A MULTILEVEL SOLUTION STRATEGY FOR THE STOCHASTIC GALERKIN METHOD FOR PARTIAL DIFFERENTIAL EQUATIONS WITH RANDOM INPUT DATA

SARAH V. OSBORN*, ERIC T. PHIPPS†, AND VICTORIA E. HOWLE‡

Abstract. We discuss solving partial differential equations (PDEs) with random input data using the stochastic Galerkin method. This method can often be computationally demanding as the method suffers from the *curse of dimensionality* where the computational effort increases greatly as the stochastic dimension increases. We consider a multilevel solution strategy for the stochastic Galerkin method that employs hierarchies of spatial and stochastic approximations in an attempt to diminish some of the computational burden. Analysis of the proposed multilevel method and numerical results are presented that compare the multilevel approach to the traditional, single-level stochastic Galerkin method.

1. Introduction. When using partial differential equations (PDEs) to model physical problems, the exact values of coefficients are often not known exactly. To obtain more realistic results, the coefficients are modeled by a set of random variables which induce variability in the solution of the physical model. We consider the numerical solution of elliptic PDEs with random input data using the stochastic Galerkin (SG) method [9, 11, 25]. The SG approach provides a powerful method to obtain highly accurate solutions by transforming a PDE with random coefficients into a set of coupled deterministic PDEs using a spectral finite element approach. However when a large number of uncertain parameters are required to characterize the uncertainty in the system, the method suffers from the *curse of dimensionality*, where the computational cost grows exponentially as a function of the number of random variables and the degree of the multivariate stochastic basis polynomials used to approximate the underlying probability space of the problem.

We are interested in exploring a multilevel based solution strategy to alleviate some of the computational burden associated with the SG method. Similar to the ideas behind multigrid solvers for linear equations, the proposed multilevel method uses a hierarchical sequence of spatial approximations to the underlying PDE model, combined with varying stochastic discretizations in order to reduce computational cost. Multilevel approaches have been successfully applied to sampling methods for solving PDEs with random coefficients. Sampling methods, for instance Monte Carlo [7, 15] and stochastic collocation [2, 24], solve a sequence of deterministic PDEs at given sample points in the stochastic space in contrast to non-sampling methods like the SG method. Multilevel acceleration methods have been proposed and investigated for the Monte Carlo method in [4, 5, 6, 21], and for the stochastic collocation method by Teckentrup et al. in [22]. We propose a multilevel approach similar to the method in [22] for the SG method that employs hierarchies of spatial and stochastic approximations. The goal is to reduce the computational effort and considerable memory requirements inherent to the SG method while maintaining the overall accuracy of the solution.

We start by describing the mathematical model and SG formulation for the model problem in Section 2. In Section 3, the formulation of the multilevel method is given along with error and cost analysis. In Section 4, numerical results are provided that illustrate the performance of the proposed multilevel method compared against the standard, single-level SG method. We conclude in Section 5 summarizing the results and outline further research directions and improvements.

2. Problem Formulation. Let D be an open subset of $\mathbb{R}^d$ and let $(\Omega, \mathcal{F}, \mu)$ be a complete probability space, where Ω is the sample space, $\mathcal{F}$ is the σ-algebra generated by Ω and $\mu : \mathcal{F} \to [0, 1]$ is the probability measure. Given the random field $a(\mathbf{x}, \omega) : \Sigma \times \overline{D} \to \mathbb{R}$ and function $f(\mathbf{x}) \in L^2(D)$, the stochastic steady-state diffusion equation with homogeneous Dirichlet boundary conditions is given by

$$\begin{aligned} -\nabla \cdot (a(\mathbf{x}, \omega)\nabla u(\mathbf{x}, \omega)) &= f(\mathbf{x}) \quad && \text{in } D \times \Omega \\ u(\mathbf{x}, \omega) &= 0 && \text{on } \partial D \times \Omega. \end{aligned} \tag{2.1}$$

We are interested in finding a random function $u(\mathbf{x}, \omega) : \overline{D} \times \Omega \to \mathbb{R}$ which satisfies (2.1). We make the assumption that the input random field can be represented in terms of a finite number $M \in \mathbb{N}$ of

*Texas Tech University, sarah.osborn@ttu.edu,
†Sandia National Laboratories, etphipp@sandia.gov,
‡Texas Tech University, victoria.howle@ttu.edu

independent random variables with given probability density functions.

ASSUMPTION 2.1. *(Finite-Dimensional Noise Assumption) Assume $a(\mathbf{x},\omega)$ can be represented by a finite number $M \in \mathbb{N}$ of independent real-valued random variables $\boldsymbol{\xi} = [\xi_1,\dots,\xi_M]^T$ where $\xi_m : \Omega \to \Gamma_m$, with given probability density functions $\rho_M : \Gamma_M \to \mathbb{R}_0^+$, $m = 1,\dots,M$.*

Due to the independence of the random variables $\{\xi_m\}_{m=1}^M$, their joint probability density function $\rho : \Gamma_1 \times \cdots \times \Gamma_M \to \mathbb{R}_0^+$ is given by

$$\rho(\mathbf{y}) = \rho_1(y_1)\dots\rho_M(y_M),$$

where the vector $\mathbf{y} = (y_1,\dots,y_M)^T \in \mathbb{R}^M$ is introduced with $y_m := \xi_m(\omega)$. In particular, we assume that the diffusion coefficient $a(\mathbf{x},\mathbf{y})$ is given by a truncated Karhunen-Loève expansion [11, 13] given by

$$a(\mathbf{x},\mathbf{y}) = \hat{a}_0(\mathbf{x}) + \sum_{i=1}^{M} \hat{a}_i(\mathbf{x})\mathbf{y}_i, \tag{2.2}$$

where $\hat{a}_0(\mathbf{x})$ is the mean of the random field and $\hat{a}_i(\mathbf{x})$ for $i > 0$ involves the eigenfunctions and eigenvalues of the assumed known covariance kernel of the random field, and y_i for $i = 1,\dots,M$ are identically distributed, uncorrelated random variables.

