International Journal for Uncertainty Quantification, 4 (6): 455–477 (2014)

ANALYSIS OF VARIANCE-BASED MIXED MULTISCALE FINITE ELEMENT METHOD AND APPLICATIONS IN STOCHASTIC TWO-PHASE FLOWS

Jia Wei,[1,2] Guang Lin,[2,3,] Lijian Jiang,[4] & Yalchin Efendiev[1]*

[1] *Department of Mathematics, Texas A&M University, College Station, Texas 77840, USA*

[2] *Computational Science & Mathematics Division, Pacific Northwest National Laboratory, Richland, Washington 99352, USA*

[3] *Department of Mathematics, School of Mechanical Engineering, Purdue University, West Lafayette, Indiana, USA*

[4] *College of Mathematics and Econometrics, Hunan University, China*

Original Manuscript Submitted: 09/03/2012; Final Draft Received: 04/29/2014

The stochastic partial differential systems have been widely used to model physical processes, where the inputs involve large uncertainties. Flows in random and heterogeneous porous media is one of the cases where the random inputs (e.g., permeability) are often modeled as a stochastic field with high-dimensional random parameters. To treat the high dimensionality and heterogeneity efficiently, model reduction is employed in both stochastic space and physical space. An analysis of variance (ANOVA)-based mixed multiscale finite element method (MsFEM) is developed to decompose the high-dimensional stochastic problem into a set of lower-dimensional stochastic subproblems, which require much less computational complexity and significantly reduce the computational cost in stochastic space, and the mixed MsFEM can capture the heterogeneities on a coarse grid to greatly reduce the computational cost in the spatial domain. In addition, to enhance the efficiency of the traditional ANOVA method, an adaptive ANOVA method based on a new adaptive criterion is developed, where the most active dimensions can be selected to greatly reduce the computational cost before conducting ANOVA decomposition. This novel adaptive criterion is based on variance-decomposition method coupled with sparse-grid probabilistic collocation method or multilevel Monte Carlo method. The advantage of this adaptive criterion lies in its much lower computational overhead for identifying the active dimensions and interactions. A number of numerical examples in two-phase stochastic flows are presented and demonstrate the accuracy and performance of the adaptive ANOVA-based mixed MsFEM.

KEY WORDS: *analysis of variance, uncertainty quantification, polynomial chaos, mixed multiscale finite element method, two-phase flow, stochastic partial differential equation, adaptivity*

1. INTRODUCTION

An important challenge in modeling flows in porous media is the treatment of complex heterogeneities and uncertainties in the permeability field. The high and low permeability may be connected at different scales. The uncertainty may arise from measurement corruption and incomplete knowledge of the physical properties. One way to describe the uncertainty is to model the permeability as a random field, which is often experimentally determined by a covariance function. Stochastic partial differential equations (SPDEs) are often used in modeling complex physical and engineering systems with uncertainties, which are usually characterized by a random field with high-dimensional parameters.

*Correspond to Guang Lin, E-mail: guanglin@purdue.edu

2152–5080/14/$35.00

To simultaneously tackle the high dimensionality and the heterogeneities, the analysis of variance (ANOVA)-based mixed multiscale finite element method (MsFEM) is employed.

Sampling in high-dimensional random space is very difficult. If the sampling of random space is conducted in full random space, then the number of samples increases drastically with respect to the dimensions of the random space. This is the notorious *curse of dimensionality*, which poses great difficulties for the stochastic approximation in a high-dimensional stochastic space. Instead of dealing with the full high-dimensional random space, the ANOVA representation can decompose a high-dimensional model into a set of low-dimensional models [1–7]. It was first introduced by Fisher in [2]. The decomposition is motivated by observing that there are dominant dimensions and significant dimensional interactions in many practical physical systems. In this case, the system accuracy will not be harmed too much if only the dominant dimensions and the significant dimensional interactions are taken into account. The ANOVA decomposition then splits a high-dimensional stochastic model into a set of low-dimensional stochastic submodels, which need much less computational effort. The curse of dimensionality can be considerably suppressed using the ANOVA approach.

For each low-dimensional subproblem, a sparse grid probabilistic collocation method (PCM) can be employed. The PCM was first introduced in [8] and has been studied extensively over the years [9]. The efficiency of the Clenshaw-Curtis-based sparse grid PCM was demonstrated by comparing it with other stochastic methods on an elliptic problem in [10]. In [11], an adaptive hierarchical sparse grid collocation algorithm was developed. In [12–14], a multielement PCM was employed to study the random roughness problem, stochastic compressible flow, and plasma flow problems.

At each collocation point, deterministic flow equations in porous media are solved; see examples in [15–17]. To treat the heterogeneity of porous media and recover the mass conservative velocity field, the mixed MsFEM [18–20] is employed. The main idea of the mixed MsFEMs is to incorporate the small-scale information into finite element basis functions and couple them through a global mixed formulation of the problem. The mixed MsFEMs share some similarities with a number of multiscale numerical methods such as the multiscale finite volume method [21], residual-free bubbles [22, 23], two-scale conservation subgrid method [24], variational multiscale method [25], and multiscale mortar method [26].

In this paper, a new approach is developed by combining ANOVA decomposition, sparse-grid based PCM for moderate dimensions and multilevel Monte Carlo [27] for high dimensions, and mixed MsFEM for efficiently solving stochastic two-phase flow equations. The combination of these model reduction techniques serves as a remedy to handle the large-scale problems in both stochastic and spatial spaces. However, because the ANOVA decomposition has a large number of terms if the number of random parameters is high, the total computational cost can still be prohibitive. To improve the efficiency and reduce computational efforts, an adaptive ANOVA technique based on a new adaptive criterion is developed. The advantage of the proposed adaptive technique is to identify the active random dimensions with respect to the function of interest through variance decomposition on a generalized polynomial chaos (gPC) expansion of the function of interest with relatively small computation efforts. The proposed adaptive ANOVA technique is different from the previous adaptive ANOVA methods [5, 6, 28]. In particular, it was noted that some adaptive ANOVA techniques, in both spatial and random spaces, were developed in [5] in the framework of the heterogeneous multiscale method. This work focuses on analyzing the errors introduced by ANOVA decomposition and mixed MsFEM, respectively. It is critical to understand the behavior of the two error contributions and find a good trade-off between them to achieve accurate results with relatively low computational cost.

In our numerical examples, the permeability fields with different statistical properties and heterogeneous structure are considered. The reference solution is computed on the fine grid and using the Monte Carlo (MC) method to sample the random space. Comparison of the solutions using ANOVA-based mixed finite element method (FEM) and ANOVA-based mixed MsFEM with the reference solution is studied. From these computations, errors from both ANOVA truncation and mixed MsFEM discretization are reported, respectively. It has also been observed that, in our numerical examples, the error introduced by the mixed MsFEM method is the dominant error source. Analyzing different functions is also helping us to better understand the effectiveness of ANOVA-based mixed MsFEM for different quantities of interest in oil reservoir simulations. The novel adaptive ANOVA-based MsFEM method is numerically comparable with existing adaptive ANOVA techniques with less online computations. Our methods are developed based on both ANOVA for dimension reduction, and variance-decomposition-based active dimension identification

for adaptivity. Sparse grid PCM for moderate random dimensions and multilevel Monte Carlo [27] for high random dimensions are employed to generate samples for active dimension identification. There are limitations for our developed adaptive ANOVA-MsFEM methods. For example, the accuracy of our methods depends on the accuracy of PCM or multilevel MC.

The rest of the paper is organized as follows. In Section 2, the background for two-phase flow in stochastic fields is presented. Section 3 is devoted to the description of mixed MsFEM and PCM. In Section 4, ANOVA-based and adaptive ANOVA-based mixed MsFEM methods are introduced. In Section 5, the numerical results using the methods introduced in Section 4 are presented for flow in different random porous media. Finally, some comments and conclusions are drawn in Section 6.

2. BACKGROUND AND NOTATIONS

Let D be a convex bounded domain in $\mathbb{R}^d$ ($d = 2, 3$) and $(\Omega, \mathcal{F}, P)$ be a probability space, where Ω is the set of outcomes, $\mathcal{F}$ is the σ algebra generated by Ω, and P is a probability measure.

Two-phase flow in random porous media is considered under the assumption that the displacement is dominated by viscous effects, with the effects of gravity, compressibility, and capillary pressure neglected. Porosity is considered to be constant. Here, the two phases will be referred to as water and oil, designated by subscripts w and o, respectively. The Darcy's law for each phase can be written as

$$v_j = -\frac{k_{rj}(S)}{\mu_j} k(x, \omega) \nabla p, \tag{1}$$

where v_j is the phase velocity, k is the permeability tensor, k_{rj} is the relative permeability to phase j ($j = o, w$), S is the water saturation, and p is the pressure. Combining Darcy's law with a statement of conservation of mass allows us to express the governing equations in terms of pressure and saturation equations

$$-\text{div}\big(\lambda(S) k(x, \omega) \nabla p\big) = Q_s, \tag{2}$$

$$\frac{\partial S}{\partial t} + v \cdot \nabla f(S) = 0, \tag{3}$$

where λ is the total mobility, Q_s is a source term, f is the fractional flux of water, and v is the total velocity, which are, respectively, given by

$$\lambda(S) = \frac{k_{rw}(S)}{\mu_w} + \frac{k_{ro}(S)}{\mu_o}, \quad f(S) = \frac{k_{rw}(S)/\mu_w}{k_{rw}(S)/\mu_w + k_{ro}(S)/\mu_o},$$

$$v = v_w + v_o = -\lambda(S)\, k \cdot \nabla p.$$

From Eqs. (1)–(3), the following stochastic two-phase flow system is formulated: find random fields $p(x, \omega) : \bar{D} \times \Omega \longrightarrow \mathbb{R}$, $v(x, \omega) : \bar{D} \times \Omega \longrightarrow \mathbb{R}$ and $S(x, \omega, t) : \bar{D} \times \Omega \times [0, T] \longrightarrow \mathbb{R}$ such that they almost surely (a.s) satisfy the following equations:

$$\begin{cases} \text{div}\big(v(x; \omega)\big) = Q_s \\ v(x; \omega) = -\lambda(S) k(x; \omega) \nabla p(x; \omega) \\ \dfrac{\partial S(x, t; \omega)}{\partial t} + v(x; \omega) \cdot \nabla f(S(x, t; \omega)) = 0. \end{cases} \tag{4}$$

Let the coefficient $k(x, \omega)$ of Eq. (4) be a stochastic field with second moment. To make $k(x, \omega)$ positive, $k(x, \omega)$ is considered to be a logarithmic stochastic field, i.e., $k(x, \omega) := \exp(a(x, \omega))$. Here, $a(x, \omega)$ is a stochastic field and its covariance function $\text{cov}[a] : \bar{D} \times \bar{D} \longrightarrow R$ of $a(x, \omega)$ is given by

$$\text{cov}[a](x_1, x_2) = \text{cov}[a(x_1), a(x_2)] = E[(a(x_1) - E[a(x_1)])(a(x_2) - E[a(x_2)])].$$

For computation, truncated Karhunen-Loève expansion (KLE) [29, 30] is employed to parametrize $a(x,\omega)$ by a finite-dimensional random field $a(x,Y)$. The function cov$[a]$ induces an integral operator $T_a : L^2(D) \longrightarrow L^2(D)$ by

$$T_a g(\cdot) = \int_D \text{cov}[a](x,\cdot)g(x)dx \;\; \forall g \in L^2(D).$$

The operator T_a is compact and self-adjoint. Consequently, there exist the eigenpairs $(\lambda_m, b_m(x))_{m\geq 1}$ of T_a such that

$$(b_i,b_j)_{L^2(D)} = \delta_{ij}, \;\; \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \cdots, \;\; \lim_{m\longrightarrow\infty} \lambda_m = 0,$$

where $(\cdot,\cdot)_{L^2}$ is the usual L^2 inner product. Define the mutually uncorrelated random variables by

$$y_i(\omega) := \frac{1}{\sqrt{\lambda_i}} \int_D (a(x,\omega) - E[a](x))b_i(x)dx, \;\; i = 1,2,\cdots.$$

Then, it follows the KLE of $a(x,\omega)$. Here, $a(x,\omega)$ is assumed to admit the following truncated KLE, i.e.,

$$a(x,Y) = E[a] + \sum_{i=1}^{n} \sqrt{\lambda_i} b_i(x) y_i(\omega),$$

where $Y := (y_1, y_2, \cdots, y_n) \in \mathbb{R}^n$. By the truncated KLE, $k(x,\omega) \approx k(x,Y) = \exp\big(a(x,Y)\big)$. In this paper, it was assumed that the stochastic field $k(x,\omega)$ can be accurately parametrized by $k(x,Y)$.

