_s)$ and $\tilde{\phi}^i(x,\omega_s)$ are bounded in H^1 norm. If we choose $m^* = \min\{m_k\}_{k=1}^K$, then the Proposition 3.3 implies

$$\begin{aligned}
||\nabla\phi^k - \nabla\tilde{\phi}^k||_{L^2(D)} \le ||\phi^k - \tilde{\phi}^k||_{H^1(D)} &= \Big|\Big| y_s^k - \sum_{i=1}^{m_k} (y_s^k, \zeta_i^k)_{H^1(D)} \zeta_i^k \Big|\Big|_{H^1(D)}^2 \\
&= \sum_{j=m_k+1}^{Q} \lambda_j \le \sum_{j=m^*+1}^{Q} \lambda_j.
\end{aligned} \tag{49}$$

Combine Eqns.(48) and (49), we obtain $|a_{ij} - \tilde{a}_{ij}| \le C\sum_{s=m^*+1}^Q \lambda_s$, where C is a generic constant and may depend on a_{min} and a_{max}.

Now we obtain the error between A and $\tilde{A}$, i.e., $\tilde{A} = A + E_a$, where E_a is the error term and is bounded by $C\sum_{s=m^*+1}^Q \lambda_s$. Let $F = (f_i)$ and $\tilde{F} = (\tilde{f}_i)$ denote the right-hand side vectors, with $f_i = \int \phi^i(x,\omega_s)f(x)dx$ and $\tilde{f}_j = \int \tilde{\phi}^i(x,\omega_s)f(x)dx$, respectively. Using a similar argument, we can get $\tilde{F} = F + E_f$, where E_f is the error term and is bounded by $C\sum_{s=m^*+1}^Q \lambda_s$.

Recall that $u_H^{gPC}(x,\omega_s) = \sum_{k=1}^K U_k\phi^k(x,\omega_s)$ and $u^{POD}(x,\omega_s) = \sum_{k=1}^K \tilde{U}_k\tilde{\phi}^k(x,\omega_s)$. Let $U = (U_1,...U_K)^T$ and $\tilde{U} = (\tilde{U}_1,...\tilde{U}_K)^T$ and they satisfy $AU = F$ and $\tilde{A}\tilde{U} = \tilde{F}$, respectively. After simple calculations, we get $U - \tilde{U} = A^{-1}(E_f + E_aU)$ and the estimate

$$||U - \tilde{U}||_{l^2} \le ||A^{-1}||_{l^2} ||E_f + E_aU||_{l^2} \le C||f(x)||_{L^2(D)} \sum_{s=m^*+1}^{Q} \lambda_s, \tag{50}$$

where C depends on a_{min} and a_{max} and we have used the fact A^{-1} is bounded above by $\frac{a_{max}}{a_{min}}$.

Let $\Phi = (\phi^1(x,\omega_s),...,\phi^K(x,\omega_s))$ and $\tilde{\Phi} = (\tilde{\phi}^1(x,\omega_s),...,\tilde{\phi}^K(x,\omega_s))$. Then $u_H^{gPC}(x,\omega_s) = \Phi U$ and $u^{POD}(x,\omega_s) = \tilde{\Phi}\tilde{U}$ and we have the estimate

$$||u_H^{gPC}(x,\omega_i) - u^{POD}(x,\omega_i)||^2_{L^2(D)} = ||\Phi U - \tilde{\Phi}\tilde{U}||_{L^2(D)} \tag{51}$$

$$\leq ||\Phi U - \tilde{\Phi}U||_{L^2(D)} + ||\tilde{\Phi}U - \tilde{\Phi}\tilde{U}||_{L^2(D)} \tag{52}$$

$$\leq ||U||_{l^2}||\Phi - \tilde{\Phi}||_{L^2(D)} + ||\tilde{\Phi}||_{L^2(D)}||U - \tilde{U}||_{l^2} \tag{53}$$

According to the estimates (49) and (50), we obtain

$$||u_H^{gPC}(x,\omega_i) - u^{POD}(x,\omega_i)||^2_{L^2(D)} \leq C_4||f(x)||_{L^2(D)} \sum_{s=m^*+1}^{Q} \lambda_s, \tag{54}$$

where C_4 depends on a_{min} and a_{max}. To achieve a given error tolerance ρ, we can choose the integer m^* according to the decay property of $\sum_{s=m^*+1}^{Q} \lambda_s$. Finally, integrating in the random space we can complete our proof

$$||u_H^{gPC} - u^{POD}||_{L^2(D)\otimes L^2(\Omega)} \leq C_4||f(x)||_{L^2(D)} \sum_{s=m^*+1}^{Q} \lambda_s. \tag{55}$$

□

Theorem 4.6. *If u_h is the reference solution to Eqns.(10) and (11) and $u^{POD}(x,\omega)$ is the solution obtained using our method, then we have the error estimate as follows*

$$||u_h - u^{POD}||_{L^2(D)\otimes L^2(\Omega)} \leq \prod_{i=1}^{r} N_i^{-2\nu_i}||\mathscr{L}^{\nu}u_h||_{L^2(\Omega)} + \Big(C_2H^2 + C_4 \sum_{s=m^*+1}^{Q} \lambda_s\Big)||f(x)||_{L^2(D)\otimes L^2(\Omega)} + C_3\frac{\varepsilon}{H} \tag{56}$$

where h and H are fine and coarse mesh size respectively, C_i, $i = 2,3,4,5$ are constants depend on a_{min} and a_{max} and diameter of the domain but do not depend on ε, H and h, and λ_s are eigenvalues obtained in the POD method with m^* chosen such that $\sum_{s=m^*+1}^{Q} \lambda_s / \sum_{s=1}^{Q} \lambda_i$ is smaller than a given error tolerance ρ.