After making Assumption 2.1, we have by the Doob-Dynkin Lemma that the solution u of the stochastic elliptic boundary value problem (2.1) can also be represented by a finite number of number of random variables $\boldsymbol{\xi}$ or, equivalently, the vector $\mathbf{y} \in \mathbb{R}^M$. The solution $u(\mathbf{x},\omega)$ thus has a deterministic equivalent $u(\mathbf{x},\mathbf{y})$, with the probability space $(\Gamma, \mathcal{B}(\Gamma), \rho(\mathbf{y})d\mathbf{y})$ replacing the abstract probability space $(\Omega,\mathcal{F},\mu)$.

2.1. Variational Formulation. Define $L^2_\rho(\Gamma)$ to be the set of real-valued square-integrable functions on Γ and $H_0^1(D)$ to be the subspace of the Sobolev space $H^1(D)$ of real-valued functions of $\bar{D}$ that vanish on the boundary ∂D. Define

$$L^2_\rho(\Gamma; H_0^1(D)) := \left\{ v : \bar{D}\times\Gamma \to \mathbb{R} : \int_\Gamma \rho(\mathbf{y}) \left\| v(\cdot,\mathbf{y}) \right\|^2_{H_0^1(D)} d\mathbf{y} < \infty \right\} \tag{2.3}$$

which is isomorphic to the tensor space $H_1^0(D) \otimes L^2_\rho(\Gamma)$. For ease of notation, we denote $V := H_0^1(D)$, $S := L^2_\rho(\Gamma)$, and $W := L^2_\rho(\Gamma; H_0^1(D))$ with norm

$$\|v\|^2_W = \int_\Gamma \|v\|^2_V \rho(\mathbf{y})d\mathbf{y} = \int_\Gamma\int_D |\nabla v(\mathbf{x},\mathbf{y})|^2 d\mathbf{x}\rho(\mathbf{y})d\mathbf{y}.$$

We consider weak solutions $u \in W$ on $D\times\Gamma$ satisfying

$$b(u,v) = \ell(v), \quad \forall v \in W, \tag{2.4}$$

where $b(\cdot,\cdot) : W\times W \to \mathbb{R}$ and $\ell : W \to \mathbb{R}$ are defined as

$$b(u,v) = \int_\Gamma \rho \int_D a\nabla u\cdot\nabla v d\mathbf{x}d\mathbf{y}, \;\; \ell(v) = \int_\Gamma \rho\int_D fv d\mathbf{x}d\mathbf{y}. \tag{2.5}$$

In order to have a unique solution, we assume that the approximate diffusion coefficient given by (2.2) satisfies

$$0 < a_{min} \le a(\mathbf{x},\mathbf{y}) \le a_{max} < \infty \quad \text{ a.e. in } D\times\Gamma, \tag{2.6}$$

where $a_{min}, a_{max} \in \mathbb{R}$. In particular, $a \in L^\infty_\rho(\Gamma; L^\infty(D))$. The Lax-Milgram theorem can be used to show that there exists a unique solution to this problem providing the above assumption on a holds and $f \in L^2_\rho(\Gamma; L^2(D))$.

2.2. Discretization. Following the SG method [9, 11, 25], the variational formulation (2.4) is discretized using a Galerkin projection onto a finite-dimensional subspace of $W \simeq V \otimes S$. The finite-dimensional subspace for the spatial domain, denoted $V^h \subset V$, is a spatial finite element basis and the finite-dimensional subspace of the stochastic domain, denoted $S^p \subset S$, is constructed using global polynomials.

After suitable spaces V^h and S^p have been constructed, the $h \times p$-SG approximation is the tensor product finite element function $u_p^h \in W_p^h := V^h \otimes S^p$, that satisfies

$$b(u_p^h, v) = l(v) \quad \forall v \in V^h \otimes S^p. \tag{2.7}$$

Applying standard finite element analysis and assuming (2.6) holds and $f \in L^2(D) \otimes S$ with $W_p^h \subset W$, (2.7) has a unique solution $u_p^h \in W_p^h$. Also, we have that u_p^h is the best approximation in the energy norm, that is,

$$\left\|u - u_p^h\right\|_E = \inf_{v \in W_p^h} \|u - v\|_E , \tag{2.8}$$

where

$$\|v\|_E^2 := \int_\Gamma \int_D a(\mathbf{x}, \mathbf{y}) \left|\nabla v(\mathbf{x}, \mathbf{y})\right|^2 d\mathbf{x}\rho(\mathbf{y})d\mathbf{y}. \tag{2.9}$$

For $u \in W$, we now proceed with discussing the particular finite-dimensional subspaces V^h and S^p used and analyze the discretization error of the SG approximation $u_p^h \in V^h \otimes S^p$. Note that this error estimate does not take into account the error of using M random variables to represent the random input parameter a, nor the error of the iterative method.

2.2.1. Spatial Discretization. For the spatial discretization, the finite element method is used where $V^h \subset V$ is defined to be the space of basis functions consisting of continuous, piecewise linear functions on a shape-regular triangulation $\mathcal{T}_h$ of D that satisfies the Dirichlet boundary conditions with maximum mesh spacing parameter h. We consider the finite element approximation of u in $V^h \otimes S$ denoted by u^h.