3. MIXED MSFEM, SPARSE GRID COLLOCATION, AND MULTILEVEL MONTE CARLO

In this section, mixed MsFEM for spatial discretization and sparse grid collocation and multilevel Monte Carlo [27] for discretization of stochastic space are presented.

3.1 Mixed MsFEM

The mixed multiscale finite element methods (MsFEM) used for spatial discretization is introduced in this section. To this end, a second-order elliptic equation is considered,

$$\begin{cases} -\text{div}(k\nabla p) = f \quad \text{in} \quad D \\ -k\nabla p \cdot n = g \quad \text{on} \quad \partial D. \end{cases} \tag{5}$$

Equation (5) describes the single-phase flow equation in porous media. The p refers to pressure, f refers to source (well or sink), and velocity $v = -k\nabla p$. Let $(\cdot,\cdot)$ denote the usual L^2 inner product. For mixed formulation, Eq. (5) is rewritten as

$$\begin{cases} k^{-1}v + \nabla p = 0 \;\; \text{in } D \\ \text{div}(v) = f \;\; \text{in } D \\ v \cdot n = g(x) \;\; \text{on } \partial D. \end{cases} \tag{6}$$

The weak formulation of Eq. (6) reads: seek $(v,p) \in H(\text{div}, D) \times L^2(D)/R$ such that $v \cdot n = g$ and

$$\begin{cases} (k^{-1}v, u) - \big(\text{div}(u), p\big) = 0 \quad \forall u \in H_0(\text{div}, D) \\ \big(\text{div}(v), q\big) = (f,q) \quad \forall q \in L^2(D). \end{cases}$$

Let $V_h \subset H(\text{div}, D)$ and $Q_h \subset L^2(D)/R$ be the finite element spaces for velocity and pressure, respectively. Define $V_h^0 = V_h \cap H_0(\text{div}, D)$. The numerical mixed formulation is to find $(v_h, p_h) \in V_h \times Q_h$ such that $v_h \cdot n = g_h$ on ∂D and

$$\begin{cases} \big(k^{-1}v_h, u_h\big) - \big(\text{div}(u_h), p_h\big) = 0 \quad \forall u_h \in V_h^0 \\ \big(\text{div}(v_h), q_h\big) = (f, q_h) \quad \forall q_h \in Q_h. \end{cases} \tag{7}$$

The mixed MsFEM for Eq. (7) is employed. It means that the mixed finite element approximation is performed on the coarse grid, where the finite element basis functions are defined. In the mixed MsFEM, the piecewise constant basis functions are applied on a coarse grid for pressure. For the velocity, the multiscale velocity basis functions are defined. The degree of freedom of the multiscale velocity basis function is defined on the interface of the coarse grid. Let e_i^K be a generic edge or face of the coarse block K. The multiscale basis equation associated with e_i^K is defined by

$$\begin{cases} -\text{div}(k\nabla w_i^K) = \dfrac{1}{|K|} \quad \text{in } K \\ -k\nabla w_i^K \cdot n = \begin{cases} b_i^K \ \text{ on } e_i^K \\ 0 \ \text{ else.} \end{cases} \end{cases}$$

For local mixed MsFEM [19], $b_i^K = 1/|e_i^K|$. If the media demonstrate strong nonlocal features including channels, fracture, and shale barriers, some global information is needed to define the boundary condition b_i^K for better accuracy of approximation [31, 32]. Then, $\psi_i^K = -k\nabla w_i^K$ defines the multiscale velocity basis function associated with e_i^K, and the multiscale finite dimensional space for velocity is defined by

$$V_h = \bigoplus_{K,i} \psi_i^K.$$

For stochastic systems (4), it is crucial to obtain the statistic properties (e.g., mean and variance) of solutions efficiently. In this paper, our discussion focuses on mean and variance, as they are fundamental for obtaining approximations of higher order moments. These properties could be obtained by sparse-grid PCM for moderate number of random dimensions or by multilevel MC method [27] and analyzing the corresponding results.

3.2 Sparse Grid PCM

Sparse grids have been successfully applied to PCM in many recent works, e.g., [15–17, 33, 34]. Based on the Smolyak formula [35], a set of Clenshaw-Curtis collocation points is chosen. With these chosen collocation points and corresponding weights, the statistic properties of the solutions can be obtained. For instance, assume that $\{y^{(j)}\}$ is the set of collocation points and $\{w^{(j)}\}$ is the corresponding weights, $j = 1, \cdots, N_c$. At each of the collocation points, the deterministic system is solved and the output $S(x, y^{(j)})$ is obtained. Then, the moments of $S(x, Y)$ can be estimated, e.g.,

$$E[S(x,t;Y)] = \int_\Omega S(x,t;\xi)dF(\xi) \approx \sum_{j=1}^{N_c} S(x,t;y^{(j)})w^{(j)},$$

$$\sigma^2\big(S(x,t;Y)\big) = \int_\Omega \big(S(x,t;\xi) - E[S(x,t;Y)]\big)^2 dF(\xi) \approx \sum_{j=1}^{N_c} S^2(x,t;y^{(j)})w^{(j)} - E^2[S(x,t;Y)].$$

4. MULTILEVEL MONTE CARLO METHOD

The multilevel Monte Carlo (MMC) method [27] can greatly reduce computational complexity through the use of a multilevel approach that combines results obtained using two levels of timestep while reducing the variance. The MMC method uses a geometric sequence of timesteps similar to the multigrid method and is proven to be efficient and reliable in achieving the desired accuracy.

Oftentimes in statistics rules are created to quantify measured data; these rules are referred to as estimators. In the standard MC method the estimator of an expected mean value is

$$\hat{P}_N^{MC} = \frac{1}{N}\sum_{i=1}^{N} P^i, \tag{8}$$

where N is the number of sample points and P^i is the expected value of each sample point.

Assuming that the discretized expected value approaches the exact solution for some random coefficient, as the discretization is refined, the root-mean-square error (RMSE) is given by

$$e(\hat{P}_M)^2 = V[\hat{P}^{MC}_{m_l,N}] + (E[P_{m_l} - P])^2 . \tag{9}$$

The first term represents the sampling error, $V[\hat{P}^{MC}_{m_l,N}] = N^{-1}V[P_{m_l}]$ is the variance, and the second term is the discretization error. To achieve RMSE less than ϵ both terms need to be less than $\epsilon^2/2$, which translates to $N \geq \epsilon^{-2}$ for the first term (large number of samples) and $m_l \geq \epsilon^{-1/\alpha}$, where α is the discretization convergence rate for the second term (high discretization).

The MMC method is based on a multilevel approach, which is used in the solution of elliptic partial differential equations. The estimator for the MMC comes from the same random sample but at different refinement levels. Then the multilevel estimator $\hat{P}^{ML}_M$ is

$$\hat{P}^{ML}_{m_l} = \sum_{l=0}^{L} \hat{Y}_l, \tag{10}$$

where

$$\hat{Y}_l = \frac{1}{N_l} \sum_{i=1}^{N_l} (P^i_{m_l} - P^i_{m_{l-1}}), \tag{11}$$

and $P^i_{m_{-1}} = 0$. Because the expectations, $\hat{Y}_l$, are estimated independently for each level, the multilevel estimator variance is

$$V[\hat{P}^{ML}_{m_l}] = \sum_{l=0}^{L} N^{-1}V[Y_l]. \tag{12}$$

Following Eq. (9), the error in multilevel MC is expressed as

$$e(\hat{P}^{ML}_M)^2 = \sum_{l=0}^{L} N^{-1}V[Y_l] + (E[P_{m_l} - P])^2 . \tag{13}$$

The multilevel variance $V[Y_l] = V[P_{m_l} - P_{m_l-1}] \to 0$ as $l \to \infty$ and if the variance is decreasing, less sample data will be needed; consequently, $N_l \to 1$ as $l \to \infty$. The cost at the coarsest level is fixed for all levels of accuracy. Achieving an overall RMSE of ϵ with MMC is easier than that achieved with the standard MC methods.

The computational cost for the multilevel and standard MC at each discretization level, l, is proportional to the refinement multiplier $\propto M^2$. A refinement multiplier of 2 would double the discretization level from level l to $l-1$. The costs are given as

$$C^{ML} = N_o + \sum_{l=1}^{L} N_l(M^l + M^{l-1}), \tag{14}$$

$$C^{MC} = \sum_{l=0}^{L} 2\epsilon^{-2}V[P_l]M^l. \tag{15}$$

The $M^l + M^{l-1}$ term in C^{ML} is necessary to account for computations at different levels for each multilevel sample, and the $2\epsilon^{-2}V[P_l]$ term in the C^{MC} calculation is necessary is to account for the fact that the variance of the estimator must be less than $1/2\epsilon^2$ as is the case for the multilevel method.

5. ADAPTIVE ANOVA-BASED MIXED MSFEMS

In this section, an ANOVA-based mixed MsFEM, an adaptive ANOVA-based mixed MsFEM, and a novel adaptive ANOVA-based mixed MsFEM with a new adaptive criterion are presented.

5.1 ANOVA-Based Mixed MsFEM

ANOVA decomposition is a general set of quantitative assessment and analysis tools for capturing the high-dimensional relationships between model inputs and model outputs. ANOVA decomposition has been employed for improving the efficiency of deducing high-dimensional input-output system behavior, and can be employed to relieve the computation efforts. The ANOVA decomposition method has been used in high-dimensional stochastic systems [3, 4, 6, 7]. ANOVA is based on the assumption that only relatively low-order correlations of the input variables are important, which is valid in many physical systems. ANOVA decomposition splits a high-dimensional system into a set of low-dimensional systems to reduce the computation cost in high-dimensional systems.

A multivariate output function $S(x,t;Y):\mathbb{R}^n \longrightarrow \mathbb{R}$. $S(x,t;Y)$ is taken to be water saturation or other functions of interest in the two-phase flow system (4). The statistic properties of $S(x,t;Y)$ can be obtained by solving the system (4) using the mixed MsFEM, adaptive ANOVA, PCM, and multilevel MC in Section 3. For simplicity, $S(x,t;Y)$ is used in this section, instead of $S_{\text{MsFEM}}(x,t;Y)$, to denote the solutions obtained in the multiscale framework.