Proof. For notation simplicity, we use $||\cdot||$ to denote the norm $||\cdot||_{L^2(D)\otimes L^2(\Omega)}$. Using the triangle inequality, we get

$$||u_h - u^{POD}|| \leq ||u_h - u_h^{gPC}|| + ||u_h^{gPC} - u_H^{gPC}|| + ||u_H^{gPC} - u^{POD}|| \tag{57}$$

and Theorem 4.6 is a simple conclusion of the three lemmas we have proved above. □

5. Numerical results

5.1. An example with an oscillatory coefficient

We consider the following multiscale elliptic PDE with random coefficient on $D = [0,1]\times[0,1]$

$$-\nabla\cdot(a^{\varepsilon}(x,y,\omega)\nabla u^{\varepsilon}(x,y,\omega)) = f(x,y), \quad (x,y)\in D, \omega\in\Omega, \tag{58}$$

$$u^{\varepsilon}(x,y,\omega) = 0, \qquad (x,y)\in \partial D. \tag{59}$$

The multiscale and random features are described by the coefficient

$$a^{\varepsilon}(x,y,\omega) = 0.2 + \frac{1}{2+p_1\sin(\frac{2\pi(x-y)}{\varepsilon_1})}\xi_1(\omega) + \frac{1}{4+p_2\sin(\frac{2\pi x}{\varepsilon_2})}\xi_2(\omega) + \frac{2+p_3\cos(\frac{2\pi(x-0.5)}{\varepsilon_3})}{2-p_3\sin(\frac{2\pi(y-0.5)}{\varepsilon_3})}\xi_3(\omega), \tag{60}$$

where $[p_1,p_2,p_3] = [1.6, 1.4, 1.5]$, $[\varepsilon_1,\varepsilon_2,\varepsilon_3] = [\frac{1}{19.3}, \frac{1}{15.4}, \frac{1}{23.7}]$, and $\{\xi_i(\omega)\}_{i=1}^3$ are independent uniform random variables in $[0,1]$. We choose the parameters of the coefficient (60) in such a way that it will generate oscillatory features in the solution.

Multiquery results in the online stage. We shall show that our multiscale reduced basis can be used to solve a multi-query problem. We use the standard FEM to discretize the physical space. We choose a 256×256 fine mesh to well resolve the physical space of the solution $u^{\varepsilon}(x,y,\omega)$. Since the solution $u^{\varepsilon}(x,y,\omega)$ is smooth in random space, we use a sparse-grid based stochastic collocation (SC) method to discretize the random space. First, we conduct a convergence study and find that the relative errors of the mean and the standard deviation (STD) between the solutions obtained by a seven-level sparse grids in the SFEM and higher-level sparse grids are smaller than 0.1% both in the L^2 and the H^1 norm. Therefore, we choose the solution obtained using the seven-level sparse grids (with 207 points) as the reference solution.

To implement our method, in the physical space we take the coarse grids to be 16×16 and fine grids to be 256×256. For the coefficient (60) with 3 random variables, we choose the gPC basis functions of total order 4 and keep the POD modes to be 4, which means that we use 5 multiscale reduced basis in each interior coarse mesh. The force function $f(x,y)$ should be well-resolved by the coarse mesh. We choose $f(x,y) \in \mathfrak{F} = \{\sin(k_i\pi x + \phi_i)\cos(l_i\pi y + \varphi_i)\}_{i=1}^{20}$, where k_i and l_i are uniformly distributed over the interval $[0,4]$, while ϕ_i and φ_i are uniformly distributed over the interval $[0,1]$.

In Fig.2 and Fig.3, we show the relative errors of the mean and the STD of our method in the L^2 norm and the H^1 norm, respectively. One can see that our method is efficient in solving multi-query problem. In Fig.4, we show the mean and STD of the solution corresponding to $f(x,y) =$

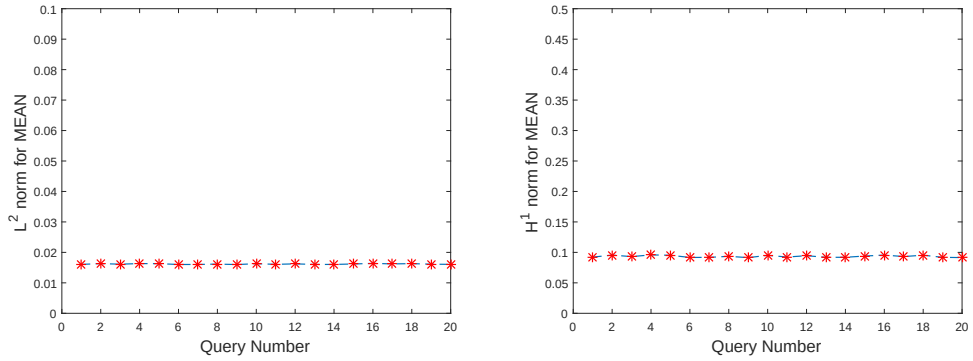

Figure 2: The mean error of our method in the L^2 norm and the H^1 norm.

$\sin(1.3\pi x + 0.1)\cos(2.1\pi y + 0.12)$. We use the notation SC to denote the reference solution and the reduction method to denote the solution obtained using our method. One can see that the mean and the STD of the DSM solution match with the mean and the STD of the reference solution very well.

Verification of the convergence rate with respect to coarse mesh. We choose the coefficient as follows

$$a^{\varepsilon}(x,y,\omega) = 0.1 + \frac{2+p_1\sin(\frac{2\pi(x-y)}{\varepsilon_1})}{2-p_1\cos(\frac{2\pi(x-y)}{\varepsilon_1})}\xi_1(\omega), \tag{61}$$

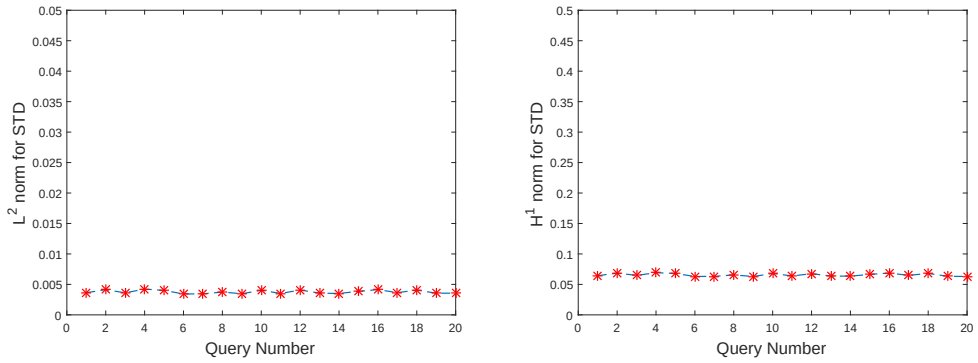

Figure 3: The STD error of our method in the L^2 norm and the H^1 norm.

