International Journal for Uncertainty Quantification, 9(2):161–186 (2019)

NUMERICAL APPROXIMATION OF ELLIPTIC PROBLEMS WITH LOG-NORMAL RANDOM COEFFICIENTS

Xiaoliang Wan[1] & Haijun Yu[2,*]

[1]*Department of Mathematics, Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803, USA*

[2]*NCMIS & LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Beijing 100190; School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*

*Address all correspondence to: Haijun Yu, NCMIS & LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Beijing 100190; School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China, E-mail: hyu@lsec.cc.ac.cn

Original Manuscript Submitted: 10/11/2018; Final Draft Received: 2/18/2019

In this work, we consider a non-standard preconditioning strategy for the numerical approximation of the classical elliptic equations with log-normal random coefficients. In earlier work, a Wick-type elliptic model was proposed by modeling the random flux through the Wick product. Due to the lower-triangular structure of the uncertainty propagator, this model can be approximated efficiently using the Wiener chaos expansion in the probability space. Such a Wick-type model provides, in general, a second-order approximation of the classical one in terms of the standard deviation of the underlying Gaussian process. Furthermore, when the correlation length of the underlying Gaussian process goes to infinity, the Wick-type model yields the same solution as the classical one. These observations imply that the Wick-type elliptic equation can provide an effective preconditioner for the classical random elliptic equation under appropriate conditions. We use the Wick-type elliptic model to accelerate the Monte Carlo method and the stochastic Galerkin finite element method. Numerical results are presented and discussed.

KEY WORDS: *Wiener chaos expansion, Wick product, stochastic elliptic PDE, uncertainty quantification, log-normal random coefficient*

1. INTRODUCTION

Numerical approximation of elliptic problems with log-normal random coefficients has received a lot of attention. We consider the following mathematical model,

$$\text{Model I:} \quad \begin{cases} -\nabla \cdot (a(\boldsymbol{x},\omega)\nabla u(\boldsymbol{x},\omega)) = f(\boldsymbol{x}), & \boldsymbol{x} \in D, \\ u(\boldsymbol{x},\omega) = 0, & \boldsymbol{x} \in \partial D, \end{cases} \tag{1}$$

where $\ln a(\boldsymbol{x},\omega)$ is a second-order homogeneous Gaussian random process, and the force term is assumed to be deterministic for simplicity. We call problem (1) model I in this paper. Theoretical difficulties of problem (1) are mainly related to the lack of uniform ellipticity, where the Lax-Milgram lemma is not applicable. The existence and uniqueness of the solution of problem (1) are usually established with respect to a weighted norm [1–3] or a weighted measure [4], or by using the Fernique theorem [5,6]. Considering the Wiener chaos approach and Galerkin projection [1,7], the difficulties of numerical approximation of problem (1) are two-fold: First, if we start from the theoretical study [2,4], a different test space rather than $L_2(\mathbb{F}; H_0^1(D))$ is required, which may be not easy to construct. Here

 © 2019 by Begell House, Inc. www.begellhouse.com

$\mathbb{F} := (\Omega, \mathcal{F}, P)$ is the probability space for ω; a detailed presentation of $\mathbb{F}$ is given in Section 2. Second, if we choose $L_2(\mathbb{F}; H_0^1(D))$ as the test space and use Wiener chaos as the basis for the probability space, although no divergence with respect to the $L_2(\mathbb{F}; H_0^1(D))$ norm has been numerically observed (the solution of problem (1) actually belongs to $L_2(\mathbb{F}; H_0^1(D))$ [6]), the stiffness matrix is full and dense. In other words, an efficient preconditioner is required. Studies of elliptic problems with other types of random coefficients can be found in [8–10], etc.

The elliptic equation with a log-normal random coefficient has been studied by means of the perturbation technique (see, e.g., [11,12]), which has been also employed for other types of random coefficients (see, e.g., [13]). However, the perturbation method only works for small variability of the random coefficient and a low degree of the Taylor polynomial [11].

Another approach is to construct an auxiliary problem as some sort of preconditioner of the original problem; e.g., the idea of using a smoother version of the original problem (generated by a smoothing kernel) in a Monte Carlo control variate approach has been discussed by Nobile et al. [14,15]. Other known preconditioning skills include the traditional algebraic preconditioner [16,17] and the bifidelity method [18].

In this paper we take a new approach to construct an auxiliary problem used as a preconditioner of model I. From the modeling point of view, the randomness can be introduced in different ways. A typical strategy is to replace the flux $a\nabla u$ as $a \diamond \nabla u$ with $\diamond$ being the Wick product [19–21], motivated by the observations that the Wick product is consistent with the Skorokhod stochastic integral in a Hilbert space and can smooth the irregularity induced by white noise. Once the Wick product is adopted, the equations for the coefficients of Wiener chaos expansion are decoupled and can be solved one-by-one. Although this is a very nice property for numerical computation, the original equation is changed and the model difference becomes the main concern. In [22,23], a new Wick-type model was proposed by modeling the flux as $\left(a^{-1}\right)^{\diamond(-1)} \diamond \nabla u$:

$$\text{Model II:} \quad \begin{cases} -\nabla \cdot \left(\left(a^{-1}\right)^{\diamond(-1)}(\boldsymbol{x}, \omega) \diamond \nabla u(\boldsymbol{x}, \omega)\right) = f(\boldsymbol{x}), & \boldsymbol{x} \in D, \\ u(\boldsymbol{x}, \omega) = 0, & \boldsymbol{x} \in \partial D, \end{cases} \tag{2}$$

which we call model II in this paper. In general, both fluxes $a \diamond \nabla u$ and $\left(a^{-1}\right)^{\diamond(-1)} \diamond \nabla u$ will introduce a second-order approximation of the solution of model I in terms of the standard deviation ($\sigma < 1$) of the underlying Gaussian process. However, the latter choice provides a much smaller difference. Actually, when the correlation length of the underlying Gaussian process goes to infinity, model II has the same solution as model I. In addition, the uncertainty propagator of model II is also lower-triangular, which can be solved efficiently. Another way to approximate the flux $a\nabla u$ using the Wick product is to employ the Mikulevicius-Rozovskii (M-R) formula [24], which shows that the product of two random variables, say X and Y, has a Taylor-like expansion,

$$XY = X \diamond Y + \sum_{n=1}^{\infty} \frac{\mathcal{D}^n X \diamond \mathcal{D}^n Y}{n!}, \tag{3}$$

where $\mathcal{D}$ indicates the Malliavin derivative [25]. It is seen that $X \diamond Y$ is the lowest-order term in this expansion. We can include more terms from the M-R formula to get a better approximation of $a\nabla u$ [26,27]. It is shown in [27] that with respect to the truncation order Q of the Malliavin derivative and the standard deviation of the underlying Gaussian process such a strategy provides a difference of $\mathcal{O}(\sigma^{2(Q+1)})$ from the solution of model I. However, upon doing so, the corresponding uncertainty propagator will not be lower-triangular anymore, although the coupling in the upper-triangular part will be weak if the truncation order in the M-R formula is relatively small.

In this work, we will explore the possibility to use model II as a predictor to improve some algorithms for model I since model II can be approximated efficiently and the difference between models I and II can be very small. Depending on the properties of the random coefficient, we mainly consider the Monte Carlo method and the Wiener chaos approach with Galerkin projection for model I.

This paper is organized as follows. In Section 2, we define the Wiener chaos space and the Wick product. Stochastic elliptic models are discussed in Section 3 and the corresponding uncertainty propagators are given in Section 4. Numerical algorithms are proposed in Section 5. We present numerical results in Section 6, followed by a summary section.

2. WIENER CHAOS SPACE AND WICK PRODUCT

Since the underlying random variables of the model are i.i.d. Gaussian, whose corresponding stochastic orthogonal polynomials are Hermite, we first introduce the basic properties of Hermite polynomials.

2.1 Hermite Polynomials

The one-dimensional (probabilistic) Hermite polynomials of degree n are defined as

$$\mathrm{H}_n(\xi) := (-1)^n e^{\xi^2/2} \frac{d^n}{d\xi^n} e^{-\xi^2/2}. \tag{4}$$

$\mathrm{H}_n(\xi)$ are orthogonal with respect to the weight $(1/\sqrt{2\pi})e^{-\xi^2/2}$, in the sense

$$\int_{-\infty}^{\infty} \mathrm{H}_m(\xi)\mathrm{H}_n(\xi) \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2} d\xi = n!\delta_{nm}. \tag{5}$$

The values of Hermite polynomials can be evaluated using the following three-term recurrence formula:

$$\begin{aligned} &\mathrm{H}_0(\xi) = 1, \qquad \mathrm{H}_1(\xi) = \xi, \\ &\mathrm{H}_{n+1}(\xi) = \xi \mathrm{H}_n(\xi) - n\mathrm{H}_{n-1}(\xi), \quad n \geq 2. \end{aligned}$$

Hermite polynomials satisfy a very simple derivative relation:

$$\mathrm{H}_n'(\xi) = n\mathrm{H}_{n-1}(\xi), \quad \forall\, n \geq 0. \tag{6}$$

We list below in Lemma 1 several properties of Hermite polynomials, which will be used later.

Lemma 1. *For one-dimensional Hermite polynomials, the following properties hold:*

$$\exp\left(s\xi - \frac{1}{2}s^2\right) = \sum_{i=0}^{\infty} \frac{s^i}{i!} \mathrm{H}_i(\xi), \tag{7}$$

$$\mathrm{H}_n(\xi + s) = \sum_{i=0}^{n} \binom{n}{i} s^{n-i} \mathrm{H}_i(\xi), \tag{8}$$

$$\mathrm{H}_i(\xi)\mathrm{H}_j(\xi) = \sum_{k \leq i \wedge j} \chi(i,j,k) \mathrm{H}_{i+j-2k}(\xi). \tag{9}$$

where $s \in \mathbb{R}$, $i \wedge j := \min\{i, j\}$ *and*

$$\chi(i,j,k) = \frac{i!j!}{k!(i-k)!(j-k)!}.$$

2.2 Wick Product

Now we list the definition and some basic properties of the Wick product, which can be found in the existing literature (e.g., [19,28]).

The Wick product of a set of random variables with finite moments is defined recursively as follows:

$$\langle \emptyset \rangle = 1, \qquad \frac{\partial \langle X_1, \ldots, X_k \rangle}{\partial X_i} = \langle X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_k \rangle, \quad k \geq 1,$$

together with the constraint that the average is zero,

$$\mathbb{E}\langle X_1, \ldots, X_k \rangle = 0, \quad k \geq 1.$$

It follows that

$$\langle X\rangle = X - \mathbb{E}[X], \quad \langle X, Y\rangle = XY - \mathbb{E}[Y]X - \mathbb{E}[X]Y + 2\mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[XY].$$

If X, Y are independent, from the above formula, we know

$$\langle X, Y\rangle = \langle X\rangle \langle Y\rangle.$$

On the other hand, if $Y = X$, we get

$$\langle X, X\rangle = X^2 - 2\mathbb{E}[X]X + 2\mathbb{E}[X]^2 - \mathbb{E}[X^2].$$

Define $X \diamond Y := \langle X, Y\rangle$ and

$$P_n(X) := X^{\diamond n} = \underbrace{\langle X, \ldots, X\rangle}_{n \text{ times}},$$

then $P_n'(x) = nP_{n-1}(x)$.

The Wick product is closely related to Hermite polynomials. If ξ is a normally distributed variable with variance 1, then

$$\xi^{\diamond n} = \mathrm{H}_n(\xi), \tag{10}$$

and

$$\mathrm{H}_n(\xi) \diamond \mathrm{H}_m(\xi) = \mathrm{H}_{n+m}(\xi). \tag{11}$$

Using a Taylor series, one can define the exponential function of the Wick product as

$$e^{\diamond X} := \sum_{n=0}^{\infty} \frac{1}{n!} X^{\diamond n}. \tag{12}$$

For a normally distributed variable ξ, it can be checked that [19]

$$e^{\diamond[\sigma\xi]} = e^{\sigma\xi - \sigma^2/2}, \tag{13}$$

$$e^{\diamond[\sigma\xi]} \diamond e^{\diamond[-\sigma\xi]} = 1, \tag{14}$$

and the following statistics hold:

$$\mathbb{E}\left[e^{\diamond[\sigma\xi]}\right] = 1, \quad \mathrm{Var}\left[e^{\diamond[\sigma\xi]}\right] = e^{\sigma^2} - 1. \tag{15}$$

2.3 Wiener Chaos Space

We define $\mathbb{F} := (\Omega, \mathcal{F}, P)$ as a complete probability space, where $\mathcal{F}$ is the σ-algebra generated by the countably many i.i.d. Gaussian random variables $\{\xi_k\}_{k\geq 1}$. Define $\boldsymbol{\xi} := (\xi_1, \xi_2, \ldots)$. Let $\mathcal{J}$ be the collection of multi-indices $\boldsymbol{\alpha}$ with $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots)$ so that $\alpha_k \in \mathbb{N}_0$ and $|\boldsymbol{\alpha}| := \sum_{k\geq 1} \alpha_k < \infty$. For $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathcal{J}$, we define

$$\boldsymbol{\alpha} + \boldsymbol{\beta} = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots), \quad \boldsymbol{\alpha}! = \prod_{k\geq 1} \alpha_k!, \quad \binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}} = \prod_{k\geq 1} \binom{\alpha_k}{\beta_k}.$$

We use (0) to denote the multi-index with all zero entries: $(0)_k = 0$ for all k. Define the collection of random variables Ξ as follows:

$$\Xi := \{\mathrm{h}_{\boldsymbol{\alpha}}, \boldsymbol{\alpha} \in \mathcal{J}\}, \quad \mathrm{h}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) := \prod_{k\geq 1} \frac{1}{\sqrt{\alpha_k!}} \mathrm{H}_{\alpha_k}(\xi_k), \tag{16}$$

where $\mathrm{H}_n(\xi)$ are the one-dimensional (probabilistic) Hermite polynomials. For convenience, we also define

$$\mathrm{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) := \prod_{k\geq 1} \mathrm{H}_{\alpha_k}(\xi_k). \tag{17}$$

For any fixed k-dimensional i.i.d. Gaussian random variable $\boldsymbol{\xi}$, the following relations hold:

$$\mathbb{E}[\mathrm{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})\mathrm{H}_{\boldsymbol{\beta}}(\boldsymbol{\xi})] = \delta_{\boldsymbol{\alpha}\boldsymbol{\beta}}\boldsymbol{\alpha}!, \quad \mathbb{E}[\mathrm{h}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})\mathrm{h}_{\boldsymbol{\beta}}(\boldsymbol{\xi})] = \delta_{\boldsymbol{\alpha}\boldsymbol{\beta}}. \tag{18}$$

The set Ξ forms an orthonormal basis for $L_2(\mathbb{F})$ [29]; that is, if $\eta \in L_2(\mathbb{F})$, then

$$\eta = \sum_{\boldsymbol{\alpha}\in\mathcal{J}} \eta_{\boldsymbol{\alpha}}\mathrm{h}_{\boldsymbol{\alpha}}, \quad \eta_{\boldsymbol{\alpha}} = \mathbb{E}[\eta\mathrm{h}_{\boldsymbol{\alpha}}] \tag{19}$$

and

$$\mathbb{E}[\eta^2] = \sum_{\boldsymbol{\alpha}\in\mathcal{J}} \eta_{\boldsymbol{\alpha}}^2. \tag{20}$$

The Wick products of multidimensional stochastic Hermite polynomials are

$$\mathrm{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) \diamond \mathrm{H}_{\boldsymbol{\beta}}(\boldsymbol{\xi}) = \mathrm{H}_{\boldsymbol{\alpha}+\boldsymbol{\beta}}(\boldsymbol{\xi}), \qquad \mathrm{h}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) \diamond \mathrm{h}_{\boldsymbol{\beta}}(\boldsymbol{\xi}) = \sqrt{\frac{(\boldsymbol{\alpha}+\boldsymbol{\beta})!}{\boldsymbol{\alpha}!\boldsymbol{\beta}!}}\mathrm{h}_{\boldsymbol{\alpha}+\boldsymbol{\beta}}(\boldsymbol{\xi}). \tag{21}$$

Note that if we consider the expansion of $\mathrm{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})\mathrm{H}_{\boldsymbol{\beta}}(\boldsymbol{\xi})$ using the base set Ξ, it is obvious that there exist low-order terms in addition to $\mathrm{H}_{\boldsymbol{\alpha}+\boldsymbol{\beta}}(\boldsymbol{\xi})$; however, in the definition of the Wick product, all these low-order terms are removed; cf. Eqs. (9) and (21). Such a difference of the Wick product from the regular multiplication stems from the fact that the Wick product should be interpreted from the viewpoint of the stochastic integral. The correspondence between the Wick product and the Ito-Skorokhod integral can be found in [19,21,25,30].

