

Solving the Stochastic Steady-State Diffusion Problem using Multigrid¹

Howard Elman² and Darran Furnival³

February 20, 2006

ABSTRACT. We study multigrid for solving the stochastic steady-state diffusion problem. We operate under the mild assumption that the diffusion coefficient takes the form of a finite Karhunen-Loève expansion. The problem is discretized using a finite element methodology using the polynomial chaos method to discretize the stochastic part of the problem. We apply a multigrid algorithm to the stochastic problem in which the spatial discretization is varied from grid to grid while the stochastic discretization is held constant. We then show, theoretically and experimentally, that the convergence rate is independent of the spatial discretization, as in the deterministic case.

1. INTRODUCTION

Mathematical models often contain partial differential equations (PDEs). The constituent parts of the PDE, i.e. the differential coefficients and the source function, are most easily modeled as functions of the spatial domain. However, uncertainty might exist as to the most appropriate functions to use in the model. A more sophisticated model might therefore represent the differential coefficients and source function not only as functions on the spatial domain but also as functions on some sample space, i.e. as random fields. This gives rise to stochastic partial differential equations (SPDEs).

In this paper we consider the stochastic steady-state diffusion equation along with homogeneous Dirichlet boundary value conditions. We are interested in the case when the diffusion coefficient is stochastic and the source function is deterministic, i.e. the diffusion coefficient is a random field and the source function is defined on the spatial domain only. However, we also treat the source function as a random field as this is required for purposes of analysis and incorporates, as a special case, the fact that the source function may be deterministic.

We will assume the diffusion coefficient to be of the form of a finite Karhunen-Loève expansion. This is common in the literature, e.g. see Babuška, Tempone & Zouraris (2004), Ghanem & Spanos (1991), and Xiu & Karniadakis (2002). We require the diffusion coefficient to be of this form for analytic as well as computational purposes.

We are interested in using a finite element methodology to find an approximate solution to the problem. We therefore obtain a weak formulation to the boundary value problem and proceed to look in a finite-dimensional subspace of the infinite-dimensional space that contains the weak solution in order to obtain a matrix problem.

¹This work was supported in part by the U.S. National Science Foundation under grant DMS0208015 and by the U.S. Department of Energy under grant DOEG0204ER25619.

²Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742. elman@cs.umd.edu.

³Applied Mathematics and Scientific Computing Program, University of Maryland, College Park, MD 20742. furnival@math.umd.edu

The finite-dimensional subspace in which we look for the approximate solution will be a tensor product of a space of functions defined on the spatial domain and a space of functions defined on the sample space. For the finite-dimensional space of functions on the spatial domain we will choose the set of piecewise linear polynomials defined on a triangulation. For the finite-dimensional space of functions defined on the sample space we will choose the set of polynomials that are of degree no greater than a chosen parameter.

For computational purposes the basis of the space of functions defined on the sample space will correspond to that obtained by employing the polynomial chaos method as developed by Ghanem & Spanos (1991) and generalized by Xiu & Karniadakis (2002). Indeed, this method provided the foundation on which this work was developed. However, this is not the only basis which is possible. Another choice would be the doubly orthogonal polynomials as discussed in Babuška et al. (2004). Although we will not compute with the doubly orthogonal polynomials they will be vital to our analytic argument.

Theoretically, we apply a two-grid correction scheme to solve the matrix problem. In this scheme the spatial discretization is varied from grid to grid, giving a fine grid on the upper level and a coarse grid on the lower level, while the stochastic discretization is kept constant. We show that the convergence factor of this method is independent of the spatial mesh parameter. An induction argument can then be used to show that the multigrid algorithm obtained by applying the two-grid correction scheme recursively has a convergence factor that is also independent of the spatial mesh parameter. We do not give this inductive argument but note that the reasoning would be the same as that for the analogous deterministic problem which is discussed e.g. in Braess (2001) and Elman, Silvester & Wathen (2005).

Experimentally, we consider two problems. These are obtained by defining the diffusion coefficient to be a Karhunen-Loève expansion consisting of random variables that are for the first problem uniformly distributed and for the second problem normally distributed. In both these problems the source function is set to unity. We then apply multigrid to a selection of matrix problems associated with different discretization parameters, both spatial and stochastic, and tabulate the number of iterates required for convergence. It will be seen from these tables that the number of iterates it requires multigrid to converge remains, to within a close approximation, constant as the spatial parameter is varied. This provides experimental evidence that the convergence factor associated with the applied multigrid algorithm is indeed independent of the spatial mesh parameter.

2. THE STOCHASTIC STEADY-STATE DIFFUSION PROBLEM

In this section we introduce the stochastic steady-state diffusion problem along with its weak formulation and finite element discretization. We then obtain a number of properties associated with the discretized system. These properties are analogous to the properties of the system of equations resulting from the deterministic steady-state diffusion problem and are proven similarly. Finally, we introduce the doubly orthogonal polynomials which can be used as a basis for the stochastic part of the problem. The doubly orthogonal polynomials will be vital to our theoretical arguments concerning the convergence of multigrid in §3.

2.1. Boundary Value Problem. The stochastic steady-state diffusion equation with homogeneous Dirichlet boundary value conditions is given by

$$(1) \quad \begin{cases} -\nabla \cdot (c\nabla u) = f & \text{in } D \times \Omega, \\ u = 0 & \text{on } \partial D \times \Omega, \end{cases}$$

where D is the spatial domain, Ω is a sample space, $c: D \times \Omega \rightarrow \mathbb{R}$ is the diffusion coefficient, and $f: D \times \Omega \rightarrow \mathbb{R}$ is the source function. The sample space in turn belongs to a probability space (Ω, \mathcal{F}, P) where \mathcal{F} is a σ -algebra and P is a probability measure.⁴

The spatial domain, D , is assumed to be a two-dimensional simply connected bounded open set with piecewise smooth boundary. In particular we take D to be the interior of a polygon.

We will let $\Omega = \Gamma^m$ where $\Gamma = (a, b)$. We will assume the diffusion coefficient to be of the form

$$(2) \quad c(x, \omega) = c_0(x) + \sum_{r=1}^m \sqrt{\lambda_r} c_r(x) \xi_r(\omega)$$

where $\xi_r: \Omega \rightarrow \mathbb{R}$ are identically distributed independent random variables with zero mean and for $\omega = (\omega_1, \dots, \omega_m) \in \Omega$, $\xi_r(\omega) = \omega_r$. Note that the distribution of $\xi = (\xi_1, \dots, \xi_m)$ will dictate the probability measure to be used. For example, if ξ_r is uniformly distributed on $(-1, 1)$ then P will be the probability measure associated with an m -dimensional uniform distribution.

We necessarily expect the solution to be a random field, $u: \overline{D} \times \Omega \rightarrow \mathbb{R}$, such that for each value of $\omega \in \Omega$ the resulting PDE is satisfied in the classical sense.