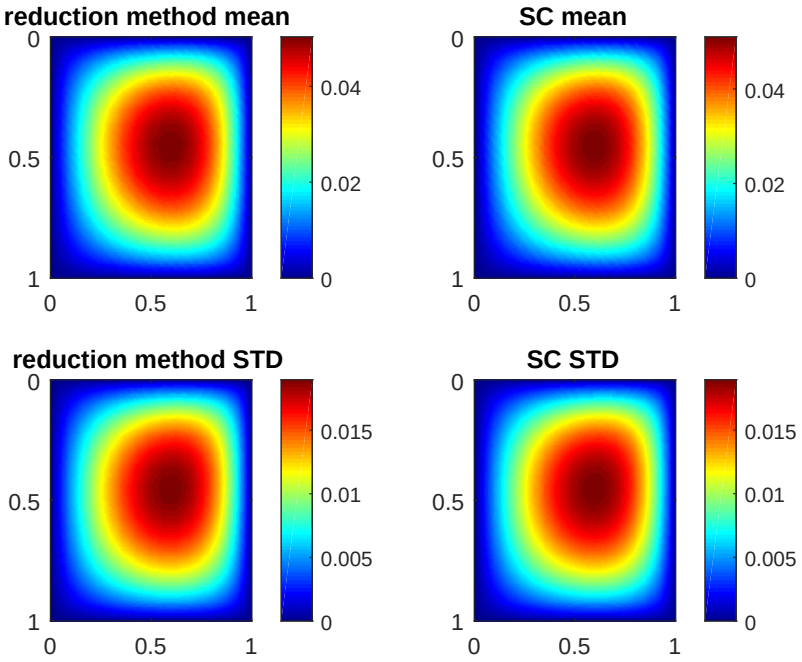

Figure 4: Profiles of the mean and STD solution.

where $p_1 = 1.6$ and $\varepsilon_1 = \frac{1}{14}$. In our test, the highest order of the gPC basis functions is 6 and the number of the POD modes is 4. We change the coarse mesh grid from 4×4 to 32×32, compare the results on different meshes, and calculate the numerical error with respect to the reference solution obtained on the fine mesh with size $\frac{1}{256}$. In Fig.5, we plot the convergence results with respect to meshsize H in both the L^2 norm and the H^1 norm. We approximately obtain a second-order convergence for the error in the L^2 norm and a first-order convergence for the error in the H^1 norm.

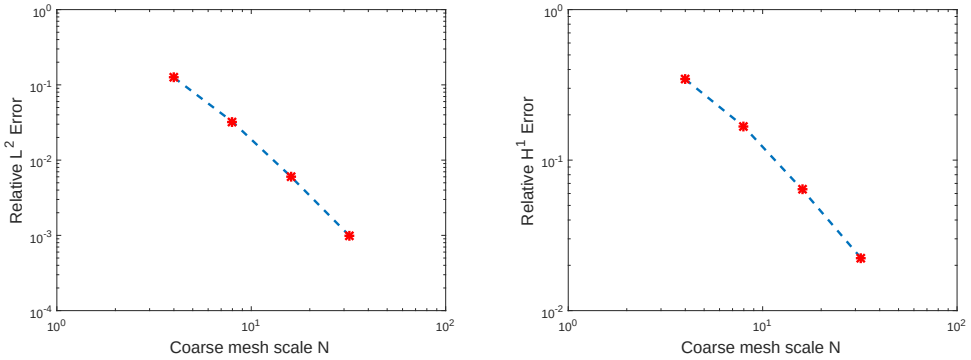

Figure 5: Convergence results with respect to mesh size, where $H = \frac{1}{4}$, $\frac{1}{8}$, $\frac{1}{16}$ and $\frac{1}{32}$. Left is for the mean error in the L^2 norm and the slope is -2.3. Right is for the mean error in the H^1 norm and the slope is -1.3.

Verification of the convergence rate with respect to the polynomial order. We fix the course mesh size to be $H = \frac{1}{16}$ and number of the POD modes to be 4. We choose the coarse mesh and POD modes in such a way that the error from physical space and the POD error are small and the error is dominant by the random space. We test two different types of coefficients. In the first case, the coefficient is given by Eq.(61) and we change the polynomial order from 1 to 7. In Fig.6, we plot convergence results with respect to different polynomial orders. In the second case, the coefficient is given by Eq.(60), which is parameterized by three random variables and we change the highest order of the gPC basis from 1 to 6. One can see the exponential decay of the error with respect to the polynomial orders.

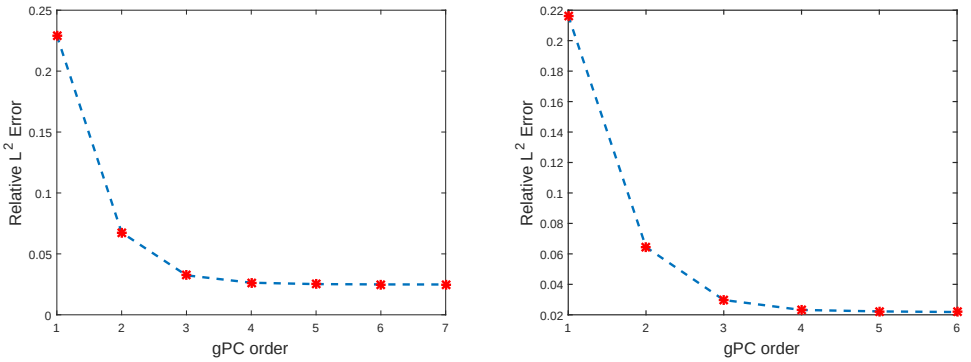

Figure 6: Convergence results with respect to the polynomial chaos order, where $H = \frac{1}{16}$. Left is for the one r.v. case. Right is for the three r.v.s case.