Definition 2.2. *A semi-discrete weak solution on $D \times \Gamma$ to* (2.4) *is a function $u^h \in V^h \otimes S$ that satisfies*

$$b(u^h(\cdot, \mathbf{y}), v) = l(v) \quad \forall v \in V^h \otimes S, \tag{2.10}$$

where $V^h \subset H_0^1(D)$.

If (2.6) holds for the approximate diffusion coefficient and $f \in L^2(D) \otimes S$, then (2.10) has a unique solution $u^h \in V^h \otimes S$. Also under the same conditions, u^h is the best approximation in the energy norm, that is,

$$\left\|u - u^h\right\|_E = \inf_{v \in V^h \otimes S} \|u - v\|_E . \tag{2.11}$$

For our multilevel analysis, we introduce a mapping P^h from $H_0^1(D) \to V^h$ that acts only on the spatial components of $u \in W$ and make the following assumption on the spatial approximation of the semi-discrete solution.

Definition 2.3. *The Galerkin projection $P^h : W \to L^2_\rho(\Gamma; V^h)$ is the orthogonal projection satisfying $\left\|u - u^h\right\|_E = \left\|u - P^h u\right\|_E$ or equivalently,*

$$b(P^h u, v) = b(u, v) \quad \forall v \in L^2_\rho(\Gamma; V^h). \tag{2.12}$$

Assumption 2.4. *(spatial approximation error) Assume $u \in W$ is the unique solution to* (2.4) *and $u^h \in V^h \otimes S$ is the unique solution to* (2.10) *for a particular $h > 0$, then there exist positive constants s and C_D, independent of h, such that*

$$\left\|u - u^h\right\|_W \leq C_D h^s.$$

When using piecewise linear finite elements, this assumption is valid, see [14] for further details.

2.2.2. Stochastic Discretization. The finite-dimensional space S^p for the stochastic dimension is constructed using global polynomials orthogonal with respect to the joint probability density function ρ, called generalized polynomial chaos (gPC) [23, 25]. Consider the set of univariate polynomials $\{\psi_k^i\}$ in y_i on the interval $\Gamma_i \subset \mathbb{R}$ orthonormal with respect to the measure ρ_i for each $i = 1, \cdots, M$ with inner product given by

$$\left\langle \psi_j^i(y_i),\, \psi_k^i(y_i) \right\rangle := \int_{\Gamma_i} \psi_j^i(y_i)\psi_k^i(y_i)\rho_i(y_i)dy_i = \delta_{jk}, \quad j,k = 0,1,\ldots \tag{2.13}$$

where j and k are the degrees of the polynomials ψ_j^i and ψ_k^i, respectively.

Given a multi-index $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_M) \in \mathbb{N}_0^M$ where $|\boldsymbol{\alpha}| = \sum_{i=1}^M \alpha_i$, the M-variate tensor product polynomial is constructed as

$$\Psi_{\boldsymbol{\alpha}}(\mathbf{y}) = \prod_{i=1}^{M} \psi_{\alpha_i}^i(y_i). \tag{2.14}$$

These polynomials are orthonormal with respect to the inner product defined by

$$\langle \Psi_{\boldsymbol{\alpha}}(\mathbf{y}),\, \Psi_{\boldsymbol{\beta}}(\mathbf{y}) \rangle := \int_{\Gamma} \Psi_{\boldsymbol{\alpha}}(\mathbf{y})\Psi_{\boldsymbol{\beta}}(\mathbf{y})\rho(\mathbf{y})d\mathbf{y} = \delta_{\boldsymbol{\alpha}\boldsymbol{\beta}}. \tag{2.15}$$

Then for $p \geq 0$, define S^p to be the complete polynomial space of total order at most p in M dimensions given by

$$S^p = \operatorname{span}\left\{ \Psi_{\boldsymbol{\alpha}}(\mathbf{y}) : |\boldsymbol{\alpha}| \leq p \right\}, \tag{2.16}$$

where the dimension of S^p is

$$Q := \frac{(p+M)!}{p!M!}. \tag{2.17}$$

We now write $S^p = \operatorname{span}\{\Psi_i(\mathbf{y}) : 0 \leq i \leq Q-1\}$ where each scalar index $i \in 0, \ldots, Q-1$ is assigned to some multi-index $\boldsymbol{\alpha}$ whose components provide the degrees of the univariate basis polynomials.

We assume the input random field follows a uniform probability distribution, so the Legendre polynomial basis is used to construct the basis. Now we discuss the error in the stochastic approximation of the solution when using a truncated gPC expansion and introduce a projection associated with the finite-dimensional subspace S^p.

DEFINITION 2.5. *The pth-order gPC approximation of the solution $u(\mathbf{x}, \mathbf{y}) \in W$ can be obtained by projecting u onto the subspace S^p, that is,*

$$P_p u := u_p(\mathbf{x}, \mathbf{y}) = \sum_{i=0}^{Q-1} u_i(\mathbf{x})\Psi_i(\mathbf{y}) \quad \forall \mathbf{x} \in D,$$

where $P_p u$ denotes the orthogonal projection operator from $L^2_\rho(\Gamma)$ onto S_p.

We make the following assumption on the stochastic approximation to be used in the multilevel analysis.

ASSUMPTION 2.6. *(stochastic approximation error) Assume $u \in W$ and that there exist a positive constant C_Γ and a sequence of positive numbers $\{\lambda_k\}_{k \in \mathbb{N}_0}$ tending to zero monotonically, such that for all $k \in N_0$,*

$$\|u - P_k u\|_W \leq C_\Gamma \lambda_k,$$

where C_Γ is a constant independent of k.

When using uniform random variables to represent the input random field, this assumption is valid using total order Legendre polynomials, see [1] for further details.

