*IMA Journal of Numerical Analysis* (2007) **27**, 675−688
doi:10.1093/imanum/drm006
Advance Access publication on April 9, 2007

# Solving the stochastic steady-state diffusion problem using multigrid

HOWARD ELMAN†
*Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742, USA*

AND

DARRAN FURNIVAL‡
*Applied Mathematics and Scientific Computing Program, University of Maryland, College Park, MD 20742, USA*

[Received on 20 February 2006; revised on 5 March 2007]

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, and the stochastic discretization.

*Keywords*: Karhunen–Loève expansion; multigrid; polynomial chaos.

## 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 modelled 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 *et al.* (2004), Ghanem & Spanos (1991) and Xiu & Karniadakis (2002).

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

†Email: elman@cs.umd.edu
‡Corresponding author. Email: furnival@cs.umd.edu

finite-dimensional subspace of the infinite-dimensional space that contains the weak solution in order to obtain a matrix problem.

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 use polynomial chaos of a given order. This corresponds to the polynomial chaos method as pioneered by Ghanem & Spanos (1991) and generalized in Xiu & Karniadakis (2002).

Theoretically, we apply a two-grid correction scheme to solve the finite-dimensional problem. In this scheme, the spatial discretization is varied from grid to grid while the stochastic discretization is kept constant. We show that the convergence rate of this method is independent of the discretization parameters. Multigrid algorithms, obtained by applying the two-grid correction scheme recursively, can then be shown to have convergence rates independent of the discretization parameters via inductive arguments. We do not give these inductive arguments but note that the reasoning would be the same as that for the analogous deterministic problem.

Experimentally, we consider two problems. These are obtained by defining the diffusion coefficient to be a finite Karhunen–Loève expansion consisting of random variables that are for the first problem uniformly distributed and for the second problem normally distributed. A multigrid algorithm consisting of a full V-cycle is then applied to solve a selection of problems associated with different discretization parameters, both spatial and stochastic. The number of iterations it requires to solve the problem within a given tolerance are tabulated. It will be seen that these tables provide experimental evidence that the converge rate of the multigrid algorithm is indeed independent of the discretization parameters.

We note that an algorithm of this nature was employed in Le Maître *et al.* (2003), where multigrid was applied to steady and unsteady diffusion problems and in which the conclusions reached are in agreement with our own. This paper builds on Le Maître *et al.* (2003) by providing theoretical analysis supporting the numerical results. We also note that multigrid was applied to the steady-state problem in Seynaeve *et al.* (2007), where classical Fourier mode analysis was extended to the stochastic case.

## 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.

### 2.1 *Boundary-value problem*

The stochastic steady-state diffusion equation with homogeneous Dirichlet boundary-value conditions is given by

$$\begin{cases} -\nabla \cdot (c \nabla u) = f & \text{in } D \times \Omega, \\ u = 0 & \text{on } \partial D \times \Omega, \end{cases} \tag{2.1}$$

where $D$ is the spatial domain, $\Omega$ is a sample space, $c\colon D \times \Omega \to \mathbb{R}$ is the diffusion coefficient and $f\colon D \times \Omega \to \mathbb{R}$ is the source function. The sample space in turn belongs to a probability space

$(\Omega, \mathscr{F}, P)$, where $\mathscr{F}$ is a $\sigma$-algebra and $P$ is a probability measure. Note that the divergence and gradient operators are considered to act on spatial components only.

The spatial domain, $D$, is assumed to be a 2D simply connected bounded open set with piecewise smooth boundary. In particular, we take $D$ to be the interior of a polygon.

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

$$c(x, \omega) = c_0(x) + \sum_{r=1}^{m} \sqrt{\lambda_r} c_r(x) \xi_r(\omega), \tag{2.2}$$

where $\xi_r \colon \Omega \to \mathbb{R}$ are identically distributed independent random variables with zero mean and for $\omega = (\omega_1, \ldots, \omega_m) \in \Omega$, $\xi_r(\omega) = \omega_r$. Note that the distribution of $\xi = (\xi_1, \ldots, \xi_m)$ will dictate the probability measure to be used. For example, if each $\xi_r$ 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, $u \colon \overline{D} \times \Omega \to \mathbb{R}$, 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 modelling 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 (2.1), 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 (2.1) is given by the following: find $u \in H_0^1(D) \otimes L_2(\Omega)$ such that

$$a(u, v) = l(v) \quad \forall v \in H_0^1(D) \otimes L_2(\Omega), \tag{2.3}$$

where

$$a(u, v) = \int_\Omega \int_D c \nabla u \cdot \nabla v, \tag{2.4}$$

$$l(v) = \int_\Omega \int_D f v. \tag{2.5}$$

Note that the integral over $\Omega$ is with respect to the probability measure, i.e. $\int_\Omega \int_D = \int_\Omega \int_D \mathrm{d}x \, \mathrm{d}P$.

