Probabilistic Engineering Mechanics 41 (2015) 60–72

Contents lists available at ScienceDirect

Probabilistic Engineering Mechanics

journal homepage: www.elsevier.com/locate/probengmech

On the efficacy of stochastic collocation, stochastic Galerkin, and stochastic reduced order models for solving stochastic problems

R.V. Field Jr.[a,1,*], M. Grigoriu [b], J.M. Emery [c]

[a] Sandia National Laboratories, Albuquerque, NM 87185-0346, USA
[b] Cornell University, Ithaca, NY 14853, USA
[c] Sandia National Laboratories, Albuquerque, NM 87185-0840, USA

ARTICLE INFO

Article history:
Received 11 December 2014
Received in revised form
14 May 2015
Accepted 15 May 2015
Available online 19 May 2015

Keywords:
Approximation theory
Monte Carlo simulation
Random variables and fields
Stochastic differential equations
Uncertainty propagation

ABSTRACT

The stochastic collocation (SC) and stochastic Galerkin (SG) methods are two well-established and successful approaches for solving general stochastic problems. A recently developed method based on stochastic reduced order models (SROMs) can also be used. Herein we provide a comparison of the three methods for some numerical examples; our evaluation only holds for the examples considered in the paper. The purpose of the comparisons is not to criticize the SC or SG methods, which have proven very useful for a broad range of applications, nor is it to provide overall ratings of these methods as compared to the SROM method. Rather, our objectives are to present the SROM method as an alternative approach to solving stochastic problems and provide information on the computational effort required by the implementation of each method, while simultaneously assessing their performance for a collection of specific problems.

1. Introduction

Consider Fig. 1, which illustrates a general framework for solving stochastic problems. Input Y is mapped to output U by mapping f which, in general, is nonlinear and time-dependent. The mapping f can be defined by a collection of differential, algebraic, and/or integral equations; it is often complex and must be solved numerically. Further, for most problems of practical interest, there is some uncertainty associated with our knowledge of Y and/or f. If we represent this uncertainty with random functions, then output U is also a random function. This class of problem is referred to as a stochastic problem [1]. The solution to the stochastic problem and, hence, our objective in this study, is to estimate the probability law of output U in an accurate, but computationally efficient, manner.

A Gaussian random function is one type of random function that is often used to represent uncertainty; Gaussian functions have nice mathematical properties, making them a desirable choice for modeling uncertainty in complex problems. However, Gaussian models are not always appropriate for representing uncertainty in "real-world" systems; examples include uncertainty that is known to be non-negative, bounded, and/or not symmetric about its mean. We therefore focus our study on the solution of stochastic problems involving non-Gaussian random functions.

Consider Fig. 2, which defines one class of non-Gaussian random functions Y as a nonlinear mapping g of W, a finite collection of independent and identically distributed (iid) standard Gaussian random variables. A broad class of non-Gaussian random functions can be represented in this way, that is, as $Y = g(W)$. For example, if W is a standard Gaussian random variable with CDF Φ, then $\exp(W)$ and $\Phi(W)$ are random variables that have a lognormal and uniform distribution, respectively. In addition, any Gaussian random function can be approximated by a collection of iid standard Gaussian random variables by its spectral representation [2, Section 3.9.4]; suitable nonlinear mappings of these approximations produce non-Gaussian random functions.

Our objective herein is to compare and contrast three methods that are commonly used to solve stochastic problems involving non-Gaussian random functions, that is, problems of the type described by Figs. 1 and 2. These methods include (1) stochastic collocation (SC) [3–5]; (2) stochastic Galerkin (SG) [5–7]; and (3) stochastic reduced order models (SROMs) [8,9]. The comparisons will be based on both accuracy and computational cost for a variety of examples.

The SC and SG methods are well-established approaches to solving stochastic problems. Our objective is not to criticize these methods, which have proven very useful for a broad range of

* Corresponding author. Tel.: +1 505 284 4060; fax: +1 505 844 6620.
E-mail addresses: rvfield@sandia.gov (R.V. Field Jr.), mdg12@cornell.edu (M. Grigoriu), jmemery@sandia.gov (J.M. Emery).

[1] Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000. SAND2015-20740 J.

http://dx.doi.org/10.1016/j.probengmech.2015.05.002
0266-8920/

Fig. 1. Framework for stochastic problems.

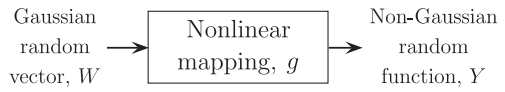

Fig. 2. Non-Gaussian random functions as nonlinear mappings of Gaussian random vectors.

applications, nor is it to provide overall ratings of these methods as compared to the SROM method. Rather, our objectives are to: (1) provide information on the computational effort required by the implementation of two well-established methods, *i.e.*, SC and SG, and of a new method based on SROM and (2) assess the performance of all three methods in some numerical examples. Our evaluation only holds for the examples considered in the paper.

The paper is organized as follows. In Section 2, we define the generalities of the three methods for analysis considered, *i.e.*, stochastic collocation, stochastic Galerkin, and stochastic reduced order models. Applications of the methods are presented in Section 3. For our study, we will consider several different mappings f and g defined by Figs. 1 and 2 to illustrate a variety of results. For the first set of examples, described in Section 3.1, we consider the case where f is a simple algebraic equation mapping input Y to output U; three different mappings g are considered, resulting in three different non-Gaussian inputs Y. Due to its simplicity, these examples prove to be very instructive since many of the required calculations can be completed in closed-form. A second, more complex set of examples are considered in Section 3.1, where mapping f is a stochastic differential equation depending on a random field.

2. Methods for analysis

We consider stochastic problems of the form

$$U = f(Y)$$
$$Y = g(W), \quad W \in \mathbb{R}^d, \tag{1}$$

where random function Y is mapped to random function U by a measurable and deterministic function f. Further, $Y = g(W)$ is defined by a different measurable and deterministic mapping of W, a collection of standard independent Gaussian random variables. Hence U depends on W, that is, $U = U(W) = f \circ g(W)$. We note that herein Y and U are assumed to be scalar-valued functions. The discussion that follows can be generalized to the case of vector-valued mappings $f: \mathbb{R}^p \to \mathbb{R}^q$ and $g: \mathbb{R}^d \to \mathbb{R}^p$, where $p, q \geq 1$ are integers; we will assume $p = q = 1$ to simplify the discussion.

In the context of solving problems of practical interest in science and engineering, we interpret Y and U to be input to and output from a model for a complex system; the input/output mapping is represented by f and oftentimes is a collection of differential, algebraic, and/or integral equations that is solved using, for example, the finite element or finite difference methods. Mapping f can therefore require significant computational effort to solve. In contrast g is generally quite simple and can be evaluated with minimal effort. Hence, when comparing the computational cost associated with each method, we simply report the number of times we solve or evaluate f, herein denoted by n.

We next provide an overview of the stochastic collocation, stochastic Galerkin, and stochastic reduced order modeling approaches to solving stochastic problems of the type described by Eq. (1). While the details and properties of each approach are quite different, the fundamental concept is the same: each approach develops an approximation for U, herein denoted by $\tilde{U}$, that requires less computational effort than solving Eq. (1) exactly. Traditional random sampling of the approximation $\tilde{U}$ can then be used to provide estimates for the probability law of U, as well as for functionals of U, that are of interest.

2.1. Stochastic collocation (SC)

The stochastic collocation approach seeks to approximate output $U(W)$, $W = (W_1, \ldots, W_d)^T$, by first computing U at a finite number of sample points, then interpolating over the entire domain of W. Denote by $w_1, \ldots, w_d$ the coordinates of $w \in \mathbb{R}^d$, and let $w_{k,i}$, $i = 0, \ldots, m_k$, be unique collocation points along coordinate k of w, where $m_k \geq 1$ is an integer. Assuming a rectangular grid in $\mathbb{R}^d$, the stochastic collocation approximation for $U(W)$ is given by

$$\tilde{U}_{sc}(W) = \sum_{k_1=0}^{m_1} \cdots \sum_{k_d=0}^{m_d} U(w_{1,k_1}, \ldots, w_{d,k_d}) \prod_{i=1}^{d} \psi_{k_i}(W_i) \tag{2}$$

where

$$\psi_k(\xi) = \prod_{\substack{j=0 \\ j \neq k}}^{m_k} \frac{\xi - w_{k,j}}{w_{k,k} - w_{k,j}} \tag{3}$$

defines Lagrange interpolation in $\mathbb{R}^d$. For calculations, we follow three steps:

1. Select collocation points $w_{k,0}, \ldots, w_{k,m_k}$, $k = 1, \ldots, d$;
2. Calculate the solutions $\{U(w_{1,k_1}, \ldots, w_{d,k_d})\} = \{f \circ g(w_{1,k_1}, \ldots, w_{d,k_d})\}$ at these points; and
3. Construct the interpolating polynomials $\{\psi_k(\xi)\}$ defined by Eq. (3).

By Steps 1 and 2, it is evident that the computational cost of the stochastic collocation approach, *i.e.*, the number of times we evaluate f defined by Eq. (1), is quantified by

$$n_{sc} = \prod_{k=1}^{d} (m_k + 1). \tag{4}$$

By Eqs. (2) and (4), a total of n_{sc} collocation points are selected. To study the performance of the collocation method as a function of n_{sc}, we first compute $\tilde{U}_{sc}$ based on all points, then proceed to remove points from the original collection and update the value for $\tilde{U}_{sc}$. By this procedure, each solution based on $n < n_{sc}$ collocation points is a subset of the original solution.

Herein we apply a standard scheme to choose collocation points on a rectangular grid in $\mathbb{R}^d$. Many variations of the SC method have been introduced to improve its efficiency including, but not limited to, sparse grid collocation, anisotropic sparse grid collocation [10] and a multi-element probabilistic collocation method [11]. It is quite possible that one or more of these variations can produce better performance of the stochastic collocation method for some or all of the examples we consider in Section 3, but an exhaustive analysis of every collocation technique is beyond the scope of this paper.

2.2. Stochastic Galerkin (SG)

Let U be a random function with finite variance. The stochastic Galerkin approach approximates U by a sum of deterministic functions multiplied by a collection of orthogonal polynomials evaluated over a space of random variables.

Let $Q \in \mathbb{R}^d$ be a vector of real-valued iid random variables, each

with finite mean and variance, and let $\{h_k : \mathbb{R}^d \to \mathbb{R}\}$ be a collection of orthogonal polynomials. The polynomials $\{h_k\}$ are orthogonal with respect to the probability measure of Q so that

$$E[h_k(Q)] = \delta_{1k}$$
$$E[h_j(Q)h_k(Q)] = \delta_{jk}E[h_k(Q)^2], \tag{5}$$

where δ_{jk} denotes the Kronecker delta function and $E[A]$ denotes the expected value of random variable A. The stochastic Galerkin approximation for U is given by

$$\tilde{U}_{sg}(Q) = \sum_{k=0}^{n_u} b_k h_k(Q), \tag{6}$$

where the $\{b_k\}$ are deterministic functions to be determined and $n_u + 1$ denotes the number of terms in the approximation. The method can be implemented in one of two ways, as will be discussed in Sections 2.2.1 and 2.2.2.

