Commun. Comput. Phys.
doi: 10.4208/cicp.270913.020414a

Vol. **16**, No. 3, pp. 571-598
September 2014

An Adaptive ANOVA-Based Data-Driven Stochastic Method for Elliptic PDEs with Random Coefficient

Zhiwen Zhang[1,†], Xin Hu[1,†], Thomas Y. Hou[1,*], Guang Lin[2,*] and Mike Yan[1]

[1] *Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA 91125, USA.*
[2] *Pacific Northwest National Laboratory, Richland, WA 99352, USA.*

Received 27 September 2013; Accepted (in revised version) 2 April 2014

Available online 24 June 2014

Abstract. In this paper, we present an adaptive, analysis of variance (ANOVA)-based data-driven stochastic method (ANOVA-DSM) to study the stochastic partial differential equations (SPDEs) in the multi-query setting. Our new method integrates the advantages of both the adaptive ANOVA decomposition technique and the data-driven stochastic method. To handle high-dimensional stochastic problems, we investigate the use of adaptive ANOVA decomposition in the stochastic space as an effective dimension-reduction technique. To improve the slow convergence of the generalized polynomial chaos (gPC) method or stochastic collocation (SC) method, we adopt the data-driven stochastic method (DSM) for speed up. An essential ingredient of the DSM is to construct a set of stochastic basis under which the stochastic solutions enjoy a compact representation for a broad range of forcing functions and/or boundary conditions.

Our ANOVA-DSM consists of offline and online stages. In the offline stage, the original high-dimensional stochastic problem is decomposed into a series of low-dimensional stochastic subproblems, according to the ANOVA decomposition technique. Then, for each subproblem, a data-driven stochastic basis is computed using the Karhunen-Loève expansion (KLE) and a two-level preconditioning optimization approach. Multiple trial functions are used to enrich the stochastic basis and improve the accuracy. In the online stage, we solve each stochastic subproblem for any given forcing function by projecting the stochastic solution into the data-driven stochastic basis constructed offline. In our ANOVA-DSM framework, solving the original high-dimensional stochastic problem is reduced to solving a series of ANOVA-decomposed stochastic subproblems using the DSM. An adaptive ANOVA strategy is also provided to further reduce the number of the stochastic subproblems and speed up our method. To demonstrate the accuracy and efficiency of our method, numerical examples are presented for one- and two-dimensional elliptic PDEs with random coefficients.

[†]The first two authors contribute equally to this work.
[*]Corresponding author. *Email addresses:* `hou@cms.caltech.edu` (T. Hou), `Guang.Lin@pnnl.gov` (G. Lin)

AMS subject classifications: 60H35, 65C30, 35R60, 62H99

Key words: Analysis of variance, stochastic partial differential equations, data-driven methods, Karhunen-Loeve expansion, uncertainty quantification, model reduction.

1 Introduction

Over the past few decades, there has been growing interest and significant progress in modeling complex physical and engineering systems with uncertainties. Many physical and engineering applications involving uncertainty quantification can be described by stochastic partial differential equations (SPDEs). One of the essential challenges in these applications is how to solve SPDEs efficiently when the dimension of stochastic input variables is high. These problems are computationally prohibitive for some of the existing numerical methods, such as stochastic finite element method [14], Wiener chaos expansion method [19, 28], generalized polynomial chaos (gPC) methods [29, 35, 37, 38], and stochastic collocation method [1, 39]. One of the reasons is that these methods use a problem-independent basis, which produces a very large coupled system when the dimension of the input stochastic variables is high.

For stochastic problems with high stochastic input dimensions, we employ the functional Analysis of Variance, or ANOVA method [4, 16] as a dimension-reduction technique. This is motivated by the observation that for many real physical systems, only a relatively small number of stochastic dimensions is important and will significantly impact the stochastic systems' outputs. The ANOVA decomposition was introduced by Fisher [9]. Later in 1948, Hoeffding successfully applied ANOVA decomposition to study U-statistics [17]. ANOVA also was used for uncertainty quantification in [36] and was employed in gPC for solving high-dimensional stochastic PDE systems in [5, 10, 12, 26, 27, 40, 42]. In [10] ANOVA was integrated with a multi-element stochastic collocation method. In [26], an adaptive version of ANOVA was developed to automatically detect the important dimensions. In [40], adaptive ANOVA methods based on three different adaptive criteria were proposed and compared.

ANOVA decomposition of the original high-dimensional stochastic problem results in a set of low-dimensional subproblems in stochastic space, which are efficiently solved by the sparse-grid stochastic collocation method. The stochastic collocation method was first introduced by Tatang and McRae in [34]. The properties of stochastic collocation method have been extensively studied in the past 10 years. In [3,30,31], the errors of integrating or interpolating functions with Sobolev regularity were analyzed for Smolyak constructions based on one-dimensional (1D) nested Clenshaw-Curtis rules. In [31], the degree of exactness of the Smolyak quadrature using Clenshaw-Curtis and Gaussian one-dimensional rules was investigated. In 2003, Gerstner and Griebel [13] introduced the dimension-adaptive tensor-product quadrature method. Recently Xiu and Hesthaven [39] have used Lagrange polynomial interpolation to construct high-order

stochastic collocation methods. In [39], the efficiency of Clenshaw-Curtis-based sparse grid stochastic collocation was demonstrated in comparison to other stochastic methods on an elliptic problem. In [11], sparse grid collocation schemes were applied to solving stochastic natural convection problems. In [21–24], a multi-element stochastic collocation method was employed to study the random roughness problem, stochastic compressible flow, and plasma flow problems. In [15, 26, 41], some adaptive hierarchical sparse grid collocation algorithms were developed.

In traditional numerical methods for SPDEs, such as the gPC and stochastic collocation methods, the basis are determined depending on the probabilistic distribution of stochastic inputs, which are problem independent. However, the stochastic outputs may not share the same probabilistic distribution as the stochastic inputs. Hence, these basis may not be optimal for such systems and cause the slow convergence. In applications, we often need to solve the same SPDE many times with multiple forcing functions or boundary conditions. This is also known as the *multi-query problem*. In this case, these traditional numerical methods will be computationally expensive or even infeasible. In [7], the authors proposed and developed a data-driven stochastic method (DSM) to study the multi-query problem of SPDEs. The motivation for the DSM is to obtain a problem-dependent stochastic basis, which can provide a compact representation of the solution to the SPDEs with multiple forcing functions or boundary conditions. To illustrate the main idea of the DSM, we consider a SPDE of the form:

$$\mathcal{L}(x,\omega)u(x,\omega)=f(x,\theta), \quad x\in D,\ \omega\in\Omega, \tag{1.1a}$$

$$u(x,\omega)=0, \quad x\in\partial D,\ \omega\in\Omega, \tag{1.1b}$$

where $D\in R^d$ is a bounded spatial domain, $\mathcal{L}(x,\omega)$ is a stochastic differential operator, and $f(x,\theta)$ is the deterministic forcing function parameterized by θ.

The DSM method proposed in [7] consists of an offline and an online stage. Below we will give a brief review of this method. In the offline stage, we use the Karhunen-Loève expansion (KLE) [20, 25] of the SPDE solutions to construct stochastic basis $\{A_i(\omega)\}_{i=0}^m$, where $A_0(\omega)=1$ and m is the number of elements in the basis. Specifically, we expand the stochastic solution in terms of this stochastic basis $u(x,\omega)=\sum_{i=0}^m u_i(x)A_i(\omega)$ and solve a coupled system of PDEs by the Galerkin method for the deterministic coefficients, $\{u_i(x)\}_{i=0}^m$. The detailed construction of the stochastic basis will be elaborated in detail in Section 2, see also [7].

The KLE is well known for generating the optimal basis in the sense that it minimizes the total mean squared error. As a result, it gives the optimal representation of a stochastic solution. First, a compact representation of $f(x,\theta)$ is constructed by expanding it into a finite dimensional basis $f_i(x)$, i.e., $f(x,\theta)\approx\sum_{i=0}^K c_i(\theta)f_i(x)$. Such expansion can be obtained by applying singular value decomposition (SVD) or the empirical interpolation method (EIM) [2] to $f(x,\theta)$. With such parametrization of f, we begin the construction of the stochastic basis $\{A_i(\omega)\}_{i=0}^m$ based on the KLE of the SPDE solution of Eq. (1.1) with $f_0(x)$ as a forcing function. An error analysis was proposed in [7] to evaluate the

completeness of the data-driven basis $\{A_i(\omega)\}_{i=0}^m$. When the dimension of the stochastic solution is low, the stochastic basis $\{A_i(\omega)\}_{i=0}^m$ is approximately complete with just one forcing function. However, when the dimension of the stochastic solution is high, we need to use multiple forcing functions to construct the stochastic basis $\{A_i(\omega)\}_{i=0}^m$. To reduce the computational cost in computing the KLE of the SPDE solution, the randomized SVD algorithm [7,18] was used in [7] to directly calculate the KLE of the stochastic solution instead of forming a covariance matrix and solving an expensive eigenvalue problem.

To ensure the stochastic basis $\{A_i\}$ is applicable to a broad range of forcing functions $f(x)$, an algorithm was designed to enrich stochastic basis based on the trial functions $f_k(x)$, $k=1,2,\cdots,K$. More specifically, a greedy-type algorithm has been proposed, which is used together with a two-level preconditioning [8] to reduce the computational cost. First, one derives an error equation for the stochastic solution obtained by the most recently enriched basis. On the coarse grid, the error equation is solved for each trial function $f_k(x)$, $k=1,2,\cdots,K$, and the maximum error τ_{k^*} is identified along with the corresponding trial function f_{k^*}. Subsequently, the error equation for this trial function is solved again, but on the fine grid. After that, the KLE of the residual error can be used to enrich the stochastic basis. This process is repeated until the maximum residual error is below the prescribed threshold ϵ. When this updating process terminates, we obtain our data-driven basis $\{A_i(\omega)\}_{i=0}^m$, which provides a compact representation of the SPDE solutions that can be used to solve this parameterized family of forcing functions. This enriching algorithm is illustrated in Fig. 1 in Section 2. The detailed implementation of this enriching algorithm depends on the specific numerical representation of the stochastic basis, which will be elaborated about at length in Section 2.

In the online stage, we expand the SPDE solution under the data-driven stochastic basis and solve a set of coupled deterministic PDEs to obtain the coefficients. By exploring the low-dimensional structure of the solution, our DSM offers considerable computational saving over some traditional methods, especially when the dimension of input stochastic variables is high, but the effective dimension of the output stochastic solution is low. Depending on the representations of the data-driven stochastic basis $\{A_i(\omega)\}$, three versions of DSM have been proposed, i.e., *ensemble representation*, *stochastic collocation representation*, and *spectral representation*. However, it should be pointed out that when the input stochastic variables are high, how to effectively represent the data-driven stochastic basis is still an open issue. To address this issue, in this paper we present an adaptive, ANOVA-based, data-driven stochastic method (ANOVA-DSM) to study the SPDEs in the multi-query setting, which integrates the advantages of both the ANOVA decomposition technique and the DSM. For notation simplicity, we only consider the DSM in stochastic collocation representation in this paper.

