

A Data-Driven Stochastic Method

Mulin Cheng, Thomas Y. Hou, Pengchong Yan
*Applied and Computational Mathematics,
California Institute of Technology, Pasadena, CA 91106*

Abstract

Stochastic partial differential equations (SPDE) arise in many physical and engineering applications. Solving these SPDE efficiently and accurately presents a great challenge. The Wiener Chaos Expansion method (WCE) and its variants have shown some promising features but still suffer from the slow convergence of polynomial expansion and the curse of dimensionality. The Monte Carlo method (MC) has an attractive property that its convergence is independent of the number of stochastic dimensions. However, its convergence rate is only of order $O(1/\sqrt{N})$ (N is the number of realizations). Inspired by Multi-scale Finite Element methods (MsFEM) and Proper Orthogonal Decomposition methods, we propose a data-driven stochastic method, which consists of offline step and online step and uses both MC and WCE. A random basis $\{A_i(\omega)\}_{i=1}^M$ is computed in the offline step based on the Karhunen-Loeve(KL) expansion, which can be viewed as partially inversion of the random differential operator. In the online step, the polynomial basis used in Polynomial Chaos expansion is replaced by the problem-dependent stochastic basis that we constructed in the offline step. By solving a set of coupled deterministic PDEs of coefficients, we obtain the numerical solutions of SPDEs. The proposed framework is applied to 1D and 2D elliptic problems with random coefficients and a horn problem. Numerical examples show that 1. The stochastic basis obtained in the offline computation can be repeatedly used in the online computation for different deterministic right hand sides. 2. A much smaller number of bases are necessary in the online computation to achieve certain error tolerance compared to WCE.

1 Introduction

Uncertainty arises in many complex real-world problems of physical and engineering interests, such as wave, heat and pollution propagation through random media, randomly forced Burgers or Navier-Stokes equations. Additional examples can be found extensively in other branches of science and engineering, such as fluid dynamics, geophysics, quantum mechanics, statistical physics, biology and etc. Stochastic Partial differential equation(SPDE) is an effective mathematical tool to model these complicated problems, which manifests randomness in many aspects, such as initial/boundary conditions, model parameters, measurements, to name a few. Solving these SPDEs poses a great challenge to the scientific computing community.

In many applications, one is mostly interested in statistical information of the stochastic solution. For this purpose, people have derived moment equations and probability density function (PDF) methods in which effective equations such as Fokker-Planck equations, have been derived for the moments and the PDF. For nonlinear problems, such methods suffer from the closure problem. Because of stability and non-intrusive nature of the Monte Carlo method(MC), it remains the most popular numerical method for SPDEs. Due to the law of large number, the convergence rate of MC is merely $\frac{1}{\sqrt{N}}$ [?], where N is the number of realizations. There are various ways to accelerate the convergence rate of MC by using variation reduction techniques, such as anti-thetic variables, conditioning, control variates, and importance sampling. For SPDEs, one can also use collocation and sparse grid methods to reduce the complexity of the method, but none of these techniques can change the nature of slow convergence of MC.

While MC method represents the solution of SPDE as an ensemble of realizations, another way to represent the solution is to expand it in an orthonormal basis, or polynomial chaos which was originally proposed by N. Wiener in 1938[?]. The method did not receive much attention until Cameron and Martin developed an explicit

and intuitive formulation for Wiener Chaos Expansion(WCE) in their elegant work in 1947. Compared to MC methods, dramatic improvements can be achieved in some problems, but WCE method has two main disadvantages

1. The number of terms in truncated WCE grows exponentially fast as the order of Hermite polynomials and the number of random variables increase, especially for long time integration.
2. The solution may depend sensitively on random dimensions, such as parameter shocks or bifurcation, which further demands more terms in WCE.

To reduce dimensionality and achieve sparse representation, several generalizations of the classical WCE method have been proposed to circumvent the drawbacks, such as generalized polynomial chaos via Wiener-Askey polynomial [?], Wiener-Haar Chaos Expansion [?], Multi-element gPC [?, ?, ?] and etc. Improvements are observed in some problems, but not all, especially for time-dependent problems. It is important to note that the basis used in the polynomial chaos expansion is determined *a priori* and problem independent, which we believe is the ultimate reason of giving rise to a non-sparse representation of the stochastic solution when the solution is not smooth in the stochastic space or when the stochastic dimension is high.

A typical scenario occurring frequently in practice is that we are interested in investigating what will happen when deterministic parts of the problem change and the random parts remain unchanged. For example, the properties of certain kind of composite material satisfy some stochastic processes with known parameters and we want to know its responses under different external forces. Mathematically, consider the SPDE in the general form of

$$\mathcal{L}(\omega)u(x, \omega) = f(x)$$

where $x \in D$, $\omega \in \Omega$ and $\mathcal{L}(\omega)$ is a stochastic differential operator. Clearly, $\mathcal{L}(\omega)$ represents the random parts of the problem while $f(x)$ is the deterministic part of the problem. Our framework can be extended to time-dependent SPDE which we should address in a separate paper. The method we proposed consists two parts:

- **Offline:** The stochastic basis $\{A_i(\omega)\}$ is constructed based on the Karhunen-Loeve(KL) expansion of MCFEM solutions $u(x, \omega)$ for a specific right-hand side $f(x) = f_0(x)$.
- **Online:** For the right-hand side $f(x)$ of interests, the solution is projected on the constructed basis,

$$u(x, \omega) \approx \sum_{i=0}^M u_i(x)A_i(\omega)$$

where $A_0 = 1$. By Galerkin projection, a system of coupled deterministic PDEs of $u_i(x, t)$ can be obtained and solved numerically. The mean, variance and other statistical quantities can be recovered from the solution of $u_i(x, t)$.

This paper is organized as follows. In the next two section, Offline and online algorithms will be described and analyzed in details along with Multi-Level Monte Carlo Method to accelerate offline computation. In sec.5, we applied our method to elliptic partial differential equations with random elliptic coefficients and compared the results with those by PC method to show the advantages. Some cases of physical and engineering interests, such as large variance or jumps at random locations, are also included and analyzed numerically in details. To demonstrate the potential application of our method in industrial/realistic scenario, a horn problem is considered. Some concluding remarks are given in Section 6.

2 Theory

2.1 Stochastic Partial Differential Equation

Consider the stochastic PDE, linear or non-linear, in the general form of

$$\begin{aligned}\mathcal{L}(\mathbf{x}, \omega)u(\mathbf{x}, \omega) &= f_0(\mathbf{x}), \quad \mathbf{x} \in D, \quad \omega \in \Omega \\ u(\mathbf{x}, \omega) &= 0, \quad \mathbf{x} \in \partial D,\end{aligned}\tag{2.1}$$

or

$$\begin{aligned}\mathcal{L}(\mathbf{x}, \omega)u(\mathbf{x}, \omega) &= 0, \quad \mathbf{x} \in D, \quad \omega \in \Omega \\ u(\mathbf{x}, \omega) &= f_0(\mathbf{x}), \quad \mathbf{x} \in \Gamma \subseteq \partial D,\end{aligned}\tag{2.2}$$

where D is a d -dimensional physical space. Here the deterministic function $f_0(x)$ exists either at the right-hand-side of the PDE, or as part of the boundary condition.

The Monte-Carlo solution of the elliptic equation, given one sample ω , $u(\mathbf{x}, \omega)$ is solved. The Karhunen-Loève decomposition of the PDE solution gives the dominant components in the random spaces. We denote the KL decomposition as

$$u(\mathbf{x}, \omega; f_0) = \bar{u}(\mathbf{x}; f_0) + \sum_{i=1}^{\infty} \sqrt{\lambda_i} A_i(\omega) \phi_i(\mathbf{x}; f_0),\tag{2.3}$$

where $\bar{u}(\mathbf{x}; f_0) = \mathbb{E}u(\mathbf{x}, \omega; f_0)$. In the current paper, we intend to derive the empirical basis $\{A_i(\omega)\}_{i=1, \dots, M}$ based on a test function $f_0(x)$, where M is a small truncation number. Notice that the basis $\{A_i(\omega)\}_{i=1, \dots, M}$ in the KL decomposition are orthonormal basis. We don't need to store all information of $A_i(\omega)$, but some moments, such as $\mathbb{E}[a(\mathbf{x}, \omega)A_i(\omega)A_j(\omega)]$. Those computations can be done off-line. Then for a different function $f_1(\mathbf{x})$, we are looking for a Galerkin projection of the solution, i.e.,

$$u(\mathbf{x}, \omega; f_1) = \bar{u}(\mathbf{x}; f_1) + \sum_{i=1}^M A_i(\omega) u_i(\mathbf{x}; f_1).\tag{2.4}$$

For easy notation, we denote $A_0(\omega) = 1$ and $u_0(\mathbf{x}; f_1) = \bar{u}(\mathbf{x}; f_1)$.

It is important to point that if the correlation length of the solution is small, then the number of basis M may be large due to the large dimension of the stochastic solution space. Then our method is not an optimal choice to solve the problem, although it still has its big advantages over other method, such as WCE. In this paper, we will exclude this situation and consider only the case with large correlation length of the solution.

In addition to that, there exists some puzzles here we need to solve here. As for the Wiener Chaos expansion, the dimension of the stochastic domain is infinite, while we only choose a PDE-driven stochastic domain, which is very small (if the variance of the solution is not large). In order to validate the Galerkin method, we will

1. verify the dominant solution space is finite dimensional;
2. define a class of recoverable function $f_1(x)$ on the basis generated off-line by $f_0(x)$.