It is common to take $\{h_k\}$ defined by Eq. (6) to be Hermite polynomials with $Q=W$ a Gaussian random vector with iid coordinates having zero mean and unit variance, *i.e.*, $W_k \sim \mathcal{N}(0, 1)$. Alternative choices include Legendre polynomials with Q having iid coordinates $Q_k \sim \mathcal{U}(-1, 1)$ distributed uniformly over $[-1, 1]$, and Laguerre polynomials with Q having iid coordinates, each with an exponential distribution with unit mean and unit variance [5,12]. In either case, the approximation defined by Eq. (6) becomes more accurate in the sense that $E[(U - \tilde{U}_{sg})^2] \to 0$ as the number of terms in the sum, n_u, increases.

2.2.1. Non-intrusive implementation

As mentioned, the mapping f defined by Eq. (1) is often complex and must be solved by numerical approximation; finite element and finite difference solutions are common approaches. In this case, it is advantageous to utilize the stochastic Galerkin method in a way that requires no modification to the existing finite element or finite difference model, that is, no modification to f. We refer to this scenario as the non-intrusive implementation of the stochastic Galerkin approach.

To solve for the unknown $\{b_k\}$ defined by Eq. (6), we exploit the orthogonality properties of the basis, *i.e.*, Eq. (5). Multiplying both sides of Eq. (6) by $h_l(Q)$ and taking expectations, we can show that

$$b_k = \frac{E[U(Q)h_k(Q)]}{E[h_k(Q)^2]}, \quad k = 0, \ldots, n_u. \tag{7}$$

Approximations for Eq. (7) are needed in general; one such approximation is by simple Monte Carlo integration. Let q_j, $j = 1, \ldots, n_{sg}$, denote iid random samples of random vector Q. Then

$$b_k \approx \frac{1}{n_{sg}E[h_k(Q)^2]} \sum_{j=1}^{n_{sg}} U(q_j)h_k(q_j), \tag{8}$$

where $U(q_j)$ is computed by evaluating Eq. (1) at $W = q_j$, and we note that values for $E[h_k(Q)^2]$ are available in closed form [6]. To study the performance of the stochastic Galerkin method as a function of n_{sg}, we first compute the $\{b_k\}$ based on all samples $\{q_1, \ldots, q_{n_{sg}}\}$, then proceed to remove points at random from the original collection and update the value for the $\{b_k\}$. By this procedure, each solution based on $n < n_{sg}$ samples of Q is a subset of the original solution.

Herein, we utilize standard Monte Carlo integration, that is, Eq. (8), to compute the coefficients of the stochastic Galerkin approximation defined by Eq. (6). There are alternative approaches that can be very useful, including numerical quadrature, sparse polynomial bases [13], wavelet expansions [14], and an adaptive, multi-element approach [15]. It is quite possible that one or more of these variations can produce better performance of the stochastic Galerkin method for some or all of the examples we consider in Section 3, but an exhaustive analysis of every variation of the Galerkin technique is beyond the scope of this paper.

2.2.2. Intrusive implementation

An alternative approach is to express input Y via an expansion similar to the expansion for U, and substitute both expansions into the expression $U = f(Y)$ to solve for the unknown $\{b_k\}$. This approach is referred to as an intrusive implementation of the stochastic Galerkin method because any existing finite element or finite difference formulation for mapping f must be modified to solve for U in this way.

To illustrate, consider the case where U and Y are random variables and $f(Y) = 1/Y$. An approximation for $Y = g(Q)$ in this case is

$$\tilde{Y}_{sg}(Q) = \sum_{k=0}^{n_y} a_k h_k(Q), \tag{9}$$

where $\{h_k : \mathbb{R} \to \mathbb{R}\}$ are univariate orthogonal polynomials, Q is a random variable with zero mean and finite variance, and $\{a_k\}$ are known deterministic constants. Substituting these expansions into $U\,Y = 1$, we have

$$\left(\sum_{l=0}^{n_u} b_l h_l(Q)\right)\left(\sum_{k=0}^{n_y} a_k h_k(Q)\right) = 1. \tag{10}$$

Multiplying both sides by $h_r(Q)$ and taking expectations, we obtain the following set of equations:

$$\sum_{k=0}^{n_y}\sum_{l=0}^{n_u} a_k b_l E[h_k(Q)h_l(Q)h_r(Q)] = E[h_r(Q)], \quad r = 0, 1, \ldots, n_u. \tag{11}$$

Assuming Hermite polynomials and $Q = W \sim \mathcal{N}(0, 1)$, Eq. (11) simplifies to

$$\sum_{k=0}^{\min(n_y,n_u)} a_k b_k k! = 1 \quad \text{and}$$
$$\sum_{k=0}^{n_y}\sum_{l=0}^{n_u} a_k b_l E[h_k(W)h_l(W)h_r(W)] = 0, \quad r = 1, \ldots, n_u, \tag{12}$$

where [16]

$$E[h_k(W)h_l(W)h_r(W)] = \begin{cases} \dfrac{k!l!r!}{(s-k)!(s-l)!(s-r)!} & k+l+r = 2s, \quad k, l, r \le s \\ 0 & \text{else} \end{cases} \tag{13}$$

Likewise, assuming Legendre polynomials and $Q \sim \mathcal{U}(-1, 1)$, Eq. (11) simplifies to

$$\sum_{k=0}^{\min(n_y,n_u)} \frac{a_k b_k}{2k+1} = 1 \quad \text{and}$$
$$\sum_{k=0}^{n_y}\sum_{l=0}^{n_u} a_k b_l E[h_k(Q)h_l(Q)h_r(Q)] = 0, \quad r = 1, \ldots, n_u, \tag{14}$$

where [17]

$$E[h_k(Q)h_l(Q)h_r(Q)] = \begin{cases} \dfrac{\eta(s-k)\eta(s-l)\eta(s-r)}{(2s+1)\eta(s)} & k+l+r = 2s, \quad k, l, r \le s \\ 0 & \text{else} \end{cases} \tag{15}$$

and $\eta(s) = (2s)!/(2^s s!s!)$.

For the non-intrusive implementation, the cost associated with the stochastic Galerkin is simply n_{sg} by Eq. (8). In addition, the

accuracy of this implementation depends on n_{sg}, as well as n_u defined by Eq. (6). The accuracy of the intrusive approach depends only on n_u because no sampling is needed to compute the coefficients. However, because the method is intrusive and requires modification to the mapping f, it is difficult to compute a cost associated with this implementation in a manner that is consistent with the other methods considered in this study. We therefore limit any further discussion to the non-intrusive implementation of the stochastic Galerkin solution method.

2.3. Stochastic reduced order models (SROM)

The stochastic reduced order model is another approach for approximating output $U(W)$ and involves two steps. First W, the vector of iid standard Gaussian random variables, is approximated by a simple random vector taking values $w_1, \ldots, w_{q_y}$ with probabilities $p_1, \ldots, p_{q_y}$, referred to as a stochastic reduced order model for W. Second, the input/output mapping f is approximated by linear and/or higher order local approximations. We define these two steps in more detail in the following sections. A detailed discussion on the errors introduced by the SROM approach is provided in [9].

2.3.1. Reduced order model for W

Let W be a vector of iid Gaussian random variables, where each coordinate has zero mean, unit variance, and marginal CDF Φ. This implies the marginal moments of each coordinate are $\mu_W(r) = E[W_1^r] = 1{\cdot}3{\cdots}(r-1)$ for r even and zero otherwise, and $E[W_k\ W_l] = \delta_{kl}$. Let $\{\Gamma_k,\ k = 1, \ldots, q_y\}$ denote a partition of $\mathbb{R}^d$, that is, the Γ_k are pairwise disjoint and $\cup_k \Gamma_k = \mathbb{R}^d$.

The SROM approximation for W is a simple random vector defined by

$$\tilde{W}_{\text{srom}} = \{(w_k, p_k),\ k = 1, \ldots, q_y\}, \tag{16}$$

where w_k denote samples of W, and $p_k = \Pr(\tilde{W}_{\text{srom}} = w_k) = \Pr(W \in \Gamma_k)$. It follows that

$$\begin{aligned}
\mu_{\tilde{W}_j}(r) &= \sum_{k=1}^{q_y} p_k\, w_{k,j}^r \\
c_{\tilde{W}_{i,j}} &= \sum_{k=1}^{q_y} p_k\, w_{k,i}\, w_{k,j} \\
F_{\tilde{W}_j}(z) &= \sum_{k=1}^{q_y} p_k\, 1(w_{k,j} \le z)
\end{aligned} \tag{17}$$

are the corresponding moments, correlation function, and marginal CDF of $\tilde{W}_{\text{srom}}$ defined by Eq. (16), where $1(A)$ denotes the indicator function and is equal to one if event A is true and zero otherwise.

The elements (w_k, p_k), $k = 1, \ldots, q_y$, defining the SROM are selected to minimize the discrepancy between properties of W and $\tilde{W}$, subject to constraints that each $p_k \ge 0$ and $p_1 + \cdots + p_{q_y} = 1$. The functional form for this discrepancy function is problem dependent; one example is given by

$$\begin{aligned}
&\gamma_1 \sum_{j=1}^{d} \sum_{r=1}^{\bar{r}} (\mu_W(r) - \mu_{\tilde{W}_j}(r))^2 + \gamma_2 \sum_{i,j=1}^{d} (\delta_{ij} - c_{\tilde{W}_{i,j}})^2 \\
&+ \gamma_3 \sum_{j=1}^{d} \int_{\mathbb{R}} (\Phi(z) - F_{\tilde{W}_j}(z))^2\, \mathrm{d}z,
\end{aligned} \tag{18}$$

where $\bar{r}$ is an integer defining the largest moment of interest, and $\gamma_k \ge 0$, $k = 1, 2, 3$, are scalars used to weight one term of this function more than the others. We note that, for the case when W is a random variable, we set $\gamma_2 = 0$ because W does not have a covariance function in this case. A more detailed discussion of some possible techniques that can be used to solve this optimization problem is discussed in [8].

Parameter q_y is referred to as the size of the SROM. For calculations, we introduce the concept of a refining SROM, that is, we first construct the SROM for the largest value for q_y considered, then select subsets of its samples to define lower order SROMs.

