An adaptive dynamically low-dimensional approximation method for multiscale stochastic diffusion equations

Eric T. Chung[a], Sai-Mang Pun[a], Zhiwen Zhang[b,*]

[a] *Department of Mathematics, The Chinese University of Hong Kong, Shatin, Hong Kong.*
[b] *Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong.*

Abstract

In this paper, we propose a dynamically low-dimensional approximation method to solve a class of time-dependent multiscale stochastic diffusion equations. In [5, 6], a dynamically bi-orthogonal (DyBO) method was developed to explore low-dimensional structures of stochastic partial differential equations (SDPEs) and solve them efficiently. However, when the SPDEs have multiscale features in the physical space, the original DyBO method becomes expensive. To address this issue, we construct multiscale basis functions within each coarse grid block for dimension reduction in the physical space. To further improve the accuracy, we also perform online procedure to construct online adaptive basis functions. In the stochastic space, we use the generalized polynomial choas (gPC) basis functions to represent the stochastic part of the solutions. Numerical results are presented to demonstrate the efficiency of the proposed method in solving the time-dependent PDE with multiscale and random features.

Keywords: Uncertainty quantification; dynamically low-dimensional approximation; online adaptive method; stochastic partial differential equations (SPDEs); generalized multiscale finite element method (GMsFEM).

1. Introduction

Uncertainty arises in many complicated real-world problems of physical and scientific applications, such as wave, heat, and pollution propagation through random media or flow driven by stochastic forces. These kind of problems usually have multiple scale features involved in the spatial domain. For example, to simulate flows in heterogeneous porous media, the permeability field is often parameterized by random fields with multiple-scale structures.

Stochastic partial differential equations (SPDEs), which contain random variables or stochastic processes, play important roles in modeling complex problems and quantifying the corresponding uncertainties. Considerable amounts of efforts have been devoted to study SPDEs, see [17, 27, 2, 29, 20, 36, 31, 39, 30, 34, 26] and references therein. These methods are effective when the dimension of stochastic input variables is low. However, their performance deteriorates when the dimension of stochastic input variables is high because of the curse of dimensionality. In addition, when SPDEs have multiscale features, the SPDE problems become more challenging as it requires tremendous computational resources to resolve the small scales of the SPDE solutions. This motivates us to develop efficient numerical methods to solve these challenging problems.

*Corresponding author
Email addresses: tschung@math.cuhk.edu.hk (Eric T. Chung), smpun@math.cuhk.edu.hk (Sai-Mang Pun), zhangzw@hku.hk (Zhiwen Zhang)

In this paper, we consider the time-dependent SPDEs with multiscale coefficients as follows

$$\frac{\partial u^\varepsilon}{\partial t}(x,t,\omega) = \mathcal{L}^\varepsilon u^\varepsilon(x,t,\omega), \quad x \in \mathcal{D}, \quad t \in (0,T], \quad \omega \in \Omega, \tag{1}$$

where suitable boundary and initial conditions are imposed, $\mathcal{D} \subset \mathbb{R}^d$ is a bounded spatial domain, Ω is a sample space, and $\mathcal{L}^\varepsilon$ is an elliptic operator that contains multiscale and random coefficient, where the smallest-scale is parameterized by ε. The major difficulties in solving (1) come from two parts. In the physical space, we need a mesh fine enough to resolve the small-scale features. In the random space, we need extra degree of freedom to represent the random features. Moreover, the problem (1) becomes more difficult if the dimension of the random input is high.

To address these challenges, we shall explore low-dimensional structures hidden in the solution $u^\varepsilon(x,t,\omega)$. Specifically, if the solution $u^\varepsilon(x,t,\omega)$ is a second-order stochastic process at each time $t > 0$, i.e., $u^\varepsilon(x,t,\omega) \in L^2(\mathcal{D} \times \Omega)$, we can approximate the solution $u^\varepsilon(x,t,\omega)$ by its m-term truncated Karhunen-Loève (KL) expansion

$$u^\varepsilon(x,t,\omega) \approx \bar{u}^\varepsilon(x,t) + \sum_{i=1}^{m} u_i^\varepsilon(x,t) Y_i(\omega,t) = \bar{u}^\varepsilon(x,t) + \mathbf{U}(x,t)\mathbf{Y}^T(\omega,t), \tag{2}$$

where $\mathbf{U}(x,t) = (u_1^\varepsilon(x,t), \cdots, u_m^\varepsilon(x,t))$ and $\mathbf{Y}(\omega,t) = (Y_1(\omega,t), \cdots, Y_m(\omega,t))$. The KL expansion explores the intrinsic low-dimensional structures of the SPDE solutions and gives the compact representation of the solutions. However, the computation of the KL expansion can be quite expensive since we need to form a covariance kernel and solve a large-scale eigenvalue problem. More details will be given in Section 2.1.

In [5, 6], a dynamically bi-orthogonal (DyBO) method was developed. This new method derives an equivalent system that can faithfully track the KL expansion of the SPDE solution. In other words, the DyBO method gives the evolution equations for $\bar{u}$, $\mathbf{U}$, and $\mathbf{Y}$. The DyBO method can accurately and efficiently solve many time-dependent SPDEs, such as the stochastic Burger's equation and stochastic Navier-Stokes equation, with considerable savings over existing stochastic methods. However, when the SPDEs have multiscale features in the physical space, the original DyBO method becomes expensive as one needs enormous degree of freedom to represent the multiscale features in the physical space. To overcome this difficulty, we shall apply the generalized multiscale finite element methods (GMsFEM) [10, 15] to construct multiscale basis functions within each coarse grid block for model reduction in the physical space.

In the GMsFEM, we divide the computation into two stages: the offline computing and the online computing. In the offline stage, we first construct global snapshot functions within each coarse neighborhood based on the given coarse and fine meshes and seek multiscale basis functions to represent the local heterogeneities. When the snapshot functions are computed, one can construct the multiscale basis functions in each coarse patch by solving some well-designed local spectral problems and identify the crucial multiscale basis functions to form the offline function space. In the online stage, we add more online multiscale basis functions that are constructed using the offline space. These online basis functions are computed adaptively in some selected spatial regions based on the current local residuals and their construction is motivated by the analysis in [11]. In general, it can guarantee that additional online multiscale basis functions will reduce the error rapidly if one chooses a sufficient number of offline basis functions. We should point out that there are many existing methods in the literature to solve multiscale problems, see [3, 13, 14, 22, 23, 32, 1, 38, 19, 28] and references therein. However, most of these methods are designed for multiscale problems with deterministic coefficients.

In our new method, we first derive the DyBO formulation for the multiscale SPDEs (1), which consists of deterministic PDEs for $\bar{u}$ and $\mathbf{U}$ respectively and ODE system for the stochastic basis

$\mathbf{Y}$. For the deterministic PDEs (for $\bar{u}$ and $\mathbf{U}$) in the formulation, we shall apply the GMsFEM to construct multiscale basis functions and use these multiscale basis functions to represent $\bar{u}$ and $\mathbf{U}$, which lead to considerable savings over original DyBO method. For the ODE system, the memory cost is relatively small and we shall apply a suitable ODE solver to compute the numerical solution. The GMsFEM enables us to significantly improve the efficiency of the DyBO method in solving time-dependent multiscale PDEs with random input.

The rest of the paper is organized as follows. In Section 2, we give a brief introduction of the KL expansion and the generalized polynomial chaos (gPC) method. In Section 3, we introduce the framework of DyBO formulation. The GMsFEM with the online adaptive construction will be introduced in Section 4. The implementation issues of the algorithm and the numerical results will be given in Section 5. Finally, some concluding remarks are given in Section 6.

2. Some preliminaries

2.1. The Karhunen-Loève expansion

In the theory of stochastic processes, the Karhunen-Loève (KL) expansion [24, 25] is a representation of a stochastic process as an infinite linear combination of orthogonal functions, analogous to a Fourier series representation of a function on a bounded interval. The importance of the KL expansion is that it yields an optimal basis in the sense that it minimizes the total mean square error.

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, whose sample space is Ω and is equipped with σ-algebra $\mathcal{F}$ and probability measure $\mathbb{P}$. Suppose $u(x,t,\omega)$, defined on a compact spatial domain $\mathcal{D} \subset \mathbb{R}^d$, $t \in [0,T]$, is a second-order stochastic process, i.e., $u(\cdot,t,\cdot) \in L^2(\mathcal{D}\times\Omega)$, where we omit ε in the solution u to simplify notations. Its KL expansion reads as follows

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

where $\bar{u}(x,t) = \mathbb{E}[u(x,t,\omega)]$, and $\{\mu_i(t), \phi_i(x,t)\}_{i=1}^{\infty}$ are the eigenpairs of the covariance kernel $C(x,y,t)$, i.e.,

$$\int_{\mathcal{D}} C(x,y,t)\phi_i(y,t)\,\mathrm{d}y = \mu_i(t)\phi_i(x,t), \quad i = 1,2,\cdots. \tag{3}$$

The covariance kernel $C(x,y,t)$ is defined as

$$C(x,y,t) = \mathbb{E}\left[(u(x,t,\omega) - \bar{u}(x,t))(u(y,t,\omega) - \bar{u}(y,t))\right]. \tag{4}$$

The random variables $\{\xi_i(\omega,t)\}_{i=1}^{\infty}$ are defined as

$$\xi_i(\omega,t) = \frac{1}{\sqrt{\mu_i(t)}}\int_{\mathcal{D}} (u(x,t,\omega) - \bar{u}(x,t))\phi_i(x,t)\,\mathrm{d}x, \quad i = 1,2,\cdots,. \tag{5}$$

Moreover, these random variables $\{\xi_i\}_{i=1}^{\infty}$ are of zero-mean and uncorrelated, i.e. $\mathbb{E}[\xi_i] = 0$, $\mathbb{E}[\xi_i\xi_j] = \delta_{ij}$. Generally, the eigenvalues μ_i's are sorted in a descending order. Their decay rates depend on the regularity of the covariance kernel $C(x,y)$. It has been proven that an algebraic decay rate, i.e. $\mu_k = O(k^{-\gamma})$, is achieved asymptotically if the covariance kernel is of finite Sobolev regularity or an exponential decay, i.e., $\mu_k = O(e^{-\gamma k})$ for some $\gamma > 0$, if the covariance kernel is piecewise analytic [35]. In general, the decay rate depends on the correlation length of the stochastic solution. Small correlation length results in slow decay of the eigenvalues.