Instead of solving the two-phase flow system (4) for $S(x,t;Y)$ directly, ANOVA decomposition represents $S(x,t;Y)$ as finite hierarchical correlated functions of input variables in the form

$$S(x,t;Y) = S_0 + \sum_{1\le j_1\le n} S_{j_1}(x,t;y_{j_1}) + \sum_{1\le j_1<j_2\le n} S_{j_1,j_2}(x,t;y_{j_1},y_{j_2}) \tag{16}$$
$$+ \cdots + \sum_{(j_1<j_2<\cdots<j_n\le n)} S_{j_1,j_2,\cdots,j_n}(x,t;y_{j_1},\cdots,y_{j_n}),$$

where $S_{j_k}(x,t;y_{j_k})$ is the first-order term, $S_{j_k,j_l}(x,t;y_{j_k},y_{j_l})$ is the second-order term, etc. Each of these terms is solved by the mixed MsFEM method.

In the standard ANOVA decomposition (16), the constant term is taken to be the mean of function $S(x,t;Y)$, i.e.,

$$S_0 = \int_{\Gamma^n} S(x,t;Y)d\mu(Y).$$

This gives that the means of all higher-order terms are zero, i.e.,

$$\int_{\Gamma^n} S_{j_1,\cdots,j_s} d\mu(Y) = 0,$$

which leads to orthogonality among all the terms, and the variance of $S(x,t;Y)$ is the sum of variances of all terms, i.e.,

$$\int_{\Gamma^n} S_{j_1,\cdots,j_s} S_{k_1,\cdots,k_l} d\mu(Y) = 0, \ \text{ for } \ (j_1,\cdots,j_s) \neq (k_1,\cdots,k_l),$$

$$\sigma^2(S) = \sum_{j=1}^{n} \sum_{|\mathbf{J}|=j} \sigma^2(S_{\mathbf{J}}).$$

To avoid the computation of high-dimension integration, the Dirac measure is often used instead of the Lebesgue measure. The Dirac measure is defined as $d\mu(Y) = \delta(Y-\mathbf{c})dY$, where $\mathbf{c}$ is called the anchor point. If $\mathbf{c}$ satisfies that $S(\mathbf{c}) = S_0$, then the ANOVA representation is the same as Eq. (16), otherwise it becomes an approximation of $S(Y)$, i.e.,

$$S(x,t;Y) \approx S(x,t;\mathbf{c}) + \sum_{1\le j_1\le n} S_{j_1}(x,t;y_{j_1}) + \sum_{1\le j_1<j_2\le n} S_{j_1,j_2}(x,t;y_{j_1},y_{j_2}) \tag{17}$$
$$+ \cdots + \sum_{(j_1<j_2<\cdots<j_n\le n)} S_{j_1,j_2,\cdots,j_n}(x,t;y_{j_1},\cdots,y_{j_n}),$$

where

$$S_j(x,t;y_j) = S(x,t;Y)|_{Y=\mathbf{c}\setminus y_j} - S(x,t;\mathbf{c}),$$

$$S_{j,k}(x,t;y_j,y_k) = S(x,t;Y)|_{Y=\mathbf{c}\setminus(y_j,y_k)} - S_j(x,t;y_j) - S_k(x,t;y_k) - S(x,t;\mathbf{c}).$$

The accuracy of this anchored-ANOVA depends on the choice of anchor point $\mathbf{c}$. Discussions about this can be found in [7]. In this paper, the anchor point is chosen to be the mean of random variable Y as discussed in [4]. Because low-order terms in ANOVA expansion usually have the dominant contribution, the ANOVA expansion is truncated to only keep the low-order (e.g., second- or third-order) terms for approximation, which can keep relatively good accuracy and significantly reduce the computational cost.

5.2 Adaptive ANOVA-Based Mixed MsFEM

The ANOVA-based mixed MsFEM reduces the computational complexity in stochastic space by dividing a high-dimensional stochastic problem into a set of lower-dimensional subproblems, while the mixed MsFEM reduces computation cost in spatial space. The computational cost of solving a low-dimensional system is reduced, but solving a large number of such subsystems can keep the computation cost still high. For example, if the dimensions of input parameter space $n = 100$, and the ANOVA expansion is truncated up to second order, there are a total of $1 + \binom{100}{1} + \binom{100}{2} = 5051$ components (terms) in the truncated ANOVA expansion. This computation cost is still high. To reduce the total number of terms, the adaptive ANOVA method is often applied. The dimensions with dominant interactions are called active or important dimensions. The idea of adaptive ANOVA is to retain the active dimensions and interactions and neglect the contributions from the less-active dimensions and interactions. The following equation describes the adaptive ANOVA representation:

$$\begin{aligned} S(x,t;Y) \approx S_0 + \sum_{j_1\in D_1} S_{j_1}(x,t;y_{j_1}) + \sum_{(j_1<j_2)\in D_2} S_{j_1,j_2}(x,t;y_{j_1},y_{j_2}) \\ + \cdots + \sum_{(j_1<j_2<\cdots<j_\nu)\in D_\nu} S_{j_1,j_2,\cdots,j_\nu}(x,t;y_{j_1},\cdots,y_{j_\nu}). \end{aligned}$$

In [4, 6, 7], $\nu = 2$ and $D_1 = \{1,\cdots,n\}$; then, D_i, $2 \le i \le \nu$ are selected according to the statistical properties of the computed expansion terms. There are two criteria used to find the active dimensions based on the first-order terms in the ANOVA decomposition.

Criterion 1: Use the mean of component function S_j as the indicator to decide the active ANOVA terms (See [4, 6]). Let

$$\eta_j^{(1)} = \frac{E[S_j]}{\sum_{j\in D_1} E[S_j]},\ \ j \in D_1;$$

then, the active dimensions D_1' can be chosen by

$$\sum_{j\in D_1'} \eta_j^{(1)} \ge p, \tag{18}$$

where p is a proportionality constant with $0 < p < 1$ and close to 1. In computation, $\{\eta_j^{(1)}\}$'s are sorted first and summed up to the active dimensions D_1' when (18) satisfies.

Criterion 2: Use the variance of component function S_j as the indicator to decide the active ANOVA terms. Define

$$\eta_j^{(2)} = \frac{\sigma^2(S_j)}{\sum_{j\in D_1} \sigma^2(S_j)},\ \ j \in D_1.$$

The active dimensions D_1' should satisfy

$$\sum_{j\in D_1'} \eta_j^{(2)} \ge p, \tag{19}$$

where p is a proportionality constant with $0 < p < 1$ and close to 1. This criterion is similar to the criterion used in [36], where $\sigma^2(S)$ instead of $\sum_{j\in D_1} \sigma^2(S_j)$ is used.

The active dimensions for second-order ANOVA terms can be found by the above criteria. If the active dimensions are needed for higher-order terms, similar criterion can be used,

$$\eta^{(1)}_{j_1,j_2} = \frac{E[S_{j_1,j_2}]}{\sum_{(j_1,j_2)\in D_2} E[S_j]},\ \eta^{(2)}_{j_1,j_2} = \frac{\sigma^2(S_{j_1,j_2})}{\sum_{(j_1,j_2)\in D_2} \sigma^2(S_j)}. \tag{20}$$

5.3 Novel Adaptive Anova-Based Mixed MsFEM

In the above criteria, D_1 is taken to be $\{1,\cdots,n\}$, i.e., all the dimensions are always considered to be active in the ANOVA first-order term computation. The selection of active dimensions is then conducted for second-order terms based on the ratio (e.g., $\eta^{(1)}_j, \eta^{(2)}_j$) associated with each dimension. The dimensions with small ratios can be neglected in not only the computation of second-order terms, but also first-order terms. Here a variance-decomposition-based method is presented to preselect active dimensions and simplify the computation of ANOVA decomposition starting from the first-order terms.

To preselect the active dimensions, the original high-dimensional system is solved on the sparse-grid collocation points of Ω or multilevel MC method. The level of sparse grid is determined by two factors: (1) the affordable computational cost; and (2) the desired approximation accuracy. Then, a polynomial chaos approximation for $S(x,t;Y)$ is built based on the sparse-grid collocation points. The total variance of $S(x,t;Y)$ is carried by the summation of the square of the coefficients of the basis functions except the zero-order term in the approximation. Since the basis functions in Ω are tensor products of the basis functions of dimension 1, the basis functions can be viewed to be the interactions of different random dimensions. In this case, the square of the coefficients presents the corresponding variance contribution due to the interaction of corresponding dimensions. By following this procedure, the variance of $S(x,t;Y)$ can be decomposed with respect to the set of random dimensions. In particular, the variance decomposition contributed by each random dimension can be obtained through the orthogonal gPC Galerkin projection. So, the importance of each random dimension can be estimated before ANOVA decomposition, which can be used to guide the selection of ANOVA terms adaptively. In this case, D_1 can be selected to be a subset of $\{1,\cdots,n\}$. Since the sparse grid collocation method is applied in a low level, the computational cost is low. If a higher level sparse grid collocation is affordable, more information about the interactions between dimensions can be obtained, and additional guidance on selection of active dimensions for higher-order terms is provided.

To be specific about the variance decomposition-based adaptive ANOVA method, let us consider a scalar function

$$S(x,t;Y) : \bar{D}\times\mathbb{R}^n \longrightarrow \mathbb{R}. \tag{21}$$

Let $\{\phi_k(Y)\}$ be the gPC basis function satisfying $E[\phi_i(Y)\ \phi_j(Y)] = \delta_{ij}\gamma_i$, where $\gamma_i = E[\phi_i^2(Y)]$ and let $\mathbb{P}^n_N(Y)$ be the space of all polynomials of $Y\in\mathbb{R}^n$ of degree up to N. Then, the orthogonal gPC projection of Eq. (21), for any fixed x, is

$$S_N(x,t;Y) = \mathbb{P}_N S = \sum_{|\mathbf{k}|=0}^{N} s_{\mathbf{k}}(x,t)\phi_{\mathbf{k}}(Y),$$

where the expansion coefficients are obtained as

$$s_{\mathbf{k}}(x,t) = \frac{1}{\gamma_{\mathbf{k}}}E[S(x,t;Y)\phi_{\mathbf{k}}(Y)] = \frac{1}{\gamma_{\mathbf{k}}}\int S(x,t;y)\phi_{\mathbf{k}}(y)dF_Y(y),\ \ \forall|\mathbf{k}|\le N,$$

where $\gamma_{\mathbf{k}} = E[\phi^2_{\mathbf{k}}]$ is the normalization constant of the basis $\phi_{\mathbf{k}}$, and $F_Y(y)$ is the probability distribution of Y.