2.3.2. Local approximation for f

A local approximation for output $U(W)$ is given by [9]

$$\begin{aligned}
\tilde{U}_{\text{srom}}(W) = \sum_{k=1}^{q_y} 1(W \in \Gamma_k) \Bigg[& U(w_k) + (\nabla U(W))^{\mathrm{T}}\Big|_{W=w_k} \cdot (W - w_k) + \\
& \frac{1}{2!}(W - w_k)^{\mathrm{T}} \nabla U(W) \nabla^{\mathrm{T}} (W - w_k)\Big|_{W=w_k} + \cdots \Bigg],
\end{aligned} \tag{19}$$

where $\nabla(\cdot)$ denotes differentiation of $U(W)$ with respect to the coordinates of $W = (W_1, \ldots, W_d)^{\mathrm{T}}$, and $U(w_k)$ denotes the solution to Eq. (1) at $W = w_k$. The approximation is based on a Taylor series expansion of the multivariate function $U(W)$ over each cell Γ_k, $k = 1, \ldots, q_y$. For calculations, we truncate the expression within the square brackets in Eq. (19) at a finite number of terms. If we choose to retain one, two, or three terms, $\tilde{U}_{\text{srom}}(W)$ is a piecewise constant, piecewise linear, or piecewise quadratic function of W, respectively. We quantify this level of truncation with symbol q_u=0, 1, or 2.

In practice, the derivatives used in Eq. (19) must be computed numerically. Assuming central differences are used, the computational cost associated with the SROM method, that is, the number of times we must evaluate mapping f defined by Eq. (1), is given by

$$n_{\text{srom}} = (2q_u d + 1)q_y, \tag{20}$$

where q_y and q_u are defined by Eqs. (16) and (19), respectively, and d is the dimension of W.

3. Applications

In this section, we apply the stochastic collocation (SC), stochastic Galerkin (SG), and stochastic reduced order modeling (SROM) approaches to a variety of examples, as illustrated in Table 1. The first row of the table corresponds to $UY = 1$ [1, Example 8.5] so that the mapping from input Y to output U is simply $U = 1/Y$. By Table 1, Examples 1–3 correspond to this stochastic algebraic equation for the case when input Y is equal to $\exp(W)$, $|W|$, and $\Phi(W)$, respectively, where W is a standard Gaussian random variable. The second row in the table is more complex and provides an example of a stochastic ordinary differential equation. Examples 4 and 5 correspond to this input/output mapping for the case when $Y(x)$ is equal to $\exp(V(x))$ and $\Phi(V(x))$, respectively, where $V(x)$ is a homogeneous Gaussian random field with zero

Table 1
Examples #1–#5 considered in this study.

$f(y)$	$g(w)$		
	$\alpha + \beta\ \exp(w)$	$\lvert w\rvert$	$\alpha + (\beta - \alpha)\Phi(w)$
$1/y$	#1	#2	#3
$\int \mathrm{d}\xi / y(\xi)$	#4		#5

mean, unit variance, and prescribed covariance function.

3.1. Algebraic equation with random variables

In this section, we consider the simple stochastic algebraic equation $UY = 1$ with solution $U = 1/Y$. Approximate solutions are provided using each of the methods presented in Section 2. Further, the methods are applied to three examples, corresponding to the case where $Y = \alpha + \exp(\beta W)$ is lognormal, $Y = |W|$, and $Y = \alpha + (\beta - \alpha)\Phi(W)$ is uniform on $[\alpha, \beta]$, where α and β are known parameters.

To assess the performance of the stochastic collocation (SC), stochastic Galerkin (SG), and stochastic reduced order model (SROM) methods, we define a collection of accuracy metrics. Let U be a random function with CDF F_U defined on a probability space $(\Omega, \mathcal{F}, \Pr)$ and let $\tilde{U}_n$ denote an approximation for U when using one of SG, SC, or SROM to solve $UY = 1$. In each case, the approximation depends on n, the number of times we evaluate $f = 1/Y$. The following metrics can be used to assess the accuracy of each of the methods (taken from [18])

$$\begin{aligned}
\mu_1(n) &= E[(U - \tilde{U}_n)^2]/E[U^2] \\
\mu_2(n) &= \Pr(\{\tilde{U}_n \le \inf_{\omega\in\Omega} U(\omega)\} \bigcup \{\tilde{U}_n > \sup_{\omega\in\Omega} U(\omega)\}) \\
\mu_3(n) &= \Pr(\tilde{U}_n > F_U^{-1}(0.99)) \\
\mu_4(n) &= \mathrm{Kurt}[\tilde{U}_n]/\mathrm{Kurt}[U]
\end{aligned} \tag{21}$$

where $E[A]$ and $\mathrm{Kurt}[A]$ denote the expected value and kurtosis coefficient, respectively, of random variable A, and we use the notation $U(\omega)$ to denote a particular random sample of U corresponding to element $\omega \in \Omega$ of the sample space. These metrics quantify the relative mean square error, the probability of realizing a non-physical sample, the error in the 1% upper fractile of the distribution, and the relative error in kurtosis, respectively, and each depends on n. Note that, for examples where the distribution of U is not available analytically, we estimate the probability law for U by Monte Carlo simulation. Further, for examples where moments do not exist, we do not report values for metrics μ_2 and μ_4.

Example 1. For this example, we consider the case where input random variable Y is given by $Y = \alpha + \exp(\beta W)$, where $\alpha \ge 0$, $|\beta| > 0$ are deterministic parameters. The exact solution to the stochastic algebraic equation for this case is $U = 1/Y = 1/(\alpha + \exp(\beta W))$.

It can be shown that random variable $Y = g(W) = \alpha + \exp(\beta W)$ with parameters $\alpha \ge 0$, $|\beta| > 0$ has an exponential distribution with cumulative distribution function (CDF)

$$\begin{aligned}
F_Y(y) &= \Pr(Y \le y) = \Pr(\beta W \le \ln(y - \alpha)) \\
&= \Phi\left(\frac{1}{|\beta|} \ln(y - \alpha)\right), \quad y > \alpha,
\end{aligned} \tag{22}$$

where Φ denotes the CDF of a standard Gaussian random variable, that is, $\Phi(z) = \int_{-\infty}^{z} \phi(u)\,\mathrm{d}u$ where $\phi(u) = \exp(-u^2/2)/\sqrt{2\pi}$ is the PDF of a standard Gaussian random variable.

It follows that output random variable $U = f \circ g(W) = 1/(\alpha + \exp(\beta W))$ takes values in $(0, 1/\alpha)$ with CDF

$$\begin{aligned}
F_U(u) &= \Pr(U \le u) = \Pr\left(\beta W \ge \ln\left(\frac{1}{u} - \alpha\right)\right) \\
&= 1 - \Phi\left(\frac{1}{|\beta|} \ln\left(\frac{1}{u} - \alpha\right)\right), \quad 0 < u < 1/\alpha.
\end{aligned} \tag{23}$$

Further, U has moments $\mu_U(r) = E[U^r] = \int_0^{1/\alpha} u^r f_U(u)\,\mathrm{d}u$, where

$$f_U(u) = \frac{\mathrm{d}}{\mathrm{d}u} F_U(u) = \frac{1}{|\beta| u^2 (u^{-1} - \alpha)} \phi\left(\frac{1}{\beta} \ln\left(\frac{1}{u} - \alpha\right)\right). \tag{24}$$

Approximation by stochastic collocation: An approximation for U using stochastic collocation is given by a special case of Eq. (2), that is

$$\tilde{U}_{\mathrm{sc}}(W) = \sum_{k=0}^{m} \left(\frac{1}{\alpha + \exp(\beta w_k)}\right) \psi_k(W), \tag{25}$$

where $(w_0, w_1, \ldots, w_m)$ are selected collocation points and

$$\psi_k(W) = \prod_{j=0,\, j\ne k}^{m} \frac{W - w_j}{w_k - w_j}, \quad k = 0, \ldots, m, \tag{26}$$

are Lagrange polynomials. We note that the selection of the collocation points requires one to specify a bounded range for W, herein denoted by (α_W, β_W). The accuracy of the approximation defined by Eq. (25) will depend on the number of terms m, as well as the bounds (α_W, β_W). We use barycentric Lagrange interpolation [19] over Chebyshev nodes [20, Section 2.4] to define the collocation points, *i.e.*,

$$w_k = \frac{1}{2}(\alpha_W + \beta_W) + \frac{1}{2}(\beta_W - \alpha_W) \cos\left(\frac{2k - 1}{2(m + 1)}\pi\right), \tag{27}$$

so as to minimize numerical issues such as Runge's phenomenon.

Approximation by stochastic Galerkin: An approximation for input $Y = \alpha + \exp(\beta W)$ using the stochastic Galerkin method is given by Eq. (9). It is common to take $\{h_k\}$ to be Hermite polynomials with $W \sim \mathcal{N}(0, 1)$ a Gaussian random variable with zero mean and unit variance. In this case

$$\begin{aligned}
a_k &= \frac{1}{k!} \int_{\mathbb{R}} g(w) h_k(w) \phi(w)\,\mathrm{d}w \\
&= \frac{1}{k!} \int_{\mathbb{R}} (\alpha + e^{\beta w}) h_k(w) \phi(w)\,\mathrm{d}w \\
&= \begin{cases} \alpha + e^{\beta^2/2}, & k = 0 \\ \dfrac{\beta^k}{k!} e^{\beta^2/2}, & k = 1, \ldots, n_y \end{cases}
\end{aligned} \tag{28}$$

The corresponding Galerkin approximation for output random variable $U = f(Y)$ is

$$\tilde{U}_{\mathrm{sg}} = \sum_{l=0}^{n_u} b_l h_l(W), \tag{29}$$

where

$$\begin{aligned}
b_l &= \frac{1}{l!} E[f(\tilde{Y}_{\mathrm{sg}}) h_l(W)] \\
&= \frac{1}{l!} \int_{\mathbb{R}} \frac{h_l(w)\phi(w)}{\sum_{k=0}^{n_y} a_k h_k(w)}\,\mathrm{d}w, \\
& l = 0, \ldots, n_u.
\end{aligned} \tag{30}$$

Non-intrusive estimates of the $\{b_l\}$ can be obtained by Eq. (8).