The m-term truncated KL expansion converges in $L^2(\mathcal{D}\times\Omega)$ to the original stochastic process $u(x,t,\omega)$ as m tends to infinity. If we denote by ε_m the truncation error, we have

$$\|\varepsilon_m\|^2_{L^2(\mathcal{D}\times\Omega)} = \left\| \sum_{i=m+1}^{\infty} \sqrt{\mu_i(t)}\xi_i(\omega,t)\phi_i(x,t) \right\|^2_{L^2(\mathcal{D}\times\Omega)} = \sum_{i=m+1}^{\infty} \mu_i(t) \to 0, \quad m\to\infty, \tag{6}$$

where we have used the bi-orthogonality of the KL expansion. In practice, we truncate the KL expansion into its first m terms and obtain the following truncated KL expansion

$$u(x,t,\omega) \approx \bar{u}(x,t) + \sum_{i=1}^{m} \sqrt{\mu_i(t)}\xi_i(\omega,t)\phi_i(x,t). \tag{7}$$

The truncation error analysis in (6) reveals the most important property of KL expansion. Specifically, given any integer m and orthonormal basis $\{\varphi_i(x,t)\}_{i=1}^m$, we may approximate the stochastic process $u(x,t,\omega)$ by

$$u_{m,\psi}(x,t,\omega) = \bar{u}(x,t) + \sum_{i=1}^{m} \zeta_i(\omega,t)\varphi_i(x,t), \tag{8}$$

where $\zeta_i(\omega,t)$, $i=1,...,m$ are the expansion coefficients. Among all m-term approximations using basis $\{\varphi_i\}$, the KL expansion given by (7) is the one that minimizes the total mean square error. In this sense, we say that the KL expansion explores the low-dimensional structures of SDPE solutions.

2.2. The generalized Polynomial Chaos (gPC) basis

One of the conceptual difficulties, from the viewpoint of the classical numerical PDEs, involves representations of random variables or functions defined on the abstract probability space Ω. Essentially, there are different ways to represent the stochastic basis $\mathbf{Y}(\omega,t)$. In this work, we use the generalized polynomial chaos (gPC) basis to represent the stochastic basis, see [17, 40, 20]. Let $\{\mathrm{H}_i(\xi)\}_{i=1}^{\infty}$ denote the one-dimensional, $\rho(\xi)$-orthonormal polynomials, i.e.,

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

For the Gaussian distribution and the uniform distribution, such orthonormal polynomial sets are formed by Hermite polynomials and Legendre polynomials, respectively. For general distributions, such polynomial set can be obtained by numerical methods [37]. Furthermore, by a tensor product representation, we can use the one-dimensional polynomial $\mathrm{H}_i(\xi)$ to construct a complete set of orthonormal basis functions $\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$'s of $L^2(\Omega)$ as follows

$$\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \prod_{i=1}^{r} \mathrm{H}_{\alpha_i}(\xi_i), \quad \boldsymbol{\alpha}\in\mathfrak{J}_r^{\infty}, \tag{9}$$

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

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

The zero multi-index corresponding to $\mathbf{H}_{\mathbf{0}}(\boldsymbol{\xi}) = 1$, which is used to represent the mean of the solution, will be excluded in this work since we have an independent dynamic equation for the evolution of the mean. 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.,

$$\mathfrak{J}_r^p = \left\{ \boldsymbol{\alpha} \,|\, \boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_r), \alpha_i \geq 0, \alpha_i \in \mathbb{N}, |\alpha| = \sum_{i=1}^{r} \alpha_i \leq p \right\}. \tag{10}$$

The number of polynomial basis functions $|\mathfrak{J}_r^p|$ is equal to $\frac{(p+r)!}{p!r!}$. We may simply write such a truncated set as $\mathfrak{J}$ when no ambiguity arises. The orthonormal basis $\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$ is the standard generalized Polynomial Chaos (gPC) basis, see [7, 18, 21, 40] for more details.

3. The DyBO formulation for multiscale time-dependent SPDEs

3.1. Problem setting

To demonstrate the main idea of our method, we consider the following multiscale stochastic diffusion equation with suitable boundary and initial conditions,

$$\frac{\partial u^\varepsilon}{\partial t} = \mathcal{L}^\varepsilon u^\varepsilon \equiv \nabla \cdot (a^\varepsilon(x,\omega)\nabla u^\varepsilon) + f(x,\omega), \quad x \in \mathcal{D},\ t \in (0,T],\ \omega \in \Omega, \tag{11}$$

where $\mathcal{D} \subset \mathbb{R}^d$ $(d = 2, 3)$ is a bounded spatial domain and Ω is a sample space. The operator $\mathcal{L}^\varepsilon$ is uniformly elliptic almost surely, namely, there exist $a_{\min}, a_{\max} > 0$, such that

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

Also, the multiscale information is described by the parameter ε. The force $f(x,\omega)$ is assumed to be in $L^2(\mathcal{D})$ for each realization $\omega \in \Omega$.

In reservoir simulation, $a^\varepsilon(x,\omega)$ is used to model the permeability field, which is assumed to be known at every location of the physical domain. Realistically, however, only a handful of permeability measurements may be available. Thus the uncertainties in the inputs of the reservoir model could be quite large. It is important to analyze the uncertainties in the reservoir models and to quantify their influences on the prediction values. In this paper, $a^\varepsilon(x,\omega)$ is assumed to be parameterized by r independent random variables, i.e., $a^\varepsilon(x,\omega) \equiv a^\varepsilon(x, \xi_1(\omega), ..., \xi_r(\omega))$. By the Doob-Dynkin lemma [33], the solution $u^\varepsilon(x,t,\omega)$ to (11) is a functional of these random variables, which can be written as $u^\varepsilon = u^\varepsilon(x,t,\xi_1(\omega), ..., \xi_r(\omega))$.

3.2. The dynamically bi-orthogonal (DyBO) formulation

To make this paper self-contained, we give a brief review of the basic ideas and the main result of the DyBO method [5, 6]. We assume the stochastic solution $u^\varepsilon(x,t,\omega)$ to (11) is a second-order stochastic process at each fixed time $t > 0$, i.e., $u^\varepsilon(\cdot,t,\cdot) \in L^2(\mathcal{D} \times \Omega)$. We consider the following m-term truncated KL expansion

$$\widetilde{u}^\varepsilon(x,t,\omega) = \bar{u}^\varepsilon(x,t) + \sum_{i=1}^{m} u_i^\varepsilon(x,t) Y_i(\omega,t) = \bar{u}^\varepsilon(x,t) + \mathbf{U}(x,t)\mathbf{Y}^T(\omega,t) \approx u^\varepsilon(x,t,\omega), \tag{13}$$

where $\bar{u}^\varepsilon(x,t)$ is the mean of the solution,

$$\mathbf{U}(x,t) = (u_1^\varepsilon(x,t), \cdots, u_m^\varepsilon(x,t)) \quad \text{and} \quad \mathbf{Y}(x,t) = (Y_1(x,t), \cdots, Y_m(x,t))$$

are the spatial and stochastic modes, respectively. By using the DyBO method, we obtain the evolution equations for $\bar{u}^\varepsilon$, $\mathbf{U}$, and $\mathbf{Y}$ as follows

$$\frac{\partial \bar{u}^\varepsilon}{\partial t} = \mathbb{E}\left[\mathcal{L}\widetilde{u}^\varepsilon\right], \tag{14a}$$

$$\frac{\partial \mathbf{U}}{\partial t} = -\mathbf{U}\mathbf{D}^T + \mathbb{E}\left[\tilde{\mathcal{L}}\widetilde{u}^\varepsilon \mathbf{Y}\right], \tag{14b}$$

$$\frac{\mathrm{d}\mathbf{Y}}{\mathrm{d}t} = -\mathbf{Y}\mathbf{C}^T + \left\langle \tilde{\mathcal{L}}\widetilde{u}^\varepsilon, \mathbf{U}\right\rangle \mathbf{\Lambda}_{\mathbf{U}}^{-1}, \tag{14c}$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product in the space $L^2(\mathcal{D})$, $\mathbf{\Lambda}_{\mathbf{U}} = \mathrm{diag}(\langle \mathbf{U}^T, \mathbf{U}\rangle) \in \mathbb{R}^{m\times m}$, and $\tilde{\mathcal{L}}\widetilde{u}^\varepsilon = \mathcal{L}\widetilde{u}^\varepsilon - \mathbb{E}\left[\mathcal{L}\widetilde{u}^\varepsilon\right]$. We define an anti-symmetrization operator $\mathcal{Q}: \mathbb{R}^{k\times k} \to \mathbb{R}^{k\times k}$ and a partial anti-symmetrization operator $\tilde{\mathcal{Q}}: \mathbb{R}^{k\times k} \to \mathbb{R}^{k\times k}$ as follows:

$$\mathcal{Q}(\mathbf{M}) = \frac{1}{2}\left(\mathbf{M} - \mathbf{M}^T\right), \qquad \tilde{\mathcal{Q}}(\mathbf{M}) = \frac{1}{2}\left(\mathbf{M} - \mathbf{M}^T\right) + \mathrm{diag}(\mathbf{M}),$$

where $\mathbf{M} \in \mathbb{R}^{k\times k}$ is a matrix and $\mathrm{diag}(\mathbf{M})$ is a diagonal matrix whose diagonal entries are equal to those of matrix $\mathbf{M}$. Then, the matrices $\mathbf{C}, \mathbf{D} \in \mathbb{R}^{m\times m}$ in (14) can be solved uniquely from the following linear system

$$\mathbf{C} - \mathbf{\Lambda}_{\mathbf{U}}^{-1}\tilde{\mathcal{Q}}\left(\mathbf{\Lambda}_{\mathbf{U}}\mathbf{C}\right) = 0, \tag{15a}$$

$$\mathbf{D} - \mathcal{Q}\left(\mathbf{D}\right) = 0, \tag{15b}$$

$$\mathbf{D}^T + \mathbf{C} = G_*(\bar{u}^\varepsilon, \mathbf{U}, \mathbf{Y}), \tag{15c}$$

where the matrix $G_*(\bar{u}^\varepsilon, \mathbf{U}, \mathbf{Y}) = \mathbf{\Lambda}_{\mathbf{U}}^{-1}\left\langle \mathbf{U}^T, \mathbb{E}\left[\tilde{\mathcal{L}}\widetilde{u}^\varepsilon \mathbf{Y}\right]\right\rangle \in \mathbb{R}^{m\times m}$.

The Eqns.(14a)(14b) in the DyBO formulation (14) are time-dependent deterministic PDEs for the mean solution $\bar{u}^\varepsilon$ and the spatial basis modes $\mathbf{U}$. They are coupled to the third equation, a system of stochastic ODEs for the stochastic basis functions $\mathbf{Y}$. Various spatial discretization schemes, such as finite element method, finite difference method, and spectral method, along with ODE solvers can be used to solve the DyBO system (14). In this paper, however, we need to use fine mesh to represent $\bar{u}^\varepsilon$ and $\mathbf{U}$ as they contain multiscale features. To address this issue, we apply the generalized multiscale finite element method (GMsFEM) [15] in the physical space and the implicit Euler scheme in the temporal space to discretize the Eqns.(14a)(14b). The framework of GMsFEM will be presented in Section 4.