Integration rules can be used to approximate the integrals in the expansion coefficients of the continuous gPC. Let $\{y^{(1)},\cdots,y^{(m)}\}$ be the sparse-grid collocation points; then, discrete projection of the solution is

$$S_N(x,t;Y) = \sum_{|\mathbf{k}|=0}^{N} \hat{s}_{\mathbf{k}}(x,t)\phi_{\mathbf{k}}(Y),$$

where the expansion coefficients are $\hat{s}_{\mathbf{k}}(x,t) = 1/\gamma_{\mathbf{k}} \sum_{j=1}^{m} S(x,t;y^{(j)})\phi_{\mathbf{k}}(y^{(j)})w^{(j)}$, and $\{w^{(j)}\}$ are the collocation weights. The coefficients $\{\hat{s}_{\mathbf{k}}\}$ are approximations to the exact projection coefficients $\{s_{\mathbf{k}}\}$. The moments of $S(x,t;Y)$ can be approximated by the moments of the approximation $S_N(x,t;Y) = \sum_{|\mathbf{k}|=0}^{N} \hat{s}_{\mathbf{k}}(x,t)\phi_{\mathbf{k}}(Y)$. The mean

$$\mu = E[S(x,t;Y)] \approx E[S_N(x,t;Y)] = \sum_{|\mathbf{k}|=0}^{N} \hat{s}_{\mathbf{k}}(x,t)E[\phi_{\mathbf{k}}(Y)] = \hat{s}_{\mathbf{0}},$$

by orthogonality of $\{\phi_k(Y)\}$. Then, the variance

$$\sigma^2(S) \;=\; E[S(x,t;Y) - \mu]^2 \approx E[S_N - \hat{s}_{\mathbf{0}}]^2 = \sum_{|\mathbf{k}|=1}^{N} \hat{s}_{\mathbf{k}}^2 \gamma_{\mathbf{k}}.$$

The variance can be further written as

$$\sigma^2(S) = \sum_{|\mathbf{k}|=1} \hat{s}_{\mathbf{k}}^2 \gamma_{\mathbf{k}} + \cdots + \sum_{|\mathbf{k}|=N} \hat{s}_{\mathbf{k}}^2 \gamma_{\mathbf{k}}, \tag{22}$$

where $\mathbf{k} = (k_1, \cdots, k_n) \in \mathbb{N}_0^n$ is a multiindex with $|\mathbf{k}| = k_1 + \cdots + k_n$. Here n is the number of random dimensions, and N is the total degree of multivariate gPC. Denote $v = \sigma^2(S)$, and $v_i = \hat{s}_{\mathbf{k}}^2 \gamma_{\mathbf{k}}$, where $|\mathbf{k}| = 1$ and i denotes the index of nonzero dimension. Similarly, let $v_{ij} = \hat{s}_{\mathbf{k}}^2 \gamma_{\mathbf{k}}$, where $|\mathbf{k}| = 2$, i and j denote the indexes of nonzero dimensions. The expansion (22) can be rewritten as

$$v = \sum_{1 \le i \le n} v_i + \sum_{1 \le i \le j \le n} v_{ij} + \cdots. \tag{23}$$

Remark: v_i/γ_i, $1 \le i \le n$ is the square of the coefficient of the gPC basis function, which can be easily obtained through Galerkin projection. γ_i is precomputed and v_i is considered as the variance contribution associated with the ith dimension. Similarly, v_{ij} is also viewed as the variance contribution associated with interaction of ith and jth dimensions. The other terms can be explained in a similar way.

Based on the variance decomposition (23), the active dimensions can be selected. Define

$$\eta_j^{(3)} = \frac{v_j}{v} \;\text{ and }\; \eta_{ij}^{(3)} = \frac{v_{ij}}{v}.$$

Then, the active dimensions can be selected similarly to Criteria (18) and (19).

Criterion 3: The active dimension D_1 should satisfy

$$\sum_{j \in D_1} \eta_j^{(3)} \ge p, \tag{24}$$

where p is a proportionality constant with $0 < p < 1$ and close to 1.

The advantage of the proposed method here is that the information on the sensitivities of each random dimension and their interactions can be efficiently obtained based on the variance decomposition technique by building a gPC expansion of the function of interest $S(x,t;Y)$ with relative low computational cost. Then, the adaptivity can start from the first-order term in ANOVA expansion based on criterion (24). Further information about the interactions among dimensions can be obtained if more computation cost is affordable, i.e., when the gPC expansion is built in a higher level of sparse grid points or larger number of multilevel Monte Carlo simulations. In [4, 6], it was assumed that the active dimension interactions only happen between active dimensions. But, it is possible that the less-active dimensions may have large interactions with the active or less-active dimensions. The variance decomposition-based adaptive method we proposed can provide additional information on the active dimension interactions among all dimensions to avoid losing some active interactions over adaptive process.

6. NUMERICAL RESULTS

In this section, it is assumed that the random permeability field $k(x,\omega) = \exp\big(a(x,\omega)\big)$ is a log-normal stochastic process. Here, the covariance function of $a(x,\omega)$ has the form

$$\mathrm{cov}[a](x,y) = \sigma^2 \exp\Big(-\frac{|x_1-y_1|^2}{2l_1^2} - \frac{|x_2-y_2|^2}{2l_2^2}\Big). \tag{25}$$

l_1 and l_2 are the correlation lengths in the horizontal and vertical dimensions, respectively. The stochastic field $a(x,\omega)$ can be approximated by a truncated KLE approximation. In practice, the KLE of $a(x,\omega)$ can be written as

$$a(x,\omega) = E[a] + \alpha \sum_{i=1}^{N} \sqrt{\lambda_i} y_i(\omega) b_i(x), \tag{26}$$

where α is a constant. In our numerical examples, $a(x,\omega)$ is sampled in Eq. (26) by generating random variables y_i from a uniform distribution on $[-1,1]$. The uniform distribution on $[-1,1]$ is used here in order to get compact support for sparse grids. The study in [16] shows that the log permeability with uniform, Beta and Gaussian distributions on the mean and standard deviation has close peak values of standard deviation. Therefore, we assume y_i here to be uniform distribution on $[-1,1]$, instead of Gaussian distribution.

$k_{rw}(S) = S^2$, $\mu_w = 0.1$, $k_{ro}(S) = (1-S)^2$, and $\mu_o = 1$ are taken in simulations. The permeability field $a(x,Y)$ is given on a fine grid. The water is injected at the lower-left corner, and the producer is at the upper-right corner. To validate ANOVA-based mixed MsFEM (ANOVA-MsFEM) on coarse grid, the results of solving stochastic two-phase flow (4) by ANOVA-MsFEM are compared with the results obtained by the following methods:

- MC method associated with mixed FEM on fine grid (MC-FEM);
- ANOVA-based mixed FEM on fine grid (ANOVA-FEM);

When mixed MsFEM is used, the fine grid is coarsened to form a uniform coarse grid. The pressure equation on the coarse grid is solved using the mixed MsFEM and then reconstructs the fine-scale velocity field as a superposition of the multiscale basis functions. The reconstructed field is used to solve the saturation equation by the finite volume method in the fine grid. The two-phase flow system is solved by the classical IMPES (implicit pressure explicit saturation). The MsFEM solver used in the numerical experiments is developed in [20]. The boundary condition is $g = 0$ and the source term f is a vector with 1 in injector, –1 in producer, and 0 at other locations. The initial condition and the boundary condition of the saturation are both zero. For the reference mixed FEM solution, we use the lowest Raviart-Thomas element. In the numerical experiments, we use the MATLAB code by Jorg Aarnes from SINTEF for the mixed MsFEM solver.

MC results are obtained from 10^4 MC simulations. The high-dimensional system in the numerical examples are decomposed by ANOVA into a set of low-dimensional subsystems up to second-order terms. These stochastic systems are solved based on level-two Smolyak sparse-grid collocation points. Various production characteristics are compared. The saturation S is computed at 0.2 PVI and 0.6 PVI and the water-cut curve $w(t)$, defining the fraction of water in the produced fluid as a function of PVI, i.e.,

$$w(t) = \frac{q_w(t)}{q_w(t) + q_o(t)},$$

where $w(t)$ is the water cut, q_w and q_o are flow rates of water and oil at the producer at time t. The breakthrough time T_w defined as $w^{-1}(10^{-5})$ at the producer and the cumulative oil production at 0.6 PVI are monitored, i.e.,

$$Q_o = -\frac{1}{\int_D \varphi dx} \int_0^{0.6\ \mathrm{PVI}} \left(\int_D \min(q_o(x,\tau),0) dx \right) d\tau,$$

where Q_o is the cumulative oil production, and φ represents the porosity. In our numerical examples, φ is set to constant 1. All the results computed by MC-FEM are considered to be the reference solutions, and errors are measured

which are defined as $e(S) = |\|E[S]\|_{L^1} - \|E[S_{ref}]\|_{L^1}|/\|E[S_{ref}]\|_{L^1}$, $e(w) = \|E[w] - E[w_{ref}]\|_{L^2}/\|E[w_{ref}]\|_{L^2}$, $e(T_w) = |E[T_w] - E[(T_w)_{ref}]|$ and $e(Q_o) = |E[Q_o] - E[(Q_o)_{ref}]|$, where $E[\cdot]$ stands for mean solutions. The standard deviation errors are defined in a similar manner.

There are different kinds of errors when the listed three methods are used to solve the governing equations. e_{total} denotes the total error, which is the expectation of absolute error between MC-FEM and ANOVA-MsFEM. The e_{stoch} is defined as the error from dimension reduction and numerical approximation by stochastic collocation method or multilevel Monte Carlo in stochastic domain, i.e., the error between MC-FEM and ANOVA-FEM, and the e_{ms} is defined to be the error from mixed MsFEM discretization in spatio-temporal domains. For water-cut function w, for instance, the definitions of these errors are listed as follows: $e_{total} = \|E[w_{\text{MC-FEM}}(t)] - E[w_{\text{ANOVA-MsFEM}}(t)]\|_{L^2}$, $e_{stoch} = \|E[w_{\text{MC-FEM}}(t)] - E[w_{\text{ANOVA-FEM}}(t)]\|_{L^2}$, and $e_{ms} = \|E[w_{\text{ANOVA-FEM}}(t)] - E[w_{\text{ANOVA-MsFEM}}(t)]\|_{L^2}$. Errors for the other functions are defined in the similar manner.

6.1 Random Permeability Field

In this example, $l_1 = 0.2$, $l_2 = 0.05$, and $\sigma^2 = 1$ are taken in Eq. (25), and $E[a] = 1$ and $\alpha = 0.05$ in Eq. (26). The permeability field $a(x, Y)$ is given on a 80×80 fine grid. The fine grid is coarsened to form a uniform 8×8 coarse grid, so that each block in the coarse grid contains a 10×10 cell partition from the fine grid. KLE is truncated to be $N = 20$ terms, so that the number of random dimensions of this stochastic system is 20.

Saturations at different times and water-cut results are shown in Figs. 1–3. The quantitative errors are reported in Tables 1 and 2. For saturation function S at different times, ANOVA-FEM behaves better than ANOVA-MsFEM compared with the reference ones. The relative large variance of the saturation function is in the flow front as shown in Figs. 1 and 2. The variance behaves differently on horizontal and vertical directions too, since the correlation lengths are different in these two directions. When time increases from 0.2 PVI to 0.6 PVI, the accuracy of the solutions is decreasing. For both ANOVA-FEM and ANOVA-MsFEM, the relative errors increase. This could due to the uncertainty accumulation in time. At 0.2 PVI, ANOVA-MsFEM is as good as ANOVA-FEM in mean value, with the same magnitude of standard deviation. When time increases, ANOVA-MsFEM results are not comparable to ANOVA-FEM. The same situation can be discovered by looking at the standard deviations at 0.2 PVI and 0.6 PVI, respectively in Figs. 1 and 2. The standard deviations have the similar pattern at 0.2 PVI, while at 0.6 PVI, the standard deviation from ANOVA-MsFEM is quite different from the above two cases. Water-cut results are the same as saturation at 0.6 PVI in the sense of the behavior of these two methods. The standard deviation of water cut has a large value around water breakthrough time for all three cases in Fig. 3. From Table 2, the absolute errors of water breakthrough time T_w and cumulative oil production Q_o, it is clear that ANOVA-MsFEM provides good approximations to these function values, while ANOVA-FEM behaves better than ANOVA-MsFEM. However, the computation of the mixed FEM is much more expensive than the mixed MsFEM.

There are errors introduced by ANOVA-MsFEM as shown from the above analysis. It is also critical to investigate the dominant error in the simulation process. Fully understanding the structure of errors will shed the light needed to develop new methods targeting the dominant error. The error introduced by MsFEM, e_{ms}, has the same magnitude as the total one e_{total} in Table 3. In summary, the error introduced by the multiscale method is the dominant one compared with the error introduced by ANOVA representation and stochastic collocation methods. Seeking better multiscale methods for better approximations is one of our future research directions.