2.2 Solution space of 1D Elliptic equation with random coefficients

Let's first consider the stochastic Elliptic equation of the form

$$\begin{aligned}-\operatorname{div}(a(\mathbf{x}, \omega)\nabla u) &= f_0(\mathbf{x}), \quad \mathbf{x} \in D, \quad \omega \in \Omega \\ u(\mathbf{x}, \omega) &= 0, \quad \mathbf{x} \in \partial D\end{aligned}\tag{2.5}$$

where $D = [0, 1]^d$, d is the space dimension, and $a(\mathbf{x}, \omega)$ is a random variable satisfying the ellipticity, $0 < a_0 \leq a(x, \xi) \leq a_1 < \infty$. In particular, we consider the simple 1-d case with homogeneous Dirichlet boundary condition:

$$\frac{\partial}{\partial x}(a(x, \omega) \frac{\partial}{\partial x} u(x, \omega)) = -f(x).$$

Then we derive

$$a(x, \omega) \frac{\partial}{\partial x} u(x, \omega) = - \int_0^x f(s_1) ds_1 + C(\omega),$$

where $C(\omega)$ is determined by the boundary conditions. In addition, we denote $b(x, \omega) = (a(x, \omega))^{-1}$. Then the solution is given by

$$u(x, \omega) = - \int_0^x b(s_2, \omega) \left(\int_0^{s_2} f(s_1) ds_1 - C(\omega) \right) ds_2.$$

Because of the boundary condition $u(1, \omega) = 0$, we have

$$C(\omega) = \frac{\int_0^1 b(s_1, \omega) \int_0^{s_1} f(s_2) ds_2 ds_1}{\int_0^1 b(s, \omega) ds}. \quad (2.6)$$

Also we can write the Karhunen-Loève decomposition of $b(x, \omega)$ as

$$b(x, \omega) = \sum_{i=0}^M b_i(x) B_i(\omega),$$

in which we only keep the first M leading terms. Then the solution is

$$\begin{aligned} u(x, \omega) &= - \int_0^x \sum_i b_i(s_2) B_i(\omega) \int_0^{s_2} f(s_1) ds_1 \\ &+ \left(\int_0^x \sum_i b_i(s) B_i(\omega) ds \right) \frac{\int_0^1 \sum_j b_j(s_1) B_j(\omega) \int_0^{s_1} f(s_2) ds_2 ds_1}{\int_0^1 b(s, \omega) ds}. \end{aligned}$$

Apparently, the stochastic part of first term lies in the space spanned by $\{B_i(\omega)\}_{i=1, \dots, M}$, while that of the second one lies in the space spanned by

$$\left\{ \frac{B_i(\omega) B_j(\omega)}{\int_0^1 b(s, \omega) ds} \right\}_{i,j} \quad \text{with } i, j = 0, 1, \dots, M.$$

Again we denote $B_0 = 1$. Then the solution of the 1-d Stochastic elliptic equation lies in the same stochastic space, which is independent of the right-hand-side function $f(x)$.

2.3 Solution space of 2D and higher dimensional elliptic equation with random coefficients

In high-dimension case, we consider the Green's function $G(x, y; \omega)$, which satisfies

$$\begin{aligned} -\operatorname{div}(a(\mathbf{x}, \omega) \nabla G(\mathbf{x}, \mathbf{y}; \omega)) &= \delta(\mathbf{x} - \mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in D, \quad \omega \in \Omega \\ G(\mathbf{x}, \mathbf{y}; \omega) &= 0, \quad \mathbf{x} \in \partial D, \quad \mathbf{y} \in D. \end{aligned} \quad (2.7)$$

Again we can write out the Karhunen-Loève decomposition of $G(\mathbf{x}, \mathbf{y}; \omega)$

$$G(\mathbf{x}, \mathbf{y}; \omega) = \sum_{i=0}^M G_i(\mathbf{x}, \mathbf{y}) \xi_i(\omega),$$

in which we keep the leading M terms. Then solution of the elliptic equation is

$$\begin{aligned} u(\mathbf{x}, \omega) &= \int_D G(\mathbf{x}, \mathbf{y}; \omega) f(\mathbf{y}) d\mathbf{y} \\ &= \sum_{i=0}^M \left(\int_D G_i(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y} \right) \xi_i(\omega). \end{aligned}$$

So again the solution lies in a finite dimensional stochastic space spanned by $\{\xi_i(\omega)\}_{i=1, \dots, M}$, regardless of the right-hand-side $f(\mathbf{x})$.

Although this part is derived for Elliptic equation, the results can be generalized to a large class of stochastic PDEs (e.g., those PDE with a Green's function, see Helmholtz equation in Sec.5.3).

3 Data-Driven Stochastic Multi-scale Method

The central task of our method is to look for a Data-Driven or Problem-Dependent random basis under which solutions enjoy a sparse expansion. The Error-Minimizing property of Karhunen-Loeve (KL) expansion makes it a natural choices. Specifically, for $u(x, \omega) \in \mathbb{L}^2(D \times \Omega)$, its Karhunen-Loeve expansion is

$$u(x, \omega) = \mathbb{E}[u] + \sum_{i=1}^{\infty} \sqrt{\lambda_i} \xi_i(\omega) \phi_i(x)$$

where λ_i and $\phi_i(x)$ are the eigenvalue and the eigenfunction of the covariance kernel $C(x, y)$, i.e.,

$$\int_D C(x, y) \phi(y) dy = \lambda \phi(x)$$

where $C(x, y) = \mathbb{E}[(u(x, \omega) - \mathbb{E}[u(x, \cdot)])(u(y, \omega) - \mathbb{E}[u(y, \cdot)])]$. Under some general conditions, the eigenvalues in KL expansion decay exponentially fast [?], which means the first few terms, or modes, of the KL expansion can give very good approximation of $u(x, \omega)$. In other words, the dimension of input random space may be large, but the dimension of the effective output random space may be much smaller as shown in fig.1.

For elliptic PDE, a heuristic argument based on Green's function shows the properties of solutions are mainly controlled by the properties of stochastic differential operator $\mathcal{L}(\omega)$. Therefore, the solutions $u(x, \omega)$ for a board class of right-hand side $f(x)$ may enjoy sparse representation under some random basis which certainly depends on the stochastic operator $\mathcal{L}(\omega)$. In other words, a random basis computed for a specific deterministic right-hand side $f(x) = f_0(x)$ may provide sparse representations for a board class of deterministic right-hand side $f(x)$. Similar philosophy has been adopted in Multi-scale Finite Element methods [?] and Proper Orthogonal Decomposition (POD) methods [?, ?, ?], so we name our method as Data-Driven Stochastic Multi-scale method.

3.1 Offline and Online Parts

If we follow this philosophy, the framework we proposed naturally breaks into two parts, the construction of a random basis $\{A_i(\omega)\}_{i=1}^M$ and the numerical solution of a coupled system of PDEs of the expansion coefficients. For a fixed stochastic differential operator $\mathcal{L}(\omega)$, the computation of the random basis may be expansive depending on the problem and the accuracy requirements. Once the random basis are constructed, they can be used repeated for different right-hand side function $f(x)$. Therefore, we call the construction of random basis the offline computation part, meaning this step is a preprocess step and can be completed in an "offline" manner contrast to the numerical solution of a coupled system of PDEs in an "online" manner. The following two algorithms reveal the details of each parts.

The goal of offline computation is to construct the random basis based on KL expansion of the random solution. The algorithm formulated below is essentially two-pass Monte Carlo method. In the first pass, the

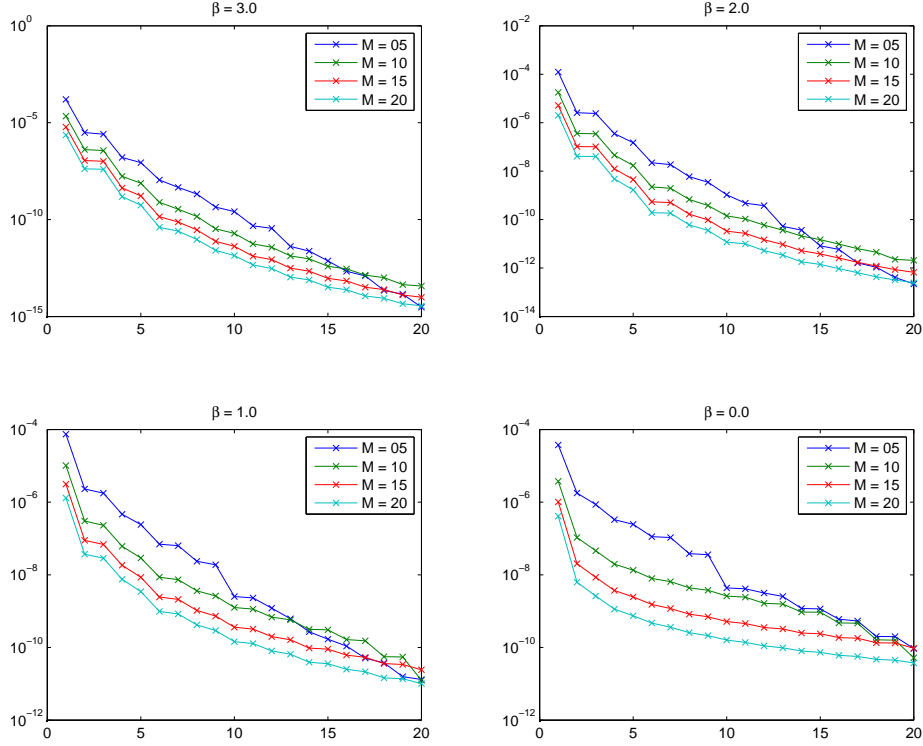


Figure 1: Eigenvalues of the covariance function of solutions for elliptic PDE with random coefficients $a(x, \omega) = \sum_{m=1}^M m^{-\beta} \xi_m (\sin(2\pi m x + \eta_i) + 1)$, where ξ_i 's, η_i 's are independent uniform random variables on $[0, 1]$. We observe that the dimensionality of input random space may be large, but effective output random space may be much smaller due the exponential convergence of eigenvalues of covariance function.

mean and the covariance function of the solution are computed. In the second pass, the random basis and their relevant statistical information are computed. Various Monte Carlo acceleration techniques can be integrated into this step and reduce the offline computational cost, which is relevant, but not the central focus of this paper.