We note that this problem is extensively discussed from a modeling perspective in Ghanem & Spanos (1991) and from an analytic perspective in Babuška et al. (2004).

2.2. Weak Formulation. In stating the weak formulation of (1) we will use tensor products of Hilbert spaces which are defined and discussed in Babuška et al. (2004) and Treves (1967). Let $c \in L_\infty(D) \otimes L_\infty(\Omega)$ and $f \in L_2(D) \otimes L_2(\Omega)$. The weak formulation of (1) is given by: find $u \in H_0^1(D) \otimes L_2(\Omega)$ such that

$$(3) \quad a(u, v) = l(v) \quad \forall v \in H_0^1(D) \otimes L_2(\Omega)$$

where⁵

$$(4) \quad a(u, v) = \int_{\Omega} \int_D c \nabla u \cdot \nabla v,$$

$$(5) \quad l(v) = \int_{\Omega} \int_D f v.$$

The *Lax-Milgram lemma* can be used to show that there exists a unique solution to this problem providing that there exist positive constants α and β such that

$$(6) \quad \alpha \leq c(x, \omega) \leq \beta \quad P\text{-a.e. } \forall x \in D.$$

2.3. Finite Element Formulation. We are interested in applying a finite element methodology to find an approximation to the solution of the variational problem given in §2.2. This entails a discretization of both the spatial and stochastic parts of the problem.

The spatial domain is discretized using a triangulation $\mathcal{T} = \{\Delta_1, \dots, \Delta_K\}$. Denoting the longest side of the t -th triangle in the triangulation as h_t , we define the mesh parameters $h = \max h_t$ and $\underline{h} = \min h_t$. We also denote the smallest angle in the t -th triangle as θ_t . We assume that any triangulation used belongs to a family of triangulations that is quasi-uniform, i.e. there exists $\rho > 0$

⁴Note that the nabla operator only operates on the spatial components of the function, that is to say, that if $f: D \times \Omega \rightarrow \mathbb{R}$ and $x = (x_1, x_2) \in D$, $\omega \in \Omega$, then

$$\nabla f(x, \omega) = \left(\frac{\partial f(x, \omega)}{\partial x_1}, \frac{\partial f(x, \omega)}{\partial x_2} \right).$$

⁵Note that the integral over Ω is with respect to the probability measure, i.e.

$$\int_{\Omega} \int_D = \int_{\Omega} \int_D dx dP.$$

such that $\underline{h} > \rho h$ for every triangulation in the family, and shape regular, i.e. there exists $\theta_* > 0$ such that $\theta_* \leq \min \theta_t$ for every triangulation in the family. The finite-dimensional subspace of $H_0^1(D)$ is then taken to be $S = \text{span}\{\phi_1, \dots, \phi_N\}$ where $\phi_k: \overline{D} \rightarrow \mathbb{R}$, $k = 1, \dots, N$, are the usual piecewise linear basis functions defined at the nodes of \mathcal{T} . (Here N is the number of internal nodes in the triangulation.)

For the subspace of $L_2(\Omega)$ we take the space of all $v \in L_2(\Omega)$ such that v is a polynomial of degree at most p . We denote this space by T . For the purposes of analysis any basis of T is sufficient. We will use the notation $T = \text{span}\{\psi_1, \dots, \psi_M\}$ where $M = M(m, p)$ and is in fact given by

$$(7) \quad M = \frac{(m+p)!}{m!p!}.$$

The basis with which we choose to compute in §4 is that derived from the generalized polynomial chaos method as discussed in Xiu & Karniadakis (2002). In this method the functions chosen to be a basis for T are those polynomials from the Askey scheme of hypergeometric polynomials that satisfy

$$(8) \quad \int_{\Omega} \psi_k \psi_l = d_k \delta_{kl}.$$

For example, if ξ_r , $r = 1, \dots, m$, are uniformly distributed on $(-1, 1)$ then the basis of T would be the set of m -dimensional Legendre polynomials of degree at most p .

We thus have $S \otimes T \subset H_0^1(D) \otimes L_2(\Omega)$ which leads to the finite element formulation: find $u_{hp} \in S \otimes T$ such that

$$(9) \quad a(u_{hp}, v) = l(v) \quad \forall v \in S \otimes T$$

where $a(\cdot, \cdot)$ and $l(\cdot)$ are as in (4) and (5). This will possess a unique solution under the same conditions that apply to the weak formulation.

2.4. Matrix Formulation. By substituting the expansion

$$(10) \quad u_{hp} = \sum_{j=1}^N \sum_{l=1}^M u_{jl} \phi_j \psi_l$$

into (9) and varying v over the basis functions of $S \otimes T$ we find that we can obtain the finite element approximation by solving the matrix problem: find $\mathbf{u} \in \mathbb{R}^{MN}$ such that

$$(11) \quad \mathbf{A}\mathbf{u} = \mathbf{f}$$

where

$$(12) \quad A = \begin{bmatrix} A_{11} & \cdots & A_{1M} \\ \vdots & & \vdots \\ A_{M1} & \cdots & A_{MM} \end{bmatrix}, \quad [A_{kl}]_{ij} = \int_{\Omega} \int_D c \nabla \phi_i \cdot \nabla \phi_j \psi_k \psi_l,$$

and

$$(13) \quad \mathbf{f} = \begin{bmatrix} \mathbf{f}_1 \\ \vdots \\ \mathbf{f}_M \end{bmatrix}, \quad [\mathbf{f}_k]_i = \int_{\Omega} \int_D f \phi_i \psi_k,$$

with $\mathbf{u} = [u_{11}, \dots, u_{N1}, \dots, u_{1M}, \dots, u_{NM}]^T$.

Once we have computed \mathbf{u} then we have u_{hp} . From this we can calculate such things as the mean, variance, and covariance of the approximation. Formulas for these quantities in terms of the coefficients of (10) are readily obtainable once a basis for T has been chosen.

2.5. Matrix and Right Hand Side Properties. We here establish some results concerning the system matrix, A , and the right hand side vector, \mathbf{f} , that will be required for the analysis of multigrad in §3.