3. Multilevel Method Formulation. Let $\{h_\ell\}_{\ell=0}^{L}$ for $L \in \mathbb{N}$ be the mesh width of a sequence of increasingly fine triangulations of the spatial domain D, which leads to a hierarchic sequence of finite-dimensional subspaces

$$V^{h_0} \subset V^{h_1} \subset \cdots \subset V^{h_\ell} \subset \cdots \subset H_0^1(D).$$

Assume for simplicity that $h_\ell = m^{-\ell} h_0$ for some $m \in \mathbb{N}$, $m > 1$. For $\ell = 0, \ldots, L$, let $u^{h_\ell}(\mathbf{x}, \mathbf{y}) \in V^{h_\ell} \otimes S$ denote the semi-discrete weak solution satisfying (2.10) and assume that each semi-discrete weak solution satisfies the spatial approximation error bound given in Assumption 2.4.

For any $L \in \mathbb{N}$, we have the identity

$$u^{h_L} = \sum_{\ell=0}^{L} (u^{h_\ell} - u^{h_{\ell-1}}),$$

with the notation $u^{h_{-1}} := 0$.

The multilevel SG method then approximates the semi-discrete weak solutions at each level using a truncated gPC expansion of total order $p_{L-\ell}$. It follows from Assumption 2.4 that as $h \to 0$, lower order stochastic approximations are necessary to achieve a required accuracy. Thus, our multilevel solution strategy can be written as

$$u_L^{ML} = \sum_{\ell=0}^{L} (u_{p_{L-\ell}}^{h_\ell} - u_{p_{L-\ell}}^{h_{\ell-1}}), \tag{3.1}$$

where p_ℓ are the polynomial orders used in the gPC expansion and $0 < p_0 \leq p_1 \leq \cdots \leq p_L \leq \infty$.

3.1. Error Analysis. We now analyze the convergence of the proposed multilevel method.

LEMMA 3.1. *Let $\{P_{p_\ell}\}_{\ell=0}^{L}$ be the $L^2_\rho(\Gamma)$ projection operators $P_{p_\ell} : S \to S^{p_\ell}$ for $\ell = 0, \ldots, L$ defined in Definition 2.5. Let $u \in W$ and $u^{h_L} \in V^h \otimes S$ satisfy* (2.4) *and* (2.10) *respectively and let $u_L^{ML} \in V^{h_L} \otimes S^{p_L}$ be the L-level multilevel approximation defined in* (3.1)*. Then,*

$$\left\| u - u_L^{ML} \right\|_W \leq C_a \left(\left\| u - u^{h_L} \right\|_W + 2 \sum_{l=0}^{L} \left\| u - P_{p_{L-\ell}} u \right\|_W \right) := C_a \left(e_h + e_p \right), \tag{3.2}$$

where $C_a = \sqrt{\frac{a_{\max}}{a_{\min}}}$.

Proof. Let the spatial projection operators $\left\{P^{h_\ell}\right\}_{\ell=0}^{L}$ be defined as in (2.3). Then for $P^h u \in V^h$ it follows that $P_p(P^h u) \in V^h \otimes S^p$ since P_p acts only on the $\mathbf{y}$-components of $P^h u$. Using the optimality property given by (2.8) yields

$$\begin{aligned}
\left\| u - u_L^{ML} \right\|_E &= \left\| u - \sum_{\ell=0}^{L} \left(u_{p_{L-\ell}}^{h_\ell} - u_{p_{L-\ell}}^{h_{\ell-1}} \right) \right\|_E \\
&\leq \left\| u - P^{h_L} u + P^{h_L} u - \sum_{\ell=0}^{L} \left(P_{p_{L-\ell}} P^{h_\ell} u - P_{p_{L-\ell}} P^{h_{\ell-1}} u \right) \right\|_E \\
&\leq \left\| u - u^{h_L} \right\|_E + \sum_{\ell=0}^{L} \left\| P^{h_\ell} \left(u - P_{p_{L-\ell}} u \right) - P^{h_{\ell-1}} \left(u - P_{p_{L-\ell}} u \right) \right\|_E \\
&\leq \left\| u - u^{h_L} \right\|_E + \sum_{\ell=0}^{L} 2 \left\| u - P_{p_{L-\ell}} u \right\|_E .
\end{aligned}$$

Since the energy norm is equivalent to the norm on W, we have the bound in the desired norm. ∎

This result shows that the total error allows for a splitting of the error into a spatial error component e_h and a stochastic error component e_p. Now, we seek to show that when the polynomial orders are chosen

appropriately, the stochastic discretization error e_p of the multilevel approximation converges at the same rate as the spatial discretization error e_h, thus resulting in a convergence result for the total error.

LEMMA 3.2. *Under the spatial and stochastic approximation error assumptions given in Assumptions 2.4 and 2.6, the SG multilevel approximation* (3.1) *admits the error bound:*

$$\left\|u - u_L^{ML}\right\| \leq C_a \left(C_D h_L^s + 2 \sum_{\ell=0}^{L} C_\Gamma \lambda_{L-\ell} \right) \leq 2 K_a C_D h_L^s, \tag{3.3}$$

when the polynomial orders are chosen so that

$$\lambda_{L-\ell} \leq \left(2C_\Gamma (L+1)\right)^{-1} C_D h_L^s.$$

Proof. For the spatial discretization error e_h, we have from Assumption 2.4 that $e_h \leq C_D h_L^s$.
Now for the stochastic discretization error e_p, using Assumption 2.6 we have

$$e_p = 2 \sum_{\ell=0}^{L} \left\|u - P_{p_{L-\ell}} u\right\|_W \leq 2 \sum_{\ell=0}^{L} C_\Gamma \lambda_{L-\ell}.$$