6.2 Random Permeability Field with Channelized Structure

For our second numerical example, $l_1 = 0.2$, $l_2 = 0.05$, and $\sigma^2 = 1$ are taken in Eq. (25), the same as in the first example. In Eq. (26), $\alpha = 1$, and $E[a]$ is chosen to have the channelized feature and shows the dominant feature of the permeability $a(x, \omega)$. $N = 20$ terms are chosen in the truncated KLE, i.e., $a(x, \omega) = E[a] + \sum_{i=1}^{20} \sqrt{\lambda_i} Y_i \Phi_i$. Consequently,

$$k(x, \omega) = \exp(a(x, \omega)) = \exp(E[a]) \exp\left(\sum_{i=1}^{20} \sqrt{\lambda_i} Y_i \Phi_i\right) := k_1(x) k_2(x, \omega).$$

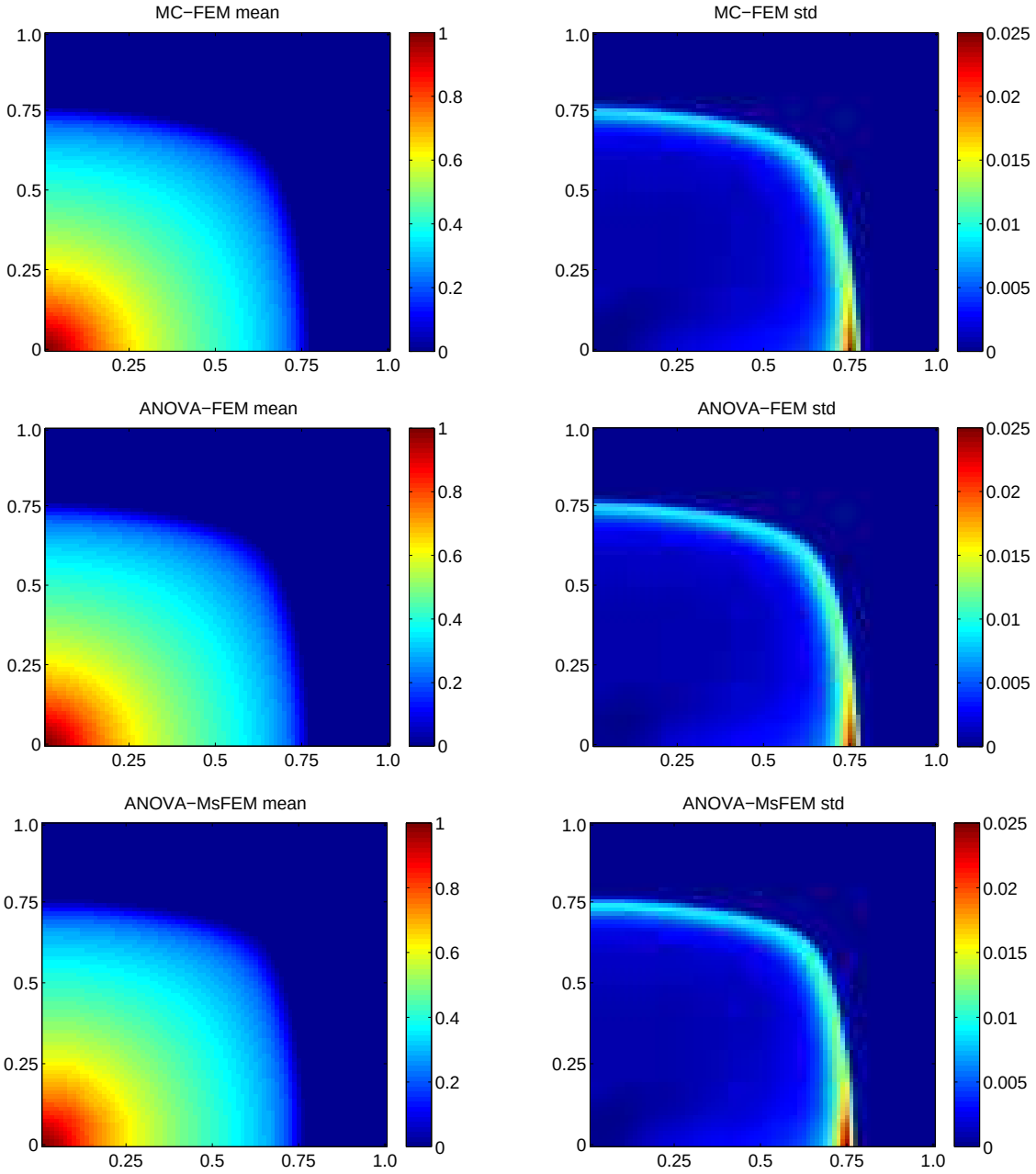

FIG. 1: Mean (left) and standard deviation (right) of saturation S at 0.2 PVI. Left column: mean; right column: standard deviation. Top row: MC-FEM; middle row: ANOVA-FEM; bottom row: ANOVA-MsFEM.

For each realization, the permeability $k(x, \omega)$ is defined on a 60×60 fine grid. Figure 4 depicts the logarithm of $k_1(x)$ (left) and an arbitrary realization of logarithm of $k_2(x, \omega)$ (right). The channelized structured $k_1(x)$ is taken from part of the layer in SPE10 data. From the permeability, it is observed that $k_1(x)$ represents a main feature of the random permeability $k(x, \omega)$. Here, the mixed MsFEM is performed on a 6×6 uniform coarse grid.

The variance in this example is larger than the previous example. The results in Tables 4 and 5 are larger in magnitude than in Tables 1 and 2. However, the trends are similar to the first example. The ANOVA-FEM generally

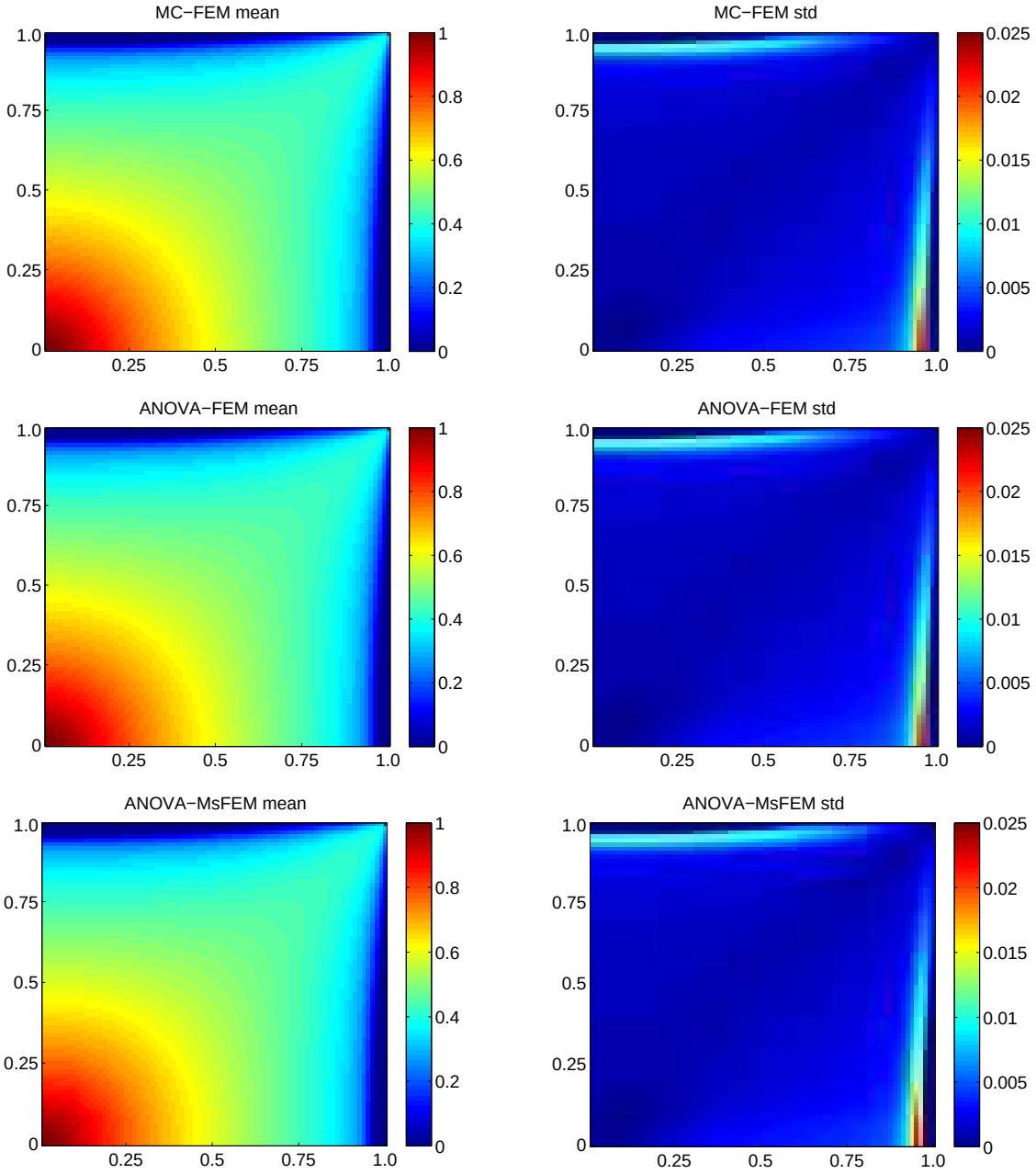

FIG. 2: Mean (left) and standard deviation (right) of saturation S at 0.6 PVI. Left column: mean; right column: standard deviation. Top row: MC-FEM; middle row: ANOVA-FEM; bottom row: ANOVA-MsFEM.

provides better accuracy than ANOVA-MsFEM. The accuracy is decreasing as time increases. But the standard deviation of saturations computed by ANOVA-MsFEM is more accurate than the result computed by ANOVA-FEM. The natural inference is that ANOVA-FEM is more accurate than ANOVA-MsFEM, because the mixed MsFEM results are approximations of FEM solutions with less computation cost. In addition, since ANOVA applied here is only up to second-order terms, the process of approximating a nonlinear operator by finite linear operations introduces extra

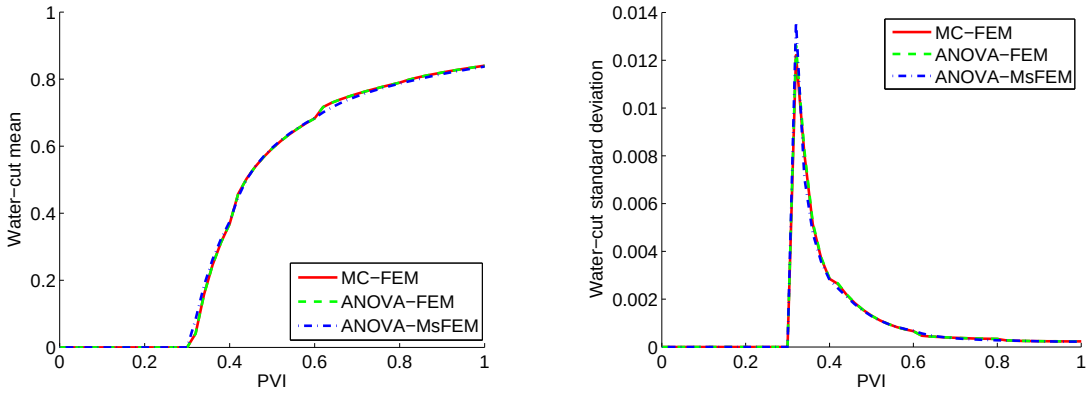

FIG. 3: Comparison of mean (left) and standard deviation (right) of water-cut w.