In the following, E refers to the stochastic mass matrix and B refers to the (deterministic) mass matrix, which are defined by

$$(14) \quad [E]_{kl} = \int_{\Omega} \psi_k \psi_l, \quad [B]_{ij} = \int_D \phi_i \phi_j,$$

respectively. By $E \otimes B$ we will mean the matrix Kronecker product as given by

$$(15) \quad E \otimes B = \begin{bmatrix} e_{11}B & \cdots & e_{1M}B \\ \vdots & & \vdots \\ e_{M1}B & \cdots & e_{MM}B \end{bmatrix}.$$

We now introduce some notation for the coefficient vector of a function in $S \otimes T$. Let $v \in S \otimes T$, then we have the expansion

$$(16) \quad v = \sum_{i=1}^N \sum_{k=1}^M v_{ik} \phi_i \psi_k.$$

We define the coefficient vector $\mathbf{v} \in \mathbb{R}^{MN}$ of v by

$$(17) \quad \mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_M \end{bmatrix}, \quad [\mathbf{v}_k]_i = v_{ik}.$$

Theorem 1. *Let $f \in S \otimes T$ with coefficient vector $\hat{\mathbf{f}}$, then $\mathbf{f} = (E \otimes B)\hat{\mathbf{f}}$, where \mathbf{f} is as in (13).*

Proof. This follows upon substituting the expansion of f into (13). □

Theorem 2. *Let $f \in S \otimes T$ with coefficient vector $\hat{\mathbf{f}}$, then $\|f\|_{L_2(D) \otimes L_2(\Omega)}^2 = ((E \otimes B)\hat{\mathbf{f}}, \hat{\mathbf{f}})$.*

Proof. This follows upon substituting the expansion of f into $\|\cdot\|_{L_2(D) \otimes L_2(\Omega)}$. □

Theorem 3. *The inequality*

$$(18) \quad C_1 h^2 \leq \frac{((E \otimes B)\mathbf{v}, \mathbf{v})}{(\mathbf{v}, \mathbf{v})} \leq C_2 h^2$$

holds for all $\mathbf{v} \in \mathbb{R}^{MN}$, where $C_1 = C_1(p)$ and $C_2 = C_2(p)$.

Proof. Let $\mathbf{v} \in \mathbb{R}^{MN}$ be the coefficient vector of some $v \in S \otimes T$. Then,

$$(19) \quad \begin{aligned} ((E \otimes B)\mathbf{v}, \mathbf{v}) &= \sum_{i=1}^N \sum_{k=1}^M [\mathbf{v}_k]_i \sum_{j=1}^N \sum_{l=1}^M [E]_{kl} [B]_{ij} [\mathbf{v}_l]_j \\ &= \sum_{i=1}^N \sum_{k=1}^M v_{ik} \sum_{j=1}^N \sum_{l=1}^M v_{jl} \int_{\Omega} \psi_k \psi_l \int_D \phi_i \phi_j \\ &= \sum_{t=1}^K \sum_{i=1}^N \sum_{k=1}^M \sum_{j=1}^N \sum_{l=1}^M v_{ik} v_{jl} \int_{\Omega} \psi_k \psi_l \int_{\Delta_t} \phi_i \phi_j. \end{aligned}$$

Let B_t be the local mass matrix and $\mathbf{v}_t \in \mathbb{R}^{3M}$ be a vector that contains only those values of \mathbf{v} that are associated with the t -th triangle in the mesh, viz.,

$$(20) \quad \mathbf{v}_t = \begin{bmatrix} \mathbf{v}_1^{(t)} \\ \vdots \\ \mathbf{v}_M^{(t)} \end{bmatrix}, \quad [\mathbf{v}_k^{(t)}]_i = v_{ik}^{(t)}$$

where $\mathbf{v}_k^{(t)} \in \mathbb{R}^3$ contains the values of \mathbf{v} at those vertices of the t -th triangle associated with the k -th basis function of T . Let B_* be the mass matrix on the canonical triangle (with vertices $(0,0)$, $(1,0)$, $(0,1)$). Then

$$(21) \quad \begin{aligned} ((E \otimes B)\mathbf{v}, \mathbf{v}) &= \sum_{t=1}^K \sum_{i=1}^3 \sum_{k=1}^M \sum_{j=1}^3 \sum_{l=1}^M v_{ik}^{(t)} v_{jl}^{(t)} \int_{\Omega} \psi_k \psi_l \int_{\Delta_t} \phi_i^{(t)} \phi_j^{(t)} \\ &= \sum_{t=1}^K 2|\Delta_t| \sum_{i=1}^3 \sum_{k=1}^M [\mathbf{v}_k^{(t)}]_i \sum_{j=1}^3 \sum_{l=1}^M [E]_{kl} [B_*]_{ij} [\mathbf{v}_l^{(t)}]_j \\ &= \sum_{t=1}^K 2|\Delta_t| ((E \otimes B_*)\mathbf{v}_t, \mathbf{v}_t) \end{aligned}$$

where $|\Delta_t|$ is the area of the t -th triangle.

For any triangle with longest side h_t and smallest angle θ_t we have the inequality

$$(22) \quad \frac{1}{4} h_t^2 \sin \theta_t \leq |\Delta_t| \leq \frac{1}{2} h_t^2 \sin \theta_t,$$

and given that $E \otimes B_*$ is a symmetric positive definite matrix there exists $\gamma = \gamma(p)$ and $\delta = \delta(p)$ such that

$$(23) \quad \gamma(\mathbf{v}_t, \mathbf{v}_t) \leq ((E \otimes B_*)\mathbf{v}_t, \mathbf{v}_t) \leq \delta(\mathbf{v}_t, \mathbf{v}_t)$$

for all \mathbf{v}_t .

Using (22) and (23) along with the assumptions of quasi-uniformity and shape regularity, and continuing with (21) we have

$$(24) \quad \begin{aligned} (E \otimes B\mathbf{v}, \mathbf{v}) &\geq \sum_{t=1}^K \frac{1}{2} \gamma h_t^2 \sin \theta_t (\mathbf{v}_t, \mathbf{v}_t) \\ &\geq \frac{1}{2} \gamma \rho^2 h^2 \sin \theta_* \sum_{t=1}^K (\mathbf{v}_t, \mathbf{v}_t) \\ &\geq \frac{1}{2} \gamma \rho^2 h^2 \sin \theta_* (\mathbf{v}, \mathbf{v}) \end{aligned}$$

and noting that there exists $\eta > 0$ such that $\sum_{t=1}^K (\mathbf{v}_t, \mathbf{v}_t) \leq \eta(\mathbf{v}, \mathbf{v})$, we also have

$$(25) \quad \begin{aligned} (E \otimes B\mathbf{v}, \mathbf{v}) &\leq \sum_{t=1}^K \delta h_t^2 \sin \theta_t (\mathbf{v}_t, \mathbf{v}_t) \\ &\leq \delta h^2 \sum_{t=1}^K (\mathbf{v}_t, \mathbf{v}_t) \\ &\leq \delta h^2 \eta (\mathbf{v}, \mathbf{v}). \end{aligned}$$

The inequalities (24) and (25) prove the theorem. \square

Theorem 4. Let $f \in S \otimes T$. Then $h\sqrt{C_1}\|f\|_{L_2(D) \otimes L_2(\Omega)} \leq \|f\|_2$, where C_1 is as in Theorem 3.