Algorithm 1 (Offline Computation).

1. Compute the expectation $\bar{u} = \mathbb{E}[u]$ and the covariance function $R(x, x') = \mathbb{E}[(u(x, \omega) - \bar{u}(x))(u(x', \omega) - \bar{u}(x')))]$ by Monte Carlo Method with N_1 realizations, i.e.,

$$\bar{u}(x) = \frac{1}{N_1} \sum_{n=1}^{N_1} u(x, \omega_n)$$

$$R(x, x') = \frac{1}{N_1} \sum_{n=1}^{N_1} u(x, \omega_n) u(x', \omega_n) - \bar{u}(x) \bar{u}(x')$$

2. Solve the first M pairs of eigenvalues and eigenfunctions

$$\lambda_i \phi_i(x) = \int_D R(x, x') \phi_i(x') dx', \quad i = 1, \dots, M$$

3. Compute the stochastic basis $A_i(\omega)$ by

$$A_i(\omega) = \frac{1}{\sqrt{\lambda_i}} \int_D u(x, \omega) \phi_i(x) dx$$

and statistical information of $A_i(\omega)$, such as

$$\mathbb{E}[A_i A_j A_k] = \frac{1}{N_2} \sum_{n=1}^{N_2} A_i(\omega_n) A_j(\omega_n) A_k(\omega_n)$$

Remark 3.1. Guard-lines of selections for computational parameters. The decaying property of eigenvalues can be used to select parameter M , i.e., the number of random basis M can be chosen such as $\frac{\lambda_{M+1}}{\lambda_1}$ is small. In general, N_1 can be chosen to be smaller than N_2 .

Remark 3.2. In the online part, Galerkin projection reduces the stochastic PDE to a coupled system of PDEs whose coefficients involve only of statistical information of $A_i(\omega)$'s, such as $\mathbb{E}[A_i A_j A_k]$, which suggests only such information be stored in the offline part.

Remark 3.3. Our method is semi-non-intrusive in the sense that certified legacy codes can be used with minimum changes in offline part. This property makes it very attractive to industrial community as it reuses codes verified by possibly expansive and time-consuming experiments and avoid another life cycle of developing new programs.

Remark 3.4. We can apply multi-level techniques to speed-up the construction of random basis which will be detailed in the next section.

Once the random basis is constructed and stored, we are ready to apply it to cases which we really interested in. The constructed random basis spans a finite-dimensional subspace in $\mathbb{L}^2(\Omega)$ and we can project the solution $u(x, \omega)$ on this subspace, and write

$$u(x, \omega) \approx \sum_{i=0}^M u_i(x) A_i(\omega) \quad (3.1)$$

where $A_0 = 1$. We include A_0 in the random basis to make the formulation concise. The expanded random basis $\{A_i(\omega)\}_{i=0}^M$ can be verified to be orthonormal. In this paper, we mainly consider elliptic-type PDE with random coefficients. To demonstrate the main idea of online part, we consider elliptic PDE with random coefficients and homogeneous boundary conditions are assumed, i.e.,

$$\begin{aligned} -\frac{\partial}{\partial x_p} \left(a_{pq}(x, \omega) \frac{\partial u}{\partial x_q}(x, \omega) \right) &= f(x), \quad x \in D \subset \mathbb{R}^d \\ u(x, \omega) &= 0, \quad x \in \partial D \end{aligned}$$

where $p, q = 1, \dots, d$, Einstein summation is assumed and

$$\mathcal{L}(\omega) = -\frac{\partial}{\partial x_p} \left(a_{pq}(x, \omega) \frac{\partial}{\partial x_q}(x, \omega) \cdot \right)$$

Similar formulations can be obtained for other types of stochastic differential operator.

Algorithm 2 (Online Computation).

1. expand the solution under the random basis $\{A_i(\omega)\}$, substitute into the SPDE, multiply both side by $A_j(\omega)$ and take expectations, i.e., Galerkin projection, which gives a set of PDEs for $u_i(x)$.

$$\begin{aligned} -\frac{\partial}{\partial x_p} \left(\mathbb{E}[a_{pq} A_i A_j] \frac{\partial u_i}{\partial x_q} \right) &= f(x), \quad x \in D, \quad i, j = 0, 1, \dots, M \\ u_i(x) &= 0, \quad x \in \partial D \end{aligned}$$

2. Solve the coupled system of deterministic PDEs by numerical methods, such as Finite Element Method, Finite Difference Method and etc.
3. Statistical information of the solution which are of primary interests can be calculated from deterministic coefficients. For example, the centered fourth moment can be computed as

$$\mathbb{E}[(u - \bar{u})^4] = \sum_{i,j,k,l} \mathbb{E}[A_i A_j A_k A_l] u_i u_j u_k u_l$$

where $\mathbb{E}[A_i A_j A_k A_l]$ is readily available in the offline part.

Remark 3.5. Besides statistical information of solutions, sometimes other types of information, such as extreme events and etc, is desired, which requires distribution information of solutions. In such cases, instead of storing only the statistical information of the random basis $\{A_i(\omega)\}_{i=1}^M$, we may exploring their sparse representation under some predefined random basis, such as polynomial chaos basis, Haar chaos basis and etc and store the offline-constructed random basis completely. This part constitutes parts of ongoing work.

3.2 Complexity Analysis

The applicability of our method depends on the dimension of the stochastic space that the solution lies in. If the dimension is large, we have to choose more stochastic bases and hence increase the complexity of the Galerkin projection.

Proposition 3.1. [?] If we consider the Gaussian covariance kernels of the form

$$R(x, x') = \sigma^2 \exp(-|x - x'|^2 / \gamma^2),$$

then the eigenvalues λ_m has the properties

$$0 \leq \lambda_m \leq \sigma^2 \frac{(1/\gamma)^{m^{1/d}+2}}{\Gamma(0.5m^{1/d})}, \quad \forall m \geq 1,$$

where x lies in a unit d -dimensional domain, and Γ is the gamma function.

It is clear that the larger correlation length γ , the slower decaying of the eigenvalues. Consequently in our DSMM, we have to adopt more bases because of the effective dimension increasing. Hence it will worse the efficiency of our method.

4 Multi-Level Monte Carlo Acceleration

Although the offline computation may be expansive, the constructed random basis can be stored and used repeatedly in online computation for different cases. Various variance reduction techniques can be used to accelerate the offline computation, among which Multi-Level Monte Carlo method (MLMC) has been proved to be very effective for stochastic differential equation [?]. Here we will demonstrate the applicability of the MLMC to SPDE.

We denote $v(x, \omega)$ the stochastic function given by the SPDE. It should be pointed out that $v(x, \omega)$ can any of the following:

- a MC solution of SPDE, in which $v(x, \omega) \in \mathbb{L}^2(D \times \Omega)$,
- a covariance matrix of the solution, in which $v(x, \omega) \in \mathbb{L}^2(D^2 \times \Omega)$.
- a function of the random basis (e.g., $a(x, \omega)A_i(\omega)A_j(\omega)$, or $A_i(\omega)A_j(\omega)A_k(\omega)$), in which $v(x, \omega) \in \mathbb{L}^2(D \times \Omega)$ or $v(\omega) \in \mathbb{L}^2(\Omega)$.

In addition, we define Mean-Squared Error (MSE) of Y with respect to the mean solution \bar{v} as

$$\text{MSE} = \mathbb{E} [\|Y - \bar{v}\|_2^2].$$

If $v(x, \omega)$ is in the space of $\mathbb{L}^2(D^2 \times \Omega)$ or $\mathbb{L}^2(D \times \Omega)$, following the Fubini's theorem, we have $\text{MSE} = \int \mathbb{E} [(Y - \bar{v})^2] dx$. While for $v(\omega) \in \mathbb{L}^2(\Omega)$, we have $\text{MSE} = \mathbb{E} [(Y - \bar{v})^2]$. For easy notation, we consider the first case only, and the theorem of second case $v(\omega)$ is easy to derive.