For the numerical approximation, the number of Gaussian random variables and the polynomial order need to be truncated. We define

$$\mathcal{J}_{M,p} = \{\boldsymbol{\alpha} | \boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_M),\ |\boldsymbol{\alpha}| \leq p\}, \tag{22}$$

where $p \in \mathbb{N}_0$ is the maximum total degree. (To reduce the number of stochastic bases, one can also consider the sparse grids or sparse spectral Galerkin method; see, e.g., [13,15,31–33], where the overall procedure is similar.) Correspondingly, $\boldsymbol{\xi}$ is split into two parts:

$$\boldsymbol{\xi} = \boldsymbol{\xi}_1 \oplus \boldsymbol{\xi}_2 = (\xi_1, \ldots, \xi_M) \oplus (\xi_{M+1}, \ldots).$$

For simplicity, we use $\boldsymbol{\xi}$ for both finite-dimensional and infinite-dimensional cases, and the dimensionality will be indicated by the set $\mathcal{J}$ or $\mathcal{J}_{M,p}$ for the index. Let $N_{M,p}$ be the cardinality of $\mathcal{J}_{M,p}$. It is obvious that there exists a one-to-one correspondence between $1 \leq i \leq N_{M,p}$ and $\boldsymbol{\alpha} \in \mathcal{J}_{M,p}$. We use $i(\boldsymbol{\alpha})$ or $\boldsymbol{\alpha}(i)$ to indicate such a one-to-one mapping whenever necessary.

Given a real separable Hilbert space X, we denote by $L_2(\mathbb{F}; X)$ the Hilbert space of square-integrable $\mathcal{F}$-measurable X-valued random elements f. When $X = \mathbb{R}$, we write $L_2(\mathbb{F})$ instead of $L_2(\mathbb{F}; \mathbb{R})$. Given a collection $\mathcal{R} = \{r_{\boldsymbol{\alpha}},\ \boldsymbol{\alpha} \in \mathcal{J}\}$ of positive real numbers with an upper bound R, i.e., $r_{\boldsymbol{\alpha}} < R$ for all $\boldsymbol{\alpha}$, we define the space $\mathcal{R}L_2(\mathbb{F}; X)$ as the closure of $L_2(\mathbb{F}; X)$ in the norm

$$\|u\|^2_{\mathcal{R}L_2(\mathbb{F};X)} = \sum_{\boldsymbol{\alpha}\in\mathcal{J}} r_{\boldsymbol{\alpha}}\|u_{\boldsymbol{\alpha}}\|^2_X, \tag{23}$$

where $u = \sum_{\boldsymbol{\alpha}\in\mathcal{J}} u_{\boldsymbol{\alpha}}\mathrm{h}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$. The space $\mathcal{R}L_2(\mathbb{F}; X)$ is called a weighted chaos space, it is a natural norm for the stochastic space using the Karhunen-Loéve expansion. In this work, X is chosen as $H^1_0(D)$ for elliptic problems with homogeneous boundary conditions.

3. STOCHASTIC ELLIPTIC MODELS

In this paper, we consider the following two stochastic elliptic models:

$$\text{Model I: } -\nabla\cdot(a(\boldsymbol{x},\omega)\nabla u_{\text{I}}(\boldsymbol{x},\omega)) = f(\boldsymbol{x}), \tag{24a}$$

$$\text{Model II: } -\nabla\cdot\left(\left(a^{-1}\right)^{\diamond(-1)}(\boldsymbol{x},\omega)\diamond\nabla u_{\text{II}}(\boldsymbol{x},\omega)\right) = f(\boldsymbol{x}), \tag{24b}$$

with boundary condition $u(\boldsymbol{x},\omega)=0$ on ∂D, where $a^{-1}(\boldsymbol{x},\omega)\diamond\left(a^{-1}(\boldsymbol{x},\omega)\right)^{\diamond(-1)}=1$. In particular, we assume that the force term $f(\boldsymbol{x})$ is deterministic for simplicity and the random coefficient $a(\boldsymbol{x},\omega)$ takes the following form:

$$a(\boldsymbol{x},\omega) = e^{\diamond(\sigma G(\boldsymbol{x},\omega))} = e^{\sigma G(\boldsymbol{x},\omega)-(1/2)\sigma^2}, \tag{25}$$

where $G(\boldsymbol{x},\omega)$ is a stationary Gaussian random process with zero mean and unit variance, subject to a normalized covariance kernel $K(\boldsymbol{x}_1,\boldsymbol{x}_2)=K(|\boldsymbol{x}_1-\boldsymbol{x}_2|)=\mathbb{E}[G(\boldsymbol{x}_1,\omega)G(\boldsymbol{x}_2,\omega)]$. According to the Mercer theorem [34], $K(\boldsymbol{x}_1,\boldsymbol{x}_2)$ has an expansion as

$$K(\boldsymbol{x}_1,\boldsymbol{x}_2) = \sum_{i=1}^{\infty}\lambda_i\phi_i(\boldsymbol{x}_1)\phi_i(\boldsymbol{x}_2), \tag{26}$$

where $\{\lambda_i,\phi_i(\boldsymbol{x})\}_{i=1}^{\infty}$ are eigenpairs of $K(\boldsymbol{x}_1,\boldsymbol{x}_2)$ satisfying

$$\int_D K(\boldsymbol{x}_1,\boldsymbol{x}_2)\phi_i(\boldsymbol{x}_2)d\boldsymbol{x}_2 = \lambda_i\phi_i(\boldsymbol{x}_1), \quad \int_D \phi_i(\boldsymbol{x})\phi_j(\boldsymbol{x})d\boldsymbol{x}=\delta_{ij}. \tag{27}$$

Then $G(\boldsymbol{x},\omega)$ has the following Karhunen-Loève (K-L) expansion:

$$G(\boldsymbol{x},\omega) = \sum_{i=1}^{\infty}\sqrt{\lambda_i}\phi_i(\boldsymbol{x})\xi_i, \tag{28}$$

where ξ_k are independent Gaussian random variables. Furthermore,

$$\sum_{i=1}^{\infty}\lambda_i\phi_i^2(\boldsymbol{x}) = K(\boldsymbol{x},\boldsymbol{x}) = \mathbb{E}[G^2(\boldsymbol{x},\omega)] = 1, \quad \forall\boldsymbol{x}\in D. \tag{29}$$

Using Eqs. (28), (29), and (7), we can obtain the Wiener chaos expansion of the log-normal random process $a(\boldsymbol{x},\omega)$,

$$a(\boldsymbol{x},\omega) = e^{\sum_{i=1}^{\infty}\sigma\sqrt{\lambda_i}\phi_i(\boldsymbol{x})\xi_i-(\sigma^2/2)\lambda_i\phi_i^2(\boldsymbol{x})} = \sum_{\boldsymbol{\alpha}\in\mathcal{J}}\frac{\Phi^{\boldsymbol{\alpha}}}{\boldsymbol{\alpha}!}\mathrm{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}), \tag{30}$$

where $\Phi(\boldsymbol{x})=\left(\sigma\sqrt{\lambda_1}\phi_1(\boldsymbol{x}),\sigma\sqrt{\lambda_2}\phi_2(\boldsymbol{x}),\ldots\right)$.

From Eq. (14), it can be easily derived that

$$(a^{-1}(\boldsymbol{x},\omega))^{\diamond(-1)} = e^{-\sigma^2}e^{\diamond(\sigma G(\boldsymbol{x},\omega))}. \tag{31}$$

Hence, the difference between Wiener chaos expansions of $\left(a(\boldsymbol{x},\omega)^{-1}\right)^{\diamond(-1)}$ and $a(\boldsymbol{x},\omega)$ is just a scaling factor $e^{-\sigma^2}$.

To make the difference between models I and II clearer, we look at the following two linear systems:

$$\text{I}: \quad \begin{cases} \nabla u_{\text{I}} = a^{-1}*\boldsymbol{F}_{\text{I}}, \\ -\nabla\cdot\boldsymbol{F}_{\text{I}} = f, \end{cases} \qquad \text{II}: \quad \begin{cases} \nabla u_{\text{II}} = a^{-1}\diamond\boldsymbol{F}_{\text{II}}, \\ -\nabla\cdot\boldsymbol{F}_{\text{II}} = f, \end{cases} \tag{32}$$

where $*$ denotes the operation of the regular product. Thus, model II is basically making the gradient "smoother" through the Wick product. Then the equation for $u_{\mathrm{I}} - u_{\mathrm{II}}$ can be obtained as

$$\begin{cases} \nabla(u_{\mathrm{I}} - u_{\mathrm{II}}) = a^{-1} * (\boldsymbol{F}_{\mathrm{I}} - \boldsymbol{F}_{\mathrm{II}}) + a^{-1}(* - \diamond)\boldsymbol{F}_{\mathrm{II}}, \\ -\nabla \cdot (\boldsymbol{F}_{\mathrm{I}} - \boldsymbol{F}_{\mathrm{II}}) = 0, \end{cases} \tag{33}$$

which corresponds to a second-order elliptic equation for $u_{\mathrm{I}} - u_{\mathrm{II}}$ as

$$-\nabla \cdot (a\nabla(u_{\mathrm{I}} - u_{\mathrm{II}})) = -\nabla \cdot \left(a * \left(a^{-1}(* - \diamond)\boldsymbol{F}_{\mathrm{II}}\right)\right). \tag{34}$$

Note that we express explicitly the regular products on the right-hand side since the regular and Wick products do not commute. It is seen that Eq. (34) corresponds to model I while the force term is related to model II through $\boldsymbol{F}_{\mathrm{II}}$.

Theorem 1 ([23]). *Let $F = -\nabla \cdot \left(a * \left(a^{-1}(* - \diamond)\boldsymbol{F}_{\mathrm{II}}\right)\right)$, where $*$ indicates the regular product. Assume that $F \in \mathcal{R}L^2(\mathbb{F}; H^{-1}(D))$, where $D \in \mathbb{R}^d$, $d = 1, 2, 3$. Then there exists a set of weights $\tilde{\mathcal{R}} = \{\tilde{r}_{\boldsymbol{\alpha}}, \boldsymbol{\alpha} \in \mathcal{J}\}$, such that*

$$\|u_{\mathrm{I}} - u_{\mathrm{II}}\|_{\tilde{\mathcal{R}}L^2(\mathbb{F}; H_0^1(D))} = C(l_c)\sigma^2 = \mathcal{O}(\sigma^2), \tag{35}$$

where l_c is the correlation length. Furthermore, $C(l_c) \to 0$ as $l_c \to \infty$.

Remark 1. It can be shown theoretically that for one-dimensional cases $D \in \mathbb{R}^1$, $C(l_c) \to 0$ as $l_c \to 0$. For high-dimensional cases, according to the Landau-Lifshitz-Matheron conjecture [35,36] in the homogenization theory for log-normal random coefficients, when $l_c \to 0$, $C(l_c) \to 1/2$ if $d = 2$, and $C(l_c) \to 1/3$ if $d = 3$.

Remark 2. By noting the Mikulevicius-Rozovskii formula [24],

$$\mathrm{h}_{\boldsymbol{\alpha}}\mathrm{h}_{\boldsymbol{\beta}} = \sum_{n=0}^{\infty} \frac{\mathcal{D}^n \mathrm{h}_{\boldsymbol{\alpha}} \diamond \mathcal{D}^n \mathrm{h}_{\boldsymbol{\beta}}}{n!}, \tag{36}$$

where $\mathcal{D}^n$ denotes the nth-order Malliavin derivative, model I can be approximated arbitrarily well as

$$-\nabla \cdot \left(\sum_{n=0}^{\infty} \frac{\mathcal{D}^n a(\boldsymbol{x}, \omega) \diamond \nabla \mathcal{D}^n u}{n!} \right) = f(\boldsymbol{x}). \tag{37}$$

When $n = 0$, Eq. (37) recovers the Wick-type model:

$$-\nabla \left(a(\boldsymbol{x}, \omega) \diamond \nabla u(\boldsymbol{x}, \omega)\right) = f(\boldsymbol{x}). \tag{38}$$

More discussions about the new Wick-type model given by Eq. (37) can be found in [27].

4. STOCHASTIC GALERKIN METHOD

4.1 Uncertainty Propagators

We now look at the uncertainty propagator of model I. Substituting the Wiener chaos expansion

$$u_{\mathrm{I}}(\boldsymbol{x}, \omega) \approx \sum_{\boldsymbol{\alpha} \in \mathcal{J}_{M,p}} u_{\mathrm{I},\boldsymbol{\alpha}}(\boldsymbol{x})\mathrm{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$$

into Eq. (24a) and implementing Galerkin projection in the probability space, we obtain the uncertainty propagator for model I as

$$-\sum_{\boldsymbol{\alpha} \in \mathcal{J}_{M,p}} \nabla \cdot \left(\mathbb{E}\left[a(\boldsymbol{x}, \omega)\mathrm{H}_{\boldsymbol{\alpha}}\mathrm{H}_{\boldsymbol{\gamma}}\right] \nabla u_{\mathrm{I},\boldsymbol{\alpha}}(\boldsymbol{x})\right) = f(\boldsymbol{x})\delta_{(0),\boldsymbol{\gamma}}, \quad \forall \boldsymbol{\gamma} \in \mathcal{J}_{M,p}. \tag{39}$$

It is seen that all chaos coefficients in Eq. (39) are coupled together, which means that they must be solved together. From the numerical point of view, a proper choice would be iterative methods. Before we look into the numerical algorithms, we now address the properties of the matrix $\mathbb{E}\left[a(\boldsymbol{x}, \omega)\mathrm{H}_{\boldsymbol{\alpha}}\mathrm{H}_{\boldsymbol{\gamma}}\right]$ for any $\boldsymbol{x} \in D$.

Lemma 2. *For any given $\boldsymbol{x} \in D$, the matrix $B_{\mathsf{l},ij}(\boldsymbol{x}) = \mathbb{E}\left[a(\boldsymbol{x},\omega)\mathrm{H}_{\boldsymbol{\alpha}(i)}\mathrm{H}_{\boldsymbol{\gamma}(j)}\right]$ is symmetric and positive definite, where $a(\boldsymbol{x},\omega)$ is a log-normal random process defined in Eq.* (25) *and $\boldsymbol{\alpha},\boldsymbol{\gamma} \in \mathcal{J}_{M,p}$.*

Proof. Apparently, the matrix $B_{\mathsf{l}}(\boldsymbol{x})$ is symmetric for any $\boldsymbol{x} \in D$. For any nonzero vector $\boldsymbol{c} = (c_1, c_2, \ldots, c_{N_{M,p}}) \neq 0$, the following inequality holds for any $\boldsymbol{x} \in D$:

$$\begin{aligned}\boldsymbol{c}^T B_{\mathsf{l}}(\boldsymbol{x})\boldsymbol{c} &= \sum_{i,j=1}^{N_{M,p}} c_i c_j \mathbb{E}\left[e^{\diamond\sigma G(\boldsymbol{x},\omega)}\mathrm{H}_{\boldsymbol{\alpha}(i)}\mathrm{H}_{\boldsymbol{\gamma}(j)}\right] = \mathbb{E}\left[\sum_{i,j}^{N_{M,p}} c_i c_j e^{\diamond\sigma G(\boldsymbol{x},\omega)}\mathrm{H}_{\boldsymbol{\alpha}(i)}\mathrm{H}_{\boldsymbol{\gamma}(j)}\right] \\ &= \mathbb{E}\left[\left(\sum_{i=1}^{N_{M,p}}\left(e^{\diamond\sigma G(\boldsymbol{x},\omega)}\right)^{1/2}\mathrm{H}_{\boldsymbol{\alpha}(i)}c_i\right)^2\right] \geq 0.\end{aligned}$$

In other words, B_{l} is non-negative definite.