Then, we choose the polynomial orders so that

$$\lambda_{L-\ell} \leq \left(2C_\Gamma (L+1)\right)^{-1} C_D h_L^s. \tag{3.4}$$

Thus, the error e_p is bounded by

$$e_p \leq 2 \sum_{\ell=0}^{L} C_\Gamma \left(2C_\Gamma (L+1)\right)^{-1} C_D h_L^s = C_D h_L^s$$

as claimed. □

3.2. Cost Analysis. In this section, we investigate a cost metric for the multilevel SG method that provides formulas for determining the number of levels, the spatial discretization parameters, and stochastic parameters that should be used. We assume that the stochastic approximation error bound given in Assumption 2.6 holds with $\lambda_\ell = p_\ell^{-r}$ for some $r > 0$, where p_ℓ is the total order degree used in the stochastic discretization. This is a reasonable assumption when using total order multivariate Legendre polynomials where r is related to the stochastic regularity of the solution u, see [1].

When approximating the SG solution a linear system of size $N_{\mathbf{x}} \cdot Q$ must be solved, where $N_{\mathbf{x}}$ is the number of spatial degrees of freedom and $Q = \frac{(M+p)!}{M!p!}$ is the number of stochastic degrees of freedom with an order p approximation space with M random variables. We estimate the computational cost of the method by determining the degrees of freedom that must be resolved. Making the estimate $Q = \mathcal{O}\left(p^M\right)$, we assume the computational cost for the multilevel approximation can be estimated with

$$\mathrm{COST}^{ML} = \sum_{\ell=0}^{L} p_{L-\ell}^M C_\ell^{\mathbf{x}}, \tag{3.5}$$

where $C_\ell^{\mathbf{x}}$ is the number of degrees of freedom for solving the deterministic problem with mesh size h_ℓ plus the number of degrees of freedom for solving the deterministic problem with mesh size $h_{\ell-1}$, i.e. $C_\ell^{\mathbf{x}} = h_\ell^{-d} + h_{\ell-1}^{-d}$, where d is the spatial dimension. Since we assume that the hierarchy of meshes are obtained using uniform refinement, we have $h_\ell = m^{-\ell} h_0 = m^{-\ell}$ as we assume $h_0 = 1$ for simplicity. Now we analyze the cost of the multilevel method to achieve a desired level of accuracy and derive formulas for the parameters in the multilevel formulation. We denote by $\mathrm{COST}_\epsilon^{ML}$ the computational cost required to achieve a desired accuracy ϵ.

THEOREM 3.3. *Suppose Assumptions 2.4 and 2.6 hold with $\lambda_\ell = p_\ell^{-r}$ for some $r > 0$. Then, for any $\epsilon > 0$, there exists an integer L such that*

$$\left\|u - u_L^{ML}\right\|_W \leq \epsilon \quad \text{and} \quad \mathit{COST}_\epsilon^{ML} \lesssim \epsilon^{\frac{-M}{r} - \frac{d}{s}}.$$

Proof. Following [22], we consider the spatial and stochastic error contributions, e_h and e_r, separately as shown in Lemma 3.1. To achieve the desired accuracy, it is sufficient to bound both the error contributions by $\frac{\epsilon}{2}$. Recall that we have $h_\ell = m^{-\ell} h_0 = m^{-\ell}$. In addition, we assume without loss of generality that $h_0 = 1$. If this is not the case, this will only scale the constants C_D and C_Γ. We start by considering the spatial error and require $e_h \leq \frac{\epsilon}{2}$. It is sufficient to require $C_D h_L^s \leq \frac{\epsilon}{2}$, so we choose L to be

$$L = \left\lceil \frac{1}{s} \log_m \left(\frac{2C_D}{\epsilon} \right) \right\rceil < \frac{1}{s} \log_m \left(\frac{2C_D}{\epsilon} \right) + 1, \tag{3.6}$$

so that $h_L \leq \left(\frac{\epsilon}{2C_D} \right)^{1/s}$. This requirement fixes the number of levels L.

Now, we want to minimize $\sum_{\ell=0}^{L} p_{L-\ell}^M C_\ell^{\mathbf{x}}$ subject to the constraint $e_p \leq \frac{\epsilon}{2}$ which is equivalent to $\sum_{\ell=0}^{L} p_{L-\ell}^{-r} \leq \frac{\epsilon}{4C_\Gamma}$. Treating the polynomial orders $\{p_\ell\}$ as continuous variables, the Lagrange multiplier method yields the optimal choice for the polynomial order at each level given by

$$p_{L-\ell} = \left\lceil (4C_\Gamma)^{1/r} \, \epsilon^{-1/r} g(m, L)^{1/r} m^{\frac{-\ell d}{r+M}} \right\rceil, \tag{3.7}$$

where $g(m, L) = \sum_{\ell=0}^{L} m^{\frac{\ell d r}{r+M}}$. Now that the polynomial orders have been chosen, we examine the complexity of the multilevel approximation:

$$\begin{aligned} \mathrm{COST}_\epsilon^{ML} &= \sum_{\ell=0}^{L} p_{L-\ell}^M C_\ell^{\mathbf{x}} \eqsim \sum_{\ell=0}^{L} p_{L-\ell}^M m^{\ell d} \lesssim \sum_{\ell=0}^{L} \left(m^{\frac{-\ell d}{r+M}} g(m, L)^{\frac{1}{r}} \epsilon^{-\frac{1}{r}} + 1 \right)^M m^{\ell d} \\ &\lesssim \left(\sum_{\ell=0}^{L} m^{\frac{\ell d}{M}} \left(m^{\frac{-\ell d}{r+M}} g(m, L)^{\frac{1}{r}} \epsilon^{\frac{-1}{r}} + 1 \right) \right)^M \\ &\eqsim \left(g(m, L)^{\frac{1}{r}} \epsilon^{-\frac{1}{r}} \sum_{\ell=0}^{L} m^{\frac{\ell d r}{M(r+M)}} + \sum_{\ell=0}^{L} m^{\frac{\ell d}{M}} \right)^M. \end{aligned}$$