We write Y as the MLMC estimate of the mean function \bar{v} and it can be decomposed as

$$Y = \sum_{k=0}^L Y_k,$$

with

$$\begin{aligned} Y_k &= \frac{1}{N_k} \sum_{i=1}^{N_k} [v_k(x, \omega_i^{(k)}) - v_{k-1}(x, \omega_i^{(k)})], \quad k = 1, \dots, L; \\ Y_0 &= \frac{1}{N_0} \sum_{i=1}^{N_0} v_0(x, \omega_i^{(0)}). \end{aligned}$$

Here $L + 1$ is the number of levels used in MLMC, and N_k is the number of MC simulations at the k -th level. Eventually MSE becomes

$$\begin{aligned} \text{MSE} &= \int \mathbb{E} \left[\left(\sum_{k=0}^L Y_k(x) - \bar{v}(x) \right)^2 \right] dx \\ &= \int \left(\mathbb{E} \left[\sum_{k=0}^L Y_k(x) \right] - \bar{v}(x) \right)^2 dx + \int \text{Var} \left(\sum_{k=0}^L Y_k(x) \right) \\ &= \int \left(\mathbb{E} \left[\sum_{k=1}^L \frac{1}{N_k} \sum_{i=1}^{N_k} [v_k(x, \omega_i^{(k)}) - v_{k-1}(x, \omega_i^{(k)})] + \frac{1}{N_0} \sum_{i=1}^{N_0} v_0(x, \omega_i^{(0)}) \right] - \bar{v}(x) \right)^2 dx \\ &+ \int \text{Var} \left(\sum_{k=1}^L \frac{1}{N_k} \sum_{i=1}^{N_k} [v_k(x, \omega_i^{(k)}) - v_{k-1}(x, \omega_i^{(k)})] + \frac{1}{N_0} \sum_{i=1}^{N_0} v_0(x, \omega_i^{(0)}) \right) \\ &= \int \left(\sum_{k=1}^L [\bar{v}_k - \bar{v}_{k-1}] + \bar{v}_0 - \bar{v}(x) \right)^2 dx + \sum_{k=1}^L \frac{1}{N_k} \int \text{Var}(v_k - v_{k-1}) dx + \frac{1}{N_0} \int \text{Var}(v_0) dx \\ &= \int (\mathbb{E}[v_L] - \bar{v})^2 dx + \sum_{k=1}^L \frac{1}{N_k} \int \text{Var}(v_k - v_{k-1}) dx + \frac{1}{N_0} \int \text{Var}(v_0) dx. \end{aligned}$$

Apparently the first term of MSE gives the bias error introduced by the numerical discretization at the finest level L . In order to compare the performance of MLMC to the single-level MC (SLMC), we assume h_L is the smallest step-size that one can solve the MC solution. Hence for SLMC simulations, we choose h_L as the step size, and N as the number of simulations to achieve ϵ^2 -MSE. Then SLMC shares the same bias error as that of MLMC. In particular, the error is $\gamma\epsilon^2/(\alpha + \gamma)$, which will be clear after we state the following theorem.

Theorem 4.1. *Let Y be the estimate of the mean \bar{v} and $h_k = M^{l-k} h_L$ be the step size of k -th level with fixed integer $M > 1$. We assume the following estimate:*

1. $\int (\mathbb{E}[u_k] - \bar{u})^2 dx = c_1 h_k^\alpha;$
2. $\int \text{Var}(u_0) dx = c_2;$

3. $\int \text{Var}(u_k - u_{k-1}) dx = c_3 h_k^\beta;$

4. the computational complexity of a k -th level MC simulation is $c_4 h_k^{-\gamma}$.

Then the multilevel estimator

$$Y = \sum_{k=0}^L Y_k$$

has the mean-square error with bound

$$\text{MSE} = \mathbb{E} [\|Y - \bar{u}\|_2^2] \leq \epsilon^2$$

with a computational cost \mathcal{C}_2

$$\mathcal{C}_2 \leq \begin{cases} c_5 \epsilon^{-2} & = \tilde{c}_5 \epsilon^{2\gamma/\alpha} \mathcal{C}_1, & \text{if } \beta > \gamma; \\ c_6 \epsilon^{-2} (\ln \epsilon)^2 & = \tilde{c}_6 \epsilon^{2\gamma/\alpha} |\ln \epsilon|^2 \mathcal{C}_1, & \text{if } \beta = \gamma; \\ c_7 \epsilon^{-2-2(\gamma-\beta)/\alpha} & = \tilde{c}_7 \epsilon^{2\beta/\alpha} \mathcal{C}_1, & \text{if } 0 \leq \beta < \gamma. \end{cases}$$

Here the finest step size h_L is chosen as the optimal step size for the Single-Level Monte Carlo simulation, i.e.,

$$h_L = \left(\frac{\gamma \epsilon^2}{c_1 (\alpha + \gamma)} \right)^{1/\alpha},$$

and the optimal computational cost \mathcal{C}_1 , subjected to ϵ^2 -MSE condition, is

$$\mathcal{C}_1 = c_4 N h_L^{-\gamma} = \frac{c_2 c_1^{\gamma/\alpha} c_4 (\alpha + \gamma)^{1+\gamma/\alpha}}{\alpha \gamma^{\gamma/\alpha}} \epsilon^{-2(1+\gamma/\alpha)}.$$

5 Numerical Examples

5.1 1D Elliptic PDE with Random Coefficients

The random basis $\{T_\alpha(\xi(\omega))\}_{\alpha \in \mathcal{J}}$ used in Polynomial Chaos expansion can be constructed analytically by tensor product of one-dimensional polynomial chaos, i.e.,

$$T_\alpha(\xi(\omega)) = \prod_{i=1}^{+\infty} H_{\alpha_i}(\xi_i(\omega))$$

where $\xi = (\xi_1, \xi_2, \dots)$, ξ_i 's are independent random variables with identical distribution μ which induces a unique system of orthogonal polynomial $H_i(x)$'s [?, ?, ?]. The index set \mathcal{J} contains sequences of integers whose number of non-zero entries is finite, i.e.,

$$\mathcal{J} = \{(\alpha_1, \alpha_2, \dots) \mid \alpha_i = 0, 1, \dots, \sum_{i=1}^{\infty} \alpha_i < \infty\}$$

We call $\{T_\alpha(\xi(\omega))\}_{\alpha \in \mathcal{J}}$ PC basis and $\{A_i(\omega)\}_{i=1}^{\infty}$ DSMM basis. It is seen immediately that the PC basis is a double-infinity sequence, i.e., both the order of one-dimensional PC basis and the dimensionality of random space can go to infinity, which leads to the so-called the cursor of dimensionality. To use the PC basis, we need to truncate the expansion on both directions which, due to the nature of tensor product, will generally results in a large number of coefficients to solve.

We have shown in the previous section that it is possible that the effective dimensionality of the random space where the solution resides can be much smaller than it appears to be in PC basis. Therefore we focus us on the order of polynomials used in the PC basis since polynomials of high order are known to be very

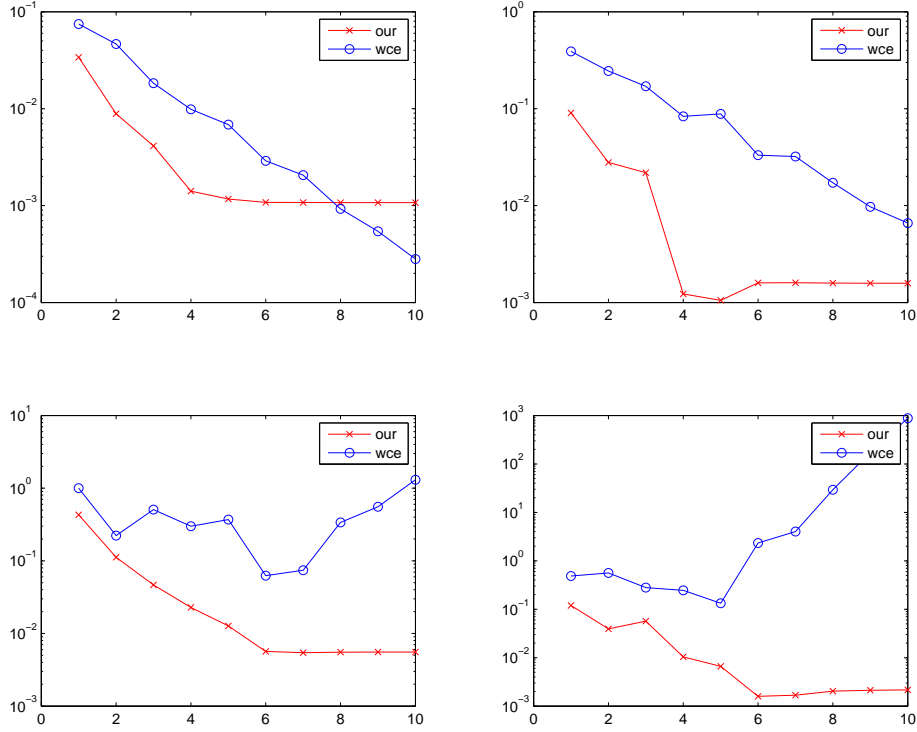


Figure 2: Error Comparison between PC and DDSMM. The errors of the mean, the variance, the third and the fourth central moments are shown, respectively from the left to the right and from the top to the bottom.

oscillatory and have bad reputations in global interpolation. To this end, we restrict us in the case of one-dimensional random space in the following example, i.e., $\xi = (\xi_1)$. The random elliptic coefficient adopted is moderately oscillatory in space and has some random phase,

$$a(x, \omega) = 2(\sin(3\pi x + 10\xi) + 1.01) \quad (5.1)$$

where ξ is a Gaussian random variable with 0 mean and unit variance. The DDSMM basis is constructed for $f(x) = 1$ and both online computation and PC method are applied to $f(x) = \sin(2\pi x) + 5 \cos(4\pi x)$ with errors shown in fig.2

The accuracy of our method depends on the accuracy of offline MC simulations, thus the decay of the error of the mean solution slows down noticeably after the certain number of random basis. Before this number, our method generally achieves accuracy of one or two orders higher than PC method with the same number of basis. Since the online computational cost of DDSMM is about the same as that of PC method for the same number of random basis. Thus our method may use much less computational resource to achieve the same accuracy requirement. For higher order (≥ 3) moments, PC method suffers from numerical instability and tends to blow up as the number of basis increases which is not a surprising result considering the oscillatory nature of high order polynomial. Therefore, high order Polynomial Chaos should be avoided especially high order moments are of interests.

