SIAM/ASA J. Uncertainty Quantification
Vol. 4, pp. 475–494

Stochastic Collocation Methods for Nonlinear Parabolic Equations with Random Coefficients*

David A. Barajas-Solano† and Daniel M. Tartakovsky†

Abstract. We evaluate the performance of global stochastic collocation methods for solving nonlinear parabolic and elliptic problems (e.g., transient and steady nonlinear diffusion) with random coefficients. The robustness of these and other strategies based on a spectral decomposition of stochastic state variables depends on the regularity of the system's response in outcome space. The latter is affected by statistical properties of the input random fields. These include variances of the input parameters, whose effect on the computational efficiency of this class of uncertainty quantification techniques has remained unexplored. Our analysis shows that if random coefficients have low variances and large correlation lengths, stochastic collocation strategies outperform Monte Carlo simulations (MCS). As variance increases, the regularity of the stochastic response decreases, which requires higher-order quadrature rules to accurately approximate the moments of interest and increases the overall computational cost above that of MCS.

Key words. uncertainty quantification, stochastic collocation, Richards equation, random coefficients

AMS subject classifications. 15A15, 15A09, 15A23

DOI. 10.1137/130930108

1. Introduction.

Nonlinear parabolic partial differential equations (PDEs), and their elliptic (steady-state) counterparts, describe a wide range of physical phenomena that range from heat conduction in solids to multiphase flow in porous media to electrodynamics (see [30] and the references therein). The ability of these equations to predict the underlying phenomena is often, if not always, compromised by uncertainty in their parameterizations. This uncertainty arises from ubiquitous heterogeneity of ambient environments in which such phenomena occur, scarcity of parametric data, and imprecise knowledge of forcings (sources and initial and boundary conditions). Quantification of the impact of parametric uncertainty on the veracity of model predictions is an integral part of modern scientific computing.

Probabilistic frameworks provide a standard approach to quantification of parametric uncertainty. They treat uncertain, spatially distributed system parameters and forcings as random fields. Solutions of the corresponding PDEs with random coefficients are given in terms of probability density functions or statistical moments of state variables or other quantities of interest (QoIs). Efficient computation of these statistics is the focus of uncertainty quantification (UQ).

*Received by the editors December 2, 2013; accepted for publication (in revised form) March 10, 2016; published electronically April 26, 2016. This research was supported in part by the Air Force Office of Scientific Research under grant FA9550-12-1-0185, by the National Science Foundation under grants EAR-1246315 and DMS-1522799, and by Defense Advanced Research Projects Agency under the EQUiPS program.
http://www.siam.org/journals/juq/4/93010.html

†Department of Mechanical and Aerospace Engineering, University of California, San Diego, La Jolla, CA 92037 (dzbarajas@ucsd.edu, dmt@ucsd.edu).

Monte Carlo simulations (MCS)—a method that consists of (i) sampling multiple realizations of input parameters from their distribution, (ii) solving deterministic PDEs for each realization, and (iii) evaluating ensemble statistics of these solutions—provide the most robust and straightforward way to solve PDEs with random coefficients. Since MCS impose no limitations on the number or statistical properties of input parameters, they usually serve as a yardstick against which the performance of other approaches is compared. MCS entails no modifications of existing deterministic solvers and therefore is often referred to as a nonintrusive technique. Yet the convergence rate of MCS is low and, thus, a large number of realizations are required to reach a target error for the MCS estimators. This renders MCS computationally expensive (often prohibitively so). Research in the field of UQ is driven by the goal of designing numerical techniques that are computationally more efficient than MCS. In other words, the simple MCS provides a baseline, which any UQ tool must outperform. Accelerated Monte Carlo (MC) techniques, such as quasi-MC and multilevel MC, are examples of such tools.

One alternative to MCS is to derive deterministic PDEs that govern the evolution of either statistical moments [16, 36] or probability density functions [21, 29] of dependent variables. The reliance on new governing equations implies that these approaches are intrusive, even though their underlying structure often remains the same and existing solvers can be used. In most implementations, these methods do not rely on finite-term approximations of random parameter fields, e.g., on truncated Karhunen–Loève (K-L) expansions, and thus do not suffer from "the curse of dimensionality." When random parameters enter PDEs as multiplicative noise, these approaches require a closure approximation. A systematic approach to obtaining such closures is based on perturbation expansions of relevant quantities into series in the powers of variances of the input parameters (see e.g., [30] and the references therein). This formally limits the applicability of such techniques to PDEs whose random coefficients exhibit low noise-to-signal ratios (coefficients of variation), even though in linear diffusion-type problems they might remain accurate for variances of log-transformed input fields as high as 4 [38].

Various flavors of stochastic finite element methods (FEMs) provide another alternative to MCS. Stochastic FEMs start by characterizing (e.g., by means of truncated K-L expansions) random parameter fields in terms of a finite set of random variables. This approximation introduces a bias of the estimators of statistics of QoIs with respect to their exact counterparts (i.e., without relying on finite-term approximate representations). The finite set of random variables defines a finite-dimensional outcome space on which a stochastic PDE(SPDE) solution is defined. The Galerkin FEM [11, 3], often equipped with h-type and p-type adaptivity, approximates such solutions in the resulting composite outcome-physical space. Stochastic Galerkin and collocation methods [37] employ orthogonal basis expansions of an SPDE solution in the chosen finite-dimensional outcome space.

These methods are often referred to as nonperturbative, even though their applications to systems whose parameters exhibit large coefficients of variation are scarce. They outperform MCS when random fields exhibit long correlations and, therefore, can be accurately represented by a few terms in their K-L expansions. As correlation length of an input parameter decreases, its K-L expansion requires more terms to maintain the same accuracy, thus increasing the dimension of the outcome space on which the solution is defined. Once the number of

random variables exceeds a certain threshold, stochastic FEMs become computationally less efficient than MCS.

Nonlinear problems of the kind considered in the present analysis pose additional challenges to the stochastic Galerkin and collocation methods and other spectral representations. In particular, the convergence rate of the quadrature rules employed to approximate integrals over the outcome space depends on the smoothness of the state variable with respect to all the random variables used to represent the input parameters. A problem's nonlinearity degrades its solution's regularity in the outcome space, undermining the performance of the stochastic Galerkin and collocation methods.

Another issue affecting the performance of these methods is their scalability with the dimension of the chosen outcome space, which must be taken into account to control the computational cost while maintaining the accuracy. Strategies for dealing with the occurrence of discontinuities and/or the loss of regularity include direction-adaptive quadrature rules [10, 8], multielement generalized polynomial chaos (gPC) [34, 32], multielement probabilistic collocation [6], locally adaptive stochastic collocation (SC) [1, 23], and wavelet expansions [19, 20]. These methods have been shown to be efficient for problems with a low-dimensional outcome space. Their applicability is curtailed by the exponential increase in the computational cost associated with adaptivity in high-dimensional spaces [23].

We investigate the relative performance of collocation techniques for solving nonlinear advection-diffusion equations with random coefficients, a class of problems that includes the Richards and viscous Burgers equations. Section 2 contains a problem formulation and its statistical parameterization. SC approaches for solving this problem are discussed in section 3. In section 4 we test these strategies on the steady (elliptic) and transient (parabolic) versions of the nonlinear Richards equations. The results are reported in terms of estimates of both the error due to truncation of K-L expansions of the random parameter fields and the error of the SC method. We also compare the rates of decay, as a function of the number of samples or realizations, of the bound on the total error of the SC method and MCS. Section 5 presents recommendations on the appropriateness of the stochastic Galerkin and collocation methods for this type of problem.

2. Problem formulation.

We consider a nonlinear parabolic equation,

$$\frac{\partial \theta(u)}{\partial t} = \nabla \cdot [K(u)(\nabla u - \mathbf{g})], \qquad \mathbf{x} = (x_1, \ldots, x_d)^\top \in D, \quad t \in (0, T], \tag{2.1}$$

defined on a d-dimensional spatial domain $D \subset \mathbb{R}^d$ $(1 \leq d \leq 3)$ during time interval $(0, T]$. Here $\theta(u)$ and $K(u)$ are continuous functions of the state variable $u(\mathbf{x}, t) : D \times [0, T] \to \mathbb{R}$, both of which are parameterized with elements of a set of P spatially heterogeneous model parameters $\Lambda(\mathbf{x}) = \{\Lambda_1(\mathbf{x}), \ldots, \Lambda_P(\mathbf{x})\}$. In the context of two-phase fluid flow in porous media, (2.1) is referred to as the Richards equation, wherein the state variable $u(\mathbf{x}, t)$ is fluid pressure, $\theta(u; \Lambda)$ is the fluid content (saturation) of a porous medium, $K(u; \Lambda(\mathbf{x}))$ is the saturation-dependent hydraulic conductivity of the medium, and $\mathbf{g}$ is the unit vector collinear with the gravitational force. A plethora of constitutive relations for $\theta(u; \Lambda)$ and $K(u; \Lambda)$ share at least three parameters (see, e.g., [30] and the references therein): saturated hydraulic conductivity K_s $(\equiv \Lambda_1)$, a shape parameter α $(\equiv \Lambda_2)$, and fluid content at full saturation

(porosity) θ_s ($\equiv \Lambda_3$). Specifically, the saturation-dependent hydraulic conductivity $K(u;\Lambda)$ and fluid retention curve $\theta(u;\Lambda)$ can be written in a generic form,

$$K(u;\Lambda) = K_s \begin{cases} K_r(u;\alpha,\ldots), & u<0, \\ 1, & u\geq 0, \end{cases} \qquad \theta(u;\Lambda) = \begin{cases} \theta(u;\alpha,\ldots), & u<0, \\ \theta_s, & u\geq 0, \end{cases} \tag{2.2}$$

where the dots denote the remaining set of parameters specific to a particular constitutive law, and the monotonically increasing function $K_r(u;\alpha,\ldots) : \mathbb{R}^- \to [0,1]$ is called relative hydraulic conductivity and reaches its maximum value of $K_r = 1$ at fluid pressure $u = 0$.