The Eq.(14c) in the DyBO formulation (14) is a random or stochastic ODE system (depending on the setting of the operator $\mathcal{L}$), which can be solved using existing ODE solves, such as Euler scheme or Runge-Kutta scheme. To represent the stochastic basis functions $Y_i(\omega, t)$, one can choose several different approaches, including ensemble representations in sampling methods (e.g., Monte carlo method, sparse-grid based stochastic collocation method, etc.) and spectral representations, such as the gPC basis or wavelet basis. In this work, we shall use the gPC basis to represent the stochastic basis functions $Y_i(\omega, t)$. The Cameron-Martin theorem [4] implies the stochastic modes $Y_i(\omega, t)$'s in the KL expansion (13) can be approximated by the linear combination of the polynomial chaos, i.e.,

$$Y_i(\omega, t) = \sum_{\boldsymbol{\alpha}\in\mathfrak{J}} \mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}(\omega)) A_{\boldsymbol{\alpha}i}(t), \quad i = 1, 2, \cdots, m, \tag{16}$$

or in a matrix form, if we write $\mathbf{H}\left(\boldsymbol{\xi}\right) = \left(\mathbf{H}_{\boldsymbol{\alpha}}\left(\boldsymbol{\xi}\right)\right)_{\boldsymbol{\alpha}\in\mathfrak{J}}$ as a row vector,

$$\mathbf{Y}(\omega, t) = \mathbf{H}\left(\boldsymbol{\xi}(\omega)\right)\mathbf{A}(t), \tag{17}$$

where $\mathbf{A} \in \mathbb{R}^{N_p \times m}$ and N_p is the number of polynomial basis functions. The KL expansion (13) now reads

$$\widetilde{u} = \bar{u} + \mathbf{U}\mathbf{A}^T\mathbf{H}^T.$$

We can derive equations for $\bar{u}$, $\mathbf{U}$ and $\mathbf{A}$, instead of $\bar{u}$, $\mathbf{U}$ and $\mathbf{Y}$. In other words, the stochastic modes $\mathbf{Y}$ are identified with a matrix $\mathbf{A} \in \mathbb{R}^{N_p \times m}$, which leads to the DyBO-gPC formulation of SPDE (11),

$$\frac{\partial \bar{u}}{\partial t} = \mathbb{E}\left[\mathcal{L}\widetilde{u}\right], \tag{18a}$$

$$\frac{\partial \mathbf{U}}{\partial t} = -\mathbf{U}\mathbf{D}^T + \mathbb{E}\left[\tilde{\mathcal{L}}\widetilde{u}\mathbf{H}\right]\mathbf{A}, \tag{18b}$$

$$\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} = -\mathbf{A}\mathbf{C}^T + \left\langle \mathbb{E}\left[\mathbf{H}^T\tilde{\mathcal{L}}\widetilde{u}\right], \mathbf{U}\right\rangle \mathbf{\Lambda}_{\mathbf{U}}^{-1}, \tag{18c}$$

where $\mathbf{C}(t)$ and $\mathbf{D}(t)$ can be solved from the linear system (15) with

$$G_*(\bar{u}, \mathbf{U}, \mathbf{Y}) = \mathbf{\Lambda}_{\mathbf{U}}^{-1}\left\langle \mathbf{U}^T, \mathbb{E}\left[\tilde{\mathcal{L}}\widetilde{u}\mathbf{Y}\right]\right\rangle = \mathbf{\Lambda}_{\mathbf{U}}^{-1}\left\langle \mathbf{U}^T, \mathbb{E}\left[\tilde{\mathcal{L}}\widetilde{u}\mathbf{H}\right]\right\rangle \mathbf{A}. \tag{19}$$

By solving the system (18), we have an approximate solution to SPDE (11)

$$u^{\text{DyBO-gPC}} = \bar{u} + \mathbf{U}\mathbf{A}^T\mathbf{H}^T.$$

The orthonormal property of $\mathbf{Y}$ implies that the columns of $\mathbf{A}$ are orthonormal, i.e., $\mathbf{A}^T\mathbf{A} = I_{m\times m} \in \mathbb{R}^{m\times m}$. We would like to point out that $\mathbf{A}\mathbf{A}^T \in \mathbb{R}^{N_p \times N_p}$ in general is not an identity matrix as $m \ll N_p$ if the SPDE solution has a low-dimensional structure.

Remark 3.1. The boundary conditions and initial conditions for each physical component, and the initial condition for each stochastic component can be obtained by projection of the initial and boundary conditions of $u(x,t,\omega)$ on the corresponding components. In fact, the generalization of the preceding DyBO formulation for SPDE systems is pretty straightforward.

3.3. The DyBO formulation for multiscale stochastic diffusion equations

In this section, we shall derive the DyBO formulation for the following multiscale stochastic diffusion equations

$$u_t(x,\omega) = \mathcal{L}u(x,t,\omega) \equiv \nabla\cdot(a(x,\omega)\nabla u(x,\omega)) + f(x), \quad x \in D, \omega \in \Omega, \tag{20}$$

where initial and boundary conditions will be given later. For notation simplicity, we omit the superscript ε in $a(x,\omega)$ and $u(x,t,\omega)$. Assume that the m-term KL expansion of the solution of (20) is given by,

$$u = \bar{u} + \mathbf{U}\mathbf{Y}^T = \bar{u} + \mathbf{U}\mathbf{A}^T\mathbf{H}^T, \tag{21a}$$

where $\mathbf{U} = (u_1, u_2, \cdots, u_m)$ and matrices $\mathbf{A} \in \mathbb{R}^{N_p \times m}$. In addition, we assume that

$$a(x,\omega) = \bar{a}(x) + \tilde{a}(x,\omega), \tag{22}$$

where $\bar{a}(x) = \mathbb{E}[a(x,\omega)]$ is the mean and $\tilde{a}(x,\omega)$ is the fluctuation with zero mean. For notation simplicity, we will write $a = \bar{a} + \tilde{a}$ in our derivation. By substituting the above expansion into

(20), we obtain the DyBO-gPC formulation for the multiscale stochastic diffusion equation (20) (see Appendix A for the details of the derivation)

$$\frac{\partial \bar{u}}{\partial t} = \nabla \cdot (\bar{a}\nabla \bar{u}) + \nabla \cdot (\mathbb{E}\left[\tilde{a}\nabla \mathbf{U}\mathbf{A}^T\mathbf{H}^T\right]) + f, \tag{23a}$$

$$\frac{\partial \mathbf{U}}{\partial t} = -\mathbf{U}\mathbf{D}^T + \nabla \cdot (\mathbb{E}\left[\tilde{a}\nabla \bar{u}\mathbf{H}\right])\mathbf{A} + \nabla \cdot (\bar{a}\nabla \mathbf{U}) + \nabla \cdot (\mathbb{E}\left[\tilde{a}\nabla \mathbf{U}\mathbf{A}^T\mathbf{H}^T\mathbf{H}\right])\mathbf{A}, \tag{23b}$$

$$\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} = -\mathbf{A}\mathbf{C}^T + \left\langle \nabla \cdot (\mathbb{E}\left[\mathbf{H}^T\tilde{a}\nabla \bar{u}\right]) + \nabla \cdot (\bar{a}\nabla \mathbf{A}\mathbf{U}^T) + \nabla \cdot (\mathbb{E}\left[\tilde{a}\mathbf{H}^T\mathbf{H}\mathbf{A}\nabla \mathbf{U}^T\right]),\, \mathbf{U}\right\rangle \mathbf{\Lambda}_{\mathbf{U}}^{-1}, \tag{23c}$$

where matrices $\mathbf{C}$ and $\mathbf{D}$ can be solved from the linear system (15) with G_*,

$$G_* = \mathbf{\Lambda}_{\mathbf{U}}^{-1}\left\langle \mathbf{U}^T, \mathbb{E}\left[\tilde{\mathcal{L}}u\mathbf{H}\right]\right\rangle \mathbf{A}.$$

Remark 3.2. In this paper, we assume that the fluctuation part $\tilde{a}$ is parameterized r random variables, i.e., $\tilde{a}\big(x, \xi_1(\omega), \cdots, \xi_r(\omega)\big) = \sum_{i=1}^{r} a_i(x)\xi_i(\omega)$, where a_i are spatial functions and $\{\xi_i\}_{i=1}^{r}$ are independent and identically distributed (i.i.d.) random variables with zero mean. The mean function $\bar{a}$ is chosen such that $a = \bar{a} + \tilde{a}$ is positive.

4. Model reduction for DyBO using GMsFEM

4.1. Motivations

The DyBO formulation (23) enables us to explore low-dimensional structures of the solution $u(x, t, \omega)$, which provides an efficient numerical method for solving (20). However, when the problem (20) has multiscale features in physical space, the DyBO method becomes expensive as one needs to use fine mesh to represent the multiscale features in the physical space.

To further reduce the computational cost in solving (23a)-(23c), we shall apply the GMsFEM to discretize $\bar{u}$ and $\mathbf{U}$. Note that, Eq.(23a) and each component of Eq. (23b) have the following deterministic time-dependent PDE form

$$\frac{\partial w}{\partial t} = \nabla \cdot (\bar{a}\nabla w) + \mathcal{G}, \tag{24}$$

$$w|_{t=0} = w_0. \tag{25}$$

for some functions $\mathcal{G}$. For example, in (23a) we have $w = \bar{u}$ and $\mathcal{G} = \nabla \cdot (\mathbb{E}\left[\tilde{a}\nabla \mathbf{U}\mathbf{A}^T\mathbf{H}^T\right]) + f$.

In order to discretize the equation (24) in time, we apply the implicit Euler scheme with time step Δt and obtain the difference equation for each time $t_n = n\Delta t$, $n = 1, 2, \cdots, N$ $(T = N\Delta t)$,

$$\frac{w^n - w^{n-1}}{\Delta t} = \nabla \cdot (\bar{a}\nabla w_n) + \mathcal{G}$$

where $w^n = w(t_n)$ and the above equation is equivalent to the following

$$-\nabla \cdot (\bar{a}\nabla w^n) + cw^n = \tilde{\mathcal{G}}, \tag{26}$$

where $c = 1/\Delta t$ and $\tilde{\mathcal{G}} = cw^{n-1} + \mathcal{G}$. Hence, for each fixed $t_n > 0$, we can use the GMsFEM to solve the second order elliptic PDE (26) with multiscale coefficient $\bar{a}$.

4.2. The general framework of the GMsFEM

In this subsection, we present the general framework of the GMsFEM for solving the multiscale PDE (26). First, we introduce the following notations. Let $\mathcal{T}^H$ denote a conforming partition of the spatial domain $\mathcal{D}$ with mesh size $H \gg \varepsilon$. The interior nodes of $\mathcal{T}^H$ are x_i, $i = 1, 2, \cdots, N_{in}$, where N_{in} is the number of interior nodes and the coarse elements of $\mathcal{T}^H$ are K_j, $j = 1, 2., \cdots, N_e$, where N_e is the number of the coarse elements. To build multiscale basis functions on the coarse elements, we denote D_i as the neighborhood corresponding to the coarse node x_i, which is defined as $D_i = \bigcup\{K_j \in \mathcal{T}^H : x_i \in \overline{K_j}\}$. Moreover, we define a finer partition $\mathcal{T}^h$ with mesh size $h \ll \varepsilon$, which will be used to compute local cell problems and represent the multiscale basis functions.

