Fourier mode analysis of multigrid methods for partial differential equations with random coefficients ☆

Bert Seynaeve [a], Eveline Rosseel [a], Bart Nicolaï [b], Stefan Vandewalle [a,*]

[a] *Katholieke Universiteit Leuven, Department of Computer Science, Celestijnenlaan 200A, B-3001 Leuven, Belgium*
[b] *Katholieke Universiteit Leuven, Division of Mechatronics, Biostatistics and Sensors, Willem de Croylaan 42, B-3001 Leuven, Belgium*

Dedicated to Professor Pieter Wesseling, a true scholar, whose work provides a continuous source of inspiration, on the occasion of his retirement.

Abstract

Partial differential equations with random coefficients appear for example in reliability problems and uncertainty propagation models. Various approaches exist for computing the stochastic characteristics of the solution of such a differential equation. In this paper, we consider the spectral expansion approach. This method transforms the continuous model into a large discrete algebraic system. We study the convergence properties of iterative methods for solving this discretized system. We consider one-level and multi-level methods. The classical Fourier mode analysis technique is extended towards the stochastic case. This is done by taking the eigenstructure into account of a certain matrix that depends on the random structure of the problem. We show how the convergence properties depend on the particulars of the algorithm, on the discretization parameters and on the stochastic characteristics of the model. Numerical results are added to illustrate some of our theoretical findings.
© 2007 Elsevier Inc. All rights reserved.

1. Introduction

Mathematical models for scientific and engineering problems often take the form of systems of ordinary or partial differential equations (PDEs). Such models usually involve certain parameters, for example the coefficients in the differential operator, the initial and boundary conditions, and the forcing function. Typically, all model parameters are assumed to be known exactly. However, in realistic situations this assumption often does not hold, as many of the parameters may have a random or stochastic character. More advanced models take this stochastic nature into account. Certain components of the system are then modelled as random variables, random fields or random processes. In that case one is interested in answering the so-called uncertainty propagation question: "Given the stochastic properties of the mathematical model, what are the stochastic characteristics of the solution?"

The numerical solution of PDEs with random coefficients can be calculated by using statistical or deterministic methods [1]. The most well-known statistical method is the Monte Carlo simulation method [2]. The stochastic model parameters are sampled repeatedly, reducing the problem to the solution of a deterministic PDE for each sample. The stochastic properties of the solution are subsequently determined by a statistical analysis of the obtained set of PDE solutions. Deterministic methods involve the direct treatment of the stochastic PDE, i.e., by directly computing certain stochastic properties of the solution. Examples are the perturbation method [3], the variance propagation algorithm [4], and the spectral expansion approach [5,6]. In the first two methods, deterministic equations are constructed for the expectation and for the variance of the solution, and for a covariance matrix. In this work, we will consider the third approach. It transforms a system of PDEs with random parameters into a deterministic linear system by means of a projection onto a certain set of random basis polynomials. The method allows the extraction, in a post-processing step, of any statistical information desired, such as moments of the solution or the probability of certain events related to the PDE.

Basically, the linear system generated by the spectral expansions approach, corresponds to the discretization matrix of a large system of coupled deterministic PDEs. The size of the system grows rapidly with the required spatial accuracy and with the required accuracy in the stochastic dimension. Different ways of enhancing the performance of the expansion method have been investigated in recent years. For example, by carefully choosing the polynomials to include in the random polynomial basis the size [7] and the complexity of the resulting system can be reduced [8,9]. By developing multi-level solution algorithms the cost of solving the system can be optimized [10,11].

In this work, we analyze the convergence behavior of multigrid applied to PDEs with random coefficients. This is done by extending the local Fourier analysis (LFA) technique for deterministic problems [12–14] to the stochastic case. The LFA accurately predicts the convergence factors observed in practice, and it provides valuable insights in the dependence of the convergence on the model characteristics, the discretization parameters and the algorithmic components. The paper is organized as follows. In Section 2 the model problem is presented. A brief summary is given of the different steps in the spectral expansion method, and the basic ideas underlying the considered iterative methods are explained. In Section 3 we suggest an alternative set of random basis polynomials. The use of these polynomials leads to a complete decoupling of the stochastic PDE into a set of deterministic PDEs of the same type. Although perhaps only of limited use from a practical point of view, this set of polynomials does prove to have interesting theoretical properties. Section 4 explains how the LFA can be extended to the stochastic case. A series of two-level convergence factors are derived for various multigrid components and strategies. The sharpness of the theoretical convergence factors are verified in Section 5 by extensive numerical experiments. Finally, in Section 6, we formulate some conclusions.

2. The model problem and its solution by stochastic expansion

2.1. A two-dimensional diffusion equation

We consider a two-dimensional steady-state diffusion equation where the diffusion coefficient $\alpha(\mathbf{x}, \omega)$ multiplying the y-derivative is a random field,

$$\frac{\partial^2 u(\mathbf{x},\omega)}{\partial x^2} + \alpha(\mathbf{x},\omega)\frac{\partial^2 u(\mathbf{x},\omega)}{\partial y^2} = \text{RHS}(\mathbf{x}), \tag{1}$$

with $\mathbf{x} = (x,y) \in \mathbf{D}$, the spatial domain, and $\omega \in \Omega$, a sample space. The sample space Ω is part of a complete probability space $(\Omega, \mathcal{F}, \mathcal{P})$, with $\mathcal{F}$ a σ-algebra and $\mathcal{P}$ a probability measure. The *random field* $\alpha(\mathbf{x},\omega)$ is defined as the mapping $\mathbf{D} \times \Omega \rightarrow \mathbb{R}$. For a given $\mathbf{x} \in \mathbf{D}$, $\alpha(\mathbf{x},\omega)$ is a *random variable* with respect to $(\Omega, \mathcal{F}, \mathcal{P})$. The boundary conditions and the forcing term are allowed to have random parameters as well. The convergence rates obtained through *local* Fourier analysis do not depend on the boundary conditions and the forcing term directly. Next, in Sections 2.2 and 2.3, we will detail the discretization of (1). In Section 2.4 we will recall the multigrid approach towards solving the discrete stochastic problem [10,11].

2.2. Discretization of the random part of the problem

The random field $\alpha(\mathbf{x},\omega)$ can be expanded into an infinite series of terms of decreasing importance, by means of a so-called Karhunen–Loève (KL) expansion [5,6,15],

$$\alpha(\mathbf{x},\omega) = \alpha_0(\mathbf{x}) + \sum_{i=1}^{\infty} \alpha_i(\mathbf{x})\xi_i(\omega). \tag{2}$$

The function α_0 is the mean of the random field; the functions α_i are suitably scaled (deterministic) eigenfunctions appearing in the spectral representation of the covariance function of the random field. The random coefficients $\xi_i(\omega)$ are uncorrelated random variables with zero mean, and are assumed to be linearly independent. For notational simplicity we will further on write the random variable $\xi_i(\omega)$ as ξ_i. For practical use the KL-expansion is truncated, say after the term in ξ_N, ignoring the least significant terms. This truncated KL-expansion will be denoted as $\alpha_{\text{KL}}(\mathbf{x},\omega)$. Any boundary conditions and/or forcing terms that are modelled as random fields are treated similarly. This may result in an additional set of independent random variables $\{\xi_i\}_{i=N+1,\ldots,N^*}$.

Next, we consider the discretization of the PDE solution $u(\mathbf{x},\omega)$ and we start with the random space dimension. Let $L^2(\Omega, \mathcal{F}, \mathcal{P})$ be a Hilbert space of square integrable functions of all N^* random variables ξ_i on $(\Omega, \mathcal{F}, \mathcal{P})$. We shall construct a finite dimensional subspace S of $L^2(\Omega, \mathcal{F}, \mathcal{P})$ defined through a set of Q basis functions $\{\Psi_l\}_{l=1,\ldots,Q}$ in the random variables $\xi_1, \ldots, \xi_{N^*}$. Let ξ denote a vector containing the random variables $\xi_1, \ldots, \xi_{N^*}$. The space S is equipped with an inner product defined by $\langle a(\xi)b(\xi) \rangle = \int_\Omega a(\xi)b(\xi)w(\xi)\mathrm{d}\xi$, with $w(\xi)$ denoting the probability density corresponding to ξ. This inner product actually corresponds to the expectation of the product of its arguments. Several approaches have been proposed to construct S, e.g. [5,8,16,17]. Here, we shall employ an orthonormal basis of multivariate polynomials Ψ_l that are globally defined in each random variable ξ_i. These multivariate polynomials are built using series of univariate polynomials $\{\varphi_n\}_{n=1,2,\ldots}$ of degree n in ξ_i and orthonormal w.r.t. the probability measure corresponding to ξ_i. That is,

$$\langle \varphi_m \varphi_n \rangle_i := \int \varphi_m(\xi_i)\varphi_n(\xi_i)w(\xi_i)\,\mathrm{d}\xi_i = \delta_{mn},$$

with δ_{mn} the Kronecker delta. In particular, for some N^*-tuple $(m_1, \ldots, m_{N^*})$ of positive integers, we have

$$\Psi_l(\xi_1, \ldots, \xi_{N^*}) = \prod_{i=1}^{N^*} \varphi_{m_i}(\xi_i). \tag{3}$$

Two criteria are often considered to determine basis functions of the form (3). One may limit the total degree of the polynomial to a given value P, i.e., $\sum_{i=1}^{N^*} m_i \leqslant P$. The total number of basis functions, Q, is then given by $(N^* + P)!/N^*!P!$ [10]. Alternatively, one may limit the degrees of the univariate factors separately, i.e. $m_i \leqslant n_i, i = 1, \ldots, N^*$ for a given set of n_i-values. In this case $Q = \prod_{i=1}^{N^*}(n_i + 1)$ [9].