We subsequently show that if $\boldsymbol{c}^T B_{\mathsf{l}}(\boldsymbol{x})\boldsymbol{c} = 0$, then $\boldsymbol{c} = 0$. Let $\boldsymbol{b} \in \mathbb{R}^M$. It is easy to generalize Eq. (8) to the high-dimensional case:

$$\mathrm{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}+\boldsymbol{b}) = \prod_{k=1}^{M}\mathrm{H}_{\alpha_k}(\xi_k+b_k) = \prod_{k=1}^{M}\sum_{i=0}^{\alpha_k}\binom{\alpha_k}{i}b_k^{\alpha_k-i}\mathrm{H}_i(\xi_k) = \sum_{\boldsymbol{\beta}\leq\boldsymbol{\alpha}}\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}\boldsymbol{b}^{\boldsymbol{\alpha}-\boldsymbol{\beta}}\mathrm{H}_{\boldsymbol{\beta}}(\boldsymbol{\xi}). \tag{40}$$

Let $\Phi(\boldsymbol{x}) = \Phi_1(\boldsymbol{x}) \oplus \Phi_2(\boldsymbol{x})$, where

$$\Phi_1(\boldsymbol{x}) = (\sigma\sqrt{\lambda_1}\phi_1(\boldsymbol{x}), \cdots, \sigma\sqrt{\lambda_M}\phi_M(\boldsymbol{x})) \text{ and } \Phi_2(\boldsymbol{x}) = (\sigma\sqrt{\lambda_{M+1}}\phi_{M+1}(\boldsymbol{x}), \sigma\sqrt{\lambda_{M+2}}\phi_{M+2}(\boldsymbol{x}), \cdots).$$

Let $\hat{\boldsymbol{\xi}} = (\xi_{M+1}, \xi_{M+2}, \ldots)$. We then have

$$\begin{aligned}\boldsymbol{c}^{\mathsf{T}} B_{\mathsf{l}}(\boldsymbol{x})\boldsymbol{c} &= \mathbb{E}\left[\left(\sum_{i=1}^{N_{M,p}}\left(e^{\diamond\sigma G(\boldsymbol{x},\omega)}\right)^{1/2}\mathrm{H}_{\boldsymbol{\alpha}(i)}c_i\right)^2\right] = \mathbb{E}\left[e^{\Phi_1^{\mathsf{T}}\boldsymbol{\xi}+\Phi_2^{\mathsf{T}}\hat{\boldsymbol{\xi}}-(1/2)\sigma^2}\left(\sum_{i=1}^{N_{M,p}}\mathrm{H}_{\boldsymbol{\alpha}(i)}c_i\right)^2\right] \\ &= \mathbb{E}\left[e^{\Phi_2^{\mathsf{T}}\hat{\boldsymbol{\xi}}-(1/2)\sigma^2}\right]\mathbb{E}\left[e^{\Phi_1^{\mathsf{T}}\boldsymbol{\xi}}\left(\sum_{i=1}^{N_{M,p}}\mathrm{H}_{\boldsymbol{\alpha}(i)}c_i\right)^2\right] \\ &= e^{(1/2)\Phi_2^{\mathsf{T}}\Phi_2-(1/2)\sigma^2}e^{(1/2)\Phi_1^{\mathsf{T}}\Phi_1}\mathbb{E}\left[\left(\sum_{i=1}^{N_{M,p}}\mathrm{H}_{\boldsymbol{\alpha}(i)}(\boldsymbol{\xi}+\Phi_1)c_i\right)^2\right] \\ &= \mathbb{E}\left[\left(\sum_{i=1}^{N_{M,p}}\sum_{\boldsymbol{\beta}\leq\boldsymbol{\alpha}(i)}\binom{\boldsymbol{\alpha}(i)}{\boldsymbol{\beta}}\Phi_1^{\boldsymbol{\alpha}(i)-\boldsymbol{\beta}}\mathrm{H}_{\boldsymbol{\beta}}(\boldsymbol{\xi})c_i\right)^2\right] \\ &= \mathbb{E}\left[\left(\sum_{\boldsymbol{\beta}\in\mathcal{J}_{M,p}}\left(\sum_{\boldsymbol{\alpha}(i)\geq\boldsymbol{\beta}}\binom{\boldsymbol{\alpha}(i)}{\boldsymbol{\beta}}\Phi_1^{\boldsymbol{\alpha}(i)-\boldsymbol{\beta}}c_i\right)\mathrm{H}_{\boldsymbol{\beta}}(\boldsymbol{\xi})\right)^2\right] = \sum_{\boldsymbol{\beta}\in\mathcal{J}_{M,p}}\left(\sum_{\boldsymbol{\alpha}\geq\boldsymbol{\beta}}\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}\Phi_1^{\boldsymbol{\alpha}-\boldsymbol{\beta}}c_{i(\boldsymbol{\alpha})}\right)^2\boldsymbol{\beta}!.\end{aligned}$$

If $\boldsymbol{c}^T B_{\mathsf{l}}(\boldsymbol{x})\boldsymbol{c} = 0$, we have

$$\sum_{\boldsymbol{\alpha}\geq\boldsymbol{\beta}}\binom{\boldsymbol{\alpha}}{\boldsymbol{\beta}}\Phi_1^{\boldsymbol{\alpha}-\boldsymbol{\beta}}(\boldsymbol{x})c_{i(\boldsymbol{\alpha})} = 0, \quad \forall\,\boldsymbol{\beta} \in \mathcal{J}_{M,p},\, \boldsymbol{x} \in D.$$

We note that the matrix in the above linear system is an upper-triangular matrix and the entries on the diagonal line are 1. In other words, the solution of the above linear system is $\boldsymbol{c} = 0$. To this end, we can conclude that the matrix B is symmetric and positive definite. □

Remark 3. In numerical computation, we often take

$$B_{\mathsf{l},ij}(\boldsymbol{x}) = \mathbb{E}\left[e^{\Phi_1(\boldsymbol{x})^T\boldsymbol{\xi}-(1/2)\sigma^2}\mathrm{H}_{\boldsymbol{\alpha}(i)}\mathrm{H}_{\boldsymbol{\gamma}(j)}\right],$$

which is the truncated version of the matrix B_{l} in Lemma 2. From the proof of Lemma 2, such a matrix is also symmetric and positive definite.

Actually $\mathbb{E}\left[a(\boldsymbol{x},\omega)\mathrm{H}_{\boldsymbol{\alpha}}\mathrm{H}_{\boldsymbol{\beta}}\right]$ can be computed exactly as in the following lemma.

Lemma 3. *Let* $a(\boldsymbol{x},\omega) = \exp^{\diamond}(\sigma G(\boldsymbol{x},\omega))$. *We then have*

$$\mathbb{E}\left[a(\boldsymbol{x},\omega)\mathrm{H}_{\boldsymbol{\alpha}}\mathrm{H}_{\boldsymbol{\beta}}\right] = \sum_{\boldsymbol{\kappa}\le\boldsymbol{\alpha}\wedge\boldsymbol{\beta}}\chi(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\kappa})\Phi^{\boldsymbol{\alpha}+\boldsymbol{\beta}-2\boldsymbol{\kappa}}(\boldsymbol{x}), \tag{41}$$

where $(\boldsymbol{\alpha}\wedge\boldsymbol{\beta})_k = \alpha_k \wedge \beta_k$, $k = 1, 2, \ldots$.

Proof. First, Eq. (9) can be generalized straightforwardly to the multidimensional case as

$$\mathrm{H}_{\boldsymbol{\alpha}}\mathrm{H}_{\boldsymbol{\beta}} = \sum_{\boldsymbol{\kappa}\le\boldsymbol{\alpha}\wedge\boldsymbol{\beta}}\chi(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\kappa})\mathrm{H}_{\boldsymbol{\alpha}+\boldsymbol{\beta}-2\boldsymbol{\kappa}}$$

with

$$\chi(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\kappa}) = \frac{\boldsymbol{\alpha}!\boldsymbol{\beta}!}{\boldsymbol{\kappa}!(\boldsymbol{\alpha}-\boldsymbol{\kappa})!(\boldsymbol{\beta}-\boldsymbol{\kappa})!}.$$

Using Eq. (30), we have

$$\begin{aligned}\mathbb{E}\left[a(\boldsymbol{x},\omega)\mathrm{H}_{\boldsymbol{\alpha}}\mathrm{H}_{\boldsymbol{\beta}}\right] &= \sum_{\boldsymbol{\gamma}\in\mathcal{J}}\frac{\Phi^{\boldsymbol{\gamma}}(\boldsymbol{x})}{\boldsymbol{\gamma}!}\mathbb{E}\left[\mathrm{H}_{\boldsymbol{\gamma}}\mathrm{H}_{\boldsymbol{\alpha}}\mathrm{H}_{\boldsymbol{\beta}}\right] = \sum_{\boldsymbol{\gamma}\in\mathcal{J}}\frac{\Phi^{\boldsymbol{\gamma}}}{\boldsymbol{\gamma}!}\sum_{\boldsymbol{\kappa}\le\boldsymbol{\alpha}\wedge\boldsymbol{\beta}}\chi(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\kappa})\mathbb{E}\left[\mathrm{H}_{\boldsymbol{\alpha}+\boldsymbol{\beta}-2\boldsymbol{\kappa}}\mathrm{H}_{\boldsymbol{\gamma}}\right]\\ &= \sum_{\boldsymbol{\kappa}\le\boldsymbol{\alpha}\wedge\boldsymbol{\beta}}\chi(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\kappa})\Phi^{\boldsymbol{\alpha}+\boldsymbol{\beta}-2\boldsymbol{\kappa}}.\end{aligned}$$

□

Remark 4. When $\boldsymbol{\alpha} = \boldsymbol{\beta}$, we have

$$\mathbb{E}\left[a(\boldsymbol{x},\omega)\mathrm{H}_{\boldsymbol{\alpha}}^2\right] = \sum_{\boldsymbol{\kappa}\le\boldsymbol{\alpha}}\chi(\boldsymbol{\alpha},\boldsymbol{\alpha},\boldsymbol{\kappa})\Phi^{2(\boldsymbol{\alpha}-\boldsymbol{\kappa})}(\boldsymbol{x}) \ge \chi(\boldsymbol{\alpha},\boldsymbol{\alpha},\boldsymbol{\alpha}) = \boldsymbol{\alpha}!.$$

Remark 5. Lemma 3 implies that to compute $\mathbb{E}[a(\boldsymbol{x},\omega)\mathrm{H}_{\boldsymbol{\beta}(i)}\mathrm{H}_{\boldsymbol{\gamma}(j)}]$ exactly, we require the coefficients of the Wiener chaos expansion of $a(\boldsymbol{x},\omega)$ up to order $2\boldsymbol{\beta}(N_{M,p})$.

We now look at the uncertainty propagators of model II. Let $\hat{a}(\boldsymbol{x},\omega) = \left(a^{-1}\right)^{\diamond(-1)}$. Using Eqs. (30) and (31), the Wiener chaos expansion of $\hat{a}(\boldsymbol{x},\omega)$ can be explicitly derived as

$$\hat{a}(\boldsymbol{x},\omega) = \sum_{\boldsymbol{\alpha}\in\mathcal{J}}\hat{a}_{\boldsymbol{\alpha}}(\boldsymbol{x})\mathrm{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \sum_{\boldsymbol{\alpha}\in\mathcal{J}}e^{-\sigma^2}\frac{\Phi^{\boldsymbol{\alpha}}}{\boldsymbol{\alpha}!}\mathrm{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}). \tag{42}$$

Following the same procedure for model I, we can obtain the uncertainty propagator of model II as

$$-\sum_{\boldsymbol{\alpha}\le\boldsymbol{\gamma}}\nabla\cdot(\hat{a}_{\boldsymbol{\gamma}-\boldsymbol{\alpha}}(\boldsymbol{x})\nabla u_{\mathsf{II},\boldsymbol{\alpha}}(\boldsymbol{x})) = f(\boldsymbol{x})\delta_{(0),\boldsymbol{\gamma}}, \quad \forall\boldsymbol{\gamma}\in\mathcal{J}_{M,p}. \tag{43}$$

It is seen that $u_{\mathrm{II},\boldsymbol{\gamma}}$ only depends on the chaos coefficients $u_{\mathrm{II},\boldsymbol{\alpha}}$ with $\boldsymbol{\alpha} < \boldsymbol{\gamma}$, which introduces a lower-triangular structure into the matrix $B_{\mathrm{II},ij}(\boldsymbol{x}) = \hat{a}_{\boldsymbol{\gamma}(j)-\boldsymbol{\alpha}(i)}(\boldsymbol{x})$. In other words, the deterministic PDEs for $u_{\mathrm{II},\boldsymbol{\gamma}}$ are naturally decoupled and can be solved one-by-one. Furthermore, Eq. (43) can be rewritten as

$$-\nabla\cdot\left(\hat{a}_{(0)}(\boldsymbol{x})\nabla u_{\mathrm{II},\boldsymbol{\gamma}}(\boldsymbol{x})\right)=\sum_{\boldsymbol{\alpha}<\boldsymbol{\gamma}}\nabla\cdot\left(\hat{a}_{\boldsymbol{\gamma}-\boldsymbol{\alpha}}(\boldsymbol{x})\nabla u_{\mathrm{II},\boldsymbol{\alpha}}(\boldsymbol{x})\right)+f(\boldsymbol{x})\delta_{(0),\boldsymbol{\gamma}}.$$

Thus, if we employ the finite element method to solve the PDE system (43), the bilinear form remains the same for all chaos coefficients $u_{\mathrm{II},\boldsymbol{\gamma}}$, which only depends on $\hat{a}_{(0)}(\boldsymbol{x})$.

4.2 Finite Element Discretization of Uncertainty Propagators

We now look at the finite element discretization of uncertainty propagators of models I and II. Let $\mathscr{T}_h$ be a family of triangulations of D with straight edges and h the maximum size of the elements in $\mathscr{T}_h$. We assume that the family is regular; in other words, the minimal angle of all the elements is bounded from below by a positive constant. We define the finite element space as

$$V_{h,q}^K=\left\{v\,\middle|\,v\circ F_K^{-1}\in\mathscr{P}_q(R)\right\},\quad V_{h,q}=\left\{v\in H_0^1(D)\,\middle|\,v|_K\in V_{h,q}^K,\,K\in\mathscr{T}_h\right\},$$

where F_K is the mapping function for the element K which maps the reference element R (for example, an equilateral triangle or an isosceles right triangle) to the element K and $\mathscr{P}_q(R)$ denotes the set of polynomials of degree at most q on R. We assume that $v|_{\partial D}=0$ for any $v\in V_{h,q}$. Thus, $V_{h,q}$ is an approximation of $H_0^1(D)$ by piecewise polynomial functions. There exist many choices of basis functions on the reference elements, such as h-type finite elements [37], spectral/hp elements [38,39], etc. Let

$$V_{h,q}=\mathrm{span}\{\theta_1(\boldsymbol{x}),\theta_2(\boldsymbol{x}),\ldots,\theta_{N_x}(\boldsymbol{x})\}\subset H_0^1(D),$$

where N_x is the total number of basis functions in the finite element space $V_{h,q}$.