Now we consider bounding $\sum_{\ell=0}^{L} m^{c\ell}$ for a constant $c > 0$ where $L \leq \frac{1}{s} \log_m(\frac{2C_D}{\epsilon}) + 1$ from (3.6),

$$\sum_{\ell=0}^{L} m^{c\ell} = \frac{m^{c(L+1)} - 1}{m^c - 1} < \frac{m^{c(L+1)}}{m^c - 1} \leq \frac{(2C_D)^{c/s} \, m^{2c}}{m^c - 1} \epsilon^{\frac{-c}{s}}. \tag{3.8}$$

Then we have the following bounds in terms of ϵ:

$$\sum_{\ell=0}^{L} m^{\frac{\ell d}{M}} \lesssim \epsilon^{\frac{-d}{sM}}, \sum_{\ell=0}^{L} m^{\frac{\ell d r}{M(r+M)}} \lesssim \epsilon^{\frac{-dr}{sM(r+M)}}, \text{ and } g(m, L) \lesssim \epsilon^{\frac{-dr}{s(r+M)}}. \tag{3.9}$$

Therefore using (3.9), we have

$$\begin{aligned} \mathrm{COST}_\epsilon^{ML} &\lesssim \left(\epsilon^{\frac{-1}{r}} \epsilon^{\frac{-d}{s(r+M)}} \epsilon^{\frac{-dr}{sM(r+M)}} + \epsilon^{\frac{-d}{sM}} \right)^M \\ &\eqsim \left(\epsilon^{\frac{-1}{r} - \frac{-d}{sM}} \right)^M \eqsim \epsilon^{\frac{-M}{r} - \frac{d}{s}}. \end{aligned} \tag{3.10}$$

□

Now, we consider the standard single-level SG approximation to construct a similar cost metric for a given accuracy. Under the same assumptions on the spatial and stochastic approximation error, the error for the $h_{SL} \times p_{SL}$ SG approximation can be bounded by

$$\left\| u - u_{p_{SL}}^{h_{SL}} \right\|_W \leq C_D h_{SL}^s + C_\Gamma p_{SL}^{-r}.$$

Then for a total error of ϵ, we require each contribution to be of order $\epsilon/2$. Thus, we choose $h_{SL} \eqsim \epsilon^{1/s}$ and $p_{SL} \eqsim \epsilon^{-1/r}$. The computational cost to achieve a total error of ϵ is then

$$\text{COST}_\epsilon^{SL} \eqsim p_{SL}^M h_{SL}^{-d} \eqsim \epsilon^{\frac{-M}{r} - \frac{d}{s}}. \tag{3.11}$$

This is the same computational cost derived in Theorem 3.3. In practice, as demonstrated in Section 4, we find that the multilevel method offers computational savings compared to the single-level approach, even though this is not apparent in the theoretical cost bound. This is due, in part, to several known factors. First, the stochastic approximation error using truncated polynomial chaos expansions offers exponential convergence, under certain conditions [3, 25]. Thus, we are over-estimating the stochastic discretization error, leading to higher than necessary polynomial orders being specified in the formula. Also, we are making an initial weak assumption on the cost metric, which leads to a pessimistic bound on the actual computational cost.

4. Numerical Results. In this section, numerical experiments are reported comparing the multilevel approach with the standard, single-level approach. We consider the stochastic diffusion equation (2.1) with $D = [0,1]^3$, $f(\mathbf{x}) = 1$, and boundary conditions $u(\mathbf{x}, \omega) = g(\mathbf{x})$ on $\partial D \times \Gamma$ where

$$g(\mathbf{x}) = \begin{cases} 1 & [0,1] \times [0,1] \times \{0\} \\ 2 & [0,1] \times [0,1] \times \{1\} \\ 0 & \text{otherwise.} \end{cases}$$

The diffusion coefficient $a(\mathbf{x}, \mathbf{y})$ is modeled with a truncated Karhunen-Loève expansion with mean $\mu = 1$ and standard deviation $\sigma = 0.2$ with an underlying exponential covariance function with correlation length $c = 1/5$. We assume the $M = 3$ random variables are uniform over $[-1, 1]$ so the Legendre multivariate polynomials are used as the stochastic basis functions. The finite element method is used to spatially discretize the problem using tri-linear finite element basis functions on a hexahedral mesh of D.

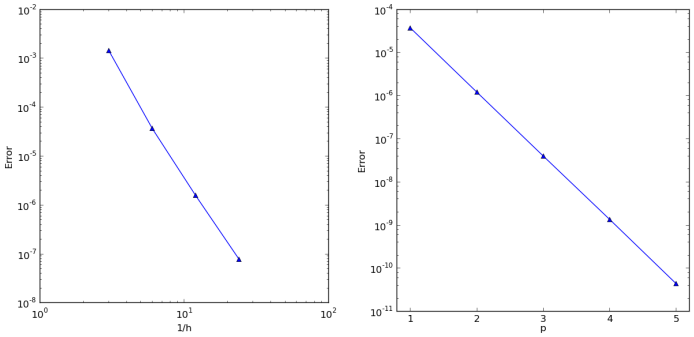
The multilevel method implementation proceeds as follows: for each level, the problem is discretized and the resulting SG linear system is solved using the preconditioned conjugate gradient (PCG) method with mean-based preconditioning [10, 16, 20]. The iterative method is terminated when the relative residual error is reduced to 10^{-8}. The error of the approximate multilevel solution is computed with respect to a reference solution $u_{ref} = u_{\hat{p}}^{\hat{h}}$, computed with a fine spatial mesh width $\hat{h} = 1/48$ and high order stochastic basis functions of order $\hat{p} = 6$.

