REGULAR ARTICLE

Computational aspects of the stochastic finite element method

**Michael Eiermann · Oliver G. Ernst ·
Elisabeth Ullmann**

Abstract   We present an overview of the stochastic finite element method with an emphasis on the computational tasks involved in its implementation.

Keywords   Uncertainty quantification · Stochastic finite element method · Hierarchical matrices · Thick-restart Lanczos method · Multiple right hand sides

Mathematics Subject Classification (2000)   65C30 · 65F10 · 65F15 · 65N30

1 Introduction

The current scientific computation paradigm consists of mathematical models—often partial differential equations (PDEs)—describing certain physical phenomena under study whose solutions are approximated by numerical schemes carried out by computers. Among these four components, represented by boxes in Fig. 1, great progress has resulted in the area of computer implementation due both to the rapid advance in computer speed and storage capacity as well as improvements in software aspects such as floating point

M. Eiermann · O. G. Ernst (✉) · E. Ullmann
Institut für Numerische Mathematik und Optimierung,
TU Bergakademie Freiberg, 09596 Freiberg, Germany
e-mail: ernst@math.tu-freiberg.de

M. Eiermann
e-mail: eiermann@math.tu-freiberg.de

E. Ullmann
e-mail: ullmann@math.tu-freiberg.de

standardization and programming methodology. Similarly, advances in numerical methods such as basic linear algebra libraries, discretization schemes and adaptivity make it possible to solve many nontrivial PDEs quickly and to as great an accuracy as desired.

An aspect of this general approach which deserves more attention is the fact that the data required by the model—various parameters such as the spatial distribution of material properties as well as source or boundary terms—are invariably assumed as known. In practice, however, such data is obtained from measurements or based on various assumptions, all subject to uncertainty. Indeed, it is quite possible for the effect of such uncertainty in the data to outweigh that of rounding or discretization errors. One usually distinguishes two types of uncertainty: the first, *aleatoric uncertainty*, which refers to an intrinsic variablity of certain quantities, such as the wind stress on a structure. In contrast, *epistemic uncertainty* refers to a lack of knowledge about certain aspects of a system which, in contrast to aleatoric uncertainty, can be reduced through additional information.

The idea of *uncertainty quantification* (UQ), i.e. quantifying the effects of uncertainty on the result of a computation, has received much interest of late. The objective is usually that of propagating quantitative information on the data through a computation to the solution. It should be obvious that technological or political decisions based on simulation results can benefit greatly when uncertainty in these results is quantified in a meaningful way.

Among the different techniques of UQ, the most common is certainly to ignore the issue and to deal with the variability of data by using averaged quantities. Other non-stochastic techniques of UQ include worst-case analysis and fuzzy set theory. In stochastic
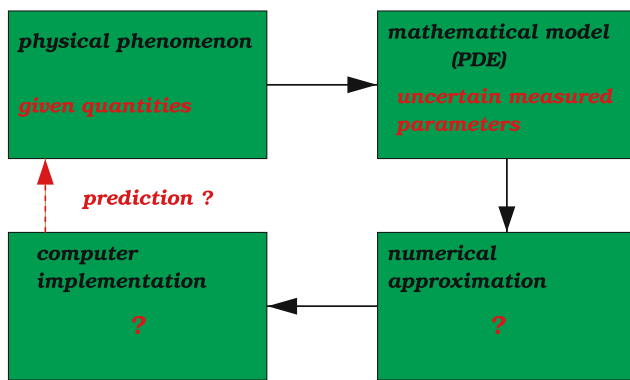
Fig. 1 Uncertainty in the typical computational science framework. The given quantities in the physical phenomenon to be modelled enter into the mathematical model as measured parameters subject to uncertainty. The *question marks* signify that it is unclear how this uncertainty can be accounted for in the numerical approximation and what computational challenges this poses in the computer implementation. The final goal, then, is incorporating the uncertainty quantified in the computation into predictive statements on the phenomenon being modelled

approaches to UQ, the uncertain quantities are modelled as random variables, so that PDEs become stochastic PDEs (SPDEs). The most straightforward way of doing this is the Monte Carlo Method [24], in which many realizations of the random variables are generated, each leading to a deterministic problem, which is then solved using whatever methods are appropriate for the deterministic problem. The resulting ensemble of solutions can then be post-processed to obtain statistical information on the variability of the solution. A more ambitious approach is to solve the SPDE, the solution of which is a stochastic process, and to derive quantitative statements on the effect of data uncertainty from the distribution of this process.

Much of the literature on stochastic differential equations, particularly stochastic ordinary differential equations, allows for processes with zero correlation length, known as white noise [18]. A rigorous theory of SPDEs based on white noise analysis requires defining the product of stochastic processes as so-called Wick products [19], which in turn leads to solutions which in general do not agree with the Monte Carlo approximation in the limit of an infinite number of realizations. For this reason, and also because it occurs in many engineering contexts, it is of interest to consider instead the case where the processes involved display significant correlations, which is sometimes referred to as *coloured noise*, and we shall do so in this paper.