TABLE 1: Relative errors of saturation S at 0.2, 0.6 PVI and water-cut w

	ANOVA-FEM		ANOVA-MsFEM	
	Mean error	Std error	Mean error	Std error
$e(S)$ at 0.2 PVI	3.422170e-09	2.184847e-03	3.422172e-09	6.116023e-03
$e(S)$ at 0.6 PVI	3.135168e-07	2.745868e-03	3.616744e-03	5.136930e-03
$e(w)$	8.544170e-05	2.620252e-02	1.571879e-02	1.050685e-01

TABLE 2: Absolute errors of water breakthrough time T_w and cumulative oil-production Q_o

	ANOVA-FEM		ANOVA-MsFEM	
	Mean error	Std error	Mean error	Std error
$e(T_w)$	7.408498e-07	2.345094e-11	2.290520e-06	2.629305e-11
$e(Q_o)$	1.592182e-07	6.972235e-09	1.662888e-03	4.961346e-10

TABLE 3: Different errors of water-cut w, water breakthrough time T_w, and cumulative oil-production Q_o

	e_{total}	e_{stoch}	e_{ms}
w	1.527618e-02	1.147299e-03	1.412888e-02
T_w	2.290520e-06	7.408498e-07	3.031370e-06
Q_o	1.662888e-03	1.592182e-07	1.662729e-03

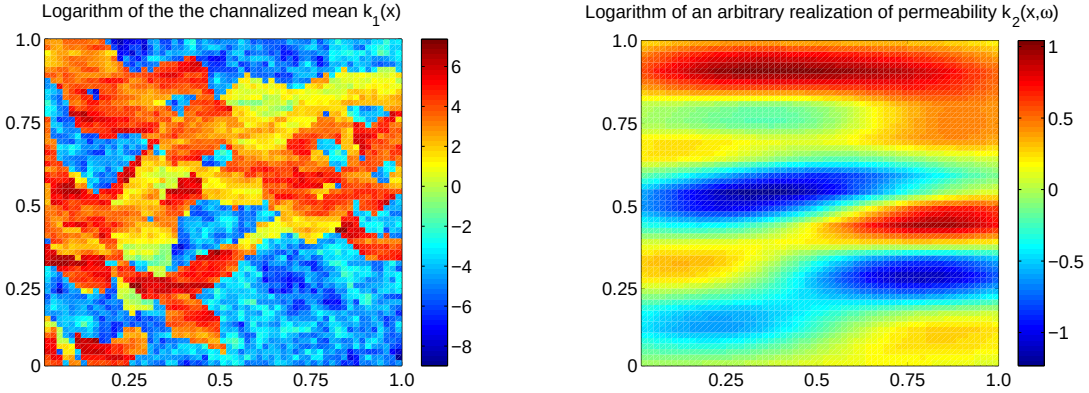

FIG. 4: The logarithm of $k_1(x)$ (left) and an arbitrary realization of logarithm of $k_2(x, \omega)$ (right).

TABLE 4: Relative errors of saturation S at 0.2, 0.6 PVI, and water-cut w

	ANOVA-FEM		ANOVA-MsFEM	
	Mean error	Std error	Mean error	Std error
$e(S)$ at 0.2 PVI	9.472140e-03	3.138906e-01	7.615875e-03	2.625666e-01
$e(S)$ at 0.6 PVI	4.296857e-03	3.612245e-01	9.758884e-03	2.605877e-01
$e(w)$	1.515789e-02	5.825024e-01	2.001007e-02	3.293715e-01

TABLE 5: Absolute errors of water breakthrough time T_w and cumulative oil-production Q_o

	ANOVA-FEM		ANOVA-MsFEM	
	Mean error	Std error	Mean error	Std error
$e(T_w)$	6.797262e-02	4.350505e-03	6.884711e-02	4.416420e-03
$e(Q_o)$	3.126008e-04	2.948469e-05	1.858780e-03	6.022319e-06

errors. If ANOVA is expanded to high-order terms, the results will converge to the true value, and ANOVA-FEM should be more accurate than ANOVA-MsFEM.

ANOVA approximations have better results for the mean solutions than the standard deviation. The uncertainty of saturations S are coming from where the flow front is (Figs. 5 and 6). The large standard deviation of water-cut w is again around water breakthrough time (Fig. 7), while the magnitude of standard deviation in this example is larger than the previous example (Fig. 3), as the variance of the parameter is larger. This example illustrates the error due to MsFEM approximation is still the dominant error in the computation as shown in Table 6.

6.3 Adaptive ANOVA Case Studies

This simulation focuses on the analysis of the adaptive ANOVA based on the new variance decomposition criterion proposed in Section 5.3. Before showing the numerical results, several observations and supportive numerical results are stated as follows. The adaptive Criteria 1 and 2 are to some extent comparable [6]. Some numerical examples demonstrate that the active dimensions are similar for both criteria. But for special functions, for example water-cut w, Criterion 2 based on variance usually gives a lower number of active dimensions. To obtain a better approximation through any kind of adaptive ANOVA method, the criterion has to be applied to the results of the interested functions.

In this numerical test case, $l_1 = 0.25$, $l_2 = 0.1$, and $\sigma^2 = 1$ are taken in Eq. (25), and $E[a] = 0$ and $\alpha = 1$ are used in Eq. (26). The KLE approximation is truncated to be $N = 50$ terms. Since the correlation lengths in horizontal and vertical directions are larger than the previous examples, the covariance function is smoother and the eigenvalues decay faster. The fine grid is 80×80 and the coarse grid is 8×8. Table 7 shows that when $p = 0.9$, for saturation S at 0.2 PVI, Criteria 1 and 2 give almost the same active dimensions. The second-order adaptive ANOVA is applied here. In fact, the active dimensions are the same when $p = 0.85$, and 0.9. And, the active dimensions are the dimensions corresponding to the largest 21 or 22 eigenvalues, while if water-cut function w is considered, Criteria 1 and 2 give 17 and 10 active dimensions, respectively.

Now, reconsidering the example with $l_1 = 0.2$, $l_2 = 0.05$, and $\sigma^2 = 1$ in Eq. (25), and $E[a] = 0$ and $\alpha = 1$ in Eq. (26), with 20 random dimensions. In this case, the 20 eigenvalues have similar magnitude to some extent. Since Criteria 1 and 2 are comparable, and our adaptive criterion is based on variance decomposition; comparison is made only to Criterion 2, which is also based on variance.

The saturation S at 0.2 PVI is taken into account. To apply variance decomposition based analysis, level-one Smolyak sparse-grid collocation points are used to build variance decomposition (23). Forty-one level-one sparse-grid points are used here. Variance decomposition can be obtained up to first-order in (23), i.e., $v = \sum_{1 \le i \le 20} v_i$. Each random dimension corresponds to one v_i. Similar variance decomposition analysis through gPC expansion can be done to the MC results based on 10^4 number of MC samples. The variance computed by large-number MC samples are treated as the true value as before. At the same time, the anchored-ANOVA discussed in previous sections provides the variance contribution by each individual random dimension. To compare these three sets of variance, a

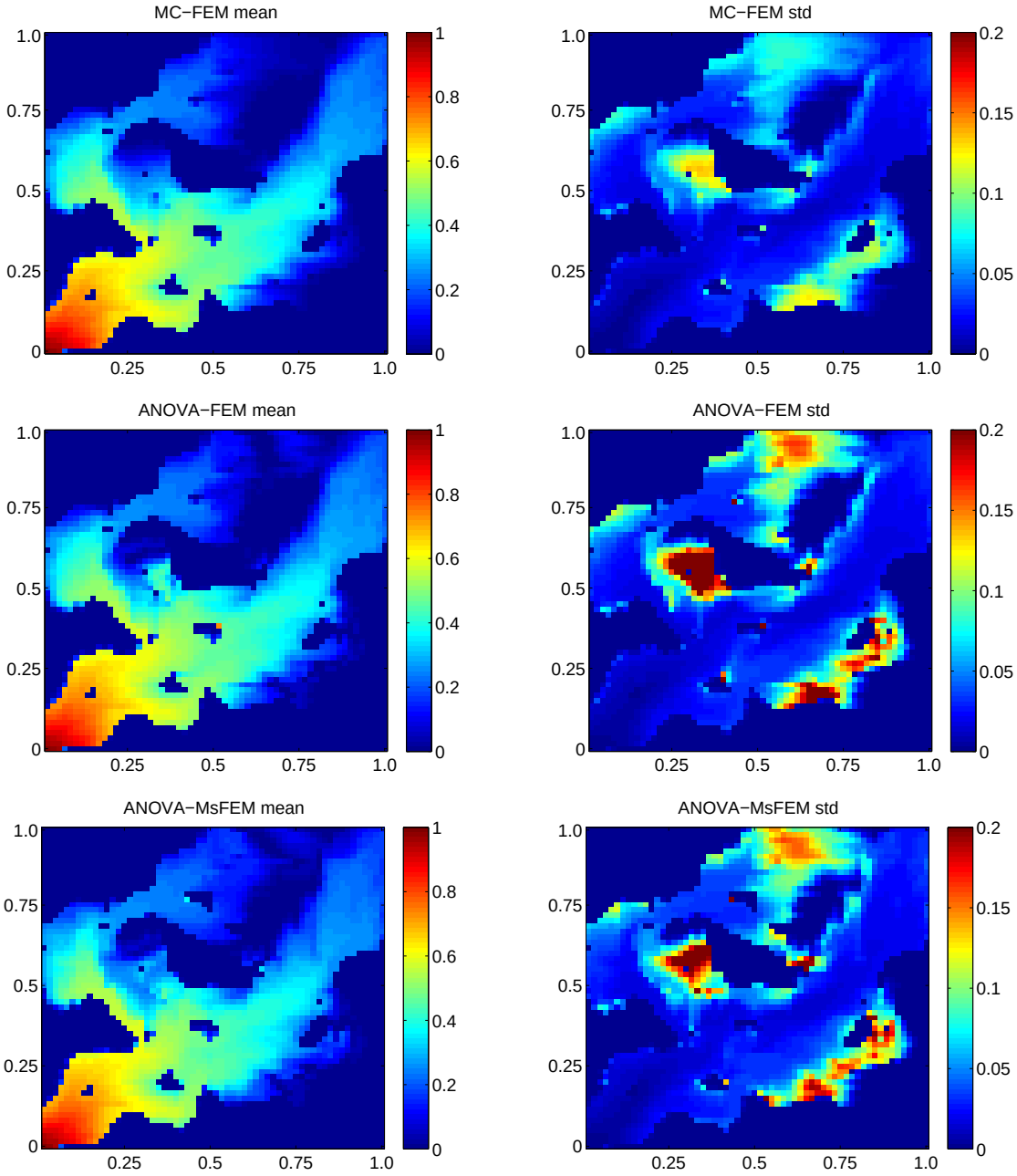

FIG. 5: Mean (left) and standard deviation (right) of saturation S at 0.2 PVI. Left column: mean; right column: standard deviation. Top row: MC-FEM; middle row: ANOVA-FEM; bottom row: ANOVA-MsFEM.

network graph plot is adopted to show the sensitivity analysis results in Fig. 8. As shown in Fig. 8, the 20 dimensions are labeled, with the radius proportional to the magnitude of variance contribution by each random dimension with respect to saturation S at 0.2 PVI, respectively.