This paper is organized as follows; in Section 2, we provide the detailed derivation of DSM in stochastic collocation representation. In Section 3, we give a preliminary introduction about standard ANOVA decomposition. The adaptive ANOVA-DSM is discussed in Section 4. In Section 5, we apply our method to both the one- and two-

dimensional elliptic PDEs with random elliptic coefficients to demonstrate its computational efficiency. Finally, some concluding remarks are provided in Section 6.

2 The data-driven stochastic method

2.1 General framework of the DSM

In this section, we give a brief review of DSM and its general framework. The primary driver for the data-driven stochastic method is to obtain a problem-dependent stochastic basis under which the solution of an SPDE enjoys a compact expansion. Clearly, such a stochastic basis should be constructed by learning some information from the solution in a data-driven fashion. The KL expansion is well known for generating the optimal basis in the sense that it minimizes the total mean squared error. Therefore we choose the KLE for post-processing of the SPDE solution and constructing a problem-dependent stochastic basis.

We first outline the DSM's general framework, which consists of offline and online stages. In the offline stage, an effective strategy was proposed to construct a data-driven basis $\{A_i(\omega)\}_{i=0}^m$, where $A_0(\omega)=1$ and m is the number of elements in the basis. The DSM method is a greedy-type algorithm combined with a two-level preconditioning [8] to reduce the computational cost. Once the data-driven basis is constructed, we can use it in the standard Galerkin method to solve the SPDEs (1.1) in the online stage. As the online stage is rather straightforward, we only state the offline computation algorithm as follows (refer to Fig. 1 for illustration of the main ideas):

DSM Offline Computation

- Step 0 (**Preparations**):
  - Set the error threshold ϵ_0. Partition the spatial domain D into a fine grid D_h and a coarse grid D_H;
  - Approximate $f(x,\theta)$ by a finite dimensional basis $\{f_k(x)\}_{k=0}^K$, i.e. $f(x,\theta)\approx\sum_{k=0}^K c_k(\theta)f_k(x)$.
- Step 1 (**Initial learning on the fine grid** D_h):
  - Solve Eq. (1.1) with $f_0(x)$ as a forcing function to obtain $u(x,\omega;f_0)$;
  - Calculate the truncated KLE of $u(x,\omega;f_0)$ and use the first m_1 terms of the stochastic modes to obtain the current data-driven basis $\{A_i(\omega)\}_{i=0}^{m_1}$, where $A_0(\omega)=1$.
- Step 2 (**Preconditioning on the coarse grid** D_H):
  - For each trial function $f_k(x)$, solve Eq. (1.1) using the current stochastic basis $\{A_i(\omega)\}_{i=0}^{m_1}$ and the stochastic Galerkin method to obtain DSM solution $u_{DSM}(x,\omega;f_k)$;

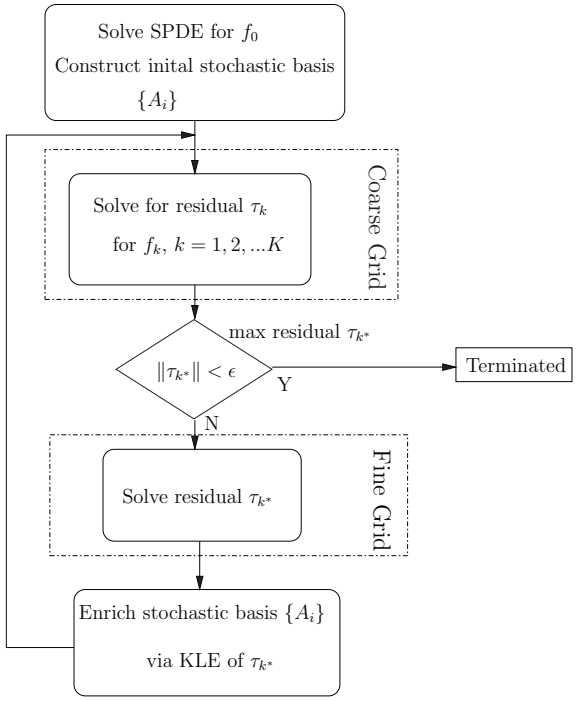

Figure 1: Greedy stochastic basis enriching algorithm on a coarse-fine grid hierarchy.

  - For each trial function $f_k(x)$, solve a residual error equation to obtain the approximate residual error $\tau_k = \tau(x,\omega;f_k)$;
  - If $\max_{1\le k\le K}||\tau_k|| < \epsilon_0$, go to Step 4; else set $k^* = arg\max_{0\le k\le K}||\tau_k||$ and $f_{k^*}(x)$, go to Step 3.

- Step 3 (**Update on fine grid** D_h):
  - Solve the residual equation associated with $f_{k^*}(x)$ to obtain the residual error $\tau_{k^*} = \tau(x,\omega;f_{k^*})$;
  - Enrich the current stochastic basis $\{A_i(\omega)\}_{i=0}^{m_1}$ by the KLE of τ_{k^*} and use $\{A_i(\omega)\}_{i=0}^{m_2}$ to denote the updated stochastic basis. Go to Step 2.

- Step 4 (**Termination**):
  - Save the data-driven stochastic basis, denoted by $\{A_i(\omega)\}_{i=0}^{m}$ and relevant statistic quantities.

The detailed implementation of this greedy-type algorithm depends on the numerical representation of the stochastic basis. In [7], we proposed three ways to represent the stochastic basis:

- Ensemble representation, i.e, sampling method, such as Monte Carlo method, quasi-Monte Carlo method, etc.
- Stochastic collocation representation, such as the sparse grid based stochastic collocation (SC) basis.
- Spectral representation, such as the gPC basis.

Different representations have their own advantages and disadvantages. The DSM in ensemble representation has the advantage that its accuracy does not depend on the dimension of the input random variables. However, its convergence rate is rather slow. The DSM in collocation representation and gPC representation depends on the multi-index of the orthonormal polynomial basis. Both are very accurate but could be expensive when the dimension of the input random variables is large. Under the same computational condition, the collocation representation can handle a larger multi-index of the orthonormal polynomial basis than the gPC representation due to its non-intrusive nature. Herein, we only introduce the DSM in collocation representation.

2.2 The generalized Polynomial Chaos basis

We assume the randomness in the differential operator $\mathcal{L}(x,\omega)$ in SPDEs (1.1) is given in terms of r independent random variables, i.e., $\boldsymbol{\xi}(\omega)=(\xi_1(\omega),\xi_2(\omega),\cdots,\xi_r(\omega))$. Without the loss of generality, we can further assume such independent random variables have the same distribution function $\rho(x)$. We get $\mathcal{L}(x,\omega)=\mathcal{L}(x,\xi_1(\omega),\cdots,\xi_r(\omega))$. By the Doob-Dynkin lemma [32], the solution of Eq. (1.1) can still be represented by these random variables, i.e. $u(x,\omega)=u(x,\xi_1(\omega),\cdots,\xi_r(\omega))$.

Let $\{\mathrm{H}_i(\xi)\}_{i=1}^{\infty}$ denote the 1D polynomials that are orthogonal to each other with respect to the distribution $\rho(\xi)$, i.e.,

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

For some commonly used distributions, such as the Gaussian and uniform distributions, these orthogonal polynomial sets are Hermite and Legendre polynomials, respectively. For general distributions, such polynomial sets can be obtained by numerical methods [35]. Furthermore, by a tensor product representation, we can use the 1D polynomial $\mathrm{H}_i(\xi)$ to construct a sufficient orthonormal basis $\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$'s of $\mathbb{L}^2(\Omega)$ as follows:

$$\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})=\prod_{i=1}^{r}\mathrm{H}_{\alpha_i}(\xi_i),\quad \boldsymbol{\alpha}\in\mathfrak{J}_r^{\infty}, \tag{2.1}$$

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

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

The zero multi-index corresponding to $\mathbf{H_0}(\boldsymbol{\xi})=1$, which is used to represent the mean of the solution. Clearly, the cardinality of $\mathfrak{J}_r^\infty$ is infinite. For the purpose of numerical computations, we prefer a finite set of polynomials. One possible choice is the set of polynomials whose total orders are, at most p, i.e.,

$$\mathfrak{J}_r^p=\left\{\boldsymbol{\alpha}\,\middle|\,\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\cdots,\alpha_r),\ \alpha_i\geq 0,\ \alpha_i\in\mathbb{N},|\alpha|=\sum_{i=1}^{r}\alpha_i\leq p\right\}. \tag{2.2}$$

The cardinality of $\mathfrak{J}_r^p$ in (2.2) or the number of polynomial basis functions, denoted by $N_p=\left|\mathfrak{J}_r^p\right|$, is equal to $\frac{(p+r)!}{p!r!}$. When no ambiguity arises, we may simply write such a truncated set as $\mathfrak{J}$. The orthonormal basis $\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi})$ is the gPC basis (refer to [14, 19, 37] for more details).

2.3 Data-driven stochastic basis via a collocation representation

In this section, we introduce the DSM method via a collocation representation. For notation convenience, we consider the following 1D SPDEs:

$$-\nabla\cdot(a(x,\omega)\nabla u(x,\omega))=f(x,\theta),\quad x\in D,\ \omega\in\Omega, \tag{2.3}$$

$$u(x,\omega)=0,\qquad x\in\partial D, \tag{2.4}$$

where the coefficient $a(x,\omega)$ is assumed to be positive with upper and lower bounds almost surely. If the coefficient $a(x,\omega)$ is given in terms of r independent random variables, i.e., $a(x,\omega)=a(x,\boldsymbol{\xi}(\omega))=a(x,\xi_1(\omega),\cdots,\xi_r(\omega))$, the solution of Eq. (2.3) can be represented by these random variables, i.e., $u(x,\omega)=u(x,\xi_1(\omega),\cdots,\xi_r(\omega))$. The forcing function $f(x,\theta)$ is approximated using a finite basis, $\{f_k(x)\}_{k=0}^K$, i.e., $f(x,\theta)=\sum_{k=0}^K c_k(\theta)f_k(x)$.