The 'Lax–Milgram lemma' can be used to show that there exists a unique solution to this problem provided that there exist positive constants $\alpha$ and $\beta$ such that

$$\alpha \leqslant c(x, \omega) \leqslant \beta \quad P\text{-a.e. } \forall x \in D, \tag{2.6}$$

where by $P$-a.e. we mean that there exists a subset $F \in \mathscr{F}$ with $P(F) = 0$ such that the inequality holds on the complement of $F$.

Assuming this condition holds, we can define the energy norm $\|\cdot\|_a^2 = a(\cdot, \cdot)$, which plays an analogous role to the energy norm defined in the analysis of the deterministic diffusion problem.

### 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 Section 2.2. This entails a discretization of both the spatial and the stochastic parts of the problem.

The spatial domain is discretized using a conforming triangulation, $\mathscr{T}$. We assume that any triangulation used belongs to a family of triangulations that is quasi-uniform and shape regular (see Braess (2001) for the definitions of quasi-uniformity and shape regularity). The finite-dimensional subspace of $H_0^1(D)$ is then taken to be $S = \text{span}\{\phi_1, \ldots, \phi_N\}$, where $\phi_k \colon \overline{D} \to \mathbb{R}$, $k = 1, \ldots, N$, are the usual piecewise linear basis functions defined at the nodes of $\mathscr{T}$. (Here, $N$ is the number of internal nodes in the triangulation.)

To construct the finite-dimensional subspace of $L_2(\Omega)$, the 'polynomial chaos method', as originally given in Ghanem & Spanos (1991) and generalized in Xiu & Karniadakis (2002), is used. This essentially involves using the set of $m$-dimensional functions (recalling that $m$ is the number of terms in (2.2)) from the 'Askey scheme of hypergeometric polynomials' that satisfy the orthogonality relationship

$$\int_\Omega \psi_k \psi_l = \delta_{kl}, \tag{2.7}$$

where we have assumed that the polynomials have been normalized. We note that if the polynomials have not been normalized, the convergence analysis given in Section 3 still holds, though the algebra is a little less tidy. The finite-dimensional subset of $L_2(\Omega)$ is then defined to be the set of all such polynomials that are of degree $p$ or less, which will be denoted as $T = \text{span}\{\psi_1, \ldots, \psi_M\}$, where

$$M = \frac{(m+p)!}{m!\,p!}. \tag{2.8}$$

Such a set is often referred to as the 'm-dimensional polynomial chaos of order $p$'.

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

$$a(u_{hp}, v) = l(v) \quad \forall\, v \in S \otimes T, \tag{2.9}$$

where $a(\cdot,\cdot)$ and $l(\cdot)$ are as in (2.4) and (2.5). This will possess a unique solution under the same conditions that apply to the weak formulation.

### 2.4 *Matrix formulation*

By substituting the expansion

$$u_{hp} = \sum_{j=1}^{N} \sum_{l=1}^{M} u_{jl} \phi_j \psi_l \tag{2.10}$$

into (2.9) 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

$$A\mathbf{u} = \mathbf{f}, \tag{2.11}$$

where

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1M} \\ \vdots & & \vdots \\ A_{M1} & \cdots & A_{MM} \end{bmatrix}, \quad [A_{kl}]_{ij} = \int_\Omega \int_D c \nabla \phi_i \cdot \nabla \phi_j \psi_k \psi_l, \tag{2.12}$$

and

$$\mathbf{f} = \begin{bmatrix} \mathbf{f}_1 \\ \vdots \\ \mathbf{f}_M \end{bmatrix}, \quad [\mathbf{f}_k]_i = \int_\Omega \int_D f \phi_i \psi_k. \tag{2.13}$$

The solution vector, $\mathbf{u}$, contains the coefficients in (2.10) stacked columnwise, i.e. $\mathbf{u} = [u_{11}, \ldots, u_{N1}, \ldots, u_{1M}, \ldots, u_{NM}]^{\mathrm{T}}$.