Proof. Using Theorem 1 we have $\|f\|_2^2 = ((E \otimes B)\hat{\mathbf{f}}, (E \otimes B)\hat{\mathbf{f}})$. Now setting $\mathbf{g} = (E \otimes B)^{\frac{1}{2}}\hat{\mathbf{f}}$ and using Theorems 2 and 3 we have

$$(26) \quad C_1 h^2 \leq \frac{((E \otimes B)\mathbf{g}, \mathbf{g})}{(\mathbf{g}, \mathbf{g})} = \frac{((E \otimes B)\hat{\mathbf{f}}, (E \otimes B)\hat{\mathbf{f}})}{((E \otimes B)\hat{\mathbf{f}}, \hat{\mathbf{f}})} = \frac{\|f\|_2^2}{\|f\|_{L_2(D) \otimes L_2(\Omega)}^2}$$

as required. \square

Theorem 5. *The inequality*

$$(27) \quad C_3 h^2 \leq \frac{(A\mathbf{v}, \mathbf{v})}{(\mathbf{v}, \mathbf{v})} \leq C_4$$

holds for all $\mathbf{v} \in \mathbb{R}^{MN}$, where $C_3 = C_3(p)$ and $C_4 = C_4(p)$.

Proof. We omit the proof. Suffice to say that it follows the proof for the analogous deterministic problem as given in Elman et al. (2005). \square

Note that Theorem 5 implies that the maximum eigenvalue of A is bounded above by a constant, viz., C_4 .

2.6. Doubly Orthogonal Polynomials. We here establish two results concerning a basis for T comprised of the so-called doubly orthogonal polynomials. The doubly orthogonal polynomials are discussed in Babuška et al. (2004). They are the set $\{\Psi_1, \dots, \Psi_M\}$ that satisfy

$$(28) \quad \int_{\Omega} \Psi_k \Psi_l = \delta_{kl}, \quad \int_{\Omega} \xi_r \Psi_k \Psi_l = \kappa_{rk} \delta_{kl},$$

where δ_{kl} is the Kronecker delta function.

An important consequence of the doubly orthogonal polynomials is that when they are used as a basis for T the system matrix in §2.4 becomes block diagonal. This allows us to think of the discrete problem as M decoupled (deterministic) steady-state diffusion problems. As this is an important point that will be used in the proof of the approximation property in §3.4 we state it explicitly in the following theorem. This is then followed by a theorem concerning the source functions of these decoupled problems that will also be used in proving the approximation property.

Theorem 6. *If the finite element approximation u_{hp} is expanded using the doubly orthogonal polynomials, viz.,*

$$(29) \quad u_{hp} = \sum_{k=1}^M u_k \Psi_k,$$

then each $u_k \in S$ is the finite element approximation to a steady-state diffusion problem.

Proof. If $\{\Psi_1, \dots, \Psi_M\}$ is used for the computational basis then the system matrix in §2.4 will be of the form

$$(30) \quad A = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_M \end{bmatrix}, \quad [A_k]_{ij} = \int_D b_k \nabla \phi_i \cdot \nabla \phi_j,$$

where

$$(31) \quad b_k = c_0 + \sum_{r=1}^m \sqrt{\lambda_r} \kappa_{rk} c_r$$

where c_r, κ_{rk} , $r = 1, \dots, m$ are from (2) and (28) respectively. This proves the theorem. \square

Theorem 7. Let $f \in S \otimes T$ and consider the expansion of the finite element approximation u_{hp} in terms of the doubly orthogonal polynomials as given in (29). Then the source function f_k of the steady-state diffusion problem to which u_k is the solution satisfies

$$(32) \quad \|f_k\|_{L_2(D)} \leq \|f\|_{L_2(D) \otimes L_2(\Omega)}.$$

Proof. As $f \in S \otimes T$ we can expand it in the doubly orthogonal polynomials as

$$(33) \quad f = \sum_{l=1}^M \hat{f}_l \Psi_l.$$

Using $\{\Psi_1, \dots, \Psi_M\}$ as the computational basis and substituting (33) into (13) we have

$$(34) \quad [\mathbf{f}_k]_i = \int_{\Omega} \int_D \left(\sum_{l=1}^M \hat{f}_l \Psi_l \right) \phi_i \Psi_k = \sum_{l=1}^M \int_{\Omega} \Psi_l \Psi_k \int_D \hat{f}_l \phi_i = \int_D \hat{f}_k \phi_i.$$

This shows that $f_k = \hat{f}_k$. Now

$$(35) \quad \begin{aligned} \|f\|_{L_2(D) \otimes L_2(\Omega)}^2 &= \int_{\Omega} \int_D \left(\sum_{l=1}^M \hat{f}_l \Psi_l \right)^2 = \sum_{l=1}^M \int_{\Omega} \Psi_l^2 \int_D \hat{f}_l^2 \\ &= \sum_{l=1}^M \int_D \hat{f}_l^2 = \sum_{l=1}^M \|f_l\|_{L_2(D)}^2 \geq \|f_k\|_{L_2(D)}^2, \end{aligned}$$

for $k = 1, \dots, M$, which proves the theorem. \square

3. MULTIGRID

In this section we give a two-grid correction scheme for solving the system of equations given in §2.4. This scheme varies the mesh parameter from grid to grid, i.e. there is a coarse grid and a fine grid, while the stochastic discretization parameter, p , is held constant. Thereby, the scheme resembles that that would be applied to the regular deterministic problem. It is known that such a scheme when applied to the deterministic problem will converge at a rate independent of the value of the mesh parameter, h . We show that this is also the case for the stochastic problem. To show this we follow a regular multigrid analysis, as given e.g. in Braess (2001) or Elman et al. (2005), and show that a smoothing property and an approximation property hold. In order to establish the approximation property we make use of the doubly orthogonal polynomials introduced in §2.6. Once the convergence of the two-grid scheme has been shown to be independent of the mesh parameter it follows, by an inductive argument, that the convergence of a multigrid algorithm, which applies the two-grid algorithm recursively, is also independent of the mesh parameter.

3.1. Stationary Iteration. Central to the idea of multigrid is the understanding that certain stationary iterations when applied to particular matrix problems tend to *smooth* the associated error.

Consider a general matrix problem $A\mathbf{u} = \mathbf{f}$. Then the matrix splitting $A = M - N$ inspires the stationary iteration

$$(36) \quad \begin{aligned} \mathbf{u}^{(k+1)} &= M^{-1}N\mathbf{u}^{(k)} + M^{-1}\mathbf{f} \\ &= M^{-1}(M - A)\mathbf{u}^{(k)} + M^{-1}\mathbf{f} \\ &= (I - M^{-1}A)\mathbf{u}^{(k)} + M^{-1}\mathbf{f}. \end{aligned}$$

The matrix $I - M^{-1}A$ is the iteration matrix of the method and in the context of multigrid is called the *smoother*.