Recently, a systematic approach for formulating and discretizing PDEs with random data known as the *Stochastic Finite Element Method* (SFEM) has become popular in the engineering community [15] and subsequently

analysed by numerical analysts (see [5] and the references therein). The results of a SFEM approximation allow one to compute a large variety of statistical information via post processing, such as moments of the solution as well as the probability of certain events related to the PDE. The method is, however, computationally expensive, and it is the objective of this paper to present computational approaches for solving the main tasks arising in the implementation of certain SFEM formulations.

The remainder of this paper is organized as follows. Section 2 introduces a model elliptic boundary value problem with random data as well as necessary stochastic terminology. In Sect. 3 the basic discretization steps of the SFEM is given with an emphasis on the structure of the resulting Galerkin equations. Section 4 identifies the two main computation tasks involved in implementing the SFEM and presents some computational schemes which exploit structure and analytical properties of the problem. Section 5 contains some numerical examples followed by some conclusions in Sect. 6.

2 Elliptic boundary value problems with random data

In this section we give a brief overview of the SFEM by starting from the following (deterministic) elliptic boundary value problem: given a domain $D \subset \mathbb{R}^d$, we seek a function u which satisfies

$$-\nabla \cdot \big(\kappa(\boldsymbol{x})\nabla u(\boldsymbol{x})\big) = f(\boldsymbol{x}), \quad \boldsymbol{x} \in D, \tag{1a}$$

$$u(\boldsymbol{x}) = g(\boldsymbol{x}), \quad \boldsymbol{x} \in \partial D, \tag{1b}$$

where the coefficient function κ is uniformly positive definite and bounded and both it as well as the source term f and boundary data g, are sufficiently smooth functions defined on D and its boundary ∂D, respectively. The data of problem (1) consists of the functions κ, f and g and we shall model possible uncertainty in these by allowing them to be *random fields* rather than deterministic functions.[1]

2.1 Random fields

Given a complete probability space $(\Omega, \mathfrak{A}, P)$ with sample space Ω, σ-algebra $\mathfrak{A}$ on Ω and probability measure P on $\mathfrak{A}$, a real-valued random field κ defined on a set D is a mapping $\kappa : D \times \Omega \to \mathbb{R}$ such that, for each $\boldsymbol{x} \in D$, $\kappa(\boldsymbol{x}, \cdot)$ is a random variable with respect to $(\Omega, \mathfrak{A}, P)$. In other words, rather than a given real number $\kappa(\boldsymbol{x})$, the

---

[1] It is also possible to treat uncertainty in the domain D, see [3,16].

random field κ at the point $x \in D$ is a random variable, and one obtains a real number $\kappa(x, \omega)$ for each realization $\omega \in \Omega$. An alternative point of view regards $\kappa(\cdot, \omega)$ as a sample drawn from an appropriate function space such that each realization ω yields a function on D. Thus, a random field (sometimes also called a *random function*) is a stochastic process with the spatial coordinate x as its index variable. The theory of random fields is treated in [1,2,7,41].

We further introduce some stochastic terminology: we shall denote the *mean* of a random variable $X : \Omega \to \mathbb{R}$ by

$$\langle X \rangle := \int_{\Omega} X(\omega) \, dP(\omega)$$

and we denote the *mean* of the random field κ at the point $x \in D$ by $\overline{\kappa}(x) := \langle \kappa(x, \cdot) \rangle$. The *covariance* of κ at $x, y \in D$ is denoted by

$$\mathrm{Cov}_{\kappa}(x, y) := \langle (\kappa(x, \cdot) - \overline{\kappa}(x))(\kappa(y, \cdot) - \overline{\kappa}(y)) \rangle ,$$

the *variance* of κ at $x \in D$ by $\mathrm{Var}_{\kappa}(x) := \mathrm{Cov}_{\kappa}(x, x)$ and the *standard deviation* of κ at x by $\sigma_{\kappa}(x) := \sqrt{\mathrm{Var}_{\kappa}(x)}$. The space of all random variables with finite variance is denoted $L^2_P(\Omega)$ with inner product $(X, Y) := \mathrm{Cov}(X, Y)$. Recall that all *covariance functions* $(x, y) \mapsto \mathrm{Cov}_{\kappa}(x, y)$ are positive semidefinite functions [26].

Often problem parameters are modelled as *Gaussian* random fields, i.e. random fields whose finite-dimensional distributions are all jointly Gaussian. Such random fields are convenient in that, e.g. they are completely specified by their first and second order statistics $\overline{\kappa}$ and Cov_{κ} and for Gaussian random variables independence is equivalent with uncorrelateness. Furthermore, Gaussian random fields occur naturally as a result of the central limit theorem and, whenever only second order statistical information is available, Gaussian random fields are the model of maximal entropy [30]. However, Gaussian random fields are not appropriate models for physical quantities which are positive and bounded, as is the case, e.g. when κ models a diffusion coefficient in (1).