Investigate the decay of error for POD modes. We consider the multiscale problem with the coefficient (60) and choose the coarse mesh grid as 16×16, the fine mesh grid as 256×256, and the polynomial order of the gPC as 4. In Fig.7, we show the relative error of the mean with respect to the

increasing of the number of the POD modes. Meanwhile, we show the cumulative sum of the eigenvalue ratio in the POD decomposition. One can find that when we increases the number of the POD modes (namely the multiscale reduced basis), we observe a fast decay in the relative error of the mean and fast increasing in the cumulative sum of the eigenvalue ratio in the POD decomposition, which implies that the POD basis can efficiently explore the low-dimensional structures of the multiscale random solution. We observe the qualitative decay of the error for the STD (not shown here).

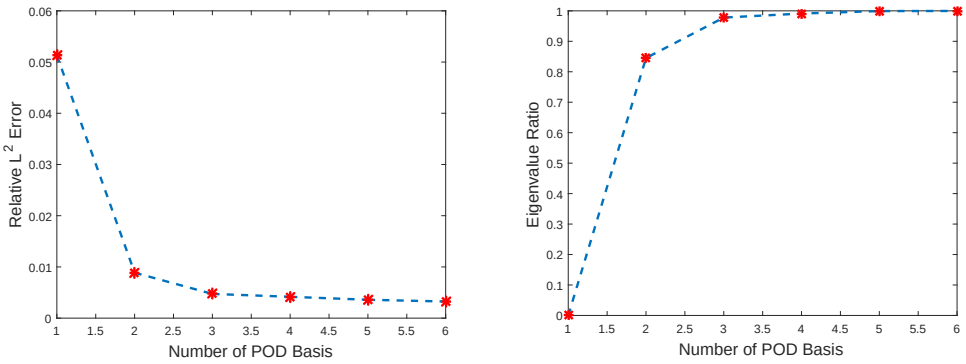

Figure 7: Convergence results with respect to the number of the POD modes.

5.2. An example whose coefficient does not have scale separation

We consider the multiscale elliptic PDE (58)-(59) with a coefficient $a^\varepsilon(x,y,\omega)$, which is a random linear combination of three deterministic coefficient fields plus a constant, i.e.,

$$a^\varepsilon(x,y,\omega) = \sum_{i=1}^{3} \xi_i(\omega) k_i(x,y) + 0.5 \tag{62}$$

where $\{\xi_i\}_{i=1}^3$ are independent uniform random variables in $[0,1]$, and $k_i(x,y)$, $i = 1,...,3$ are deterministic coefficient fields without scale separation. Specifically, $k_i(x,y) = |\theta_i(x,y)|$, where $\theta_i(x,y)$, $i = 1,...,3$ are defined on a 5×5, 9×9, and 17×17 grids over the domain D. For each grid cell, the value of $\theta_i(x,y)$ is normally distributed. In Fig.8, we show four samples of the coefficient $a^\varepsilon(x,y,\omega)$. One can see that the coefficient does not have any periodic structure.

Multiquery results in the online stage. The implementation of the SFEM and our method are exactly the same as in the previous example. In the online stage we use them to solve the problem (58) with different $f(x,y)$. We randomly generate 20 force functions $f(x,y) \in \{\sin(k_i\pi x + l_i)\cos(m_i\pi x + n_i\}_{i=1}^{20}$, where k_i, l_i, m_i, and n_i are random numbers. In Fig.9, we show the relative error for the mean obtained using our method in the L^2 norm and the H^1 norm, respectively. The results for the STD error are similar (not shown here).

Verification of the convergence rate with respect to meshsize H. We test the problem (58)-(59) with coefficient (62). We choose the highest order of the gPC basis functions as 4 and the number of POD modes as 4. We choose the coarse mesh grids as 4×4 , 8×8, 16×16, and 32×32. We compare the results on different grids and compute the numerical error with respect to the reference solution obtained on the fine mesh $\frac{1}{256}$. In Fig.10, we plot the convergence results with respect to meshsize H. The convergence rate is 0.84 in the L^2 norm and is 0.65 in the H^1 norm since other sources of error may become dominant in this example.

Verification of the convergence rate with respect to the polynomial order. We fix the course mesh size to be $H = \frac{1}{16}$ and keep the number of the POD modes to be 4. Then we test two different