Randomness messes up the elegant deterministic world in many possible ways. One typical scenario arising in practice is that multiple sources of randomness coexist in the system. A typical example of such scenario

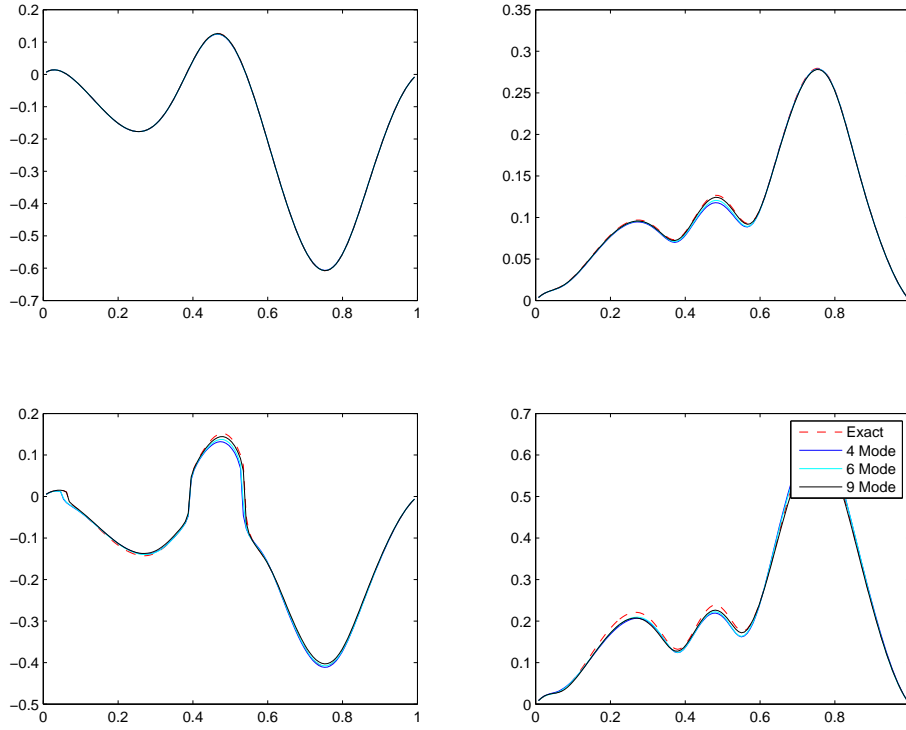


Figure 3: Moments of the solution of one-dimensional elliptic PDE with random coefficients of large variance. The mean, the variance, the third and the fourth central moments of the random solution are shown, respectively from the left to the right and from the top to the bottom. The exact solution is computed with 1.0×10^7 realizations.

is atmospheric turbulence in which many eddies of intermediate length scale constitute multiple sources of randomness. The variance of quantities we are interested in may be comparable to the mean value in magnitude. To demonstrate the effectiveness of DDSMM for this scenario, we considered

$$a(x, \omega) = 1 + \sum_{i=1}^4 D_i \xi_i (\sin(C_i \pi x) + 1) \quad (5.2)$$

where $D = (100, 120, 200, 150)$, $C = (1.2, 2.3, 3.1, 4.3)$ and ξ_i 's are iid random variables. Clearly, this elliptic coefficient has four sources of random. The results are compared with the exact solution in fig.3. Obviously, our method achieves very good approximation for the mean and moments even with only $M = 4$, four random basis.

If the stochastic dimensionality and the highest order of polynomials in the PC basis are r and p , respectively, PC method requires solving $P = \frac{(r+p)!}{r!p!}$ coefficients from a coupled PDE system. Further, if the coupled PDE system is discretized along each direction by N elements, the size of stiffness matrix is $O((PN^d)^2)$, which is astronomically large even for moderate values. In this example, the dimension of random space is $r = 4$ and, even for $p = 3$, the truncated PC expansion has $P = 35$ terms, while our method gives very good approximations by using 4 basis and nearly indistinguishable results by using 9 basis.

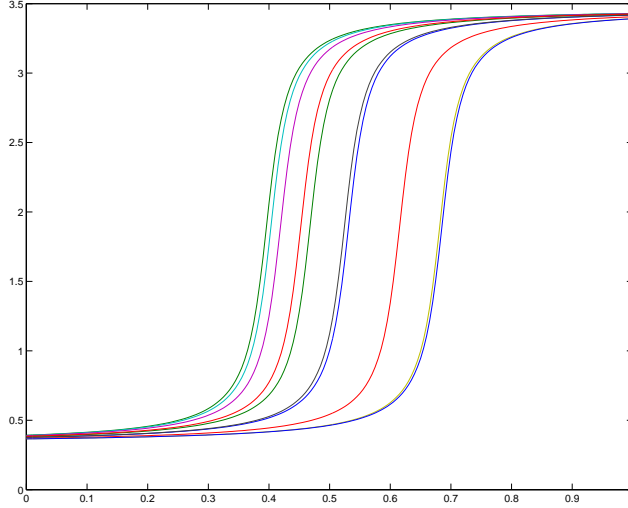


Figure 4: Elliptic coefficient with single jump at random location.

Another typical scenario arising in industrial world is that the randomness is constrained in some spatial region, and some properties may change dramatically, but the exact location of such jump is random. For example, the interfaces of composite material may be random and the properties of materials on both sides of the interface may differ dramatically. Another example is the shocks surrounding air-crafts whose onset locations are affected by many factors and thus the exact locations of shocks can be viewed random. Our method can deal with such cases without any substantial modification. To this end, we consider

$$a(x, \omega) = \arctan(40(x - 0.5 + 0.4(\xi - 0.5))) + 1.9 \quad (5.3)$$

where ξ is a uniformly distributed random variable $[0, 1]$. Some typical realizations of the elliptic coefficients are shown in fig.4 with results by DDSMM shown in fig.5

5.2 2D Elliptic PDE with Random Coefficients

We further apply our method to two-dimensional elliptic problem with random coefficients. Similar to one-dimensional cases, two typical scenarios are considered here, solutions of large variance and elliptic coefficients jumping at random locations. For the first case, the following isotropic elliptic coefficient is used with results by DDSMM shown in fig.6.

$$a(x, y, \omega) = 0.1 + \sum_{i=1}^2 D_i \xi_i (\sin(E_i \pi x + F_i \pi y) + 1) \quad (5.4)$$

where $D = (10, 30)$, $E = (2.3, 3.4)$, $F = (3.7, -5.1)$ and ξ_i 's are uniform random variable on $[0, 1]$. Again, the DDSMM random basis $\{A_i(\omega)\}_{i=1}^M$ is constructed for $f(x) = 1$ and applied to the case $f(x) = \sin(1.3\pi x + 3.4\pi y) + 8 \cos(4.3\pi x - 3.1\pi y)$

For the second case, we consider a line-type jump at random locations, i.e.,

$$a(x, y, \omega) = \arctan\{20[(x - 0.5 + 0.2\xi_1) + (y - 0.5 + 0.2\xi_2)]\} + 1.9 \quad (5.5)$$

where ξ_i 's are uniform random variable on $[-0.5, +0.5]$. Clearly, our method can capture moments of the random solutions with a few of basis.

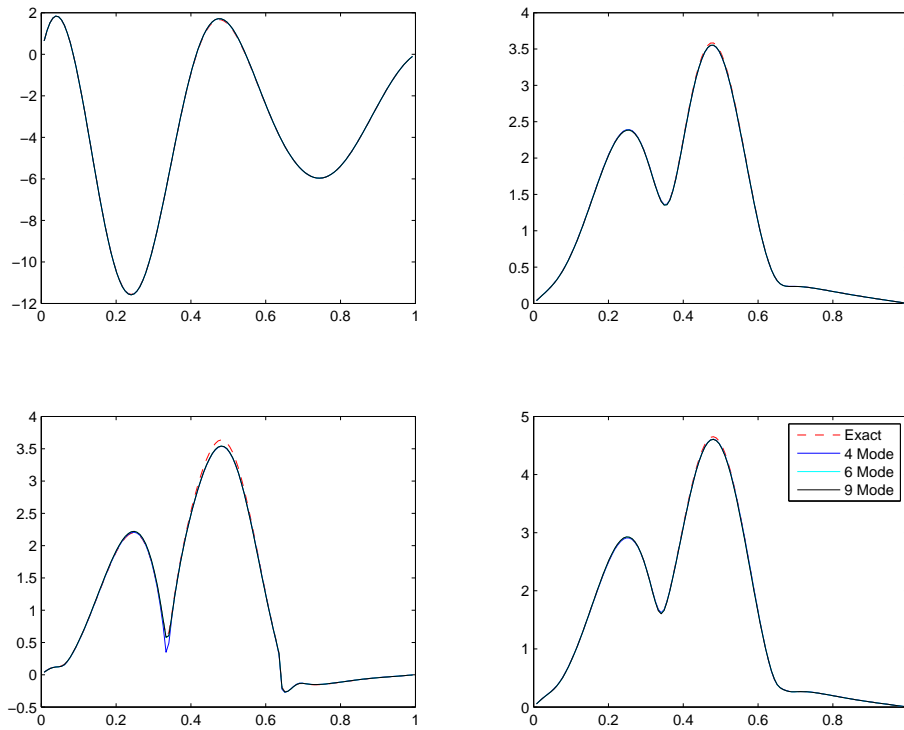


Figure 5: Moments of the solution of one-dimensional elliptic PDE with coefficient jumping at a single random location. The mean, the variance, the third and the fourth central moments of the random solution are shown, respectively from the left to the right and from the top to the bottom. The exact solution is computed with 1.0×10^7 realizations.