Using the first criterion, a so-called Generalized Polynomial Chaos basis [16,1] can be constructed. The univariate polynomials are chosen from the Wiener–Askey scheme according to the probability distributions of the uncorrelated random variables in (2). In case of a set of Gaussian variables ξ, this procedure results in a basis of multivariate Hermite polynomials. This basis is also called the Polynomial Chaos (PC) basis and is used in the original spectral approach of Ghanem [5,6]. The second criterion can be used to create a double orthogonal polynomial basis, as proposed by Babuška [8].

Having specified an appropriate random basis, the solution $u(\mathbf{x}, \omega)$ is approximated by a linear combination of basis functions with deterministic coefficients $u_l(\mathbf{x})$. When the basis functions are collected together in the column vector Ψ and the coefficients in a column vector $\mathbf{u}(\mathbf{x})$, we can write

$$u(\mathbf{x}, \omega) = \sum_{l=1}^{Q} u_l(\mathbf{x}) \Psi_l(\xi_1, \ldots, \xi_{N^*}) = \Psi^{\mathrm{T}} \mathbf{u}(\mathbf{x}). \tag{4}$$

2.3. Reduction to a fully discrete deterministic linear system

We replace the random field $\alpha(\mathbf{x}, \omega)$ by the truncated expansion $\alpha_{\mathrm{KL}}(\mathbf{x}, \omega)$ and look for a solution of the form (4) such that the residual is minimized in some appropriate sense. Typically one computes the solution by using a least squares approach, requiring orthogonality of the residual to the space determined by the chosen Polynomial Chaos basis. Taking the linearity of the inner product into account gives

$$\langle \Psi \Psi^{\mathrm{T}} \rangle \frac{\partial^2 \mathbf{u}(\mathbf{x})}{\partial x^2} + \langle \alpha_{\mathrm{KL}}(\mathbf{x}, \omega) \Psi \Psi^{\mathrm{T}} \rangle \frac{\partial^2 \mathbf{u}(\mathbf{x})}{\partial y^2} = \langle \mathrm{RHS}(\mathbf{x}) \Psi \rangle.$$

The real symmetric matrix $\langle \alpha_{\mathrm{KL}}(\mathbf{x}, \omega) \Psi \Psi^{\mathrm{T}} \rangle$ is independent of the solution and will be denoted as $\mathscr{M}(\mathbf{x})$,

$$\mathscr{M}(\mathbf{x}) := \langle \alpha_{\mathrm{KL}}(\mathbf{x}, \omega) \Psi \Psi^{\mathrm{T}} \rangle = \alpha_0(\mathbf{x}) \langle \Psi \Psi^{\mathrm{T}} \rangle + \sum_{i=1}^{N} \alpha_i(\mathbf{x}) \langle \xi_i \Psi \Psi^{\mathrm{T}} \rangle. \tag{5}$$

For Gaussian random variables this matrix is easily determined [18]. Using the orthonormality of the basis functions, the above equation can further be simplified to

$$\frac{\partial^2 \mathbf{u}(\mathbf{x})}{\partial x^2} + \mathscr{M}(\mathbf{x}) \frac{\partial^2 \mathbf{u}(\mathbf{x})}{\partial y^2} = \langle \mathrm{RHS}(\mathbf{x}) \Psi \rangle. \tag{6}$$

Conventional spatial discretization methods can be applied to reduce this coupled system of PDEs to an algebraic linear system. We shall use a standard five-point finite difference scheme on a rectangular grid $G_h = \{(jh, kh)\}_{j,k \in \mathbb{Z}}$ with spacing h in x- and y-directions. The value $\mathrm{RHS}(jh, kh)$, the random variable $\alpha(jh, kh, \omega)$ and the corresponding matrix $\mathscr{M}(jh, kh)$ will be denoted by $\mathrm{RHS}_{j,k}$, $\alpha_{j,k}$ and $\mathscr{M}_{j,k}$. The discrete approximation to $\mathbf{u}(\mathbf{x})$ evaluated at point (jh, kh) is denoted as $\mathbf{u}_{j,k}$. Applying the spatial discretization to the system of deterministic PDEs (6) leads to

$$(\mathbf{u}_{j-1,k} - 2\mathbf{u}_{j,k} + \mathbf{u}_{j+1,k}) + \mathscr{M}_{j,k}(\mathbf{u}_{j,k-1} - 2\mathbf{u}_{j,k} + \mathbf{u}_{j,k+1}) = h^2 \langle \mathrm{RHS}_{j,k} \Psi \rangle. \tag{7}$$

When the equations are collected over all grid points a classical linear system results,

$$L_h \mathbf{u}_h = \mathbf{f}_h. \tag{8}$$

The dimension of the system equals the number of spatial grid points multiplied by Q. When the unknowns are ordered pointwise, that is, all unknowns at a particular grid point are grouped together, then L_h is a block matrix with blocks of size $Q \times Q$. Its block structure is identical to the classical five-point star matrix structure. When the unknowns are ordered variable-wise, then L_h is a $Q \times Q$ block matrix with blocks having the size and structure of a discretized scalar PDE.

2.4. Iterative solution of the fully discrete system

We consider iterative methods for solving (8). For example, in iteration step $m + 1$ of a pointwise Gauss–Seidel (GS-LEX) scheme the new value $u_{j,k}^{(m+1)} = \Psi^{\mathrm{T}} \mathbf{u}_{j,k}^{(m+1)}$ in grid point (jh, kh) is determined from,

$$2(I + \mathscr{M}_{j,k}) \mathbf{u}_{j,k}^{(m+1)} = (\mathbf{u}_{j-1,k}^{(m+1)} + \mathbf{u}_{j+1,k}^{(m)}) + \mathscr{M}_{j,k}(\mathbf{u}_{j,k-1}^{(m+1)} + \mathbf{u}_{j,k+1}^{(m)}) - h^2 \langle \mathrm{RHS}_{j,k} \Psi \rangle, \tag{9}$$

assuming that $I + \mathscr{M}_{j,k}$ is an invertible matrix. We shall use the previous scheme as a smoother in a two-level iterative method. A typical two-grid iteration scheme $TG(v_1, v_2)$, which updates the approximation $\mathbf{u}_h$, can be formulated as follows [14]:

- pre-smoothing: smooth v_1 times $\mathbf{u}_h$, i.e., $\mathbf{u}_h \leftarrow S_h^{v_1} \mathbf{u}_h$
- coarse grid correction:
    - calculate residual: $\mathbf{r}_h = \mathbf{f}_h - L_h \mathbf{u}_h$
    - restrict residual: $\mathbf{f}_{2h} = I_h^{2h} \mathbf{r}_h$
    - solve for coarse grid correction: $L_{2h} \mathbf{e}_{2h} = \mathbf{f}_{2h}$
    - correct: $\mathbf{u}_h \leftarrow \mathbf{u}_h + I_{2h}^h \mathbf{e}_{2h}$
- post-smoothing: smooth v_2 times $\mathbf{u}_h$, i.e., $\mathbf{u}_h \leftarrow S_h^{v_2} \mathbf{u}_h$

Next to the *fine* grid G_h also a *coarse* grid G_{2h} is used, for example obtained by doubling the grid spacing. Residuals ($\mathbf{r}_h$) and corrections ($\mathbf{e}_{2h}$) are transferred from one grid to another by using restriction (I_h^{2h}) and prolongation (I_{2h}^h) operators. By using a recursion the two-level algorithm can be extended towards a true multilevel method that uses a hierarchy of coarser and coarser grids.

3. A polynomial basis that decouples the stochastic PDE

It has been shown in [8,9] that the use of a certain carefully constructed set of orthogonal polynomials $\{\Psi_l\}_{l=1,\ldots,Q}$ leads to a complete decoupling of system (8). Here we suggest an alternative (and easier) derivation of those polynomials. This decoupling method is of practical importance mainly for low degree polynomials and a limited number of random variables. Otherwise, the method is computationally too costly. For theoretical purposes, however, a study of this algorithm is interesting as it provides with little effort a set of convergence factor bounds for the more general case.

As before, the Q basis functions are generated as products of univariate polynomials

$$\Psi'_{l_1,\ldots,l_{N^*}}(\xi_1,\ldots,\xi_{N^*}) = \prod_{i=1}^{N^*} \psi_{n_i l_i}(\xi_i), \quad l_i \in \{0,\ldots,n_i\}, \tag{10}$$

i.e., each of the $\prod_{i=1}^{N^*}(n_i + 1)$ basis functions consist of N^* univariate factors. In the classical Polynomial Chaos procedure these univariate polynomials constitute a set of orthogonal polynomials of increasing degree. Here, we consider an alternative set. The univariate polynomials in ξ_i ($i = 1,\ldots,N$), i.e., in the random variables generated by the KL-expansion (2), are selected to be exactly of a specified degree n_i, and are given by

$$\psi_{n_i l_i}(\xi_i) = (-1)^{n_i - l_i} \rho_{n_i l_i} \prod_{r \neq l_i} (\xi_i - \zeta_{n_i r}), \qquad l_i = 0,\ldots,n_i. \tag{11}$$

The values $\zeta_{n_i r}$, for $r = 0,\ldots,n_i$, are the roots of the orthogonal polynomial φ_{n_i+1} of degree $n_i + 1$, i.e.,

$$\varphi_{n_i+1}(\xi_i) = \rho_{n_i+1}^* \prod_{r=0}^{n_i} (\xi_i - \zeta_{n_i r}).$$

The polynomial φ_{n_i+1} corresponds to a typical univariate factor used in the construction of a classical Generalized Polynomial Chaos basis. The positive normalization constants $\rho_{n_i l_i}$ and $\rho_{n_i+1}^*$ are chosen so that $\langle \psi_{n_i l_i}^2 \rangle_i = 1$ and $\langle \varphi_{n_i+1}^2 \rangle_i = 1$. In case of a standard Gaussian random variable ξ_i, the values $\rho_{n_i l_i}^*$ and $\rho_{n_i l_i}$ are equal to $1/\sqrt{(n_i + 1)!}$.