The Trilinos framework [12] is used to implement and numerically test the multilevel method. The Stokhos package offers tools for embedded uncertainty quantification methods and is used to assemble the algebraic equations for the SG method [19]. In order to compute the polynomial chaos coefficients, an approach similar to automatic differentiation that incorporates template-based generic programming and operator overloading is used, see [17, 18] for further details. We use the CG iterative solver provided by the Belos package with a mean-based preconditioner implemented using algebraic multigrid from the MueLu package [8].

We now present an experiment using the multilevel method for approximating the solution of (2.1) where the spatial discretization is refined by a factor of two in each spatial dimension starting with an initial mesh width of $h_0 = 1/5$. The reported error is measured by $\|\cdot\|_*$ which is defined for a gPC expansion $v(\mathbf{x}, \mathbf{y}) = \sum_{i=0}^{Q-1} v_i(\mathbf{x})\Psi_i(\mathbf{y})$ as,

$$\|v\|_* := \left(\sum_{i=0}^{Q-1} \|v_i(\mathbf{x})\|_{L^2(D)} \right)^{1/2}.$$

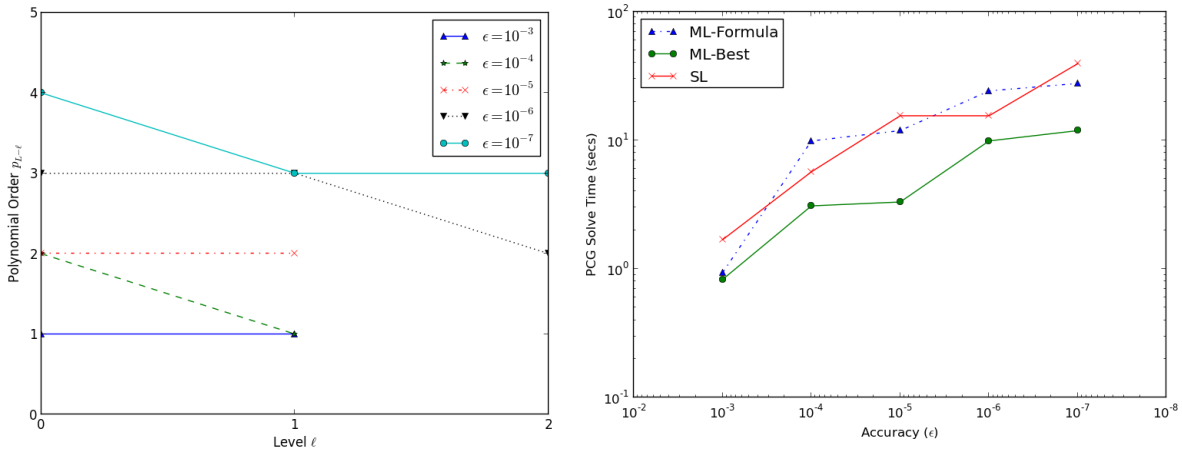
Figure 4.1a shows the finite element error, and confirms Assumption 2.4 holds with approximate values $C_D = 0.002$ and $s = 7$. The stochastic approximation error, shown in Figure 4.1b, satisfies Assumption 2.6 with approximate values $C_\Gamma = 1.2 \times 10^{-4}$ and $r = 9$. Using these computed constants, Figure 4.2a shows the polynomial orders for each level, $p_{L-\ell}$, determined from (3.7) with number of levels computed from (3.6). Figure 4.2b shows the total iterative solve time, measuring the computational cost, of the multilevel method and standard single-level SG method. In the plot, the data labeled ML-Formula is the result of



(a) Spatial Finite Element Error (b) Stochastic Approximation Error

Fig. 4.1: Left: The spatial finite element approximation error $\left\|u_6^{1/48} - u_6^h\right\|_*$ versus $1/h$. Right: The stochastic approximation error $\left\|u_6^{1/48} - u_p^{1/48}\right\|_*$ versus p.

the multilevel method with the polynomial orders and mesh sizes determined using the formulas derived in Theorem 3.3. For data labeled ML-Best, the parameters are chosen manually to achieve an accuracy of ϵ while minimizing the computational cost. The data labeled SL is the result of the standard single-level method with hand-tuned parameters chosen to achieve an accuracy of ϵ. For this problem formulation, for all desired accuracy levels tested ML-Best, using hand-tuned polynomial orders, outperforms the standard single level method in overall solve time. The multilevel method using polynomial orders dictated by the formula is competitive with the standard single-level method.



(a) Formula Derived Polynomial Orders (b) Total Solve Time

Fig. 4.2: Left: Polynomial order as determined by (3.7) at each level L of the multilevel method for varying accuracies ϵ. Right: The total iterative solve time versus the level of accuracy ϵ for the proposed multilevel method and standard single-level method. The multilevel formulation using hand-tuned polynomial orders, ML-Best, offers computational savings for all desired accuracy levels tested. ML-Formula, using the polynomial orders dictated by the formula, is competitive with the standard, single-level method in overall solve time.

5. Conclusions. The stochastic Galerkin method is a powerful method to obtain highly accurate approximate solutions to PDEs with random input coefficients, yet it is expensive, especially as the stochastic dimension increases. We propose and analyze a multilevel solution strategy based on the idea that the solution can be written as a telescoping sum of a hierarchy of spatial approximations. The ability to use lower-order stochastic approximations within the sum allows for computational savings. Our numerical results show that this method can achieve considerable computational savings while maintaining overall accuracy in the solution.