In the **initial learning step** of our DSM method, we first use the stochastic collocation method to generate J collocation points $\mathbf{z}_j\in R^r$ according to the distribution of randomness in the coefficient $a(x,\boldsymbol{\xi}(\omega))$, as well as the associated weights $w_j\in R$. Then, we solve Eqs. (2.3)-(2.4) with the random variable evaluated at the collocation grid points and $f_0(x)$ as the right-hand side

$$-\nabla\cdot(a(x,\mathbf{z}_j)\nabla u(x,\mathbf{z}_j))=f_0(x),\quad x\in D,\ j=1,\cdots,J, \tag{2.5}$$

$$u(x,\mathbf{z}_j)=0,\qquad x\in\partial D. \tag{2.6}$$

By solving Eqs. (2.5)-(2.6), we can obtain the values of the stochastic solution $u(x,\omega;f_0)$ on the collocation points, i.e., $\{u(x,\mathbf{z}_j;f_0)\}_{j=1}^J$. The m_1-term KLE of the solution $u(x,\omega;f_0)$ gives the dominant components in the random space. We use the decaying property of eigenvalues to select parameter m_1, i.e., the number of stochastic basis m_1 that can be chosen, such as $\lambda_{m_1+1}/\lambda_1$, is smaller than some pre-defined threshold, say 10^{-4}. We denote the truncated KLE as;

$$u(x,\omega;f_0)\approx\bar{u}(x;f_0)+\sum_{i=1}^{m_1}\sqrt{\lambda_i}A_i(\omega)\phi_i(x;f_0). \tag{2.7}$$

We call the stochastic basis $\{A_i(\omega)\}_{i=0}^{m_1}$ in Eq. (2.7) the *data-driven stochastic basis*, where $A_0(\omega)=1$. Furthermore, we would like to expand the stochastic basis $A_i(\omega)$ in a gPC basis, i.e.,

$$A_i(\omega)=\sum_{\boldsymbol{\alpha}} A_{\boldsymbol{\alpha}i}\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}(\omega)), \tag{2.8}$$

where $\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}(\omega))$ are the orthonormal polynomial basis decided by the distribution of the random variables in the coefficient $a(x,\boldsymbol{\xi}(\omega))$. The expansion coefficient $A_{\boldsymbol{\alpha}i}$ is given by:

$$A_{\boldsymbol{\alpha}i}=E[A_i(\omega)\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}(\omega))]\approx\sum_{j=1}^{J}A_i(\mathbf{z}_j)\mathbf{H}_{\boldsymbol{\alpha}}(\mathbf{z}_j)w_j,\quad \boldsymbol{\alpha}\in\mathfrak{J}_r^p, \tag{2.9}$$

where $\mathbf{z}_j\in R^r$ and $w_j\in R$ are the sparse grid points and associated weights, respectively. We use the N_p-by-(m_1+1) matrix $\mathbf{A}$ to denote the expansion coefficient $A_{\boldsymbol{\alpha}i}$, which is essentially the data-driven stochastic basis in the stochastic collocation representation. In general, the stochastic basis constructed by using f_0 may not be adequate to give an accurate approximation of the SPDE for another right-hand side, $f(x,\theta)$. We need to supplement the stochastic basis by using multiple trial functions involving other f_k.

In the **preconditioning and update step** of DSM, a greedy-type algorithm was proposed which is use together with a two-level preconditioning strategy to enrich the stochastic basis. First of all, we consider the error analysis. Given a new right-hand side $f_1(x)=f(x,\theta)$ for some choice of θ, we expand the solution in terms of the stochastic basis, $\{A_i(\omega)\}_{i=0}^{m}$,

$$u(x,\omega;f_1)\approx\bar{u}(x;f_1)+\sum_{i=1}^{m_1}A_i(\omega)u_i(x;f_1)\equiv\sum_{i=0}^{m_1}A_i(\omega)u_i(x;f_1). \tag{2.10}$$

In the rest part of this subsection, we also use $u_i(x)\equiv u_i(x;f_1)$ for simplification. We use the standard stochastic Galerkin method to obtain the coefficient $u_i(x)$. Specifically, we substitute the expansion (2.10) into the SPDE (2.3), multiply both sides by $A_j(\omega)$ and take expectations. This gives rise to a coupled PDEs system for the expansion coefficient $u_i(x)$,

$$-\nabla\cdot\big(E[aA_iA_j]\nabla u_i(x,\omega)\big)=f_1(x)E[A_j],\quad x\in D,\ j=0,1,\cdots,m_1, \tag{2.11}$$

$$u_i(x)=0,\qquad x\in\partial D, \tag{2.12}$$

where Einstein summation is assumed. The term $E[aA_iA_j]$ can be calculated as follows:

$$E[aA_iA_j]=a_{\boldsymbol{\alpha}}(x)A_{\boldsymbol{\beta}i}A_{\boldsymbol{\gamma}j}E[\mathbf{H}_{\boldsymbol{\alpha}}(\boldsymbol{\xi}(\omega))\mathbf{H}_{\boldsymbol{\beta}}(\boldsymbol{\xi}(\omega))\mathbf{H}_{\boldsymbol{\gamma}}(\boldsymbol{\xi}(\omega))], \tag{2.13}$$

and

$$E[\mathbf{H}_{\boldsymbol{\alpha}}(\omega)\mathbf{H}_{\boldsymbol{\beta}}(\omega)\mathbf{H}_{\boldsymbol{\gamma}}(\omega)]\approx\sum_{j=1}^{J}\mathbf{H}_{\boldsymbol{\alpha}}(\mathbf{z}_j)\mathbf{H}_{\boldsymbol{\beta}}(\mathbf{z}_j)\mathbf{H}_{\boldsymbol{\gamma}}(\mathbf{z}_j)w_j,\quad \boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}\in\mathfrak{J}_r^p. \tag{2.14}$$

We can solve the coupled deterministic PDEs system (2.11)-(2.12) by any standard numerical method, such as finite element method (FEM) or finite difference method (FDM). We then obtain the expansion coefficient $\{u_i(x)\}_{i=0}^{m_1}$ and the approximate solution for $u(x,\omega;f_1)$ in (2.10). We know the exact solution can be written as:

$$u(x,\omega;f_1)=\sum_{i=0}^{m_1}A_i(\omega)u_i(x;f_1)+\tau(x,\omega;f_1), \tag{2.15}$$

where $\tau(x,\omega;f_1)$ is the error. Simple calculations show that the error satisfies the following equation:

$$-\nabla\cdot(a(x,\omega)\nabla\tau(x,\omega;f_1))=f_1(x)+\sum_{i=0}^{m_1}\nabla\cdot(a(x,\omega)A_i(\omega)\nabla u_i(x)). \tag{2.16}$$

To verify the effectiveness of the stochastic basis, we solve the residual Eq. (2.16) on a *coarse grid* for each $f_k(x)$ $(k=1,\cdots,K)$ and obtain the error $\{\tau(x,\omega;f_k)\}_{i=k}^{K}$. If $\max_{1\le k\le K}||\tau(x,\omega;f_k)||<\epsilon_0$, then this stochastic basis is sufficient, where $||\cdot||$ is the L^2 norm of the variance of the stochastic solution. Otherwise, we identify the maximum error $\tau_{k^*}=\max_{1\le k\le K}||\tau(x,\omega;f_k)||$ along with the corresponding trial function $f_{k^*}(x)$. Subsequently, we solve the residual Eq. (2.16) for this trial function $f_{k^*}(x)$ again on a *fine grid*. We also do the KLE for the residual solution $\tau(x,\omega;f_{k^*})$ and extract several dominant components in the random space, supplementing them to the current stochastic basis. We use $\{A_i(\omega)\}_{i=0}^{m_2}$ to denote the updated stochastic basis. This process is repeated until the maximum residual is below the prescribed threshold ϵ_0. We project the stochastic basis denoted by $\{A_i(\omega)\}_{i=0}^{m}$ into the gPC basis according to Eqs. (2.8)-(2.9) and only save the N_p-by-(m+1) matrix **A**.

In **the online stage**, for each query $f(x,\theta)$ with our data-driven stochastic basis **A**, we use the standard stochastic Galerkin method to solve the SPDEs (2.3)-(2.4). Specifically, we expand the solution in terms of the stochastic basis $\{A_i(\omega)\}_{i=0}^{m}$, i.e.,

$$u(x,\omega)\equiv\sum_{i=0}^{m}A_i(\omega)u_i(x). \tag{2.17}$$

Then, we substitute the expansion (2.17) into the SPDE (2.3), multiply both side by $A_j(\omega)$ and take expectations. This gives rise to a coupled PDEs system for the expansion coefficient $u_i(x)$,

$$-\nabla\cdot\big(E[aA_iA_j]\nabla u_i(x,\omega)\big)=f(x,\theta)E[A_j],\quad x\in D,\ j=0,1,\cdots,m_1, \tag{2.18}$$

$$u_i(x)=0,\qquad x\in\partial D, \tag{2.19}$$

where Einstein summation is assumed. Solving the coupled deterministic PDEs system (2.18)-(2.19) by the FEM or FDM, we obtain the expansion coefficient $\{u_i(x)\}_{i=0}^{m_1}$ and thus the solution $u(x,\omega)$.

Construction of the stochastic basis could be expensive. However, once the stochastic basis is constructed, it can be used repeatedly for different right-hand side function $f(x,\theta)$ in the online stage. In [7], a complexity analysis has been carried out for DSM. It was demonstrated that in the multi-query problem it offers considerable computational savings than traditional methods, such as Monte Carlo, gPC and stochastic collocation methods. See [7] for more details. We will also demonstrate the effectiveness of this method in Section 5.

3 ANOVA decomposition

In statistics, the ANOVA method can be used to describe the interactions between a large number of variables when only few samples are available. The same idea can be adopted in the interpolation and integration of high-dimensional problems, as well as stochastic systems. For most well-defined physical systems, only relatively low-order correlations of the input variables are expected to be important for the output of the system. The ANOVA expansion utilizes this property, and, at each new level of ANOVA expansion, higher-order correlation effects of the input variables are accounted for. Consider a Lebesgue integrable multivariate stochastic function $f(\mathbf{Y}):\mathbb{R}^d\to\mathbb{R}$, and d is the dimension of stochastic space we are interested in. The ANOVA expansion represents $f(\mathbf{Y})$ as finite hierarchical correlated functions of input variables in the form of:

$$f(\mathbf{Y})=f_0+\sum_{s=1}^{d}\sum_{j_1<\cdots<j_s} f_{j_1,\cdots,j_s}(Y_{j_1},\cdots,Y_{j_s}), \tag{3.1}$$

or equivalently

$$\begin{aligned} f(\mathbf{Y})=f_0&+\sum_{1\le j_1\le d} f_{j_1}(Y_{j_1})+\sum_{1\le j_1<j_2\le d} f_{j_1,j_2}(Y_{j_1},Y_{j_2})+\cdots\\ &+f_{1,2,\cdots,d}(Y_1,Y_2,\cdots,Y_d). \end{aligned} \tag{3.2}$$

We call $f_{j_k}(Y_{j_k})$, $1\le j_1\le d$ the first-order term, $f_{j_k,j_l}(Y_{j_k},Y_{j_l})$, $1\le j_1<j_2\le d$ the second-order term, etc. The ANOVA components have the following properties:

1. The constant term is the mean of function, i.e.,

$$f_0=\int_{\Gamma^d} f(\mathbf{Y})d\mu(\mathbf{Y}), \tag{3.3}$$

which means that all higher-order components have zero mean:

$$\int_{\Gamma^d} f_{j_1,\cdots,j_s}d\mu(\mathbf{Y})=0,\quad \text{for } 1\le s\le d. \tag{3.4}$$

2. The other important property of ANOVA expansion is the orthogonality among its terms:
$$\int_{\Gamma^d} f_{j_1,\cdots,j_s} f_{k_1,\cdots,k_l} d\mu(\mathbf{Y}) = 0, \tag{3.5}$$
if $(j_1,\cdots,j_s) \neq (k_1,\cdots,k_l)$. This is the direct consequence of (3.4).