For the univariate polynomials in ξ_i with $i = N + 1,\ldots,N^*$, both this alternative procedure and the usual Generalized Polynomial Chaos procedure with polynomials of increasing degree can be applied. In both cases these polynomials will be denoted by $\psi_{n_i l_i}$, $l_i = 0,\ldots,n_i$, $i = N + 1,\ldots,N^*$. Let the column vector Ψ represent the (Generalized) Polynomial Chaos basis functions, then we collect the alternative basis functions in a column vector Ψ'.

The alternative *univariate* polynomials satisfy some interesting properties.

Property 3.1.

(a) $\langle \psi_{n_i l_i} \psi_{n_i m_i} \rangle_i = \delta_{l_i m_i}, \quad i = 1,\ldots,N^*$;

(b) $\langle \xi_i \psi_{n_i l_i} \psi_{n_i m_i} \rangle_i = \zeta_{n_i l_i} \delta_{l_i m_i}, \quad i = 1,\ldots,N$.

Proof. The results for the case $l_i \neq m_i$ follow immediately from (11) by taking into account that φ_{n_i+1} is orthogonal to every polynomial in ξ_i of degree at most n_i. For example, for property (b),

$$\langle \xi_i \psi_{n_i l_i} \psi_{n_i m_i} \rangle = (-1)^{l_i + m_i} \frac{\rho_{n_i l_i} \rho_{n_i m_i}}{\rho_{n_i+1}^*} \left\langle \varphi_{n_i+1} \left(\xi_i \prod_{r \neq l_i, m_i} (\xi_i - \zeta_{n_i r}) \right) \right\rangle = 0.$$

For $l_i = m_i$, property (a) follows from the normalization condition. The proof for $l_i = m_i$ of property (b) is based on the following identities:

$$\langle \xi_i \psi_{n_i l_i}^2 \rangle - \zeta_{n_i l_i} \langle \psi_{n_i l_i}^2 \rangle = \langle (\xi_i - \zeta_{n_i l_i}) \psi_{n_i l_i}^2 \rangle = \frac{(-1)^{n_i - l_i} \rho_{n_i l_i}}{\rho_{n_i+1}^*} \langle \varphi_{n_i+1} \psi_{n_i l_i} \rangle = 0. \quad \square$$

We can derive some analogous properties for the new *multivariate* polynomials (10), under the assumption that $\{\xi_i\}_{i=1,\dots,N^*}$ are independent random variables.

Property 3.2.

(a) $\langle \Psi'_{l_1,\dots,l_{N^*}} \Psi'_{m_1,\dots,m_{N^*}} \rangle = \delta_{(l_1,\dots,l_{N^*})(m_1,\dots,m_{N^*})}$;

(b) $\langle \xi_i \Psi'_{l_1,\dots,l_{N^*}} \Psi'_{m_1,\dots,m_{N^*}} \rangle = \zeta_{n_i l_i} \delta_{(l_1,\dots,l_{N^*})(m_1,\dots,m_{N^*})}, \quad \forall i \in \{1,\dots,N\}.$

These properties will allow us to discuss the performance of certain iterative methods in terms of the roots $\{\zeta_{n_i l_i}\}_{l_i=0,\dots,n_i; i=1,\dots,N}$ (see Section 4).

The random basis functions (10) can be used to decouple the model problem (1). As before, we replace the random field α by its truncated Karhunen–Loève expansion, and we write the approximate solution as a linear combination of the random basis functions,

$$u(\mathbf{x}, \omega) = \sum_{l_1=0}^{n_1} \cdots \sum_{l_{N^*}=0}^{n_{N^*}} u'_{l_1,\dots,l_{N^*}}(\mathbf{x}) \Psi'_{l_1,\dots,l_{N^*}}(\xi_1,\dots,\xi_{N^*}),$$

where the deterministic coefficient functions $u'_{l_1,\dots,l_{N^*}}(\mathbf{x})$ are to be determined. Using the properties derived in this section, it can easily be checked that the matrix $\mathcal{M}(\mathbf{x})$ from Eq. (6) becomes a diagonal matrix. More directly, the result can be derived from an orthogonalization of the residual of the model problem w.r.t. any of the basis functions $\Psi'_{m_1,\dots,m_{N^*}}$. For example, orthogonality w.r.t. ξ leads to an equation of the form

$$\frac{\partial^2 u'_{m_1,\dots,m_{N^*}}}{\partial x^2} + \left(\alpha_0(\mathbf{x}) + \sum_{i=1}^{N} \alpha_i(\mathbf{x}) \zeta_{n_i m_i} \right) \frac{\partial^2 u'_{m_1,\dots,m_{N^*}}}{\partial y^2} = \langle \text{RHS}(\mathbf{x}) \Psi'_{m_1,\dots,m_{N^*}} \rangle. \tag{12}$$

This is a deterministic PDE in a single unknown variable $u'_{m_1,\dots,m_{N^*}}(\mathbf{x})$. A similar PDE is obtained for every $(m_1,\dots,m_{N^*}) \in \{0,1,\dots,n_1\} \times \cdots \times \{0,1,\dots,n_{N^*}\}$. As such the stochastic problem is decomposed into $\prod_{i=1}^{N^*}(n_i+1)$ deterministic PDEs. Note that if the number of Karhunen–Loève terms increases or if the degrees n_i are increased, the number of deterministic PDEs rapidly becomes prohibitively large for practical purposes.

The roots $\{\zeta_{n_i m_i}\}_{m_i=0,\dots,n_i}$ of the polynomials φ_{n_i+1} play an important role in this decomposition. They determine the $\prod_{i=1}^{N}(n_i+1)$ realizations of the random field α for which deterministic versions of the model problem have to be solved. For each of those realizations the PDE must be solved for $\prod_{i=N+1}^{N^*}(n_i+1)$ different right hand sides.

4. Local Fourier mode analysis

Fourier mode analysis is a powerful tool to analyze multigrid methods quantitatively [12,19]. It is frequently used to optimize the different algorithmic components for a specific problem. It was originally developed for linear discrete operators, but it can also be applied to more general discrete operators by linearizing the operator locally and by 'freezing' its coefficients to a constant value. In this section we apply the LFA to study multigrid for PDEs with random coefficients. In Section 4.1 we analyze the properties of a number of one-level iterative methods that are frequently used as smoothers inside a multigrid algorithm. In Section 4.2

the two-grid algorithm is studied. There, we consider the case where the classical (Generalized) Polynomial Chaos functions are used for discretization. In Section 4.3 we treat the special case of basis functions that decouple the stochastic problem into a set of deterministic ones, as explained in Section 3.

4.1. Analysis of one-level iterative methods

We shall work out the details of the local Fourier analysis for the GS-LEX method (9). An infinite grid is assumed to eliminate the effect of boundary conditions. We define the errors in the coefficient vectors as

$$\mathbf{e}_{j,k}^{(m)} = \mathbf{u}_{j,k} - \mathbf{u}_{j,k}^{(m)}.$$

At every grid point the error is a Q-vector. From 7 and 9 it follows that

$$2(I + \mathcal{M}_{j,k})\mathbf{e}_{j,k}^{(m+1)} - \mathbf{e}_{j-1,k}^{(m+1)} - \mathcal{M}_{j,k}\mathbf{e}_{j,k-1}^{(m+1)} = \mathbf{e}_{j+1,k}^{(m)} + \mathcal{M}_{j,k}\mathbf{e}_{j,k+1}^{(m)}. \tag{13}$$

In the case of a variable coefficients problem, we shall perform the LFA to the problem by freezing the random field α to its value $\alpha_{j,k}$ in the considered grid point (jh, kh) [20]. In the model problem (1), this corresponds to the replacement of the random field $\alpha(\mathbf{x}, \omega)$ by a random variable $\alpha(\omega)$. As such, $\mathcal{M}_{j,k}$ can be considered to be a fixed known matrix, which will be denoted as $\mathcal{M}$. It is a real and symmetric matrix, but not necessarily positive. Theoretically, this might prevent GS-LEX to be well-defined or to converge. In practical situations however, this will not turn out to be an issue (see Section 5).

We shall decompose the error into a sum of exponential Fourier modes of the form

$$\mathbf{e}_{j,k}(\theta, \mathbf{z}) = \exp(\iota(j\theta_x + k\theta_y))\mathbf{z}, \tag{14}$$

where ι represents the imaginary unit and with $\theta := (\theta_x, \theta_y) \in [-\pi, \pi)^2$. The Q-vector $\mathbf{z} = [A_1, A_2, \ldots, A_Q]^{\mathrm{T}}$ contains the amplitudes of the Fourier mode in the random functions $\{\Psi_l\}_{l=1,\ldots,Q}$. Application of (13) to mode $\mathbf{e}_{j,k}(\theta, \mathbf{z}^{(m)})$, e.g., one of the components of the error at the mth iteration, generates mode $\mathbf{e}_{j,k}(\theta, \mathbf{z}^{(m+1)})$ with

$$((2 - \exp(-\iota\theta_x))I + (2 - \exp(-\iota\theta_y))\mathcal{M})\mathbf{z}^{(m+1)} = (\exp(\iota\theta_x)I + \exp(\iota\theta_y)\mathcal{M})\mathbf{z}^{(m)}.$$

The convergence of the above iteration depends on the spectrum of $\mathcal{M}$, denoted as $\sigma(\mathcal{M})$, i.e., on the random structure of the model problem. Let $\mathbf{z}_q$ be an eigenvector of $\mathcal{M}$ with corresponding eigenvalue λ_q, and set $\mathbf{z}^{(m)} = \mathbf{z}_q$. Then, we immediately find that $\mathbf{z}^{(m+1)} = \widehat{S}(\theta, \lambda_q)\mathbf{z}^{(m)}$ with

$$\widehat{S}(\theta, \lambda_q) = \frac{\exp(\iota\theta_x) + \exp(\iota\theta_y)\lambda_q}{(2 - \exp(-\iota\theta_x)) + (2 - \exp(-\iota\theta_y))\lambda_q}. \tag{15}$$
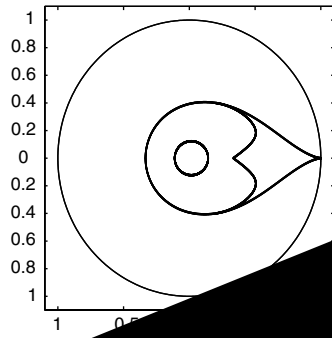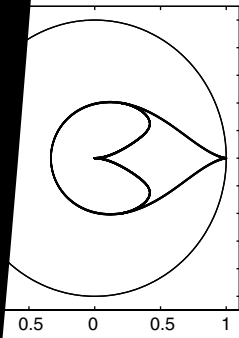
We can decompose the iteration error into a sum of independent components of the form $\exp(\iota(j\theta_x + k\theta_y))\mathbf{z}_q$, which are eigenvectors of the GS-LEX iteration operator, with $\widehat{S}(\theta, \lambda_q)$ the corresponding eigenvalues. The distribution of the eigenvalues of the GS-LEX iteration operator for several values of λ_q is shown in Fig. 1a–c. The eigenvalues lie inside the unit circle for every $\lambda_q > 0$ and every $\theta \in [-\pi, \pi)^2$.