Approximation by stochastic reduced order models: Let $\tilde{U}_{\mathrm{srom}}$ denote an approximation for U using the stochastic reduced order model approach; the approximation is based on (1) a reduced order model for input W and (2) a local approximation for mapping $U(W) = 1/(\alpha + \exp(\beta W))$. The SROM for W is defined by Eq. (16), computed by minimizing Eq. (18) with $\gamma_2 = 0$. By Eq. (19), the local approximation for $U(W)$ yields

$$\tilde{U}_{\mathrm{srom}}(W) = \sum_{k=1}^{q_y} \left[\frac{1}{\alpha + e^{\beta w_k}}\right] 1(W \in I_k) \tag{31a}$$

$$\tilde{U}_{\text{srom}}(W) = \sum_{k=1}^{q_y} \left[\frac{1}{\alpha + e^{\beta w_k}} - \frac{\beta e^{\beta w_k}}{(\alpha + e^{\beta w_k})^2}(W - w_k) \right] 1(Y \in \Gamma_k) \tag{31b}$$

$$\tilde{U}_{\text{srom}}(W) = \sum_{k=1}^{q_y} \left[\frac{1}{\alpha + e^{\beta w_k}} - \frac{\beta e^{\beta w_k}}{(\alpha + e^{\beta w_k})^2}(W - w_k) + \left(\frac{2\beta^2 e^{2\beta w_k}}{(\alpha + e^{\beta w_k})^3} - \frac{\beta^2 e^{\beta w_k}}{(\alpha + e^{\beta w_k})^2} \right) \frac{(W - w_k)^2}{2} \right] 1(Y \in \Gamma_k) \tag{31c}$$

for the case of $q_u = 0, 1, 2$, respectively, that is, for the case of a piecewise constant, piecewise linear, and piecewise quadratic local approximation.

Results: Fig. 3 illustrates the performance of the three methods for $n \approx 100$, where n represents one of n_{sc}, n_{sg}, and n_{srom} defined by Eqs. (4), (8), and (20), respectively. Parameters $\alpha = \beta = 1$, $n_y = n_u = 10$, $q_u = 1$, $\gamma_1 = 1$, $\gamma_3 = 3$, and $\bar{r} = 6$ were used for calculations. In each plot, the PDF of U defined by Eq. (24) is illustrated by a solid black line. It is evident that the stochastic collocation method (panel (a)) provides a very accurate estimate for the PDF of U for $0 \le u \le 1$, but permits non-physical samples of U, that is, samples outside the support of the distribution of U. The stochastic Galerkin method (panel (b)) provides a poor estimate for the PDF of U within $[0, 1]$ and also allows for a significant number of samples in $(-\infty, 0)$. The SROM method (panel (c)) provides an excellent approximation for the PDF of U.

The metrics defined by Eq. (21) are illustrated in Figs. 4 and 5 as a function of n. When possible, black dashed lines are used to denote perfect performance as quantified by a particular metric. For example, a method delivers perfect estimates of $\Pr(\tilde{U}_n > F_U^{-1}(0.99))$ when $\mu_3 = 0.01$; the dashed black line in Fig. 5 (a) is used to represent this. Further, a method delivers perfect performance as quantified by metric μ_2 when $\mu_2 = 0$. For clarity, results for metric μ_2 are plotted on a log scale. Therefore, any zero values for μ_2 were changed to 1×10^{-8} so as to appear in the plot; a dashed black line drawn at 1×10^{-8} is used in Fig. 4(b) to represent this.

Results indicate the stochastic reduced order model is superior for this example for most of the performance metrics considered. The relative mean square error (*cf.* Fig. 4(a)) and probability of non-physical samples using the SROM (*cf.* Fig. 4(b)) both decrease rapidly with increasing n. We note that, when using the local approximation defined by Eq. (31a), that is, when $q_u = 0$, the probability of realizing a non-physical sample by the SROM method is zero for any n by construction.

By Fig. 5(a), the SROM provides accurate estimates of tail probabilities for the case of $q_u = 1, 2$. The method delivers poor estimates of metric μ_3 when $q_u = 0$, however, because the local approximation and, therefore, the CDF of U, is piecewise constant in this case. It follows that, in general, piecewise linear or piecewise quadratic local approximations are needed for accurate tail estimates when using the SROM method. Accurate estimates of kurtosis (*cf.* Fig. 5(b)) can be achieved for $q_u = 0$, 1, and 2, but only when the number of terms in the SROM q_y is large.

The performance of the stochastic collocation and stochastic Galerkin methods can also be assessed by the metrics defined by

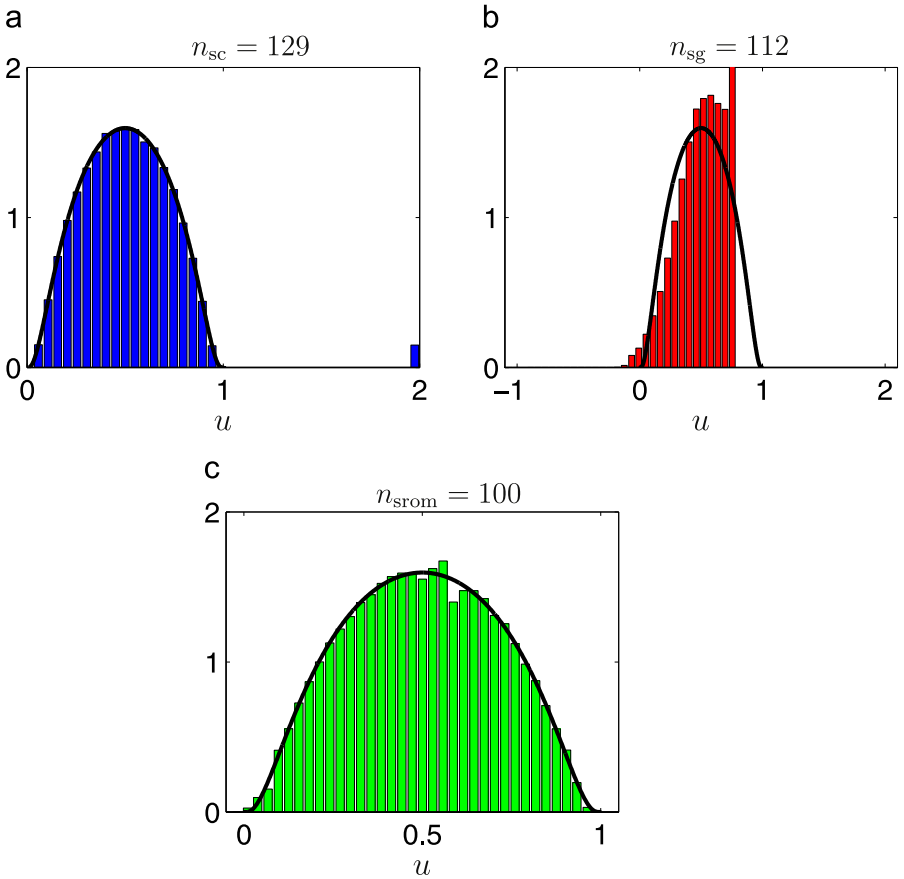

Fig. 3. Results for Example #1: estimates of the PDF of U using (a) stochastic collocation, (b) stochastic Galerkin, and (c) stochastic reduced order models. The estimates are obtained with approximately the same computational cost.

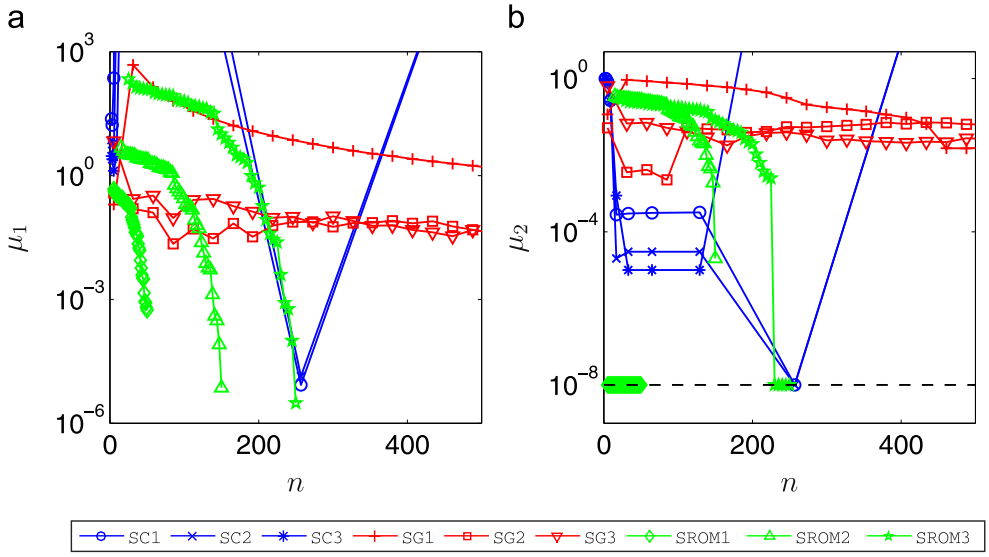

Fig. 4. Results for Example #1: metrics $\mu_1(n)$ and $\mu_2(n)$ defined by Eq. (21) as a function of the total number of function evaluations, n. Note: SC1 – SC3 denote stochastic collocation with $(\alpha_W, \beta_W) = (-3.4, 3.4)$, $(-3.9, 3.9)$, and $(-4.2, 4.2)$; SG1–SG3 denote stochastic Galerkin with $n_u = n_y = 5$, 10, and 15; and SROM1–SROM3 denote stochastic reduced order model with $q_u = 0$, 1, and 2.

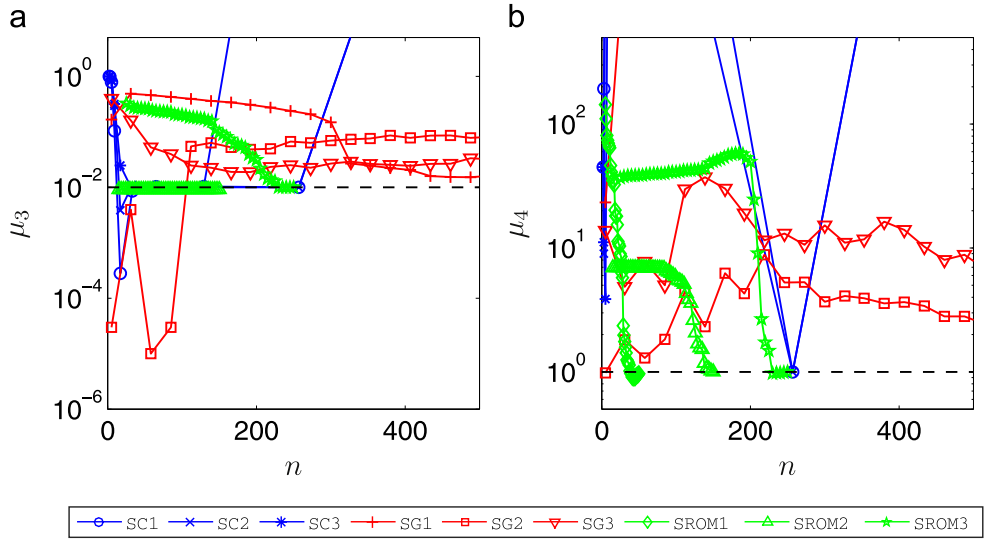

Fig. 5. Results for Example #1: metrics $\mu_3(n)$ and $\mu_4(n)$ defined by Eq. (21) as a function of the total number of function evaluations, n. Note: SC1 – SC3 denote stochastic collocation with $(\alpha_W, \beta_W) = (-3.4, 3.4)$, $(-3.9, 3.9)$, and $(-4.2, 4.2)$; SG1–SG3 denote stochastic Galerkin with $n_u = n_y = 5$, 10, and 15; and SROM1–SROM3 denote stochastic reduced order model with $q_u = 0$, 1, and 2.