We define χ_i as the standard multiscale basis functions [22]. To obtain χ_i, we need to solve the following cell problem for each coarse note x_i,

$$-\nabla \cdot (\bar{a}\nabla\chi_i) = 0 \qquad \text{in } K \subset D_i, \tag{27}$$

$$\chi_i = g_i \qquad \text{on } \partial K, \tag{28}$$

where g_i is a continuous function on ∂K and is linear on each edge of ∂K. The bilinear form $\mathcal{A}(\cdot,\cdot)$ associated with (28) is defined as follows

$$\mathcal{A}(u, v) := \int_{\mathcal{D}} \Big(\bar{a}\nabla u \cdot \nabla v + cuv\Big). \tag{29}$$

The concept of the offline basis functions is important in the framework of GMsFEM. To obtain these offline basis functions, one needs to perform a space reduction by solving a local spectral problem over a suitable function space V. To define such function space V, we recall the definition of the snapshot space $V_{snap}^{D_i}$ based on the harmonic extensions. Let $J_h(D_i)$ be the set of all nodes of the fine mesh $\mathcal{T}^h$ lying on ∂D_i. For each fine-grid node $x_j \in J_h(D_i)$, we define a discrete delta function $\delta_j^h(x)$ in $J_h(D_i)$ as follows

$$\delta_j^h(x_k) = \begin{cases} 1, & k = j, \\ 0, & k \neq j, \end{cases}$$

where $x_k \in J_h(D_i)$. Next, we define the snapshot function ψ_j^{snap} in a coarse neighborhood D_i as follows: find ψ_j^{snap} such that

$$-\nabla \cdot (\bar{a}\nabla\psi_j^{D_i,snap}) = 0 \qquad \text{in } D_i, \tag{30}$$

$$\psi_j^{D_i,snap} = \delta_j^h \qquad \text{on } \partial D_i. \tag{31}$$

We denote $J_i := |J_h(D_i)|$ and for each coarse node of coarse x_i, $i = 1, \cdots, N_{in}$, define the local snapshot space $V_{snap}^{(i)}$ as follows:

$$V_{snap}^{(i)} := \text{snap}\{\psi_j^{D_i,snap} : j = 1, \cdots, J_i\}.$$

Next, we define the snapshot space V_{span} to be the direct sum of all local snapshot spaces $V_{snap}^{(i)}$:

$$V_{snap} := \bigoplus_{i=1}^{N_{in}} V_{snap}^{(i)}.$$

Since the snapshot space defined above is usually of large dimension, a reduction technique will be performed on the snapshot space V_{snap} to get a smaller space V_{off}. This reduction is achieved by

solving a local spectral problem on each coarse neighborhood D_i. The analysis in [16] motivates the following construction. This spectral problem is defined as follows: find $(\psi, \lambda) \in V_{snap}^{(i)} \times \mathbb{R}$ such that

$$\int_{D_i} \bar{a}\nabla\psi \cdot \nabla v = \lambda \int_{D_i} \hat{a}\psi v, \quad \forall v \in V_{snap}^{(i)}, \tag{32}$$

where $\hat{a} := \bar{a}\sum_{i=1}^{N_{in}} H^2|\nabla\chi_i|^2$. After solving the spectral problem (32) in a coarse neighborhood D_i, we arrange these eigenvalues $\lambda_k^{D_i}$, $k = 1, 2, \cdots, J_i$ in ascending order $\lambda_1^{D_i} \leq \lambda_2^{D_i} \leq \cdots \leq \lambda_{J_i}^{D_i}$, and denote the corresponding eigenfunctions by $\Psi_k^{D_i}$, $k = 1, 2, \cdots, J_i$. After that, we take the first l_i eigenfunctions in each coarse neighborhood. Using these eigenfunctions we can define

$$\phi_k^{D_i,\text{off}} = \sum_{j=1}^{l_i} (\Psi_k^{D_i})_j \psi_j^{D_i,\text{snap}}, \quad k = 1, 2, \cdots, l_i,$$

where $(\Psi_k^{D_i})_j$ denotes the j-th component of $\Psi_k^{D_i}$. Eventually, the offline basis functions $\psi_k^{D_i,\text{off}}$ for the coarse neighborhood D_i is defined by

$$\psi_k^{D_i,\text{off}} = \chi_i \phi_k^{D_i,\text{off}}, \quad k = 1, \cdots, l_i,$$

where χ_i is the standard multiscale basis function for the coarse neighborhood D_i. We also define the local offline space $V_{\text{off}}^{(i)}$, $i = 1, \cdots, N_{in}$ and the global multiscale basis function space V_{off} as follows:

$$V_{\text{off}}^{(i)} := \text{span}\{\psi_k^{D_i,\text{off}} : k = 1, \cdots, l_i\}, \quad i = 1, \cdots, N_{in},$$

$$V_{\text{off}} := \bigoplus_{i=1}^{N_{in}} V_{\text{off}}^{(i)}.$$

These $\psi_k^{D_i,\text{off}}$ will be used as the local offline basis functions. Once the global multiscale basis function space V_{off} is constructed, we can find the GMsFEM solution $u_{\text{off}}^n \in V_{\text{off}}$ at time $t = t_n$ by solving the following equation,

$$\mathcal{A}(u_{\text{off}}^n, v) = \left\langle c u_{\text{off}}^{n-1} + \mathcal{G},\, v \right\rangle, \quad \forall v \in V_{\text{off}}. \tag{33}$$

4.3. Online Process of GMsFEM

In order to achieve a rapid convergence in the GMsFEM, we need to build some online basis functions to enrich the global multiscale basis function space V_{off}. In this subsection, we shall provide a brief introduction of the online computation of the GMsFEM, which will be used in the DyBO method.

Let $u_{\text{off}}^n \in V_{\text{off}}$ be the numerical solution obtained in (33) at time $t = t_n$. Consider a given coarse neighborhood D_i, we define a space $V_i = H_0^1(D_i) \cap V_{snap}$ equipped with the norm $\|v\|_{V_i}^2 := \int_{D_i} \hat{a}(x)|\nabla v|^2$, where the weighted function $\hat{a}(x) = \big(H^2 \sum_{i=1}^{N_{in}} |\nabla\chi_i|^2\big)\bar{a}(x)$. We also define the following residual operator on V_i by

$$\mathcal{R}_i^n(v; u_{\text{off}}^n) := \int_{D_i} \big(c u_{\text{f}}^{n-1} + \mathcal{G}\big)v - \int_{D_i} \big(\hat{a}\nabla u_{\text{off}}^n \cdot \nabla v + c u_{\text{off}}^n v\big), \quad \forall v \in V_i, \tag{34}$$

where u_{f}^{n-1} is the fine-scale solution at time $t = t_{n-1}$. This is called the H^{-1}-residual on the coarse neighborhood D_i. The operator norm of $\mathcal{R}_i^n$, denoted by $\|\mathcal{R}_i^n\|_{V_i^*}$ gives a measure of the quantity of the residual. In this work, the online basis functions mean the basis functions that are computed during the iterative process for a given fixed time $t = t_n$, contrary to offline basis functions [12] that

are computed before the iterative process. The online basis functions are obtained based on the local residuals $\mathcal{R}_i^n$ for the multiscale solution u_{off}^n.

Suppose that one needs to add a basis function $\phi \in V_i$ on the coarse neighborhood D_i. We use the index $\tau \geq 1$ to represent the enrichment level and denote the current multiscale basis function space to be $V_{\text{off}}^{n,\tau}$ at time $t = t_n$. Let $V_{\text{off}}^{n,\tau+1} = V_{\text{off}}^{n,\tau} + \text{span}\{\phi\}$ be the new approximation space at time $t = t_n$ and $u_{\text{off}}^{n,\tau+1}$ be the corresponding new GMsFEM solution. The analysis in [11] suggests that the required online basis function $\phi \in V_i$ is the solution to the following equation:

$$\mathcal{A}(\phi, v) = \mathcal{R}_i^n(v; u_{\text{off}}^{n,\tau}) \quad \forall v \in V_i. \tag{35}$$

Hence, the new online basis function $\phi \in V_i$ is obtained by solving (35). To summarize, let $\mathcal{I} \subset \{1, 2, \cdots, N_{in}\}$ be the index set of the coarse neighborhoods where the online enhancements are required. For each $i \in \mathcal{I}$, we obtain some basis functions $\phi_i \in V_i$ by solving (35). Then, we define $V_{\text{off}}^{n,\tau+1} = V_{\text{off}}^{n,\tau} + \text{span}\{\phi_i : i \in \mathcal{I}\}$. Consequently, following the arguments in [11], we have at time $t = t_n$,

$$\|u_{\text{f}}^n - u_{\text{off}}^{n,\tau+1}\|_V^2 \leq \left(1 - \frac{\Lambda_{\min}^{(\mathcal{I})}}{C_{\text{err}}} \frac{\sum_{i\in\mathcal{I}} \|\mathcal{R}_i^n\|_{V_i^*}(\lambda_{l_i+1}^{D_i})^{-1}}{\sum_{i=1}^{N} \|\mathcal{R}_i^n\|_{V_i^*}(\lambda_{l_i+1}^{D_i})^{-1}}\right) \|u_{\text{f}}^n - u_{\text{off}}^{n,\tau}\|_V^2, \tag{36}$$

where C_{err} is a uniform constant and $\Lambda_{\min}^{(\mathcal{I})} = \min_{i\in\mathcal{I}} \lambda_{l_i+1}^{D_i}$. Inequality (36) shows that it is able to obtain a better convergence of the online adaptive GMsFEM by adding more online basis functions at each time $t = t_n$ and the rate of convergence depends on the constant C_{err} and $\Lambda_{\min}^{(\mathcal{I})}$. The details of implementation for the online adaptive GMsFEM will be presented in Section 4.4.