In Fig. 1d and e, we compare the eigenvalues of a stochastic problem, with a related deterministic problem, namely the problem with the mean of $\alpha(\omega)$ as deterministic coefficient. We observe that each eigenvalue in the deterministic case is split into Q eigenvalues in the stochastic case. This can be explained intuitively as follows. Consider L_{xx} and L_{yy}, the spatial discretizations of the operators $\frac{\partial^2}{\partial x^2}$ and $\frac{\partial^2}{\partial y^2}$. We can rewrite the GS-LEX iteration operator S in tensor product notation as

$$S = (I_Q \otimes L_{xx}^+ + \mathcal{M} \otimes L_{yy}^+)^{-1}(I_Q \otimes L_{xx}^- + \mathcal{M} \otimes L_{yy}^-), \tag{16}$$

with the matrix splittings $L_{xx} = L_{xx}^+ - L_{xx}^-$, $L_{yy} = L_{yy}^+ - L_{yy}^-$ and L_{xx}^+, L_{yy}^+ the lower triangular parts of L_{xx} and L_{yy}, respectively. Based on the eigenvalue-eigenvector decomposition of the matrix $\mathcal{M}$, i.e., $\mathcal{M} = V\Lambda V^{\mathrm{T}}$, we can apply a similarity transformation to (16) and obtain

$$\sigma(S) = \sigma(V^{\mathrm{T}} \otimes I_h \cdot S \cdot V \otimes I_h) = \sigma(I_Q \otimes L_{xx}^+ + \Lambda \otimes L_{yy}^+)^{-1}(I_Q \otimes L_{xx}^- + \Lambda \otimes L_{yy}^-)$$

$$= \bigcup_{q=1}^{Q} \sigma(L_{xx}^+ + \lambda_q L_{yy}^+)^{-1}(L_{xx}^- + \lambda_q L_{yy}^-).$$

max (18)
x=(jh,k...

...nalytical e... ...15). The LFA spectr... ...given by ...= 1 sinc... ...implies that GS-LEX ...nverge at ...nfinite... ...s, GS-LEX does converg... ...situations, the def... ...alues and the eigenfunctions... ...S-LEX oper-ator... ...g Dirichlet boundary conditions, ...ble estimates fo... ...n be obtained from the GS-LEX eige... ...s on an infinite

grid (15) by sampling θ values from a finite grid and excluding the zero frequencies ($\theta_x = 0$ or $\theta_y = 0$) from the analysis [14]. Based on these heuristics, the definition of the spectral radius (17) applies to the finite domain, $\theta \in \Theta^D := \{\theta | \theta = (\theta_x, \theta_y) \text{ with } \theta_d = \pi m_d / n_d, 1 \leqslant m_d \leqslant n_d - 1, n_d = 1/h(d = x, y)\}$. The largest eigenvalue in modulus then occurs at $(\theta_x, \theta_y) = (\pi h, \pi h)$ with $\widehat{S}((\pi h, \pi h), \lambda_q) = \frac{\exp(\iota \pi h)}{2 - \exp(-\iota \pi h)}$; and thus the spectral radius is given by

$$\rho = 1 - \pi^2 h^2 + \mathcal{O}(h^4). \tag{19}$$

Since this expression is independent of λ_q, we also have

$$\bar{\rho} = 1 - \pi^2 h^2 + \mathcal{O}(h^4). \tag{20}$$

The smoothing factor and the spectral radius of GS-LEX are shown as a function of λ_q in Fig. 4. For a grid with different grid spacing in the two directions, (h_x, h_y), the eigenvalues of the GS-LEX iteration operator become

$$\widehat{S}_{h_x, h_y}(\theta, \lambda_q) = \frac{h_y^2 \exp(\iota \theta_x) + h_x^2 \exp(\iota \theta_y) \lambda_q}{h_y^2 (2 - \exp(-\iota \theta_x)) + h_x^2 (2 - \exp(-\iota \theta_y)) \lambda_q}.$$

In this case, the modulus of the largest eigenvalue does depend on λ_q. The spectral radius is given by $\rho = 1 - \pi^2 \frac{1 + \lambda_q}{h_y^2 + h_x^2 \lambda_q} h_x^2 h_y^2 + \mathcal{O}(h_x^4 h_y^4)$.

The LFA can also be performed for other classical iterative methods. In Table 1, we consider the red–black version of the point Gauss–Seidel method, the x-line lexicographic Gauss–Seidel scheme, and the pointwise ω-Jacobi scheme, also known as damped or weighted Jacobi. The table shows for each method the formula for the eigenvalues as a function of θ and λ_q, and the smoothing factor as a function of λ_q. The analysis of red–black Gauss–Seidel is somewhat more elaborate than the analysis of lexicographic Gauss–Seidel since the red–black updating scheme results in an intermixing of the high frequency and low frequency Fourier modes. This necessitates a generalized definition of the smoothing factor in which the smoothing factor depends on the number of fine-grid relaxation sweeps, see [19,21]. In Table 1, the results for a single smoothing step per cycle are given. The analysis of red–black Gauss–Seidel can straightforwardly be extended to the analysis of red–black successive overrelaxation (SOR) iterations using the results in [22, Th. 2.1]. If $s_{\text{GS–RB}}$ represents the smoothing factor of red–black Gauss–Seidel, then a good approximation to the optimal SOR damping parameter and the corresponding smoothing factor are given by $\omega_{\text{SOR,opt}}$ and $s_{\text{SOR,opt}}$:

Table 1
Eigenvalues and smoothing factor for iterative methods applied to the (frozen) discrete differential operator (7) (GS-LEX: lexicographic Gauss–Seidel, GS-RB: pointwise red–black Gauss–Seidel, x-line GS-LEX: line GS-LEX in the x-direction and Ω-Jacobi: weighted Jacobi)

	Eigenvalues	Smoothing factor
Point GS-LEX	$\dfrac{\exp(\iota \theta_x) + \exp(\iota \theta_y) \lambda_q}{(2 - \exp(-\iota \theta_x)) + (2 - \exp(-\iota \theta_y)) \lambda_q}$	$\sqrt{\dfrac{\beta(\lambda_q^2 + 1) + 4\lambda_q^2 \sqrt{\beta}}{\beta(5\lambda_q^2 + 8\lambda_q + 5) - 4\sqrt{\beta}}}$ if $0 < \lambda_q \leqslant 1$
		$\sqrt{\dfrac{\gamma(\lambda_q^2 + 1) + 4\lambda_q \sqrt{\gamma}}{\gamma(5\lambda_q^2 + 8\lambda_q + 5) - 4\lambda_q^3 \sqrt{\gamma}}}$ if $\lambda_q > 1$
		With $\beta = 5\lambda_q^2 - 2\lambda_q + 1$ and $\gamma = \lambda_q^2 - 2\lambda_q + 5$
Point GS-RB	0 or $\left(\dfrac{\cos(\theta_x) + \cos(\theta_y) \lambda_q}{1 + \lambda_q}\right)^2$	$\begin{cases} \dfrac{1}{(1 + \lambda_q)^2} & \text{if } 0 < \lambda_q \leqslant 1 \\ \dfrac{\lambda_q^2}{(1 + \lambda_q)^2} & \text{if } \lambda_q > 1 \end{cases}$
x-line GS-LEX	$\dfrac{\exp(\iota \theta_y) \lambda_q}{(2(1 - \cos(\theta_x))) + (2 - \exp(-\iota \theta_y)) \lambda_q}$	$\begin{cases} \dfrac{\sqrt{5}}{5} & \text{if } 0 < \lambda_q \leqslant \dfrac{1 + \sqrt{5}}{2} \\ \dfrac{\lambda_q}{2 + \lambda_q} & \text{if } \lambda_q > \dfrac{1 + \sqrt{5}}{2} \end{cases}$
ω-Jacobi	$1 - \omega + \omega \dfrac{\cos(\theta_x) + \cos(\theta_y) \lambda_q}{1 + \lambda_q}$	$\begin{cases} \dfrac{2 + \lambda_q}{2 + 3\lambda_q} & \text{if } 0 < \lambda_q \leqslant 1 \\ \dfrac{1 + 2\lambda_q}{3 + 2\lambda_q} & \text{if } \lambda_q > 1 \end{cases}$

$$\omega_{\text{SOR,opt}} = \frac{2}{1 + \sqrt{1 - s_{\text{GS-RB}}}} \quad \text{and}$$

The smoothing factor for ω-Jacobi is ba[...] optimal [...]
smoothing factor. Some calculations show [...] ptimal v[...]

$$\omega_{\text{opt},\lambda_q} = \frac{2 + 2\lambda_q}{2 + 3\lambda_q} \quad \text{if } 0 < \lambda_q \leqslant 1 \quad \text{and}$$