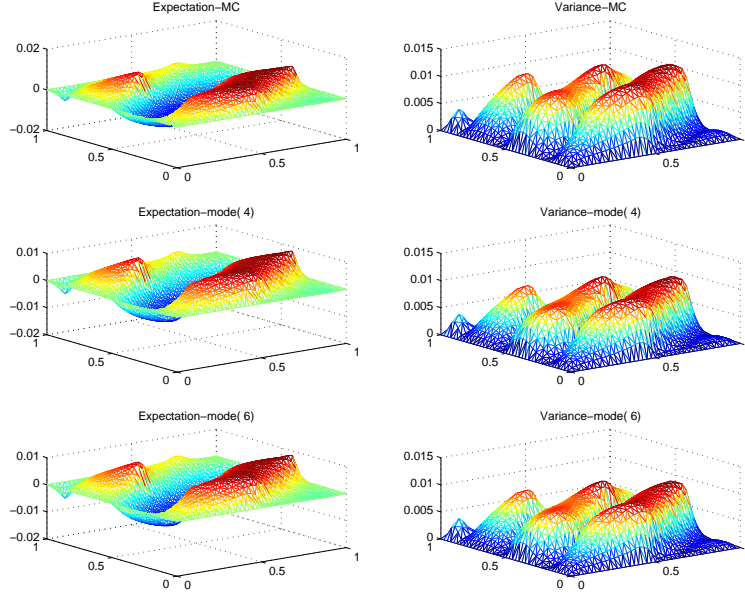


Figure 6: Moments of the solution of two-dimensional elliptic PDE with coefficient jumping at random locations. The mean, the variance, the third and the fourth central moments of the random solution are shown, respectively from the left to the right and from the top to the bottom. The exact solution is computed with 1.0×10^5 realizations.

5.3 2D Helmholtz equation in random media

In this example, we test our method on a different PDE, Helmholtz equation, in the form of

$$\nabla^2 u(x, \omega) + k^2(1 + n^2(x, \omega))u(x, \omega) = 0,$$

with boundary conditions

$$\begin{aligned} \frac{\partial u}{\partial \vec{n}} - iku &= 0 \quad \text{on } \Gamma_1, \\ \frac{\partial u}{\partial \vec{n}} &= 0 \quad \text{on } \Gamma_2, \\ u(x, \omega) &= f(x) \quad \text{on } \Gamma_3, \end{aligned}$$

where \vec{n} is the unit normal of the boundary, k is the wave number, and $n^2(x, \omega)$ is the random reflectivity of the media. Now the deterministic function $f(x)$ appears on part of the boundary, instead of the right-hand-side of the PDE. We also call this the Horn problem because of its derivation and physical structure (see Fig.8(a)). For numerical testing, we choose the random reflectivity of the media

$$n^2(x, \omega) = [\xi_1(\omega) \sin^2(2\pi x) \sin^2(2\pi y) + \xi_2(\omega) \cos^2(6\pi x) \sin^2(8\pi y)]^2,$$

with ξ_1 and ξ_2 are i.i.d uniformly distributed random variables in $[0, 1]$.

Again we pick $f(x) = 1$ to construct the basis off-line, and recover the Horn solution with $f(x) = \sin(2\pi) \sin(2\pi y) + 2$ on-line. For the comparison purpose, we also ran the Monte-Carlo simulations to get

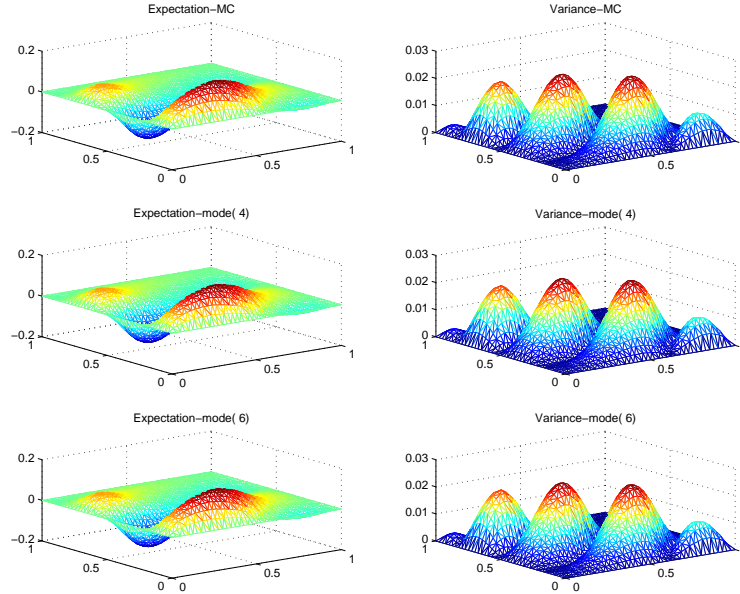


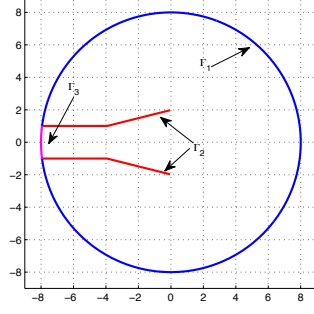
Figure 7: Moments of the solution of two-dimensional elliptic PDE with coefficient jumping at random locations. The mean, the variance, the third and the fourth central moments of the random solution are shown, respectively from the left to the right and from the top to the bottom. The exact solution is computed with 1.0×10^5 realizations.

the moments of the solution, e.g., mean and standard deviation of the Horn solution. It is important to point out that when wave number k increases, the variance of the solution increases, so do the eigenvalues of the K-L expansion. Consequently for a fixed error tolerance, the problem of high-frequency needs significantly more bases than that of low-frequency (see Fig.8(b)). For example, if we discard all eigen-directions with eigenvalues less than 10^{-4} , we need as much as 11 bases for $k = 2$, comparing to 4 bases for $k = 0.7$. In particular, if we further increase the level to 10^{-6} , we may need as much as 20 bases for high-frequency case, which makes our method impractical.

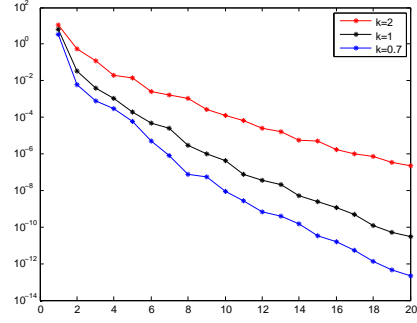
6 Conclusions

In this paper, the problem we target takes the general form $\mathcal{L}(\omega)u(x, \omega) = f(x)$ where randomness only appears in the stochastic differential operator. Thus, it is reasonable, at least heuristically, to conjecture that there exists a random basis which depends only, or largely, on $\mathcal{L}(\omega)$ under which the solution for a board class of right-hand side $f(x)$ has some sparse representation. We have verified this conjecture by many numerical examples presented in the previous sections and are working on rigorous analysis.

Under the framework of Monte Carlo or Polynomial Chaos expansion, the computation must be re-done from the very beginning every time the right-hand-side function $f(x)$ is changed which happens often in the design phase. Actually, many information from previous computations may be re-usable for the future computation. However, the questoin is how! Our method constructs a random basis based on KL expansion to extract the effective subspace of the output random space in the offline part and then reuse it in the online computation. In this sense, DDSMM can be viewed as partially inversion of the random differential operator. Obviously, the online computation cost of DDSMM is only a fraction of that of Monte Carlo methods. Furthermore, because



(a) Horn problem.



(b) Eigenvalue distribution of K-L expansion of the Horn solution for different wave number k .

Figure 8: (Left) Physical domain of the horn problem, where we have far-field boundary condition on Γ_1 , non-flow boundary condition on Γ_2 , and Dirichlet boundary condition on Γ_3 . (right) Display the leading eigenvalue distribution of K-L expansion for different wave number k .

the random basis in DDSMM is a sparse basis, it certainly requires much fewer random modes to achieve specified accuracy and much less computational resources compared to Polynomial Chaos expansion.

Our method is also semi-non-intrusive in the sense that certified legacy codes can be used with minimum changes in offline part. This property makes it very attractive to industrial community as it can reuse codes verified by possibly expansive and time-consuming experiments and avoid another life cycle of developing new programs.

We should point out that DDSMM can be extended to the cases where the right-hand-side function involves of randomness, i.e., $f(x, \omega)$ and the SPDE is time dependent, which is the ongoing research work. Offline and online computations in this framework may have different characteristics of parallelism. We intend to exploit for 3D problems these difference on appropriate parallel architectures, for example, offline on GPU and online part on distributed-memory cluster.

Acknowledgements

This work was in part supported by the AFOSR MURI grant FA9550-09-1-0613.