In the current implementation, the accuracy of the computed solution is determined by comparing it to a reference solution computed with a fine mesh and high polynomial order, however in practice, this comparison is infeasible. A more practical implementation method is needed to determine when the multilevel solution has achieved the desired accuracy. Also, the bounds used for the stochastic approximation error and the computational cost metric are weak, and result in a bound on the computational cost of the multilevel method that is not sharp. In particular, the computational cost is currently an estimate on the number of unknowns in the system. A more accurate cost estimate would include a measurement of the number of non-zeros in the linear operator which is a better estimator of the computational effort.

References.

[1] C. Audouze and P. Nair, *Some a priori error estimates for finite element approximations of elliptic and parabolic linear stochastic partial differential equations*, International Journal for Uncertainty Quantification, 4 (2014).

[2] I. Babuška, F. Nobile, and R. Tempone, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM Journal on Numerical Analysis, 45 (2007), pp. 1005–1034.

[3] I. Babuška, R. Tempone, and G. Zouraris, *Galerkin finite element approximations of stochastic elliptic partial differential equations*, SIAM Journal on Numerical Analysis, 42 (2004), pp. 800–825.

[4] A. Barth, C. Schwab, and N. Zollinger, *Multi-level monte carlo finite element method for elliptic pdes with stochastic coefficients*, Numerische Mathematik, 119 (2011), pp. 123–161.

[5] J. Charrier, R. Scheichl, and A. L. Teckentrup, *Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods*, SIAM Journal on Numerical Analysis, 51 (2013), pp. 322–352.

[6] K. Cliffe, M. Giles, R. Scheichl, and A. L. Teckentrup, *Multilevel monte carlo methods and applications to elliptic pdes with random coefficients*, Computing and Visualization in Science, 14 (2011), pp. 3–15.

[7] G. Fishman, *Monte Carlo*, Springer Series in Operations Research, Springer-Verlag, New York, 1996. Concepts, algorithms, and applications.

[8] J. Gaidamour, J. Hu, C. Siefert, and R. Tuminaro, *Design considerations for a flexible multigrid preconditioning library*, Scientific Programming, 20 (2012), pp. 223–239.

[9] R. Ghanem and P. D. Spanos, *Polynomial chaos in stochastic finite elements*, Journal of Applied Mechanics, 57 (1990).

[10] R. G. Ghanem and R. M. Kruger, *Numerical solution of spectral stochastic finite element systems*, Computer Methods in Applied Mechanics and Engineering, 129 (1996), pp. 289 – 303.

[11] R. G. Ghanem and P. D. Spanos, *Stochastic finite elements: a spectral approach*, Springer-Verlag New York, 1991.

[12] M. Heroux, R. Bartlett, V. Howle, R. Hoekstra, J. Hu, T. Kolda, R. Lehoucq, K. Long, R. Pawlowski, E. Phipps, A. Salinger, H. Thornquist, R. Tuminaro, J. Willenbring, A. Williams, and K. Stanley, *An overview of the Trilinos package*, ACM Trans. Math. Softw., 31 (2005). `http://trilinos.sandia.gov/`.

[13] M. Loeve, *Probability theory, vol. ii*, Graduate texts in mathematics, 46 (1978), pp. 0–387.

[14] G. J. Lord, C. E. Powell, and T. Shardlow, *An Introduction to Computational Stochastic PDEs*, Cambridge Texts in Applied Mathematics, Cambridge University Press, 2014.

[15] N. Metropolis and S. Ulam, *The Monte Carlo method*, Journal of the American Statistical Association, 44 (1949), pp. 335–341.

[16] M. Pellissetti and R. Ghanem, *Iterative solution of systems of linear equations arising in the context of stochastic finite elements*, Advances in engineering software, 31 (2000), pp. 607–616.

[17] E. Phipps, H. C. Edwards, J. Hu, and J. T. Ostien, *Exploring emerging manycore architectures for uncertainty quantification through embedded stochastic Galerkin methods*, International Journal of Computer Mathematics, 91 (2014), pp. 707–729.

[18] ———, *Exploring heterogeneous multicore architectures for advanced embedded uncertainty quantification*, Tech. Rep. SAND2014-17875, Sandia National Laboratories, October 2014.

[19] E. T. Phipps, *Stokhos Stochastic Galerkin Uncertainty Quantification Methods.* `http://trilinos.sandia.gov/packages/stokhos/`, 2011.

[20] C. E. Powell and H. Elman, *Block-diagonal preconditioning for spectral stochastic finite-element systems*, IMA Journal of Numerical Analysis, 29 (2009), pp. 350–375.

[21] A. L. Teckentrup, *Multilevel Monte Carlo methods for highly heterogeneous media*, in proceedings of the Winter Simulation Conference, Winter Simulation Conference, 2012, p. 32.

[22] A. L. Teckentrup, P. Jantsch, C. G. Webster, and M. Gunzburger, *A multilevel stochastic collocation method for partial differential equations with random input data*, arXiv preprint arXiv:1404.2647, (2014).

[23] N. Wiener, *The homogeneous chaos*, Am J Math, 60 (1938), pp. 897–936.

[24] D. Xiu and J. Hesthaven, *High-order collocation methods for differential equations with random inputs*, SIAM Journal on Scientific Computing, 27 (2005), pp. 1118–1139.

[25] D. Xiu and G. Karniadakis, *The wiener-askey polynomial chaos for stochastic differential equations*, SIAM Journal on Scientific Computing, 24 (2002), pp. 619–644.