4.4. The complete algorithm and implementation

We shall summarize the overall computational scheme for our problem in this section. Recall that the multiscale coefficient is $a(x, \omega) = \bar{a}(x) + \tilde{a}(x, \omega)$ and the fluctuation part $\tilde{a}(x, \omega)$ is represented as $\tilde{a}(x, \omega) = \tilde{a}(x, \xi_1, \xi_2, \cdots, \xi_r) = \sum_{i=1}^r a_i \xi_i = a_i \xi_i$, where the Einstein notation is used. Therefore, we rewrite the DyBO formulation for the multiscale stochastic diffusion equations (20) as follows

$$\frac{\partial \bar{u}}{\partial t} = \nabla \cdot (\bar{a} \nabla \bar{u}) + \nabla \cdot (\mathbb{E}\left[a_i \xi_i \nabla \mathbf{U} \mathbf{A}^T \mathbf{H}^T\right]) + f, \tag{37}$$

$$\frac{\partial \mathbf{U}}{\partial t} = -\mathbf{U}\mathbf{D}^T + \nabla \cdot (\mathbb{E}\left[a_i \xi_i \nabla \bar{u} \mathbf{H}\right])\mathbf{A} + \nabla \cdot (\bar{a} \nabla \mathbf{U}) + \nabla \cdot (\mathbb{E}\left[a_i \xi_i \nabla \mathbf{U} \mathbf{A}^T \mathbf{H}^T \mathbf{H}\right])\mathbf{A}, \tag{38}$$

$$\frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} = -\mathbf{A}\mathbf{C}^T + \left\langle \nabla \cdot (\mathbb{E}\left[\mathbf{H}^T a_i \xi_i \nabla \bar{u}\right]) + \nabla \cdot (\bar{a} \nabla \mathbf{A} \mathbf{U}^T) + \nabla \cdot (\mathbb{E}\left[a_i \xi_i \mathbf{H}^T \mathbf{H} \mathbf{A} \nabla \mathbf{U}^T\right]), \mathbf{U} \right\rangle \mathbf{\Lambda}_{\mathbf{U}}^{-1}, \tag{39}$$

where the computational domain is $\mathcal{D} \times [0, T]$ with $\mathcal{D} = [0, 1]^2 \subset \mathbb{R}^2$. We assume that homogeneous boundary condition is imposed on the Eq.(20) so the solutions $\bar{u}$ and $\mathbf{U} = (u_1, \cdots, u_m)$ will vanish on $\partial\mathcal{D}$. The initial conditions for $\bar{u}$, $\mathbf{U}$ and $\mathbf{A}$ depend on the initial condition of u, which will be discussed in Section 5.

We denote the set of the offline basis functions as $\{\eta_i\}_{i=1}^{N_d}$, where $N_d = \dim(V_{\text{off}})$. We write $V_{\text{off}} = \text{span}\{\eta_i : i = 1, \cdots, N_d\}$ and the row vector $\mathcal{B} = \mathcal{B}(x) = \big(\eta_1(x), \cdots, \eta_{N_d}(x)\big)$. In the setting of GMsFEM, for each time $t > 0$, we seek the approximations for both the functions $\bar{u}$ and $\mathbf{U}$ using the multiscale basis functions and we assume that the following representations hold

$$\bar{u}(x, t) = \mathcal{B}(x)\hat{u}_0(t), \quad \hat{u}_0(t) \in \mathbb{R}^{N_d},$$

$$\mathbf{U}(x, t) = \mathcal{B}(x)\hat{U}_m(t), \quad \hat{U}_m(t) := (\hat{u}_1(t), \cdots, \hat{u}_m(t)) \in \mathbb{R}^{N_d \times m}.$$

Then, the variational form of (37) becomes

$$\mathcal{M}\frac{d\hat{u}_0}{dt} = -\mathcal{S}_0\hat{u}_0 - \mathcal{S}_i\hat{U}_m\mathbf{A}^T\mathbb{E}\left[\xi_i\mathbf{H}^T\right] + \hat{f}, \tag{40}$$

where

$$\mathcal{M} = (\langle \eta_j, \eta_k \rangle) \in \mathbb{R}^{N_d \times N_d}, \quad \mathcal{S}_0 = (\langle \bar{a}\eta_j, \eta_k \rangle) \in \mathbb{R}^{N_d \times N_d},$$
$$\mathcal{S}_i = (\langle a_i \eta_j, \eta_k \rangle) \in \mathbb{R}^{N_d \times N_d}, \quad \hat{f} = (\langle f, \eta_1 \rangle \cdots \langle f, \eta_{N_d} \rangle)^T \in \mathbb{R}^{N_d}.$$

Similarly, the variational form of (38) becomes

$$\mathcal{M}\frac{d\hat{U}_m}{dt} = -\mathcal{M}\hat{U}_m\mathbf{D}^T - \mathcal{S}_i\hat{u}_0\mathbb{E}\left[\xi_i\mathbf{H}\right]\mathbf{A} - \mathcal{S}_0\hat{U}_m - \mathcal{S}_i\hat{U}_m\mathbf{A}^T\mathbb{E}\left[\xi_i\mathbf{H}^T\mathbf{H}\right]\mathbf{A}. \tag{41}$$

To obtain the fully discretized scheme, we apply the implicit Euler method to approximate the time derivatives in (40) and (41). Combining with the variational forms, we obtain the following difference equations for each fixed time $t = t_n = n\Delta t$, $n = 1, \cdots, N$

$$\mathcal{S}_0\hat{u}_0^n + c\mathcal{M}\hat{u}_0^n = \mathcal{G}_1^{n-1}, \tag{42}$$
$$\mathcal{S}_0\hat{U}_i^n + c\mathcal{M}\hat{U}_i^n = \mathcal{G}_2^{n-1}, \quad i = 1, ..., m, \tag{43}$$

where $c = 1/\Delta t$ and the right hand sides $\mathcal{G}_1$ and $\mathcal{G}_2$ are defined as follows

$$\mathcal{G}_1^{n-1} = c\mathcal{M}\hat{u}_0^{n-1} - \mathcal{S}_i\hat{U}_m^{n-1}\mathbf{A}_{n-1}^T\mathbb{E}\left[\xi_i\mathbf{H}^T\right] + \hat{f},$$
$$\mathcal{G}_2^{n-1} = c\mathcal{M}\hat{U}_m^{n-1} - \mathcal{M}\hat{U}_m^{n-1}\mathbf{D}_{n-1}^T - \mathcal{S}_i\hat{u}_0^{n-1}\mathbb{E}\left[\xi_i\mathbf{H}\right]\mathbf{A}_{n-1} - \mathcal{S}_i\hat{U}_m^{n-1}\mathbf{A}_{n-1}^T\mathbb{E}\left[\xi_i\mathbf{H}^T\mathbf{H}\right]\mathbf{A}_{n-1},$$

where $\mathbf{A}_{n-1} = \mathbf{A}(t_{n-1})$, $\hat{U}_m^n = \hat{U}_m(t_n)$, $\hat{u}_0^n = \hat{u}_0(t_n)$, and $\mathbf{D}_n = \mathbf{D}(t_n)$. Then, we apply integration by part and representation of $\bar{u}$ and $\mathbf{U}$ in (39) to obtain the following ODE for $\mathbf{A}$

$$\frac{d\mathbf{A}}{dt} = -\mathbf{A}\mathbf{C}^T - \left(\mathbb{E}\left[\xi_i\mathbf{H}^T\right]\hat{u}_0^T\mathcal{S}_i\hat{U}_m + \mathbf{A}\hat{U}_m^T\mathcal{S}_0\hat{U}_m + \mathbb{E}\left[\xi_i\mathbf{H}^T\mathbf{H}\right]\mathbf{A}\hat{U}_m^T\mathcal{S}_i\hat{U}_m\right)\mathbf{\Lambda}_{\mathbf{U}}^{-1}. \tag{44}$$

In this work, we use implicit Euler scheme to approximate the time derivative and get

$$\mathbf{A}_n = \mathbf{A}_{n-1} - \Delta t\left(\mathbf{A}_{n-1}\mathbf{C}_{n-1}^T + \mathcal{G}_3^{n-1}\right), \tag{45}$$

where $\mathbf{C}_{n-1} = \mathbf{C}(t_{n-1})$ and

$$\mathcal{G}_3^{n-1} = \left(\mathbb{E}\left[\xi_i\mathbf{H}^T\right](\hat{u}_0^{n-1})^T\mathcal{S}_i\hat{U}_m^{n-1} + \mathbf{A}_{n-1}(\hat{U}_m^{n-1})^T\mathcal{S}_0\hat{U}_m^{n-1} + \mathbb{E}\left[\xi_i\mathbf{H}^T\mathbf{H}\right]\mathbf{A}_{n-1}(\hat{U}_m^{n-1})^T\mathcal{S}_i\hat{U}_m^{n-1}\right)\mathbf{\Lambda}_{\mathbf{U}}^{-1}.$$

To summarize, we solve the following dicrete system to obtain $\hat{u}_0^n$, $\hat{U}_m^n$, and $\mathbf{A}_n$ at each time $t = t_n$, $n = 1, \cdots, N$,

$$\mathcal{S}_0\hat{u}_0^n + c\mathcal{M}\hat{u}_0^n = \mathcal{G}_1^{n-1}, \tag{46}$$
$$\mathcal{S}_0\hat{U}_m^n + c\mathcal{M}\hat{U}_m^n = \mathcal{G}_2^{n-1}, \tag{47}$$
$$\mathbf{A}_n = \mathbf{A}_{n-1} - \Delta t\left(\mathbf{A}_{n-1}\mathbf{C}_{n-1}^T + \mathcal{G}_3^{n-1}\right), \tag{48}$$

where the matrices $\mathbf{C}_{n-1}$ and $\mathbf{D}_{n-1}$ in (46)-(48) can be computed using the system (15) and the corresponding $G_*(\bar{u}, \mathbf{U}, \mathbf{Y}) = -\mathbf{\Lambda}_{\mathbf{U}}^{-1}\left(\hat{U}_m^T\mathcal{S}_i^T\hat{u}_0\mathbb{E}\left[\xi_i\mathbf{H}\right] + \hat{U}_m^T\mathcal{S}\hat{U}_m\mathbf{A}^T + \hat{U}_m^T\mathcal{S}_i^T\hat{U}_m\mathbf{A}^T\mathbb{E}\left[\xi_i\mathbf{H}^T\mathbf{H}\right]\right)\mathbf{A}$.

To further improve the accuracy of our DyBO method, we use the online adaptive GMsFEM [11]. Assume that for every coarse neighborhood D_i, we pick l_i offline basis functions to form the global space V_{off} and denote $V_{ms}^{n,0} = V_{\text{off}}$. We use τ to represent the enrichment level of online adaptive GMsFEM and $V_{ms}^{n,\tau}$ to denote the enriched space at level τ for time $t = t_n$. For each fixed time $t = t_n$, the online enrichment is performed for the coarse regions having residuals that are θ fraction of the total residual.

The online adaptive GMsFEM is stated as follows: fix the parameter $\theta \in (0, 1)$. Given a fixed time $t = t_n$ and start with enrichment level $\tau = 0$. Then go to the step 1 below.

Step 1. Solve the desired variational equation to obtain the multiscale solution $\bar{u}_{\text{off}}^{n,\tau}$ or $\mathbf{U}_m^{n,\tau}$.

Step 2. Compute the residual $r_i^n = \|R_i^n\|_{V_i^*}$ for each coarse neighborhood D_i. Rearrange the residuals in descending order $r_1^n \geq r_2^n \geq \cdots$.

Step 3. Select the coarse neighborhoods that need to be enriched. Take the smallest integer k such that

$$\theta \sum_{i=1}^{N_{in}} (r_i^n)^2 \leq \sum_{i=1}^{k} (r_i^n)^2.$$

Step 4. Add online basis functions $\phi_1, \cdots, \phi_k$ for the selected coarse neighborhoods $D_1, \cdots, D_k$. The online basis function ϕ_i can be obtained by solving (35) over D_i.

After step 4, repeat from step 1 until the global residual $\sum_{i=1}^{N_{in}} (r_i^n)^2$ is smaller than a certain tolerance. If the online procedure is terminated, one can move to another time level t_{n+1}.