Appendix (Proof of Stability)

Consider elliptic equation with random coefficients

$$-\frac{\partial}{\partial x_i} \left(a_{ij}(x, \omega) \frac{\partial u}{\partial x_j}(x, \omega) \right) = f(x), \quad x \in D \subset \mathbb{R}^d \quad (6.1)$$

$$u(x, \omega) = 0 \quad x \in \partial D \quad (6.2)$$

The solution $u(x, \omega)$ can be written in terms of Green's function

$$u(x, \omega) = \int_D G(x, y, \omega) f(y) dy \quad (6.3)$$

If the expectation of Green's function

$$\bar{G}(x, y) = \mathbb{E} [G(x, y, \omega)] \quad (6.4)$$

exists and the covariance function

$$\text{Cov}_G(x, y, x', y') = \mathbb{E} [(G(x, y, \omega) - \bar{G}(x, y))(G(x', y', \omega) - \bar{G}(x', y')))] \quad (6.5)$$

exists and continuous, Green's function $G(x, y, \omega)$ has KL expansion

$$G(x, y, \omega) = \bar{G}(x, y) + \sum_{i=1}^{\infty} \sqrt{\lambda_i} G_i(x, y) B_i(\omega) \quad (6.6)$$

where $\{\lambda_i, G_i(x, y)\}_{i=1}^{\infty}$ are eigenpairs of covariance function $\text{Cov}_G(x, y, x', y')$

$$\int_{D \times D} \text{Cov}_G(x, y, x', y') G_i(x', y') dx' dy' = \lambda_i G_i(x, y) \quad (6.7)$$

with orthogonality conditions

$$\int_{D \times D} G_i(x, y) G_j(x, y) dx dy = \delta_{ij} \quad (6.8)$$

$$\mathbb{E} [B_i(\omega) B_j(\omega)] = \delta_{ij} \quad (6.9)$$

and

$$B_i(\omega) = \frac{1}{\sqrt{\lambda_i}} \int_{D \times D} (G(x, y, \omega) - \bar{G}(x, y)) G_i(x, y) dx dy \quad (6.10)$$

Then the solution $u(x, \omega)$ can be written as

$$u(x, \omega) = \sum_{i=0}^{\infty} \sqrt{\lambda_i} F_i(x) B_i(\omega) \quad (6.11)$$

where

$$F_i(x) = \int_D G_i(x, y) f(y) dy \quad (6.12)$$

Lemma 6.1. $\|F_i(x)\|_2 \leq \|f(x)\|_2$

Proof. For each x ,

$$\begin{aligned} |F_i(x)|^2 &= \left| \int_D G_i(x, y) f(y) dy \right|^2 \\ &\leq \int_D G_i(x, y)^2 dy \int_D f^2(y) dy \\ &= \int_D G_i(x, y)^2 dy \|f(y)\|_2^2 \end{aligned}$$

Integrate on both sides

$$\begin{aligned} \|F_i(x)\|_2^2 &= \int_D F_i^2(x) dx \\ &\leq \int_D \int_D G_i(x, y)^2 dy dx \|f(y)\|_2^2 \\ &= \|f(y)\|_2^2 \end{aligned}$$

The orthogonality of G_i is used in the last equality. □

Now consider KL expansion of $u(x, \omega)$

$$u(x, \omega) = \bar{u}(x) + \sum_{i=1}^{\infty} \sqrt{\theta_i} \phi_i(x) A_i(\omega) \quad (6.13)$$

where $\{\sqrt{\theta_i}, \phi_i(x)\}_{i=1}^{\infty}$ are eigenpairs of the covariance function $\text{Cov}_u(x, y)$, i.e.,

$$\int_D \text{Cov}_u(x, y) \phi_i(y) dy = \theta_i \phi_i(x) \quad (6.14)$$

and

$$A_i(\omega) = \frac{1}{\sqrt{\theta_i}} \int_D u(x, \omega) \phi_i(x) dx \quad (6.15)$$

with bi-orthogonality

$$\begin{aligned} \int_D \phi_i(x) \phi_j(x) dx &= \delta_{ij} \\ \mathbb{E}[A_i(\omega) A_j(\omega)] &= \delta_{ij} \end{aligned}$$

Combine eqn.6.11 and eqn.6.13, we have

$$\sum_{j=1}^{\infty} \sqrt{\lambda_j} F_j(x) B_j(\omega) = \sum_{j=1}^{\infty} \sqrt{\theta_j} \phi_j(x) A_j(\omega)$$

in which the mean $\bar{u}(x)$ has been cancelled on both sides. Multiple both sides by $F_i(x)$ and integrate over D ,

$$\sum_{j=1}^{\infty} \alpha_{ij} \sqrt{\lambda_j} B_j(\omega) = \sum_{j=1}^{\infty} \beta_{ij} \sqrt{\theta_j} A_j(\omega) \quad (6.16)$$

where

$$\begin{aligned} \alpha_{ij} &= \int_D F_i(x) F_j(x) dx \\ \beta_{ij} &= \int_D F_i(x) \phi_j(x) dx \end{aligned}$$

The equality can be written in matrix form

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \begin{pmatrix} \Lambda_1^{\frac{1}{2}} & 0 \\ 0 & \Lambda_2^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix} \begin{pmatrix} \Theta_1^{\frac{1}{2}} & 0 \\ 0 & \Theta_2^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \quad (6.17)$$

where $\alpha_{11} \in \mathbb{R}^{N_1 \times N_1}$, $\Lambda_1 = \text{diag}(\lambda_1, \dots, \lambda_{N_1})$, $\Lambda_2 = \text{diag}(\lambda_{N_1+1}, \dots)$, $\beta_{11} \in \mathbb{R}^{N_1 \times N_2}$, $\Theta_1 = \text{diag}(\theta_1, \dots, \theta_{N_2})$, $\Theta_2 = \text{diag}(\theta_{N_2+1}, \dots)$, $B_1 \in \mathbb{R}^{N_1}$ and $A_1 \in \mathbb{R}^{N_2}$. or

$$\alpha_{11} \Lambda_1^{\frac{1}{2}} B_1 + \alpha_{12} \Lambda_2^{\frac{1}{2}} B_2 = \beta_{11} \Theta_1^{\frac{1}{2}} A_1 + \beta_{12} \Theta_2^{\frac{1}{2}} A_2 \quad (6.18)$$

$$\alpha_{21} \Lambda_1^{\frac{1}{2}} B_1 + \alpha_{22} \Lambda_2^{\frac{1}{2}} B_2 = \beta_{21} \Theta_1^{\frac{1}{2}} A_1 + \beta_{22} \Theta_2^{\frac{1}{2}} A_2 \quad (6.19)$$

Before proceeding to the error analysis, we have the following bounds

Lemma 6.2.

$$\begin{aligned} |\alpha_{ij}| &\leq \|F_i(x)\|_2 \|F_j(x)\|_2 \leq \|f(x)\|_2^2 \\ |\beta_{ij}| &\leq \|F_i(x)\|_2 \leq \|f(x)\|_2 \end{aligned}$$

Suppose we can construct $\{A_i(\omega)\}_{i=1}^{N_2}$ exactly without numerical errors in offline computation and apply this basis to the case with right hand side function $g(x) \in \mathbb{L}^2(D)$. By solving the coupled equations resulting from Galerkin projection, we have the numerical solution

$$\tilde{v}(x, \omega) = \sum_{i=0}^{N_2} \tilde{v}_i(x) A_i(\omega) \quad (6.20)$$

The solution can be written in terms of Green's function

$$v(x, \omega) = \sum_{i=0}^{\infty} \sqrt{\lambda_i} H_i(x) B_i(\omega) \quad (6.21)$$

where

$$H_i(x) = \int_D G_i(x, y) g(y) dy. \quad (6.22)$$

By Lemma (6.1), it is clear that $\|H_i(x)\|_2 \leq \|g(x)\|_2$. In addition, the expansion of $v(x, \omega)$ on the basis $A_i(\omega)$ is

$$v(x, \omega) = \sum_{i=0}^{\infty} v_i(x) A_i(\omega). \quad (6.23)$$

Thus the error is

$$\begin{aligned} \mathcal{E}(x, \omega) &= v(x, \omega) - \tilde{v}(x, \omega) \\ &= \mathcal{E}_A(x, \omega) + \mathcal{E}_B(x, \omega), \end{aligned}$$

where

$$\mathcal{E}_A(x, \omega) = \sum_{i=0}^{N_2} [v_i(x) - \tilde{v}_i(x)] A_i(\omega), \quad (6.24)$$

$$\mathcal{E}_B(x, \omega) = \sum_{i=1}^{\infty} \sqrt{\lambda_i} H_i(x) B_i(\omega) - \sum_{i=1}^{N_2} v_i(x) A_i(\omega). \quad (6.25)$$

The first error is due to Galerkin projection, and the second is due to approximation of B_i by A_i . For simplicity, we assume $\|g(x)\|_2 = 1$. To bound the second error term, we recall

$$v_j(x) = \mathbb{E} [v(x, \omega) A_j(\omega)] = \sum_{i=1}^{\infty} \sqrt{\lambda_i} H_i(x) \mathbb{E} [B_i(\omega) A_j(\omega)], \quad (6.26)$$

and plug it into $\mathcal{E}_B(x, \omega)$. We have, after changing the order of summation,

$$\mathcal{E}_B(x, \omega) = \mathcal{E}_{B1}(x, \omega) + \mathcal{E}_{B2}(x, \omega), \quad (6.27)$$

and

$$\begin{aligned} \mathcal{E}_{B1}(x, \omega) &= \sum_{i=1}^{N_1} \sqrt{\lambda_i} H_i(x) \left(B_i(\omega) - \sum_{j=1}^{N_2} \mathbb{E} [B_i(\omega) A_j(\omega)] A_j(\omega) \right), \\ \mathcal{E}_{B2}(x, \omega) &= \sum_{i=N_1+1}^{\infty} \sqrt{\lambda_i} H_i(x) \left(B_i(\omega) - \sum_{j=1}^{N_2} \mathbb{E} [B_i(\omega) A_j(\omega)] A_j(\omega) \right). \end{aligned}$$

Theorem 6.3. *Given the condition that N_1 satisfies*

$$\sum_{i=N_1+1}^{\infty} \sqrt{\lambda_i} \leq \epsilon,$$

there exists \bar{N}_2 , such that for $\forall N_2 \geq \bar{N}_2$, we have

$$\mathbb{E} \left[\int |\mathcal{E}_B(x, \omega)|^2 dx \right] \leq 2\epsilon^2.$$

Proof. Proof is given by the following inequality

$$\mathbb{E} \left[\int |\mathcal{E}_B(x, \omega)|^2 dx \right] \leq \mathbb{E} \left[\int |\mathcal{E}_{B_1}(x, \omega)|^2 dx \right] + \mathbb{E} \left[\int |\mathcal{E}_{B_1}(x, \omega)|^2 dx \right],$$

and Lemma (6.4), (6.5). □

Lemma 6.4. *Given the condition that N_1 satisfies*

$$\sum_{i=N_1+1}^{\infty} \sqrt{\lambda_i} \leq \epsilon,$$

we have

$$\mathbb{E} \left[\int |\mathcal{E}_{B_1}(x, \omega)|^2 dx \right] \leq \epsilon^2.$$