3.2. Two-grid Correction Scheme. Let $T \subset L_2(\Omega)$ and $S^{2h} \subset S^h \subset H_0^1(D)$ be as defined in §2.3. Then defining $V^{2h} = S^{2h} \otimes T$ and $V^h = S^h \otimes T$ we have $V^{2h} \subset V^h \subset H_0^1(D) \otimes L_2(\Omega)$. Finite element formulations in V^h and V^{2h} give rise to matrix equations which we represent as $\mathbf{A}\mathbf{u} = \mathbf{f}$ and $\bar{A}\bar{\mathbf{u}} = \bar{\mathbf{f}}$ respectively.

We now define a prolongation operator $I_{2h}^h : V^{2h} \rightarrow V^h$ via natural inclusion, i.e. for $v_{2h} \in V^{2h}$, $I_{2h}^h v_{2h} = v_{2h}$. To see how I_{2h}^h can be represented we note that any basis function ϕ_j^{2h} of S^{2h} can be expanded in the basis functions of S^h , viz.,

$$(37) \quad \phi_j^{2h} = \sum_{i=1}^{N_h} p_{ij} \phi_i^h, \quad j = 1, \dots, N_{2h}.$$

We define a matrix P using the coefficients above, i.e. $[P]_{ij} = p_{ij}$. Now we have, for $v_{2h} \in V^{2h}$,

$$(38) \quad \begin{aligned} v_{2h} &= \sum_{j=1}^{N_{2h}} \sum_{k=1}^M v_{jk}^{2h} \phi_j^{2h} \psi_k = \sum_{j=1}^{N_{2h}} \sum_{k=1}^M v_{jk}^{2h} \sum_{i=1}^{N_h} p_{ij} \phi_i^h \psi_k \\ &= \sum_{i=1}^{N_h} \sum_{k=1}^M \left(\sum_{j=1}^{N_{2h}} p_{ij} v_{jk}^{2h} \right) \phi_i^h \psi_k = \sum_{i=1}^{N_h} \sum_{k=1}^M [P\mathbf{v}_k^{2h}]_i \phi_i^h \psi_k. \end{aligned}$$

As $v_{2h} \in V^h$ we also have the expansion

$$(39) \quad v_{2h} = \sum_{i=1}^{N_h} \sum_{k=1}^M v_{ik}^h \phi_i^h \psi_k.$$

Comparing (38) and (39) we see that $[P\mathbf{v}_k^{2h}]_i = v_{ik}^h$ or that $P\mathbf{v}_k^{2h} = \mathbf{v}_k^h$. From this it follows that if \mathbf{v}^{2h} is the coefficient vector of v_{2h} in V^{2h} , then $(I \otimes P)\mathbf{v}^{2h}$ is the coefficient vector of v_{2h} in V^h . (Here I is an $M \times M$ identity matrix.) We therefore call $I \otimes P$ the prolongation matrix and introduce the notation $\mathcal{P} = I \otimes P$.

We next define a restriction operator $I_h^{2h} : V^h \rightarrow V^{2h}$ such that the corresponding restriction matrix \mathcal{R} satisfies $\mathcal{R} = \mathcal{P}^T$ (or equivalently $\mathcal{R} = I \otimes R$ where $R = P^T$). That is to say, that if I_h^{2h} maps $v_h \in V^h$ to $v_{2h} \in V^{2h}$ and \mathbf{v}^h and \mathbf{v}^{2h} are the respective coefficient vectors of these functions, then $\mathbf{v}^{2h} = \mathcal{R}\mathbf{v}^h = \mathcal{P}^T \mathbf{v}^h$. With the prolongation and restriction operators related in this way we have the desirable relationships $\bar{\mathbf{f}} = \mathcal{R}\mathbf{f}$ and $\bar{A} = \mathcal{R}A\mathcal{P}$.

Using these definitions we have the following algorithm for a two-grid iterative correction scheme.

```

choose initial guess  $\mathbf{u}$ 
for  $i = 0, 1, \dots$ 
  for  $j = 1 : k$ 
     $\mathbf{u} \leftarrow (I - M^{-1}A)\mathbf{u} + M^{-1}\mathbf{f}$ 
  end
   $\bar{\mathbf{r}} = \mathcal{R}(\mathbf{f} - A\mathbf{u})$ 
  solve  $\bar{A}\bar{\mathbf{e}} = \bar{\mathbf{r}}$ 
   $\mathbf{u} \leftarrow \mathbf{u} + \mathcal{P}\bar{\mathbf{e}}$ 
end

```

The success of this algorithm necessarily depends on how well the smoother works and how well the functions are passed between the coarse and fine grids.

3.3. Convergence of Two-Grid Correction Scheme. We wish to establish that the two-grid algorithm given in §3.2 converges and that the contraction rate is independent of h . This can be

shown to be true providing the *smoothing property* and the *approximation property* are satisfied, as is shown in the following theorem.

Theorem 8. *Providing the smoothing property,*

$$(40) \quad \|A(I - M^{-1}A)^k \mathbf{y}\|_2 \leq \eta(k) \|\mathbf{y}\|_A \quad \forall \mathbf{y} \in \mathbb{R}^{MN_h},$$

with $n(k) \rightarrow 0$ as $k \rightarrow \infty$, and the approximation property,

$$(41) \quad \|(A^{-1} - \mathcal{P}\bar{A}^{-1}\mathcal{R})\mathbf{y}\|_A \leq C_5 \|\mathbf{y}\|_2 \quad \forall \mathbf{y} \in \mathbb{R}^{MN_h},$$

with $C_5 = C_5(p)$, are satisfied, then, providing k is sufficiently large, the two-grid algorithm given in §3.2 converges and the contraction rate is independent of h .

Proof. It can be shown that the error associated with the two-grid algorithm obeys the recursive relationship

$$(42) \quad \mathbf{e}^{(i+1)} = (A^{-1} - \mathcal{P}\bar{A}^{-1}\mathcal{R})A(I - M^{-1}A)^k \mathbf{e}^{(i)}.$$

Hence,

$$(43) \quad \begin{aligned} \|\mathbf{e}^{(i+1)}\|_A &= \|(A^{-1} - \mathcal{P}\bar{A}^{-1}\mathcal{R})A(I - M^{-1}A)^k \mathbf{e}^{(i)}\|_A \\ &\leq C_5 \|A(I - M^{-1}A)^k \mathbf{e}^{(i)}\|_2 \\ &\leq C_5 \eta(k) \|\mathbf{e}^{(i)}\|_A. \end{aligned}$$

Since $\eta(k) \rightarrow 0$ as $k \rightarrow \infty$ there exists some minimal number of smoothing steps such that $C_5 \eta(k) < 1$. \square

The proof that the smoothing property holds is dependent on the choice of smoother, i.e. the value of M . For the case of $M = \theta I$, $\theta \in \mathbb{R}$, which choice corresponds to Richardson's iterative method, the proof follows that given in Braess (2001) and Elman et al. (2005). We will prove that the approximation property holds in §3.4.