As illustrated in Table 1, for $\lambda_q = 1$, the smo[...] rs coin[...]
the central discretization of the two-dimensi[...] ian. A[...]
that shows the smoothing properties of the [...] rative

4.2. Local Fourier analysis for the general app

Application of the (frozen) discrete different[...] cha[...]
vector-valued Fourier mode of the form (14) l[...]

$$L_h \mathbf{e}_{j,k}(\theta, \mathbf{z}) = \frac{1}{h^2}\left(\mathbf{e}_{j-1,k}(\theta, \mathbf{z}) - 2\mathbf{e}_{j,k}(\theta, \mathbf{z}) + \mathbf{e}_{\cdots}\right)$$

$$= -\frac{4}{h^2}\exp(\iota(j\theta_x + k\theta_y))\left(\sin^2\left(\frac{\theta_{\cdots}}{2}\right) \cdots \right)$$

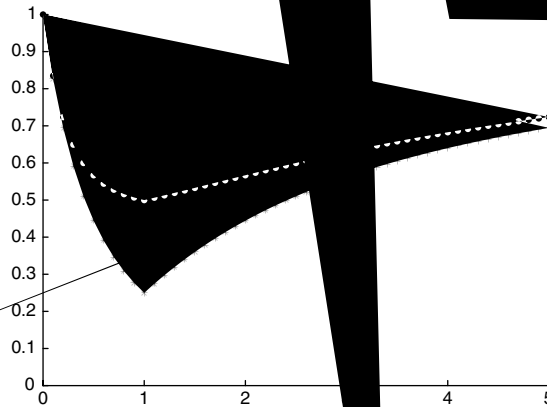If $\mathbf{z}$ is selected to be one of the eigenvectors $\mathbf{z}_q$ of [...]
plifies to

$$L_h \mathbf{e}_{j,k}(\theta, \mathbf{z}_q) = \widehat{L}_h(\theta, \lambda_q)\mathbf{e}_{j,k}(\theta, \mathbf{z}_q)$$

with

$$\widehat{L}_h(\theta, \lambda_q) = -\frac{4}{h^2}\left(\sin^2\left(\frac{\theta_x}{2}\right) + \sin^2\left(\frac{\theta_y}{2}\right)\lambda_q\right).$$

Hence, the Fourier mode $\mathbf{e}_{j,k}(\theta, \mathbf{z}_q)$ is an eigenfunction [...] ozen[...]
eigenfunction for the operator L_{2h} on grid G_{2h} with e[...]

$$\widetilde{L}_{2h}(\theta, \lambda_q) = -\frac{1}{h^2}(\sin^2(\theta_x) + \sin^2(\theta_y)\lambda_q).$$

The Fourier mode $\mathbf{e}_{j,k}(\theta, \mathbf{z}_q)$ is not an eigenfunction of the restriction and prolongation operator. In order to analyze those operators, one divides the Fourier space into four-dimensional subspaces of harmonics, $\mathscr{E}(\theta, \mathbf{z}_q) := \mathrm{span}[e(\theta^1, \mathbf{z}_q)\, e(\theta^2, \mathbf{z}_q)\, e(\theta^3, \mathbf{z}_q)\, e(\theta^4, \mathbf{z}_q)]$, for a given $(\theta_x, \theta_y) \in [-\frac{\pi}{2}, \frac{\pi}{2}]^2$ with

$$\theta^1 = (\theta_x, \theta_y), \quad \theta^2 = (\theta_x, \theta_y - \mathrm{sign}(\theta_y)\pi),$$
$$\theta^3 = (\theta_x - \mathrm{sign}(\theta_x)\pi, \theta_y), \quad \theta^4 = (\theta_x - \mathrm{sign}(\theta_x)\pi, \theta_y - \mathrm{sign}(\theta_y)\pi).$$

These spaces of harmonics are invariant under the fine and coarse grid discrete differential operators, and under certain smoothing operators. The action of such an operator on an element of such a space can be described by a four by four matrix, the so-called symbol of that operator. For example, the symbol of the pointwise lexicographic Gauss–Seidel method, denoted as $\widetilde{S}(\theta, \lambda_q)$, is given by

$$\widetilde{S}(\theta, \lambda_q) = \mathrm{diag}(\widehat{S}(\theta^1, \lambda_q), \widehat{S}(\theta^2, \lambda_q), \widehat{S}(\theta^3, \lambda_q), \widehat{S}(\theta^4, \lambda_q)),$$

with $\widehat{S}(\theta, \lambda_q)$ defined by (15). The Fourier symbol of L_h becomes

$$\widetilde{L}_h(\theta, \lambda_q) = \mathrm{diag}(\widehat{L}_h(\theta^1, \lambda_q), \widehat{L}_h(\theta^2, \lambda_q), \widehat{L}_h(\theta^3, \lambda_q), \widehat{L}_h(\theta^4, \lambda_q)).$$

The prolongation operator maps the mode $\mathbf{e}_{j,k}(2\theta, \mathbf{z}_q)$ on G_{2h} onto the space $\mathscr{E}(\theta, \mathbf{z}_q)$ [19]. It is characterized by the symbol $\widetilde{I}_{2h}^h(\theta)$, which for bilinear interpolation is given by

$$\widetilde{I}_{2h}^h(\theta) = \frac{1}{4}\begin{bmatrix}(1 + \cos(\theta_x))(1 + \cos(\theta_y)) \\ (1 + \cos(\theta_x))(1 - \cos(\theta_y)) \\ (1 - \cos(\theta_x))(1 + \cos(\theta_y)) \\ (1 - \cos(\theta_x))(1 - \cos(\theta_y))\end{bmatrix}.$$

Using standard coarsening, the restriction maps the space $\mathscr{E}(\theta, \mathbf{z}_q)$ onto the single mode $\mathbf{e}(\theta^1, \mathbf{z}_q)$. The corresponding Fourier representation satisfies $\widetilde{I}_h^{2h}(\theta) = (\widetilde{I}_{2h}^h(\theta))^{\mathrm{T}}$.

To conclude, the action of this two-grid operator, corresponding to the frozen differential operator on the space $\mathscr{E}(\theta, \mathbf{z}_q)$, is characterized by the following symbol:

$$\widetilde{TG}(\theta, \lambda_q) = (\widetilde{S}(\theta, \lambda_q))^{v_2}(\widetilde{I} - \widetilde{I}_{2h}^h(\theta)(\widetilde{L}_{2h}(2\theta, \lambda_q))^{-1}\widetilde{I}_h^{2h}(\theta)\widetilde{L}_h(\theta, \lambda_q))(\widetilde{S}(\theta, \lambda_q))^{v_1},$$

where v_1, v_2 are the number of pre-smoothing, respectively post-smoothing steps and $\widetilde{I}$ corresponds to a (4×4) identity matrix. The asymptotic convergence factor of the two-grid algorithm, applied to the differential operator with frozen coefficients at a fixed grid point (jh, kh), is given by,

$$\rho(TG) = \max_{\lambda_q \in \sigma(\mathscr{M})} \max_{\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]^2} \rho(\widetilde{TG}(\theta, \lambda_q)).$$
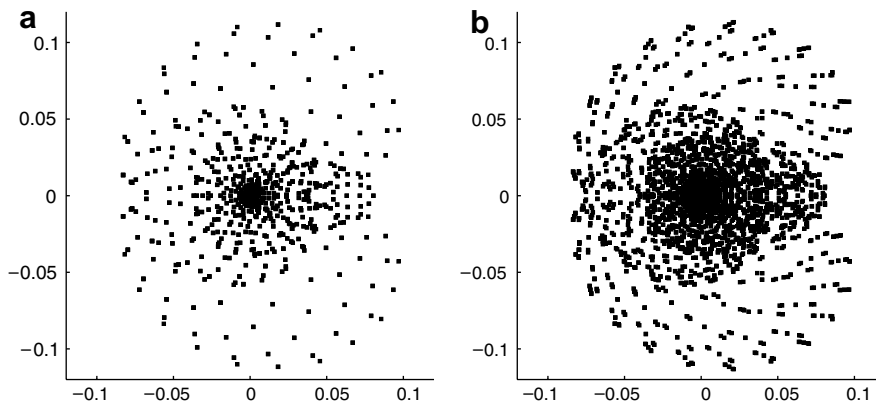

Fig. 3. Eigenvalues of the two-grid iteration matrix for a deterministic and a stochastic problem (400 equidistant samples of $(\theta_x, \theta_y) \in [-\frac{\pi}{2}, \frac{\pi}{2}]^2$). (a) $\alpha = 1.0$ (deterministic). (b) $\alpha \sim N(1.0, (0.1)^2)$, second-order PC-basis.

Fig. 3 compares the eigenvalues of the two-grid iteration matrix for a stochastic problem and a related deterministic problem. Similar conclusions hold as for Fig. 1 in Section 4.1. Each of the eigenvalues of the deterministic problem is again split into Q eigenvalues for the stochastic problem.

In Fig. 4, the convergence properties of one-level and two-level iterative methods are illustrated as a function of λ_q. The smoothing properties of the one-level method GS-LEX and the convergence factors of the multi-level methods show a strong dependence on λ_q. Especially small values of λ_q result in a very poor multigrid performance. These small values occur when a high order (Generalized) Polynomial Chaos basis is used or when the variance of the random field increases. Also other random field statistics, such as the correlation length, influence the spectrum of $\mathscr{M}$, but less severe than the influence of the variance. Increasing the number of Karhunen–Loève terms has only a small influence on the range of eigenvalues λ_q, asymptotically this influence disappears. The convergence properties of the iterative methods are illustrated by numerical experiments in Section 5.

When the two-grid operator contains spatially varying coefficients, the asymptotic convergence factor of the two-grid operator can be approximated by the maximum of the convergence factors $\rho(TG_{j,k})$ for all grid points (jh, kh) [20],

$$\rho(TG) = \max_{\mathbf{x}=(jh,kh)\in\mathbf{D}} \max_{\lambda_q\in\sigma(\mathscr{M}(\mathbf{x}))} \max_{\theta\in[-\frac{\pi}{2},\frac{\pi}{2}]^2} \rho(\widetilde{TG}(\theta, \lambda_q)), \tag{21}$$

if the variation of the coefficients is sufficiently smooth [23, Section 8.2.2].