Another common assumption on random fields is that they are *homogeneous*, i.e. that their finite dimensional distributions are invariant under translation; in particular, this implies $\overline{\kappa} \equiv$ const and $\mathrm{Cov}_{\kappa}(x, y) = c(x - y)$, where the *covariance function* c can be represented by a Fourier transform via Bochner's Theorem. In addition, random fields are often assumed to be *isotropic*, i.e. invariant under orthogonal transformations, which restrict the covariance function further to satisfy $c(x - y) = c(r), r := \|x - y\|$. Finally, assumptions on c are often made which guarantee that the realizations of

the associated random field are continuous or differentiable in a mean-square or almost sure sense. Figure 2 shows some covariance functions commonly used in modelling random fields. Each covariance function contains the parameters σ, the standard deviation and a, which is proportional to the *correlation length*[2] defined as $1/c(0) \int_0^\infty c(r) \, dr$. This quantity gives a length scale for the distance over which the random field exhibits significant correlations. See [42] for a discussion of exponential versus Bessel correlation models.

2.2 Karhunen–Loève expansion

The SFEM is based on separate treatment of the deterministic and stochastic independent variables, in this case x and ω. To this end, the random fields modelling the data are expanded in a sum of products of functions of x and ω only, respectively. While there are other possibilities (some are mentioned in [28]), the most common approach for achieving this is the *Karhunen–Loève expansion* (KL) [20,25].

Any random field $\kappa : D \times \Omega \to \mathbb{R}$ with a continuous covariance function possesses the representation

$$\kappa(x, \omega) = \overline{\kappa}(x) + \sum_{j=1}^{\infty} \sqrt{\lambda_j} \kappa_j(x) \xi_j(\omega), \qquad (2)$$

where the series converges in $L^\infty(D) \otimes L^2_P(\Omega)$ (see, e.g. [39] for a definition of the tensor product of Hilbert spaces). Here $\{\xi_j\}_{j=1}^{\infty}$ is a sequence of mutually uncorrelated random variables in $L^2_P(\Omega)$ with zero mean and unit variance determined by

$$\xi_j(\omega) = \frac{1}{\sqrt{\lambda_j}} \int_D (\kappa(x, \omega) - \overline{\kappa}(x)) \, \kappa_j(x) \, dx.$$

The numbers λ_j and functions $\kappa_j : D \to \mathbb{R}$ are the eigenpairs of the compact, nonnegative-definite and self-adjoint covariance integral operator $C : L^2(D) \to L^2(D)$, defined by

$$u \mapsto Cu = \int_D \mathrm{Cov}_{\kappa}(x, \cdot) u(x) \, dx \in L^2(D).$$

The eigenfunctions are orthogonal in $L^2(D)$ and the eigenvalues satisfy

$$\sum_{j=1}^{\infty} \lambda_j = \int_D \mathrm{Var}_{\kappa}(x) \, dx. \qquad (3)$$

---

[2] The *correlation* of κ at x and y is defined as $\mathrm{Cov}_{\kappa}(x, y)/\sigma_{\kappa}(x)\sigma_{\kappa}(y)$, which for homogeneous isotropic fields becomes $c(r)/c(0)$.
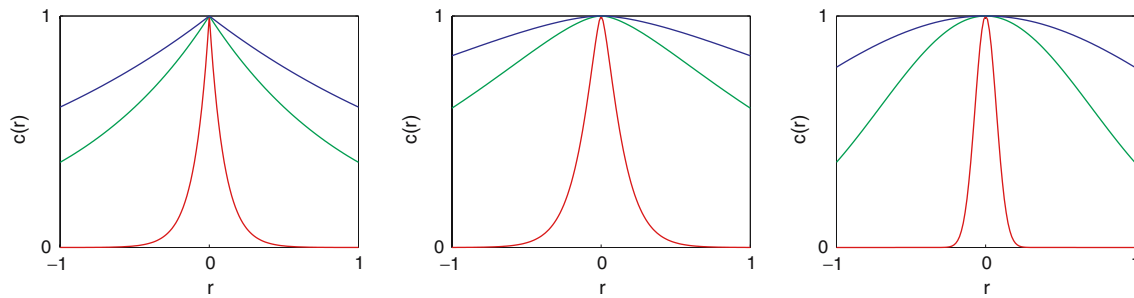
Fig. 2 Some common covariance functions of homogeneous, isotropic random fields: exponential correlation $c(r) = \sigma^2 e^{-r/a}$ (*left*), Bessel correlation $c(r) = \sigma^2 \frac{r}{a} K_1(\frac{r}{a})$ (*middle*) and the smooth correlation $c(r) = \sigma^2 e^{-r^2/a^2}$ (*right*). In each case the parameter $a > 0$ has the values $a = 0.1$ (*lower curve*), $a = 1$ (*middle curve*) and $a = 2$ (*upper curve*). K_1 is the second-kind Bessel function of order one

The eigenvalues of C form a descending sequence of nonnegative numbers converging to zero, and hence partial sums of (3) capture more and more of the total variance of the random field κ and truncating the KL expansion (2) yields an approximation of κ.