3.4. Approximation Property. We here wish to show that the approximation property given in (41) is satisfied.

Theorem 9. *For the problem under consideration, the approximation property given in Theorem 8 holds.*

Proof. Given $\mathbf{y} \in \mathbb{R}^{MN_h}$ we can find some $f \in S^h \otimes T$ such that $\mathbf{y} = \mathbf{f}$. Let u_{hp} and $u_{2h,p}$ be the fine and coarse grid solutions respectively with coefficient vectors $\mathbf{u} = A^{-1}\mathbf{f}$ and $\bar{\mathbf{u}} = \bar{A}^{-1}\mathbf{f} = \bar{A}^{-1}\mathcal{R}\mathbf{f}$. Then we have

$$(44) \quad \begin{aligned} \|(A^{-1} - \mathcal{P}\bar{A}^{-1}\mathcal{R})\mathbf{y}\|_A^2 &= \|\mathbf{u} - \mathcal{P}\bar{\mathbf{u}}\|_A^2 = (\mathbf{u} - \mathcal{P}\bar{\mathbf{u}}, \mathbf{u} - \mathcal{P}\bar{\mathbf{u}})_A \\ &= a(u_{hp} - I_{2h}^h u_{2h,p}, u_{hp} - I_{2h}^h u_{2h,p}) \\ &= a(u_{hp} - u_{2h,p}, u_{hp} - u_{2h,p}). \end{aligned}$$

Utilizing the fact that c is bounded above P -a.e. by β and expanding u_{hp} and $u_{2h,p}$ using the doubly orthogonal polynomials we have

$$(45) \quad \|(A^{-1} - \mathcal{P}\bar{A}^{-1}\mathcal{R})\mathbf{y}\|_A^2 \leq \beta \int_{\Omega} \int_D \left| \nabla \left(\sum_{k=1}^M u_k^h \Psi_k - \sum_{k=1}^M u_k^{2h} \Psi_k \right) \right|^2.$$

Now using the fact that the doubly orthogonal polynomials are orthonormal we get

$$\begin{aligned}
(46) \quad \|(A^{-1} - \mathcal{P}\bar{A}^{-1}\mathcal{R})\mathbf{y}\|_A^2 &\leq \beta \sum_{k=1}^M \int_{\Omega} \Psi_k^2 \int_D |\nabla(u_k^h - u_k^{2h})|^2 \\
&= \beta \sum_{k=1}^M \int_D |\nabla(u_k^h - u_k^{2h})|^2 \\
&\leq \beta \sum_{k=1}^M \|u_k^h - u_k^{2h}\|_{H^1(D)}^2.
\end{aligned}$$

From Theorem 6 we know that $u_k^h \in S^h$, $u_k^{2h} \in S^{2h}$, $k = 1, \dots, M$, are finite element approximations to (deterministic) steady-state diffusion problems. Let $u_k \in H_0^1(D)$, $k = 1, \dots, M$, be the weak solutions to these problems. Furthermore, let $f_k \in S^h$, $k = 1, \dots, M$, be the source functions of the associated problems. Then there exists constant $C_d > 0$ such that $\|u - u_k^h\|_{H^1(D)} \leq C_d h \|f_k\|_{L_2(D)}$ for all k . Therefore,

$$\begin{aligned}
(47) \quad \|(A^{-1} - \mathcal{P}\bar{A}^{-1}\mathcal{R})\mathbf{y}\|_A^2 &\leq \beta \sum_{k=1}^M (\|u_k - u_k^h\|_{H^1(D)} + \|u_k - u_k^{2h}\|_{H^1(D)})^2 \\
&\leq \beta \sum_{k=1}^M (C_d h \|f_k\|_{L_2(D)} + 2C_d h \|f_k\|_{L_2(D)})^2 \\
&= \beta \sum_{k=1}^M (3C_d h \|f_k\|_{L_2(\Omega)})^2.
\end{aligned}$$

Now using Theorems 4 and 7 we have

$$\begin{aligned}
(48) \quad \|(A^{-1} - \mathcal{P}\bar{A}^{-1}\mathcal{R})\mathbf{y}\|_A^2 &\leq \beta \sum_{k=1}^M (3C_d h \|f\|_{L_2(D) \otimes L_2(\Omega)})^2 \\
&\leq \beta M \left(\frac{3C_d}{\sqrt{C_1}} \|\mathbf{f}\|_2 \right)^2 \\
&= \beta M \left(\frac{3C_d}{\sqrt{C_1}} \|\mathbf{y}\|_2 \right)^2
\end{aligned}$$

which completes the proof. \square

3.5. Extension to Multigrid. The two-grid correction scheme given in §3.2 only contains pre-smoothing. In practice post-smoothing is often also applied. In the numerical experiments given in §4 post-smoothing is applied. We have neglected post-smoothing in the preceding analytic argument in order to keep things a little simpler. It can be shown, though we omit the details here, that the two-grid correction scheme with post-smoothing also converges with a contraction factor independent of the spatial mesh parameter.

Recursively applying the two-grid correction scheme, gives rise to a multigrid scheme. A number of variations are possible, see, for example, Briggs, Henson & McCormick (2000). That multigrid converges with a contraction factor independent of the spatial mesh parameter can be established by an inductive argument once the two-grid scheme has been shown to converge with a contraction factor independent of the spatial mesh parameter. This inductive argument will be no different for the stochastic problem than for the analogous deterministic problem and is discussed e.g. in Braess (2001) and Elman et al. (2005).

4. NUMERICAL EXPERIMENTS

We now perform some numerical experiments to provide practical support for the theoretical results obtained in §3. The model that we use follows that given in Ghanem & Spanos (1991) and Xiu & Karniadakis (2002). We would also like to direct the reader's attention to Maitre, Knio, Debusschere, Najim & Ghanem (2003) where multigrid was applied to stochastic steady and unsteady diffusion problems and in which the conclusions reached are in agreement with the conclusions that we reach in §4.6 and §4.7.

4.1. Model Problem. We take the spatial domain to be $D = (-1, 1)^2$ and consider the deterministic source function $f = 1$.

To construct the diffusion coefficient we consider a process with mean function $c_0(x)$, constant variance ν , and covariance function $r(x, y)$. Such a process will have a Karhunen-Loève expansion of the form

$$(49) \quad c(x, \omega) = c_0(x) + \sum_{k=1}^{\infty} \sqrt{\lambda_k} c_k(x) \xi_k(\omega)$$

where (ξ_k) is a sequence of uncorrelated and identically distributed random variables, and (λ_k) and (c_k) can be computed by solving the eigenvalue equation

$$(50) \quad \int_D r(x, y) c_k(x) dx = \nu \lambda_k c_k(y).$$

If need be we make the further assumption that (ξ_k) is a sequence of independent random variables. The sequence (λ_k) is ordered so as to be non-increasing.