Remark 4.1. In our numerical experiments, we set the tolerance for residuals in the online process as 10^{-4}. Under this circumstance, only 2-3 times of online iterations are required at each time level. Meanwhile, the L^2-error between the multiscale solution and the fine-scale solution is nearly less than 2% when the online procedure is terminated.

5. Numerical experiments

In this section, we shall present numerical experiments to demonstrate the efficiency of our proposed method. We consider the following multiscale stochastic diffusion equation

$$\frac{\partial u^\varepsilon}{\partial t}(x,\omega) = \nabla \cdot (a^\varepsilon(x,\omega)\nabla u^\varepsilon(x,\omega)) + f(x,\omega), \quad x \in \mathcal{D},\ t \in (0,T],\ \omega \in \Omega, \tag{49}$$

where $\mathcal{D} = [0,1]^2 \subset \mathbb{R}^2$ is the unit square and $T = 1$. In the first three examples, the coarse mesh size is $H = \frac{1}{10}$ and the fine mesh size is $h = \frac{1}{100}$. The time step is $\Delta t = 10^{-3}$. In the fourth example, we take $H = \frac{1}{20}$ and $h = \frac{1}{200}$. For the step size in time, we use $\Delta t = \frac{1}{3} \times 10^{-3}$. Due to the efficiency of the DyBO formulation [8, 9], we use the fine mesh of the spatial domain $\mathcal{D}$ to solve the DyBO systems for $\bar{u}$ and $\mathbf{U}$, obtaining the fine-scale solution as a reference solution. Throughout the numerical simulations, we will use $l_i = 5$ offline basis functions to form the global multiscale space V_{off} initially.

Recall that the multiscale coefficient $a^\varepsilon(x,\omega)$ is given by $a^\varepsilon(x,\omega) = a^\varepsilon(x,\xi_1,\cdots,\xi_r) = \bar{a}^\varepsilon(x) + \sum_{i=1}^r a_i^\varepsilon(x)\xi_i(\omega)$, where $\{\xi_i(\omega)\}_{i=1}^r$ is a set of i.i.d. random variables with the uniform distribution over the interval $[-1,1]$. The initial condition of the solution is assumed to have the truncated KL expansion of the following form

$$\tilde{u}(x,t,\omega) = \bar{u}(x,t) + \sum_{i=1}^{m} u_i(x,t)Y_i(t,\omega), \tag{50}$$

where m is the number of mode in the KL expansion.

The stochastic basis $Y_i(t,\omega)$ can be expanded as $Y_i(t,\omega) = \sum_{j=1}^{N_p} H_j(\omega)A_{ji}(t), \quad i = 1,\cdots,m$, where $\{H_j(\omega)\}_{j=1}^{N_p}$ is a set of tensor products of Legendre polynomials in the interval $[-1,1]$, $N_p = \frac{(p+r)!}{p!r!} - 1$ is the total number of basis, p is the maximum degree of polynomial, and r is the number of the random variables. Let $\mathbf{A}(t) = (A_{ji}(t))_{N_p \times m}$ denote the expansion coefficients of all the stochastic basis. Then, the initial condition of the matrix $\mathbf{A}(t)|_{t=0} = \big(A_{ji}(0)\big)_{N_p \times m}$ should satisfy $\mathbb{E}[\mathbf{HA}] = 0$ and $\mathbf{A}^T(t)\mathbf{A}(t)|_{t=0} = I_{m\times m}$.

For each function to be approximated (e.g. $\bar{u}$, u_i or the variance function $\mathrm{var}(u) := \sum_{i=1}^{m} u_i^2$), we define the following error quantity at $t = t_n$ to estimate the accuracy of the proposed method

$$e_2^n = \frac{\|u_{\mathrm{f}}^n - u_{\mathrm{approx}}^n\|_{L^2(\mathcal{D})}}{\|u_{\mathrm{f}}^n\|_{L^2(\mathcal{D})}}$$

where u_{f}^n refers to the reference solution and u_{approx}^n is the approximation obtained by the proposed method. For simplicity, we call this quantity e_2^n to be L^2-error.

5.1. First example

In our first example, the coefficient $a^\varepsilon(x, \omega)$ is parameterized by one random variable ($r = 1$) and the degree of polynomial is chosen to be $p = 9$. The number of terms in the KL expansion is $m = 3$. The mean $\bar{a}$ of the multiscale coefficient a^ε is of high-contrast (see Figure 1) and $a_1 = 0.4$. In this case, the source function f is considered to be zero. Also, the initial conditions for $\bar{u}$ and u_i ($i = 1, 2, 3$) are given as follows

$$\begin{aligned}
\bar{u}(x_1, x_2, t)|_{t=0} &= 2\big(1 - \cos(2\pi x_1)\big)\big(1 - \cos(2\pi x_2)\big), \\
u_1(x_1, x_2, t)|_{t=0} &= 16\big(1 - \cos(2\pi x_1)\big)\big(1 - \cos(2\pi x_2)\big), \\
u_2(x_1, x_2, t)|_{t=0} &= 12\big(1 - \cos(4\pi x_1)\big)\big(1 - \cos(4\pi x_2)\big), \\
u_3(x_1, x_2, t)|_{t=0} &= 8\big(1 - \cos(6\pi x_1)\big)\big(1 - \cos(6\pi x_2)\big).
\end{aligned}$$

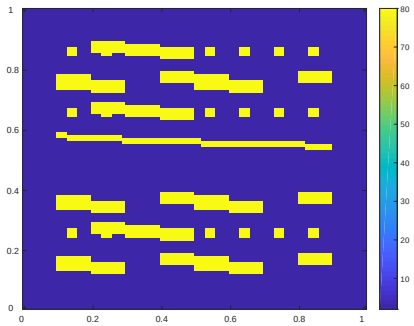

Figure 1: The coefficient $\bar{a}$. (Max: 80, Min: 1)

function	online status	$t = 0.1$	$t = 0.2$	$t = 0.4$	$t = 0.8$	$t = 1.0$
$\bar{u}$	S	10.4954%	10.4864%	10.4834%	10.4822%	10.4821%
	E	0.8865%	0.8240%	0.8234%	0.8399%	0.8398%
u_1	S	10.4822%	10.4809%	10.4815%	10.4819%	10.4820%
	E	0.8228%	0.8396%	0.8397%	0.8398%	0.8398%
u_2	S	11.7272%	11.1802%	11.0746%	11.0688%	11.0670%
	E	1.1176%	1.2291%	1.0196%	1.0178%	1.0721%
u_3	S	11.4743%	11.1082%	11.0248%	11.2790%	11.2725%
	E	1.0134%	1.0018%	1.1297%	0.9622%	0.9629%
$\mathrm{var}(u)$	S	18.3969%	18.4012%	18.4021%	18.4024%	18.4025%
	E	1.2463%	1.2451%	1.2453%	1.2454%	1.2454%

Table 1: First case: L^2-error for each functions. (S: start, E: end)

The L^2-error at several moments in the time interval $[0, T]$ are recorded and presented in Table 1. As shown in the table, the L^2-error of the functions ($\bar{u}$, u_i or var(u)) decay rapidly after performing the online procedure at each time level. For instance, in Table 1 the relative L^2-error of the variance function is nearly 18% before applying the online construction at each time level. On the other hand, at the end of the online procedure at each time level the relative error is very small, nearly 1.24%. It shows that the online adaptive construction is necessary for significantly reducing the computational error.

5.2. Second example

In this example, the mean of the multiscale coefficient $a^\varepsilon(x, \omega)$ is oscillatory, which is given by

$$\bar{a}(x_1, x_2) = 10 \times \frac{2 + P \sin(\frac{2\pi(x_1 - x_2)}{\varepsilon})}{2 - P \cos(\frac{2\pi(x_1 - x_2)}{\varepsilon})},$$

where $P = 1.7$ and $\varepsilon = 1/8$. The fluctuation part of the multiscale coefficient $a^\varepsilon(x, \omega)$ is parameterized by two independent random variables ($r = 2$) with a_1 and a_2 defined as

$$a_1(x_1, x_2) = 0.02 \times \frac{2 + P_1 \sin(\frac{2\pi(x_1 - x_2)}{\varepsilon_1})}{2 - P_1 \sin(\frac{2\pi(x_1 - x_2)}{\varepsilon_1})} \quad \text{and} \quad a_2(x_1, x_2) = 0.04 \times \frac{2 + P_2 \sin(\frac{2\pi x_1}{\varepsilon_2})}{2 - P_2 \sin(\frac{2\pi x_2}{\varepsilon_2})},$$

where $P_1 = 1.6$, $P_2 = 1.5$, $\varepsilon_1 = 1/8$ and, $\varepsilon_2 = 1/7$. The degree of polynomial is set to be $p = 9$ and let $m = 3$ in this case. The source function f is zero over the whole computational domain $\mathcal{D} \times [0, T]$. The initial conditions for the mean of the solution and the physical modes are presented below

$$\begin{aligned}
\bar{u}(x_1, x_2, t)|_{t=0} &= 16\big(1 - \cos(2\pi x_1)\big)\big(1 - \cos(2\pi x_2)\big), \\
u_1(x_1, x_2, t)|_{t=0} &= 12\big(1 - \cos(2\pi x_1)\big)\big(1 - \cos(2\pi x_2)\big), \\
u_2(x_1, x_2, t)|_{t=0} &= 8\big(1 - \cos(4\pi x_1)\big)\big(1 - \cos(4\pi x_2)\big), \\
u_3(x_1, x_2, t)|_{t=0} &= 4\big(1 - \cos(6\pi x_1)\big)\big(1 - \cos(6\pi x_2)\big).
\end{aligned}$$

function	online status	$t = 0.1$	$t = 0.2$	$t = 0.4$	$t = 0.8$	$t = 1.0$
$\bar{u}$	S	1.8428%	1.8438%	1.8456%	1.8486%	1.9224%
	E	0.4621%	0.4619%	0.4161%	0.4152%	0.4545%
u_1	S	1.8418%	1.8423%	1.8433%	1.8444%	1.8447%
	E	0.4773%	0.4616%	0.4191%	0.4165%	0.4163%
u_2	S	10.4590%	16.8297%	7.8207%	8.6577%	14.4298%
	E	0.5357%	0.6788%	0.5248%	0.4908%	0.5827%
u_3	S	5.6898%	8.9971%	9.0853%	4.1592%	14.2342%
	E	0.3386%	0.3974%	0.4520%	0.4057%	0.5167%
var(u)	S	2.8478%	2.8481%	2.8487%	2.8494%	2.5672%
	E	0.4050%	0.4065%	0.3772%	0.3805%	0.1245%

Table 2: Second case: L^2-error for each functions. (S: start, E: end)

In this example, the multiscale features are driven by both the mean part and the fluctuation part of the coefficient $a(x, \omega)$. In Table 2, the numerical results of L^2-error are reported. One can observe that the errors for both the functions are very small (nearly 0.4%) at each time t_n after the execution of the online procedure.