Both adaptive ANOVA based on Criterion 2 and variance decomposition-based adaptive ANOVA results are different from the true ones with regard to matching the radius. To further explore the behavior of the two methods, two

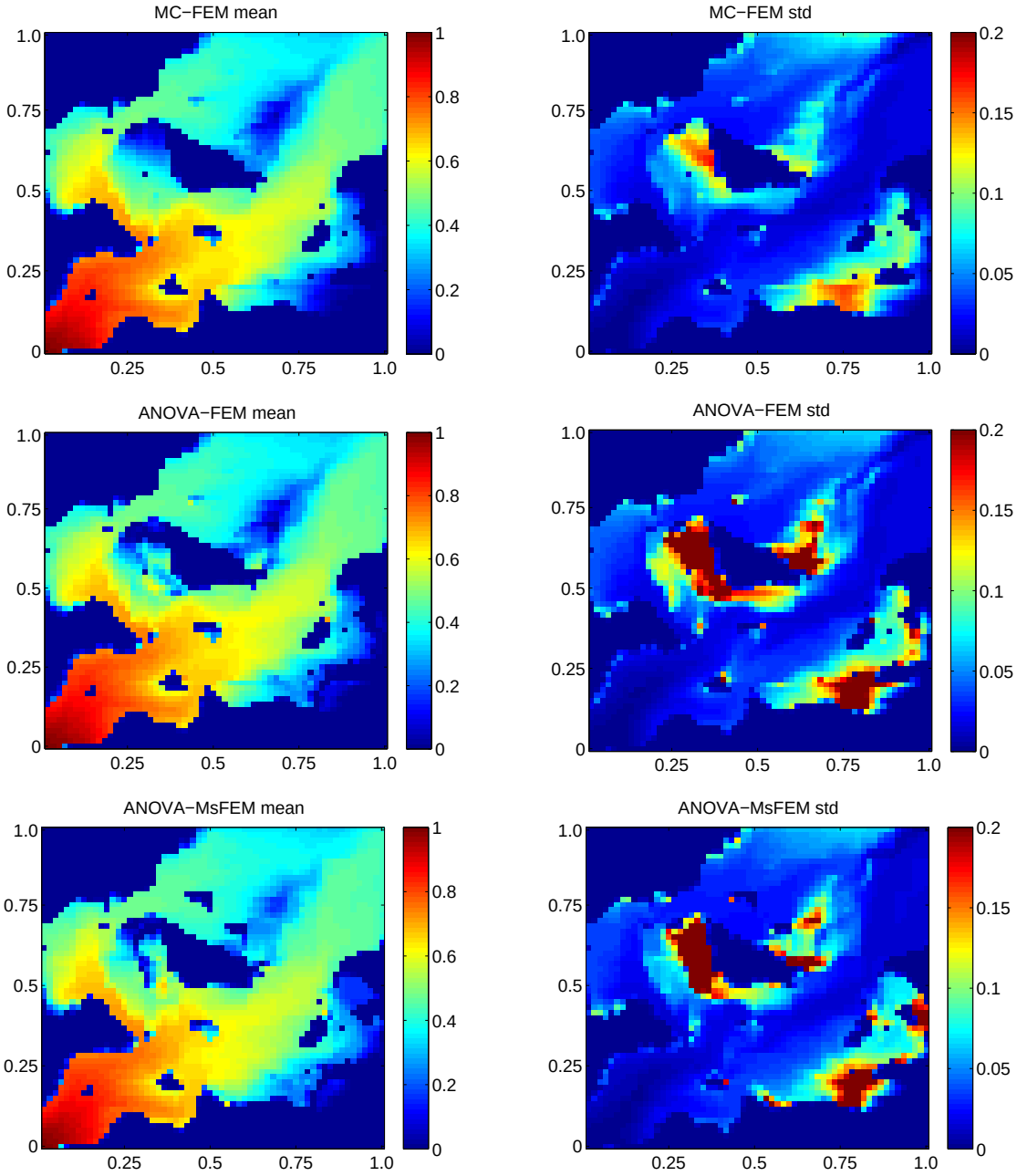

FIG. 6: Mean (left) and standard deviation (right) of saturation S at 0.6 PVI. Left column: mean; right column: standard deviation. Top row: MC-FEM; middle row: ANOVA-FEM; bottom row: ANOVA-MsFEM.

different comparisons are made. The advantage of our proposed variance decomposition-based adaptive method is that the active dimensions can be found before ANOVA approximation by only a small amount of computation, while the adaptive ANOVA can only find the active dimensions after finishing the computation of all first-order terms.

To make a "fair" comparison, it was assumed that adaptive ANOVA is expanded to first-order terms in these two methods. In the first case, the active dimensions are chosen by Criterion 2, and the mean and variance of saturation

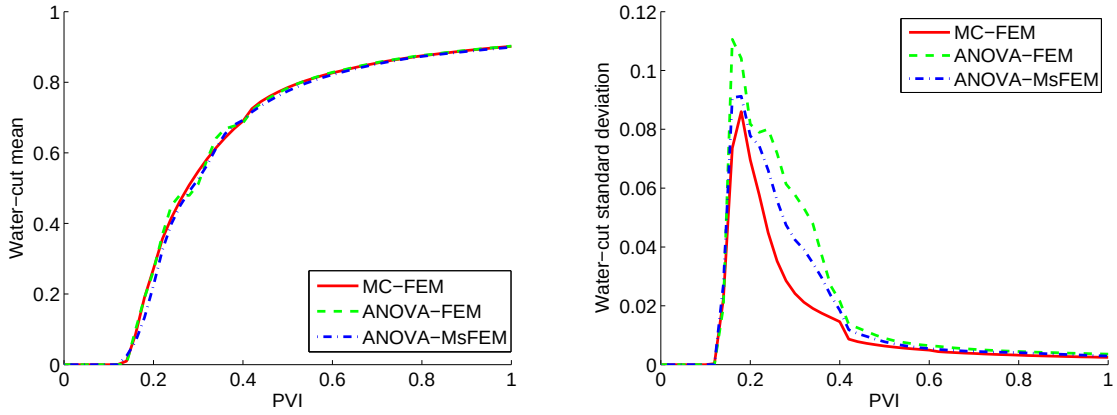

FIG. 7: Comparison of mean (left) and variance (right) of water-cut w.

TABLE 6: Different errors of water-cut w, water breakthrough time T_w, and cumulative oil-production Q_o

	e_{total}	e_{stoch}	e_{ms}
w	9.937488e-02	7.527776e-02	1.200735e-01
T_w	6.884711e-02	6.797262e-02	1.368197e-01
Q_o	1.858780e-03	3.126008e-04	1.546179e-03

TABLE 7: Comparison of adaptive ANOVA for saturation S at 0.2 PVI

	p	# of Active dim	Relative error of mean	Relative error of std
Criterion 1	0.90	21	9.314872e-02	5.618844e-02
Criterion 2	0.90	22	8.425106e-02	4.924464e-02

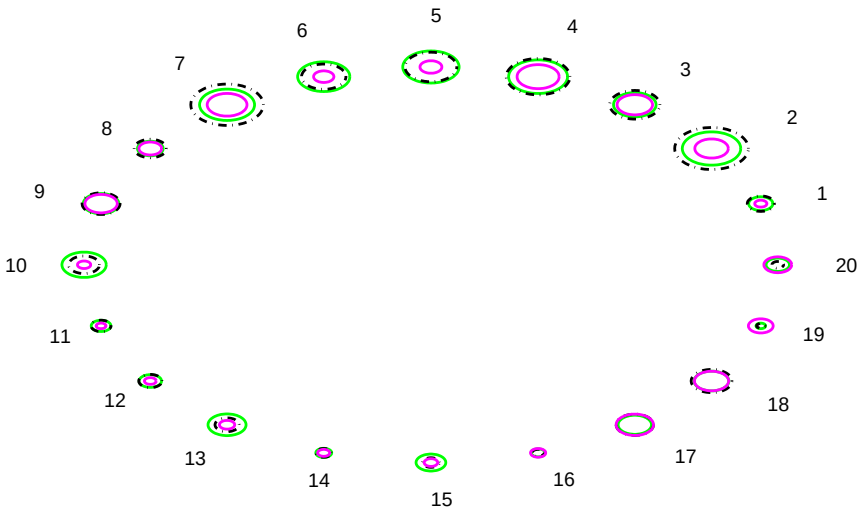

FIG. 8: Network graph of active random dimensions with respect to saturation S at 0.2 PVI. The radius is proportional to the magnitude of variance. The larger radius refers to larger variance, which is considered as more active dimension. Magenta: MC results; green: adaptive ANOVA approximation (Criterion 2); black: variance decomposition-based adaptive ANOVA (Criterion 3).

S at 0.2 PVI are formed by zero-order terms and first-order terms of active dimensions. In the second case, variance decomposition is conducted to choose active dimensions (Criterion 3), then the adaptive ANOVA includes zero-order and first-order terms of active dimensions are considered. The results are in Table 8. When $p = 0.90$, Criterion 3 gives one active dimension less than Criterion 2. These dimensions are not exactly the same. For Criterion 2, dimensions $11, 14, 16, 19$ are not included, and for Criterion 3, dimensions $14, 15, 16, 19, 20$ are not included. Criterion 3 gives results as good as Criterion 2 as shown in Table 8. The advantages of our proposed method will be more obvious, when the dimension of the problem is higher. In that case, more computations for the first-order terms can be saved.

Further, an "unfair" comparison can be made. For Criterion 2, all the first-order terms are kept and second-order terms with active dimensions are computed. And for Criterion 3, use only the first terms of pre-fixed active dimensions and compute second-order terms based on these active dimensions. Since there are less first-order terms included in Criterion 3, the relative error of std results are not as good as the ones calculated using Criterion 2. But the relative error of mean gives better approximation. Table 9 shows the difference between these two methods.

In fact, if level-two sparse grid collocation points are used, the variance function can be approximated by higher-order terms in Eq. (23). For example, 841 collocation points are taken in the random space and compute second-order in Eq. (23). Then the information between random dimensions can be obtained before ANOVA decomposition. In Fig. 9, the radius of circles corresponding to each dimension depicts the variance contribution associating with certain dimensions, and the width of the lines between any pair of dimensions depicts the correlation between that pair. The widths of the lines are normalized for display. We can select the active interactions between random dimensions by certain criteria. For example, in Fig. 10, the largest 50 correlations are shown. The thicker the line, the larger the

TABLE 8: A "fair" comparison of adaptive ANOVAs up to first-order terms for saturation S at 0.2 PVI

	p	# of Active dims	Relative error of mean	Relative error of std
Criterion 2	0.90	16	4.489370e-02	3.101351e-01
Criterion 3	0.90	15	4.519126e-02	3.079101e-01

TABLE 9: An "unfair" comparison of adaptive ANOVAs up to second-order terms for saturation S at 0.2 PVI

	p	# of Active dims	Relative error of mean	Relative error of std
Criterion 2	0.90	16	3.054183e-02	3.959672e-02
Criterion 3	0.90	15	2.065996e-02	7.027969e-02

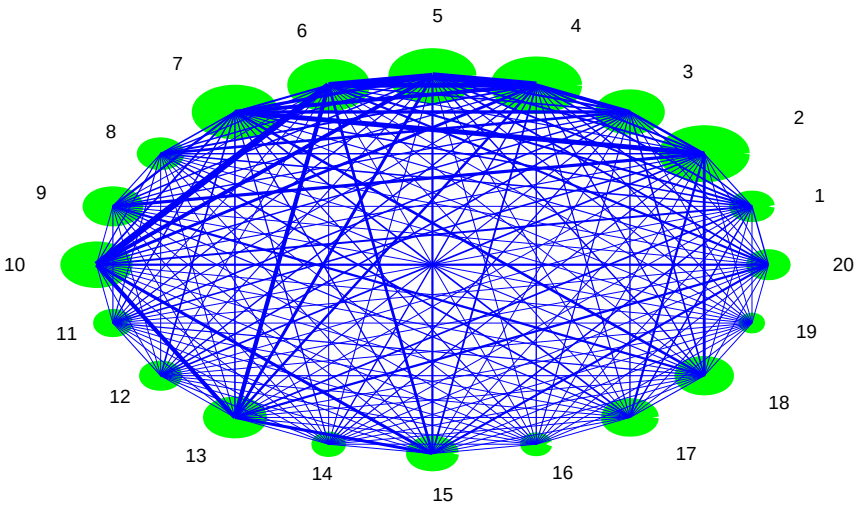

FIG. 9: Network graph of active dimensions and the interaction between pairs.