For computational purposes we need a finite term expansion so we approximate (49) by

$$(51) \quad c(x, \omega) = c_0(x) + \sum_{k=1}^m \sqrt{\lambda_k} c_k(x) \xi_k(\omega)$$

where λ_k , c_k , $k = 1, \dots, m$, still satisfy (50). From the modeling perspective the replacement of the infinite expansion with the finite expansion is justified providing (λ_k) decays rapidly. We demonstrate the decay of (λ_k) in §4.2 where we discuss the covariance function that we will use.

We will consider two cases for the distributions of the random variables ξ_k , $k = 1, \dots, m$. In §4.6 we take ξ_k , $k = 1, \dots, m$, to be uniformly distributed on $(-1, 1)$ with $c_0(x) = 10$ and $\nu = 1/3$. In §4.7 we take ξ_k , $k = 1, \dots, m$, to be normally distributed with $c_0(x) = 10$ and $\nu = 0.01$.

For information on the Karhunen-Loève expansion the reader is referred to Loève (1994). And for a discussion of its use in mathematical modeling the reader is referred to Ghanem & Spanos (1991).

4.2. Exponential Covariance. We consider for the covariance function of the diffusion coefficient the exponential covariance function given by

$$(52) \quad r(x, y) = \nu e^{-\frac{1}{b}|x_1 - y_1| - \frac{1}{b}|x_2 - y_2|}$$

where $x = (x_1, x_2)$, $y = (y_1, y_2) \in D$. The constant b is called the correlation length and will affect the decay of (λ_k) , a larger value producing faster decay.

Analytic expressions for (λ_k) and (c_k) can be obtained for this choice of covariance function. The derivation is given in Ghanem & Spanos (1991). We here will simply state the results. Given $D = (-a, a)^2$, let

$$(53) \quad \lambda_k^1 = \frac{2/b}{\theta_k^2 + (1/b)^2}, \quad k = 1, 2, \dots$$

and

$$(54) \quad c_k^1(x) = \frac{\cos(\theta_k a x)}{\sqrt{a + \frac{\sin(2\theta_k a)}{2\theta_k}}}, \quad k = 1, 3, \dots,$$

$$(55) \quad c_k^1(x) = \frac{\sin(\theta_k a x)}{\sqrt{a - \frac{\sin(2\theta_k a)}{2\theta_k}}}, \quad k = 2, 4, \dots,$$

where $c_k^1: (-a, a) \rightarrow \mathbb{R}$ and $(\theta_k) > 0$ is an increasing sequence satisfying the transcendental equations

$$(56) \quad (1/b) - \theta_k \tan(\theta_k a) = 0, \quad \theta_k + (1/b) \tan(\theta_k a) = 0.$$

Then

$$(57) \quad \lambda_k = \lambda_i^1 \lambda_j^1, \quad c_k(x) = c_i^1(x_1) c_j^1(x_2),$$

where the ordering is such that (λ_k) forms a non-increasing sequence. We note that (λ_k^1) and (c_k^1) are in fact the eigenvalues and eigenfunctions of the equivalent 1-dimensional problem.

We give these analytic expressions of (λ_k) and (c_k) so that the reader may better see the nature of the diffusion coefficient. In practice, however, the covariance function may be such that analytic expressions for (λ_k) and (c_k) are not available and (50) will need to be solved numerically. In fact, in the numerical experiments given in §4.6 and §4.7 we solved (50) numerically. This is not a trivial task as it involves solving a generalized eigenvalue problem in which the right hand side matrix, in this context sometimes called the Galerkin covariance matrix, is in general dense. This can be done efficiently by using the Lanczos algorithm together with fast methods for matrix-vector products, as discussed in Eiermann, Ernst & Ullmann (2005). This issue is also discussed in Karniadakis, Su, Xiu, Lucor, Schwab & Todor (2005).

For further discussion on appropriate choices of covariance functions see Ghanem & Spanos (1991) and Xiu & Karniadakis (2002).

4.3. Matrix Expansions. With c and f defined as in §4.1 we find that the matrices A and \mathbf{f} given in §2.4 have the form

$$(58) \quad A = G_0 \otimes A_0 + \sum_{k=1}^m \sqrt{\lambda_k} G_k \otimes A_k,$$

$$(59) \quad \mathbf{f} = \mathbf{g}_0 \otimes \mathbf{f}_0$$

where, defining $\xi_0 = 1$,

$$(60) \quad [G_k]_{ij} = \int_{\Omega} \xi_k \psi_i \psi_j, \quad [\mathbf{g}_0]_i = \int_{\Omega} \psi_i,$$

$$(61) \quad [A_k]_{ij} = \int_D c_k \nabla \phi_i \cdot \nabla \phi_j, \quad [\mathbf{f}_0]_i = \int_D \phi_i.$$

We note that the matrices A_k will be sparse due to the choice of spatial basis functions for S . These will form the blocks of the system matrix, A . The block structure will be determined by the matrices G_k and so it is important that the basis functions for T are chosen so that these too are sparse.

4.4. Polynomial Chaos. There are a number of choices on how to construct T . As mentioned in §2.3 we follow the methodology given in Xiu & Karniadakis (2002) which generalizes the method of polynomial chaos as given in Ghanem & Spanos (1991). The functions thus chosen to be the computational basis for T are those from the Askey scheme of hypergeometric polynomials that are orthogonal to with respect to the probability measure, i.e.

$$(62) \quad \int_{\Omega} \psi_k \psi_l = d_k \delta_{kl},$$

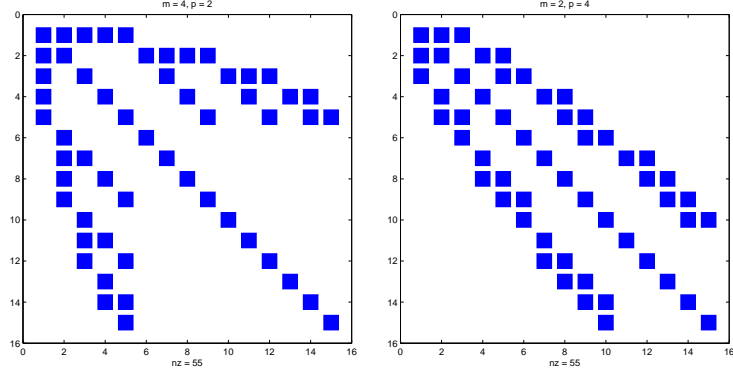


FIGURE 1. Block structure of system matrix A .

and are of degree p or less. It follows directly that the matrix G_0 will be diagonal. Furthermore, the matrices G_k , $k = 1, \dots, m$, will turn out to be sparse due to the three-term recurrence relation that all orthogonal polynomials satisfy. The resultant block structure of A is shown in Figure 1 for two choices of m and p . We emphasize that the solid block shown in the figure represents a sparse matrix having the structure of a stiffness matrix arising from a deterministic steady-state diffusion problem, as discussed in §4.3.