We note that the matrix $A$, thus defined, is symmetric and positive definite. This fact is implicitly used throughout the remainder of this paper in demonstrating the convergence of multigrid.

Though not immediately transparent from (2.12), the choice of $S$ ensures that the blocks of $A$ are sparse and the choice of $T$ ensures that $A$ has a sparse block structure. This is further discussed in Pellissetti & Ghanem (2000) where block sparsity plots of $A$ are given for various choices of $m$ and $p$.

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.

### 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 multigrid in Section 3.

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

$$[E]_{kl} = \int_\Omega \psi_k \psi_l, \quad [B]_{ij} = \int_D \phi_i \phi_j, \tag{2.14}$$

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

$$E \otimes B = \begin{bmatrix} e_{11}B & \cdots & e_{1M}B \\ \vdots & & \vdots \\ e_{M1}B & \cdots & e_{MM}B \end{bmatrix} = \begin{bmatrix} B & & \\ & \ddots & \\ & & B \end{bmatrix} \tag{2.15}$$

as, by (2.7), $E$ is an identity matrix.

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

$$v = \sum_{i=1}^{N} \sum_{k=1}^{M} v_{ik} \phi_i \psi_k. \tag{2.16}$$

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

$$\mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_M \end{bmatrix}, \quad [\mathbf{v}_k]_i = v_{ik}. \tag{2.17}$$

THEOREM 2.1 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 (2.13).

*Proof.* This follows upon substituting the expansion of $f$ into (2.13). □

THEOREM 2.2 Let $f \in S \otimes T$ with coefficient vector $\hat{\mathbf{f}}$. Then $\| f \|^2_{L_2(D)\otimes L_2(\Omega)} = ((E \otimes B)\hat{\mathbf{f}}, \hat{\mathbf{f}})$.

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

THEOREM 2.3 The inequality

$$C_1 h^2 \leqslant \frac{((E \otimes B)\mathbf{v}, \mathbf{v})}{(\mathbf{v}, \mathbf{v})} \tag{2.18}$$

holds for all $\mathbf{v} \in \mathbb{R}^{MN}$, where $C_1$ is a constant.

*Proof.* The right-hand side of (2.18) is the Rayleigh quotient of $E \otimes B$ and this is bounded below by the lowest eigenvalue of $E \otimes B$. The eigenvalues of $E \otimes B$ are the products of the eigenvalues of $E$ and the eigenvalues of $B$. Denoting the minimum eigenvalues of $E$ and $B$ as $\lambda_{\min}(E)$ and $\lambda_{\min}(B)$, respectively, we thus have

$$\lambda_{\min}(E)\lambda_{\min}(B) \leqslant \frac{((E \otimes B)\mathbf{v}, \mathbf{v})}{(\mathbf{v}, \mathbf{v})}. \tag{2.19}$$

For a quasi-uniform and shape-regular mesh, the minimum eigenvalue of $B$ is bounded below by $C_1 h^2$, where $C_1$ is a constant, as shown in, e.g. Elman *et al.* (2005), and $\lambda_{\min}(E) = 1$ as $E$ is an identity matrix. □

THEOREM 2.4 Let $f \in S \otimes T$. Then $h\sqrt{C_1}\| f \|_{L_2(D)\otimes L_2(\Omega)} \leqslant \|\mathbf{f}\|_2$, where $C_1$ is as in Theorem 2.3.