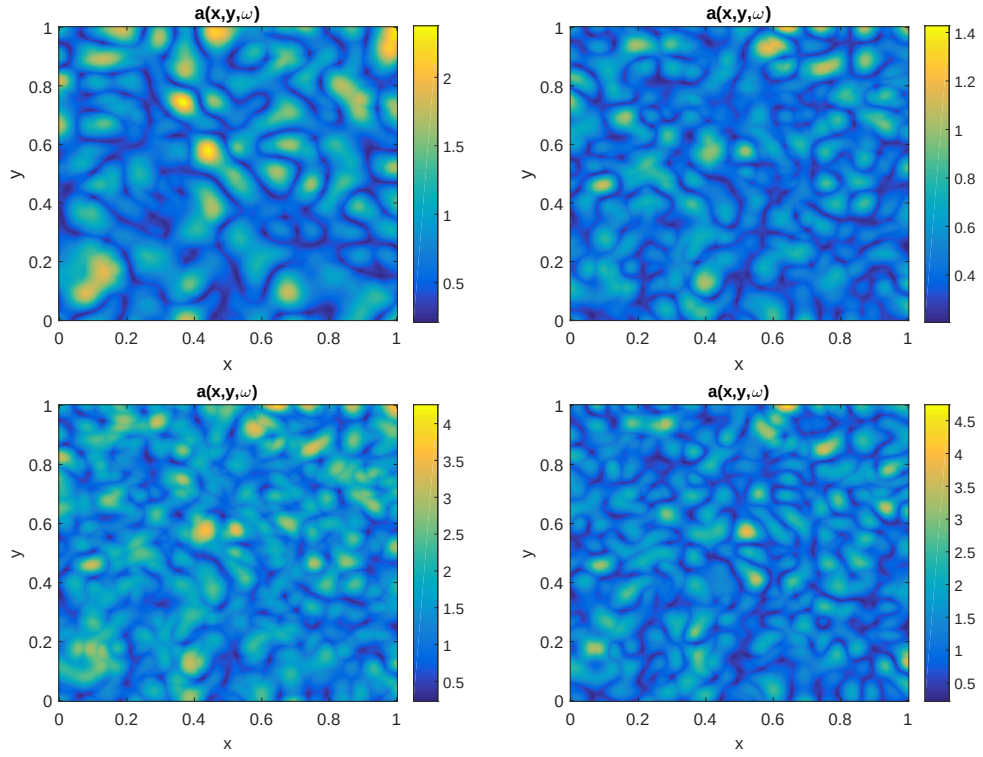

Figure 8: Four coefficient samples of $a(x, y, \omega)$.

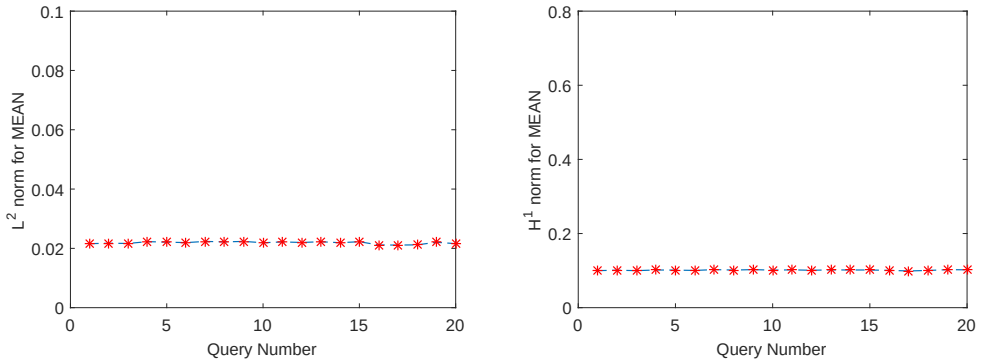

Figure 9: The mean error of our method in the L^2 norm and the H^1 norm.

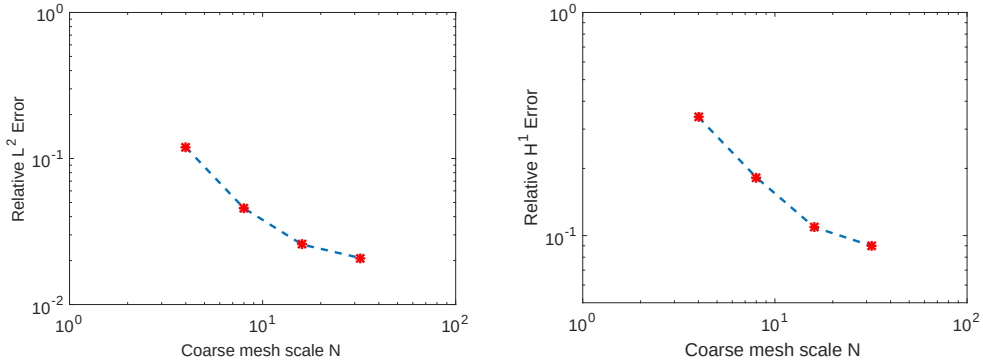

Figure 10: Convergence results with respect to mesh size. Left is for the error in the L^2 norm and the slope is -0.84. Right is for the error in H^1 norm and the slope is -0.65.

coefficients. In the first case, the coefficient is given by $a^\varepsilon(x,y,\omega) = \xi_1(\omega)k_1(x,y) + 0.5$, which is parameterized by one random variable. We change the polynomial order from 1 to 7. In Fig.11, we plot the convergence of the relative mean error with respect to polynomial orders. In the second case, the coefficient is given by (62) and we change the highest polynomial order from 1 to 6. One can see the errors decay exponentially with respect to the polynomial order. Moreover, one can observe that the errors get stagnant since other sources of error may become dominant.

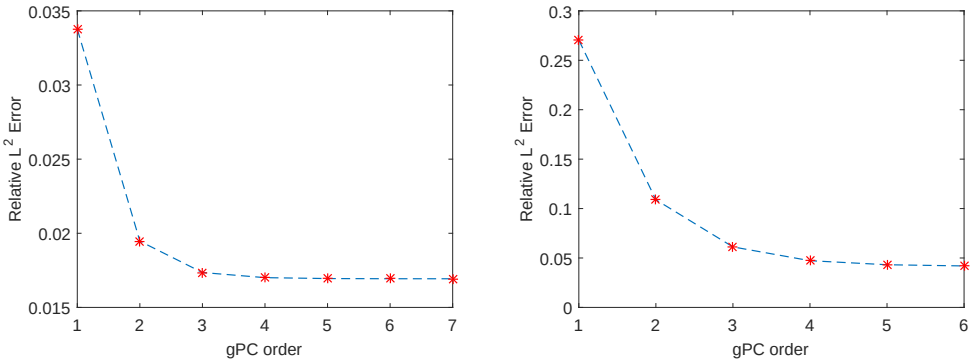

Figure 11: Convergence results with respect to the polynomial chaos order, where $H = \frac{1}{16}$ and POD modes 4. Left is for one r.v. case. Right is for three r.v.s case.

Investigate the decay of error for POD modes. We consider the multiscale problem with the coefficient (62) and choose the coarse mesh grid as 16×16, the fine mesh grid as 256×256, and the polynomial order of the gPC as 4. In Fig.12, we show the relative error of the mean with respect to the increasing of the number of the POD modes (namely the multiscale reduced basis). We find that the relative error of the mean decreases when we increases the number of the POD modes. We observe the qualitative decay of the error for the STD (not shown here).