The truncated Wiener chaos space $W_{M,p}$ is defined as

$$W_{M,p}=\left\{\sum_{\boldsymbol{\alpha}\in\mathcal{J}_{M,p}}c_{\boldsymbol{\alpha}}\mathrm{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})\,\middle|\,c_{\boldsymbol{\alpha}}\in\mathbb{R}\right\},\tag{44}$$

The stochastic finite element method for model I can be formulated as follows: Find $u_{\mathrm{I},h}\in V_{h,q}\otimes W_{M,p}$, such that for all $v\in V_{h,q}\otimes W_{M,p}$

$$\mathcal{B}_{\mathrm{I}}(u_{\mathrm{I},h},v)=\mathcal{L}(v),\tag{45}$$

where the bilinear form is

$$\mathcal{B}_{\mathrm{I}}(v_1,v_2)=\int_D\mathbb{E}\left[a(\boldsymbol{x},\omega)\nabla v_1\cdot\nabla v_2\right]d\boldsymbol{x},\tag{46}$$

and the linear form is

$$\mathcal{L}(v)=\int_D\mathbb{E}[fv]d\boldsymbol{x}.\tag{47}$$

Lemma 4. *The stiffness matrix for the stochastic finite element method of model I is symmetric and positive definite.*

Proof. Consider the approximation

$$u_{\mathrm{I},h}(\boldsymbol{x},\boldsymbol{\xi})=\sum_{\boldsymbol{\alpha}\in\mathcal{J}_{M,p}}u_{\mathrm{I},h,\boldsymbol{\alpha}}\mathrm{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})=\sum_{\substack{1\le i\le N_x,\\ \boldsymbol{\alpha}\in\mathcal{J}_{M,p}}}u_{\mathrm{I},h,\boldsymbol{\alpha},i}\theta_i(\boldsymbol{x})\mathrm{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}),\tag{48}$$

where $u_{\mathrm{I},h,\boldsymbol{\alpha},i}\neq 0$ for some i and $\boldsymbol{\alpha}$. We have

$$\begin{aligned}\mathcal{B}_{\mathsf{l}}(u_{\mathsf{l},h},u_{\mathsf{l},h}) &= \sum_{\substack{1\le i\le N_x,\\ \boldsymbol{\alpha}\in\mathcal{J}_{M,p}}}\sum_{\substack{1\le j\le N_x,\\ \boldsymbol{\beta}\in\mathcal{J}_{M,p}}}\int_D u_{\mathsf{l},h,\boldsymbol{\alpha},i}u_{\mathsf{l},h,\boldsymbol{\beta},j}\mathbb{E}\left[a(\boldsymbol{x},\omega)\mathrm{H}_{\boldsymbol{\alpha}}\mathrm{H}_{\boldsymbol{\beta}}\right]\nabla\theta_i(\boldsymbol{x})\cdot\nabla\theta_j(\boldsymbol{x})d\boldsymbol{x}\\ &= \int_D\sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathcal{J}_{M,p}}\mathbb{E}\left[a(\boldsymbol{x},\omega)\mathrm{H}_{\boldsymbol{\alpha}}\mathrm{H}_{\boldsymbol{\beta}}\right]\nabla u_{\mathsf{l},h,\boldsymbol{\alpha}}\cdot\nabla u_{\mathsf{l},h,\boldsymbol{\beta}}d\boldsymbol{x}\\ &= \int_D\left(\sum_{j=1}^{d}\partial_{x_j}\left(\hat{\mathbf{u}}_{\mathsf{l}}(\boldsymbol{x})\right)^{\mathsf{T}}B_{\mathsf{l}}(\boldsymbol{x})\partial_{x_j}\hat{\mathbf{u}}_{\mathsf{l}}(\boldsymbol{x})\right)d\boldsymbol{x},\end{aligned}$$

where the vector $\hat{\mathbf{u}}_{\mathsf{l}}(\boldsymbol{x})$ is defined as $\left(\hat{\mathbf{u}}_{\mathsf{l}}(\boldsymbol{x})\right)_k = u_{\mathsf{l},h,\boldsymbol{\alpha}(k)}(\boldsymbol{x})$, $k=1,\ldots,N_{M,p}$. Due to the homogeneous boundary conditions, a nonzero constant mode does not exist in the space $V_{h,q}$. Using Lemma 2, we know that $\mathcal{B}_{\mathsf{l}}(u_{\mathsf{l},h},u_{\mathsf{l},h})>0$, and the conclusion follows. □

4.3 Structures of Stiffness Matrices of the sFEM

Based on Eq. (48), we define some matrix notations:

$$\mathbf{u}_{\mathsf{l}} = \begin{bmatrix}\mathbf{u}^{\mathsf{l},1}\\ \mathbf{u}^{\mathsf{l},2}\\ \vdots\\ \mathbf{u}^{\mathsf{l},N_{M,p}}\end{bmatrix},\quad \mathbf{u}^{\mathsf{l},i} = \begin{bmatrix}u_{\mathsf{l},h,\boldsymbol{\alpha}(i),1}\\ u_{\mathsf{l},h,\boldsymbol{\alpha}(i),2}\\ \vdots\\ u_{\mathsf{l},h,\boldsymbol{\alpha}(i),N_x}\end{bmatrix},\ i=1,\ldots,N_{M,p}. \tag{49}$$

Obviously, the total number of unknowns is $N_x\times N_{M,p}$. The weak form (45) leads to the linear system $A_{\mathsf{l}}\mathbf{u}_{\mathsf{l}}=\mathbf{f}$ with the block structure

$$A_{\mathsf{l}} = \begin{pmatrix}A_{\mathsf{l},11} & A_{\mathsf{l},12} & \cdots & A_{\mathsf{l},1N_{M,p}}\\ A_{\mathsf{l},21} & A_{\mathsf{l},22} & \cdots & A_{\mathsf{l},2N_{M,p}}\\ \vdots & \vdots & \ddots & \vdots\\ A_{\mathsf{l},N_{M,p}1} & A_{\mathsf{l},N_{M,p}2} & \cdots & A_{\mathsf{l},N_{M,p}N_{M,p}}\end{pmatrix},\quad \mathbf{f} = \begin{pmatrix}\mathbf{f}_1\\ \mathbf{f}_2\\ \vdots\\ \mathbf{f}_{N_{M,p}}\end{pmatrix}. \tag{50}$$

We considering the approximation of $a(\boldsymbol{x},\omega)$ as [see Eq. (30)]

$$a^{M,\hat{p}}(\boldsymbol{x},\boldsymbol{\xi}) = \sum_{\boldsymbol{\alpha}\in\mathcal{J}_{M,\hat{p}}}a_{\boldsymbol{\alpha}}^{M,\hat{p}}(\boldsymbol{x})\mathrm{H}_{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\alpha}\in\mathcal{J}_{M,\hat{p}}}\frac{\Phi^{\boldsymbol{\alpha}}(\boldsymbol{x})}{\boldsymbol{\alpha}!}\mathrm{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}), \tag{51}$$

where $\hat{p}$ is the polynomial order of the Wiener chaos expansion. Then the blocks $A_{\mathsf{l},ij}$ can be expressed as

$$A_{\mathsf{l},ij} = \sum_{\boldsymbol{\alpha}\in\mathcal{J}_{M,\hat{p}}}\mathbb{E}\left[\mathrm{H}_{\boldsymbol{\alpha}}\mathrm{H}_{\boldsymbol{\beta}(i)}\mathrm{H}_{\boldsymbol{\gamma}(j)}\right]S_{\boldsymbol{\alpha}},\quad i,j=1,\ldots,N_{M,p}, \tag{52}$$

where

$$(S_{\boldsymbol{\alpha}})_{ij} = \int_D a_{\boldsymbol{\alpha}}^{M,\hat{p}}(\boldsymbol{x})\nabla\theta_i(\boldsymbol{x})\cdot\nabla\theta_j(\boldsymbol{x})d\boldsymbol{x}. \tag{53}$$

Define matrix $C_{\boldsymbol{\alpha}}$ as

$$(C_{\boldsymbol{\alpha}})_{ij} = \mathbb{E}\left[\mathrm{H}_{\boldsymbol{\alpha}}\mathrm{H}_{\boldsymbol{\beta}(i)}\mathrm{H}_{\boldsymbol{\gamma}(j)}\right]. \tag{54}$$

Then the matrix A_{l} can be rewritten in the tensor-product form as

$$A_{\mathsf{l}} = \sum_{\boldsymbol{\alpha}\in\mathcal{J}_{M,\hat{p}}}C_{\boldsymbol{\alpha}}\otimes S_{\boldsymbol{\alpha}}. \tag{55}$$

Then the matrix-vector multiplication of $A_{\mathrm{I}}\mathbf{u}_{\mathrm{I}}$ can be computed in a relatively efficient way. We rewrite the vector $A_{\mathrm{I}}\mathbf{u}_{\mathrm{I}}$ of length $N_xN_{M,p}$ to an $N_{M,p}$-by-N_x matrix and denote such a matrix as $[A_{\mathrm{I}}\mathbf{u}_{\mathrm{I}}]$. Then we have

$$[A_{\mathrm{I}}\mathbf{u}_{\mathrm{I}}] = \sum_{\boldsymbol{\alpha}\in\mathcal{J}_{M,\hat{p}}} [S_{\boldsymbol{\alpha}}\mathbf{u}^{\mathrm{I},1}\, S_{\boldsymbol{\alpha}}\mathbf{u}^{\mathrm{I},2}\ldots S_{\boldsymbol{\alpha}}\mathbf{u}^{\mathrm{I},N_{M,p}}]C_{\boldsymbol{\alpha}}^{\mathsf{T}}, \tag{56}$$

where $S_{\boldsymbol{\alpha}}\mathbf{u}^{\mathrm{I},i}$ is the ith column vector of an $N_{M,p}$-by-N_x matrix.

4.4 Comments on the Bilinear Form $\mathcal{B}_{\mathrm{I}}$

Using the log-normal random coefficient $a(\boldsymbol{x},\omega)$, we have shown that the bilinear form $\mathcal{B}_{\mathrm{I}}(\cdot,\cdot)$ is positive definite. However, we do not have the ellipticity here because $a(\boldsymbol{x},\omega)$ is not strictly positive. Instead of using the Lax-Milgram lemma, the existence and uniqueness of a solution $u(\boldsymbol{x},\omega)\in L_2(H_0^1(D))$ can be established by the Fernique theorem with appropriate regularity assumptions for the covariance function of the underlying Gaussian field [6]. The key observation is that the random variable $a_{\min}^{-1}(\omega)=\min_{\boldsymbol{x}\in D} a(\boldsymbol{x},\omega)\in L_p(\mathbb{F}), p>0$. From the theoretical point of view, an inf-sup condition can be established for the continuous bilinear form $\mathcal{B}_{\mathrm{I}}(v_1,v_2)$, where $v_1\in L_2(\mathbb{F};H_0^1(D))$ and $v_2\in L_2(\hat{\mathbb{F}}=(\Omega,\mathcal{F},a_{\min}^2(\omega)P(d\omega));H_0^1(D))$ [4,6]. Note here that the measure of the probability space for test functions v_2 is weighted by the random variable $a_{\min}^2(\omega)$. According to theoretical observations, one choice for the test functions can be

$$\left\{\frac{v}{a_{\min}(\omega)} : v\in L_2(\mathbb{F};H_0^1(D))\right\}.$$

However, it is not clear how to deal with $a_{\min}(\omega)$ numerically. For numerical studies of model I with the Galerkin projection, we usually choose test functions from $v_2\in L_2(\mathbb{F};H_0^1(D))$. Since the stiffness matrix A_{I} is symmetric and positive definite, the existence and uniqueness of solution $\mathbf{u}_{\mathrm{I}}$ is guaranteed. No divergence of the solution with respect to $L_2(\mathbb{F};H_0^1(D))$ norm has been observed for such a procedure.

5. NUMERICAL ALGORITHMS

Based on the properties of the Wick product and the assumptions of Theorem 1, we have the following asymptotic results [23] for Eq. (34) satisfied by $u_{\mathrm{I}}-u_{\mathrm{II}}$. With respect to σ, we have the following power series:

$$-\nabla\cdot\left(a*\left(a^{-1}(*-\diamond)\boldsymbol{F}_{\mathrm{II}}\right)\right)=\sigma^2\tilde{f}_2(\boldsymbol{x},\boldsymbol{\xi})+\sigma^3\tilde{f}_3(\boldsymbol{x},\boldsymbol{\xi})+\ldots.$$

Substituting

$$a(\boldsymbol{x},\omega)=a_0(\boldsymbol{x})+\sigma a_1(\boldsymbol{x},\omega)+\sigma^2 a_2(\boldsymbol{x},\omega)+\ldots,$$

and the following ansatz of $u_{\mathrm{I}}-u_{\mathrm{II}}$,

$$u_{\mathrm{I}}-u_{\mathrm{II}}=\tilde{u}_0(\boldsymbol{x})+\sigma\tilde{u}_1(\boldsymbol{x},\boldsymbol{\xi})+\sigma^2\tilde{u}_2(\boldsymbol{x},\boldsymbol{\xi})+\ldots,$$

into Eq. (34) and comparing the coefficients of σ^i, we obtain

$$\begin{aligned}
&-\nabla\cdot(a_0\nabla\tilde{u}_0)=0,\\
&-\nabla\cdot(a_0\nabla\tilde{u}_1)=\nabla\cdot(a_1\nabla\tilde{u}_0),\\
&-\nabla\cdot(a_0\nabla\tilde{u}_2)=\nabla\cdot(a_2\nabla\tilde{u}_0)+\nabla\cdot(a_1\nabla\tilde{u}_1)+\tilde{f}_2(\boldsymbol{x},\boldsymbol{\xi}),\\
&\qquad\qquad\ldots\ldots,
\end{aligned}$$

which results in

$$\tilde{u}_0(\boldsymbol{x})=\tilde{u}_1(\boldsymbol{x},\boldsymbol{\xi})=0,\quad \tilde{u}_i(\boldsymbol{x},\boldsymbol{\xi})\neq 0,\ i=2,3,\ldots.$$

Thus, $u_{\mathrm{I}}-u_{\mathrm{II}}$ has the following power series expansion with respect to σ:

$$u_{\mathrm{I}}-u_{\mathrm{II}}=\sigma^2\tilde{u}_2(\boldsymbol{x},\boldsymbol{\xi})+\sigma^3\tilde{u}_3(\boldsymbol{x},\boldsymbol{\xi})+\ldots, \tag{57}$$

which holds for any $\boldsymbol{x} \in D$. Then both the mean and standard deviation of $u_{\mathrm{I}} - u_{\mathrm{II}}$ are of $\mathcal{O}(\sigma^2)$ if they exist.

When $l_c \to \infty$, the random coefficient becomes

$$a(\boldsymbol{x}, \omega) = e^{\sigma\xi - (1/2)\sigma^2}, \tag{58}$$

where $\xi \sim \mathcal{N}(0,1)$. In other words, the noise is spatially independent. Model II becomes

$$-\nabla \cdot \left((a^{-1})^{\diamond(-1)} \diamond \nabla u\right) = -(a^{-1})^{\diamond(-1)} \diamond \Delta u = f(\boldsymbol{x}), \tag{59}$$

which is equivalent to model I, since

$$-\Delta u = a^{-1} \diamond f(\boldsymbol{x}) = a^{-1} f(\boldsymbol{x}). \tag{60}$$

We now consider a perturbation of the coefficient given in Eq. (58)

$$a(\boldsymbol{x}, \omega) = e^{\sigma(1+\epsilon\,\phi(\boldsymbol{x}))\xi - (1/2)\sigma^2}, \tag{61}$$

where ϵ is a small positive number. When $\epsilon \to 0$, $u_{\mathrm{II}} \to u_{\mathrm{I}}$. We use the random coefficient (61) to mimic the case that $l_c \to \infty$.

Example 1. Consider a one-dimensional exponential covariance kernel on $x \in [0,1]$:

$$K(x_1, x_2) = e^{-(|x_1 - x_2|/l_c)}.$$

Its eigenvalues satisfy

$$w^2 = \frac{2\epsilon - \epsilon^2\lambda_i}{\lambda_i}, \quad (w^2 - \epsilon^2)\tan(w) - 2\epsilon w = 0, \tag{62}$$

where $\epsilon = 1/l_c$. Its eigenfunctions are

$$\phi_i(x) = \frac{w\cos(wx) + \epsilon\sin(wx)}{\sqrt{(1/2)(\epsilon^2 + w^2) + (w^2 - \epsilon^2)(\sin(2w)/4w) + (\epsilon/2)(1 - \cos(2w))}}. \tag{63}$$

It can be shown that as $\epsilon \to 0$, $w \sim \sqrt{2}\epsilon^{1/2}$, which results in that $\lambda_1 = 1 + \mathcal{O}(\epsilon)$ and $\phi_1(x) = 1 + \mathcal{O}(\epsilon)$. Thus it is reasonable to consider a perturbation given in Eq. (61) with $\epsilon = 1/l_c$.