3. The variance of f is the sum of variance of all component functions:
$$\sigma^2(f) = \sum_{s=1}^{d} \sum_{|\mathbf{s}|=s} \sigma^2(f_{\mathbf{s}}). \tag{3.6}$$

It is worth pointing out that Eq. (3.6) holds only when the measure used in the calculation of variance, i.e., the integral with Lebesgue measure, is the same as that in the ANOVA decomposition.

Remark 3.1. It could be extremely expensive to compute ANOVA decomposition for high-dimensional $f(\mathbf{Y})$. Therefore, Dirac measure is adopted instead of Lebesgue measure, i.e., $d\mu(\mathbf{Y}) = \delta(\mathbf{Y}-\mathbf{c})d\mathbf{Y}, \mathbf{c} \in \Gamma^d$. The special point $\mathbf{c}$ is termed *anchor point*. However, it is difficult to calculate anchor point $\mathbf{c}$ such that $f_0 = f(\mathbf{c}) = \overline{f(\mathbf{Y})}$. In this paper, we take anchor point $\mathbf{c}$ to be the mean of random variable $\mathbf{Y}$ as an approximation. In this case, the property (3.3) and (3.4) do not hold any more. Additional terms of ANOVA decomposition are needed to improve the accuracy of the mean.

The measure $d\mu(\mathbf{Y})$ determines the particular form of each component function, following the notation in [33]. We introduce a projection operator $\mathcal{P}_{\mathbf{s}}: \Gamma^d \to \Gamma^{|\mathbf{s}|}$
$$\mathcal{P}_{\mathbf{s}} f(\mathbf{Y}_{\mathbf{s}}) := \int_{\Gamma^{d-|\mathbf{s}|}} f(\mathbf{Y}) d\mu_{I\setminus\mathbf{s}}(\mathbf{Y}), \tag{3.7}$$
where $d\mu_{I\setminus\mathbf{s}} := \prod_{i\in I, i\notin\mathbf{s}} d\mu_i(Y_i)$. Therefore, each term $f_{\mathbf{s}}$ can be recursively defined by:
$$f_{\mathbf{s}}(\mathbf{Y}_{\mathbf{s}}) = \mathcal{P}_{\mathbf{s}} f(\mathbf{Y}_{\mathbf{s}}) - \sum_{\mathbf{t}\subset\mathbf{s}} f_{\mathbf{t}}(\mathbf{Y}_{\mathbf{t}}). \tag{3.8}$$

3.1 Adaptivity in ANOVA decomposition

When the nominal dimension of the stochastic problem increases, the computational complexity of the standard ANOVA becomes prohibitive to evaluate all of the terms. For example, for nominal dimension $d = 100$, the number of terms for second-order ANOVA decomposition needed to calculate is
$$1 + \binom{100}{1} + \binom{100}{2} = 5051.$$

Nevertheless, in many stochastic problems, most of the interactions among different dimensions are usually weak and have little contribution to the stochastic outputs. This

means that the active dimension of those stochastic problems is small. Therefore, some adaptive approaches can be employed to solve those problems efficiently without losing much accuracy.

There are many "adaptive" approaches, and the one we employ in this paper is obtained by replacing the nominal dimension by an active dimension, i.e., we modify Eq. (3.2) to be:

$$f(\mathbf{Y}) \approx f_0 + \sum_{j_1 \in \mathcal{F}_1} f_{j_1}(Y_{j_1}) + \sum_{(j_1,j_2)\in \mathcal{F}_2} f_{j_1,j_2}(Y_{j_1},Y_{j_2}) + \cdots + \sum_{(j_1,j_2,,j_\nu)\in \mathcal{F}_\nu} f_{j_1,j_2,\cdots,j_\nu}(Y_{j_1},Y_{j_2},\cdots Y_{j_\nu}). \tag{3.9}$$

In practice, $\mathcal{F}_\nu$ are the active dimensions for each subgroup. In the model problems we consider here we truncate the expansion with $\nu = 2$. We then adopt adaptivity criteria to determine the active dimensions $\mathcal{F}_1$ and $\mathcal{F}_2$. We describe the adaptive criteria based on variance weight as follow, see [40].

Adaptivity based on variance weight: Let $T_1 = \sum_{j=1}^{N} \sigma^2(f_j)$, which is the sum of the variance of all the first-order terms. Assume that the first-order terms are sorted such that $\sigma^2(f_j)$ is monotonically decreasing. The active dimension $\mathcal{F}_1$ should satisfy:

$$\sum_{f_j \in \mathcal{F}_1} \sigma^2(f_j) \geq pT_1, \tag{3.10}$$

where p is a proportionality constant with $0 < p < 1$ and very close to 1. This criterion is similar to the criterion used in [6], where $\sigma^2(f)$ instead of T_1 is used on the right-hand side of (3.10) and p is set to be 0.99. The set $\mathcal{F}_2$ can be found by computing

$$\eta_{j_1,j_2} = \frac{\sigma^2(f_{j_1,j_2})}{\sum_{f_j \in \mathcal{F}_1} \sigma^2(f_j)}, \tag{3.11}$$

and bounding η_{j_1,j_2} with a predefined error threshold θ_2.

Remark 3.2. When we employ the preceding criteria to applications of solving SPDEs, we replace the variance of component function f_j with their L_2 norm values on the physical domain.

4 An adaptive ANOVA-based data-driven stochastic method

For large dimension $d \gg 1$, the number of basis or collocation points required in many stochastic methods, such as the stochastic finite element method, Wiener chaos expansion method, gPC method, and stochastic collocation method, increases exponentially, which makes these methods become expensive or infeasible. This is the well-known challenge, *curse of dimensionality*. The DSM in collocation representation also has this

problem because the number of polynomial basis index (2.2) used in constructing the DSM basis increases fast if the dimension of the stochastic problem is large. Combining the advantages of both the ANOVA decomposition technique and the DSM, we propose an adaptive ANOVA-DSM to study high-dimensional SPDEs in the multi-query setting.

Let $\boldsymbol{\xi}(\omega)=(\xi_1(\omega),\cdots,\xi_d(\omega))$ denote a d-dimensional vector of independent and identically distributed random variables. We consider the following 1D SPDE to demonstrate the main idea:

$$-\nabla\cdot(a(x,\boldsymbol{\xi}(\omega))\nabla u(x,\boldsymbol{\xi}(\omega)))=f(x,\theta), \quad x\in D,\ \omega\in\Omega, \tag{4.1}$$

$$u(x,\boldsymbol{\xi}(\omega))=0, \quad x\in\partial D, \tag{4.2}$$

where the coefficient $a(x,\boldsymbol{\xi}(\omega)$ is assumed to be positive with upper and lower bounds almost surely. For notational simplicity, we denote the solution by $u(x,\boldsymbol{\xi})$ in the rest of this section. We summarize our new method as follows:

Adaptive ANOVA-DSM offline computation

1. Expand stochastic solution $u(x,\boldsymbol{\xi})$ in an adaptive ANOVA decomposition:

$$u(x,\boldsymbol{\xi})=u_0(x)+\sum_{j_1\in\mathcal{F}_1}u_{j_1}(x,\boldsymbol{\xi}_{j_1})+\sum_{(j_1,j_2)\in\mathcal{F}_2}u_{j_1,j_2}(x,\boldsymbol{\xi}_{j_1},\boldsymbol{\xi}_{j_2})+\cdots, \tag{4.3}$$

   where $\mathcal{F}_1$ and $\mathcal{F}_2$ are the first and second active dimension sets, respectively. See 6 for a detailed discussion about how to determine them.

2. Solve each first-order ANOVA decomposition subproblem with the DSM and obtain the stochastic basis matrix $\mathbf{A}^{j_1}$, $j_1\in\mathcal{F}_1$. For example, the stochastic basis matrix $\mathbf{A}^1$ is obtained by applying the DSM for the following subproblem:

$$-\nabla\cdot(a(x,\xi_1(\omega),c_2,\cdots,c_d)\nabla u(x,\xi_1(\omega),c_2,\cdots,c_d))=f(x,\theta), \quad x\in D,\ \omega\in\Omega, \tag{4.4}$$

$$u(x,\xi_1(\omega),c_2,\cdots,c_d)=0, \quad x\in\partial D, \tag{4.5}$$

   where c_k, $k=2,\cdots,d$ are anchor points.

3. For each second-order term $u_{j_1,j_2}(x,\boldsymbol{\xi}_{j_1},\boldsymbol{\xi}_{j_2})$, apply the DSM to solve the corresponding subproblem and obtain the stochastic basis matrix $\mathbf{A}^{j_1j_2}$, $(j_1,j_2)\in\mathcal{F}_2$.

4. Repeat the preceding steps, if needed, to calculate stochastic basis for a high-order ANOVA decomposition subproblem.

Remark 4.1. In practical computation when the variance of the stochastic solution is not large, only the adaptive first- and second-order terms are needed in ANOVA expansion to obtain accurate results for mean and variance. In calculating each subproblem of the ANOVA decomposition, several DSM basis could be enough for accuracy requirement. This essentially expedites our computation.

Adaptive ANOVA-DSM online computation

1. For each query $f(x,\theta)$, expand stochastic solution $u(x,\boldsymbol{\xi})$ in an adaptive ANOVA decomposition:

$$u(x,\boldsymbol{\xi}) \approx u_0(x) + \sum_{j_1 \in \mathcal{F}_1} u_{j_1}(x,\boldsymbol{\xi}_{j_1}) + \sum_{(j_1,j_2)\in \mathcal{F}_2} u_{j_1,j_2}(x,\boldsymbol{\xi}_{j_1},\boldsymbol{\xi}_{j_2}) + \cdots, \tag{4.6}$$

   where the sets $\mathcal{F}_1$ and $\mathcal{F}_2$ are selected according to the adaptive approach in the offline stage.