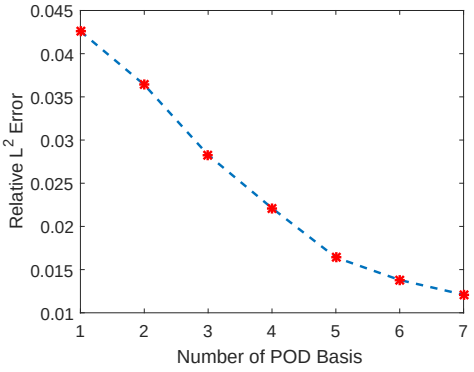

Figure 12: Convergence results with respect to the number of the POD modes.

5.3. An example with a high-contrast coefficient

Finally, we consider the multiscale elliptic PDE (58)-(59) with a high-contrast coefficient. The coefficient $a^\varepsilon(x,y,\omega)$ is given by

$$a^\varepsilon(x,y,\omega) = 0.2 + 20\xi_1(\omega)\mathbf{1}_{D_1}(x,y) + 60\xi_2(\omega)\mathbf{1}_{D_2}(x,y) + 100\xi_3(\omega)\mathbf{1}_{D_3}(x,y), \tag{63}$$

where $\xi_i(\omega) \in U[0,1]$ and $\mathbf{1}_{D_i}(x,y)$, $i = 1,2,3$ are indicator functions. The local domains $D_1 = B_1(x_1,0.1) \cup B_2(x_2,0.1)$, $D_2 = [0.125,0.875] \times [0.375,0.5]$, $D_3 = B_3(x_3,0.1) \cup B_4(x_4,0.1) \cup B_5(x_5,0.1)$,

where $B_i(x_i, r)$ are the circles with the radius $r = 0.1$ and the centers at $x_1 = (0.25, 0.75)$, $x_2 = (0.625, 0.75)$, $x_3 = (0.1875, 0.1875)$, $x_4 = (0.5, 0.1875)$, and $x_5 = (0.8125, 0.1875)$. In Fig.13, we show two samples of the coefficient $a^\varepsilon(x, y, \omega)$. One can see that the coefficient (63) has high-contrast features, where the highest contrast is $\frac{100}{0.2} = 500$.

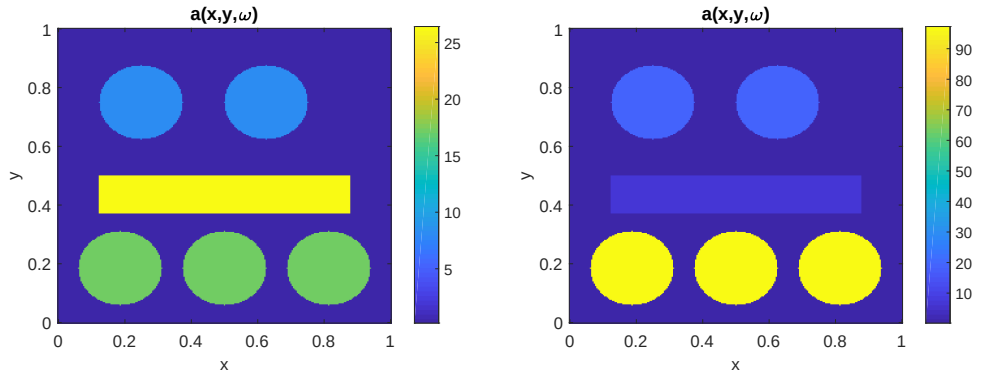

Figure 13: Two samples of the coefficient $a(x, y, \omega)$.

Multiquery results in the online stage. The implementation of the SFEM and our method are exactly the same as in the previous two examples. We randomly generate 20 force functions $f(x, y) \in \{\sin(k_i\pi x + l_i)\cos(m_i\pi x + n_i\}_{i=1}^{20}$, where k_i, l_i, m_i, and n_i are random numbers. In Fig.14, we show the relative error for the mean obtained using our method in the L^2 norm and the H^1 norm, respectively. The results for the STD error are similar (not shown here). Thus, our method is efficient in solving this problem with a high-contrast coefficient.

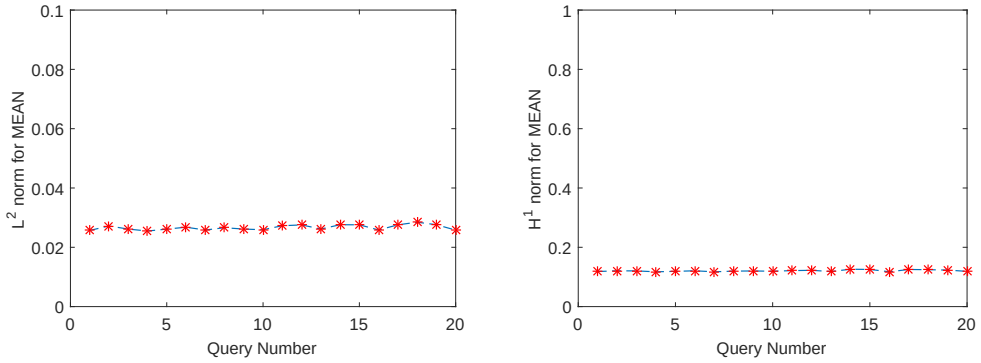

Figure 14: The mean error of our method in L^2 norm and H^1 norm.

Verification of the convergence rate with respect to the meshsize H. We test the problem (58)-(59) with coefficient (63). We choose the highest order of the gPC basis functions as 4 and the number of POD modes as 4. We choose the coarse mesh grids as 4×4 , 8×8, 16×16, and 32×32. We compare the results on different grids and compute the numerical error with respect to the reference solution obtained on the fine mesh $\frac{1}{256}$. In Fig.15, we plot the convergence results with respect to meshsize H. The convergence rate is 1.8 in the L^2 norm and is 1.12 in the H^1 norm.