We here use a one-dimensional elliptic problem to examine the random coefficient (61) and present a numerical study of the convergence behavior of $u_{\mathrm{II}} \to u_{\mathrm{I}}$ as $\epsilon \to 0$. In Fig. 1 we plot the relative difference between u_{I} and u_{II} defined as

$$\epsilon_r = \frac{\|u_{\mathrm{I}} - u_{\mathrm{II}}\|_{L_2(\Omega; H_0^1(D))}}{\|u_{\mathrm{I}}\|_{L_2(\Omega; H_0^1(D))}},$$

with respect to σ and ϵ. It is seen that the dominant error takes a form

$$\log(\epsilon_r) = \log(\epsilon) + 2\log(\sigma) + C, \tag{64}$$

i.e.,

$$\epsilon_r \sim C\epsilon\sigma^2, \tag{65}$$

where C is a general constant. This suggests that although model II provides a general second-order approximation of model I, the constant before σ^2 goes to zero linearly with respect to $1/l_c$ as l_c goes to infinity.

To accelerate the numerical algorithms for model I, such as the Monte Carlo method and the Galerkin projection method, we take advantage of the small difference between u_{I} and u_{II} either when σ is relatively small or the correlation length is relatively large such that the constant $C(l_c)$ is close to 0, and the fact that u_{II} can be obtained effectively. Based on this idea, we use the solution u_{II} as a predictor of u_{I}, or the stiffness matrix A_{II} of model II as a preconditioner of A_{I}.

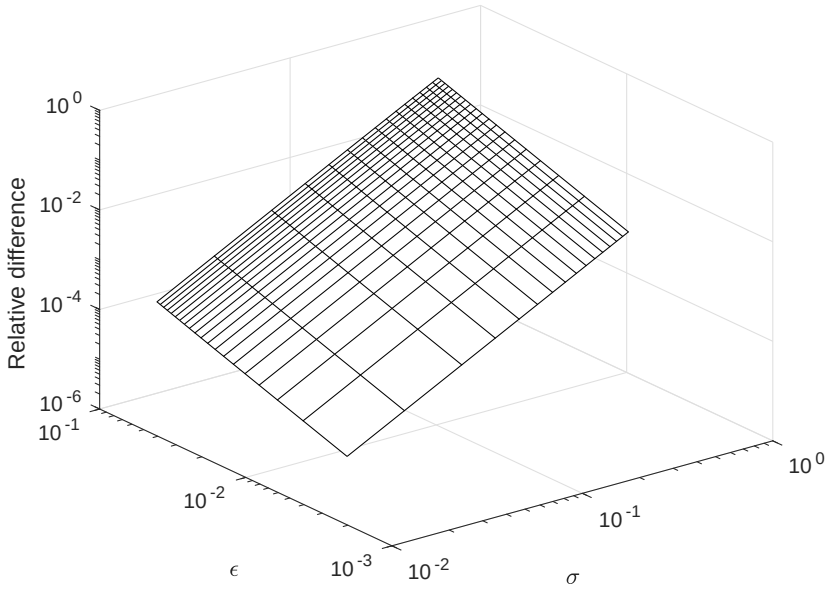

FIG. 1: Relative difference between u_{I} and u_{II} with respect to σ and ϵ for one-dimensional elliptic problem subject to the random coefficient (61)

5.1 Variance Reduction for the Monte Carlo Method

When the correlation length l_c is relatively small, eigenvalues of the covariance kernel decay slowly implying that a relatively large number of Gaussian random variables need to be kept for a good approximation of the log-normal random coefficient. For such a case, the Monte Carlo method can be more efficient than the Wiener chaos expansion. We then propose the following two-step methodology:

(i) **Predictor given by** $u_{\mathrm{II},h}$: We first consider Wiener chaos expansion of model II to obtain the numerical solution $u_{\mathrm{II},h}$. Its mean will be just the zeroth-order coefficient $u_{\mathrm{II},h,(0)}$.

(ii) **A predictor-corrector method**: Using the solution $u_{\mathrm{II},h}$ as a control variate for variance reduction, we further refine the Monte Carlo simulations of $u_{\mathrm{I},h}$ in the following way:

$$\tilde{u}_{\mathrm{I},h}(\boldsymbol{x},\boldsymbol{\xi}) := u_{\mathrm{II},h,(0)}(\boldsymbol{x}) + (u_{\mathrm{I},h}(\boldsymbol{x};\boldsymbol{\xi}) - u_{\mathrm{II},h}(\boldsymbol{x};\boldsymbol{\xi})), \tag{66}$$

$$\mathbb{E}_{\mathrm{IS}}[u_{\mathrm{I},h}](\boldsymbol{x}) := \mathbb{E}_{\mathrm{mc}}[\tilde{u}_{\mathrm{I},h}](\boldsymbol{x}) := \frac{1}{N_{\mathrm{mc}}}\sum_{i=1}^{N_{\mathrm{mc}}} \tilde{u}_{\mathrm{I},h}(\boldsymbol{x};\boldsymbol{\xi}^{(i)}), \tag{67}$$

where N_{mc} indicates the number of samples of $\boldsymbol{\xi}$, and $\boldsymbol{\xi}^{(i)}$ the ith sample.

Based on Eq. (57), we have the following lemma:

Lemma 5. *We have the following error estimate:*

$$\left\|\mathbb{E}_{\mathrm{IS}}[u_{\mathrm{I},h}] - \mathbb{E}[u_{\mathrm{I},h}]\right\|^2_{L_2(\mathbb{F};H_0^1(D))} = \int_D \mathrm{Var}(\mathbb{E}_{\mathrm{IS}}[u_{\mathrm{I},h}])(\boldsymbol{x})d\boldsymbol{x} = \mathcal{O}(\sigma^4)N_{\mathrm{mc}}^{-1}. \tag{68}$$

Proof. Firstly, it is easy to check that $\mathbb{E}\big[\mathbb{E}_{\mathrm{IS}}[u_{\mathrm{I},h}]\big] = \mathbb{E}[u_{\mathrm{I},h}]$, so the first equal sign holds. Secondly,

$$\begin{aligned}\mathrm{Var}(\mathbb{E}_{\mathrm{IS}}[u_{\mathrm{I},h}]) &= N_{\mathrm{mc}}^{-1}\,\mathrm{Var}(\tilde{u}_{\mathrm{I},h}) = N_{\mathrm{mc}}^{-1}\,\mathrm{Var}(u_{\mathrm{I},h}-u_{\mathrm{II},h}) = N_{\mathrm{mc}}^{-1}\big(\mathbb{E}[(u_{\mathrm{I},h}-u_{\mathrm{II},h})^2] - \mathbb{E}^2[u_{\mathrm{I},h}-u_{\mathrm{II},h}]\big)\\ &\le N_{\mathrm{mc}}^{-1}\mathbb{E}[(u_{\mathrm{I},h}-u_{\mathrm{II},h})^2] = N_{\mathrm{mc}}^{-1}\int (u_{\mathrm{I},h}-u_{\mathrm{II},h})^2\rho(\boldsymbol{\xi})d\boldsymbol{\xi} = \mathcal{O}(\sigma^4)N_{\mathrm{mc}}^{-1},\end{aligned}$$

where the last step is obtained using Eq. (57). Then the second equal sign of Eq. (68) is obtained by taking integration of the above equation with respect to spatial variable $\boldsymbol{x}$. □

From Eq. (68), we have

$$\left\|\mathbb{E}_{\text{IS}}[u_{\text{I},h}] - \mathbb{E}[u_{\text{I},h}]\right\|_{L_2(\mathbb{F};H_0^1(D))} = \mathcal{O}(\sigma^2)N_{\text{mc}}^{-1/2}. \tag{69}$$

Since a direct Monte Carlo method to calculate $\mathbb{E}[u_{\text{I},h}]$ has an error $\mathcal{O}(1)N_{\text{mc}}^{-1}$, so the standard deviation reduction is quadratic with respect to σ.

We now look at the computation cost. For the brute-force Monte Carlo method, the cost is $\mathcal{O}((\tau_1+\tau_2)\hat{N}_{\text{mc}})$, where τ_1 is the time for construction of the stiffness matrix and τ_2 the time for solving a linear system. For the proposed strategy, the cost is $\mathcal{O}((\tau_1+\tau_2+\tau_3)N_{\text{mc}}+\tau_4)$, where τ_3 is the time for the evaluation of $u_{\text{II},h}(\boldsymbol{x};\boldsymbol{\xi}^{(i)})$, which is much smaller than $\tau_1+\tau_2$, and τ_4 is the time to obtain $u_{\text{II},h}$. To obtain $u_{\text{II},h}$, only one stiffness matrix is needed. Since the uncertainty propagator is decoupled, $\tau_4 \approx \tau_1 + N_{M,p}\tau_2$. Then the cost for the proposed strategy is about $\mathcal{O}((\tau_1+\tau_2)N_{\text{mc}}+\tau_2 N_{M,p}+\tau_1)$. Thus, if a low-order Wiener chaos solution $u_{\text{II},h}$ serves as an effective control variate, the proposed strategy can be much more efficient than the brute-force Monte Carlo method, since N_{mc} can be much smaller than $\hat{N}_{\text{mc}}$ for the same accuracy.

Remark 6. Consider

$$\tilde{u}_{\text{I},h}(\alpha;\boldsymbol{x},\boldsymbol{\xi}) = u_{\text{I},h}(\boldsymbol{x};\boldsymbol{\xi}) - \alpha(u_{\text{II},h}(\boldsymbol{x};\boldsymbol{\xi}) - u_{\text{II},h,(0)}(\boldsymbol{x})), \tag{70}$$

where α is a real number. It is well known that for all $\alpha \in (-\infty,\infty)$, $\tilde{u}_{\text{I},h}(\alpha)$ provides an unbiased estimator of $\mathbb{E}[u_{\text{I},h}]$ through

$$\mathbb{E}_{\text{mc}}[\tilde{u}_{\text{I},h}] = \frac{1}{N_{\text{mc}}}\sum_{i=1}^{N_{\text{mc}}} \tilde{u}_{\text{I},h}(\alpha;\boldsymbol{x},\boldsymbol{\xi}^{(i)}), \tag{71}$$

which holds for any $\boldsymbol{x} \in D$. For a fixed $\boldsymbol{x} \in D$, we know that if we choose $\alpha^* = \sigma_{\text{I,II}}/\sigma_{\text{I}}^2$ with

$$\sigma_i = \mathbb{E}[(u_{i,h}-\bar{u}_{i,h})^2]^{1/2},\ i = \text{I},\ \text{II} \quad \text{and} \quad \sigma_{\text{I,II}} = \mathbb{E}[(u_{\text{I},h}-\bar{u}_{\text{I},h})(u_{\text{II},h}-\bar{u}_{\text{II},h})],$$

the variance of $\tilde{u}_{\text{I},h}$ is minimized with respect to α such that

$$\text{Var}(\tilde{u}_{\text{I},h})(\alpha^*) = \sigma_{\text{I}}^2(1-\rho_{\text{I,II}})^2,$$

where $\rho_{\text{I,II}} = \sigma_{\text{I,II}}/(\sigma_{\text{I}}\sigma_{\text{II}})$ is the autocorrelation function of $u_{\text{I},h}$ and $u_{\text{II},h}$. Due to the fact given by Eq. (57) and Theorem 1, $\rho_{\text{I,II}} \approx 1$ for small σ or large l_c, when $u_{\text{I},h}$ and $u_{\text{II},h}$ are almost linear corresponding to $\alpha^* \approx 1$ (see more numerical experiments in [23]). This is the reason we choose $\alpha = 1$ in Eq. (67).

5.2 Stochastic Galerkin Projection Method

Due to the large number of unknowns and the strong coupling between the chaos coefficients $u_{\text{I},\boldsymbol{\alpha}}$, iterative numerical methods are more appropriate for solving the linear system given by the finite element discretization of the uncertainty propagator (39) of model I. In other words, an effective preconditioner is required. Consider the linear system

$$A_{\text{I}}\mathbf{u}_{\text{I}} = \mathbf{f}. \tag{72}$$

Algorithm 1: Variance reduction for Monte Carlo simulations

Solve model II to obtain the Wiener chaos expansion of $u_{\text{II},h}(\boldsymbol{x},\boldsymbol{\xi})$.
for $i = 1, 2, \ldots, N_{mc}$ **do**
  Sample model I to obtain $u_{\text{I},h}(\boldsymbol{x},\boldsymbol{\xi}^{(i)})$;
  Sample the solution of model II to obtain $u_{\text{II},h}(\boldsymbol{x},\boldsymbol{\xi}^{(i)})$;
  Update the statistics using an unbiased estimator as Eq. (67).
end

Let $\mathbf{u}_{\text{II}}$ be a vector consisting of unknowns from the discretization of $u_{\text{II},h}$ based on the same basis as that for $u_{\text{I},h}$. Define A_{II} as the stiffness matrix corresponding to the discretization of the uncertainty propagator of model II. Then the stochastic finite element method for model II has the following matrix form:

$$A_{\text{II}}\mathbf{u}_{\text{II}} = \mathbf{f}. \tag{73}$$

Based on the structure of the uncertainty propagator of model II, we know that A_{II} is a block lower-triangular matrix,

$$A_{\text{II}} = \begin{pmatrix} A_{\text{II},11} & 0 & \dots & 0 \\ A_{\text{II},21} & A_{\text{II},22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_{\text{II},N_{M,p}1} & A_{\text{II},N_{M,p}2} & \dots & A_{\text{II},N_{M,p}N_{M,p}} \end{pmatrix}, \tag{74}$$

where the blocks $A_{\text{II},ij}$ are defined as

$$A_{\text{II},ij} = S_{\boldsymbol{\gamma}(i)-\boldsymbol{\alpha}(j)}, \quad i \geq j. \tag{75}$$

with

$$(S_{\boldsymbol{\gamma}(i)-\boldsymbol{\alpha}(j)})_{m,n} = \int_D \hat{a}_{\boldsymbol{\gamma}(i)-\boldsymbol{\alpha}(j)}(\boldsymbol{x})\nabla\theta_m(\boldsymbol{x}) \cdot \nabla\theta_n(\boldsymbol{x})d\boldsymbol{x}. \tag{76}$$

Note that

$$A_{\text{II},11} = A_{\text{II},22} = \ldots = A_{\text{II},N_{M,p},N_{M,p}} = S_{(0)}. \tag{77}$$

Lemma 6. *Consider the stiffness matrices A_{I} and A_{II}. We have that the condition number*

$$\kappa\left(A_{\text{II}}^{-1}A_{\text{I}}\right) \leq 1 + \mathcal{O}(\sigma^2). \tag{78}$$

Proof. Since the difference between u_{I} and u_{II} is of $\mathcal{O}(\sigma^2)$, we have in the matrix form

$$\|\mathbf{u}_{\text{I}} - \mathbf{u}_{\text{II}}\| = \|A_{\text{I}}^{-1}\mathbf{f} - A_{\text{II}}^{-1}\mathbf{f}\| = \mathcal{O}(\sigma^2), \tag{79}$$

which holds for any $\mathbf{f}$. Hence

$$\|A_{\text{I}}^{-1} - A_{\text{II}}^{-1}\| = \mathcal{O}(\sigma^2). \tag{80}$$

Then the condition number of $A_{\text{II}}^{-1}A_{\text{I}}$ is

$$\begin{aligned} \kappa &= \|A_{\text{II}}^{-1}A_{\text{I}}\|\|A_{\text{I}}^{-1}A_{\text{II}}\| = \|\left(A_{\text{II}}^{-1} - A_{\text{I}}^{-1} + A_{\text{I}}^{-1}\right)A_{\text{I}}\|\|\left(A_{\text{I}}^{-1} - A_{\text{II}}^{-1} + A_{\text{II}}^{-1}\right)A_{\text{II}}\| \\ &= \|I + \left(A_{\text{II}}^{-1} - A_{\text{I}}^{-1}\right)A_{\text{I}}\|\|I + \left(A_{\text{I}}^{-1} - A_{\text{II}}^{-1}\right)A_{\text{II}}\| \leq 1 + \|A_{\text{I}}\|\|A_{\text{II}}\|\left(\mathcal{O}(\sigma^2) + \mathcal{O}(\sigma^4)\right). \end{aligned} \tag{81}$$