2.3 Stochastic boundary value problem

By allowing data of the boundary value problem (1), e.g. the source term f and coefficient function κ, to be random fields, we obtain a stochastic boundary value problem (SBVP), the solution of which must then also be a random field. We thus seek $u : D \times \Omega \to \mathbb{R}$ such that, P-almost surely (P-a.s.), there holds

$$-\nabla \cdot \big(\kappa(\boldsymbol{x}, \omega) \nabla u(\boldsymbol{x}, \omega)\big) = f(\boldsymbol{x}, \omega), \quad \boldsymbol{x} \in D, \tag{4a}$$

$$u(\boldsymbol{x}, \omega) = 0, \quad \boldsymbol{x} \in \partial D, \tag{4b}$$

where we have prescribed homogeneous (deterministic) boundary values for simplicity. To obtain a well-posed problem (cf. [4,5,29]) we further assume that $\kappa \in C^1(\overline{D})$, as a function of $\boldsymbol{x}$, and that it is P-a.s. uniformly bounded above away from zero below. Finally, if $\mathfrak{B}(D)$ denotes the σ-algebra generated by the open subsets of D and likewise for $\mathfrak{B}(\mathbb{R})$, then we choose $\mathfrak{A}$ to be the smallest σ-algebra on Ω such that f and κ are continuous with respect to $\mathfrak{B}(D) \times \mathfrak{A}$ and $\mathfrak{B}(\mathbb{R})$.

For finite element discretization of (4) we recast it in a variational formulation. We begin with the deterministic version (1) and select a suitable function space X, in our example the Sobolev space $H_0^1(D)$. The usual integration by parts procedure leads to the problem of finding $u \in X$ such that

$$a(u, v) = \ell(v) \quad \forall v \in X,$$

with the bounded and coercive bilinear form $a : X \times X \to \mathbb{R}$ the bounded linear form $\ell : X \to \mathbb{R}$ given by

$$a(u, v) = \int_D \kappa \nabla u \cdot \nabla v \, d\boldsymbol{x}, \quad \ell(v) = \int_D f v \, d\boldsymbol{x} \quad u, v \in X.$$

For the variational characterization of the SBVP (4), we choose the tensor product space $X \otimes L_P^2(\Omega)$ as the function space of random fields on D and now seek $u \in X \otimes L_P^2(\Omega)$ such that

$$\langle a(u, v) \rangle = \langle \ell(v) \rangle \quad \forall v \in X \otimes L_P^2(\Omega). \tag{5}$$

The reader is again referred to [4,5,29] for discussions of well-posedness of this stochastic variational problem.

3 The stochastic finite element method

The SFEM in its current form was first introduced in the monograph by Ghanem and Spanos [15]. Although both the term and the idea of incorporating randomness in a finite element formulation have a longer history (see [27,35] for overviews of earlier work, particularly in the area of stochastic mechanics) this probably constitutes the first systematic Galerkin approximation in deterministic *and* random variables. Convergence analyses of SFEM formulations can be found in [4,5,29,36]. Excellent surveys can be found in [21,22]. See also [37], which emphasises SFEM for reliability analysis. See [38,40] for an analysis of SFEM discretizations based on white noise analysis.

3.1 Discretization steps

The SFEM discretization treats the deterministic and stochastic variables separately. For the determinitstic part, let

$$X^h = \text{span}\{\phi_1, \phi_2, \ldots, \phi_{N_x}\} \subset X \qquad (6)$$

be any suitable finite dimensional subspace of the deterministic variational space. In particular, this finite element discretization of the associated deterministic problem can be chosen completely independently of the stochastic discretization.

For the stochastic discretization, the first step is to determine a finite number M of *independent* random variables $\{\xi_m\}_{m=1}^M$ which together sufficiently capture the stochastic variability of the problem. This step, which should be regarded as part of the modelling, could, e.g. be achieved by expanding the random fields in (4) in their Karhunen–Loève series and truncating these after a sufficiently large number of terms. As a consequence, the stochastic variation of the random fields is now only through its dependence on the random variables $\xi_1, \ldots, \xi_M$, i.e.

$$\kappa(x, \omega) = \kappa(x, \xi_1(\omega), \ldots, \xi_M(\omega)) =: \kappa(x, \xi(\omega))$$

and analogously for the random field f. Let $\Gamma_m := \xi_m(\Omega)$ denote the range of ξ_m and assume each ξ_m has a probability density $\rho_m : \Gamma_m \to [0, \infty)$. Since the ξ_m was assumed independent, their joint probability density is given by

$$\rho(\xi) = \rho_1(\xi_1), \ldots, \rho_M(\xi_M), \quad \xi \in \Gamma := \Gamma_1 \times \cdots \times \Gamma_M.$$

We can now reformulate the stochastic variational problem (4) in terms of the random vector ξ, i.e. we replace $L_P^2(\Omega)$ by $L_\rho^2(\Gamma)$ and obtain the problem of finding $u \in H_0^1(D) \otimes L_\rho^2(\Gamma)$ such that

$$\langle a(u, v) \rangle = \langle \ell(v) \rangle \quad \forall v \in H_0^1(D) \otimes L_\rho^2(\Gamma), \qquad (7)$$

where now

$$\langle a(u, v) \rangle = \int_\Gamma \rho(\xi) \int_D \kappa(x, \xi) \nabla u(x, \xi) \cdot \nabla v(x, \xi) \, dx \, d\xi,$$

and

$$\langle \ell(v) \rangle = \int_\Gamma \rho(\xi) \int_D f(x, \xi) v(x, \xi) \, dx \, d\xi.$$

The variational problem (7) is thus an approximation of the SBVP (4) by a deterministic variational problem with a finite number of parameters.