*Proof.* Using Theorem 2.1, we have $\|\mathbf{f}\|_2^2 = ((E \otimes B)\hat{\mathbf{f}}, (E \otimes B)\hat{\mathbf{f}})$. Now setting $\mathbf{g} = (E \otimes B)^{\frac{1}{2}}\hat{\mathbf{f}}$ and using Theorems 2.2 and 2.3, we have

$$C_1 h^2 \leqslant \frac{((E \otimes B)\mathbf{g}, \mathbf{g})}{(\mathbf{g}, \mathbf{g})} = \frac{((E \otimes B)\hat{\mathbf{f}}, (E \otimes B)\hat{\mathbf{f}})}{((E \otimes B)\hat{\mathbf{f}}, \hat{\mathbf{f}})} = \frac{\|\mathbf{f}\|_2^2}{\| f \|^2_{L_2(D)\otimes L_2(\Omega)}} \tag{2.20}$$

as required. □

### 2.6 *Semi-discrete formulation*

In proving the approximation property used in the analysis of multigrid in Section 3.5, it will be useful to introduce the solution of a semi-discrete version of the finite-element formulation where only the stochastic space is discretized. This is given by the following: find $u_p \in H_0^1(D) \otimes T$ such that

$$a(u_p, v) = l(v) \quad \forall v \in H_0^1(D) \otimes T. \tag{2.21}$$

This has a unique solution under the same conditions as apply to the weak formulation in Section 2.2.

THEOREM 2.5 The solution to the semi-discrete problem, $u_p$, and the finite-element approximation, $u_{hp}$, defined in Section 2.3, satisfy

$$\|u_p - u_{hp}\|_a \leqslant \sqrt{\beta} C_2 h \| D^2 u_p \|_{L_2(D)\otimes L_2(\Omega)}, \tag{2.22}$$

where $\beta$ is defined in (2.6), $\|D^2 v\|_{L_2(D)\otimes L_2(\Omega)}$ is defined by

$$\|D^2 v\|^2_{L_2(D)\otimes L_2(\Omega)} = \int_\Omega |v|^2_{H^2(D)} \quad \forall\, v \in H^2(D)\otimes L_2(\Omega) \tag{2.23}$$

and $C_2$ is a constant.

*Proof.* From Galerkin orthogonality, we have

$$\|u_p - u_{hp}\|_a \leqslant \|u_p - v\|_a \quad \forall\, v \in S \otimes T. \tag{2.24}$$

Now, let $\tilde{u}_p \in S \otimes T$ be the spatial interpolant of $u_p \in H^1_0(D) \otimes T$, i.e. if $x_j$, $j = 1, \ldots, N$, are the nodes of the spatial triangulation $\mathscr{T}$, then $\tilde{u}_p(x_j, \omega) = u_p(x_j, \omega)\ \forall\, \omega \in \Omega$. Then

$$\|u_p - u_{hp}\|^2_a \leqslant \|u_p - \tilde{u}_p\|^2_a \leqslant \beta \int_\Omega |u_p - \tilde{u}_p|^2_{H^1(D)}. \tag{2.25}$$

A standard interpolation result, as given in, e.g. Johnson (1987), tells us that

$$|u_p - \tilde{u}_p|_{H^1(D)} \leqslant C_2 h |u_p|_{H^2(D)} \quad \forall\, \omega \in \Omega, \tag{2.26}$$

where $C_2$ is only dependent on the spatial domain. Therefore,

$$\|u_p - u_{hp}\|^2_a \leqslant \beta C_2^2 h^2 \int_\Omega |u_p|^2_{H^2(D)} = \beta C_2^2 h^2 \|D^2 u_p\|^2_{L_2(D)\otimes L_2(\Omega)}, \tag{2.27}$$

which proves the theorem. □

THEOREM 2.6 The $H^2$-regularity bound

$$\|D^2 u_p\|_{L_2(D)\otimes L_2(\Omega)} \leqslant C_3 \|f\|_{L_2(D)\otimes L_2(\Omega)} \tag{2.28}$$

holds, where $C_3$ is a constant.

*Proof.* From elliptic regularity, we have

$$|u_p|_{H^2(D)} \leqslant C_3 \|f\|_{L_2(D)} \quad \forall\, \omega \in \Omega, \tag{2.29}$$

where $C_3$ is dependent on the spatial domain and $\beta$. Squaring and integrating over $\Omega$ give the desired result. □

## 3. Multigrid

In this section, we give a two-grid correction scheme for solving the system of equations given in Section 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 polynomial chaos order, $p$, is held constant. Thereby, the scheme resembles that which 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. Moreover, the convergence rate for the stochastic problem will also be independent of the parameters $m$ and $p$. To show this, we follow a regular multigrid analysis, as given in, e.g. Braess (2001) or Elman *et al.* (2005), and show that a smoothing property and an approximation property hold. Once the convergence rate of the two-grid scheme has been shown to be independent of the discretization parameters, then inductive arguments can be used to show that multigrid algorithms, derived from applying the two-grid scheme recursively, also have a convergence rate independent of the discretization parameters.