□

Remark 7. When σ is relatively small, we expect that A_{II} can provide a good preconditioner for linear system (73). Instead of solving Eq. (73), we can solve

$$A_{\text{II}}^{-1}A_{\text{I}}\mathbf{u}_{\text{I}} = A_{\text{II}}^{-1}\mathbf{f}. \tag{82}$$

5.2.1 Preconditioned Richardson's Iteration

One commonly used iterative method for the uncertainty propagator (39) of model I is the block Gauss-Seidel method, which can be expressed as

$$\begin{aligned} -\nabla\cdot\left(\mathbb{E}\left[a(\boldsymbol{x},\omega)\mathrm{H}_{\boldsymbol{\gamma}}^2\right]\nabla u_{\boldsymbol{\gamma}}^{\text{I},n+1}(\boldsymbol{x})\right) &= \sum_{i=1}^{k(\boldsymbol{\gamma})-1}\nabla\cdot\left(\mathbb{E}\left[a(\boldsymbol{x},\omega)\mathrm{H}_{\boldsymbol{\alpha}(i)}\mathrm{H}_{\boldsymbol{\gamma}}\right]\nabla u_{\boldsymbol{\alpha}(i)}^{\text{I},n+1}(\boldsymbol{x})\right) \\ &+ \sum_{i=k(\boldsymbol{\gamma})+1}^{N_{M,p}}\nabla\cdot\left(\mathbb{E}\left[a(\boldsymbol{x},\omega)\mathrm{H}_{\boldsymbol{\alpha}(i)}\mathrm{H}_{\boldsymbol{\gamma}}\right]\nabla u_{\boldsymbol{\alpha}(i)}^{\text{I},n}(\boldsymbol{x})\right) + f(\boldsymbol{x})\delta_{(0),\boldsymbol{\gamma}}, \quad \forall\,\boldsymbol{\gamma}\in\mathcal{J}_{M,p}, \end{aligned} \tag{83}$$

where the superscript n indicates the iteration step. It is shown in Lemma 3 that $\mathbb{E}\left[a(\boldsymbol{x},\omega)\mathrm{H}_{\boldsymbol{\gamma}}^2\right]$ is strictly positive. We know that the block Gauss-Seidel method corresponds to a fixed point iteration on a preconditioned system,

$$M^{-1}A_{\mathrm{I}}\mathbf{u}_{\mathrm{I}} = M^{-1}\mathbf{f},$$

where M is the lower-triangular part of matrix A_{I}. Based on the comparability of models I and II, we can construct the following preconditioned Richardson's iterative method [40]:

$$\mathbf{u}_{\mathrm{I}}^{(k+1)} = \mathbf{u}_{\mathrm{I}}^{(k)} + \gamma A_{\mathrm{II}}^{-1}(A_{\mathrm{I}}\mathbf{u}_{\mathrm{I}}^{(k)} - \mathbf{f}), \tag{84}$$

where γ is the non-negative acceleration parameter. We know that the Richardson's iterative method converges when $\gamma < 2/\rho(A_{\mathrm{II}}^{-1}A_{\mathrm{I}})$, where $\rho(\cdot)$ indicates the spectral radius of a matrix. Based on the relation between A_{I} and A_{II}, we expect that $\rho((A_{\mathrm{II}})^{-1}A_{\mathrm{I}})$ is close to 1 when σ is relatively small.

5.2.2 Preconditioned GMRES Method

We also consider Krylov subspace methods. Since A_{I} is symmetric and positive definite, a common choice to solve the linear system is the preconditioned conjugate gradient (CG) method. We here consider to use A_{II} as a preconditioner, which is not symmetric. Hence we use a preconditioned GMRES method [40] instead of the CG method.

6. NUMERICAL RESULTS

We consider both one-dimensional and two-dimensional ($D = [-1,1]^d, d = 1,2$) elliptic problem with random coefficient subject to a nonzero force term

$$f(\boldsymbol{x}) = \prod_{i=1}^{d}(x_i^2 + 4x_i + 1)e^{x_i}, \tag{85}$$

and homogeneous boundary conditions. Assume the underlying Gaussian random field of the log-normal coefficient $a(\boldsymbol{x},\omega) = e^{\sigma G(\boldsymbol{x},\omega) - \frac{1}{2}\sigma^2}$, with G's correlation function is given by

$$K(\boldsymbol{x}_1,\boldsymbol{x}_2) = e^{-|\boldsymbol{x}_1-\boldsymbol{x}_2|^2/2l_c^2}, \tag{86}$$

or

$$K(\boldsymbol{x}_1,\boldsymbol{x}_2) = e^{-|\boldsymbol{x}_1-\boldsymbol{x}_2|/l_c}, \tag{87}$$

where l_c is the correlation length and σ the standard deviation. Due to the analyticity of the Gaussian kernel, the eigenvalues decay exponentially [9]. The decay rate is determined by the value of the correlation length, where a larger l_c corresponds to a faster decay rate. The physical discretization is given by 25 uniform finite elements with order $q = 4$ for the one-dimensional case, and 32×32 uniform quadratic finite elements for the two-dimensional cases. We test the parameters $\sigma = 0.2, 0.6, 1$ and $l_c = 20, 2, 0.2$. The solution differences of model I and model II are similar to the results in [23,27], so we only sketch the results for the two-dimensional case here.

The results for the two-dimensional case with Gaussian type kernel are given in Figs. 2–4 for $l_c = 20, 2, 0.2$, respectively. The results for the two-dimensional exponential kernel with $l_c = 20, 2, 0.2$ are given in Figs. 5–7, respectively. The truncation errors of the K-L expansion for the Gaussian kernel and exponential kernel are set to be 2×10^{-3} and 3×10^{-2}, respectively. For model I, if the dimension of the stochastic space M is less than 20, we use the stochastic Galerkin method, otherwise we use the Monte Carlo method. From these figures, we say for small σ values, the results of model II agree very well with the results of model I. A larger correlation length ℓ_c also makes a better agreement between the results of models I and II. This is consistent with the theoretical results.

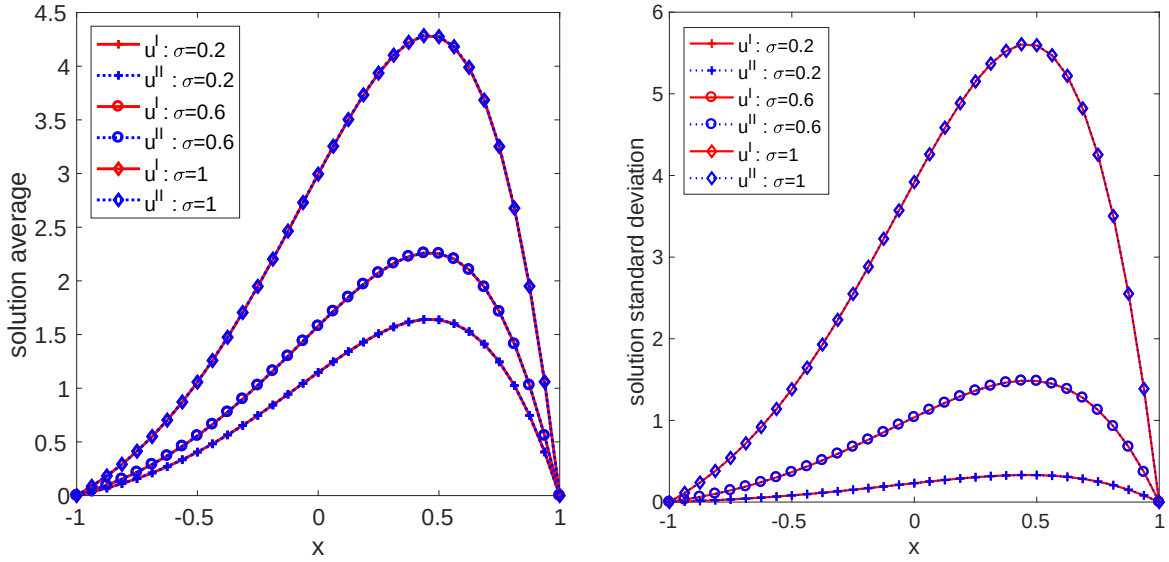

FIG. 2: The average (left) and standard deviation (right) of models I and II at the horizontal line $y = 0$: Gaussian kernel with $\ell_c = 20$, $M = 1$, and $p = 16$ is used for the stochastic Galerkin approximation of both models I and II

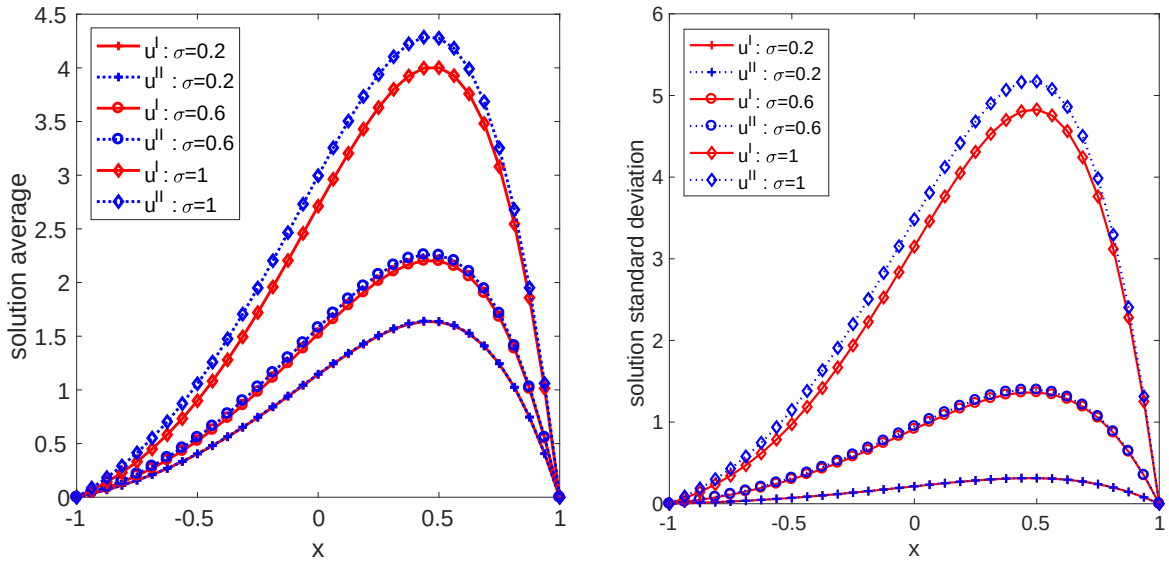

FIG. 3: The average (left) and standard deviation (right) of models I and II at the horizontal line $y = 0$: Gaussian kernel with $\ell_c = 2$, $M = 6$, and $p = 6$ is used for the stochastic Galerkin approximation of both models I and II

6.1 Using $u_{\mathrm{II},h}$ as a Control Variate

When the correlation length is relatively small, a large number of random variables are required to represent the random coefficient and the Monte Carlo method would be a better choice for computation. The mean and variance are given by the following unbiased estimators, respectively:

$$\bar{u}_{\mathrm{I},h} = \frac{1}{N_{\mathrm{mc}}} \sum_{i=1}^{N_{\mathrm{mc}}} u_{\mathrm{I},h}(\boldsymbol{x}, \boldsymbol{\xi}^{(i)}),$$

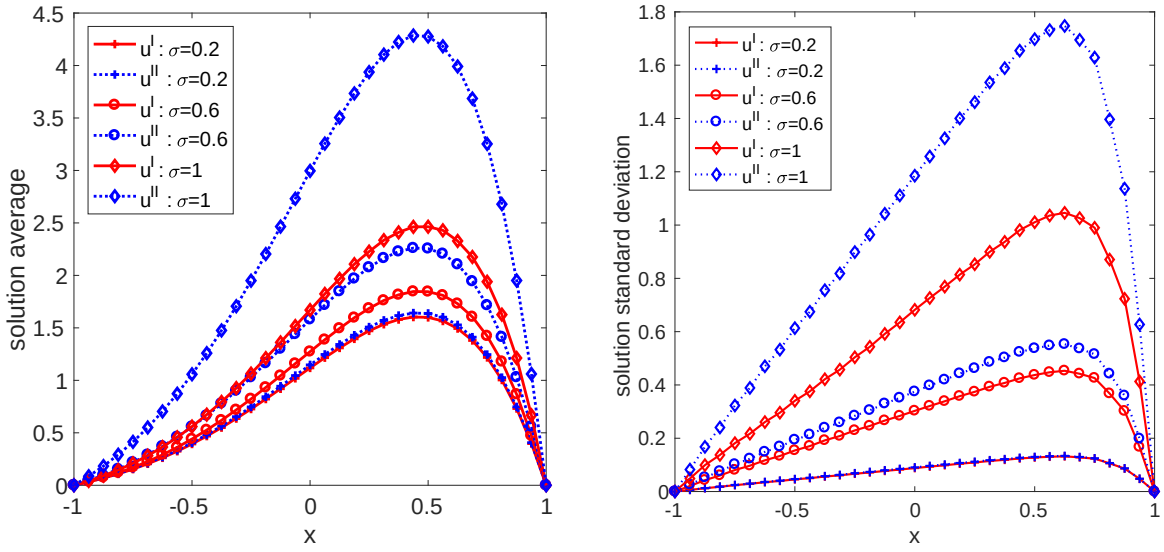

FIG. 4: The average (left) and standard deviation (right) of models I and II at the horizontal line $y = 0$: Gaussian kernel with $\ell_c = 0.2$, $M = 94$, and $p = 1$ is used for the stochastic Galerkin approximation of model II. $M = 94$ and $N_{\text{mc}} = 10{,}000$ are used for the Monte Carlo method of model I

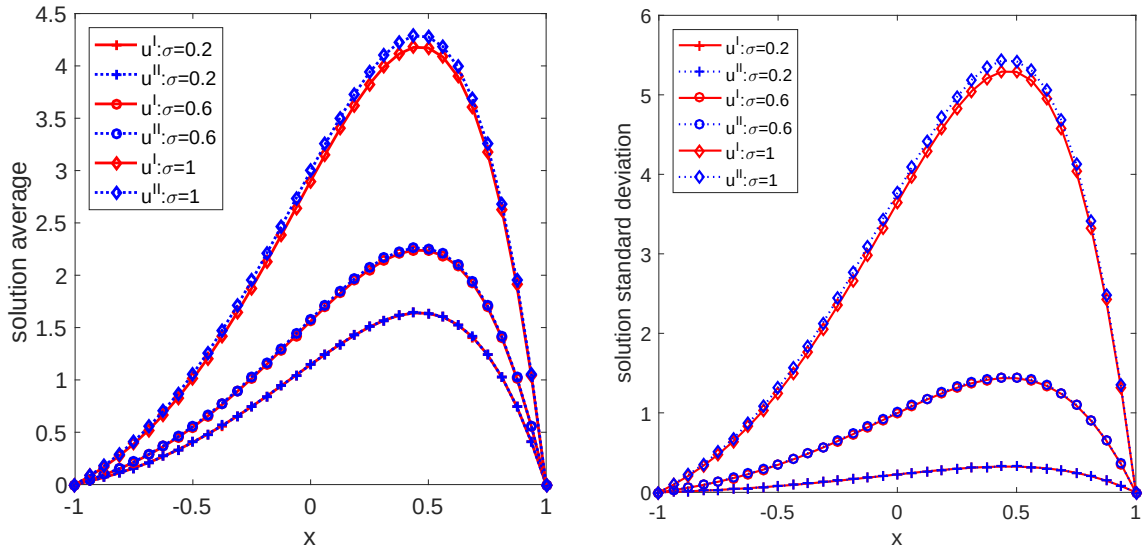

FIG. 5: The average (left) and standard deviation (right) of models I and II at the horizontal line $y = 0$: exponential kernel with $\ell_c = 20$, $M = 3$, and $p = 8$ is used for the stochastic Galerkin approximation of both models I and II

$$\text{Var}(u_{\text{l},h}) \approx \frac{1}{N_{\text{mc}} - 1} \sum_{i=1}^{N_{\text{mc}}} (u_{\text{l},h}(\boldsymbol{x}, \boldsymbol{\xi}^{(i)}) - \bar{u}_{\text{l},h}(\boldsymbol{x}))^2.$$

The average and standard deviations of Monte Carlo solutions at line $y = 0$ for model I with and without using model II as a control variate are given in Fig. 8 (exponential kernel in 1D), and Fig. 9 (exponential kernel in 2D). The results for a Gaussian kernel are similar but easier to obtain. It is seen that variance reduction is achieved for all σ, but for a small σ value, the reduction is significant. To numerically verify how the variance reduction is related to σ and l_c, we solved the two models with different parameters: $l_c = 0.2, 0.4, 0.6, 0.8, 1.0, 1.2$ and $l_c = 8, 4, 2, 1, 0.5, 0.25$.