We next introduce a finite dimensional subspace

$$W^h = \text{span}\{\psi_1(\xi), \psi_2(\xi), \ldots, \psi_{N_\xi}(\xi)\} \subset W := L_\rho^2(\Gamma) \qquad (8)$$

of the stochastic parameter space and approximate the tensor product space $X \otimes W$ by the tensor product

$$X^h \otimes W^h = \{v \in L^2(D \times \Gamma) :$$
$$v \in \text{span}\{\phi(x)\psi(\xi) : \phi \in X^h, \psi \in W^h\}\}.$$

The trial and test functions $u^h \in X^h \otimes W^h$ are thus of the form

$$u^h(x, \xi) = \sum_{i,j} u_{i,j} \phi_i(x) \psi_j(\xi)$$

with a set of $N_x \cdot N_\xi$ coefficients $u_{i,j}$. The construction of W^h can be based on the tensor product structure of $W = L_{\rho_1}^2(\Gamma_1) \otimes \cdots \otimes L_{\rho_M}^2(\Gamma_M)$, discretizing the spaces $L_{\rho_m}^2(\Gamma_m)$ of univariate functions by finite dimensional subspaces W_m^h and forming

$$W^h = \text{span}\left\{ \psi_\alpha(\xi) = \prod_{m=1}^M \psi_{\alpha_m}(\xi_m) : \psi_{\alpha_m} \in W_m^h \subset L_{\rho_m}(\Gamma_m) \right\},$$
$$\alpha \in \mathbb{N}_0^M.$$

Several constructions for W^h have been proposed in the literature. One such approach (cf. [4,5,8,9,12]) employs piecewise polynomials on a partition of each domain Γ_m into subintervals (this assumes the Γ_m are bounded, as is the case, e.g. when the ξ_m are uniformly distributed). The more widely used construction (cf. [4,15,29, 45]), however, employs global polynomials in each variable ξ_m. When all random variables ξ_m are independent and identically distributed Gaussian, a basis of tensor product Hermite polynomials is used and the resulting space is sometimes called the *polynomial chaos expansion*, a terminology originally introduced by Norbert Wiener [43] in the context of turbulence modelling. A similar construction, referred to as *generalized polynomial chaos*, employs expansions in orthogonal polynomials associated with other classical probability distributions [45]. Since tensor product polynomial spaces increase rapidly in dimension, reduced tensor product spaces bounding the total polynomial degree or sparse grid approaches have also been proposed [22,36].

3.2 Structure of Galerkin equations

Efficient solution algorithms can be obtained by exploiting the structure of the linear system of equations resulting from the Galerkin discretization, and hence we discuss this structure here.

A basis of the discrete trial and test space $X^h \otimes W^h$ is given by all functions $\phi_i(x)\psi_j(\xi)$ where ϕ_i and ψ_j belong to the given bases of X^h and W^h, respectively. Expanding the discrete solution approximation $u^h(x, \xi) = \sum_{i,j} u_{i,j} \phi_i(x) \psi_j(\xi)$ in this basis, inserting u^h

into the variational Eq. (7) along with a test function $v(\boldsymbol{x}, \boldsymbol{\xi}) = \phi_k(\boldsymbol{x})\psi_\ell(\boldsymbol{\xi})$ results in the equation

$$\sum_{i,j} \left(\int_\Gamma \rho(\boldsymbol{\xi})\psi_j(\boldsymbol{\xi})\psi_\ell(\boldsymbol{\xi})[\boldsymbol{K}(\boldsymbol{\xi})]_{i,k}\, d\boldsymbol{\xi} \right) u_{i,j}$$
$$= \int_\Gamma \rho(\boldsymbol{\xi})\psi_\ell(\boldsymbol{\xi})[\boldsymbol{f}(\boldsymbol{\xi})]_k\, d\boldsymbol{\xi} \quad \forall k, \ell, \tag{9}$$

where we have introduced the matrix $\boldsymbol{K}(\boldsymbol{\xi})$ and vector $\boldsymbol{f}(\boldsymbol{\xi})$ defined as

$$[\boldsymbol{K}(\boldsymbol{\xi})]_{i,k} := \int_D \kappa(\boldsymbol{x}, \boldsymbol{\xi})\nabla\phi_i(\boldsymbol{x}) \cdot \nabla\phi_k(\boldsymbol{x})\, d\boldsymbol{x},$$
$$i, k = 1, 2, \ldots, N_x, \tag{10}$$

$$[\boldsymbol{f}(\boldsymbol{\xi})]_k := \int_D f(\boldsymbol{x}, \boldsymbol{\xi})\phi_k(x)\, d\boldsymbol{x},$$
$$k = 1, 2, \ldots, N_x. \tag{11}$$

Equation (9) may be viewed as a semidiscrete equation, where we have left the stochastic variables continuous. The matrix $\boldsymbol{K}(\boldsymbol{\xi})$ and vector $\boldsymbol{f}(\boldsymbol{\xi})$, which have the form of the usual finite element stiffness matrix and load vector, are seen to still depend on the random vector $\boldsymbol{\xi}$, i.e. are a random matrix and vector, respectively.