Eq. (21) as a function of n; these results are also shown in Figs. 4 and 5. It can be observed that, in general, the performance of the stochastic Galerkin method improves with increasing n, but at a much slower rate than for the SROM method. In general, the stochastic collocation method delivers satisfactory estimates of metric μ_3, but the other metrics can vary dramatically. As n grows large, the response surface can become highly oscillatory and produce erroneous results. Further, it is evident that the stochastic collocation method is very sensitive to parameters α_W and β_W that define the upper and lower bounds on the collocation points.

Example 2. We next consider the solution to $UY = 1$ assuming Y takes a different distribution. Elementary calculations show that input random variable $Y = g(W) = |W|$ has CDF

$$F_Y(y) = \Pr(Y \le y) = 2\Phi(y) - 1, \quad y > 0. \tag{32}$$

It follows that the CDF and PDF of output random variable $U = f \circ g(W) = 1/|W|$ are

$$F_U(u) = \Pr(U \le u) = \Pr\left(|W| \ge \frac{1}{u}\right) = 2\left(1 - \Phi\left(\frac{1}{u}\right)\right), \quad u > 0 \quad \text{and}$$
$$f_U(u) = \frac{\mathrm{d}}{\mathrm{d}u} F_U(u) = \frac{2}{u^2}\phi\left(\frac{1}{u}\right), \quad u > 0. \tag{33}$$

By Eq. (33), $E[U]$ does not exist and $E[U^r]$ is unbounded for $r > 1$. Metrics μ_1 and μ_4 defined by Eq. (21) are therefore undefined for this example, and the stochastic Galerkin method cannot be used.

Metrics μ_2 and μ_3 defined by Eq. (21) are illustrated in Fig. 6 as a function of n. By panel (a), the stochastic collocation solution allows for non-physical samples of U in some cases. In contrast, the SROM produces zero non-physical samples of U for all cases. Both methods underestimate the 1% upper fractile (panel (b)) of U.

Example 3. Next consider the case where input random variable $Y = g(W) = \alpha + (\beta - \alpha)\Phi(W)$ is distributed uniformly on bounded interval (α, β), where $0 < \alpha < \beta < \infty$ are parameters. It follows that the CDF, PDF, and moments of output random variable

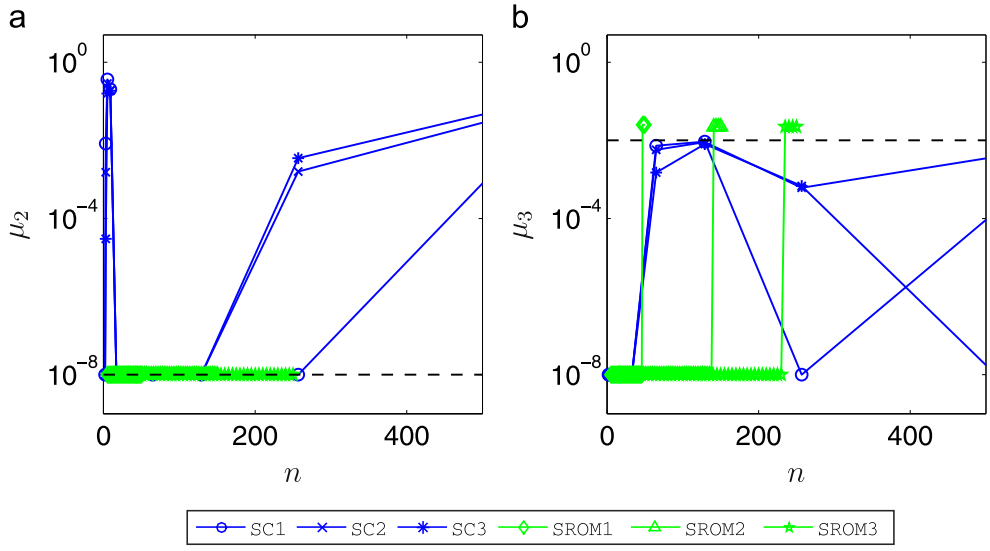

Fig. 6. Results for Example #2: metrics $\mu_2(n)$ and $\mu_3(n)$ defined by Eq. (21) as a function of the total number of function evaluations, n. Note: SC1–SC3 denote stochastic collocation with $(\alpha_W, \beta_W) = (-5, 5), (-6, 6)$, and $(-8, 8)$; and SROM1–SROM3 denote stochastic reduced order model with q_u=0, 1, and 2.

$U = f \circ g(W) = 1/(\alpha + (\beta - \alpha)\Phi(W))$ are given by

$$F_U(u) = 1 - \frac{u^{-1} - \alpha}{\beta - \alpha}, \quad 1/\beta \le u \le 1/\alpha,$$
$$f_U(u) = \frac{\mathrm{d}}{\mathrm{d}u} F_U(u) = \frac{1}{u^2(\beta - \alpha)} \quad \text{and}$$
$$\mu_U(r) = E[U^r] = \begin{cases} \dfrac{\log \alpha - \log \beta}{\alpha - \beta} & r = 1 \\ \dfrac{\alpha^{1-r} - \beta^{1-r}}{(\beta - \alpha)(r-1)} & r > 1 \end{cases} \tag{34}$$

The three methods defined in Section 2 are now used to defined approximations for $U = 1/(\alpha + (\beta - \alpha)\Phi(W))$. Because the support of U is bounded in this case, the stochastic Galerkin approximation for U is defined by Eq. (29) with $\{h_k(W)\}$ equal to the Legendre polynomials evaluated on $W \sim \mathcal{U}(-1, 1)$. For this case, we can show that

$$\begin{aligned} a_k &= \frac{2k+1}{2} \int_{-1}^{1} g \circ \Phi^{-1}\left(\frac{w+1}{2}\right) h_k(w)\, \mathrm{d}w \\ &= \frac{2k+1}{2} \int_{-1}^{1} \alpha + (\beta - \alpha)\left(\frac{w+1}{2}\right) h_k(w)\, \mathrm{d}w \\ &= \begin{cases} (\beta + \alpha)/2, & k = 0 \\ (\beta - \alpha)/2, & k = 1 \\ 0, & k = 2, 3, \ldots \end{cases} \end{aligned} \tag{35}$$

Each of the metrics defined by Eq. (21) is illustrated in Figs. 7 and 8 as a function of n for the case of $\alpha = 1$, $\beta = 3$, $n_y = n_u$, $\gamma_1 = 1$, $\gamma_3 = 3$, and $\bar{r} = 6$. The stochastic collocation method is superior for this example in all metrics considered. The stochastic reduced order model also performs well, but requires a larger value for n to achieve satisfactory performance. The stochastic Galerkin method performs rather poorly in comparison. By Fig. 7, metrics μ_1 and μ_2 decrease with increasing n, but at a much slower rate than for the other two methods. Fig. 8 illustrates that the stochastic Galerkin method can provide poor estimates for the 1% upper fractile and kurtosis of U, even for large n.

3.2. Differential equation with random fields

In this section, we apply the three methods presented in Section 2 to approximate the solution to a stochastic differential equation modeling heat transport in a random media. Let U: $D \to \mathbb{R}$ be the solution of the following boundary value problem (BVP)

$$-\nabla \cdot (Y(x) \nabla U(x)) = Z(x), \quad x \in D \subset \mathbb{R}^3, \tag{36}$$

where $Y(x)$ and $Z(x)$ are random fields. This is a version of the heat equation, where U denotes the temperature of a medium in D with thermal diffusivity Y, subjected to heat source Z (see [1, Section 9.4] and [21, Section A.3]).

To simplify the discussion, consider the heat equation in one spatial dimension of length l, that is, $D = [-l/2, l/2]$, with no heat source. Under these assumptions, the BVP simplifies to

$$(Y(x)U'(x))' = 0, \quad -l/2 \le x \le l/2. \tag{37}$$

With deterministic boundary conditions $U(-l/2) = u_{-l/2}$ and $U(l/2) = u_{l/2}$, the solution to Eq. (37) is a random field defined by

$$U(x) = u_{-l/2} + C \int_{-l/2}^{x} \frac{\mathrm{d}\xi}{Y(\xi)} \tag{38}$$

where

$$C = (u_{l/2} - u_{-l/2}) \left(\int_{-l/2}^{l/2} \frac{\mathrm{d}\xi}{Y(\xi)} \right)^{-1} \tag{39}$$

is a random variable.

Let

$$V(x) = \frac{1}{\sqrt{n_v}} \sum_{k=1}^{n_v} A_k \cos\left(\frac{2\pi k x}{l}\right) + B_k \sin\left(\frac{2\pi k x}{l}\right), \quad -l/2 \le x \le l/2 \tag{40}$$

be a Gaussian random field, where A_k and B_k are independent $\mathcal{N}(0, 1)$ random variables. It can be shown that $V(x)$ is homogeneous with zero mean, unit variance, and correlation function

$$\rho(\tau) = E[V(x)\, V(x+\tau)] = \frac{1}{n_v} \sum_{k=1}^{n_v} \cos\left(\frac{2\pi k \tau}{l}\right). \tag{41}$$

We model random field $Y(x)$ representing the thermal

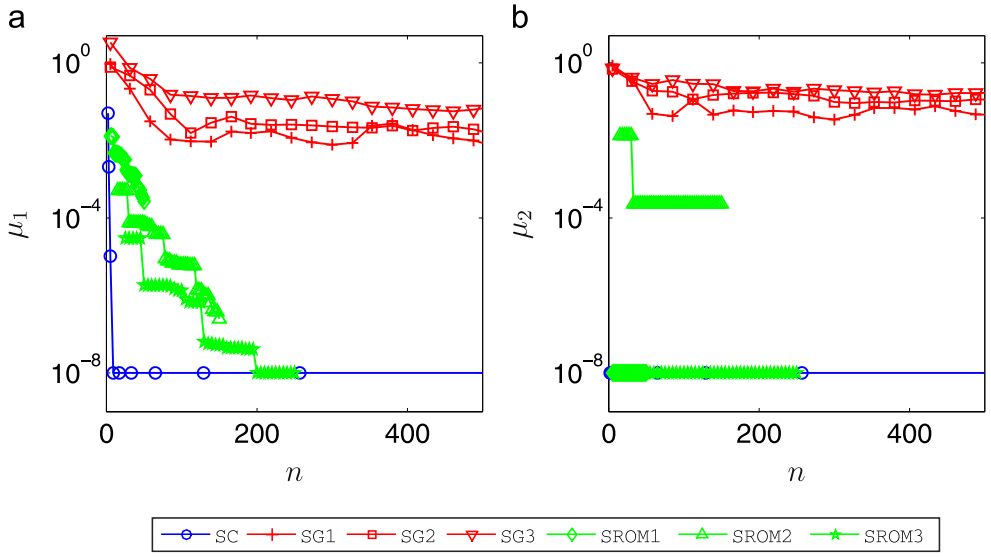

Fig. 7. Results for Example #3: metrics $\mu_1(n)$ and $\mu_2(n)$ defined by Eq. (21) as a function of the total number of function evaluations, n. Note: SC denotes stochastic collocation; SG1–SG3 denote stochastic Galerkin with $n_u = n_y = 5$, 10, and 15; and SROM1 – SROM3 denote stochastic reduced order model with q_u=0, 1, and 2.