2. Solve for the mean term $u_0(\mathbf{x})$, which satisfies a deterministic equation by replacing random variables $\boldsymbol{\xi}$ with anchor point $\mathbf{c}$ in the stochastic PDE (4.4)-(4.5),

$$-\nabla\cdot(a(x,\mathbf{c})\nabla u_0(x)) = f(x,\theta), \quad x\in D, \tag{4.7}$$
$$u_0(x) = 0, \quad x\in \partial D. \tag{4.8}$$

3. For each high-order term $u_{j_1,j_2,\cdots,j_s}(x,\boldsymbol{\xi}_{j_1},\boldsymbol{\xi}_{j_2},\cdots,\boldsymbol{\xi}_{j_s})$, with our data-driven stochastic basis $\mathbf{A}^{j_1 j_2\cdots j_s}$, we use the standard stochastic Galerkin method to solve the SPDEs (4.4)-(4.5).

4. Use the ANOVA-DSM solutions to calculate the statistical information of the solution, such as mean and variance.

It should be pointed out that the construction of an ANOVA-DSM basis can be expensive if we solve the Eqs. (4.4)-(4.5) only once for a given forcing function. However, when we need to solve the same SPDE many times with multiple forcing functions, our online algorithm offers considerable computational savings because the number of elements in the stochastic basis is, in general, much smaller than the number of orthogonal polynomials in the polynomial chaos basis, especially when the input stochastic dimension is high.

5 Numerical examples

In this section, we perform a number of numerical experiments to test the performance and accuracy of the adaptive ANOVA-DSM for elliptic SPDEs with random coefficients. We demonstrate that the DSM can offer accurate numerical solutions to SPDEs with significant computational savings in the online stage over traditional stochastic methods, such as the ANOVA based stochastic collocation method (denoted by ANOVA-SC) [40]. All of the simulations and comparisons are conducted on a single computing node with 16GB of memory at the Caltech Center for Advanced Computing Research (CACR).

5.1 1D elliptic SPDE in 20-dimensional stochastic space

We consider the following 1D elliptic SPDE with random coefficient:

$$-\frac{\partial}{\partial x}\Big(a(x,\omega)\frac{\partial}{\partial x}u(x,\omega)\Big)=f(x), \quad x\in D=(0,1), \ \omega\in\Omega, \tag{5.1}$$

$$u(0,\omega)=0, \qquad u(1,\omega)=0. \tag{5.2}$$

When modeling a whole aquifer or a whole oil reservoir, the correlation length scale for random field $a(x,\omega)$ is significantly smaller than the size of the computational region. However, the correlation is typically large enough to fall outside the domain of stochastic homogenization techniques. In addition, typical sedimentation processes lead to fairly irregular structures and pore networks. Therefore, a faithful model should assume only limited spatial regularity of $a(x,\omega)$. A convenient way is to assume the covariance function of $\log(a(x,\omega))$ is given by:

$$C(x,y)=\sigma^2 e^{-\frac{|x-y|^p}{\lambda}}, \quad x,y\in[0,1]. \tag{5.3}$$

The parameters σ^2 and λ are the variance and the correlation length, respectively. In this paper, we choose $p=1$, $\sigma^2=1$, and $\lambda=0.1$.

We use the KLE of the covariance function $C(x,y)$ to produce samples of $a(x,\omega)$. Let $k(x,\omega)=\log(a(x,\omega))$. We expand $k(x,\omega)$ in term of a countable set of uncorrelated, zero mean random variables $\{\xi_n\}_{n=1}^{\infty}$ such that:

$$k(x,\omega)=\sum_{n=1}^{\infty}\sqrt{\theta_n}\xi_n(\omega)\phi_n(x), \tag{5.4}$$

where we assume $E[k(x,\omega)]=0$ and $\{\theta_n,\phi_n(x)\}_{n=1}^{\infty}$ are the eigenpairs of the covariance function (5.3). An important point to note is that for random field $k(x,\omega)$, the random variables $\{\xi_n\}_{n=1}^{\infty}$ are assumed to be independent and uniformly distributed within range $\xi_n\in(-\sqrt{3},\sqrt{3})$. In practice, we truncate the expansion (5.4) after a finite number K of terms and define the coefficient as:

$$a(x,\omega)=e^{\sum_{n=1}^{K}\sqrt{\theta_n}\xi_n(\omega)\phi_n(x)}. \tag{5.5}$$

We choose $K=20$ in (5.5), and select the anchor point $\mathbf{c}=\mathbf{0}$. Some typical samples of the elliptic coefficients are shown in Fig. 2. The function class of the right-hand side is chosen to be $\mathfrak{F}=span\{1,\sin(i\pi x),\cos(i\pi x)\}_{i=1}^{15}$. The finite element method is used for the spatial discretization with mesh size $h=\frac{1}{256}$. The mesh size of the coarse grid in preconditioning of the data-driven method (i.e., the Step 2 of DSM Offline Computation) will be chosen as $h_c=\frac{1}{64}$. Legendre polynomials are used in the data-driven stochastic basis representation. We choose $r=1$ and $p=10$ in the orthonormal basis index (2.2) to represent the DSM basis for first-order ANOVA decomposition subproblem. Similarly, we choose $r=2$ and $p=8$ for the second-order ANOVA decomposition subproblem. Due to the modest stochastic

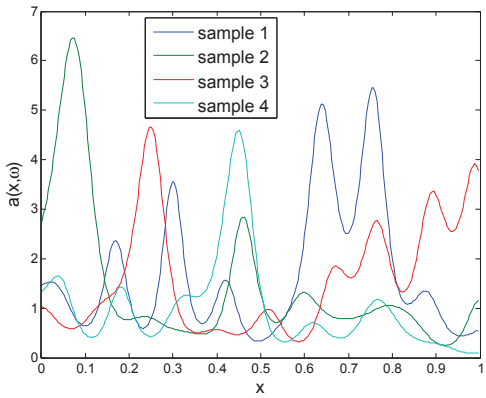

Figure 2: The random process with exponential covariance function.

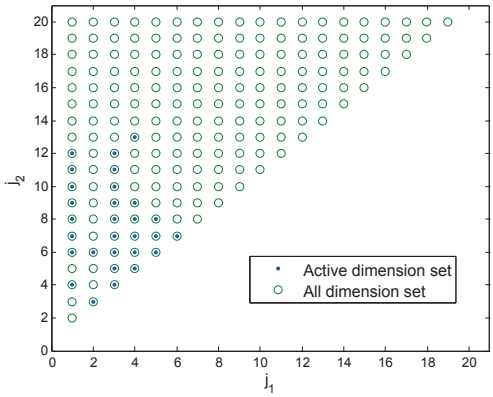

Figure 3: The active dimension set in 2nd ANOVA decomposition.

input dimension of (5.5), the "exact" solution is obtained by sparse grid based stochastic collocation method with quadrature level 5, which has 90561 quadrature points.

Let $u_{DSM}(x,\omega)$ denote the ANOVA-based data-driven solution, $u_{SC}(x,\omega)$ denote the ANOVA-based stochastic collocation solution, and $u(x,\omega)$ the exact solution. To quantify the error, we define the relative error of mean and STD (standard deviation) in $L^2(D)$. For instance, the error of the DSM solution can be defined as follows:

$$e_{mean}=\frac{||\bar{u}(x)-\bar{u}_{DSM}(x)||_{L^2(D)}}{||\bar{u}(x)||_{L^2(D)}},$$

and

$$e_{STD}=\frac{||STD(u)-STD(u_{DSM})||_{L^2(D)}}{||STD(u)||_{L^2(D)}}.$$

In Fig. 3, we show the second-order active dimension set $\mathcal{F}_2$ in the ANOVA decomposition obtained by the adaptive algorithm. Without adaptivity, the second-order ANOVA

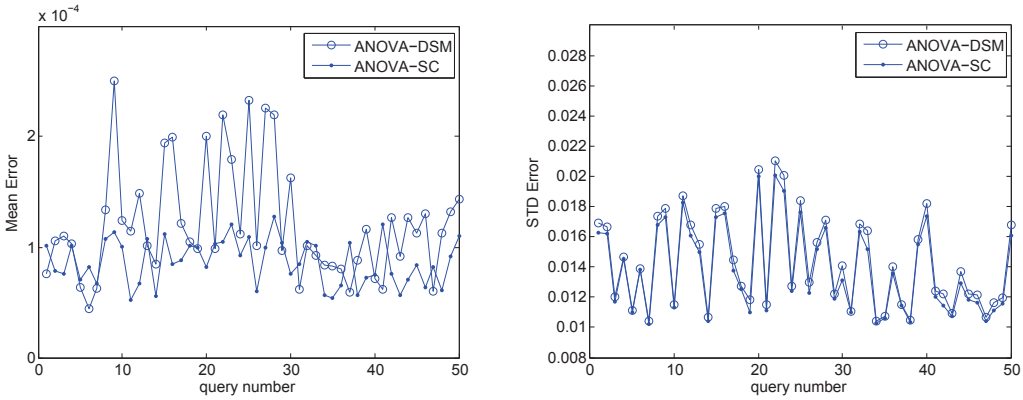

Figure 4: The multi-query results of the ANOVA-DSM and ANOVA-SC solvers.

decomposition will generate 190 subproblems. However, using our adaptive algorithm ($\eta = 0.99$), we only need to solve 29 subproblems.

We randomly generate 50 force functions, i.e., $f(x) \in \{a_i \sin(\pi k_i x) + b_i \cos(\pi l_i x)\}_{i=1}^{50}$, where a_i, b_i, k_i, and l_i are random numbers. In Fig. 4, we show the mean and STD comparisons in the online stage. We also show the computational cost of the ANOVA-based data-driven stochastic solver (denoted by ANOVA-DSM) and the ANOVA-based stochastic collocation solver (denoted by ANOVA-SC) in Fig. 11. Let n denote the query number. The computational cost of ANOVA-DSM consists of two parts. In the offline stage, it takes 30.951 seconds to train the data-driven stochastic basis. In the online stage, with our DSM basis, solving the SPDE (5.1) once will take 0.052 seconds. Therefore, the total cost will be $t_{DSM}(n) = 30.951 + 0.052n$. For the ANOVA-SC solver, we omit the cost of choosing the active dimension set in ANOVA decomposition and only consider the online stage cost. It will take 2.606 seconds to solve the SPDE (5.1) once with the adaptive

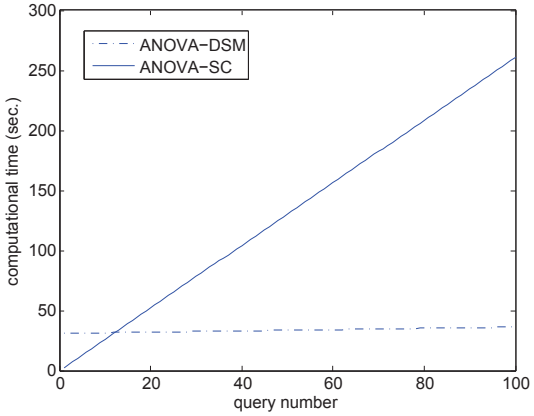

Figure 5: The computation time comparison.