Investigate the decay of error for POD modes. Finally, we choose the coarse mesh grid as 16×16, the fine mesh grid as 256×256, and the polynomial order of the gPC as 4. In Fig.16, we show the relative error of the mean with respect to the increasing of the number of the POD modes (namely the multiscale reduced basis). We find that the relative error of the mean decreases when we increases

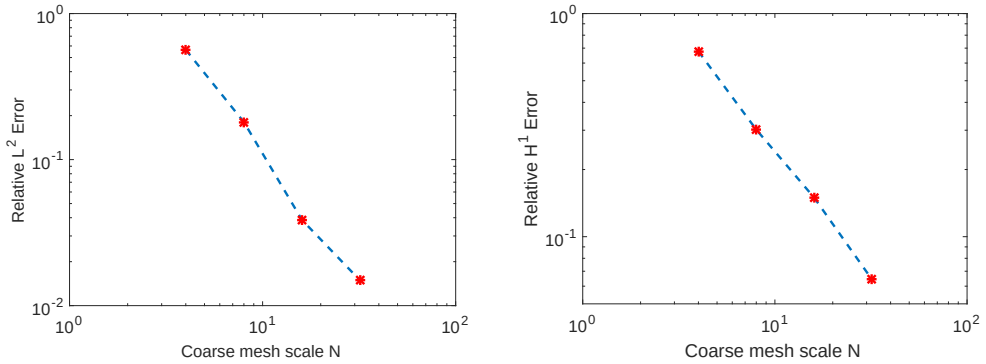

Figure 15: Convergence results with respect to mesh size. Left is for the mean error in the L^2 norm and the slope is -1.80. Right is for the mean error in the H^1 norm and the slope is -1.12.

the number of the POD modes. We observe the qualitative decay of the error for the STD (not shown here).

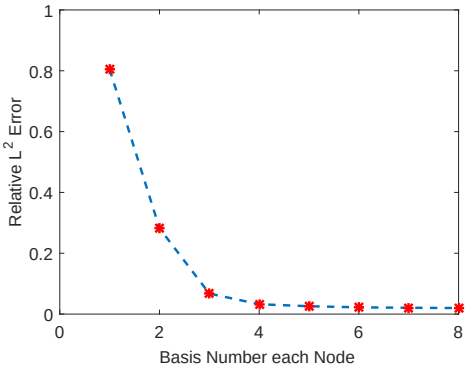

Figure 16: Convergence results with respect to the number of the POD modes.

6. Conclusion

In this paper, we proposed a proper orthogonal decomposition method for multiscale elliptic PDEs with random coefficients. Our method consists of offline and online stages. In the offline stage, we constructed a set of reduced basis functions within each coarse grid block using the POD method, which allows us to explore the low-dimensional structures hidden in the solution space. In the online stage, we can efficiently solve the multiscale SPDEs using our multiscale reduced basis functions. Under mild conditions, we analyse the error between the numerical solution obtained from our method and the exact solution. We presented several numerical examples for 2D stochastic elliptic PDEs with stochastic multiscale coefficients to demonstrate the accuracy and efficiency of our proposed method.

7. Acknowledgements

The research of Z. Zhang is supported by the Hong Kong RGC grants (Projects 27300616 and 17300318), National Natural Science Foundation of China (Project 11601457), and Seed Funding Programme for Basic Research (HKU). W. Ching and Z. Zhang are supported by the IMR and RAE Research Fund, Faculty of Science, The University of Hong Kong.

References

[1] H. Abdi and L. Williams. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4):433–459, 2010.

[2] G. Allaire and R. Brizzi. A multiscale finite element method for numerical homogenization. *SIAM Multiscale Model Simul.*, 4(3):790–812, 2005.

[3] M. Arnst and R. Ghanem. Probabilistic equivalence and stochastic model reduction in multiscale analysis. *Comput. methods Appl. Mech. Engrg*, 197(43):3584–3592, 2008.

[4] R. Askey and J. Wilson. *Some basic hypergeometric orthogonal polynomials that generalize Jacobi polynomials*, volume 319. American Mathematical Soc., 1985.

[5] B. V. Asokan and N. Zabaras. A stochastic variational multiscale method for diffusion in heterogeneous random media. *Journal of Computational Physics*, 218:654–676, 2006.

[6] I. Babuska, G. Caloz, and E. Osborn. Special finite element methods for a class of second order elliptic problems with rough coefficients. *SIAM J. Numer. Anal.*, 31:945–981, 1994.

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

[8] I. Babuska, R. Tempone, and G. Zouraris. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM J. Numer. Anal.*, 42:800–825, 2004.

[9] G. Berkooz, P. Holmes, and J. L. Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *Annual review of fluid mechanics*, 25(1):539–575, 1993.

[10] C Canuto and A Quarteroni. Approximation results for orthogonal polynomials in Sobolev spaces. *Mathematics of Computation*, 38(157):67–86, 1982.

[11] J. Charrier. Strong and weak error estimates for elliptic partial differential equations with random coefficients. *SIAM Journal on numerical analysis*, 50(1):216–246, 2012.

[12] Z. Chen and T. Y. Hou. A mixed multiscale finite element method for elliptic problems with oscillating coefficients. *Math. Comp.*, 72:541–576, 2002.

[13] M. Cheng, T. Y. Hou, M. Yan, and Z. Zhang. A data-driven stochastic method for elliptic PDEs with random coefficients. *SIAM J. UQ*, 1:452–493, 2013.

[14] A. Cohen, R. Devore, and C. Schwab. Convergence Rates of Best N-term Galerkin Approximations for a Class of elliptic sPDEs. *Found. Comput. Math.*, 10(6):615–646, 2010.

[15] W. E and B. Engquist. The heterogeneous multi-scale methods. *Comm. Math. Sci.*, 1:87–133, 2003.

[16] Y. R. Efendiev, T. Y. Hou, and X. Wu. Convergence of a nonconforming multiscale finite element method. *SIAM Journal on Numerical Analysis*, 37(3):888–910, 2000.