A problem formulation is completed by specifying initial conditions

$$u(\mathbf{x},0) = u_{in}(\mathbf{x}), \qquad \mathbf{x}\in D, \tag{2.3}$$

and boundary conditions

$$\mathcal{B}(u,\mathbf{x},t) = b(\mathbf{x},t), \qquad \mathbf{x}\in\partial D, \quad t\in(0,T]. \tag{2.4}$$

Here $u_{in}(\mathbf{x})$ is the initial distribution of fluid pressure in the flow domain D, $\mathcal{B}$ is the boundary operator representing (Dirichlet and/or Neumann) boundary conditions on various segments of the domain's boundary ∂D, and $b(\mathbf{x},t)$ denotes the corresponding boundary functions.

The constitutive laws (2.2), model parameters $\Lambda(\mathbf{x})$, forcing terms $u_{in}(\mathbf{x})$ and $b(\mathbf{x},t)$, and the domain of definition $D\times[0,T]$ are such that the BVP (2.1)–(2.4) admits a unique solution $u(\mathbf{x},t)$ (e.g., [24]).

2.1. Probabilistic model formulation. Heterogeneity of the coefficients $\Lambda(\mathbf{x})$, practical impossibility of measuring their values at every point $\mathbf{x}\in D$, and ubiquitous measurement errors at points where data are available render the spatio-temporal distribution of these input parameters uncertain. As a consequence, the task of obtaining a deterministic, unique solution of (2.1)–(2.4) is in general impossible. We therefore adopt a probabilistic framework, which treats the uncertain coefficients $\Lambda(\mathbf{x})$ as random fields whose sample statistics are inferred from spatially distributed measurements by invoking ergodicity (see, e.g., [30] and the references therein). This renders a solution $u(\mathbf{x},t)$ random as well.

Let $(\Omega,\mathcal{F},\mathcal{P})$ be a complete probability triple, where Ω is the sample space, $\mathcal{F}\subseteq 2^\Omega$ is the σ-algebra of events, and $\mathcal{P}:\mathcal{F}\to[0,1]$ is the probability measure. The domain of definition of the measurable parameters $\Lambda(\mathbf{x})$ is extended to the sample space Ω, i.e., $\Lambda\equiv\Lambda(\mathbf{x},\omega) : D\times\Omega\to\mathbb{R}$. As a consequence, (2.1)–(2.4) gives rise to the following stochastic boundary-value problem (sBVP): find a stochastic measurable function $u\equiv u(\mathbf{x},t,\omega) : D\times[0,T]\times\Omega\to\mathbb{R}$ that satisfies almost surely

$$\frac{\partial\theta(u,\mathbf{x},\omega)}{\partial t} = \nabla\cdot[K(u,\mathbf{x},\omega)(\nabla u-\mathbf{g})], \quad \mathbf{x}\in D, \quad t\in(0,T], \tag{2.5}$$

$$\mathcal{B}(u,\mathbf{x},t) = b(\mathbf{x},t), \qquad \mathbf{x}\in\partial D, \quad t\in(0,T], \tag{2.6}$$

and deterministic initial condition (2.3).

We assume that for each random parameter $\Lambda_i(\mathbf{x},\omega)$ there exists a transformation $f_i = f_i(\Lambda_i)$, where each $f_i(\mathbf{x},\omega)$ $(i=1,\ldots,P)$ is a zero-mean, square-integrable Gaussian random

field. These fields are allowed to be cross-correlated, with a known cross-covariance kernel $C_{ij}(\mathbf{x},\mathbf{y}) = \langle f_i(\mathbf{x},\omega) f_j(\mathbf{y},\omega)\rangle$. Here and below, the symbol $\langle\cdot\rangle$ denotes the ensemble average. Knowledge of the cross-covariance stems from statistical analysis of available measurements of the fields f_i (e.g., via a variogram analysis and co-Kriging, which has become standard in studies of flow and transport in heterogeneous porous media).

2.2. Approximate model parameterization. Let the random fields $f_i(\mathbf{x},\omega)$ $(i = 1,\dots,P)$ be defined over a probability space $(\Omega,\mathcal{F},\mathcal{P})$ and indexed over a closed and bounded domain D, with continuous autocovariance functions $C_i(\mathbf{x},\mathbf{y}) = \langle f_i(\mathbf{x},\omega) f_i(\mathbf{y},\omega)\rangle$. Then each $f_i(\mathbf{x},\omega)$ can be represented using a K-L expansion [5]

$$f_i(\mathbf{x},\omega) = \sum_{k=1}^{\infty} \sqrt{\gamma_k^i}\,\phi_k^i(\mathbf{x})\xi_k^i(\omega), \qquad i \in 1,\dots P, \tag{2.7}$$

where γ_k^i and $\phi_k^i(\mathbf{x})$ $(k = 1,2,\dots)$ are, respectively, the eigenvalues and the eigenfunction of the autocovariance function $C_i(\mathbf{x},\mathbf{y})$, and $\{\xi_k^i\}_{k=1}^{\infty}$ is a set of independent, identically distributed standard Gaussian random variables. The covariance between random variables from the two sets, e.g., ξ_k^i and ξ_m^j, obey the relation

$$\langle \xi_k^i \xi_m^j\rangle = (\gamma_k^i \gamma_m^j)^{-1/2} \int_D \int_D C_{ij}(\mathbf{x},\mathbf{y})\phi_k^i(\mathbf{x})\phi_m^j(\mathbf{y})\,\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}. \tag{2.8}$$

If the underlying Gaussian fields $f_i(\mathbf{x},\omega)$ and $f_j(\mathbf{x},\omega)$ are mutually independent, then $C_{ij}(\mathbf{x},\mathbf{y}) = \delta_{ij} C_i(\mathbf{x},\mathbf{y})$ and $\langle \xi_k^i \xi_m^j\rangle = \delta_{ij}\delta_{km}$, where δ_{ij} is the Kronecker delta function.

To render the ith K-L expansion (2.7) computable, one has to truncate it by retaining the S_i leading terms, i.e., to approximate $f_i(\mathbf{x})$ with

$$\hat{f}_i(\mathbf{x},\omega) = \sum_{k=1}^{S_i} \sqrt{\gamma_k^i}\,\phi_k^i(\mathbf{x})\xi_k^i(\omega), \qquad i = 1,\dots,P. \tag{2.9}$$

Without loss of generality, we set $S_1 = S_2 = \cdots = S_P \equiv S$. For a given variance of $f_i(\mathbf{x},\omega)$, the value of S necessary to approximate (2.7) with a given accuracy depends on the rate of decay of the eigenvalues γ_k^i. The decay rate is given by the regularity, in the sense of Definition 2.2 in [7], of the autocovariance kernel $C_i(\mathbf{x},\mathbf{y})$. Analyticity gives rise to an exponential or superexponential decay, while piecewise regularity yields algebraic decay. Regardless of the regularity type, for a given accuracy, the value of S increases as the autocorrelation length of $f_i(\mathbf{x},\omega)$ decreases.

If the input fields $f_i(\mathrm{x},\omega)$ $(i = 1,\dots,P)$ are mutually independent,[1] then the random variables $\{\xi_k^i\}_{k=1}^{S}$ in their K-L expansions form an N-dimensional vector

$$\boldsymbol{\xi}_N(\omega) = (\{\xi_k^1\}_{k=1}^{S}, \{\xi_k^2\}_{k=1}^{S}, \dots, \{\xi_k^P\}_{k=1}^{S})^{\top}, \qquad N = PS. \tag{2.10}$$

This vector is characterized by a standard joint multi-Gaussian probability density function

$$\rho(\mathbf{s}) = (2\pi)^{-N/2} \exp\left(-\frac{1}{2}\mathbf{s}^{\top}\mathbf{s}\right) \tag{2.11}$$

[1] The assumption of mutually independent fields can be relaxed by a proper construction of K-L expansions for multicorrelated fields (e.g., [5]).

with support $\Gamma = \mathbb{R}^N$. The number N is often referred to as a *stochastic dimension* stemming from the finite-dimensional noise approximation (2.9).

2.3. Formulation of approximate sBVP. Approximation of the parameter fields $f_i(\mathbf{x}, \omega)$ in (2.7) with their truncated counterparts $\hat{f}_i(\mathbf{x}, \boldsymbol{\xi}_N)$ in (2.9) transforms (2.5) into

$$\frac{\partial \theta(u_N, \mathbf{x}, \boldsymbol{\xi}_N)}{\partial t} = \nabla \cdot [K(u_N, \mathbf{x}, \boldsymbol{\xi}_N)(\nabla u_N - \mathbf{g})]. \qquad \mathbf{x} \in D, \quad \boldsymbol{\xi}_N \in \Gamma, \quad t \in (0, T]. \tag{2.12}$$

This equation is subject to the deterministic initial (2.3) and boundary (2.4) conditions. According to the Doob–Dynkin lemma [27], a solution u_N of this sBVP is a function of $\boldsymbol{\xi}_N(\omega)$, i.e., $u_N(\mathbf{x}, t, \omega) = u_N(\mathbf{x}, t, \boldsymbol{\xi}_N(\omega)) : D \times [0, T] \times \Gamma \to \mathbb{R}$. The probability triple $(\Omega, \mathcal{F}, \mathcal{P})$ is thus replaced with the triple $(\Gamma, B, \rho(\boldsymbol{\xi}_N)\mathrm{d}\boldsymbol{\xi}_N)$, where B is the Borel σ-algebra formed by all open subsets of Γ. When viewed as a function of $\boldsymbol{\xi}_N(\omega)$, $u_N(\mathbf{x}, t, \boldsymbol{\xi}_N)$ is referred to as a *stochastic response surface* or simply *response*.