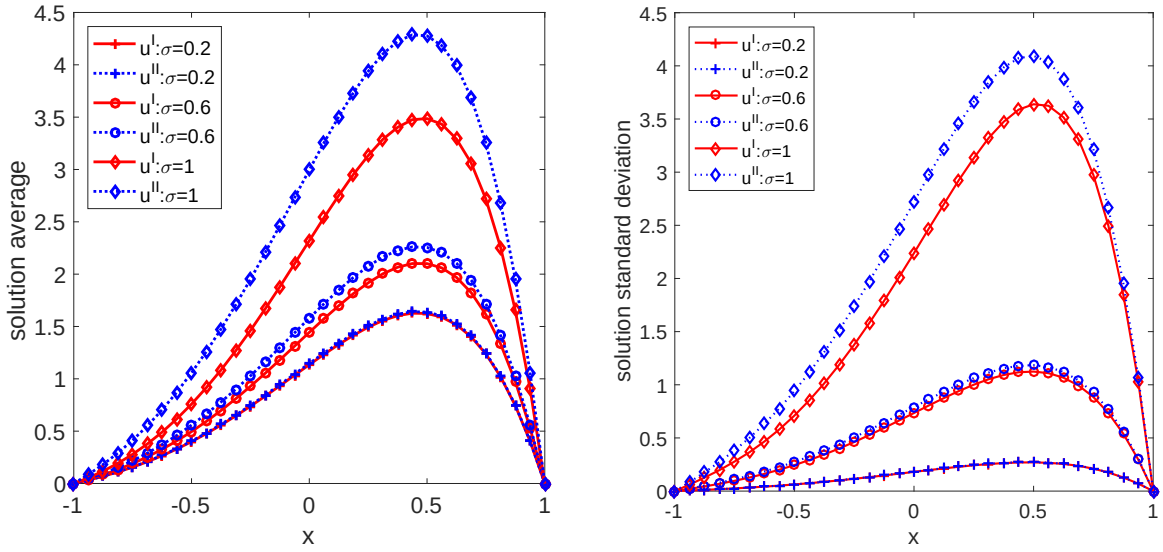

FIG. 6: The average (left) and standard deviation (right) of models I and II at the horizontal line $y = 0$: exponential kernel with $\ell_c = 2$, $M = 28$, and $p = 2$ is used for the stochastic Galerkin approximation of model II. $M = 28$ and $N_{\text{mc}} = 10{,}000$ are used for the Monte Carlo method of model I

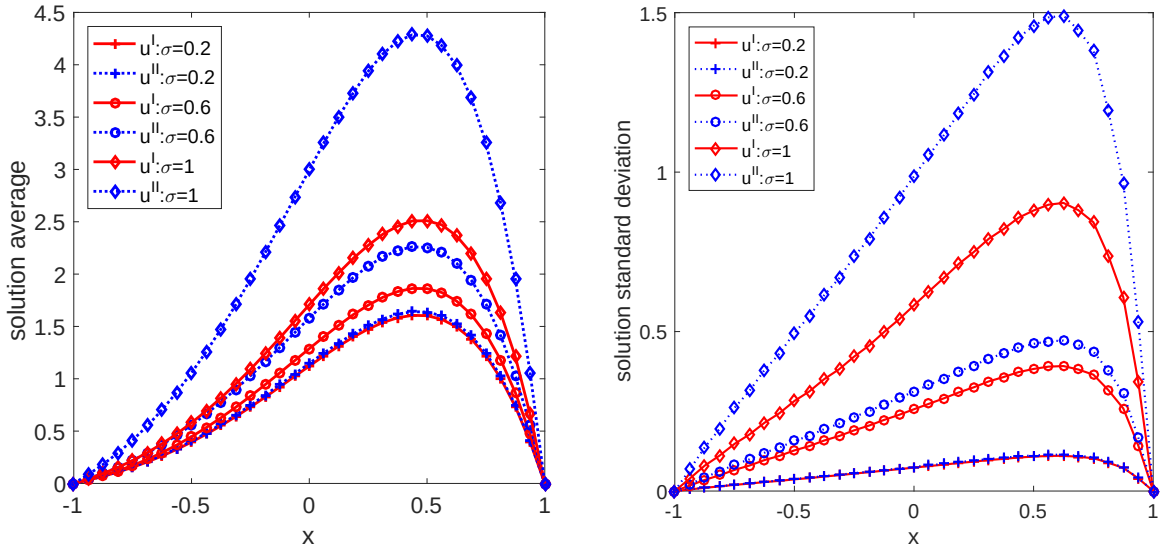

FIG. 7: The average (left) and standard deviation (right) of models I and II at the horizontal line $y = 0$: exponential kernel with $\ell_c = 0.2$, $M = 86$ and $p = 1$ is used for the stochastic Galerkin approximation of model II. $M = 86$ and $N_{\text{mc}} = 10{,}000$ are used for the Monte Carlo method of model I

The corresponding results for one-dimensional and two-dimensional cases with exponential kernel are given in Figs. 10 and 11, respectively. The standard deviation reduction (69) derived from Lemma 5 is clearly verified.

6.2 Using A_{II} as a Preconditioner

The results of using model II to precondition model I is given in Tables 1 and 2 (for 1D cases) and Tables 3 and 4 (for 2D cases). We set the default relaxation parameter in the Richardson iteration to $\gamma = 1/(1 + 3\sigma^2)$.

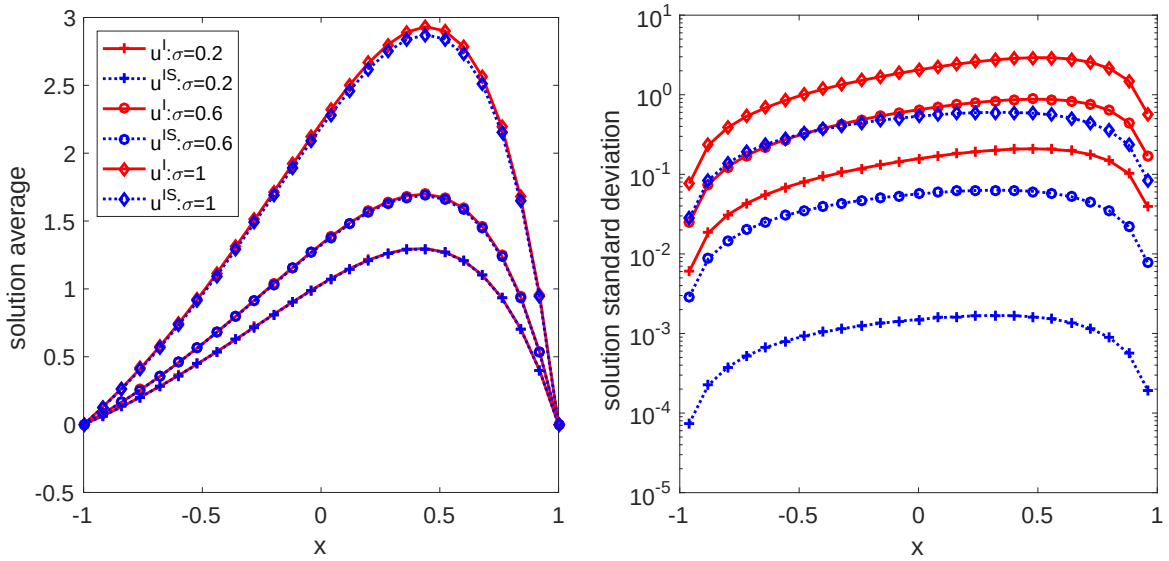

FIG. 8: The mean and standard deviation of the Monte Carlo method for model I with and without important sampling in one-dimensional case. The exponential kernel with correlation length $l_c = 1$ is used. $M = 12, p = 4$ for the stochastic Galerkin approximation of model II. $M = 12, N_{\text{mc}} = 10{,}000$ for the Monte Carlo method. Note that log scale is used for the standard deviation

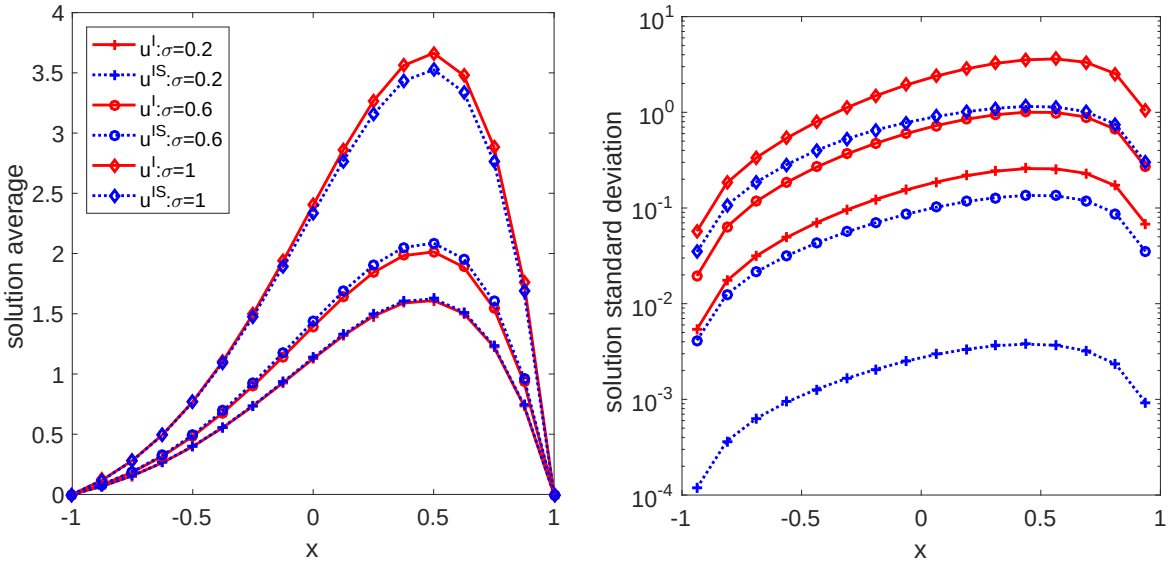

FIG. 9: The mean and standard deviation of the Monte Carlo method for model I with and without important sampling in the two-dimensional case. The exponential kernel with correlation length $l_c = 2$ is used. $M = 19, p = 2$ for the stochastic Galerkin approximation of model II. $M = 19, N_{\text{mc}} = 1000$ for the Monte Carlo method. Note that log scale is used for the standard deviation

For almost all the cases, the preconditioned Richardson iteration and GMRES are both better than the commonly used Gauss-Seidel iteration, especially for large l_c or small σ. The iteration numbers of the Richardson method and GMRES are much smaller than the Gauss-Seidel method; meanwhile their increases with respect to the standard deviation parameter σ are also slower, except for the cases with $p = 1$. For large variance, the preconditioned GMRES

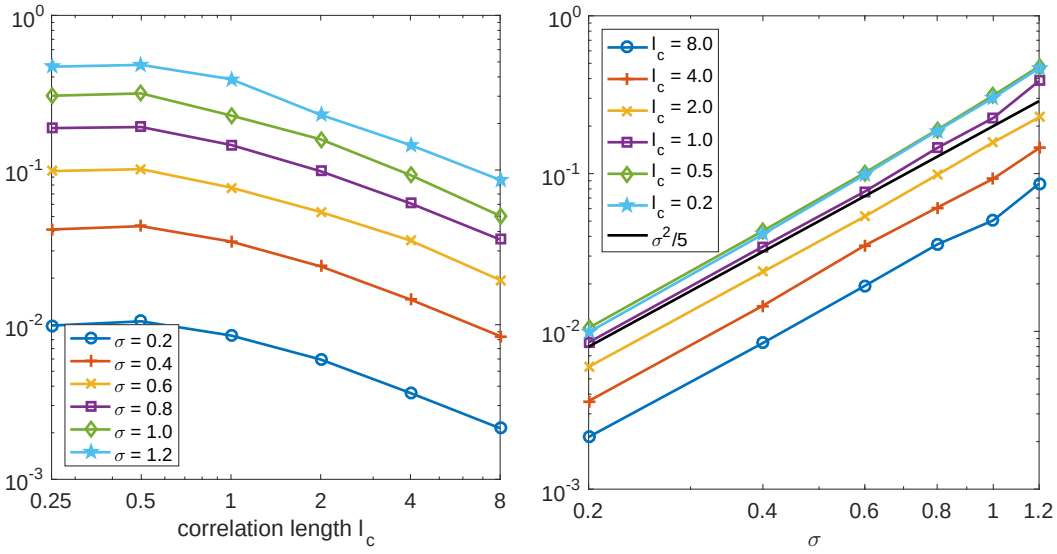

FIG. 10: The variance reduction for the one-dimensional case with an exponential kernel having different correlation lengths and different values of σ . The y axes are $\|\operatorname{Var}(\tilde{u}_{I,h})\|_{H_0^1(D)}/\|\operatorname{Var}(u_{I,h})\|_{H_0^1(D)}$. $N_{\text{mc}} = 10{,}000$ samples are used for the Monte Carlo method. The tolerance of the K-L expansion is set to 3×10^{-2}. The values of M, p corresponding to the stochastic Galerkin approximation of model II with $l_c = 8, 4, 2, 1, 0.5, 0.25$ are $(3, 6), (4, 6), (7, 5), (12, 4), (19, 3), (27, 3)$, respectively. Note that log scales are used for both x and y axes

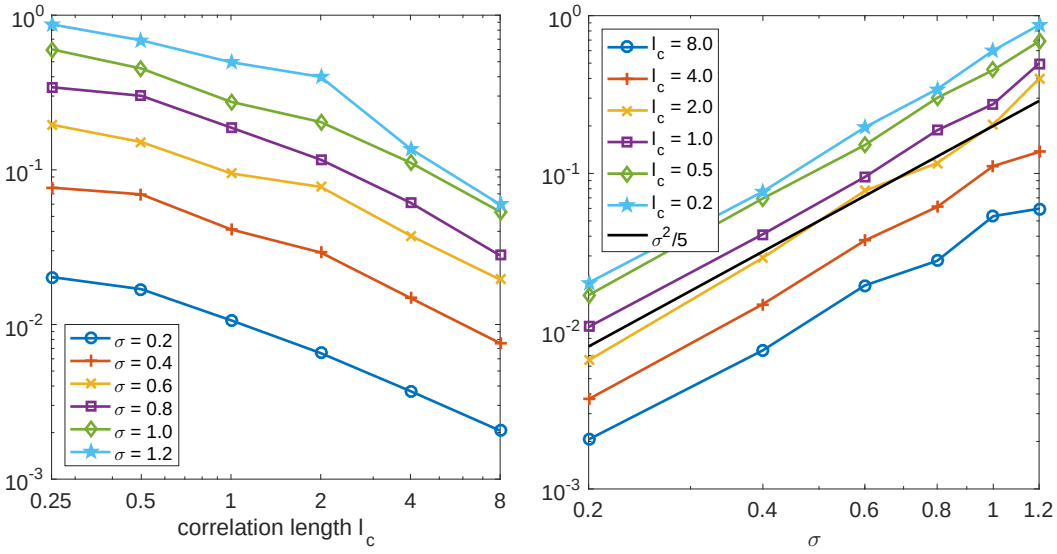

FIG. 11: The variance reduction for the two-dimensional case with an exponential kernel having different correlation lengths and different values of σ. The y axes are $\|\operatorname{Var}(\tilde{u}_{I,h})\|_{H_0^1(D)}/\|\operatorname{Var}(u_{I,h})\|_{H_0^1(D)}$. $N_{\text{mc}} = 1000$ samples are used for the Monte Carlo method. The tolerance of the K-L expansion is set to 3×10^{-2}. The values of M, p corresponding to the stochastic Galerkin approximation of model II with $l_c = 8, 4, 2, 1, 0.5, 0.25$ are $(5, 5), (11, 4), (19, 3), (28, 3), (35, 2), (40, 2)$, respectively. Note that log scales are used for both x and y axes

method behaves much better than Gauss-Seidel and Richardson methods. Note that we use the solution of model II as initial values for the Richardson and GMRES iterations, so in the cases that model II is a very good approximation of model I, the corresponding iteration numbers are 0.