### 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. Given the problem $A\mathbf{u} = \mathbf{f}$, the matrix splitting $A = Q - Z$ inspires the stationary iteration

$$\begin{aligned}\mathbf{u}^{(k+1)} &= Q^{-1}Z\mathbf{u}^{(k)} + Q^{-1}\mathbf{f} \\ &= Q^{-1}(Q - A)\mathbf{u}^{(k)} + Q^{-1}\mathbf{f} \\ &= (I - Q^{-1}A)\mathbf{u}^{(k)} + Q^{-1}\mathbf{f}.\end{aligned} \tag{3.1}$$

The matrix $I - Q^{-1}A$ 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 Section 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 $A\mathbf{u} = \mathbf{f}$ and $\bar{A}\bar{\mathbf{u}} = \bar{\mathbf{f}}$, respectively.

We now define a prolongation operator $I_{2h}^h$: $V^{2h} \to V^h$ 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 $\phi_j^{2h}$ of $S^{2h}$ can be expanded in the basis functions of $S^h$, viz.,

$$\phi_j^{2h} = \sum_{i=1}^{N_h} p_{ij}\phi_i^h, \quad j = 1, \ldots, N_{2h}. \tag{3.2}$$

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

$$\begin{aligned}v_{2h} &= \sum_{j=1}^{N_{2h}}\sum_{k=1}^{M} v_{jk}^{2h}\phi_j^{2h}\psi_k = \sum_{j=1}^{N_{2h}}\sum_{k=1}^{M} v_{jk}^{2h}\sum_{i=1}^{N_h} p_{ij}\phi_i^h\psi_k \\ &= \sum_{i=1}^{N_h}\sum_{k=1}^{M}\left(\sum_{j=1}^{N_{2h}} p_{ij}v_{jk}^{2h}\right)\phi_i^h\psi_k = \sum_{i=1}^{N_h}\sum_{k=1}^{M}[P\mathbf{v}_k^{2h}]_i\phi_i^h\psi_k.\end{aligned} \tag{3.3}$$

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

$$v_{2h} = \sum_{i=1}^{N_h}\sum_{k=1}^{M} v_{ik}^h\phi_i^h\psi_k. \tag{3.4}$$

Comparing (3.3) and (3.4), we see that $[P\mathbf{v}_k^{2h}]_i = v_{ik}^h$ or $P\mathbf{v}_k^{2h} = \mathbf{v}_k^h$. 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 $\mathscr{P} = I \otimes P$.

We next define a restriction operator $I_h^{2h}$: $V^h \to V^{2h}$ such that the corresponding restriction matrix $\mathscr{R}$ satisfies $\mathscr{R} = \mathscr{P}^{\mathrm{T}}$ (or equivalently $\mathscr{R} = I \otimes R$, where $R = P^{\mathrm{T}}$). That is to say that if

$I_h^{2h}$ maps $v_h \in V^{2h}$ to $v_{2h} \in V^{2h}$ and $\mathbf{v}^h$ and $\mathbf{v}^{2h}$ are the respective coefficient vectors of these functions, then $\mathbf{v}^{2h} = \mathscr{R}\mathbf{v}^h = \mathscr{P}^{\mathrm{T}}\mathbf{v}^h$. With the prolongation and restriction operators related in this way, we have the desirable relationships $\bar{\mathbf{f}} = \mathscr{R}\mathbf{f}$ and $\bar{A} = \mathscr{R}A\mathscr{P}$.

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

choose initial guess $\mathbf{u}$
for $i = 0, 1, \ldots$
  for $j = 1{:}\, k$
    $\mathbf{u} \leftarrow (I - Q^{-1}A)\mathbf{u} + Q^{-1}\mathbf{f}$
  end
  $\bar{\mathbf{r}} = \mathscr{R}(\mathbf{f} - A\mathbf{u})$
  solve $\bar{A}\bar{\mathbf{e}} = \bar{\mathbf{r}}$
  $\mathbf{u} \leftarrow \mathbf{u} + \mathscr{P}\bar{\mathbf{e}}$
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 the fine grids.