This problem formulation introduces two sources of error into estimators of the statistics of $u(\mathbf{x}, t, \omega)$. First, transition from (2.5) to (2.12), i.e., from u to u_N, introduces a bias error ϵ_{KL}, which can only be reduced by increasing the number S of K-L terms of the expansion of each random field. Bounds for ϵ_{KL} have been derived in [4] for the steady-state linear counterpart of (2.1), i.e., for $K(u, \mathbf{x}, \omega) = K(\mathbf{x}, \omega)$. To the best of our knowledge, no similar results exist for nonlinear elliptic or parabolic problems, for which this error must be investigated numerically. Second, numerical solution of (2.12) introduces an *estimation* or *sampling error* ϵ_{est}, which depends on the UQ method employed (e.g., MCS or SC). The third source of error is due to a numerical discretization of each deterministic solve; it can be reduced by employing a higher-order numerical scheme or a finer mesh, and it is assumed to be negligibly small relative to both ϵ_{KL} and ϵ_{est}.

Let $\langle u \rangle$ and $\langle u_N \rangle$ denote the ensemble means of $u(\mathbf{x}, t, \omega)$ in (2.5) and $u_N(\mathbf{x}, t, \boldsymbol{\xi}_N)$ in (2.12), respectively. Furthermore, let $\hat{u}_N$ be a computable (e.g., by a finite number of MC realizations) estimate of $\langle u_N \rangle$. Then, the error introduced by approximating the true $\langle u \rangle$ with its computable estimate $\hat{u}_N$ is bounded by $|\langle u \rangle - \hat{u}_N| \leq |\langle u \rangle - \langle u_N \rangle| + |\langle u_N \rangle - \hat{u}_N| = \epsilon_{\mathrm{KL}} + \epsilon_{\mathrm{est}}$. We use numerical experiments to investigate the errors ϵ_{KL} and ϵ_{est} when a SC method is used to obtain $\hat{u}_N$ from (2.12).

3. Stochastic collocation. Given the PDF (2.11), central moments of the stochastic state variable $u_N(\mathbf{x}, t, \boldsymbol{\xi}_N)$ are defined as weighted integrals over the support of $\boldsymbol{\xi}_N$. For example, the mean and variance of $u_N(\mathbf{x}, t, \boldsymbol{\xi}_N)$ are

$$\langle u_N(\mathbf{x}, t) \rangle = \int_\Gamma u_N(\mathbf{x}, t, \mathbf{s}) \rho(\mathbf{s}) \mathrm{d}\mathbf{s} \tag{3.1}$$

and

$$\sigma^2_{u_N}(\mathbf{x}, t) = \int_\Gamma u_N^2(\mathbf{x}, t, \mathbf{s}) \rho(\mathbf{s}) \mathrm{d}\mathbf{s} - \langle u_N(\mathbf{x}, t) \rangle^2, \tag{3.2}$$

respectively. SC methods approximate such N-dimensional integrals using quadrature rules with a properly chosen set of collocation points (or "nodes") in Γ and a corresponding set of

collocation weights. Specifically, an N-dimensional weighted integral of an integrable function $f(\mathbf{s})$,

$$\mathcal{I}_N[f] = \int_\Gamma f(\mathbf{s})\rho(\mathbf{s})\,\mathrm{d}\mathbf{s}, \tag{3.3}$$

is approximated via a quadrature formula

$$\mathcal{I}_N[f] \approx \mathcal{Q}_N[f] = \sum_{i=1}^{Q} w_i f(\mathbf{s}_i), \tag{3.4}$$

where $\mathbf{s}_i$ and ω_i are the nodes and weights of the quadrature rule, respectively; and Q is the number of collocation nodes. Employing (3.4) to approximate (3.1) and (3.2), we obtain SC estimates of the mean and variance,

$$\hat{u}_N^{\mathrm{SC}}(\mathbf{x},t) = \sum_{i=1}^{Q} w_i\, u(\mathbf{x},t,\mathbf{s}_i), \tag{3.5}$$

$$\hat{\sigma}_{u_N}^{2,\mathrm{SC}}(\mathbf{x},t) = \sum_{i=1}^{Q} w_i\, [u(\mathbf{x},t,\mathbf{s}_i)]^2 - [\hat{u}_N^{\mathrm{SC}}(\mathbf{x},t)]^2, \tag{3.6}$$

where $u(\mathbf{x},t,\mathbf{s}_i)$ is a solution of (2.12), subject to (2.3) and (2.4), with the random vector $\boldsymbol{\xi}_N$ taking a deterministic value or "collocation node" $\mathbf{s}_i$.

This approach is a sampling technique whose *raison d'être* is the claim that, by a judicious selection of quadrature rules, it requires a smaller number of terms, Q, than MCS to achieve the same accuracy. It is worth recalling that an MCS estimator of $\langle u_N \rangle$ is given by

$$\hat{u}_N^{\mathrm{MC}}(\mathbf{x},t) = \frac{1}{Q}\sum_{i=1}^{Q} u_N(\mathbf{x},t,\boldsymbol{\eta}_i), \tag{3.7}$$

where $\{\boldsymbol{\eta}_i\}_{i=1}^{Q}$ is a set of Q realizations of $\boldsymbol{\xi}_N$, sampled from the multivariate Gaussian distribution (2.11). The MCS estimation error decays as $\epsilon_{\mathrm{est}}^{\mathrm{MC}} = \sigma_{u_N}/\sqrt{Q}$, where σ_{u_N} is the standard deviation of u_N. It is independent of the stochastic dimension N, which is an advantage for high-dimensional problems. On the downside, MCS have a slow convergence rate, $1/\sqrt{Q}$. The goal of SC methods is to improve on this convergence rate.

3.1. Selection of quadrature rules. Construction of an appropriate interpolation quadrature rule is a key part of the SC approach. In what follows, we summarize both requirements for the selection of rules and properties of several families of quadrature rules. We focus on two types of quadrature rules based on (unidimensional) polynomial interpolation: product rules and sparse grid rules. Let $A(q,N)$ denote a family of rules for N-dimensional integrals, where q is a construction parameter that defines the family. If $H(q,N)$ is the associated set of quadrature nodes, then $Q \equiv \dim H(q,N)$.

Let $\mathbb{P}(l,N)$ be the space of all N-dimensional polynomials of total degree at most l. A quadrature rule $A(q,N)$ is said to have a degree of exactness l if [26]

$$\mathcal{I}_N[f] = A(q,N)[f] \quad \forall f \in \mathbb{P}(l,N). \tag{3.8}$$

The value of l depends on the number of stochastic dimensions, N, and the construction parameter, q, i.e., $l = l(q, N)$, such that l increases with q for a fixed N. A general stochastic response cannot be represented exactly by polynomials of any order. Therefore, the use of quadrature rules introduces an approximation error that depends on the accuracy with which a polynomial with $l(q, N)$ can interpolate $u_N(\mathbf{x}, t, \boldsymbol{\xi}_N)$ for given $\mathbf{x}$ and t. The degree of exactness required to achieve a certain accuracy depends on the smoothness of the response: the smoother the integrand, the lower the polynomial's total degree.

Estimates of the convergence rate of quadrature rules are available for certain classes of smoothness (e.g., [25, 35]), such as the space of functions with bounded mixed derivatives defined over the hyper-cube $[0,1]^N$, i.e., $F_N^r = \{f : [0,1]^N \to \mathbb{R} \mid D^{(\boldsymbol{\alpha})} f$ continuous if $\alpha_i \leq r$ for all $i\}$, where $D^{(\boldsymbol{\alpha})}$ denotes a partial derivative and $\boldsymbol{\alpha} \in \mathbb{N}_0^N$. Unfortunately, to the best of our knowledge, no estimates of the convergence rate are available for quadrature rules based on polynomial interpolation for function spaces defined over $\mathbb{R}^N$.

Ideal quadrature rules for the sBVP for (2.12) would satisfy the following conditions:

1. *Scalability*: The number of nodes Q must grow as slowly as possible with both the stochastic dimension N and the construction parameter q (or, equivalently, with the degree of exactness, l). Scalability with respect to N and l allows one to control for the bias introduced by the finite-dimensional noise approximation and for the estimation error, respectively. For odd degrees of exactness, $l = 2k + 1$, a lower bound on Q is $N^k/k!$ for fixed k and large N, and k^N for fixed N and large k [26]. We look for constructions with scalability properties as close to optimal as possible.
2. *Stability*: In general, quadrature weights can be negative, and their absolute values can be arbitrarily large. A sequence of large positive and negative weights would lead to cancellation errors in computation of the weighted averages (3.5) and (3.6). As a result, we look for quadrature rules with $R \equiv \sum_{i=1}^{M} |w_i| \geq 1$ as close to 1 as possible.
3. Not required but desirable is to use rules with embedded sets of nodes, i.e., nested rules for which $H(q, N) \subset H(q + 1, N)$. This allows one to increase the construction parameter q while preserving previous sets of nodes, so that the work invested in computing $u_N(\mathbf{x}, t, \mathbf{s}_i)$ is not lost.