4.5. Spatial Mesh. As mentioned in §4.1 the spatial domain is taken to be $D = (-1, 1)^2$. For the triangulation of D we will use uniform meshes consisting of an underlying grid of $n \times n$ squares each of which is further subdivided into two equal triangles.

4.6. Multigrid for Diffusion with Uniform Distributions. We now let ξ_k , $k = 1, \dots, m$, be uniformly distributed on $\Gamma = (-1, 1)$. Therefore, $\Omega = (-1, 1)^m$ and $dP = d\omega/2^m$. We also set $\nu = 1/3$.

Given the nature of λ_k , c_k , $k = 1, \dots, m$, as given in §4.2, and the form of the diffusion coefficient as given in §4.1, it can be shown that (6) will be satisfied providing the mean is sufficiently large. We here take $c_0(x) = 10$.

Applying the generalized polynomial chaos method as described in §4.4 the basis of T will be the set of m -dimensional Legendre polynomials of degree p or less.

Now multigrid is applied. A full V-cycle is used with an $n \times n$ finest mesh and a 2×2 coarsest mesh. For the smoother we use the damped Jacobi method with the damping parameter set to $2/3$. Three pre-smoothing and three post-smoothing iterates are carried out. The iterations stop when the relative residual reaches a tolerance of 10^{-6} . Table 1 shows the number of iterations required for convergence for varying values of m , n , and p . The results clearly support the theoretical conclusion that the contraction rate of the multigrid algorithm is mesh independent. The table also indicates that the method is apparently insensitive to the parameters m and p .

4.7. Multigrid for Diffusion with Normal Distributions. We now let ξ_k , $k = 1, \dots, m$, be normally distributed with zero mean and variance ν . Now we have $\Omega = \mathbb{R}^m$ and $dP = e^{-\omega^2/(2\nu)}/(2\pi\nu)^{m/2}$. We take $c_0(x) = 1$.

Note that the diffusion coefficient as defined in §4.1 will now fail to satisfy condition (6) no matter what the choice of ν . However, we have reason to believe that the theory still applies. We give here only a heuristic argument. Given a sufficiently small variance the probability of c being outside of two positive bounds becomes negligibly small. That is to say, that if the normal distributions were replaced by similar distributions that looked like the normal distributions with their tails cut off

$n = 4$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
$p = 1$	6	6	6	6
$p = 2$	6	6	6	6
$p = 3$	6	6	6	6
$p = 4$	6	6	6	6
$n = 8$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
$p = 1$	7	7	7	7
$p = 2$	7	7	7	7
$p = 3$	7	7	7	7
$p = 4$	7	7	7	7
$n = 16$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
$p = 1$	7	7	7	7
$p = 2$	7	7	7	7
$p = 3$	7	7	7	7
$p = 4$	7	7	7	7
$n = 32$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
$p = 1$	7	7	7	7
$p = 2$	7	7	7	7
$p = 3$	7	7	7	7
$p = 4$	7	7	7	7

TABLE 1. Number of iterations required for multigrid to converge for diffusion defined via uniform distributions.

so as to ensure that c satisfies (6), then the difference would not be noticed computationally. We emphasize that we have not pursued this reasoning analytically. We have found that sufficiently small variance results in positive definite systems that yield sensible results. We note that this problem has been tackled in Ghanem & Spanos (1991) and Xiu & Karniadakis (2002). We take $\nu = 0.01$.

Applying the generalized polynomial chaos method as described in §4.4 the basis of T will be the set of m -dimensional generalized Hermite polynomials of degree p or less.

Now multigrid is applied. A full V-cycle is used with an $n \times n$ finest mesh and a 2×2 coarsest mesh. For the smoother we use the damped Jacobi method with the damping parameter set to $2/3$. Three pre-smoothing and three post-smoothing iterates are carried out. The iterations stop when the relative residual reaches a tolerance of 10^{-6} . Table 2 shows the number of iterations required for convergence for varying values of m , n , and p . The results support the theoretical conclusion that the contraction rate of the multigrid algorithm is mesh independent. The table also indicates that the method is apparently insensitive to m and only slightly sensitive to p .

REFERENCES

- Babuška, I., Tempone, R. & Zouraris, G. E. (2004), ‘Galerkin finite element approximation of stochastic elliptic partial differential equations’, *Siam J. Numer. Anal.* **42**, 800–825.
- Braess, D. (2001), *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*, second edn, Cambridge University Press.
- Briggs, W. L., Henson, V. E. & McCormick, S. F. (2000), *A Multigrid Tutorial*, second edn, SIAM.
- Eiermann, M., Ernst, O. G. & Ullmann, E. (2005), Computational aspects of the stochastic finite element method, in ‘Proceedings of Algorithmy’.

$n = 4$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
$p = 1$	6	6	6	6
$p = 2$	7	7	7	7
$p = 3$	7	7	7	7
$p = 4$	7	7	7	7
$n = 8$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
$p = 1$	8	8	8	8
$p = 2$	8	8	8	8
$p = 3$	9	9	9	9
$p = 4$	10	10	10	10
$n = 16$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
$p = 1$	8	8	8	8
$p = 2$	8	8	8	8
$p = 3$	9	9	9	9
$p = 4$	9	10	10	10
$n = 32$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
$p = 1$	7	7	8	8
$p = 2$	8	8	8	8
$p = 3$	8	8	9	9
$p = 4$	9	9	9	9

TABLE 2. Number of iterations required for multigrid to converge for diffusion defined via uniform distributions.

- Elman, H. C., Silvester, D. J. & Wathen, A. J. (2005), *Finite Elements and Fast Iterative Solvers: with Applications in Incompressible Fluid Dynamics*, Oxford University Press.
- Ghanem, R. G. & Spanos, P. D. (1991), *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag.
- Karniadakis, G., Su, C.-H., Xiu, D., Lucor, D., Schwab, C. & Todor, R. (2005), Generalized polynomial chaos solution for differential equations with random inputs, Technical Report 2005-01, Eidgenössische Technische Hochschule Zürich.
- Loève, M. (1994), *Probability Theory II*, Springer.
- Maitre, O. L., Knio, O., Debusschere, B., Najim, H. & Ghanem, R. (2003), ‘A multigrid solver for two-dimensional stochastic diffusion equations’, *Comput. Methods Appl. Mech. Engrg.* **192**, 4723–4744.
- Treves, F. (1967), *Topological Vector Spaces, Distributions and Kernels*, Academic Press.
- Xiu, D. & Karniadakis, G. E. (2002), ‘Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos’, *Comput. Methods Appl. Mech. Engrg.* **191**, 4927–4948.