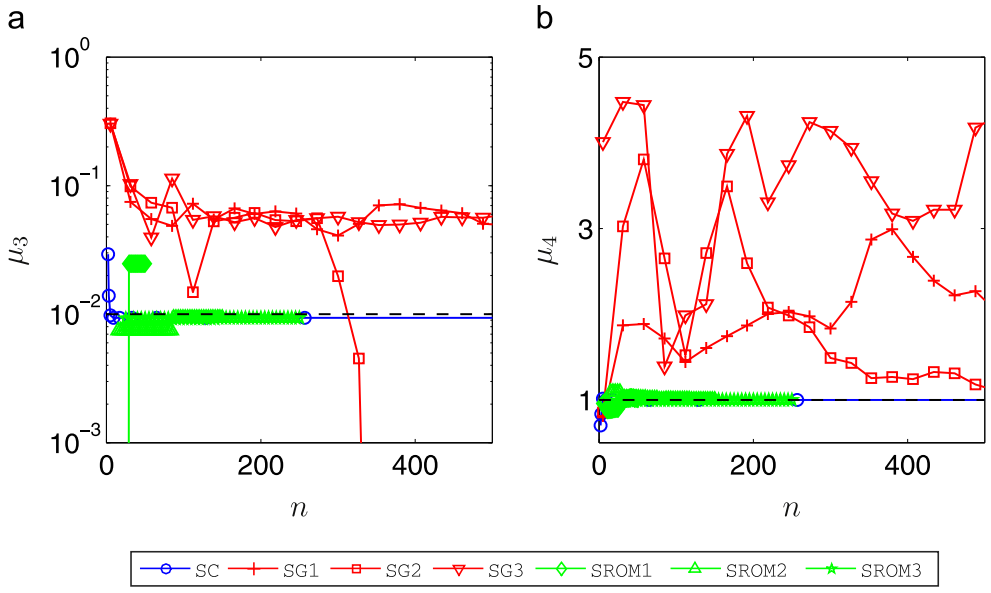

Fig. 8. Results for Example #3: metrics $\mu_3(n)$ and $\mu_4(n)$ defined by Eq. (21) as a function of the total number of function evaluations, n. Note: SC denotes stochastic collocation; SG1–SG3 denote stochastic Galerkin with $n_u = n_y = 5$, 10, and 15; and SROM1–SROM3 denote stochastic reduced order model with q_u=0, 1, and 2.

diffusivity of the random media as $Y(x) = g(V(x))$, where $V(x)$ is defined by Eq. (40) and g is one of two mappings: (i) $g(V) = \alpha + \exp(\beta V)$ or (ii) $g(V) = \alpha + (\beta - \alpha)\Phi(V)$. It follows by Eqs. (38) and (40) that the solution to the BVP depends on x and W, *i.e.*, $U = U(x, W)$, where $W = (A_1, \ldots, A_{n_\nu}, B_1, \ldots, B_{n_\nu})$ is a vector of iid standard Gaussian random variables.

Ten samples of random field $U(x)$ defined by Eq. (38) are illustrated in Fig. 9 assuming l=12, $u_{-l/2} = 0$, $u_{l/2} = 1$, and n_ν=5. Panel (a) illustrates these samples assuming the thermal diffusivity is $Y(x) = g(V(x)) = 0.01 + \exp(0.4V(x))$; samples assuming thermal diffusivity that are modeled as $Y(x) = g(V(x))$= $0.01 + 3.99\Phi(V(x))$ are illustrated in panel (b). We note that the two collections of samples are qualitatively very different.

In the following sections, we apply the methods of stochastic collocation, stochastic Galerkin, and stochastic reduced order modeling to provide approximations for solution $U(x)$ defined by Eq. (38). The accuracy of each approximation is quantified by the following metrics:

$$\begin{aligned}\mu_1(n) &= E[(U(0) - \tilde{U}_n(0))^2]/E[U(0)^2]\\ \mu_2(n) &= \Pr\left(\tilde{U}_n(-4) > \zeta_1 \bigcap \tilde{U}_n(2) \le \zeta_2\right)\\ &\quad - \Pr\left(U(-4) > \zeta_1 \bigcap U(2) \le \zeta_2\right)\end{aligned} \tag{42}$$

where metric μ_1 is the relative mean square error, and metric μ_2 quantifies how accurate the methods are at approximating a point on a joint CDF. We will assess the metrics assuming the two different mappings $Y(x) = g(V(x))$ for the thermal diffusivity discussed above; these two cases are referred to as Examples #4 and #5 below.

Example 4. For this example, we consider the case where the thermal diffusivity of the random media is given by

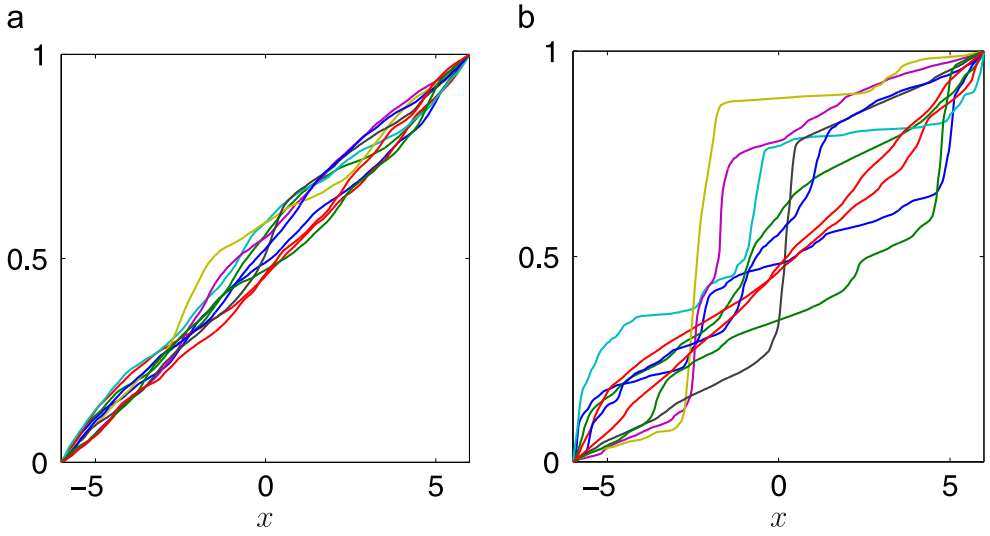

Fig. 9. 10 samples of random field $U(x)$ defined by Eq. (38) with thermal diffusivity given by (a) $Y(x) = \alpha + \exp(\beta V(x))$ and (b) $Y(x) = \alpha + (\beta - \alpha)\Phi(V(x))$.

$Y(x) = \alpha + \exp(\beta V(x))$, where $V(x)$ is a Gaussian random field defined by Eq. (40) with n_ν=5, and parameters $\alpha = 1/100$, $\beta = 2/5$. By this construction, $Y(x)$ is a non-Gaussian translation random field with marginal CDF equal to Eq. (22).

By Eq. (38) and Fig. 9, $U(x)$ is a random field whose marginal probability law depends on x. To further demonstrate this,

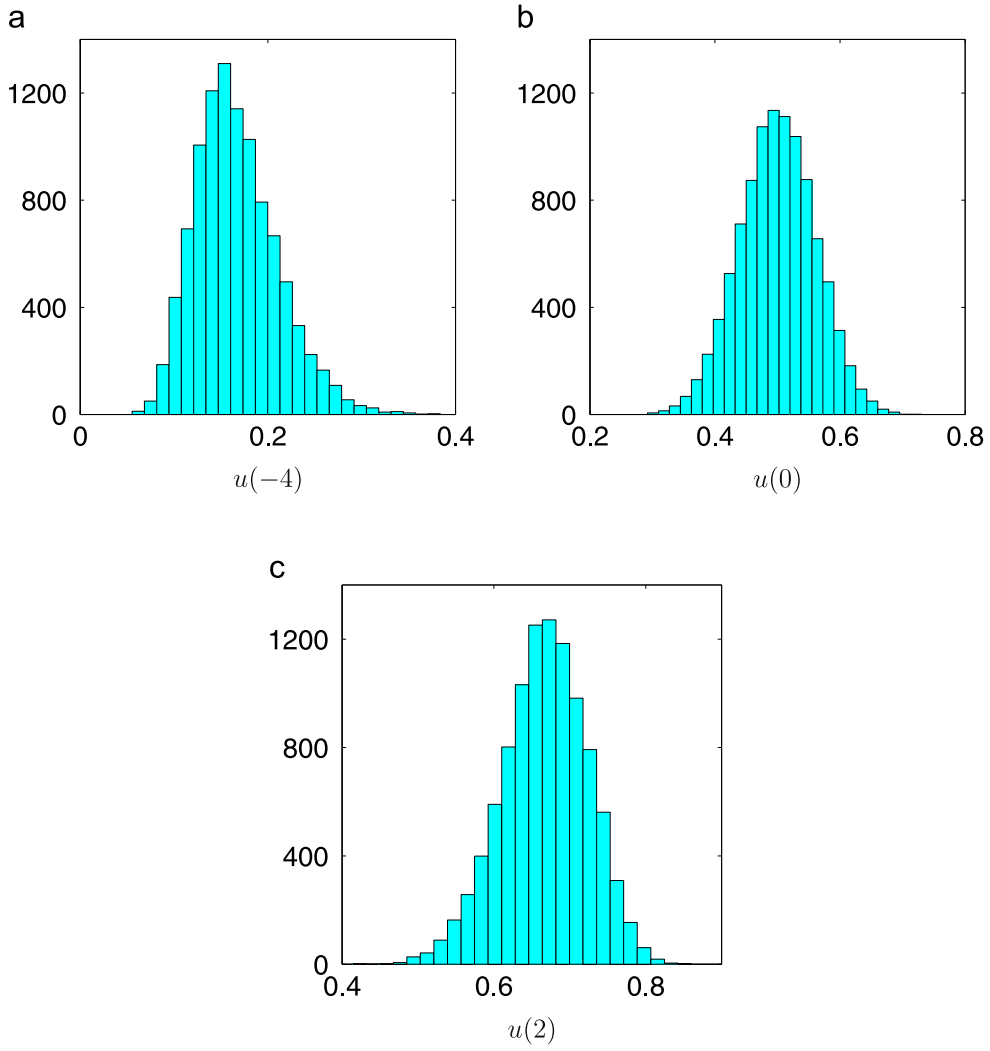

Fig. 10. Histograms of 100,000 samples of $U(x)$ for Example #4 at 3 spatial locations: (a) $x = -4$, (b) x=0, and (c) x=2.