Product quadrature rules are obtained by the tensor product of N univariate rules (e.g., Gauss rules), one for each stochastic dimension. The total number of nodes for product Gaussian rules with the degree of exactness $l = 2k+1$ and the same number of nodes in each of the N dimensions is $Q = (k+1)^N$ [26]. In other words, the total number of nodes increases exponentially with N, as opposed to the polynomial lower bound N^k. This exponential growth is sometimes called the "curse of dimensionality" because reduction of the finite-dimensional bias defined in section 2.1 exponentially increases the computational cost. On the other hand, these rules behave optimally for a fixed N and increasing k. Their convergence rate improves with higher smoothness of the integrand and deteriorates with an increasing number of stochastic dimensions. Their stability is optimal, i.e., $R = 1$, as they result from the tensor product of (optimally stable) one-dimensional rules. For functions $f \in F_N^r$, the estimation error is $\epsilon_{\text{est}} = \mathcal{O}(Q^{-r/N})$, provided q is sufficiently large.

Sparse grid rules are constructed via the Smolyak algorithm [28] from one-dimensional quadrature rules for each stochastic dimension. The total number of nodes for sparse grid rules with the degree of polynomial exactness $l = 2k + 1$ is $Q \approx 2^k N^k/k!$. This polynomial

increase of Q with N is a significant improvement over the exponential scalability of product rules [26]. For fixed N and increasing k, $Q \approx k^N(\log k)^{N-1}$, which is close to optimal. For functions $f \in F_N^r$, the estimation error is $\epsilon_{\text{est}} = \mathcal{O}(Q^{-r}(\ln Q)^{(N-1)(r+1)})$. The convergence rate of sparse grid rules depends weakly on N and strongly on the integrand's smoothness. Since the Smolyak algorithm generally introduces negative weights, R for sparse grids is larger than 1 and typically increases with both l and N.

Sparse grid quadrature rules can be either *global*, with fixed quadrature nodes, or *locally adaptive*. Global constructions are defined chiefly by the quadrature weight for which they are designed and allow for a limited level of adaptivity. For example, if the smoothness of the stochastic response exhibits a directional dependence in the sample space, one can introduce nonisotropic quadrature rules that locate quadrature nodes preferentially along more critical directions [10]. Locally adaptive rules allocate nodes throughout the sample space attending to the local features of the response, such as sharp gradients or discontinuities [1, 2, 15, 23].

3.2. Quadrature rules for multivariate-Gaussian weights. The probability density function $\rho(\mathbf{s})$ of the stochastic vector $\boldsymbol{\xi}_N$ determines the type of one-dimensional rules used to construct a multidimensional sparse grid. Selection of rules for multivariate-Gaussian $\rho(\mathbf{s})$ is more limited than that for the uniform $\rho(\mathbf{s})$ on the hyper-cube. Furthermore, as noted earlier, no good error bounds exist for rules defined on unbounded domains.

For multivariate-Gaussian $\rho(\mathbf{s})$, the natural choice is to employ one-dimensional Gauss-Hermite (GH) quadratures in each stochastic direction. While in some cases one can construct embedded Kronrod extensions of the GH rules, embedded rules of an arbitrary degree of exactness do not exist [17]. As a consequence, the resulting sparse grid GH rules are not embedded.

Alternative constructions have been developed by restricting the degree of exactness l. For example, the Genz–Keister (GK) rules [9] exist for $l \leq 51$. They are embedded and result in better stability and lower node count than their GH counterparts. More efficient rules have been developed for $l = 5$ and $l = 7$ for both general $\rho(\mathbf{s})$ [22, 14] and $\rho(\mathbf{s}) > 0$ [33, 18]. The GK rules for $l > 13$ are most efficient, in the sense that they require the fewest nodes for a given degree of accuracy.

3.3. Quadrature rules for polynomial chaos expansions. The procedure described above enables one to compute ensemble moments of QoI in a nonintrusive way, i.e., by utilizing an existing deterministic solver for the PDE under consideration. gPC expansions [37] provide an intrusive alternative (requiring modifications of the deterministic PDE solver), which allows computing a full PDF of QoI. To regain nonintrusiveness, one can approximate the coefficients in a gPC expansion with the weighted quadrature rules described above by employing the projection integral definition of these coefficients [37]. The error of the resulting gPC approach is that of the truncated gPC expansion and that of the quadrature rule.

4. Numerical experiments. In the absence of theoretical error estimators for numerical solutions of nonlinear elliptic (steady-state) and parabolic (time-dependent) advection-diffusion equations, we perform a series of numerical experiments to investigate the accuracy and robustness of SC methods as well as their efficiency vis-à-vis MCS. Our focus is on the effect of statistical characteristics, especially variance, of random coefficients on the performance

of SC methods. The impact of the stability of quadrature rules is also considered.

Grounding our investigation in applications to two-phase fluid flow in heterogeneous porous media, we focus on uncertainty in the values of saturated hydraulic conductivity $K_s(\mathbf{x})$ and shape parameter $\alpha(\mathbf{x})$, while disregarding uncertainty in the other hydraulic parameters which typically exhibit smaller spatial variability. Theoretical considerations and experimental evidence (see, e.g., [31] and the references therein) suggest that the random fields $K_s(\mathbf{x},\omega)$ and $\alpha(\mathbf{x},\omega)$ have log-normal distributions. Consequently, we model $Y(\mathbf{x},\omega) = \ln K_s(\mathbf{x},\omega)$ and $\beta(\mathbf{x},\omega) = \ln \alpha(\mathbf{x},\omega)$ as stationary, square-integrable Gaussian fields. Let $K_g = \exp(\langle Y\rangle)$ and $\alpha_g = \exp(\langle\beta\rangle)$ denote the geometric means of K_s and α, respectively. Then, $K_s = K_g \exp(Y')$ and $\alpha = \alpha_g \exp(\beta')$, where $Y'(\mathbf{x},\omega)$ and $\beta'(\mathbf{x},\omega)$ are zero-mean Gaussian fluctuations with variances σ_Y^2 and σ_β^2, and correlation lengths λ_Y and λ_β. Following the standard practice in the field (see [31] and the references therein) we assume the random fields $Y'(\mathbf{x},\omega)$ and $\beta'(\mathbf{x},\omega)$ to have exponential autocovariances

$$C_Y(\mathbf{x},\mathbf{y}) \equiv \langle Y'(\mathbf{x},\omega)Y(\mathbf{y},\omega)\rangle = \sigma_Y^2 \mathrm{e}^{-|\mathbf{x}-\mathbf{y}|/\lambda_Y} \quad \text{and} \quad C_\beta(\mathbf{x},\mathbf{y}) = \sigma_\beta^2 \mathrm{e}^{-|\mathbf{x}-\mathbf{y}|/\lambda_\beta}, \tag{4.1}$$

These autocovariances are only piecewise regular in the sense of [7], and therefore their K-L eigenvalues exhibit algebraic (power of -2) decay. Finally, we assume the random fields $Y'(\mathbf{x},\omega)$ and $\beta'(\mathbf{x},\omega)$ to be mutually uncorrelated.

4.1. Elliptic problems. We start with an elliptic problem that represents a steady one-dimensional version of (2.1),

$$\frac{\mathrm{d}}{\mathrm{d}z}\left[K(z,u,\omega)\frac{\mathrm{d}(u-z)}{\mathrm{d}z}\right] = 0, \qquad 0 < z < 1. \tag{4.2}$$

It is supplemented with Gardner's (exponential) constitutive relation (see [31] and the references therein)

$$K(z,u,\omega) = K_s(z,\omega)\begin{cases} \mathrm{e}^{\alpha(z,\omega)u(z,\omega)}, & u \le 0, \\ 1, & u > 0. \end{cases} \tag{4.3}$$

and is subject to boundary conditions

$$K\frac{\mathrm{d}u}{\mathrm{d}z}(z=0,\omega) = -r, \qquad u(z=1,\omega) = 0, \tag{4.4}$$

where r is a prescribed (deterministic) boundary flux. This sBVP was solved for various combinations of the statistical properties of the Gaussian fields $Y = \ln K$ and $\beta \ln \alpha$, which are summarized in Table 1.

Our QoI is $\langle u(0)\rangle$, mean fluid pressure $\langle u(z)\rangle$ at the porous tube's inlet $z = 0$. The goal of the subsequent numerical experiments is to estimate both the finite-dimensional noise bias, ϵ_{KL}, and the estimation error, ϵ_{est}, introduced by the SC computation of this QoI. The SC estimate, $\hat{u}_N^{\mathrm{SC}}(0)$, is evaluated by using the truncated S-term K-L representations (2.9) of the random fluctuations $Y'(z,\omega)$ and $\beta'(z,\omega)$ and by employing the GK quadrature introduced in section 3.2. Recall that the stochastic dimension of this formulation, or the total number of the random input variables $\boldsymbol{\xi}_N(\omega)$, is $N = 2S$.

Table 1
Model parameters for elliptic problem (4.2)–(4.4).