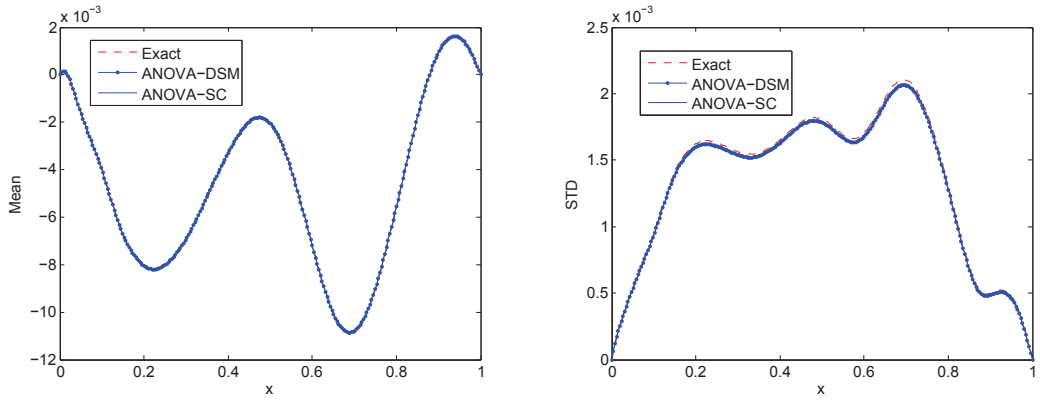

Figure 6: The snapshot solutions of mean and STD.

ANOVA-SC solver. Thus, the total cost will be $t_{SC}(n)=2.606n$. The ANOVA-DSM solution has the same accuracy as the ANOVA-SC solution. However, in a multi-query setting when we need to solve the SPDE (5.1) with more than 13 times, the ANOVA-DSM solver will be superior to the ANOVA-SC solver. Finally, in Fig. 6, we show the mean and STD of the solution corresponding to $f(x)=0.1\sin(8.6\pi x)+0.6\cos(4.3\pi x)$. It can be seen that the mean profile of the ANOVA-DSM and ANOVA-SC solutions match the exact solution quite well. However, the STD profile of the ANOVA-DSM and ANOVA-SC solutions is a little lower than the exact solution. Higher-order adaptive ANOVA decomposition can improve the accuracy of STD, but it requires extra computational cost.

5.2 1D elliptic SPDE in 50-dimensional stochastic space

We consider the 1D elliptic SPDE in high-dimensional stochastic space. The stochastic coefficient $a(x,\omega)$ of Eq. (5.1) now reads as:

$$a(x,\omega)=\sum_{i=1}^{50}C_i\xi_i(\omega)(\sin(D_i\pi x)+1), \tag{5.6}$$

where $\{\xi_i\}$ are independent uniform random variables in $[0,1]$ and $C_i\in(0,0.001)$ and $D_i\in(0,10)$ are randomly generated. We first select the anchor point $\mathbf{c}=\mathbf{0.5}$. We still use $\mathfrak{F}=span\{1,\sin(i\pi x),\cos(i\pi x)\}_{i=1}^{15}$ as right-hand side functions to train the DSM basis. The finite element method is used for the spatial discretization with mesh size $h=\frac{1}{256}$. The mesh size of the coarse grid in the preconditioning of the data-driven method will be chosen as $h_c=\frac{1}{64}$. Legendre polynomials are used in the data-driven stochastic basis representation. We choose $r=1$ and $p=10$ in the orthonormal basis index (2.2) to represent the DSM basis for the first-order ANOVA decomposition subproblem. Similarly, we choose $r=2$ and $p=6$ for the second-order ANOVA decomposition subproblem. Let u_{DSM} denote the solution obtained by the ANOVA-based DSM solver.

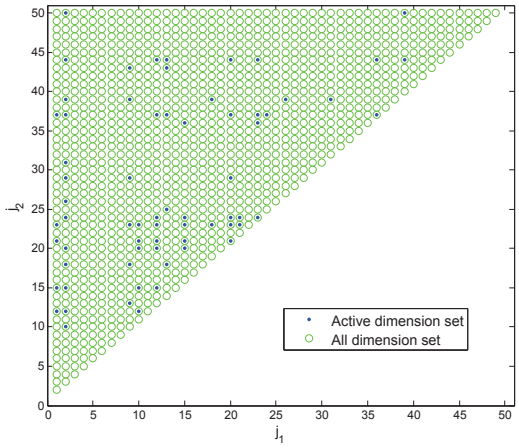

Figure 7: The active dimension set in the second-order ANOVA decomposition.

The gPC method and stochastic collocation method are computationally prohibitive due to the *curse of dimensionality*. Therefore, we use the Monte Carlo method with 10^6 samples to calculate the "exact" solution. The ANOVA decomposition of the SPDE (5.1) results in a set of low-dimensional subproblems in stochastic space. Thus, the ANOVA-based stochastic collocation method is an effective approach for solving high-dimensional SPDEs (see [5,40]). We also compare the ANOVA-SC solver's performance, as well as our ANOVA-DSM solver.

In Fig. 7, we show the second-order active dimension set $\mathcal{F}_2$ in the ANOVA decomposition obtained by the adaptive algorithm. Without adaptivity, the second-order ANOVA decomposition will generate 1225 subproblems. However, using our adaptive algorithm ($\eta=0.99$), we only need to solve 67 subproblems, which significantly reduces the computational cost.

We randomly generate 50 force functions, i.e., $f(x)\in\{a_i\sin(\pi k_i x)+b_i\cos(\pi l_i x)\}_{i=1}^{50}$, where a_i, b_i, k_i, and l_i are random numbers. In Fig. 8, we show the mean and STD comparison in the online stage. Let n denote the query number. In the offline stage of ANOVA-DSM, we spend 73.901 seconds to train the data-driven stochastic basis. In the online stage with our DSM basis, solving the SPDE (5.1) once will take 0.087 seconds. Therefore, the total cost will be $t_{DSM}(n)=73.901+0.087n$. For the ANOVA-SC solver, we omit the cost of choosing the active dimension set in the ANOVA decomposition and only consider the online stage cost. It will take 5.629 seconds to solve the SPDE (5.1) once with the adaptive ANOVA-SC solver. Thus, the total cost will be $t_{SC}(n)=5.629n$. Again, the ANOVA-DSM solution has the same accuracy as the ANOVA-SC solution. However, in a multi-query setting when we need to solve the SPDE (5.1) more than 14 times, the ANOVA-DSM solver will be superior to the ANOVA-SC solver. Finally, in Fig. 9, we show the mean and STD of the solution corresponding to $f(x)=0.2\sin(6.9\pi x)+3.3\cos(5.6\pi x)$. It is evident that the mean profile of the ANOVA-

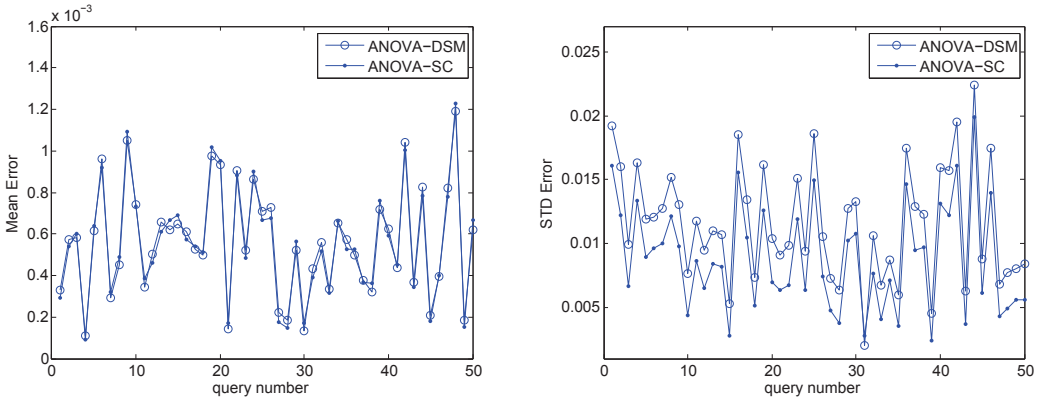

Figure 8: The multi-query results of ANOVA-DSM and ANOVA-SC solver.

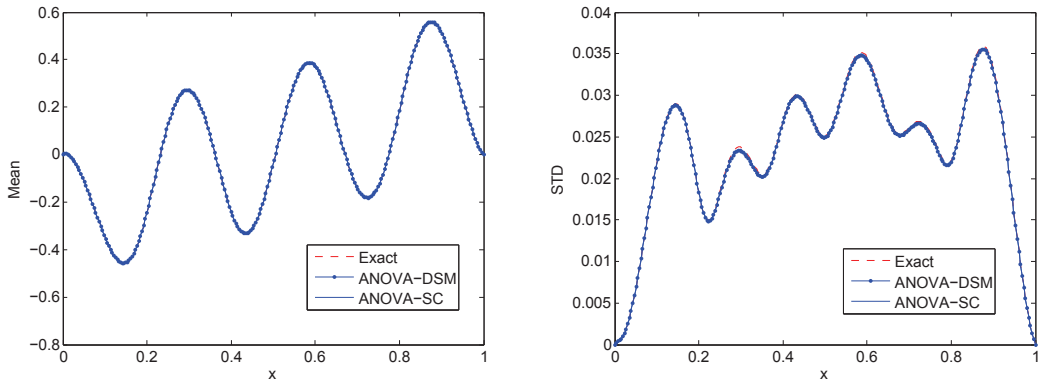

Figure 9: The snapshot solutions of mean and STD.

DSM and ANOVA-SC solutions match the exact solution quite well. However, the STD profile of ANOVA-DSM and ANOVA-SC solutions is a little lower than the exact solution. Higher-order adaptive ANOVA decompositions are needed to improve the accuracy of STD, but it requires additional computational cost.