Next, we introduce the $N_x \times N_x$ matrices

$$\boldsymbol{A}_{\ell,j} := \langle \psi_j(\boldsymbol{\xi})\psi_\ell(\boldsymbol{\xi})K(\boldsymbol{\xi}) \rangle, \quad \ell, j = 1, \ldots, N_\xi$$

along with the N_x-dimensional vectors

$$\boldsymbol{f}_\ell = \langle \psi_\ell(\boldsymbol{\xi})\boldsymbol{f}(\boldsymbol{\xi}) \rangle, \quad \ell = 1, \ldots, N_\xi,$$

and, defining the global Galerkin matrix $\boldsymbol{A}$ and vector $\boldsymbol{f}$ by

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{1,1} & \ldots & \boldsymbol{A}_{1,N_\xi} \\ \vdots & & \vdots \\ \boldsymbol{A}_{N_\xi,1} & \ldots & \boldsymbol{A}_{N_\xi,N_\xi} \end{bmatrix}, \quad \boldsymbol{f} = \begin{bmatrix} \boldsymbol{f}_1 \\ \vdots \\ \boldsymbol{f}_{N_\xi} \end{bmatrix}, \tag{12}$$

we obtain the global Galerkin system

$$\boldsymbol{A}\boldsymbol{u} = \boldsymbol{f} \tag{13}$$

with the block vector $\boldsymbol{u}$ of unknowns

$$\boldsymbol{u} = \begin{bmatrix} \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{u}_{N_\xi} \end{bmatrix}, \quad \boldsymbol{u}_j = \begin{bmatrix} u_{1,j} \\ \vdots \\ u_{N_x,j} \end{bmatrix}, \quad j = 1, \ldots, N_\xi.$$

More can be said about the basic structure of the Galerkin system by taking account the precise dependence of the random fields κ and f—and hence the $\boldsymbol{K}$ and $\boldsymbol{f}$—on the random vector $\boldsymbol{\xi}$. We discuss three cases of increasing generality.

3.2.1 Random fields linear in $\boldsymbol{\xi}$

The simplest structure results when the random fields κ and f admit expansions with terms linear in the random variables $\{\xi_m\}_{m=1}^M$. Such is the case, e.g. when the random fields are approximated by a truncated Karhunen–Loève expansion. For Gaussian random fields these random variables are also Gaussian and, being uncorrelated, independent. For nonGaussien fields the independence of the random variables in the KL-expansion is often added as an extra modelling assumption. Thus, if κ and f are of the form

$$\kappa(\boldsymbol{x}, \boldsymbol{\xi}) = \kappa_0(\boldsymbol{x}) + \sum_{m=1}^M \kappa_m(\boldsymbol{x})\xi_m,$$
$$f(\boldsymbol{x}, \boldsymbol{\xi}) = f_0(\boldsymbol{x}) + \sum_{m=1}^M f_m(\boldsymbol{x})\xi_m, \tag{14}$$

the associated matrix $\boldsymbol{K}(\boldsymbol{\xi})$ and vector $\boldsymbol{f}(\boldsymbol{\xi})$ are given by

$$\boldsymbol{K}(\boldsymbol{\xi}) = \boldsymbol{K}_0 + \sum_{m=1}^M \boldsymbol{K}_m\xi_m, \quad \boldsymbol{f}(\boldsymbol{\xi}) = \boldsymbol{f}_0 + \sum_{m=1}^M \boldsymbol{f}_m\xi_m,$$

in terms of matrices $\boldsymbol{K}_m$ and vectors $\boldsymbol{f}_m$, $m = 0, 1, \ldots, M$ given by

$$[\boldsymbol{K}_m]_{i,k} = (\kappa_m\nabla\phi_k, \nabla\phi_i)_{L^2(D)}, \quad [\boldsymbol{f}_m]_k = (f_m, \phi_k)_{L^2(D)},$$
$$i, k = 1, 2, \ldots, N_x.$$

As a result, the global Galerkin matrix $\boldsymbol{A}$ and load vector $\boldsymbol{f}$ from (12) take on the form of sums of Kronecker (tensor) products

$$\boldsymbol{A} = \boldsymbol{G}_0 \otimes \boldsymbol{K}_0 + \sum_{m=1}^M \boldsymbol{G}_m \otimes \boldsymbol{K}_m,$$
$$\boldsymbol{f} = \boldsymbol{g}_0 \otimes \boldsymbol{f}_0 + \sum_{m=1}^M \boldsymbol{g}_m \otimes \boldsymbol{f}_m, \tag{15}$$

with matrices $\boldsymbol{G}_m$ and vectors $\boldsymbol{g}_m$ given in terms of the stochastic basis from (8) and random variables $\{\xi_m\}_{m=1}^M$ as

$$[\boldsymbol{G}_0]_{\ell,k} = \langle \psi_k\psi_\ell \rangle, \qquad [\boldsymbol{g}_0]_\ell = \langle \psi_\ell \rangle, \qquad k, \ell = 1, 2, \ldots, N_\xi,$$
$$[\boldsymbol{G}_m]_{\ell,k} = \langle \xi_m\psi_k\psi_\ell \rangle, \quad [\boldsymbol{g}_m]_\ell = \langle \xi_m\psi_\ell \rangle, \quad k, \ell = 1, 2, \ldots, N_\xi,$$
$$m = 1, 2, \ldots, M.$$