Test	σ_Y^2	σ_β^2	λ_Y	λ_β	K_g	α_g	r
(a)	0.1	0.1	1.0	1.0	1.0	5.0	0.1
(b)	0.5	0.5	1.0	1.0	1.0	5.0	0.1
(c)	0.1	0.1	0.3	0.3	1.0	5.0	0.1
(d)	0.5	0.5	0.3	0.3	1.0	5.0	0.1

For a given realization (node) $\mathbf{s}_i$ of the random input vector $\boldsymbol{\xi}_N(\omega)$, the corresponding $u_N(z, \mathbf{s}_i)$ is computed by using a cell-centered finite-volume discretization of (4.2)–(4.4); the resulting system of nonlinear algebraic equations is solved by using the KINSOL library [13] implementation of the modified Newton's method with line search. The computational domain $D = [0, 1]$ is discretized into M uniform cells of length $\Delta z = 1/M = 0.01$. Fluxes across the finite volume cells are approximated by using a two-point flux-approximation formula. The pressure $u(z, \mathbf{s}_i)$ at $z = 0$ is then linearly interpolated given the flux r and $u(\Delta z/2, \mathbf{s}_i)$.

In lieu of the exact solution $\langle u(0)\rangle$, we employ its MC estimator as a surrogate. A surrogate of the exact expectation $\langle u_N(0)\rangle$, denoted by $\hat{u}_N^{\mathrm{MC}(n)}(0)$, represents an MC estimate of $\langle u_N(0)\rangle$ accurate to n significant digits with a confidence of 95%. These MC estimators, accurate to $n = 3$ and 4 significant digits, are reported in Table 2 for $N = 10$, 20, 30, and 40. To render the K-L truncation error ϵ_{KL} negligible, we conducted MC simulations for $N = 1600$. The resulting MC estimators, $\hat{u}_{\mathrm{ref}}^{\mathrm{MC}(n)}(0)$, serve as surrogates for $\langle u(0)\rangle$. It took on the order of $10^7 - 10^8$ MC realizations to obtain the MC estimators accurate to $n = 3$ and 4 significant digits.

Table 2
MC estimators of $\langle u(0)\rangle$ accurate to n significant digits with 95% confidence. Tests (a)–(d) use the corresponding parameter values listed in Table 1.

		$\hat{u}_N^{\mathrm{MC}(n)}(0)$			
S	$N = 2S$	Test (a), $n = 4$	Test (b), $n = 4$	Test (c), $n = 3$	Test (d), $n = 3$
5	10	−0.4458	−0.4306	−0.438	−0.408
10	20	−0.4443	−0.4252	−0.434	−0.392
15	30	−0.4438	−0.4234	−0.432	−0.387
20	40	−0.4437	−0.4224	−0.431	−0.384
800	ref	−0.4430	−0.4198	−0.429	−0.375

The absolute difference $|\hat{u}_N^{\mathrm{MC}(n)}(0) - \hat{u}_{\mathrm{ref}}^{\mathrm{MC}(n)}(0)|$ serves as a surrogate for the bias error ϵ_{KL}. Figure 1 reveals that it exhibits a power law behavior of the form $\epsilon_{\mathrm{KL}} = C(\sigma_Y, \sigma_\beta, \lambda_Y, \lambda_\beta)N^{-\gamma}$, where the decay rate γ is independent of the variance and correlation length of the random fields Y and β, while the prefactor C depends on these statistical properties. Specifically, C increases with the variances σ_Y^2 and σ_β^2 and decreases with the correlation lengths λ_Y and λ_β.

This experimentally observed bound on the bias error due to truncation of the K-L expansions of the random coefficients is consistent with the theoretical result, $\epsilon_{\mathrm{KL}} \le C(\sigma_Y, \lambda_Y)N^{-\gamma}$, derived in [4] for the linear sBVP corresponding to (4.3) with $\alpha(\mathrm{x}) \equiv 0$. The latter power law

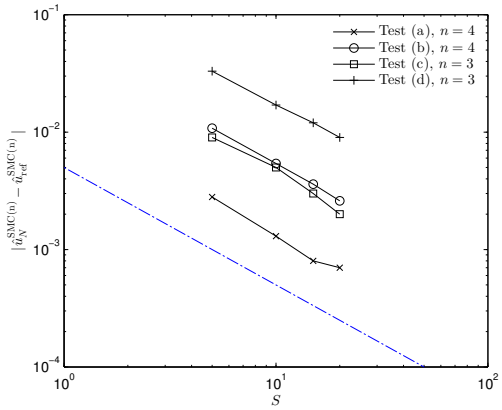

Figure 1. *Dependence of the bias error, $\epsilon_{KL} = |\hat{u}_N^{\mathrm{MC}(n)}(0) - \hat{u}_{\mathrm{ref}}^{\mathrm{MC}(n)}(0)|$, on the number of terms in the K-L expansions of Y and β for the test cases whose parameter values are reported in Table 1. The dashed line indicates the power law with the exponent $\gamma = 1$.*

behavior was proved to stem from the algebraic decay of the eigenvalues of the K-L expansion for the exponential covariance functions (4.1).

Figure 1 establishes that the power $\gamma \approx 1$, i.e., the bias error ϵ_{KL}, decays slowly with the number of terms in the K-L expansions, $\epsilon_{\mathrm{KL}} \sim 1/N$. This corresponds to the polynomial decay of the eigenvalues of the C^1-discontinuous exponential kernel reported in [4]. Such a slow decay of the bias error impairs one's ability to compute estimators with prescribed total error (see section 2.1). Indeed, it implies that the reduction of ϵ_{KL} by a factor of m necessitates the increase of the number of terms in the K-L expansions, N, by the factor of m and the number of collocation nodes, Q, by the factor of $m^* > m$. As discussed in section 2.1, the ratio m^*/m increases with the polynomial degree of exactness of the sparse grid quadrature rule employed.

The absolute difference $|\hat{u}_N^{\mathrm{SC}}(0) - \hat{u}_N^{\mathrm{MC}(n)}(0)|$ serves as a surrogate for the estimation error ϵ_{est}. Figure 2 exhibits ϵ_{est} for the four test cases (see Table 1) and for several values of N and the polynomial degree of exactness (i.e., the resulting number of collocation nodes, Q). Also shown is the estimation error of MCS, $\epsilon_{\mathrm{est}}^{\mathrm{MC}}$, defined as the root-mean-square error (RMSE). Both ϵ_{est} and the rate of its decay with the number of deterministic solves (collocation points Q) compare favorably with their MCS counterparts when variances of the input parameters are small ($\sigma_Y^2 = \sigma_\beta^2 = 0.1$ in Figures 2(a) and 2(c)). In other words, the SC method is computationally more efficient than the MCS. When the variances increase to $\sigma_Y^2 = \sigma_\beta^2 = 0.5$ (Figures 2(b) and 2(d)), the SC method outperforms the MCS only for large correlation lengths (Figure 2(b), in which $\lambda_Y = \lambda_\beta = 1.0$ coincides with the length of the computational domain). For smaller correlation lengths (e.g., $\lambda_Y = \lambda_\beta = 0.3$ in Figure 2(d), the convergence rate of the SC estimation error is less than 1/2 for $S \geq 15$; therefore, the MCS require fewer realizations than the SC method does to achieve a given estimation error.

These findings undermine the notion that the performance of stochastic FEMs vis-à-vis MCS is insensitive to the degree of uncertainty in model coefficients, i.e., to their variances or coefficients of variation.