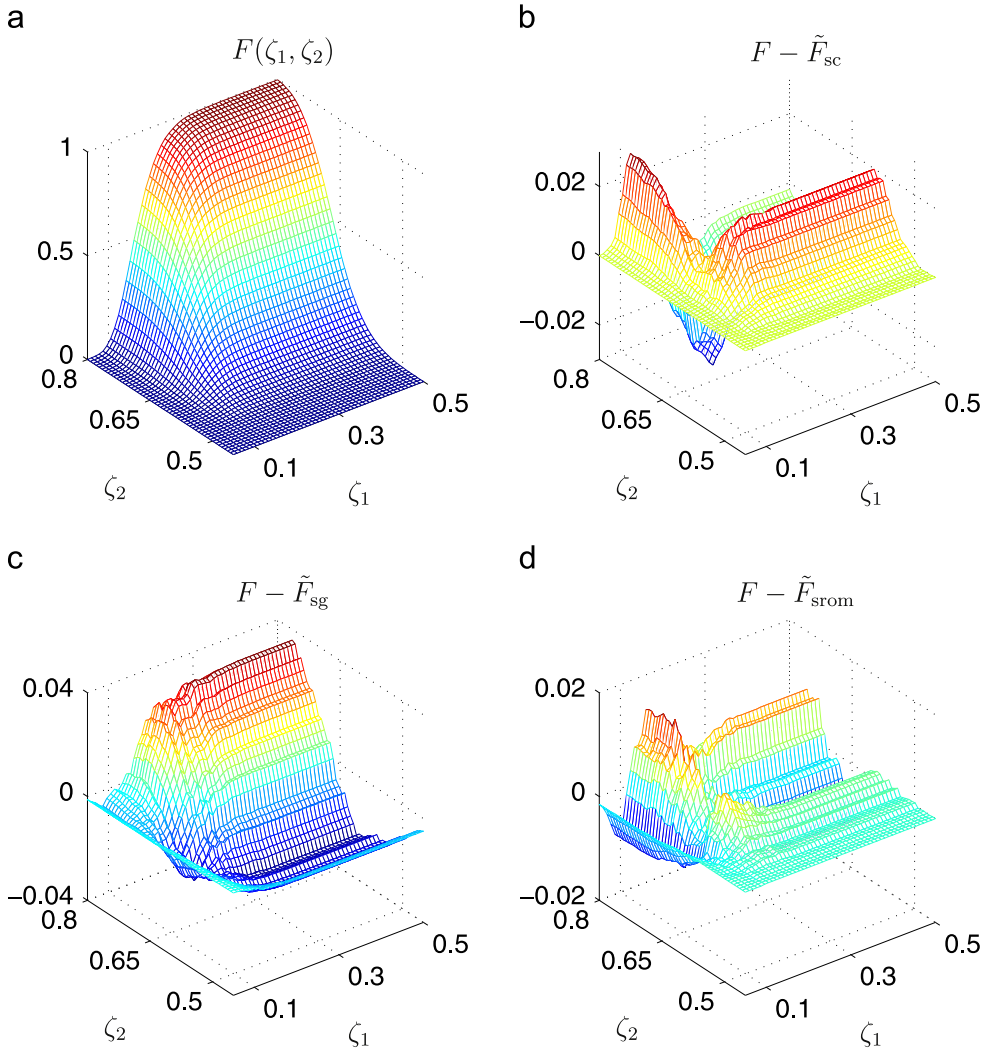

Fig. 11. Plots of $\Pr(U(-4) \le y_1 \cap U(2) \le y_2)$ for Example #4: (a) exact result, (b) error using stochastic collocation, (c) error using stochastic Galerkin, and (d) error using stochastic reduced order models.

consider Fig. 10, which illustrates histograms of 100,000 random samples of $U(x)$ at three different spatial locations; panels (a), (b), and (c) illustrate these samples for $x = -4$, $x=0$, and $x=2$, respectively. In addition, the joint CDF of random variables $U(-4)$ and $U(2)$, defined as $F(\zeta_1, \zeta_2) = \Pr(U(-4) \le \zeta_1 \cap U(2) \le \zeta_2)$, is illustrated in Fig. 11(a).

Approximation by stochastic collocation: The approximation for random field $U(x)$ by the stochastic collocation method is given by Eq. (2) with $d = 2n_v$ and $m_1 = \cdots = m_d = m$ so that, by Eq. (4), the cost associated with using the stochastic collocation method for this example is equal to $n_{\rm sc} = (m+1)^{2n_v}$. The so-called "curse of dimensionality" is evident here, as the cost associated with this approach increases exponentially with the number of underlying random variables, $d = 2\ n_v$.

Let $\tilde{F}_{\rm sc}(\zeta_1, \zeta_2)$ denote an estimate of the approximation for the joint CDF of $U(-4)$ and $U(2)$, that is, an estimate for $F(\zeta_1, \zeta_2) = \Pr(U(-4) \le \zeta_1 \cap U(2) \le \zeta_2)$, using the stochastic collocation method. The error $F(\zeta_1, \zeta_2) - \tilde{F}_{\rm sc}(\zeta_1, \zeta_2)$ is illustrated in Fig. 11 (b) assuming $m=1$, meaning that the SC method was implemented with cost $n_{\rm sc} = 2^{10} = 1,024$. Results indicate that the error ranges from -0.02 to 0.02 with the largest errors occurring for small values of ζ_1 and ζ_2.

Approximation by stochastic Galerkin: We can also utilize the non-intrusive implementation of the stochastic Galerkin method; the corresponding approximation for $U(x)$ is given by Eq. (6) with the unknown $\{b_k(x)\}$ computed by Eq. (8). The cost of using the stochastic Galerkin method in this way is $n_{\rm sg}$ defined by Eq. (8). Let $\tilde{F}_{\rm sg}$ denote an approximation for the joint CDF of $U(-4)$ and $U(2)$ by the stochastic Galerkin approach. The error $F(\zeta_1, \zeta_2) - \tilde{F}_{\rm sg}(\zeta_1, \zeta_2)$ is illustrated in Fig. 11(c) with $n_{\rm sg} = 100,000$ showing that the error is roughly twice as large as the error produced by the SC method, but with a cost that is 100 times greater. As indicated with the plot, the SG method can deliver inaccurate estimates of the joint CDF even with a relatively high cost.

Approximation by stochastic reduced order models: To approximate $U(x)$ by the stochastic reduced order modeling approach, we take two steps. First, we build an SROM for random vector $W = (A_1, \ldots, A_{n_y}, B_1, \ldots, B_{n_y})$ by minimizing the objective function defined by Eq. (18) with $\gamma_1 = \gamma_2 = \gamma_3 = 1$. Second, we build a local approximation for mapping $W \mapsto U$ using Eq. (19). We will consider the cases where we truncate the expression within the square brackets in Eq. (19) at one, two, and three terms, referred to as truncation levels of $q_u = 0$, 1, and 2, respectively. This corresponds to the cases where the local approximation for $W \mapsto U$ is piecewise constant, piecewise linear, and piecewise quadratic in W.

Let $\tilde{F}_{\rm srom}$ denote an approximation for the joint CDF of $U(-4)$ and $U(2)$ by the stochastic reduced order modeling approach. The error is illustrated in Fig. 11(d) with $q_y = 40$, $q_u = 1$ and $d = 2n_v = 10$ so that, by Eq. (20), the cost of using the stochastic reduced order modeling approach to make this estimate of the joint CDF is

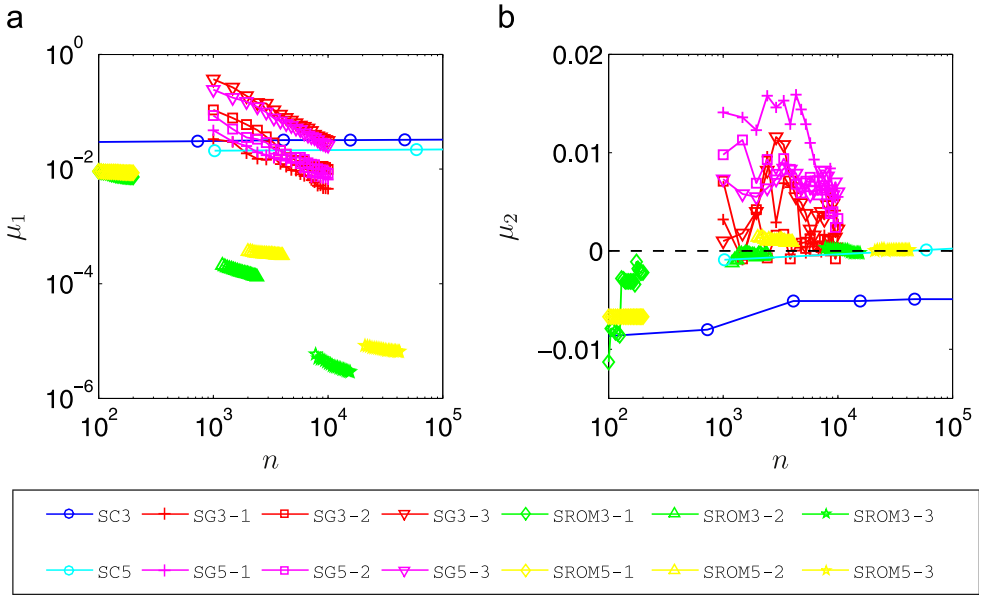

Fig. 12. Results for Example #4: metrics $\mu_1(n)$ and $\mu_2(n)$ as a function of the total number of function evaluations, n. Note: SC3 and SC5 denote stochastic collocation with $n_\nu=3$ and $n_\nu=5$; SG3-1, SG3-2, and SG3-3 denote stochastic Galerkin with $n_\nu=3$, $n_y=10$, and $n_u=50$, 100, and 250; SG5-1, SG5-2, and SG5-3 denote stochastic Galerkin with $n_\nu=5$, $n_y=10$, and $n_u=50$, 100, and 250; SROM3-1, SROM3-2, and SROM3-3 denote stochastic reduced order model with $n_\nu=3$ and $q_u=0$, 1, 2; and SROM5-1, SROM5-2, and SROM5-3 denote stochastic reduced order model with $n_\nu=5$ and $q_u=0$, 1, 2.