5.3 2D elliptic SPDE in 20-dimensional stochastic space

We further apply our ANOVA-DSM to solve the following 2D stochastic elliptic problem with a random coefficient:

$$-\nabla\cdot(a(x,y,\omega)\nabla u(x,y,\omega))=f(x,y), \quad (x,y)\in D,\ \omega\in\Omega, \tag{5.7}$$

$$u(x,y,\omega)=0, \quad (x,y)\in\partial D,\ \omega\in\Omega, \tag{5.8}$$

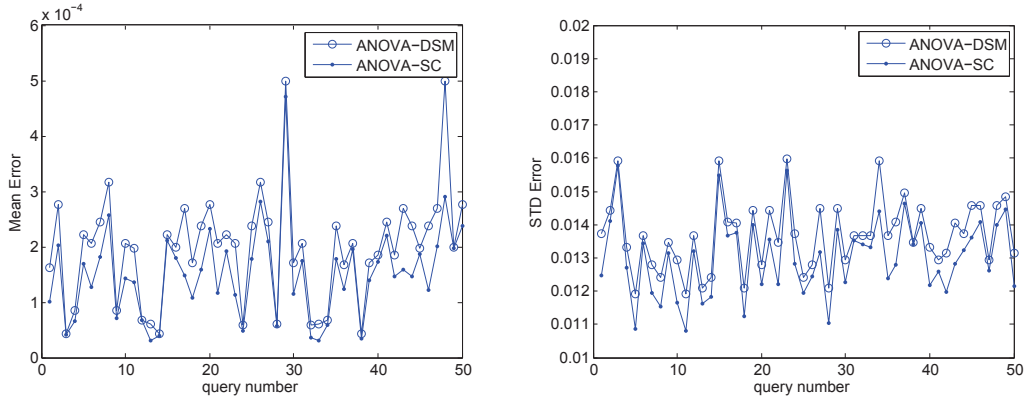

Figure 10: The multi-query results of ANOVA-DSM and ANOVA-SC solver.

where $D=[0,1]\times[0,1]$. The random coefficient is defined as:

$$a(x,y,\omega)=\sum_{i=1}^{20}\xi_i(\sin(i\pi x)\cos((21-i)\pi y)+1.001), \tag{5.9}$$

where $\{\xi_i\}$ are independent uniform random variables in $[0,1]$. We still select the anchor point $\mathbf{c}=\mathbf{0.5}$. In the offline stage, the function class of the right-hand side in the preconditioning DSM method is chosen to be $\mathfrak{F}=\{\sin(k_i\pi x+l_i\pi y)\cos(m_i\pi x+n_i\pi y)\}_{i=1}^{20}$, where k_i, l_i, m_i, and n_i are random numbers. We use this random training strategy to reduce the computational cost. The finite element method is used for the spatial discretization. We first partition the domain D into squares with mesh size $h=\frac{1}{128}$ then further partition them into triangular meshes. The mesh size of the coarse grid in the preconditioning of the data-driven method will be chosen as $h_c=\frac{1}{32}$. Legendre polynomials are used in the data-driven stochastic basis representation. We choose the same parameter as Section 5.1 in the orthonormal basis index (2.2) to represent the DSM basis for first- and second-order ANOVA decomposition subproblems. We use the Monte Carlo method with 10^5 samples to calculate the "exact" solution.

With our adaptive algorithm ($\eta=0.99$), we only need to solve 24 second-order ANOVA decomposition subproblems, which significantly reduces the computational cost. The second-order active set result is similar to Fig. 3 (not shown here). We randomly generate 50 force functions, i.e., $f(x)\in\{\sin(\pi k_i x+a_i)\cos(\pi l_i y+b_i)\}_{i=1}^{50}$, where k_i, l_i, a_i and b_i are random numbers. In Fig. 10, we show the mean and STD comparison in the online stage. Let n denote the query number. In this 2D problem, the offline stage of the ANOVA-DSM solver takes 4331.24 seconds. In the online stage with our DSM basis, solving the SPDE (5.7) once will take 26.23 seconds. Therefore, the total cost will be $t_{DSM}(n)=4331.24+26.23n$. For the ANOVA-SC solver, we omit the cost of choosing the active dimension set in the ANOVA decomposition and only consider the online stage cost. It will take 343.33 seconds to solve the SPDE (5.7) once with the adaptive ANOVA-SC solver. Thus, the total cost will be $t_{SC}(n)=343.33n$. The ANOVA-DSM solution has

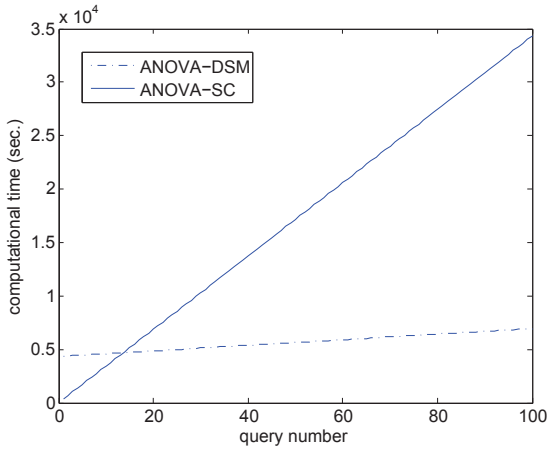

Figure 11: The computation time comparison.

the same accuracy as the ANOVA-SC solution. However, in a multi-query setting when we need to solve the SPDE (5.7) more than 14 times, the ANOVA-DSM solver will be superior to the ANOVA-SC solver. Finally, in Fig. 12, we show the mean and STD of the solution corresponding to $f(x,y)=\sin(6.8\pi x+0.1)\cos(4.3\pi y+0.2)$. It can be seen that the mean and STD of the ANOVA-DSM solution match the exact solution very well. The relative errors of the mean and STD are 0.032% and 1.34%, respectively.

Remark 5.1. The offline computing of ANOVA-DSM consists of two parts, choosing the active dimension set and training the data-driven stochastic basis. Since both the ANOVA-DSM and ANOVA-SC solver spend the same time to choose the active dimension set, we only consider the later part as the offline computational time of ANOVA-DSM.

6 Conclusion and discussion

In this paper, a novel adaptive, ANOVA-based, data-driven method has been developed to solve high-dimensional stochastic elliptic equations arising from various applications, such as the randomly heterogeneous porous media flow problem. The developed method has both an offline and online computation. In the offline computation, an adaptive ANOVA decomposition technique is applied to adaptively decompose the original high-dimensional problem into a set of low-dimensional sub-problems. By modeling the behavior of stochastic systems with only the first few lower-order terms of the high-dimensional input, adaptive ANOVA is able to efficiently represent the output response to the high-dimensional inputs with specified good accuracy. The adaptive ANOVA decomposition results in a set of low-dimensional subproblems in stochastic space, which can be efficiently solved by sparse grid based stochastic collocation method.

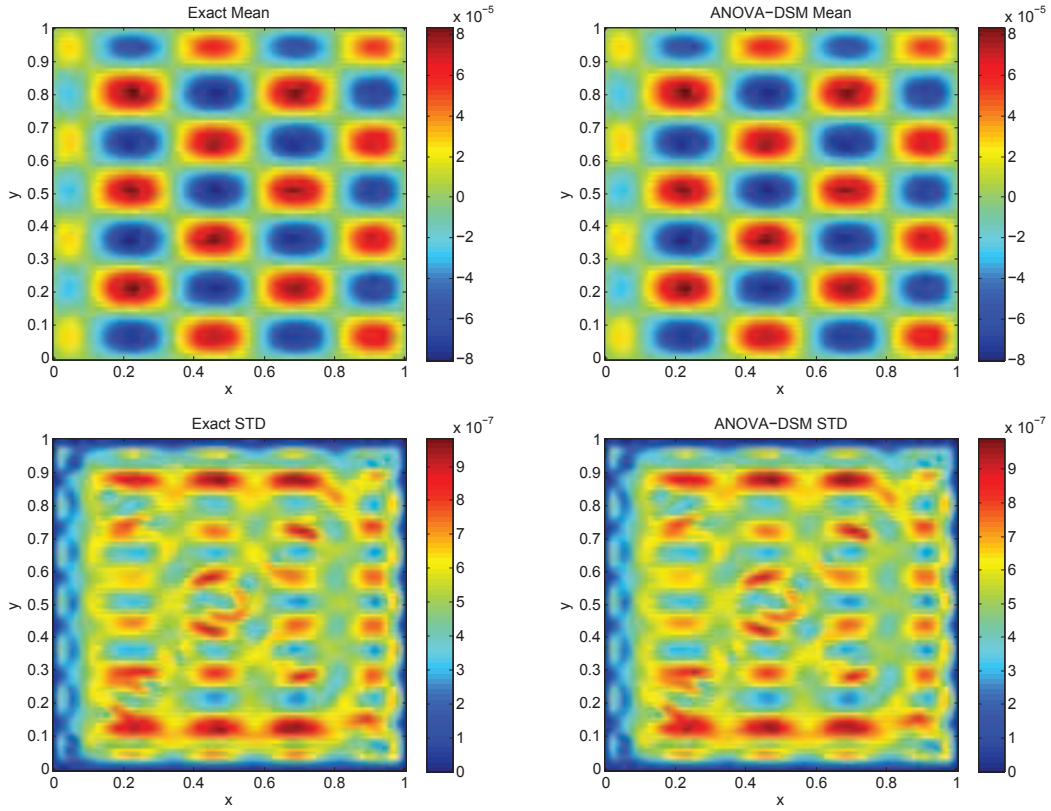

Figure 12: The snapshot solutions of mean and STD.

We adopt the DSM to improve the stochastic collocation method's slow convergence. In the offline stage, for each subproblem in the adaptive ANOVA decomposition, an optimal data-driven stochastic basis is obtained by the KLE of the covariance matrix of the stochastic output solutions computed by the stochastic collocation method. Multiple trial functions are used to enrich the stochastic basis and improve our method's accuracy. In the online computation, a Galerkin-projection-based method with the optimal data-driven basis developed in the offline part is employed, which greatly reduces the computational cost.

Numerical examples involving both 1D and 2D elliptic PDEs with high-dimensional random coefficients have been conducted to verify the accuracy and efficiency of the developed adaptive ANOVA-based DSM method. These numerical examples indicate the following three advantages of the proposed adaptive ANOVA-based DSM method: (1) by integrating with adaptive ANOVA decomposition, it can effectively solve stochastic problems within desired accuracy even for problems with high-dimensional stochastic inputs; (2) the optimal data-driven stochastic basis can be used for various deterministic forcing terms on the right-hand-side function of the elliptic PDE with random coef-

ficients; (3) in comparing with the classic numerical solver, such as the Monte Carlo or stochastic collocation methods, the ANOVA-DSM offers considerable computational savings when solving the same stochastic PDE many times with multiple forcing functions. We should point out that when the variance of the SPDE solution is large, the current version of the ANOVA-DSM method does not offer much computational savings as we need to include higher-order ANOVA decomposition subproblems to solve. Currently, we are adopting some model reduction ideas and developing a new version of DSM to handle this class of problems.

Acknowledgments

The work of Thomas Hou was supported in part by an AFOSR MURI project under Contract FA 9550-09-1-0613, a DOE Grant DE-FG02-06ER25727, and NSF FRG Grant DMS-1159138. The work of Guang Lin was supported by the U.S. Department of Energy (DOE) Office of Science Advanced Scientific Computing Research Applied Mathematics program. Pacific Northwest National Laboratory is operated by Battelle for the DOE under Contract DE-AC05-76RL01830.

Appendix: An effective algorithm to determine $\mathcal{F}_1$ and $\mathcal{F}_2$

In this appendix, we propose an adaptive approach to determine the first-order active dimension set $\mathcal{F}_1$ and second-order active dimension set $\mathcal{F}_2$ in the ANOVA decomposition. In this algorithm, the adaptivity is based on variance weight. A similar algorithm can be developed based on the mean weight, see [27].