TABLE 1: Preconditioning results of one-dimensional problem with a Gaussian kernel. n_{GS}, n_γ, n_{GMRES} mean the iteration number of Gauss-Seidel, Richardson, and GMRES, respectively. We take $\gamma = 1/(1 + 3\sigma^2)$ for the Richardson method. The tolerance of K-L expansion is set to 2×10^{-3}. The relative tolerance for the iteration solvers is set to 10^{-3}

l_c	σ	M	p	$N_{M,p}$	n_{GS}	n_γ	n_{GMRES}
20	0.2	1	10	11	3	0	0
20	0.6	1	10	11	27	0	0
20	1	1	10	11	> 100	22	5
2	0.2	3	10	286	3	1	1
2	0.6	3	10	286	22	3	1
2	1	3	10	286	> 100	19	9
0.2	0.2	11	3	364	3	1	1
0.2	0.6	11	3	364	10	5	5
0.2	1	11	3	364	29	12	9

TABLE 2: Preconditioning results of one-dimensional problem with an exponential kernel. n_{GS}, n_γ, n_{GMRES} mean the iteration number of Gauss-Seidel, Richardson, and GMRES, respectively. We take $\gamma = 1/(1 + 3\sigma^2)$ for the Richardson method. The tolerance of K-L expansion is set to 3×10^{-2}. The relative tolerance for the iteration solvers is set to 10^{-3}

l_c	σ	M	p	$N_{M,p}$	n_{GS}	n_γ	n_{GMRES}
20	0.2	2	10	66	3	0	0
20	0.6	2	10	66	24	2	1
20	1	2	10	66	> 100	16	9
2	0.2	8	5	1287	3	1	1
2	0.6	8	5	1287	17	4	3
2	1	8	5	1287	> 100	9	9
0.2	0.2	51	2	1378	3	1	1
0.2	0.6	51	2	1378	7	5	3
0.2	1	51	2	1378	15	7	6

According to our understanding of u_{II}, the worst scenario for the proposed preconditioners is when l_c is small and σ is large. In a very few cases (e.g., $l_c = 0.2$ and $\sigma = 0.6, 1$ in Tables 3 and 4), the preconditioned Richardson iteration requires more iterations to converge than Gauss-Seidel. This is probably because a first-order Wiener chaos approximation is used; the big approximation error together with the big modeling error deteriorate the performance of the preconditioning and the parameter ω in the Richardson method is not optimal.

Based on the above observations, we advocate to use GMRES with model II as a preconditioner for solving model I.

In the end, we compare our approach with some existing methods by solving a test problem studied in [17]. The physical domain is set to $[0, 1]^2$, and the force term $f(\boldsymbol{x}) = 1$. The underlying Gaussian field of the log-normal coefficient $a(\boldsymbol{x}, \omega)$ has a correlation function $K(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sigma^2 r K_1(r)$, where $r = \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_2$ and K_1 is the modified Bessel function of the second kind with order one. Set $M = 5$ in the K-L expansion, such that 97% of the Gaussian field's total variance is captured. The iteration numbers of the Richardson and GMRES methods for the stochastic Galerkin method of model I with model II as preconditioner for different σ and p are given in Table 5. From the table,

TABLE 3: Preconditioning results of two-dimensional problem with Gaussian kernel. n_{GS}, n_γ, n_{GMRES} mean the iteration number of Gauss-Seidel, Richardson, and GMRES, respectively. We take $\gamma = 1/(1+3\sigma^2)$ for the Richardson method. The tolerance of K-L expansion is set to 10^{-2}. The relative tolerance for the iteration solvers is set to 10^{-3}

l_c	σ	M	p	$N_{M,p}$	n_{GS}	n_γ	n_{GMRES}
20	0.2	1	16	17	3	0	0
20	0.6	1	16	17	25	0	0
20	1	1	16	17	29	1	1
2	0.2	4	5	126	3	0	0
2	0.6	4	5	126	17	5	4
2	1	4	5	126	48	14	7
0.2	0.2	80	1	81	2	1	1
0.2	0.6	80	1	81	3	4	2
0.2	1	80	1	81	4	7	3

TABLE 4: Preconditioning results of two-dimensional problem with exponential kernel. n_{GS}, n_γ, n_{GMRES} mean the iteration number of Gauss-Seidel, Richardson, and GMRES, respectively. We take $\gamma = 1/(1+3\sigma^2)$ for the Richardson method. The tolerance of K-L expansion is set to 3×10^{-2}. The relative tolerance for the iteration solvers is set to 10^{-3}

l_c	σ	M	p	$N_{M,p}$	n_{GS}	n_γ	n_{GMRES}
20	0.2	3	8	165	3	0	0
20	0.6	3	8	165	12	1	1
20	1	3	8	165	41	14	10
2	0.2	28	2	435	3	1	1
2	0.6	28	2	435	4	3	3
2	1	28	2	435	10	9	4
0.2	0.2	86	1	87	2	1	1
0.2	0.6	86	1	87	2	3	2
0.2	1	86	1	87	4	7	3

TABLE 5: The iteration numbers of Richardson and GMRES method solving the two-dimensional problem with Matern-tye kernel studied in [17]. We take $\gamma = 1/(1+3\sigma^2)$ for the Richardson method. The relative tolerance for the iteration solvers is set to 10^{-8}. $M = 5$

	Richardson					**GMRES**				
σ	p **= 1**	p **= 2**	p **= 3**	p **= 4**	p **= 5**	p **= 1**	p **= 2**	p **= 3**	p **= 4**	p **= 5**
0.2	5	6	5	6	6	3	3	4	4	4
0.4	10	10	11	10	10	3	4	5	6	7
0.6	14	16	17	18	19	4	5	6	7	8
0.8	16	19	21	23	25	5	6	7	8	9
1.0	16	19	21	24	26	5	7	8	9	11

we see that both the Richardson and GMRES methods are efficient. As p increases, the iteration numbers increase slowly. As σ increases, the iteration numbers also increase slowly. The preconditioning effects are still very good for the cases with $\sigma = 1$. These results are very competitive comparing to the algebraic preconditioners studied in [17] for this test example.

7. SUMMARY

In this work, we consider the Wick approximation of two stochastic elliptic problems with log-normal random coefficients, where model II is a second-order approximation of model I with respect to σ. Model II can be used as a precondition for model I in a stochastic Galerkin method. The numerical results show that the preconditioned Richardson iteration is better than the commonly used Gauss-Seidel method when σ is small or l_c is large. Meanwhile, the former method has a parameter to tune. The preconditioned GMRES method works very well for all the values of σ and l_c tested using default parameters. Model II can also be used as an efficient important sampling process for model I to reduce the variance of a Monte Carlo approach when the stochastic dimension in a Karhunen-Loève expansion is very high.

ACKNOWLEDGMENTS

The work of X. Wan was partially supported by NSF grant DMS-1620026. The work of H. Yu was partially supported by China National Program on Key Basic Research Project 2015CB856003, NNSFC Grant 11771439, and China Science Challenge Project TZ2018001.

REFERENCES

1. Lototsky, S., Rozovskii, B., and Wan, X., Elliptic Equations of Higher Stochastic Order, *ESAIM: Math. Model. Numer. Anal.*, **5**(4):1135–1153, 2010.
2. Galvis, J. and Sarkis, M., Aproximating Infinity-Dimensional Stochastic Darcy's Equations without Uniform Ellipticity, *SIAM J. Numer. Anal.*, **47**(5):3624–3651, 2009.
3. Gittelson, C.J., Stochastic Galerkin Discretization of the Log-Normal Isotropic Diffusion Problems, *Math. Models Methods Appl. Sci.*, **20**(2):237–263, 2010.
4. Mugler, A. and Starkloff, H.-J., On Elliptic Partial Differential Equations with Random Coefficients, *Stud. Univ. Babes-Bolyai Math.*, **56**(2):473–487, 2011.
5. Da Prato, G. and Zabczyk, J., *Stochastic Equations in Infinite Dimensions*, Encyclopedia of Mathematics and Its Applications, vol. 44, Cambridge, UK: Cambridge University Press, 1992.
6. Charrier, J., Strong and Weak Error Estimates for Elliptic Partial Differential Equations with Random Coefficients, *SIAM J. Numer. Anal.*, **50**(1):216–246, 2012.
7. Ghanem, R. and Spanos, P., *Stochastic Finite Element: A Spectral Approach*, New York: Springer-Verlag, 1991.
8. Babuska, I., Tempone, R., and Zouraris, G., Galerkin Finite Element Approximations of Stochastic Elliptic Differential Equations, *SIAM J. Numer. Anal.*, **42**:800–825, 2004.
9. Frauenfelder, P., Schwab, C., and Todor, R., Finite Elements for Elliptic Problems with Stochastic Coefficients, *Comput. Methods Appl. Mech. Eng.*, **194**:205–228, 2005.
10. Todor, R. and Schwab, C., Convergence Rates for Sparse Chaos Approximations of Elliptic Problems with Stochastic Coefficients, *IMA J. Numer. Anal.*, **27**(2):232–261, 2007.
11. Bonizzoni, F. and Nobile, F., Perturbation Analysis for the Darcy Problem with Log-Normal Permeability, *SIAM/ASA J. Uncertainty Quantif.*, **2**:223–244, 2014.
12. Bonizzoni, F., Nobile, F., and Kressner, D., Tensor Train Approximation of Moment Equations for Elliptic Equations with Lognormal Coefficient, *Comput. Meth. Appl. Mech. Eng.*, **308**:349–376, 2016.
13. Chkifa, A., Cohen, A., DeVore, R., and Schwab, C., Sparse Adaptive Taylor Approximation Algorithms for Parametric and Stochastic Elliptic PDEs, *ESAIM: Math. Model. Numer. Anal.*, **47**(1):253–280, 2013.

14. Nobile, F. and Tesei, F., A Multi Level Monte Carlo Method with Control Variate for Elliptic PDEs with Log-Normal Coefficients, *Stochastic PDE: Anal. Comput.*, **3**(3):398–444, 2015.
15. Nobile, F., Tamellini, L., Tesei, F., and Tempone, R., An Adaptive Sparse Grid Algorithm for Elliptic PDEs with Lognormal Diffusion Coefficient, in *Sparse Grids and Applications—Stuttgart 2014*, J. Garcke and D. Pflger, Eds., Cham, Switzerland: Springer International Publishing, vol. 109, pp. 191–220, 2016.
16. Powell, C.E. and Elman, H.C., Block-Diagonal Preconditioning for Spectral Stochastic Finite-Element Systems, *IMA J. Numer. Anal.*, **29**(2):350–375, 2009.
17. Powell, C. and Ullmann, E., Preconditioning Stochastic Galerkin Saddle Point Systems, *SIAM J. Matrix Anal. Appl.*, **31**(5):2813–2840, 2010.
18. Hampton, J., Fairbanks, H., Narayan, A., and Doostan, A., *Parametric/Stochastic Model Reduction: Low-Rank Representation, Non-Intrusive Bi-Fidelity Approximation, and Convergence Analysis*, arXiv:1709.03661 [math], Sep. 2017.
19. Holden, H., Oksendal, B., and Zhang, T., *Stochastic Partial Differential Equations: A Modeling, White Noise Functional Approach*, Boston: Birkhauser, 1996.
20. Theting, T., Solving Wick-Stochastic Boundary Value Problems using a Finite Element Method, *Stochastics: Int. J. Prob. Stochastic Proc.*, **70**:241–270, 2000.
21. Wan, X., Rozovskii, B., and Karniadakis, G., A Stochastic Modeling Methodology based on Weighted Wiener Chaos and Malliavin Calculus, *Proc. Natl. Acad. Sci.*, **106**:14189–14194, 2009.
22. Wan, X., A Note on Stochastic Elliptic Models, *Comput. Methods Appl. Mech. Eng.*, **199**(45-48):2987–2995, 2010.
23. Wan, X., A Discussion on Two Stochastic Modeling Strategies for Elliptic Problems, *Commun. Comput. Phys.*, **11**:775–796, 2012.
24. Mikulevicius, R. and Rozovskii, B.L., On Unbiased Stochastic Navier-Stokes Equations, *Probab. Theory Relat. Fields*, **154**(3-4):787–834, 2012.
25. Nualart, D., *Malliavin Calculus and Related Topics*, 2nd ed., New York: Springer, 2006.
26. Venturi, D., Wan, X., Mikulevicius, R., Rozovskii, B., and Karniadakis, G., Wick-Malliavin Approximation to Nonlinear Stochastic PDEs: Analysis and Simulations, *Proc. R. Soc. A*, **469**:20130001, 2013.
27. Wan, X. and Rozovskii, B.L., The Wick-Malliavin Approximation of Elliptic Problems with Log-Normal Random Coefficients, *SIAM J. Sci. Comput.*, **35**(5):A2370–A2392, 2013.
28. Hu, Y. and Yan, J., Wick Calculus for Nonlinear Gaussian Functionals, *Acta Math. Appl. Sin. Eng. Ser.*, **25**:399–414, 2009.
29. Cameron, R. and Martin, W., The Orthogonal Development of Nonlinear Functionals in Series of Fourier-Hermite Functionals, *Ann. Math.*, **48**:385, 1947.
30. Lototsky, S. and Rozovskii, B., Stochastic Differential Equations Driven by Purely Spatial Noise, *SIAM J Math. Anal.*, **41**(4):1295–1322, 2009.
31. Shen, J. and Yu, H., Efficient Spectral Sparse Grid Methods and Applications to High-Dimensional Elliptic Problems, *SIAM J. Sci. Comput.*, **32**:3228–3250, 2010.
32. Shen, J. and Yu, H., Efficient Spectral Sparse Grid Methods and Applications to High-Dimensional Elliptic Equations II: Unbounded Domains, *SIAM J. Sci. Comput.*, **34**:1141–1164, 2012.
33. Shen, J., Wang, L.-L., and Yu, H., Approximations by Orthonormal Mapped Chebyshev Functions for Higher-Dimensional Problems in Unbounded Domains, *J. Comput. Appl. Math.*, **265**:264–275, 2014.
34. Riesz, F. and Nagy, B.S., *Functional Analysis*, New York: Dover, 1990.
35. Landau, L.D. and Lifshitz, E.M., *Electrodynamics of Continuous Media*, Oxford: Pergamon Press, 1960.
36. Matheron, G., *Eléments pour une Théorie des Milieux Poreux*, Paris: Masson, 1967.
37. Ciarlet, P., *The Finite Element Method for Elliptic Problems*, Philadelphia: SIAM, 2002.
38. Karniadakis, G. and Sherwin, S., *Spectral/hp Element Methods for CFD*, 2nd ed., Oxford: Oxford University Press, 2005.
39. Schwab, C., *p- and hp- Finite Element Methods*, Oxford: Oxford University Press, 1998.
40. Saad, Y., *Iterative Methods for Sparse Linear Systems*, 2nd ed., Philadelphia: SIAM, 2003.