4.3. Some comments on the decoupling approach

The decoupling approach results in a number of PDEs of the form (12), each with a deterministic coefficient

$$\lambda_{m_1,\ldots,m_N}(\mathbf{x}) = \alpha_0(\mathbf{x}) + \sum_{i=1}^{N} \alpha_i(\mathbf{x})\zeta_{n_i m_i}. \tag{22}$$

These coefficients can be collected into a diagonal matrix $\Lambda(\mathbf{x})$. The set of decoupled PDEs can then be written as a system of PDEs of the form (6), with $\mathscr{M}(\mathbf{x})$ replaced by $\Lambda(\mathbf{x})$. For analysis purposes it may again be required to freeze the coefficient to a local value. We denote this frozen coefficient by $\lambda_{m_1,\ldots,m_N}$, and the
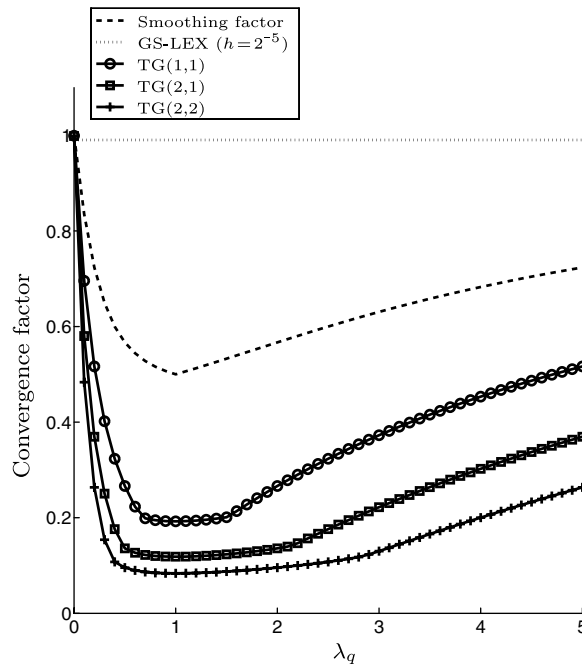

Fig. 4. Convergence factors for a one-level iterative method and for a two-grid cycle with GS-LEX smoothing.

corresponding diagonal matrix by Λ. Obviously, all of the results of Section 4.2 continue to hold for the decoupling approach, when $\mathcal{M}(\mathbf{x})$, $\mathcal{M}$, and λ_q are replaced by $\Lambda(\mathbf{x})$, Λ, and $\lambda_{m_1,\ldots,m_N}$ respectively.

Now we would like to point out two relations between the general and the decoupled approach. First, we consider the use of a Generalized Polynomial Chaos set of basis functions (3) that contains all possible products of the univariate polynomials. This set of functions Ψ spans the same vector space as the alternative basis Ψ' from Section 3 and it is also orthonormal. Hence, $\Psi = Z\Psi'$ must hold for some orthogonal matrix Z. With that matrix the eigenvalue decomposition of $\mathcal{M}(\mathbf{x})$ can be computed as

$$\mathcal{M}(\mathbf{x}) = \langle \alpha_{\mathrm{KL}}(\mathbf{x}, \omega) Z\Psi'\Psi'^{\mathrm{T}} Z^{\mathrm{T}} \rangle = Z\langle \alpha_{\mathrm{KL}}(\mathbf{x}, \omega)\Psi'\Psi'^{\mathrm{T}} \rangle Z^{\mathrm{T}} = Z\Lambda(\mathbf{x})Z^{\mathrm{T}}.$$

Thus the eigenvalues of $\mathcal{M}(\mathbf{x})$ are the diagonal entries $\lambda_{m_1,\ldots,m_N}(\mathbf{x})$ of $\Lambda(\mathbf{x})$. So, the convergence rate of the multigrid method applied to the coupled system of PDEs (6) constructed using this particular Generalized Polynomial Chaos basis, is equal to the worst case of the convergence rates of multigrid applied to the decoupled PDEs (12) separately.

Next, we consider the case when the set of basis functions (3) does not contain all possible products of the univariate polynomials. As mentioned in Section 2.2, this situation arises when Ψ is constructed by specifying the maximum degree of the multivariate polynomials. In this case, a diagonalization of $\mathcal{M}(\mathbf{x})$ is in general not possible. There is no orthogonal transformation that diagonalizes the different matrices $\langle \xi_i \Psi\Psi^{\mathrm{T}} \rangle$ appearing in the definition (5) of $\mathcal{M}(\mathbf{x})$ simultaneously. However, it is possible to find upper and lower bounds for the eigenvalues of $\mathcal{M}(\mathbf{x})$. Therefore, the set Ψ of polynomials is completed to the minimal set that contains all possible products of the required univariate polynomials, denoted by $\overline{\Psi}$. Then, $\mathcal{M}(\mathbf{x}) = \langle \alpha_{\mathrm{KL}}(\mathbf{x}, \omega)\Psi\Psi^{\mathrm{T}} \rangle$ is a submatrix of $\overline{\mathcal{M}}(\mathbf{x}) = \langle \alpha_{\mathrm{KL}}(\mathbf{x}, \omega)\overline{\Psi}\overline{\Psi}^{\mathrm{T}} \rangle$ obtained by omitting all but Q rows and columns. As a result the range of eigenvalues $[\lambda_{\min}(\mathbf{x}), \lambda_{\max}(\mathbf{x})]$ of $\mathcal{M}(\mathbf{x})$ must be contained in the range of eigenvalues $[\bar{\lambda}_{\min}(\mathbf{x}), \bar{\lambda}_{\max}(\mathbf{x})]$ of $\overline{\mathcal{M}}(\mathbf{x})$.

Hence, analysis of the convergence of the multigrid method for the decoupling approach using polynomials of degree n_i in ξ_i, provides us immediately with upper bounds for the multigrid convergence of the general approach, using any set of polynomials with degrees in ξ_i lower than or equal to n_i. The analysis of the decoupling approach is a straightforward exercise, once the zeros $\zeta_{n_i m_i}$ have been computed.

5. Numerical results

We present some test results that illustrate the convergence behavior and the sharpness of the Fourier mode analysis for the single-grid and multigrid methods discussed in the earlier sections. As a test case we consider model problem (1) with zero forcing term on an unit square domain with homogeneous Dirichlet boundary conditions. In a first test problem the random field $\alpha(\mathbf{x}, \omega)$ is actually a Gaussian random variable $\alpha_{\mathrm{var}}(\omega)$ with mean α_0 and variance σ^2. In the second test problem we selected $\alpha(\mathbf{x}, \omega)$ to be a Gaussian field $\alpha_{\mathrm{field}}(\mathbf{x}, \omega)$ with a specified mean $\alpha_0(\mathbf{x})$ and represented by a truncated Karhunen–Loève expansion based on the exponentially decaying covariance function,

$$C(\mathbf{x}, \mathbf{x}') = \sigma^2 \exp\left(-\frac{||\mathbf{x} - \mathbf{x}'||_1}{L_{\mathrm{c}}}\right). \tag{23}$$

The covariance function is parameterized by the variance σ^2 and by the correlation length L_{c}. Note that the Karhunen–Loève expansion (2) is known analytically for (23), see [5,18]. For the second test problem, the differential operator contains spatially varying coefficients. By selecting a mean value different from one, the effect of anisotropy can be investigated.

5.1. Single grid methods

We first consider pointwise GS-LEX, pointwise ω-Jacobi, pointwise GS-RB and x-line GS-LEX. Tables 2 and 3 display the convergence and smoothing factors obtained by Fourier analysis and the convergence factors obtained by numerical experiment for the first (random variable) problem and for the second (random field) problem. The results are given for an isotropic and an anisotropic problem. The theoretical convergence factors for the random variable problem are computed by using formula (19). For the random field case the

Table 2
Theoretical convergence factor (ρ_{theo}), smoothing factor (s_{theo}), and numerically observed convergence factor (ρ) for the one-level iterative solution of the model problem with $\alpha(\mathbf{x}, \omega)$ a random variable α_{var} ($h = 2^{-4}$, second-order PC-basis)

	$\alpha_{\text{var}} : \alpha_0 = 1, \sigma = 0.25$			$\alpha_{\text{var}} : \alpha_0 = 0.2, \sigma = 0.05$		
	ρ_{theo}	s_{theo}	ρ	ρ_{theo}	s_{theo}	ρ
ω-Jacobi	0.988	0.600	0.985	0.983	0.800	0.982
Point GS-LEX	0.961	0.551	0.962	0.961	0.817	0.962
Point GS-RB	0.961	0.407	0.962	0.961	0.807	0.962
x-line GS-LEX	0.935	0.447	0.937	0.827	0.447	0.845

Table 3
Theoretical convergence factors (ρ_{theo}), smoothing factor (s_{theo}), and numerically observed convergence factors (ρ) for the one-level iterative solution of the model problem ($h = 2^{-4}$, second-order PC-basis, 4-term Karhunen–Loève expansion, $L_c = 1$)

	$\alpha_{\text{field}} : \alpha_0(\mathbf{x}) = 1, \sigma = 0.25$			$\alpha_{\text{field}} : \alpha_0(\mathbf{x}) = 0.2, \sigma = 0.05$		
	ρ_{theo}	s_{theo}	ρ	ρ_{theo}	s_{theo}	ρ
ω-Jacobi	0.986	0.684	0.984	0.982	0.898	0.982
Point GS-LEX	0.961	0.544	0.962	0.961	0.808	0.962
Point GS-RB	0.961	0.390	0.962	0.961	0.797	0.962
x-line GS-LEX	0.933	0.447	0.935	0.821	0.447	0.838

estimate (20) is used. For the results in Table 2, the Polynomial Chaos basis consists of three basis functions, for the results in Table 3, of 15 basis functions.