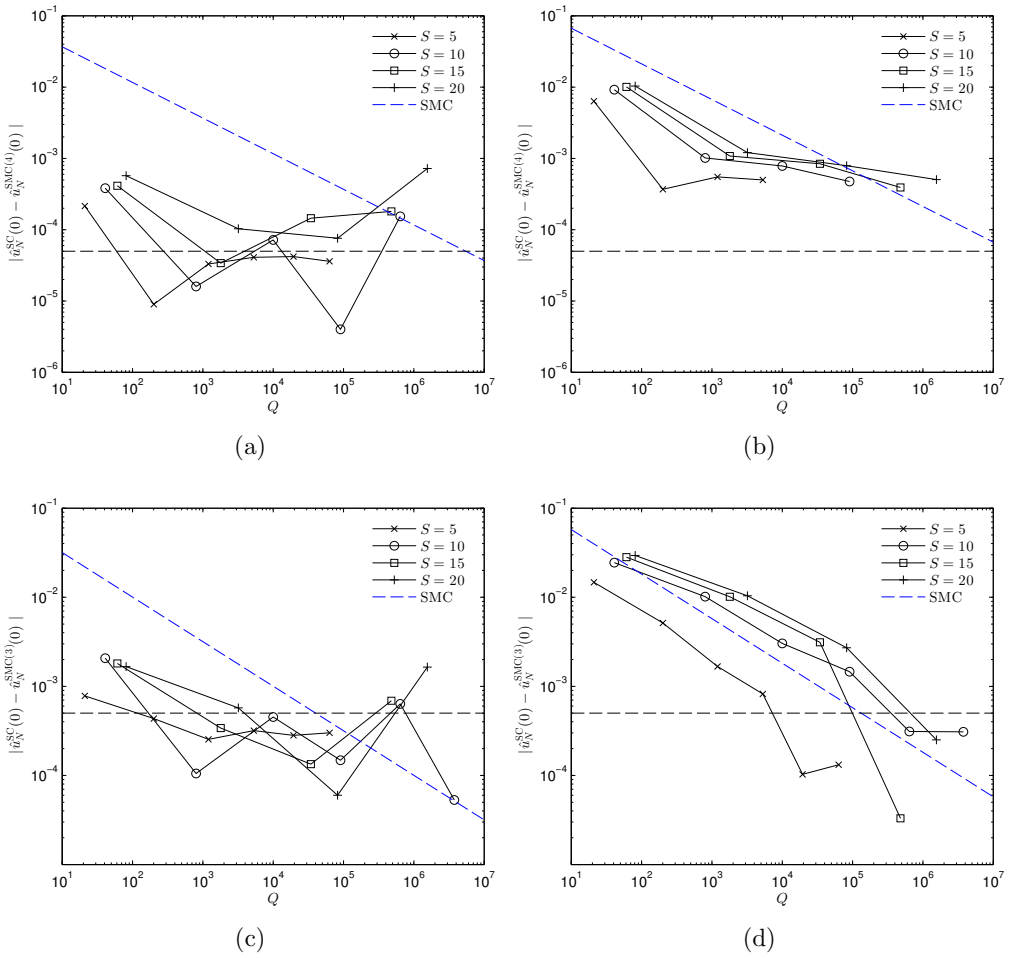

Figure 2. *Dependence of the estimation error, $\epsilon_{est} = |\hat{u}_N^{\mathrm{SC}}(0) - \hat{u}_N^{\mathrm{MC}(n)}(0)|$, on the number of SC nodes, Q (or the number of MC realizations). Subfigures* (a) *through* (d) *correspond to the test cases in Table 1. The dashed black lines indicate half-width corresponding to the number of accurate significant digits, n. The dashed blue line indicates the decay of the RMSE of the simple MC simulations (SMC), ϵ_{est}^{MC}.*

Figure 3 elaborates this point further by presenting the total error, $\epsilon = \epsilon_{\mathrm{KL}} + \epsilon_{\mathrm{est}}$, as a function of the number of collocation nodes, Q. This error is compared with the total error of the MCS for $N = 40$ (the RMS estimation error, plus the corresponding bias error). For small variances and large correlation lengths (the test cases in Figures 3(a)–3(c), the SC method requires fewer deterministic solves to reach an achievable total error target (ATET) than the MCS does. (We define an ATET as an error target that is larger than the smallest bias error considered, which in these experiments corresponds to $N = 40$.) The situation is reversed for the more typical case (correlation lengths that are a fraction of a computational domain's size, Figure 3(d), in which for any achievable error target the MCS are more efficient than the

SC method.

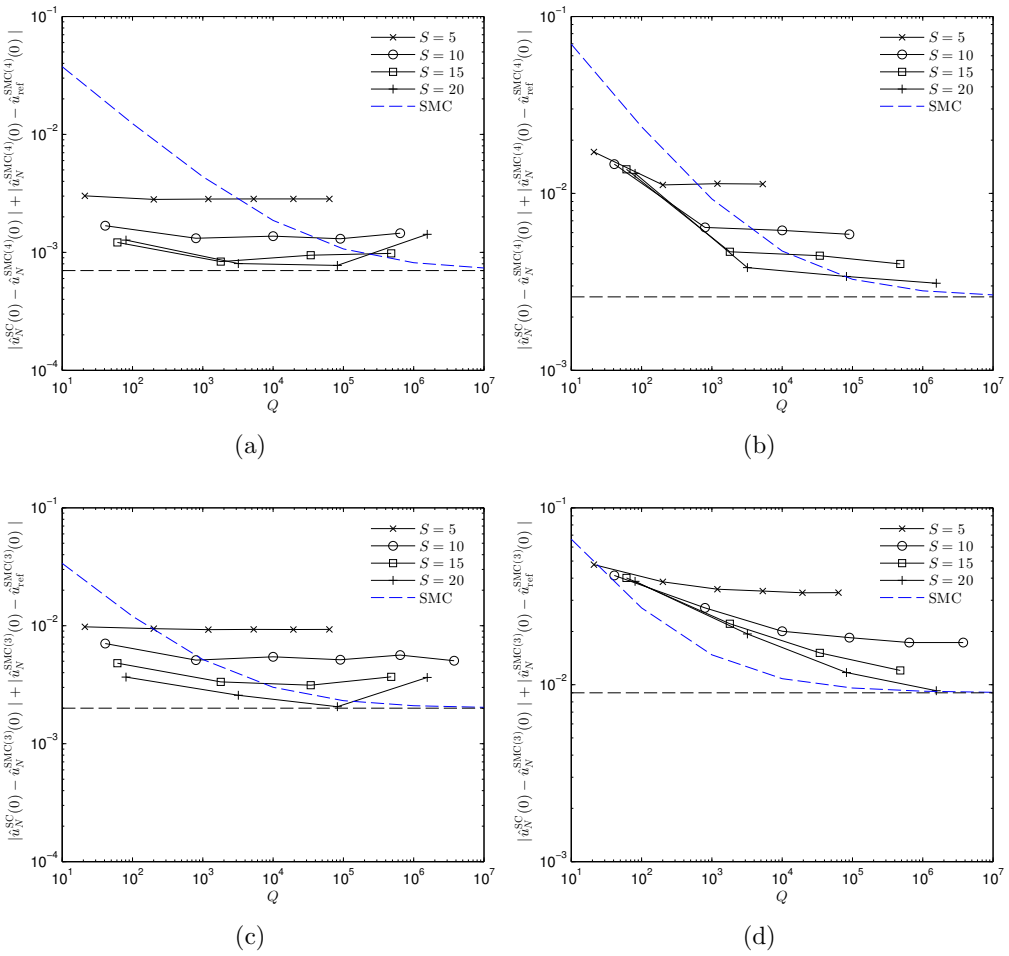

Figure 3. *Dependence of the total error, $\epsilon^{(n)} = \epsilon_{KL}^{(n)} + \epsilon_{est}^{(n)}$, on the number of SC nodes, Q. Here $\epsilon_{KL}^{(n)} = |\hat{u}_N^{\mathrm{MC}(n)}(0) - \hat{u}_{\mathrm{ref}}^{\mathrm{MC}(n)}(0)|$ and $\epsilon_{est}^{(n)} = |\hat{u}_N^{\mathrm{SC}}(0) - \hat{u}_N^{\mathrm{MC}(n)}(0)|$. Subfigures* (a) *through* (d) *correspond to the test cases in Table* 1. *The dashed black lines indicate the bias error $\epsilon_{\mathrm{KL}}^{(n)}$ for $N = 40$. The dashed blue line indicates the total error (RMSE + bias) of the simple MC simulations (SMC) for $N = 40$.*

The lack of rigorous estimates of the convergence rate of quadrature rules with Gaussian weights makes it difficult to determine a reason for the observed deterioration of the performance of the SC method vis-à-vis MCS as variances of the input parameters increase. Nevertheless, the following observations appear to be valid. First, asymptotic estimates, such as those described in section 3.1, obscure the effect of the input parameters' variance on the convergence of high-dimensional sparse grid quadratures. For such quadrature techniques, ϵ_{est} is affected by the magnitude of the derivatives of the response surface. Second, the use of K-L expansions implies that if the variance of a random field increases by a factor of a, then the

magnitude of an rth partial derivative of the response surface increases by a factor of $a^{r/2}$. Therefore, ϵ_{est} increases with variance.

It is worth emphasizing that these observations apply for any interpolatory UQ technique that employs a finite-dimensional noise approximation such as truncated K-L expansions. Therefore, we expect a similar behavior to occur for other versions of the SC method and for the stochastic Galerkin method.

4.2. Parabolic problems. Consider a parabolic problem, which is given by a one-dimensional version of (2.1),

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}\left[K \frac{\partial}{\partial z}(u - z)\right], \qquad 0 < z < 1. \tag{4.5}$$

It is supplemented with the van Genuchten–Mualem constitutive relations (see [31] and the references therein),

$$K = K_{\text{s}}(z,\omega) \begin{cases} S_{\text{e}}^{1/2}[1 - (1 - S_{\text{e}}^{1/m})^m]^2, & u \le 0, \\ 1, & u > 0, \end{cases} \tag{4.6}$$

$$S_{\text{e}} \equiv \frac{\theta - \theta_{\text{r}}}{\theta_s - \theta_{\text{r}}} = \begin{cases} (1 + |\alpha(z,\omega)u|^n)^{-m}, & u \le 0, \\ 1, & u > 0, \end{cases} \tag{4.7}$$

and is subject to initial and boundary conditions

$$u(z,0) = u_{\text{b}}, \quad u(0,t) = u_{\text{t}}, \quad u(1,t) = u_{\text{b}}. \tag{4.8}$$

Values of the deterministic model parameters m, n, θ_{r}, θ_{s} and the statistical properties of the Gaussian coefficients $Y = \ln K_{\text{s}}(z,\omega)$ and $\beta = \ln \alpha(z,\omega)$ are summarized in Table 3.

Table 3
Model parameters for parabolic problem (4.5)–(4.8).

Test	σ_Y^2	σ_β^2	λ_Y	λ_β	K_{g}	α_{g}	θ_{s}	θ_{r}	n	m	u_{t}	u_{b}
(a)	0.001	0.001	1.0	1.0	0.5532	2.01	0.368	0.102	2.0	0.5	−1.25	−16.67
(b)	0.01	0.01	1.0	1.0	0.5532	2.01	0.368	0.102	2.0	0.5	−1.25	−16.67

In each deterministic solve, the pressure profile $u(x,t)$ is approximated by using a cell-centered finite volume scheme with M degrees of freedom to discretize (4.5) in space. The discrete state vector is $\mathbf{u}(t) = [u(z_1,t), \ldots u(z_M,t)]^\top$. The finite volume centroids are located along the z direction at locations $z_i = (i - 1/2)\Delta z$ with $i \in [1, M]$ and $\Delta z = 1/M = 0.01$. The time-stepping of the state vector $\mathbf{u}(t)$ is performed using the three-stage RADAU IIA implicit Runge–Kutta scheme [12]. For the following discussion, statistics of the state vector $\mathbf{u}(t)$ are understood to be computed elementwise.