Figure 3 shows examples of the sparsity pattern of the global Galerkin matrix $\boldsymbol{A}$ when the polynomial chaos in Gaussian random variables is used as the stochastic basis. Each square in the matrix corresponds to a position at which one of the matrices $\boldsymbol{G}_m$, $m = 0, 1, \ldots, M$
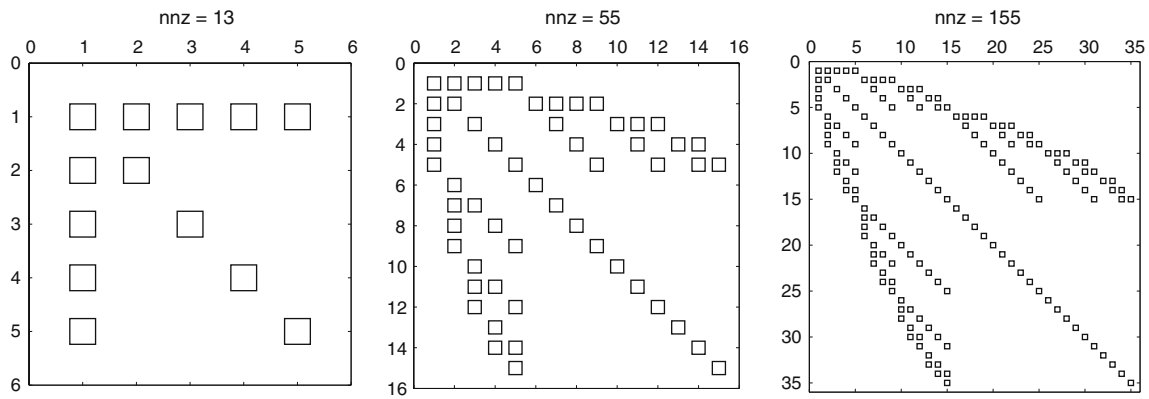
Fig. 3 Sparsity pattern of the sum of the matrices $\{G_m\}_{m=0}^{M}$ for $M = 4$ random variables and a stochastic space consisting of global polynomials of total degree $p = 1$ (left), $p = 2$ (middle) and $p = 3$ (right)

has a nonzero, hence there will be a nonzero $N_x \times N_x$ block at this position in the global matrix A.

In [4] it is shown that, when global polynomials are used for the stochastic variables, then by choosing suitable orthogonal polynomials as basis function the Kronecker factors G_m in (15) all assume diagonal form, resulting in a block diagonal Galerkin matrix A. This reduces the linear algebra problem to that of solving a sequence of uncoupled linear systems of size $N_x \times N_x$. (See also [5,13].)

3.2.2 Expansion in nonindependent random variables

If there is no justification in assuming that the uncorrelated random variables η_m in an expansion of the form

$$\kappa(\boldsymbol{x}, \omega) = \kappa_0(\boldsymbol{x}) + \sum_{m=1}^{M} \kappa_m(\boldsymbol{x})\eta_m(\omega),$$

are independent, then a popular remedy is to perform a KL expansion of each such linearly occurring random variable in a new set of independent Gaussian random variables, say $\boldsymbol{\xi}(\omega) = [\xi_1(\omega), \ldots, \xi_M(\omega)]^\top$, so that

$$\eta_m(\omega) = \sum_r \eta_m^{(r)}\psi_r(\boldsymbol{\xi}), \quad r = 1, 2, \ldots, N_\xi. \tag{16}$$

Such an expansion leads to a Galerkin matrix of the form

$$A = G_0 \otimes K_0 + \sum_m \sum_r \eta_m^{(r)} H_r \otimes K_m, \tag{17}$$

with "stiffness" matrices K_m as in Sect. 3.2.1 and stochastic matrix factors $H_r \in \mathbb{R}^{N_\xi \times N_\xi}$ given by the triple products

$$[H_r]_{\ell,j} = \langle \psi_r \psi_\ell \psi_j \rangle, \quad r, \ell, j = 1, 2, \ldots, N_\xi.$$

In comparison with Sect. 3.2.1, these matrices are less sparse than their counterparts G_m and no change of basis has yet been found in which these matrices simplify. Figure 4 shows some MATLAB spy-plots of the sparsity pattern of the sum of all matrices H_m in the case where $\{\psi_j\}$ is the polynomial chaos basis of bounded total degree.

3.2.3 Expansion in stochastic basis

Finally, an SFEM discretization can do without a KL or KL-like expansion entirely and expand any random fields in whatever basis has been chosen for the stochastic subspace W^h:

$$\kappa(\boldsymbol{x}, \boldsymbol{\xi}) = \kappa_0(\boldsymbol{x}) + \sum_r \kappa_r(\boldsymbol{x})\psi_r(\boldsymbol{\xi}).$$

This approach, along with suggestions for computing the coefficient functions $\kappa_r(\boldsymbol{x})$ are described in [22, Chapter 4].