$n_{\text{srom}} = 840$. The error is slightly smaller than the corresponding error using SC. We note that, by comparing the results illustrated in Fig. 11, the SROM approach delivers the most accurate estimate for the joint CDF of $U(-4)$ and $U(2)$, while requiring the smallest computational cost.

Fig. 12 illustrates the performance of the three methods as quantified by the two metrics defined by Eq. (42). Metrics μ_1 and μ_2 are plotted as a function of n, where n represents one of n_{sc}, n_{sg}, and n_{srom} defined above. Parameters $\alpha = 1/100$, $\beta = 2/5$, $l = 12$, $\zeta_1 = 0.2$, and $\zeta_2 = 0.6$, were used for all calculations. The stochastic reduced order model is clearly superior to the other approaches for this example. The relative mean square error (panel (a)) is much smaller than either of the other two methods for all n, and the estimate of the probability $\mu_2 = \Pr(U(-4) > 0.2 \cap U(2) \le 0.6)$ (panel (b)) is the most accurate for all n with the exception of when $n=100$. The stochastic Galerkin approach illustrates that its accuracy does improve with increasing n, but at a slower rate and with much larger error than the SROM approach. The stochastic collocation method provides fairly accurate estimates of μ_2, but poor estimates of μ_1, and the method does not improve as n increases.

Results are shown for both $n_\nu=3$ and $n_\nu=5$, the number of terms in Eq. (40) defining the underlying Gaussian random field. As expected the addition of more underlying random variables requires a greater cost from each method in order to achieve the same level of accuracy.

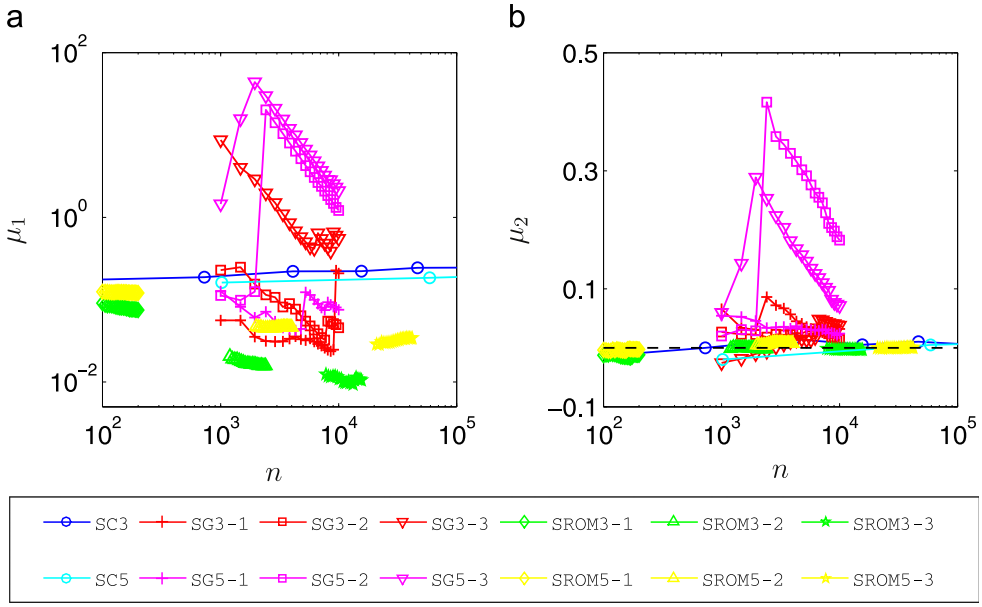

Fig. 13. Results for Example #5: metrics $\mu_1(n)$ and $\mu_2(n)$ as a function of the total number of function evaluations, n. Note: SC3 and SC5 denote stochastic collocation with $n_\nu=3$ and $n_\nu=5$; SG3-1, SG3-2, and SG3-3 denote stochastic Galerkin with $n_\nu=3$, $n_y=10$, and $n_u=50$, 100, and 250; SG5-1, SG5-2, and SG5-3 denote stochastic Galerkin with $n_\nu=5$, $n_y=10$, and $n_u=50$, 100, and 250; SROM3-1, SROM3-2, and SROM3-3 denote stochastic reduced order model with $n_\nu=3$ and $q_u=0$, 1, 2; and SROM5-1, SROM5-2, and SROM5-3 denote stochastic reduced order model with $n_\nu=5$ and $q_u=0$, 1, 2.

Example 5. As a final example, we consider the case where the thermal diffusivity is modeled by random field $Y(x) = \alpha + (\beta - \alpha)\Phi(V(x))$, where $0 < \alpha < \beta < \infty$ are parameters, and $V(x)$ is a Gaussian random field defined by Eq. (40); parameters $\alpha = 1/100$, $\beta = 4$ are used for all calculations.

Fig. 13 illustrates the performance of the three methods as quantified by the two metrics defined by Eq. (42) for this example. With $n_v = 3$, the SROM method is slightly more accurate with respect to metric μ_1 than the stochastic Galerkin approach; this difference becomes more pronounced for $n_v = 5$. The stochastic collocation method provides fairly accurate estimates of μ_2, but poor estimates of μ_1, and the method only marginally improves as n increases.

4. Conclusions

The stochastic collocation (SC) and stochastic Galerkin (SG) methods are two well-established and successful approaches for solving general stochastic problems. Herein we provided a comparison of the two methods with a recently developed method based on stochastic reduced order models (SROMs). Comparisons were made by assessing the accuracy of each method as a function of the computational effort required for a specific set of example problems; our evaluation only holds for the examples considered in the paper.

The examples considered in the paper included a stochastic algebraic equation, as well as a stochastic differential equation modeling heat transport in a random media. In all cases, non-Gaussian random functions were considered and used to model the uncertain quantities. Some general observations on the performance of the methods included (1) the stochastic Galerkin method, when limited to the case where output U had finite variance, exhibited mean square convergence as expected; (2) for problems where U was bounded, the stochastic collocation method worked well; (3) the stochastic collocation and stochastic Galerkin methods can deliver non-physical samples of the solution for the case when U is not bounded; and (4) when using the piecewise constant local approximation, the probability of realizing a non-physical sample using the SROM method is zero. In general, we conclude that, for the examples studied herein, the SROM approach outperformed the other methods in most metrics, particularly for cases where output U was unbounded. However, it was observed that first or second order local approximations were required in order for the SROM method to deliver accurate estimates of tail probabilities.

The purpose of this study was not to criticize the SC or SG methods, which have proven very useful for a broad range of applications, nor was it to provide overall ratings of these methods as compared to the SROM method. Rather, our objectives were to present the SROM approach as an alternative approach to solving stochastic problems, and to provide information on the computational effort required by the implementation of each method, while simultaneously assessing their performance for a collection of specific problems.

References

[1] M. Grigoriu, Stochastic Systems, Springer-Verlag, London, 2012.
[2] M. Grigoriu, Stochastic Calculus: Applications in Science and Engineering, Birkhäuser, Boston, MA, 2002.
[3] I. Babuška, F. Nobile, R. Tempone, A stochastic collocation method for elliptic partial differential equations with random input data, SIAM J. Numer. Anal. 45 (3) (2007) 1005–1034.
[4] F. Nobile, R. Tempone, C.G. Webster, A sparse grid stochastic collocation method for partial differential equations with random input data, SIAM J. Numer. Anal. 46 (5) (2008) 2309–2345.
[5] D. Xiu, Numerical Methods for Stochastic Computations: A Spectral Method Approach, Princeton University Press, Princeton, 2010.
[6] R.G. Ghanem, P.D. Spanos, Stochastic Finite Elements: A Spectral Approach, Revised Edition, Dover Publications, Inc., New York, NY, 2003.
[7] H.G. Matthies, A. Keese, Galerkin methods for linear and nonlinear stochastic partial differential equations, Comput. Methods Appl. Mech. Eng. 194 (12–16) (2005) 1295–1331.
[8] M. Grigoriu, Reduced order models for random functions. Application to stochastic problems, Appl. Math. Model. 33 (1) (2009) 161–175.
[9] M. Grigoriu, A method for solving stochastic equations by reduced order models and local approximations, J. Comput. Phys. 231 (2012) 6495–6513.
[10] F. Nobile, R. Tempone, C. Webster, An anisotropic sparse grid collocation method for elliptic partial differential equations with random input data, SIAM J. Numer. Anal. 46 (5) (2008) 2411–2442.
[11] J. Foo, X. Wan, G. Karniadakis, The multi-element probabilistic collocation method (ME-PCM): error analysis and applications, J. Comput. Phys. 227 (2008) 9572–9595.
[12] R. Askey, J. Wilson, Some Basic Hypergeometric Orthogonal Polynomials that Generalize Jacobi Polynomials, American Mathematical Society: Memoirs of the American Mathematical Society, vol. 54, American Mathematical Society, Providence, RI, 1985.
[13] R.A. Todor, C. Schwab, Convergence rates for sparse chaos approximations of elliptic problems with stochastic coefficients, IMA J. Numer. Anal. 27 (2) (2007) 232–261.
[14] O.P.L. Maitre, H.N. Najm, R.G. Ghanem, O.M. Knio, Multi-resolution analysis of Wiener-type uncertainty propagation schemes, J. Comput. Phys. 197 (2004) 502–531.
[15] X. Wan, G.E. Karniadakis, An adaptive multi-element generalized polynomial chaos method for stochastic differential equations, J. Comput. Phys. 209 (2) (2005) 617–642.
[16] I.W. Busbridge, Some integrals involving Hermite polynomials, J. Lond. Math. Soc. 23 (1948) 135–141.
[17] J.C. Adams, On the expression of the product of any two Legendre's coefficients by means of a series of Legendre's coefficients, Proc. R. Soc. Lond. 27 (1878) 63–71.
[18] R.V. Field Jr., M. Grigoriu, On the accuracy of the polynomial chaos approximation, Probab. Eng. Mech. 19 (1-2) (2004) 65–80.
[19] J.-P. Berrut, L.N. Trefethen, Barycentric Lagrange interpolation, SIAM Rev. 46 (3) (2004) 501–517.
[20] W. Gautschi, Numerical Analysis: An Introduction, Birkhäuser, Boston, MA, 1997.
[21] I. Stakgold, Boundary Value Problems of Mathematical Physics, vol. 1, SIAM, Philadelphia, 2000.