### 3.3 *Convergence of two-grid correction scheme*

We wish to establish that the two-grid algorithm given in Section 3.2 converges and that the convergence rate is independent of $h$, $m$ and $p$. That the scheme converges can be shown to be true provided that the 'smoothing property' and the 'approximation property' are satisfied, as is shown in the following theorem.

THEOREM 3.1 Provided that the smoothing property,

$$\|A(I - Q^{-1}A)^k\mathbf{y}\|_2 \leqslant \eta(k)\|\mathbf{y}\|_A \quad \forall\, \mathbf{y} \in \mathbb{R}^{MN_h}, \tag{3.5}$$

with $\eta(k) \to 0$ as $k \to \infty$, and the approximation property,

$$\|(A^{-1} - \mathscr{P}\bar{A}^{-1}\mathscr{R})\mathbf{y}\|_A \leqslant C_4\|\mathbf{y}\|_2 \quad \forall\, \mathbf{y} \in \mathbb{R}^{MN_h}, \tag{3.6}$$

where $C_4$ is a constant, are satisfied, then, provided that $k$ is sufficiently large, the two-grid algorithm given in Section 3.2 converges.

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

$$\mathbf{e}^{(i+1)} = (A^{-1} - \mathscr{P}\bar{A}^{-1}\mathscr{R})A(I - Q^{-1}A)^k\mathbf{e}^{(i)}. \tag{3.7}$$

Hence,

$$\begin{aligned}\|\mathbf{e}^{(i+1)}\|_A &= \|(A^{-1} - \mathscr{P}\bar{A}^{-1}\mathscr{R})A(I - Q^{-1}A)^k\mathbf{e}^{(i)}\|_A \\ &\leqslant C_4\|A(I - Q^{-1}A)^k\mathbf{e}^{(i)}\|_2 \leqslant C_4\eta(k)\|\mathbf{e}^{(i)}\|_A.\end{aligned} \tag{3.8}$$

Since $\eta(k) \to 0$ as $k \to \infty$, there exists some minimal number of smoothing steps such that $C_4\eta(k) < 1$. □

That the smoothing property and approximation hold and that $\eta(k)$ and $C_4$ are independent of $h$, $m$ and $p$ are discussed in Sections 3.4 and 3.5.

### 3.4 *Smoothing property*

The proof that the smoothing property holds is dependent on the choice of smoother, i.e. the choice of $Q$. For the case of $Q = \theta I$, $\theta \in \mathbb{R}$, whose choice corresponds to Richardson's iterative method, the proof follows that given in Braess (2001) and Elman *et al.* (2005) where the damped Jacobi iteration is applied to the deterministic diffusion problem and $\eta(k)$ is shown to be independent of $h$. Moreover, it can be shown, though we omit the details, that $\theta$ can be chosen such that $\eta(k)$ is also independent of $m$ and $p$.

### 3.5 *Approximation property*

We here wish to show that the approximation property given in (3.6) is satisfied.

THEOREM 3.2 For the problem under consideration, the approximation property given in Theorem 3.1 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 $\bar{\mathbf{u}} = \bar{A}^{-1}\mathbf{f} = \bar{A}^{-1}\mathscr{R}\mathbf{f}$. Then, we have

$$\begin{aligned}\|(A^{-1} - \mathscr{P}\bar{A}^{-1}\mathscr{R})\mathbf{y}\|_A^2 &= \|\mathbf{u} - \mathscr{P}\bar{\mathbf{u}}\|_A^2 = (\mathbf{u} - \mathscr{P}\bar{\mathbf{u}}, \mathbf{u} - \mathscr{P}\bar{\mathbf{u}})_A \\ &= a(u_{hp} - I_{2h}^h u_{2h,p}, u_{hp} - I_{2h}^h u_{2h,p}) \\ &= a(u_{hp} - u_{2h,p}, u_{hp} - u_{2h,p}) = \|u_{hp} - u_{2h,p}\|_a^2. \end{aligned} \tag{3.9}$$