Figure 4 exhibits temporal snapshots of the reference MC estimator $\hat{u}_{\text{ref}}^{\text{MC}(3)}(z,t)$, and its corresponding sample variance $\hat{\sigma}_u^2(z,t)$, at times $t = 6.0$, 12.0, and 18.0. The reference estimator is computed using $S = 800$ terms in the K-L expansion of each random field and

is accurate to $n = 3$ significant digits with 95% confidence. Each realization (deterministic solve) of (4.5)–(4.8) develops a sharp gradient (a moving front) in the physical coordinate z, across which the state variable u drops from u_{t} to the background value u_{b}. As the dynamics of the sharp front is influenced by values of the random coefficients $K_{\mathrm{s}}(z,\omega)$ and $\alpha(z,\omega)$, it is reasonable to surmise that this behavior in the physical space z translates into steep gradients of the stochastic response surface in the outcome space $\boldsymbol{\xi}_N$. Indeed, even small fluctuations of the model parameters K_{s} and α (e.g., $\sigma_Y^2 = \sigma_\beta^2 = 0.01$ in Test (b)) result in large uncertainty about the spatial position of the moving front (see, e.g., the corresponding sample variance of $u(z,t)$ in the second row of Figure 4).

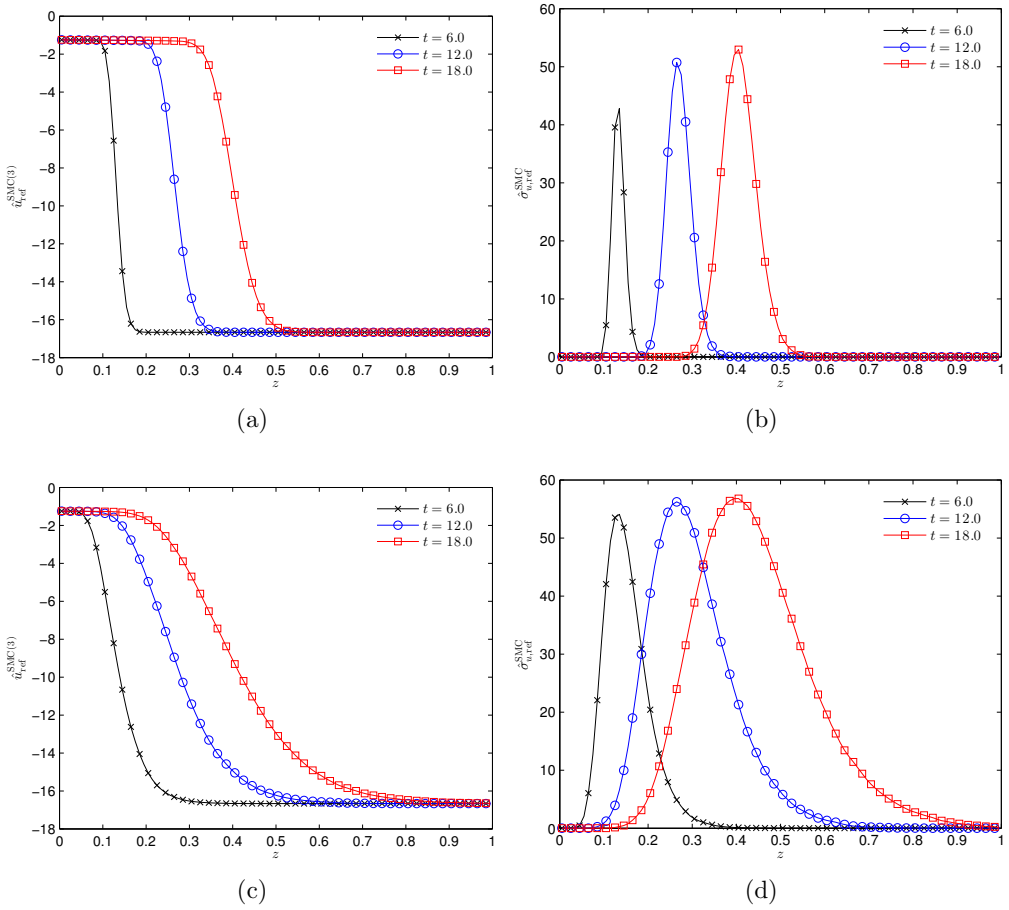

Figure 4. *Temporal snapshots of the reference MC estimator* $\hat{u}_{\mathrm{ref}}^{\mathrm{MC}(3)}(z,t)$ *and the corresponding sample variance* $\hat{\sigma}_{u,\mathrm{ref}}^{2,\mathrm{MC}}(z,t)$*, at times* $t = 6.0$*,* 12.0*, and* 18.0*. The first and second rows correspond to the test cases* (a) *and* (b) *in Table* 3*, respectively.*

We use element-wise differences $|\hat{u}_N^{\mathrm{MC}(3)}(z_i,t) - \hat{u}_{\mathrm{ref}}^{\mathrm{MC}(3)}(z_i,t)|$ and $|\hat{u}_N^{\mathrm{MC}(3)}(z_i,t) - \hat{u}_N^{\mathrm{SC}}(z_i,t)|$ as surrogates the element-wise bias, $\epsilon_{\mathrm{KL}}(z_i,t)$, and estimation, $\epsilon_{\mathrm{est}}(z_i,t)$, errors, respectively.

The maximum norms of these errors, $\epsilon_{\mathrm{KL}}^{\max}(t) = \|\epsilon_{\mathrm{KL}}(z_i,t)\|_\infty$ and $\epsilon_{\mathrm{est}}^{\max}(t) = \|\epsilon_{\mathrm{est}}(z_i,t)\|_\infty$ at time $t = 18.0$ are shown in Table 4 and Figure 5, respectively, for $N = 2$, 4, 10, and 20. Comparison of these results suggests that the estimation error, rather than the bias error, is the limiting factor in achieving a given global accuracy. In other words, construction of an adequate quadrature of the stochastic response surface, which exhibits large gradients in the outcome space, is the central challenge for this nonlinear hyperbolic sBVP. Figure 5 shows that quadrature rules of a large degree of polynomial exactness are required to accurately approximate the expectation $\langle u(z,t)\rangle$.

Table 4

Maximum norm of the bias error $\epsilon_{KL}^{max}(t = 18.0)$. The error estimators are computed for $n = 3$ significant digits with 95% confidence.

		$\epsilon_{\mathrm{KL}}^{\max}(t = 18.0)$	
S	$N = 2S$	Test (a)	Test (b)
1	2	0.51	0.65
2	4	0.09	0.16
5	10	0.06	0.09
10	20	0.02	0.04
15	30	0.03	0.04

Similar to the conclusion drawn from the elliptic case (Figure 2), construction of an accurate quadrature becomes more challenging as variance of the random model parameters increases. Again, this is because the absolute value of the mixed derivatives of the stochastic response surface increases with this variance. If the error requirements are such that $S > 2$ K-L expansion terms are required, the use of an accurate quadrature in the SC method renders the latter less efficient than MCS.

A possible approach to overcoming this challenge is to use adaptive rules for high-dimensional quadratures of response surfaces with large gradients [1, 2, 15, 23]. Such constructions might become problematic when one is interested in multiple QoI, e.g., the pressure at various locations in physical space studied in this section. That is because the region of outcome space where large gradients occur, that is, the region over which adaptivity is required, also varies in space and time. For the application studied in this section, each component of the stochastic response vector $\mathbf{u}(t,\boldsymbol{\xi}_N)$ exhibits large gradients over different regions of the outcome space, and these regions also vary with time t.

5. Conclusions. We evaluated the performance of global SC methods for solving nonlinear parabolic and elliptic problems (e.g., transient and steady nonlinear diffusion) with random coefficients. The robustness of these and other strategies based on a spectral decomposition of stochastic state variables depends on the regularity of the system's response in outcome space. The latter is affected by statistical properties of the input random fields. These include variances of the input parameters, whose effect on the computational efficiency of this class of uncertainty quantification techniques has remained unexplored. Our analysis shows that if random coefficients have low variances and large correlation lengths, SC strategies outperform

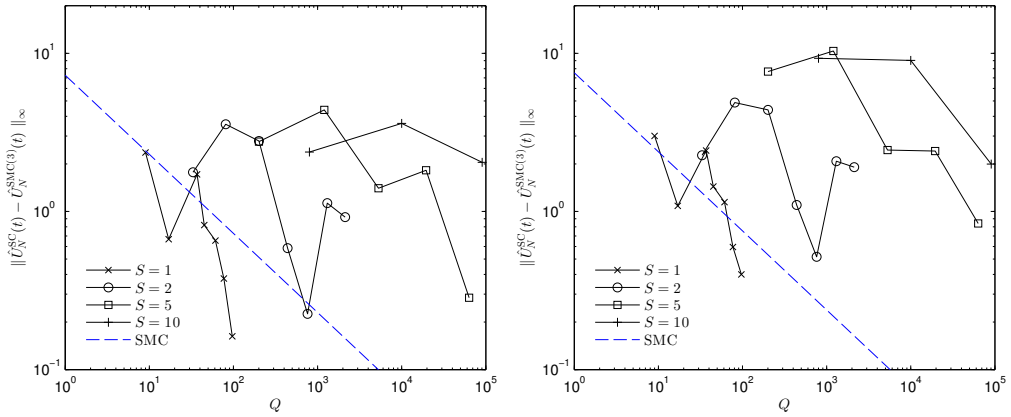

Figure 5. *Dependence of the estimation error,* $\epsilon_{est}^{max}(t) = \|\hat{u}_N^{\mathrm{MC}(3)}(z_i,t) - \hat{u}_N^{\mathrm{SC}}(z_i,t)\|_\infty$ *on the number of* SC *nodes, for the cases of Table 3, and* $t = 18.0$. *Dashed black lines indicate half-width corresponding to number of accurate significant digits* n. *Dashed blue line indicates the decay of the maximum norm of the RMSE of simple MC simulations (SMC).*

MCS. As variance increases, the regularity of the stochastic response decreases, which requires higher-order quadrature rules to accurately approximate the moments of interest and increases the overall computational cost above that of MCS.