[17] B. Ganapathysubramanian and N. Zabaras. Modelling diffusion in random heterogeneous media: Data-driven models, stochastic collocation and the variational multi-scale method. *Journal of Computational Physics*, 226:326–353, 2007.

[18] R. Ghanem and P. Spanos. *Stochastic finite elements: a spectral approach.* Springer-Verlag, New York, 1991.

[19] I. G. Graham, F. Y. Kuo, J. A. Nichols, R. Scheichl, C. Schwab, and I. H. Sloan. Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficients. *Numerische Mathematik*, 131(2):329–368, 2015.

[20] H. Han and Z. Zhang. Multiscale tailored finite point method for second order elliptic equations with rough or highly oscillatory coefficients. *Commun. Math. Sci.*, 10:945–976, 2012.

[21] A Hannachi, I. Jolliffe, and D. Stephenson. Empirical orthogonal functions and related techniques in atmospheric science: A review. *International journal of climatology*, 27(9):1119–1152, 2007.

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

[23] T. Y. Hou, W. Luo, B. Rozovskii, and H. M. Zhou. Wiener chaos expansions and numerical solutions of randomly forced equations of fluid mechanics. *J. Comput. Phys.*, 216:687–706, 2006.

[24] T. Y. Hou and X. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.*, 134:169–189, 1997.

[25] T. Y. Hou, X. Wu, and Z. Cai. Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. *Mathematics of Computation*, 68(227):913–943, 1999.

[26] T. Y. Hou, X. Wu, and Y. Zhang. Removing the cell resonance error in the multiscale finite element method via a Petrov-Galerkin formulation. *Communications in Mathematical Sciences*, 2(2):185–205, 2004.

[27] T.Y. Hou, P. Liu, and Z. Zhang. A localized data-driven stochastic method for elliptic PDEs with random coefficients. *Bull. Inst. Math. Acad. Sin. (N.S.)*, 1:179–216, 2011.

[28] T. J. Hughes, G. R. Feijoo, L. Mazzei, and J. B. Quincy. The variational multiscale methoda paradigm for computational mechanics. *Comput. Methods Appl. Mech. Eng.*, 166(1-2):3–24, 1998.

[29] T. Iliescu and Z. Wang. Variational multiscale proper orthogonal decomposition: Convection-dominated convection-diffusion-reaction equations. *Mathematics of Computation*, 82(283):1357–1378, 2013.

[30] P. Jenny, S. Lee, and Tchelepi. H. Multiscale finite volume method for elliptic problems in subsurface flow simulation. *J. Comput. Phys.*, 187:47–67, 2003.

[31] K. Karhunen. Uber lineare methoden in der Wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fennicae. Ser. A. I. Math.-Phys.*, 37:1–79, 1947.

[32] I. G. Kevrekidis, C. W. Gear, J. M. Hyman, P. G. Kevrekidid, O. Runborg, and C. Theodoropoulos. Equation-free, coarse-grained multiscale computation: Enabling mocroscopic simulators to perform system-level analysis. *Communications in Mathematical Sciences*, 1(4):715–762, 2003.

[33] M. Loève. *Probability theory. Vol. II, 4th ed. GTM. 46.* Springer-Verlag, ISBN 0-387-90262-7, 1978.

[34] A. Malqvist and D. Peterseim. Localization of elliptic multiscale problems. *Mathematics of Computation*, 83(290):2583–2603, 2014.

[35] H. G. Matthies and A. Keese. Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations. *Comput. Method Appl. Mech. Eng.*, 194:1295–1331, 2005.

[36] J.M. Melenk and I. Babuška. The partition of unity finite element method: basic theory and applications. *Comput. Methods. Appl. Mech. Eng.*, 139(1-4):289–314, 1996.

[37] H. N. Najm. Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics. *Annual Review of Fluid Mechanics*, 41:35–52, 2009.

[38] F. Nobile, R. Tempone, and C. Webster. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. Numer. Anal.*, 46:2309–2345, 2008.

[39] G. North, T. Bell, R. Cahalan, and F. Moeng. Sampling errors in the estimation of empirical orthogonal functions. *Monthly Weather Review*, 110(7):699–706, 1982.

[40] B. Oksendal. *Stochastic Differential Equations: an introduction with applications.* Springer Science and Business Media, 2013.

[41] H. Owhadi and L. Zhang. Metric based up-scaling. *Comm. Pure Appl. Math.*, LX:675–723, 2007.

[42] G. Rozza, D. B. Huynh, and A. T. Patera. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. *Archives of Computational Methods in Engineering*, 15(3):1–47, 2007.

[43] T. Sapsis and P. Lermusiaux. Dynamically orthogonal field equations for continuous stochastic dynamical systems. *Physica D: Nonlinear Phenomena*, 238:2347–2360, 2009.

[44] L. Sirovich. Turbulence and the dynamics of coherent structures. I. Coherent structures. *Quarterly of applied mathematics*, 45(3):561–571, 1987.

[45] J. Wan and N. Zabaras. A probabilistic graphical model approach to stochastic multiscale partial differential equations. *Journal of Computational Physics*, 250:477–510, 2013.

[46] X. L. Wan and G. Karniadakis. Multi-element generalized polynomial chaos for arbitrary probability measures. *SIAM J. Sci. Comp.*, 28:901–928, 2006.

[47] K. Willcox and J. Peraire. Balanced model reduction via the proper orthogonal decomposition. *AIAA journal*, 40(11):2323–2330, 2002.

[48] D. Xiu and G. Karniadakis. Modeling uncertainty in flow simulations via generalized polynomial chaos. *J. Comput. Phys.*, 187:137–167, 2003.

[49] D. X. Zhang. *Stochastic methods for flow in porous media: coping with uncertainties.* Academic press, 2001.

[50] Z. Zhang, M. Ci, and T. Y. Hou. A multiscale data-driven stochastic method for elliptic PDEs with random coefficients. *SIAM Multiscale Model. Simul.*, 13:173–204, 2015.