Although the theoretical convergence factors hold only under the LFA assumptions of an infinite grid, we observe a good correspondence between the numerical and theoretical values. The results illustrate the independence of the asymptotic convergence factors of the pointwise iterative methods on the distribution of the random parameter α. As in the deterministic case, the line Gauss–Seidel method converges faster than the pointwise schemes for the anisotropic problem. The smoothing factors, which are indicative of the obtainable multigrid convergence factor, show a clear dependence on the statistics of the random parameter in case of the pointwise iterative methods. Also here, the advantage of using a line-relaxation method for the anisotropic problem is obvious.

Next, we investigate the convergence properties of pointwise GS-LEX more thoroughly, see the numerical results in Table 4. To illustrate the $1 - \mathcal{O}(h^2)$ behavior of the convergence factors, the table displays the convergence factors subtracted from unity. In the first block of rows the convergence is presented as a function of the mesh spacing h for the model problem with random field. The total number of discrete unknowns can be computed as the product of the number of internal grid points $(1/h - 1)^2$ multiplied by the number of random

Table 4
Theoretical convergence factors (ρ_{theo}) and numerically observed convergence factors (ρ) for the pointwise lexicographic Gauss–Seidel method for the model problem with random field: difference from 1 is given

Grid spacing	$h = 2^{-3}$	$h = 2^{-4}$	$h = 2^{-5}$	$h = 2^{-6}$	$h = 2^{-7}$
$1 - \rho_{\text{theo}}$	1.542×10^{-1}	3.855×10^{-2}	9.638×10^{-3}	2.410×10^{-3}	6.024×10^{-4}
$1 - \rho$	1.465×10^{-1}	3.806×10^{-2}	9.607×10^{-3}	2.408×10^{-3}	6.023×10^{-4}
KL-terms	$N = 1$	$N = 2$	$N = 3$	$N = 4$	$N = 5$
$1 - \rho_{\text{theo}} = 9.607 \times 10^{-3}$; $1 - \rho$	9.607×10^{-3}	9.607×10^{-3}	9.607×10^{-3}	9.607×10^{-3}	9.607×10^{-3}
Order PC-basis	$P = 1$	$P = 2$	$P = 3$	$P = 4$	$P = 5$
$1 - \rho_{\text{theo}} = 9.607 \times 10^{-3}$; $1 - \rho$	9.607×10^{-3}	9.607×10^{-3}	9.607×10^{-3}	9.607×10^{-3}	9.607×10^{-3}
Variance σ^2	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.4$	$\sigma = 0.6$	$\sigma = 0.7$
$1 - \rho_{\text{theo}} = 9.607 \times 10^{-3}$; $1 - \rho$	9.607×10^{-3}	9.607×10^{-3}	9.607×10^{-3}	9.607×10^{-3}	9.607×10^{-3}
Correlation length	$L_c = 0.25$	$L_c = 0.5$	$L_c = 1$	$L_c = 2.5$	$L_c = 5$
$1 - \rho_{\text{theo}} = 9.607 \times 10^{-3}$; $1 - \rho$	9.607×10^{-3}	9.607×10^{-3}	9.607×10^{-3}	9.607×10^{-3}	9.607×10^{-3}

Unless specified differently, the following default configuration is used: grid spacing $h = 2^{-5}$, second-order PC-basis, 4-term KL-expansion of α_{field} with $\sigma = 0.2$ and $L_c = 1$.

basis functions, Q. For example, the result for the $h = 2^{-7}$ case corresponds to a system constructed with second order polynomials in 4 random variables ξ_i, this leads to $Q = 15$ and a total of 241 935 unknowns. When the order P of the random polynomials is varied (third block row of Table 4), the number of random basis functions Q increases as follows: $Q = 5, 15, 35, 70, 126$ corresponding to $P = 1, 2, 3, 4, 5$, respectively. The theory from Section 4.1 states that the asymptotic convergence of pointwise GS-LEX is independent of the eigenstructure of $\mathcal{M}$ when equal grid spacing is used in the x- and y-dimension. The numerical results confirm the independence of the convergence rate on the order of the Polynomial Chaos basis, on the number of random dimensions and on the field statistics, the variance and correlation length of the random field. The convergence rate only depends on the grid spacing.

5.2. Multigrid methods

First, we shall verify the convergence analysis of the two-grid method. Some test results are given in Table 5. Two types of theoretical convergence factors are shown. The first, ρ_{theo}, is based on Eq. (21). The second, $\bar{\rho}_{\text{theo}}$, is an upper bound for ρ_{theo} based on the connection between the general and the decoupling approach in Section 4.3 and calculated by using the roots $\zeta_{n_i li}$. Also numerically observed results are presented.

For small problems, the upper bounds are quite sharp. For larger systems the quality of the upper bounds deteriorates, especially for a high Polynomial Chaos order or a larger number of random dimensions. This is due to the fact that the completed basis $\overline{\Psi}$ contains many more basis functions than Ψ, when the Polynomial Chaos order or the number of random dimensions is sufficiently high. The eigenvalue ranges for the coefficient matrices, $\mathcal{M}$ and $\overline{\mathcal{M}}$, are given in Table 6.

The convergence analysis in Section 4.2 states that the multigrid convergence is independent of the grid spacing, but not of the eigenvalue distribution of $\mathcal{M}$. The mesh-independent convergence is illustrated by the first block row of Table 5. As for the λ_q-dependence, we observe that the range of eigenvalues λ_q is in particular sensitive to the order of the random polynomial basis and to the variance of the random field. This is illustrated in Table 6. It is to a lesser extent dependent on the correlation length. The results for the various correlation lengths in the last block row of Tables 5 and 6 are based on a 20-term KL-expansion instead of on a 4-term expansion since decreasing the correlation length broadens the KL-eigenvalue spectrum and thus requires more KL-terms to represent the random field. Asymptotically the range of eigenvalues λ_q becomes

Table 5
Numerical (ρ) and theoretical convergence factors ($\bar{\rho}_{\text{theo}}$ and ρ_{theo}) for the two-grid cycle $TG(2,1)$ with GS-LEX smoothing

Grid spacing	$h = 2^{-4}$	$h = 2^{-5}$	$h = 2^{-6}$	$h = 2^{-7}$	$h = 2^{-8}$
	$\rho_{\text{theo}} = 0.123$; $\bar{\rho}_{\text{theo}} = 0.164$				
ρ	0.105	0.111	0.112	0.113	0.113
KL-terms	$N = 1$	$N = 2$	$N = 3$	$N = 4$	$N = 5$
ρ_{theo}	0.122	0.122	0.123	0.123	0.123
$\bar{\rho}_{\text{theo}}$	0.122	0.128	0.138	0.164	0.198
ρ	0.113	0.113	0.113	0.113	0.113
Order PC-basis	$P = 1$	$P = 2$	$P = 3$	$P = 4$	$P = 5$
ρ_{theo}	0.120	0.123	0.129	0.140	0.183
$\bar{\rho}_{\text{theo}}$	0.123	0.164	0.339	0.765	>1
ρ	0.113	0.113	0.117	0.151	0.187
Variance σ^2	$\sigma = 0.05$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.4$
ρ_{theo}	0.119	0.119	0.123	0.133	0.200
$\bar{\rho}_{\text{theo}}$	0.119	0.122	0.164	0.496	>1
ρ	0.112	0.112	0.113	0.135	0.202
Correlation length	$L_c = 0.5$	$L_c = 0.75$	$L_c = 1$	$L_c = 2.5$	$L_c = 5$
ρ_{theo}	0.120	0.120	0.120	0.120	0.120
$\bar{\rho}_{\text{theo}}$	0.250	0.229	0.224	0.161	0.134
ρ	0.112	0.112	0.112	0.112	0.112

Unless specified differently, the following configuration is used: $h = 2^{-7}$, second-order PC-basis, 4-term KL-expansion of α_{field} with $\sigma = 0.2$ and $L_c = 1$. The tests on the effect of the correlation length are based on a 20-term KL-expansion and a first-order PC-basis.

Table 6
Eigenvalue ranges for the coefficient matrices $\mathscr{M}(\mathbf{x})$, $[\lambda_{min}, \lambda_{max}]$, and $\overline{\mathscr{M}}(\mathbf{x})$, $[\bar{\lambda}_{min}, \bar{\lambda}_{max}]$

KL-terms	$N = 1$	$N = 2$	$N = 3$	$N = 4$	$N = 5$
$[\lambda_{min}, \lambda_{max}]$	[0.706, 1.294]	[0.694, 1.306]	[0.687, 1.313]	[0.682, 1.318]	[0.677, 1.323]
$[\bar{\lambda}_{min}, \bar{\lambda}_{max}]$	[0.706, 1.294]	[0.585, 1.415]	[0.489, 1.511]	[0.422, 1.578]	[0.366, 1.634]
Order PC-basis	$P = 1$	$P = 2$	$P = 3$	$P = 4$	$P = 5$
$[\lambda_{min}, \lambda_{max}]$	[0.816, 1.184]	[0.682, 1.318]	[0.571, 1.429]	[0.475, 1.525]	[0.389, 1.611]
$[\bar{\lambda}_{min}, \bar{\lambda}_{max}]$	[0.666, 1.334]	[0.422, 1.578]	[0.221, 1.779]	[0.047, 1.953]	[−0.109, 2.109]
Variance σ^2	$\sigma = 0.05$	$\sigma = 0.1$	$\sigma = 0.2$	$\sigma = 0.3$	$\sigma = 0.4$
$[\lambda_{min}, \lambda_{max}]$	[0.920, 1.080]	[0.841, 1.159]	[0.682, 1.318]	[0.522, 1.478]	[0.363, 1.637]
$[\bar{\lambda}_{min}, \bar{\lambda}_{max}]$	[0.856, 1.144]	[0.711, 1.289]	[0.422, 1.578]	[0.133, 1.867]	[−1.558, 2.156]
Correlation length	$L_c = 0.5$	$L_c = 0.75$	$L_c = 1$	$L_c = 2.5$	$L_c = 5$
$[\lambda_{min}, \lambda_{max}]$	[0.814, 1.186]	[0.809, 1.191]	[0.806, 1.194]	[0.802, 1.198]	[0.801, 1.199]
$[\bar{\lambda}_{min}, \bar{\lambda}_{max}]$	[0.301, 1.699]	[0.325, 1.676]	[0.331, 1.669]	[0.428, 1.572]	[0.516, 1.484]

Unless specified differently, the following configuration is used: second-order PC-basis, 4-term KL-expansion of α_{field} with $\sigma = 0.2$ and $L_c = 1$. The tests on the effect of the correlation length are based on a 20-term KL-expansion and a first-order PC-basis.

independent of the included number of Karhunen–Loève terms. This explains the observed independence of the convergence factors on the number of KL-terms.