The lack of rigorous estimates of the convergence rate of quadrature rules with Gaussian weights makes it difficult to determine a reason for the observed deterioration of the performance of the SC method vis-à-vis MCS as variances of the input parameters increase. Nevertheless, the following observations appear to be valid. First, asymptotic estimates of an estimation error ϵ_{est} obscure the effect of the input parameters' variance on the convergence of high-dimensional sparse grid quadratures. For such quadrature techniques, ϵ_{est} is affected by the magnitude of the derivatives of the response surface. Second, the use of K-L expansions implies that if variance of a random field increases by a factor of a, then the magnitude of an rth partial derivative of the response surface increases by a factor of $a^{r/2}$. Therefore, ϵ_{est} increases with variance.

A possible approach to overcoming this challenge is to use adaptive rules for high-dimensional quadratures of response surfaces with large gradients. Such constructions might become problematic when a QoI exhibits spatio-temporal variability. That is because a stochastic response surface might possess large gradients in different regions of the outcome space, and these regions might vary in space and time.

It is worth emphasizing that these observations apply for any interpolatory UQ technique that employs a finite-dimensional noise approximation such as truncated K-L expansions. Therefore, we expect similar behavior to occur for other versions of the SC method and for the stochastic Galerkin method.

REFERENCES

[1] N. Agarwal and N. R. Aluru, *A domain adaptive stochastic collocation approach for analysis of MEMS under uncertainties*, J. Comput. Phys., 228 (2009), pp. 7662–7688.

[2] R. Archibald, A. Gelb, R. Saxena, and D. Xiu, *Discontinuity detection in multivariate space for stochastic simulations*, J. Comput. Phys., 228 (2009), pp. 2676–2689.

[3] I. Babuška, R. Tempone, and G. E. Zouraris, *Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation*, Comput. Methods Appl. Mech. Engrg., 194 (2005), pp. 1251–1294.

[4] J. Charrier, *Strong and weak error estimates for elliptic partial differential equations with random coefficients*, SIAM J. Numer. Anal., 50 (2012), pp. 216–246.

[5] H. Cho, D. Venturi, and G. E. Karniadakis, *Karhunen-Loève expansion for multi-correlated stochastic processes*, Probab. Engrg. Mech., 34 (2013), pp. 157–167.

[6] J. Foo, X. Wan, and G. E. Karniadakis, *The multi-element probabilistic collocation method (ME-PCM): Error analysis and applications*, J. Comput. Phys., 227 (2008), pp. 9572–9595.

[7] P. Frauenfelder, C. Schwab, and R. A. Todor, *Finite elements for elliptic problems with stochastic coefficients*, Comput. Methods Appl. Mech. Engrg., 194 (2005), pp. 205 – 228.

[8] B. Ganapathysubramanian and N. Zabaras, *Sparse grid collocation schemes for stochastic natural convection problems*, J. Comput. Phys., 225 (2007), pp. 652–685.

[9] A. Genz and B. D. Keister, *Fully symmetric interpolatory rules for multiple integrals over infinite regions with Gaussian weight*, J. Comput. Appl. Math., 71 (1996), pp. 299–309.

[10] T. Gerstner and M. Griebel, *Numerical integration using sparse grids*, Numer. Algorithms, 18 (1998), pp. 209–232.

[11] R. G. Ghanem and P. D. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, New York, 1991.

[12] E. Hairer and G. Wanner, *Stiff differential equations solved by radau methods*, J. Comput. Appl. Math., 111 (1999), pp. 93–111.

[13] A. C. Hindmarsh, P. N. Brown, K. E. Grant, S. L. Lee, R. Serban, D. E. Shumaker, and C. S. Woodward, *SUNDIALS: Suite of nonlinear and differential/algebraic equation solvers*, ACM Trans. Math. Software, 31 (2005), pp. 363–396.

[14] A. Hinrichs and E. Novak, *Cubature formulas for symmetric measures in higher dimensions with few points*, Math. Comp., 76 (2007), pp. 1357–1372.

[15] J. D. Jakeman, R. Archibald, and D. Xiu, *Characterization of discontinuities in high-dimensional stochastic problems on adaptive sparse grids*, J. Comput. Phys., 230 (2011), pp. 3977–3997.

[16] K. D. Jarman and T. F. Russell, *Eulerian moment equations for 2-D stochastic immiscible flow*, Multiscale Model. Simul., 1 (2003), pp. 598–608.

[17] D. K. Kahaner and G. Monegato, *Nonexistence of extended Gauss-Laguerre and Gauss-Hermite quadrature rules with positive weights*, Z. Angew. Math. Mech. (ZAMM), 29 (1978), pp. 983–986.

[18] G. Kuperberg, *Numerical cubature using error-correcting codes*, SIAM J. Numer. Anal., 44 (2006), pp. 897–907.

[19] O. P. Le Maître, O. M. Knio, H. N. Najm, and R. G. Ghanem, *Uncertainty propagation using Wiener-Haar expansions*, J. Comput. Phys., 197 (2004), pp. 28–57.

[20] O. P. Le Maître, H. N. Najm, R. G. Ghanem, and O. M. Knio, *Multi-resolution analysis of Wiener-type uncertainty propagation schemes*, J. Comput. Phys., 197 (2004), pp. 502–531.

[21] P. C. Lichtner and D. M. Tartakovsky, *Upscaled effective rate constant for heterogeneous reactions*, Stoch. Environ. Res. Risk Assess., 17 (2003), pp. 419–429.

[22] J. Lu and D. Darmofal, *Higher-dimensional integration with Gaussian weight for applications in probabilistic design*, SIAM J. Sci. Comput., 26 (2004), pp. 613–624.

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

[24] W. Merz and P. Rybka, *Strong solutions to the Richards equation in the unsaturated zone*, J. Math. Anal. Appl., 371 (2010), pp. 741–749.

[25] E. Novak and K. Ritter, *High dimensional integration of smooth functions over cubes*, Numer. Math., 75 (1996), pp. 79–97.

[26] E. Novak and K. Ritter, *Simple cubature formulas with high polynomial exactness*, Constr. Approx., 15 (1999), pp. 499–522.
[27] B. Øksendal, *Stochastic differential equations*, in Stochastic Differential Equations, Universitext, Springer-Verlag, Berlin, 2003, pp. 65–84.
[28] S. A. Smolyak, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Dokl. Akad. Nauk SSSR, 4 (1963), pp. 240–243.
[29] D. M. Tartakovsky, M. Dentz, and P. C. Lichtner, *Probability density functions for advective-reactive transport in porous media with uncertain reaction rates*, Water Resources Res., 45 (2009), W07414.
[30] D. M. Tartakovsky, A. Guadagnini, and M. Riva, *Stochastic averaging of nonlinear flows in heterogeneous porous media*, J. Fluid Mech., 492 (2003), pp. 47–62.
[31] D. M. Tartakovsky, Z. Lu, A. Guadagnini, and A. M. Tartakovsky, *Unsaturated flow in heterogeneous soils with spatially distributed uncertain hydraulic parameters*, J. Hydrol., 275 (2003), pp. 182–193.
[32] J. Tryoen, O. P. Le Maître, M. Ndjinga, and A. Ern, *Intrusive Galerkin methods with upwinding for uncertain nonlinear hyperbolic systems*, J. Comput. Phys., 229 (2010), pp. 6485–6511.
[33] N. Victoir, *Asymmetric cubature formulae with few points in high dimension for symmetric measures*, SIAM J. Numer. Anal., 42 (2004), pp. 209–227.
[34] X. Wan and G. Karniadakis, *Multi-element generalized polynomial chaos for arbitrary probability measures*, SIAM J. Sci. Comput., 28 (2006), pp. 901–928.
[35] G. W. Wasilkowski and H. Wozniakowski, *Explicit cost bounds of algorithms for multivariate tensor product problems*, J. Complexity, 11 (1995), pp. 1–56.
[36] C. L. Winter, D. M. Tartakovsky, and A. Guadagnini, *Moment equations for flow in highly heterogeneous porous media*, Surv. Geophys., 24 (2003), pp. 81–106.
[37] D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University Press, Princeton, NJ, 2010.
[38] M. Ye, S. P. Neuman, A. Guadagnini, and D. M. Tartakovsky, *Nonlocal and localized analyses of conditional mean transient flow in bounded, randomly heterogeneous porous media*, Water Resources Res., 40 (2004), W05104.