We remark that the choice of the number of pre-selected basis functions l_i will also affect the performance of the algorithm. One need to include at least l_i eigenfunctions that corresponding to small eigenvalues to form the finite element space at each time level. In practical, the number of small eigenvalues obtained in (32) depends on the multiscale features of the problem. One may take a larger $l_i(\geq 5)$ in the first example to obtain a better accuracy before performing the online procedure. In this case, $l_i = 5$ may be a reasonable choice.

5.3. Third example

We consider a numerical example where the coefficient is parameterized by three independent random variables ($r = 3$) and the number of terms in the KL expansion is $m = 4$. These random variables have uniform distribution in the interval $[-1, 1]$. Next, we set the coefficients a_1, a_2 and a_3 to be some functions with multiscale features spatially

$$
\begin{aligned}
a_1(x_1,x_2) &= 0.02 \times \frac{2 + P_1 \sin(\frac{2\pi(x_1-x_2)}{\varepsilon_1})}{2 - P_1 \sin(\frac{2\pi(x_1-x_2)}{\varepsilon_1})}, && P_1 = 1.6 \quad \text{and} \quad \varepsilon_1 = 1/8,\\
a_2(x_1,x_2) &= 0.04 \times \frac{2 + P_2 \sin(\frac{2\pi x_1}{\varepsilon_2})}{2 - P_2 \sin(\frac{2\pi x_2}{\varepsilon_2})}, && P_2 = 1.5 \quad \text{and} \quad \varepsilon_2 = 1/7,\\
a_3(x_1,x_2) &= 0.08 \times \frac{2 + P_3 \sin(\frac{2\pi(x_1-0.5)}{\varepsilon_3})}{2 - P_3 \sin(\frac{2\pi(x_2-0.5)}{\varepsilon_3})}, && P_3 = 1.4 \quad \text{and} \quad \varepsilon_3 = 1/6.
\end{aligned}
$$

The mean $\bar{a}$ of the multiscale coefficient is of high-contrast and it is plotted in Figure 2. The source function is chosen to be $f \equiv 1$ in the computational domain and the initial conditions for the mean of the solution and the physical modes are presented as follows

$$
\begin{aligned}
\bar{u}(x_1,x_2,t)|_{t=0} &= 32\big(1-\cos(2\pi x_1)\big)\big(1-\cos(2\pi x_2)\big),\\
u_1(x_1,x_2,t)|_{t=0} &= 24\big(1-\cos(2\pi x_1)\big)\big(1-\cos(2\pi x_2)\big),\\
u_2(x_1,x_2,t)|_{t=0} &= 16\big(1-\cos(4\pi x_1)\big)\big(1-\cos(4\pi x_2)\big),\\
u_3(x_1,x_2,t)|_{t=0} &= 8\big(1-\cos(6\pi x_1)\big)\big(1-\cos(6\pi x_2)\big),\\
u_4(x_1,x_2,t)|_{t=0} &= 4\big(1-\cos(8\pi x_1)\big)\big(1-\cos(8\pi x_2)\big).
\end{aligned}
$$

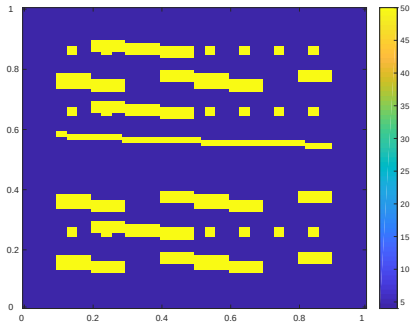

Figure 2: The coefficient $\bar{a}$. (Max: 50, Min: 4)

The convergence history is recorded in Table 3. In this example, we add one more mode function u_4 and the error of the variance function var(u) is less than 0.8% at each time level when the online procedure is terminated. Due to the linearity of the PDE and the DyBO formulation, one may easily extend this algorithm to the version with more modes in the KL expansion.

function	online status	$t = 0.1$	$t = 0.2$	$t = 0.4$	$t = 0.8$	$t = 1.0$
$\bar{u}$	S	2.9209%	2.9206%	2.9208%	2.9206%	2.9206%
	E	0.4733%	0.4730%	0.4723%	0.4730%	0.4729%
u_1	S	2.7086%	3.7209%	2.8470%	3.7136%	3.2688%
	E	0.4189%	0.4851%	0.4724%	0.4834%	0.4464%
u_2	S	3.9159%	9.2374%	4.8202%	15.6289%	6.4858%
	E	0.4884%	0.4440%	0.3512%	0.4461%	0.3756%
u_3	S	5.1971%	14.1185%	10.6377%	22.7078%	7.1893%
	E	0.6254%	0.4379%	0.4065%	0.6773%	0.4676%
u_4	S	8.5717%	13.7116%	13.0090%	10.2476%	19.6901%
	E	0.4121%	0.6560%	0.4602%	0.4305%	0.6090%
var(u)	S	4.8924%	5.5994%	4.8215%	5.5589%	5.1409%
	E	0.5711%	0.7005%	0.7088%	0.7058%	0.6608%

Table 3: Third case: L^2-error for each functions. (S: start, E: end)

5.4. Fourth example

In the last example, we assume that the coefficient is parameterized by three independent random variables in the fluctuation and let $m = 3$. We set the mean of the multiscale coefficient to be $\bar{a} = 5$ and the coefficients a_i $(i = 1, \cdots, r)$ to be more high-oscillated and given as follows

$$a_1(x_1, x_2) = 0.02 \times \frac{2 + P_1 \sin(\frac{2\pi(x_1 - x_2)}{\varepsilon_1})}{2 - P_1 \sin(\frac{2\pi(x_1 - x_2)}{\varepsilon_1})}, \qquad P_1 = 1.6 \quad \text{and} \quad \varepsilon_1 = 1/16,$$

$$a_2(x_1, x_2) = 0.04 \times \frac{2 + P_2 \sin(\frac{2\pi x_1}{\varepsilon_2})}{2 - P_2 \sin(\frac{2\pi x_2}{\varepsilon_2})}, \qquad P_2 = 1.5 \quad \text{and} \quad \varepsilon_2 = 1/15,$$

$$a_3(x_1, x_2) = 0.08 \times \frac{2 + P_3 \sin(\frac{2\pi(x_1 - 0.5)}{\varepsilon_3})}{2 - P_3 \sin(\frac{2\pi(x_2 - 0.5)}{\varepsilon_3})}, \qquad P_3 = 1.4 \quad \text{and} \quad \varepsilon_3 = 1/14.$$

Moreover, the source function and the initial conditions in this case are given as follows

$$\begin{aligned}
f(x_1, x_2) &= 1 + 0.3\big(1 - \cos(2\pi x_1)\big)\big(1 - \cos(2\pi x_2)\big), \\
\bar{u}(x_1, x_2, t)|_{t=0} &= 12\big(1 - \cos(2\pi x_1)\big)\big(1 - \cos(2\pi x_2)\big), \\
u_1(x_1, x_2, t)|_{t=0} &= 12\big(1 - \cos(2\pi x_1)\big)\big(1 - \cos(2\pi x_2)\big), \\
u_2(x_1, x_2, t)|_{t=0} &= 8\big(\sin(4\pi x_1)\big)\big(1 - \cos(4\pi x_2)\big), \\
u_3(x_1, x_2, t)|_{t=0} &= 4\big(1 - \cos(6\pi x_1)\big)\big(\sin(6\pi x_2)\big).
\end{aligned}$$

In this example, the randomness is driven by the fluctuation of the coefficient $a(x, \omega)$ with small magnitude perturbation from the random variables ξ_i. There is an observable efficiency for this type of highly-oscillated multiscale problem with random input in the sense that the numerical error between the multiscale solution and the fine-scale solution is small (less than 2%) at each time $t_n \in [0, T]$ when the online procedure is terminated.

We point out that one can obtain a better accuracy if one set the tolerance in the online procedure to be smaller (say 10^{-5}). However, it is more time-consuming as it requires more number of online refinement at each time level t_n. In general, the setting of the tolerance in the algorithm depends by the demands of the accuracy to the application.

function	online status	$t = 0.1$	$t = 0.2$	$t = 0.4$	$t = 0.8$	$t = 1.0$
$\bar{u}$	S	2.2021%	2.2030%	2.2057%	2.2057%	2.2057%
	E	1.1144%	1.1125%	1.0764%	1.0765%	1.0766%
u_1	S	2.2019%	2.2022%	2.3130%	2.3178%	2.3253%
	E	1.1132%	1.1134%	1.1200%	1.1307%	1.1818%
u_2	S	4.5920%	4.7298%	7.8030%	3.5810%	4.9615%
	E	0.9437%	0.8238%	0.9198%	0.9085%	0.7774%
u_3	S	10.5251%	10.0601%	12.6847%	3.5763%	5.4680%
	E	0.9547%	1.1308%	1.1413%	0.8894%	0.8114%
$\mathrm{var}(u)$	S	4.2831%	4.2831%	4.3376%	4.3376%	4.3429%
	E	1.5765%	1.5765%	1.4760%	1.4704%	1.5502%

Table 4: Third case: L^2-error for each functions. (S: start, E: end)

6. Conclusion

In this paper, we proposed a new framework, which combines the DyBO formulation and the online adaptive GMsFEM, to solve a time-dependent PDE with multiscale and random features. For a given multiscale PDE with random input, one can derive its corresponding DyBO formulation with the assumption that the solution has a low-dimensional structure in the sense of Karhunen-Loève expansion. The DyBO method enables us to faithfully track the KL expansion of the SPDE solution. However, the DyBO method becomes expensive when the problem contains multiscale features in the physical space.

We propose to apply the GMsFEM with online basis enrichment to improve the performance of the DyBO method. For the mean of the solution and physical modes of the solution in the truncated KL expansion, they are deterministic and dependent on time. We applied GMsFEM with implicit Euler scheme to solve the deterministic PDEs. At each time level, the online construction was applied in order to reduce the error rapidly. For the stochastic modes of the solution in the truncated KL expansion, we project them onto a set of polynomial chaos basis and obtain a ODE system, which can be solved using existing ODE solvers. Thanks to the approximation property of the multiscale basis functions obtained using the GMsFEM, the degree of freedom of our new method is relatively small compared with the original DyBO method. Therefore, our new method provides considerable computational savings over the original DyBO method.

The numerical results indicate that our new method is efficient for the time-dependent PDE problem when the multiscale coefficient consist of some independent random variables with small magnitude of perturbation. The numerical error remains to be less than 2% during the propagation in the time interval when the online construction for basis functions is executed.

Acknowledgements

The research of E. Chung is supported by Hong Kong RGC General Research Fund (Project 14304217). The research of Z. Zhang is supported by the Hong Kong RGC grants (27300616, 17300817), National Natural Science Foundation of China (Project 11601457), and Seed Funding Programme for Basic Research (HKU). S. Pun would like to thank the Issac Newton Institute for Mathematical Sciences for support and hospitality during the programme *Uncertainty quantification for complex systems: theory and methodologies* when work on this paper was undertaken. This work was supported by EPSRC Grant Numbers EP/K032208/1 and EP/R014604/1.