As such, our LFA explains some features observed in [10]: the convergence properties usually deteriorate as the variance of the random field increases, but only decrease slightly as the correlation length is decreased; the multigrid performance usually drops as the order of the Polynomial Chaos increases; the multigrid convergence rate is independent of the number of stochastic dimensions.

Fig. 5 illustrates the performance of the true *multi*-grid method with a pointwise smoother (lexicographic and red–black Gauss–Seidel, red–black successive overrelaxation), and with an *x*-line Gauss–Seidel smoother, for an isotropic and an anisotropic problem. As in the deterministic PDE case line relaxation methods can deal effectively with (certain types of) anisotropy when standard coarsening is used. Further, a red–black SOR method with suitable damping factor turns out to be an efficient smoother both for isotropic and anisotropic problems. This confirms with the behavior of the SOR smoothing factor as shown in Fig. 2. The practical multigrid convergence behavior is demonstrated in Table 7 with some numerical convergence factors for different types of multigrid cycles. The problem setup results in a linear system with about one million unknowns ($h = 2^{-8}, Q = 15$). The multigrid convergence rate of the *W*-cycles is as good as the two-grid performance.
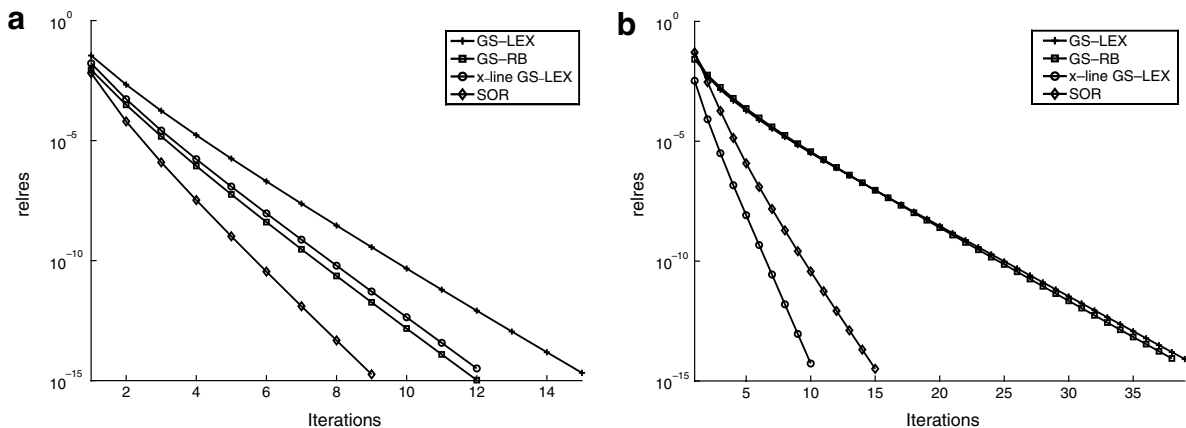

Fig. 5. Residual norms for a V(2,1) multigrid cycle with pointwise lexicographic Gauss–Seidel (GS-LEX), pointwise red–black Gauss–Seidel (GS-RB), lexicographic *x*-line Gauss–Seidel (*x*-line GS-LEX) or pointwise red–black successive overrelaxation (SOR-RB) smoothing ($h = 2^{-4}$, second-order PC-basis, 4-term Karhunen-Loève expansion of α_{field} with $L_c = 1$). (a) $\alpha_{field} : \alpha_0(\mathbf{x}) = 1, \sigma = 0.25$. (b) $\alpha_{field} : \alpha_0(\mathbf{x}) = 0.2, \sigma = 0.05$.

Table 7
Numerically observed convergence factors for the multigrid solution of the model problem

	TG-cycle			V-cycle			W-cycle		
Smoothing steps	(1,1)	(2,1)	(2,2)	(1,1)	(2,1)	(2,2)	(1,1)	(2,1)	(2,2)
Convergence factor	0.117	0.051	0.039	0.145	0.074	0.055	0.118	0.051	0.039

A red–black Gauss–Seidel smoother is used. ($h = 2^{-8}$, second-order PC-basis, 4-term KL-expansion of α_{field} with $\sigma = 0.2$ and $L_c = 1$).

6. Conclusions

We have extended and applied local Fourier analysis techniques to study the performance of iterative solvers for PDEs involving random coefficients. To this end, the special structure of the discrete algebraic systems arising from the spectral expansion approach has been exploited. The LFA shows that the convergence depends crucially on the eigenstructure of a certain matrix, which characterizes the random structure of the PDE. This analysis has enabled us to explain some of the features that are observed (also by other authors [10]) in computational experiments.

We have also specified a set of random basis functions that lead to a complete decoupling of the model problem into deterministic PDEs. The local Fourier analysis for the decoupling approach gives an indication of the convergence properties that one can expect for the Generalized Polynomial Chaos approach.

The goal of our study was to gain some insights into the convergence behavior of the stochastic expansion approach. To that end, we concentrated on a model problem that could be analyzed quantitatively. The method does however extend to more challenging problems, e.g. problems on irregular finite element meshes, to time-dependent problems, to different PDE operators, and so on.

References

[1] G. Karniadakis, C.-H. Su, D. Xiu, D. Lucor, C. Schwab, R. Todor, Generalized polynomial chaos solution for differential equations with random inputs, Research Report 2005-01, Seminar for Applied Mathematics, ETC Zürich, January 2005.
[2] G.I. Schuëller, A state-of-the-art report on computational stochastic mechanics, Prob. Engrg. Mech. 12 (4) (1997) 197–322.
[3] I. Babuška, P. Chatzipantelidis, On solving elliptic stochastic partial differential equations, Comput. Methods Appl. Mech. Engrg. 191 (2002) 4093–4122.
[4] N. Scheerlinck, P. Verboven, J.D. Stigter, J. De Baerdemaeker, J.F. Van Impe, B.M. Nicolaï, A variance propagation algorithm for stochastic heat and mass transfer problems in food processes, Int. J. Num. Methods Engrg. 51 (8) (2001) 961–983.
[5] R. Ghanem, P. Spanos, A spectral stochastic finite element formulation for reliability analysis, J. Engrg. Mech. ASCE 17 (1991) 2351–2372.
[6] R. Ghanem, P. Spanos, Stochastic Finite Elements: A Spectral Approach, Springer Verlag, New York, USA, 1991.
[7] M. Pellissetti, R. Ghanem, Iterative solution for systems of linear equations arising in the context of stochastic finite elements, Adv. Engrg. Software 31 (2000) 607–616.
[8] I. Babuška, R. Tempone, G. Zouraris, Galerkin finite element approximations of stochastic elliptic partial differential equations, SIAM J. Numer. Anal. 42 (2004) 800–825.
[9] I. Babuška, R. Tempone, G.E. Zouraris, Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation, Comput. Methods Appl. Mech. Engrg. 194 (2005) 1251–1294.

[10] O. Le Maître, O. Knio, B. Debusschere, H. Najm, R. Ghanem, A multigrid solver for two-dimensional stochastic diffusion equations, Comput. Methods Appl. Mech. Engrg. 192 (2003) 4723–4744.

[11] H. Elman, D. Furnival, Solving the stochastic steady-state diffusion problem using multigrid, Technical Report TR-4786, University of Maryland, Department of Computer Science (2006).

[12] A. Brandt, Multi-level adaptive solutions to boundary-value problems, Math. Comput. 31 (1977) 333–390.

[13] A. Brandt, Rigorous quantitative analysis of multigrid I: Constant coefficients two-level cycle with L_2-norm, SIAM J. Numer. Anal. 31 (6) (1994) 1695–1730.

[14] P. Wesseling, An Introduction to Multigrid Methods, Wiley, Chichester, England, 1992.

[15] M. Loève, Probability Theory, Springer, New York, USA, 1977.

[16] D. Xiu, G. Karniadakis, The Wiener–Askey polynomial chaos for stochastic differential equations, SIAM J. Sci. Comput. 24 (2) (2002) 619–644.

[17] X. Wan, G.E. Karniadakis, An adaptive multi-element generalized polynomial chaos method for stochastic differential equations, J. Comput. Phys. 209 (2) (2005) 617–642.

[18] B. Sudret, A. Der Kiureghian, Stochastic finite elements and reliability: a state-of-the-art report, Technical Report UCB/SEMM-2000/08, University of California, Berkeley (2000).

[19] U. Trottenberg, C. Oosterlee, A. Schüller, Multigrid, Academic Press, San Diego, USA, 2001.

[20] R. Wienands, W. Joppich, Practical Fourier Analysis for Multigrid Methods, CRC Press, Boca Raton, Florida, USA, 2005.

[21] I. Yavneh, Multigrid smoothing factors for red–black Gauss–Seidel relaxation applied to a class of elliptic operators, SIAM J. Numer. Anal. 32 (4) (1995) 1126–1138.

[22] I. Yavneh, On red–black SOR smoothing in multigrid, SIAM J. Sci. Comput. 17 (1) (1996) 180–192.

[23] W. Hackbusch, Multi-grid Methods and Applications, Springer, Berlin, Germany, 1985.