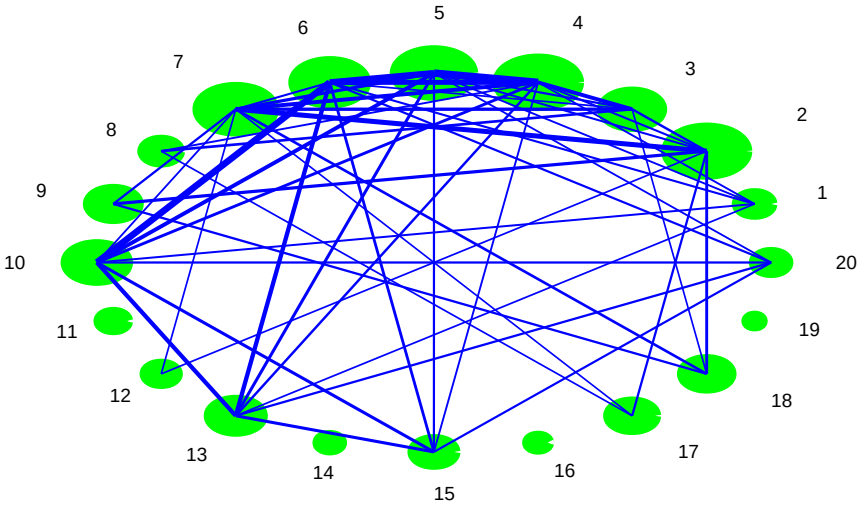

FIG. 10: Network graph of active dimensions and the largest 50 interactions between pairs.

correlation is. By applying similar criterion as listed in (20), the active interactions between pairs of dimensions can be found. Further, to avoid deleting dimensions which have small variances associated with them, but large correlation/interaction with other dimensions, high-level sparse-grid collocation points or larger number multilevel Monte Carlo samples are needed. However, this requires more computational cost.

7. CONCLUSION

In this paper, an ANOVA-based and an adaptive ANOVA-based mixed MsFEM with a novel adaptive criterion are developed for the stochastic two-phase flow problem with random permeability fields. The properties of the methods have been further studied with numerical examples on different random permeability fields. The structure and sources of the errors have been investigated. The numerical experiments show that the dominant errors are introduced by the mixed multiscale methods compared with the ANOVA decomposition errors. This motivates us to further develop multiscale methods with better accuracy to improve the developed approach in our future research. In particular, a systematic enrichment technique developed in [37] is planned for implementation.

Note that the full ANOVA decomposition will contain more and more subproblems in the stochastic space, when the dimensions increase. This fact can make the total computations as expensive as solving a high-dimensional problem directly when the number of random dimension exceeds a certain point. The carefully designed adaptive ANOVA method could be a remedy. A new adaptive ANOVA-based mixed MsFEM with a novel variance-decomposition-based adaptive criterion has been proposed and compared with the existing two adaptive criteria. Our proposed adaptive ANOVA method can determine the active dimensions and interactions among dimensions before computing the ANOVA decomposition to greatly reduce the computational cost. The numerical results show that this novel adaptive method can achieve similar accuracy as other adaptive strategies but with much lower computational cost. The advantage in saving computational time will be more obvious when the number of random dimensions of the problem becomes higher.

ACKNOWLEDGMENTS

This work was accomplished and funded by both the Applied Mathematics Program within the DOE's Office of Advanced Scientific Computing Research as part of the Collaboratory on Mathematics for Mesoscopic Modeling of

Materials and the Pacific Northwest National Laboratory's Carbon Sequestration Initiative, which is part of the Laboratory Directed Research and Development Program. A portion of the computations was performed using PNNL Institutional Computing cluster systems. PNNL is operated by Battelle for the DOE under contract DE-AC05-76RL01830.

REFERENCES

1. Cao, Y., Chen, Z., and Gunzburger, M., An ANOVA analysis for a class of partial differential equations with uncertain boundary conditions, *Int. J. Numer. Anal. Model.*, 6(2):256–273, 2009.
2. Fisher, R., *Statistical Methods for Reserach Workers*, Oliver and Boyd, Edinburgh, Scotland, 1925.
3. Foo, J. Y. and Karniadakis, G. E., Multi-element probabilistic collocation in high dimensions, *J. Comput. Phys.*, 229:1536–1557, 2009.
4. Ma, X. and Zabaras, N., An adaptive high-dimensional stochastic model representation technique for the solution of stochastic partial differential equations, *J. Comput. Phys.*, 229:3884–3915, 2010.
5. Ma, X. and Zabaras, N., A stochastic mixed finite element heterogenours multiscale method for flow in porous media, *J. Comput. Phys.*, 230:4696–4722, 2011.
6. Yang, X., Choi, M., Lin, G., and Karniadakis, G. E., Adaptive ANOVA decomposition of stochastic incompresible and compressible flows, *J. Comput. Phys.*, 231:1587–1614, 2011.
7. Zhang, Z., Choi, M., and Karniadakis, G. E., Anchor points matter in anova decomposition, In *Spectral and High Order Methods for Partial Diferential Equations Lecture Notes in Computational Science and Engineering*, Vol. 76, Springer, Berlin, pp. 347–355, 2011.
8. Tatang, M. and McRae, G., Direct treatment of uncertainty in models of reaction and transport, Technical Report, MIT Tech. Rep., 1994.
9. LeMaître, O. P. and Knio, O. M., *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*, Springer–New York, 2010.
10. Xiu, D. and Hesthaven, J. S., High order collocation methods for differential equations with random inputs, *SIAM J. Sci. Comput.*, 27:1118–1139, 2005.
11. Ma, X. and Zabaras, N., An adaptive hierarchical sparse grid collocation algorithm for solution of stochastic differential equations, *J. Comput. Phys.*, 228:3084–3113, 2009.
12. Lin, G. and Karniadakis, G. E., Sensitivity analysis and stochastic simulations of non-equilibrium plasma flow, *Int. J. Numer. Meth. Eng.*, 80:738–766, 2009.
13. Lin, G., Su, C.-H., and Karniadakis, G. E., Predicting shock dynamics in the presence of uncertainties, *J. Comput. Phys.*, 217:260–276, 2006.
14. Lin, G., Su, C.-H., and Karniadakis, G. E., Stochastic modeling of random roughness in shock scattering problems: Theory and simulations, *Comput. Methods Appl. Mech. Eng.*, 197:3420–3434, 2008.
15. Lin, G. and Tartakovsky, A. M., An efficient, high-order probabilistic collocation method on sparse grids for three-dimensional flow and solute transport in randomly heterogeneous porous media, *Adv. Water Res.*, 32(5):712–722, 2009.
16. Lin, G. and Tartakovsky, A. M., Numerical studies of three-dimensional stochastic Darcy's equation and stochastic advection-diffusion-dispersion equation, *J. Sci. Comput.*, 43(1):92–117, 2010.
17. Lin, G., Tartakovsky, A. M., and Tartakovsky, D. M., Uncertainty quantification via random domain decomposition and probabilistic collocation on sparse grids, *J. Comput. Phys.*, 229(19):6995–7012, 2010.
18. Aarnes, J. E. and Efendiev, Y., Mixed multiscale finite element for stochastic porous media flows, *SIAM Sci. Comput.*, 30:2319–2339, 2008.
19. Chen, Z. and Hou, T. Y., A mixed multiscale finite element method for elliptic problems with oscillating coefficients, *Math. Comput.*, 72:541–576, 2003.
20. Efendiev, Y. and Hou, T. Y., *Multiscale Finite Element Methods: Theory and Applications*, Springer, Berlin, 2009.
21. Jenny, P., Lee, S. H., and Tchelepi, H., Multi-scale finite volume method for elliptic problems in subsurface flow simulation, *J. Comput. Phys.*, 187:47–67, 2003.

22. Brezzi, F., Interacting with the subgrid world, *Numerical Analysis* (Dundee), Chapman & Hall/CRC, Boca Raton, FL, pp. 69–82, 2000.
23. Sangalli, G., Capturing small scales in elliptic problems using a residual-free bubbles finite element method, *Multiscale Model. Simul.*, 1:485–503, 2003.
24. Arbogast, T., Implementation of a locally conservative numerical subgrid upscaling scheme for two-phase Darcy flow, *Comput. Geosci.*, 6:453–481, 2002.
25. Hughes, T., Feijoo, G., Mazzei, L., and Quincy, J., The variational multiscale method—A paradigm for computational mechanics, *Comput. Methods Appl. Mech. Eng.*, 166:3–24, 1998.
26. Arbogast, T., Pencheva, G., Wheeler, M. F., and Yotov, I., A multiscale mortar mixed finite element method, *Multiscale Model. Simul.*, 6:319–346, 2007.
27. Giles, M. B., Multi-level Monte Carlo path simulation, *Oper. Res.*, 56(3):607–617, 2008.
28. Wang, X. and Fang, K., The effective dimension and quasi-Monte Carlo integration, *J. Complex.*, 19:101–124, 2003.
29. Grigoriu, M., *Stochastic Calculus: Applications in Science and Engineering*, Birkhäuser, Boston, Basel, Berlin, 2002.
30. Wong, E., *Stochastic Processes in Information and Dynamical Systems*, MCGraw-Hill, New York, 1971.
31. Aarnes, J. E., On the use of a mixed multiscale finite element method for greater flexibility and increased speed or imporved accuracy in reservoir simulation, *Multiscale Model. Simul.*, 2:421–439, 2004.
32. Jiang, L., Mishev, I., and Li, Y., Stochastic mixed multiscale finite element methods and their applications in reandom porous media, *Comput. Methods Appl. Mech. Eng.*, 119:2721–2740, 2011.
33. Nobile, F., Tempone, R., and Webster, C. G., An anisotropic sparse grid stochastic collocation method for elliptic partial differential equations with random input data, *SIAM J. Numer. Anal.*, 46:2411–2442, 2008.
34. Nobile, F., Tempone, R., and Webster, C. G., A sparse grid stochastic collocation method for partial differential equations with random input data, *SIAM J. Numer. Anal.*, 46:2309–2345, 2008.
35. Barthelmann, V., Novak, E., and Ritter, K., High dimensional polynomial interpolation on sparse grids, *Compuat. Math.*, 12:273–288, 2000.
36. Caflisch, R. E., Morokoff, W., and Owen, A., Valuation of mortgage-backed securities using brownian bridges to reduce the effective dimension, *J. Comput. Finance*, 1:27–46, 1997.
37. Efendiev, Y., Galvis, J., and Wu, X. H., Multiscale finite element methods for high-contrast problems using local spectral basis functions, *J. Comput. Phys.*, 230:937–955, 2011.