Now introducing the solution of the semi-discrete problem, $u_p$, defined in Section 2.6, and applying Theorems 2.5 and 2.6 give

$$\begin{aligned}\|(A^{-1} - \mathscr{P}\bar{A}^{-1}\mathscr{R})\mathbf{y}\|_A &\leqslant \|u_p - u_{hp}\|_a + \|u_p - u_{2h,p}\|_a \\ &\leqslant \sqrt{\beta}C_2 h \|D^2 u_p\|_{L_2(D)\otimes L_2(\Omega)} + 2\sqrt{\beta}C_2 h \|D^2 u_p\|_{L_2(D)\otimes L_2(\Omega)} \\ &\leqslant 3\sqrt{\beta}C_2 C_3 h \|f\|_{L_2(D)\otimes L_2(\Omega)}. \end{aligned} \tag{3.10}$$

Finally, applying Theorem 2.4 gives

$$\|(A^{-1} - \mathscr{P}\bar{A}^{-1}\mathscr{R})\mathbf{y}\|_A \leqslant \frac{3\sqrt{\beta}C_2 C_3}{\sqrt{C_1}}\|\mathbf{f}\|_2 = \frac{3\sqrt{\beta}C_2 C_3}{\sqrt{C_1}}\|\mathbf{y}\|_2, \tag{3.11}$$

which establishes the approximation property. □

### 3.6 *Extension to multigrid*

The two-grid correction scheme given in Section 3.2 only contains pre-smoothing. In practice, post-smoothing is often also applied, as in the numerical experiments given in Section 4. 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 convergence rate independent of $h$, $m$ and $p$.

Recursively applying the two-grid correction scheme gives rise to a multigrid scheme. A number of variations are possible, see, e.g. Briggs *et al.* (2000). That multigrid converges with a convergence rate independent of the parameters $h$, $m$ and $p$ can be established by an inductive argument. This inductive argument will be no different for the stochastic problem than for the analogous deterministic problem and is discussed in, e.g. 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 Section 3. The model that we use follows that given in Ghanem & Spanos (1991) and Xiu & Karniadakis (2002).

### 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 $\nu$ and covariance function $r(x, y)$. Such a process will have a Karhunen–Loéve expansion of the form

$$c(x, \omega) = c_0(x) + \sum_{k=1}^{\infty} \sqrt{\lambda_k} c_k(x) \xi_k(\omega), \tag{4.1}$$

where $(\xi_k)$ is a sequence of uncorrelated and identically distributed random variables with variance $\nu$ and mean zero and $(\lambda_k)$ and $(c_k)$ can be computed by solving the eigenvalue equation

$$\int_D \frac{r(x, y)}{\nu} c_k(x) \mathrm{d}x = \lambda_k c_k(y). \tag{4.2}$$

If need be we make the further assumption that $(\xi_k)$ is a sequence of independent random variables. The sequence $(\lambda_k)$ is ordered so as to be nonincreasing. For $r(x, y)$, we use the exponential covariance function given by

$$r(x, y) = \nu \, \mathrm{e}^{-\frac{1}{b}|x_1 - y_1| - \frac{1}{b}|x_2 - y_2|}, \tag{4.3}$$

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 $(\lambda_k)$, a larger value producing faster decay. We set $b = 10$. Analytic solutions to (4.2) for this choice of covariance function are given in Ghanem & Spanos (1991).

For computational purposes, we need a finite-term expansion, so we curtail (4.1) to obtain

$$c(x, \omega) = c_0(x) + \sum_{k=1}^{m} \sqrt{\lambda_k} c_k(x) \xi_k(\omega). \tag{4.4}$$

From the modelling perspective, the replacement of the infinite expansion with the finite expansion is justified provided that $(\lambda_k)$ decays rapidly, which is discussed in Ghanem & Spanos (1991).

We will consider two cases for the distributions of the random variables $\xi_k$. In Section 4.2, we take $\xi_k$ to be uniformly distributed on $(-1, 1)$ with $c_0(x) = 10$ and $\nu = 1/3$. In Section 4.3, we take $\xi_k$ to be normally distributed with $c_0(x) = 1$ and $\nu = 0.01$.

### 4.2 *Multigrid for diffusion with uniform distributions*

We now let $\xi_k$ to be uniformly distributed on $(-1, 1)$. Therefore, $\Omega = (-1, 1)^m$ and $\mathrm{d}P = \mathrm{d}\omega/2^m$. We also set $\nu = 1/3$ and $c_0(x) = 10$. Applying the generalized polynomial chaos method as described in Section 2.3, the basis of $T$ will be the set of $m$-dimensional Legendre polynomials of degree $p$ or less.