Adaptive algorithm for $\mathcal{F}_1$

1. Set the error threshold $\eta = 0.99$. Choose a set of test functions $\{f_k(x)\}_{k=0}^K$, which can be the basis of $f(x,\theta)$ or decided by substituting randomly selected parameters $\{\theta_k\}_{k=0}^K$ into $f(x,\theta)$.
2. For each test function $f_k(x)$, expand stochastic solution $u(x,\boldsymbol{\xi};f_k)$ into the first-order ANOVA decomposition;

$$u(x,\boldsymbol{\xi};f_k) = u_0(x;f_k) + \sum_{j_1=1}^{d} u_{j_1}(x,\boldsymbol{\xi}_{j_1};f_k). \tag{A.1}$$

   Solve each first-order ANOVA decomposition subproblem with the stochastic collocation method or DSM to obtain the corresponding variance.
3. Let:

$$\sigma^2(f_k) = \sum_{j_1=1}^{d} \sigma_{j_1}^2(f_k)$$

and assume the first-order terms $\sigma^2_{j_1}(f_k)$ are sorted in descending order. The active dimension set $\mathcal{F}_1(f_k)$ can be found by finding a set with the minimal number of elements satisfied, such as in the following condition:

$$\sum_{j_1\in\mathcal{F}_1(f_k)}^{d}\sigma^2_{j_1}(f_k)\geq\eta\sigma^2(f_k).$$

The first-order active dimension set $\mathcal{F}_1$ can be chosen as the union of all the sets $\mathcal{F}_1(f_k)$, $k=0,\cdots,K$.

Adaptive algorithm for $\mathcal{F}_2$

1. Set the error threshold $\eta=0.99$. Choose a set of test functions $\{f_k(x)\}_{k=0}^K$, which can be the basis of $f(x,\theta)$ or decided by substituting randomly selected parameters $\{\theta_k\}_{k=0}^K$ into $f(x,\theta)$.

2. For each test function $f_k(x)$, expand stochastic solution $u(x,\boldsymbol{\xi};f_k)$ in the ANOVA decomposition:

$$u(x,\boldsymbol{\xi};f_k)=u_0(x;f_k)+\sum_{j_1\in\mathcal{F}_1}u_{j_1}(x,\boldsymbol{\xi}_{j_1};f_k)+\sum_{j_1<j_2,(j_1,j_2\in\mathcal{F}_1)}u_{j_1,j_2}(x,\boldsymbol{\xi}_{j_1},\boldsymbol{\xi}_{j_2};f_k),\qquad(A.2)$$

where $\mathcal{F}_1$ is obtained by the adaptive algorithm. Solve each first- and second-order ANOVA decomposition subproblem with the stochastic collocation method or DSM to obtain the corresponding variance.

3. Let:

$$\sigma^2(f_k)=\sum_{j_1\in\mathcal{F}_1}\sigma^2_{j_1}(f_k)+\sum_{j_1<j_2,(j_1,j_2\in\mathcal{F}_1)}\sigma^2_{j_1j_2}(f_k)$$

and assume the second-order terms are sorted in descending order. The active dimension set $\mathcal{F}_2(f_k)$ can be found by finding a set with the minimal number of elements satisfied, such as in the following condition:

$$\sum_{j_1\in\mathcal{F}_1}\sigma^2_{j_1}(f_k)+\sum_{(j_1,j_2)\in\mathcal{F}_2(f_k)}\sigma^2_{j_1j_2}(f_k)\geq\eta\sigma^2(f_k).$$

The second-order active dimension set $\mathcal{F}_2$ can be selected as the union of all the sets $\mathcal{F}_2(f_k)$, $k=0,\cdots,K$.

References

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

[2] M. Barrault, Y. Maday, N.C. Nguyen, and A.T. Patera. An 'empirical interpolation' method: application to efficient reduced basis discretization of partial differential equations. C.R. Acad. Sci. Paris, Series I, 339: 667-672, 2004.
[3] V. Barthelmann, E. Novak, and K. Ritter. High dimensional polynomial interpolation on sparse grids. Adv. Comput. Math., 12:273-288, 2000.
[4] M. Bieri and C. Schwab. Sparse high order FEM for elliptic SPDEs. Comput. Methods Appl. Mech. Engrg., 198:1149-1170, 2009.
[5] Y. Cao, Z. Chen, and M. Gunzburger. ANOVA Expansions and Efficient Sampling Methods for Parameter Dependent Nonlinear PDEs. Int. J. Numer. Analysis and Modeling, 6(2):256-273, 2009.
[6] R.E. Caflisch, W. Morokoff, and A. Owen. Valuation of mortgage-backed securities using brownian bridges to reduce the effective dimension. J. Comput. Finance, 1:27-46, 1997.
[7] M.L. Cheng, T.Y. Hou, M. Yan and Z.W. Zhang. A Data-driven Stochastic Method for elliptic PDEs with random coefficients. SIAM/ASA J. UQ, 1-1, 452-493, 2013.
[8] P. Dostert, Y. Efendiev, T. Y. Hou, and W. Luo, Coarse gradient Langevin algorithms for dynamic data integration and uncertainty quantification. J. Comput. Phys., 217:123-142, 2006.
[9] R. Fisher. Statistical Methods for Research Workers. Oliver and Boyd, 1925.
[10] J. Y. Foo and G. E. Karniadakis. Multi-element probabilistic collocation in high dimensions. J. Comput. Phys., 229:1536-1557, 2009.
[11] B. Ganapathysubramanian and N. Zabaras. Sparse grid collocation schemes for stochastic natural convection problems. J. Comput. Phys., 225(1):652-685, 2007.
[12] Z. Gao and J.S. Hesthaven. On ANOVA Expansions and Stragies for Choosing the Anchor Point. App. Math. Comp., 217(7):3274-3285, 2009.
[13] T. Gerstner and M. Griebel. Dimension-adaptive tensor-product quadrature. Computing, 71(1):65-87, 2003.
[14] R.G. Ghanem and P.D. Spanos. Stochastic Finite Elements: A Spectral Approach. New York : Springer-Verlag, 1991.
[15] M. Griebel. Adaptive sparse grid multilevel methods for elliptic pdes based on finite differences. Computing, 61(2):151C180, 1998
[16] M. Griebel. Sparse grids and related approximation schemes for higher dimensional problems. Proceedings of the conference on Foundations of Computational Mathematics, Santander, Spain, 2005.
[17] W. Hoeffding. A class of statistics with asymptotically normal distributions. Annals of Math. Statist., 19:293C325, 1948.
[18] N. Halko, P.G. Martinsson and J. Tropp. Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions. SIAM Review, 53(2):217-288, 2011
[19] T.Y. Hou, W. Luo, B. Rozovskii, and H. Zhou. Wiener Chaos expansions and numerical solutions of randomly forced equations of fluid mechanics. J. Comput. Phys., 216(2):687-706, 2006.
[20] K. Karhunen. Uber lineare Methoden in der Wahrscheinlichkeitsrechnung. Ann. Acad. Sci. Fennicae. Ser. A. I. Math.-Phys. 37:1-79, 1947.
[21] G. Lin, C.H. Su, and G.E. Karniadakis. Predicting shock dynamics in the presence of uncertainties. J. Comp. Phys., 217:260C276, 2006.
[22] G. Lin and G.E. Karniadakis. Sensitivity analysis and stochastic simulations of non-equilibrium plasma flow. Int. J. Numer. Meth. Engng, 80(6-7):738C766, 2009.
[23] G. Lin, C.H. Su, and G. E. Karniadakis. Random roughness enhances lift in supersonic flow.

Physical Review Letters, 99:104501, 2007.
[24] G. Lin, C.H. Su, and G. E. Karniadakis. Stochastic modeling of random roughness in shock scattering problems: Theory and simulations. Comput. Methods Appl. Mech. Engrg., 197:3420C3434, 2008.
[25] M. Loeve. Probability theory. Vol. II, 4th ed. GTM. 46. Springer-Verlag. ISBN 0-387-90262-7.
[26] X. Ma and N. Zabaras. An adaptive hierarchical sparse grid collocation algorithm for the solution of stochastic differential equations. J. Comput. Phys., 228(8):3084-3113, 2009.
[27] X. Ma and N. Zabaras. An adaptive high-dimensional stochastic model representation technique for the solution of stochastic partial differential equations. J. Comp. Phys., 229:3884C3915, 2010.
[28] O. Le Maitre. Uncertainty propagation using Wiener-Haar expansions. J. Comput. Phys., 197(1):28-57, 2004.
[29] H.N. Najm. Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics. Ann. Rev. Fluid Mech., 41(1):35-52, 2009.
[30] E. Novak and K. Ritter. High dimensional integration of smooth funcions over cubes. Numer. Math., 75:79C97, 1996.
[31] E. Novak and K. Ritter. Simple cubature formulas with high polynomial exactness. Constr. Approx., 15:499C522, 1999.
[32] B. K. Oksendal. Stochastic Differential Equations: An Introduction with Applications. Sixth edition. Berlin: Springer. 2003.
[33] H. Rabitz and O. F. Alis. General foundations of high-dimensional model representations. J. Math. Chem., 25:197C233, 1999.
[34] M. Tatang and G. McRae. Direct treatment of uncertainty in models of reaction and transport. Technical report, MIT Tech. Rep., 1994.
[35] X. Wan and G.E. Karniadakis. An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. J. Comput. Phys., 209(2):617-642, 2005.
[36] C.L. Winter, A. Guadagnini, D. Nychka, and D.M. Tartakovsky. Multivariate sensitivity analysis of saturated flow through simulated highly heterogeneous groundwater aquifers. J. Comput. Phys., 217:166C175, 2009.
[37] D. Xiu and G.E. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. SIAM J. Sci. Comput., 24(2):614-644, 2002.
[38] D. Xiu and G.E. Karniadakis. Modeling uncertainty in flow simulations via generalized polynomial chaos. J. Comput. Phys., 187(1):137-167, 2003.
[39] D. Xiu and J.S. Hesthaven. High-order collocation methods for differential equations with random inputs. SIAM J. Sci. Comput., 27(3):1118-1139, 2005.
[40] Xiu Yang, Minseok Choi, Guang Lin, and George Em Karniadakis. Adaptive anova decomposition of stochastic incomressible and comressible flows. J. Comp. Phys., 231(4):1587-1614, 2012.
[41] H. Xu, S. Rahman, A generalized dimensionCreduction method for multidimensional integration in stochastic mechanics, Int. J. Numer. Methods Eng. 61 (12) (2004) 1992-2019.
[42] Z. Zhang, M. Choi, G.E. Karniadakis, Anchor points matter in anova decomposition, in: Spectral and High Order Methods for Partial Differential Equations Lecture Notes in Computational Science and Engineering, Springer, 76:347-355, 2011.