Appendix A. Derivations of the DyBO Formulation for the multiscale SPDE

In this appendix, we provide the details of the derivations of the DyBO-gPC formulation of multiscale SPDE (20). Substituting the KL expansion of u into Eq. (20), we get

$$\begin{aligned}\mathcal{L}u &= \nabla\cdot((\bar{a}+\tilde{a})(\nabla\bar{u}+\nabla\mathbf{U}\mathbf{A}^T\mathbf{H}^T))+f\\ &= \nabla\cdot(\bar{a}\nabla\bar{u})+\nabla\cdot(\tilde{a}\nabla\bar{u})+\nabla\cdot(\bar{a}\nabla\mathbf{U}\mathbf{A}^T\mathbf{H}^T)+\nabla\cdot(\tilde{a}\nabla\mathbf{U}\mathbf{A}^T\mathbf{H}^T)+f.\end{aligned}$$

Taking expectations on both sides yields

$$\mathbb{E}\left[\mathcal{L}u\right]=\nabla\cdot(\bar{a}\nabla\bar{u})+\nabla\cdot(\mathbb{E}\left[\tilde{a}\nabla\mathbf{U}\mathbf{A}^T\mathbf{H}^T\right])+f,$$

where we have used the fact that $\mathbb{E}\left[\tilde{a}\right]=0$ and $\mathbb{E}\left[\mathbf{H}\right]=\mathbf{0}$. Then, we obtain

$$\begin{aligned}\tilde{\mathcal{L}}u &= \mathcal{L}u-\mathbb{E}\left[\mathcal{L}u\right]\\ &= \nabla\cdot(\tilde{a}\nabla\bar{u})+\nabla\cdot(\bar{a}\nabla\mathbf{U}\mathbf{A}^T\mathbf{H}^T)+\nabla\cdot(\tilde{a}\nabla\mathbf{U}\mathbf{A}^T\mathbf{H}^T)-\nabla\cdot(\mathbb{E}\left[\tilde{a}\nabla\mathbf{U}\mathbf{A}^T\mathbf{H}^T\right])\end{aligned}$$

In addition, we compute some related terms as follows

$$\mathbb{E}\left[\tilde{\mathcal{L}}u\mathbf{H}\right]=\nabla\cdot(\mathbb{E}\left[\tilde{a}\nabla\bar{u}\mathbf{H}\right])+\nabla\cdot(\bar{a}\nabla\mathbf{U}\mathbf{A}^T)+\nabla\cdot(\mathbb{E}\left[\tilde{a}\nabla\mathbf{U}\mathbf{A}^T\mathbf{H}^T\mathbf{H}\right])$$

and

$$\left\langle\mathbf{U}^T,\,\mathbb{E}\left[\tilde{\mathcal{L}}u\mathbf{H}\right]\right\rangle_{m\times N_p}=\left\langle\mathbf{U}^T,\,\nabla\cdot(\mathbb{E}\left[\tilde{a}\nabla\bar{u}\mathbf{H}\right])+\nabla\cdot(\bar{a}\nabla\mathbf{U}\mathbf{A}^T)+\nabla\cdot(\mathbb{E}\left[\tilde{a}\nabla\mathbf{U}\mathbf{A}^T\mathbf{H}^T\mathbf{H}\right])\right\rangle$$

From Eq. (18), we have the DyBO-gPC formulation for the multiscale SPDE (20),

$$\begin{aligned}\frac{\partial\bar{u}}{\partial t} &= \nabla\cdot(\bar{a}\nabla\bar{u})+\nabla\cdot(\mathbb{E}\left[\tilde{a}\nabla\mathbf{U}\mathbf{A}^T\mathbf{H}^T\right])+f,\\ \frac{\partial\mathbf{U}}{\partial t} &= -\mathbf{U}\mathbf{D}^T+\nabla\cdot(\mathbb{E}\left[\tilde{a}\nabla\bar{u}\mathbf{H}\right])\mathbf{A}+\nabla\cdot(\bar{a}\nabla\mathbf{U})+\nabla\cdot(\mathbb{E}\left[\tilde{a}\nabla\mathbf{U}\mathbf{A}^T\mathbf{H}^T\mathbf{H}\right])\mathbf{A},\\ \frac{\mathrm{d}\mathbf{A}}{\mathrm{d}t} &= -\mathbf{A}\mathbf{C}^T+\left\langle\nabla\cdot(\mathbb{E}\left[\mathbf{H}^T\tilde{a}\nabla\bar{u}\right])+\nabla\cdot(\bar{a}\nabla\mathbf{A}\mathbf{U}^T)+\nabla\cdot(\mathbb{E}\left[\tilde{a}\mathbf{H}^T\mathbf{H}\mathbf{A}\nabla\mathbf{U}^T\right]),\,\mathbf{U}\right\rangle\mathbf{\Lambda}_{\mathbf{U}}^{-1},\end{aligned}$$

where matrices $\mathbf{C}$ and $\mathbf{D}$ can be solved from the linear system (15) with G_*,

$$G_*=\mathbf{\Lambda}_{\mathbf{U}}^{-1}\left\langle\mathbf{U}^T,\,\mathbb{E}\left[\tilde{\mathcal{L}}u\mathbf{H}\right]\right\rangle\mathbf{A}.$$

and we have used that $\mathbf{A}^T\mathbf{A}=I_{m\times m}$

Reference

[1] Assyr Abdulle, E Weinan, Björn Engquist, and Eric Vanden-Eijnden. The heterogeneous multiscale method. *Acta Numerica*, 21:1–87, 2012.

[2] 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.

[3] Alain Bourgeat. Homogenized behavior of two-phase flows in naturally fractured reservoirs with uniform fractures distribution. *Computer Methods in Applied Mechanics and Engineering*, 47(1):205–216, 1984.

[4] Robert H Cameron and William T Martin. The orthogonal development of non-linear functionals in series of fourier-hermite functionals. *Annals of Mathematics*, pages 385–392, 1947.

[5] M. Cheng, T. Y. Hou, and Z. Zhang. A dynamically bi-orthogonal method for stochastic partial differential equations I: derivation and algorithms. *J. Comput. Phys.*, 242:843–868, 2013.

[6] M. Cheng, T. Y. Hou, and Z. Zhang. A dynamically bi-orthogonal method for stochastic partial differential equations II: adaptivity and generalizations. *J. Comput. Phys.*, 242:753–776, 2013.

[7] Mulin Cheng, Thomas Y Hou, Mike Yan, and Zhiwen Zhang. A data-driven stochastic method for elliptic pdes with random coefficients. *SIAM/ASA Journal on Uncertainty Quantification*, 1(1):452–493, 2013.

[8] Mulin Cheng, Thomas Y Hou, and Zhiwen Zhang. A dynamically bi-orthogonal method for time-dependent stochastic partial differential equations i: Derivation and algorithms. *Journal of Computational Physics*, 242:843–868, 2013.

[9] Mulin Cheng, Thomas Y Hou, and Zhiwen Zhang. A dynamically bi-orthogonal method for time-dependent stochastic partial differential equations ii: Adaptivity and generalizations. *Journal of Computational Physics*, 242:753–776, 2013.

[10] Eric Chung, Yalchin Efendiev, and Thomas Y Hou. Adaptive multiscale model reduction with generalized multiscale finite element methods. *Journal of Computational Physics*, 320:69–95, 2016.

[11] Eric T Chung, Yalchin Efendiev, and Wing Tat Leung. Residual-driven online generalized multiscale finite element methods. *Journal of Computational Physics*, 302:176–190, 2015.

[12] Eric T Chung, Yalchin Efendiev, and Guanglian Li. An adaptive gmsfem for high-contrast flow problems. *Journal of Computational Physics*, 273:54–76, 2014.

[13] Manuel E Cruz and Anthony T Patera. A parallel monte-carlo finite-element procedure for the analysis of multicomponent random media. *International journal for numerical methods in engineering*, 38(7):1087–1121, 1995.

[14] BB Dukaar and PETER K Kitanidis. Determination of the effective hydraulic conductivity for heterogeneous porous media using a numerical spectral approach. *Water Resour. Res*, 28(4):1155–1178, 1992.

[15] Yalchin Efendiev, Juan Galvis, and Thomas Y Hou. Generalized multiscale finite element methods (gmsfem). *Journal of Computational Physics*, 251:116–135, 2013.

[16] Yalchin Efendiev, Juan Galvis, and Xiao-Hui Wu. Multiscale finite element methods for high-contrast problems using local spectral basis functions. *Journal of Computational Physics*, 230(4):937–955, 2011.

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

[18] Roger G Ghanem and Pol D Spanos. *Stochastic finite elements: a spectral approach.* Courier Corporation, 2003.

[19] 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.

[20] 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.

[21] Thomas Y Hou, Wuan Luo, Boris Rozovskii, and Hao-Min Zhou. Wiener chaos expansions and numerical solutions of randomly forced equations of fluid mechanics. *Journal of Computational Physics*, 216(2):687–706, 2006.

[22] Thomas Y Hou and Xiao-Hui Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *Journal of computational physics*, 134(1):169–189, 1997.

[23] Ruben Juanes. A variational multiscale finite element method for multiphase flow in porous media. *Finite elements in analysis and design*, 41(7):763–777, 2005.

[24] Kari Karhunen. *U about linear methods in probability theory*, volume 37. University of Helsinki, 1947.

[25] Michel Loeve. Probability theory, vol. ii. *Graduate texts in mathematics*, 46:0–387, 1978.

[26] X. Ma and N. Zabaras. An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations. *J. Comput. Phys.*, 228:3084–3113, 2009.

[27] O. Maıtre, O. M. Knio, H. Najm, and R. G. Ghanem. A stochastic projection method for fluid flow: I. basic formulation. *J. Comput. Phys.*, 173:481–511, 2001.

[28] Axel Målqvist and Daniel Peterseim. Localization of elliptic multiscale problems. *Mathematics of Computation*, 83(290):2583–2603, 2014.

[29] 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.

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

[31] 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.

[32] James Nolen, George Papanicolaou, and Olivier Pironneau. A framework for adaptive multiscale methods for elliptic problems. *Multiscale Modeling & Simulation*, 7(1):171–196, 2008.

[33] Bernt Øksendal. Stochastic differential equations. In *Stochastic differential equations*, pages 65–84. Springer, 2003.

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

[35] Christoph Schwab and Radu Alexandru Todor. Karhunen–loève approximation of random fields by generalized fast multipole methods. *Journal of Computational Physics*, 217(1):100–122, 2006.

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

[37] Xiaoliang Wan and George Em Karniadakis. An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. *Journal of Computational Physics*, 209(2):617–642, 2005.

[38] E Weinan, Bjorn Engquist, et al. The heterognous multiscale methods. *Communications in Mathematical Sciences*, 1(1):87–132, 2003.

[39] D. Xiu. Fast numerical methods for stochastic computations: a review. *Commun. Comput. Phys.*, 5:242–272, 2009.

[40] Dongbin Xiu and George Em Karniadakis. The wiener–askey polynomial chaos for stochastic differential equations. *SIAM journal on scientific computing*, 24(2):619–644, 2002.