For the triangulation of $D$, we will use a uniform mesh consisting of an underlying grid of $n \times n$ squares each of which is further subdivided into two equal triangles. A full V-cycle is used with a $2 \times 2$ 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 conversion rate of the multigrid algorithm is independent of $h$, $m$ and $p$.

### 4.3 *Multigrid for diffusion with normal distributions*

We now let $\xi_k$ to be normally distributed with zero mean and variance $\nu$. Now, we have $\Omega = \mathbb{R}^m$ and $\mathrm{d}P = \mathrm{e}^{-\omega^2/(2\nu)}/(2\pi\nu)^{m/2}\,\mathrm{d}\omega$. We take $c_0(x) = 1$. Applying the generalized polynomial chaos method as described in Section 2.3, the basis of $T$ will be the set of $m$-dimensional generalized Hermite polynomials of degree $p$ or less.

Note that the diffusion coefficient as defined in Section 4.1 will now fail to satisfy condition (2.6) no matter what the choice of $\nu$. However, we have reason to believe that the theory still applies. We give

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

| $n = 4$ | $m = 1$ | $m = 2$ | $m = 3$ | $m = 4$ |
|---|---|---|---|---|
| $p = 1$ | 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 2 *Number of iterations required for multigrid to converge for diffusion defined via normal distributions*

| $n = 4$ | $m = 1$ | $m = 2$ | $m = 3$ | $m = 4$ |
|---|---|---|---|---|
| $p = 1$ | 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 |

here only a heuristic argument. Given a sufficiently small variance, the probability of $c$ being outside 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 so as to ensure that $c$ satisfies (2.6), 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 take $\nu = 0.01$.

For the triangulation of $D$, we will use a uniform mesh consisting of an underlying grid of $n \times n$ squares each of which is further subdivided into two equal triangles. A full V-cycle is used with a $2 \times 2$ 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 conversion rate of the multigrid algorithm is independent of $h$, $m$ and $p$.

**Acknowledgements**

We would like to thank Catherine Powell for her discussion, insight and numerous readings of this paper. We would also like to thank two anonymous reviewers whose careful reading led to the correction of an early version of the paper. This work was supported in part by the U.S. National Science Foundation under grant DMS0208015 and the U.S. Department of Energy under grant DOEG0204ER25619.

## 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*. Cambridge: Cambridge University Press.

BRIGGS, W. L., HENSON, V. E. & MCCORMICK, S. F. (2000) *A Multigrid Tutorial*. SIAM.

ELMAN, H. C., SILVESTER, D. J. & WATHEN, A. J. (2005) *Finite Elements and Fast Iterative Solvers: with Applications in Incompressible Fluid Dynamics*. Oxford: Oxford University Press.

GHANEM, R. G. & SPANOS, P. D. (1991) *Stochastic Finite Elements: A Spectral Approach*. New York: Springer.

JOHNSON, C. (1987) *Numerical Solution of Partial Differential Equations by the Finite Element Method*. Cambridge: Cambridge University Press.

LE MAÎTRE, O. P., KNIO, O. M., DEBUSSCHERE, B. J., NAJM, H. N. & GHANEM, R. (2003) A multigrid solver for two-dimensional stochastic diffusion equations. *Comput. Methods Appl. Mech. Eng.*, **192**, 4723–4744.

PELLISSETTI, M. F. & GHANEM, R. G. (2000) Iterative solution of systems of linear equations arising in the context of stochastic finite elements. *Adv. Eng. Softw.*, **31**, 607–616.

SEYNAEVE, B., ROSSEEL, E., NICOLAÏ, B. & VANDEWALLE, S. (2007) Fourier mode analysis of multigrid methods for partial differential equations with random coefficients. *J. Comput. Phys.* (in press). Available at http://dx.doi.org/doi:10.1016/j.jcp.2006.12.011.

TREVES, F. (1967) *Topological Vector Spaces, Distributions and Kernels*. New York: Academic Press.

XIU, D. & KARNIADAKIS, G. E. (2002) Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos. *Comput. Methods Appl. Mech. Eng.*, **191**, 4927–4948.