Proof. (Lemma 6.4) Based on Cauchy-Schwarz inequality, we have

$$\begin{aligned} \mathbb{E} \left[\int |\mathcal{E}_{B_1}(x, \omega)|^2 dx \right] &= \mathbb{E} \left[\int \left| \sum_{i=N_1+1}^{\infty} \sqrt{\lambda_i} H_i(x) \left(B_i(\omega) - \sum_{j=1}^{N_2} \mathbb{E} [B_i(\omega) A_j(\omega)] A_j(\omega) \right) \right|^2 dx \right] \\ &\leq \sum_{i=N_1+1}^{\infty} \sqrt{\lambda_i} \int |H_i(x)|^2 dx \cdot \sum_{i=N_1+1}^{\infty} \sqrt{\lambda_i} \mathbb{E} \left[\left| B_i(\omega) - \sum_{j=1}^{N_2} \mathbb{E} [B_i(\omega) A_j(\omega)] A_j(\omega) \right|^2 \right] \\ &\leq \left(\sum_{i=N_1+1}^{\infty} \sqrt{\lambda_i} \right)^2. \end{aligned}$$

The last inequality is given by

$$\|H_i(x)\|_2 \leq \|g(x)\|_2 = 1,$$

and

$$\mathbb{E} \left[\left| B_i(\omega) - \sum_{j=1}^{N_2} \mathbb{E} [B_i(\omega) A_j(\omega)] A_j(\omega) \right|^2 \right] = 1 - \sum_{j=1}^{N_2} (\mathbb{E} [B_i(\omega) A_j(\omega)])^2 \leq 1.$$

□

Lemma 6.5. *For a fixed N_1 , there exists a large \bar{N}_2 such that $N_2 \geq \bar{N}_2$ we have*

$$\mathbb{E} \left[\int |\mathcal{E}_{B_2}(x, \omega)|^2 dx \right] \leq \epsilon.$$

Proof. (Lemma 6.5) Now apply Cauchy-Schwarz inequality again on the error term $\mathcal{E}_{B2}(x, \omega)$ and we have

$$\begin{aligned} \mathbb{E} \left[\int |\mathcal{E}_{B2}(x, \omega)|^2 dx \right] &= \mathbb{E} \left[\int \left| \sum_{i=1}^{N_1} \sqrt{\lambda_i} H_i(x) \left(B_i(\omega) - \sum_{j=1}^{N_2} \mathbb{E}[B_i(\omega) A_j(\omega)] A_j(\omega) \right) \right|^2 dx \right] \\ &\leq \sum_{i=1}^{N_1} \int |H_i(x)|^2 dx \cdot \sum_{i=1}^{N_1} \lambda_i \mathbb{E} \left[\left| B_i(\omega) - \sum_{j=1}^{N_2} \mathbb{E}[B_i(\omega) A_j(\omega)] A_j(\omega) \right|^2 \right] \\ &\leq N_1 \cdot \sum_{i=1}^{N_1} \lambda_i \left(1 - \sum_{j=1}^{N_2} (\mathbb{E}[B_i(\omega) A_j(\omega)])^2 \right). \end{aligned}$$

Since $\{A_i\}_{i=1, \dots}$ is a complete orthonormal basis of the space $L^2(\Omega)$, for any normalized random variable $B(\omega)$ we have

$$\lim_{N_2 \rightarrow \infty} \sum_{j=1}^{N_2} (\mathbb{E}[B(\omega) A_j(\omega)])^2 = 1.$$

Hence there exists integer $N^{(i)}$, such that

$$\sum_{j=1}^{N^{(i)}} (\mathbb{E}[B_i(\omega) A_j(\omega)])^2 \geq 1 - \min\left(1, \frac{\epsilon^2}{N_1^2 \lambda_i}\right).$$

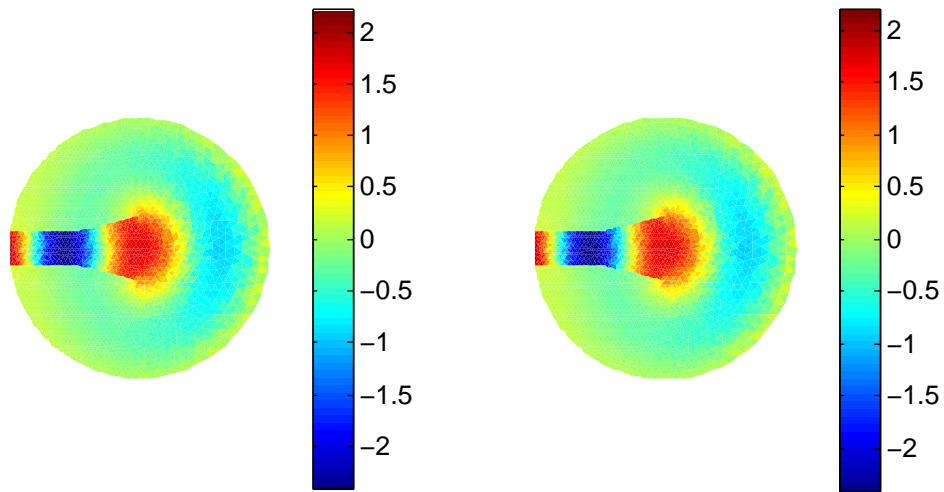
Now we choose

$$N_2 \geq \bar{N}_2 = \max_{i=1, \dots, N_1} N^{(i)},$$

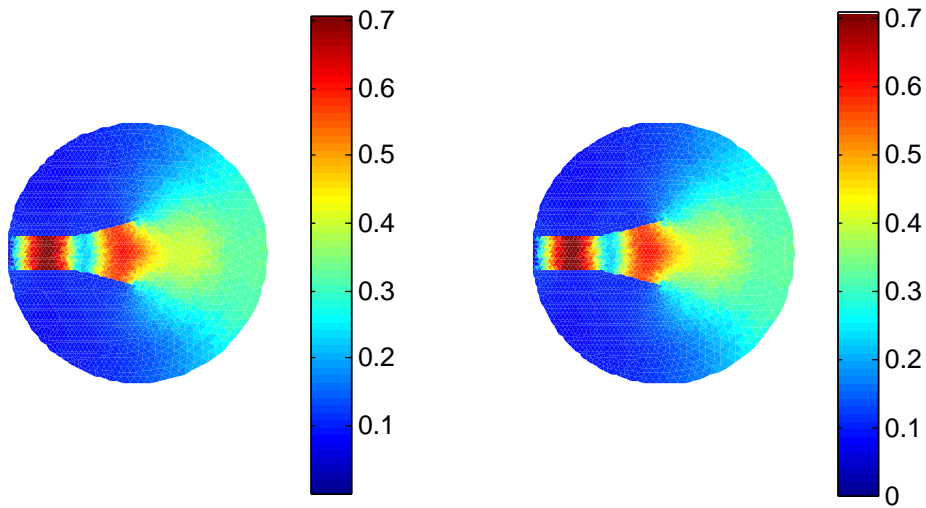
and have

$$\begin{aligned} \mathbb{E} \left[\int |\mathcal{E}_{B2}(x, \omega)|^2 dx \right] &\leq N_1 \cdot \sum_{i=1}^{N_1} \lambda_i \left(1 - \sum_{j=1}^{N_2} (\mathbb{E}[B_i(\omega) A_j(\omega)])^2 \right) \\ &\leq N_1 \cdot \sum_{i=1}^{N_1} \lambda_i \left(1 - (1 - \min\left(1, \frac{\epsilon^2}{N_1^2 \lambda_i}\right)) \right) \\ &\leq \epsilon^2 \end{aligned}$$

□



(a) Mean solution of horn problem.



(b) Standard deviation of the solution

Figure 9: (left) Monte Carlo results; (right) Recovered results via DDSMM using only 4 modes.

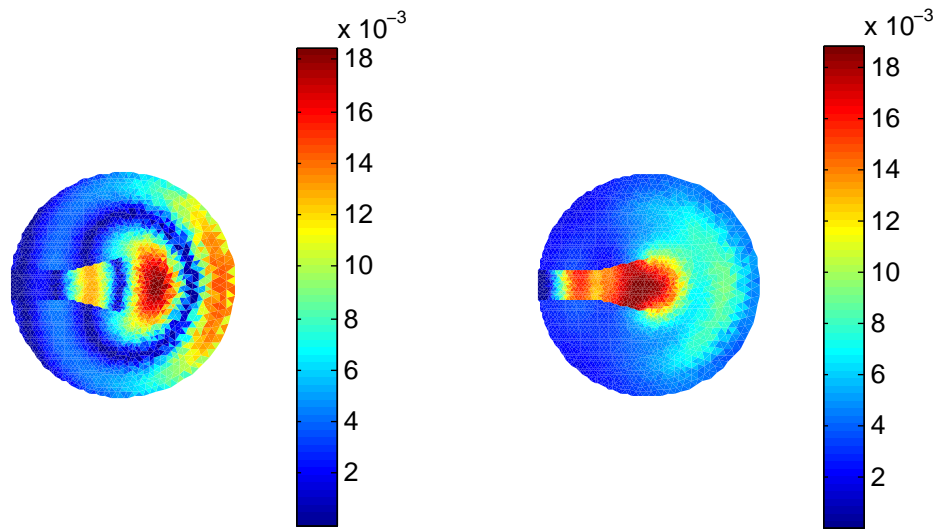


Figure 10: Relative errors of mean solution (left) and standard deviation (right).