4 Computational aspects

The typical situation in which the SFEM can be applied is when the stochastic PDE involves random fields with correlation length sufficiently large that their KL expansion yields a good approximation when truncated after a small number M of (say, at most 20) terms. As a result, the SFEM discretization would involve M independent random variables and the stochastic dimension N_ξ of the problem then results from the manner in which the M-fold tensor product space $L_\rho(\Gamma)$ (see Sect. 3.1) is discretized.
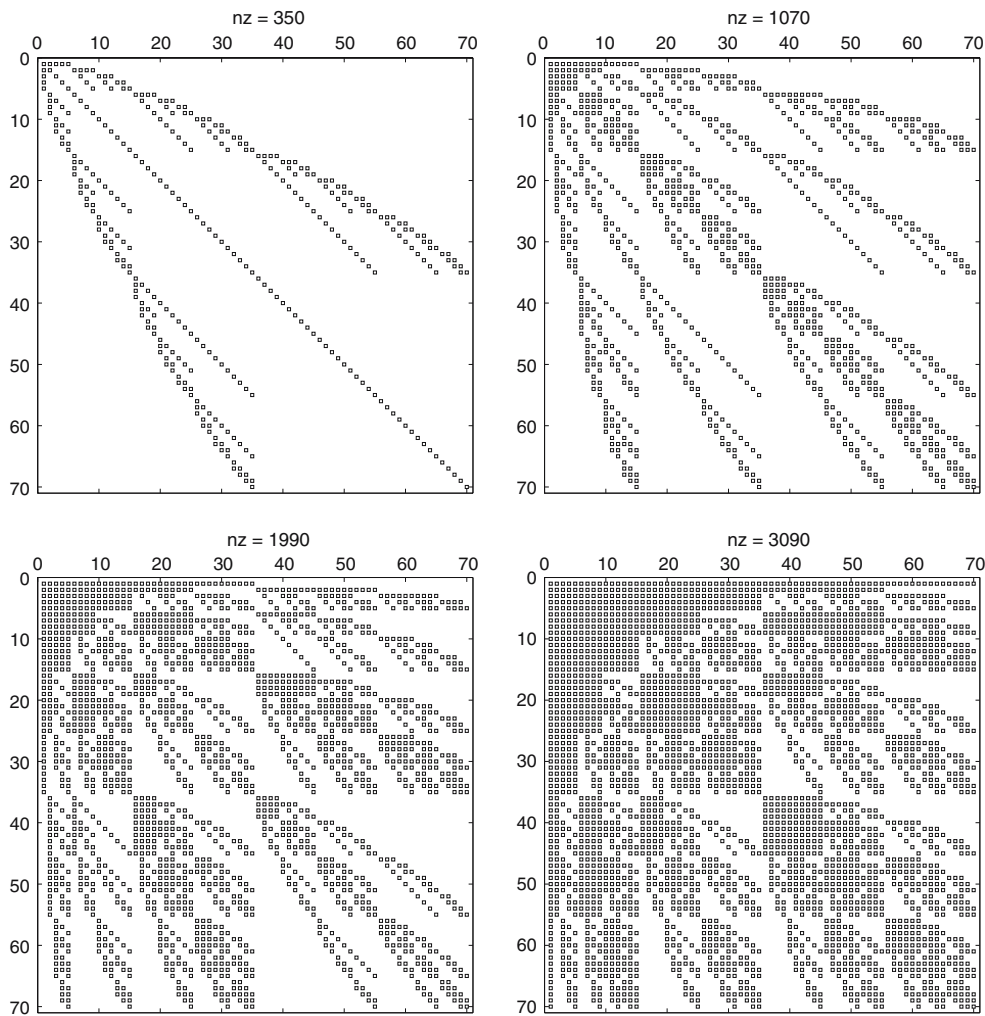
Fig. 4 Sparsity patterns of the sum of the matrices H_r for $M = 4$ random variables and basis functions of total degree $p = 4$. The index r enumerates all polynomials in $M = 4$ variables up to total degree (from *upper left to lower right*) 1, 2, 3 and 4, respectively, corresponding to expansions (16) of different orders

In this section we consider the two main compuation-al tasks involved in the implementation of the SFEM for the model problem (4) and (5), namely the approximate KL expansion, a large eigenvalue problem, and the solution of the Galerkin equations (13).

4.1 The covariance eigenvalue problem

The numerical approximation of the KL expansion (2) of a random field with known covariance function requires an approximation of the eigenpairs of the covariance operator $C : L^2(D) \to L^2(D)$ defined by

$$u \mapsto Cu, \quad (Cu)(x) = \int_D u(y)c(x, y) \, dy, \tag{18}$$

whose kernel function $c(x, y)$ is the covariance of the given random field. We shall consider covariance kernels of the form $c(x, y) = c(\|x - y\|)$ with the real-valued function of a scalar argument c one of the examples shown in Fig. 2. The associated integral operator C is self-adjoint, nonnegative definite and compact, hence the eigenvalues are countable, real, nonnegative and have zero as the only possible point of accumulation. The decay rate of the eigenvalues to zero depends on the smoothness of the kernel function, i.e. an analytic kernel results in exponential decay, whereas finite Sobolev regularity results in an algebraic decay. Moreover, the decay rate increases with the correlation length. See [14] and the references therein for detailed statements on eigenvalue decay of covariance operators.

We consider a Galerkin discretization of the operator C resulting from a finite-dimensional subspace

$$Y^h = \text{span}\{\eta_1, \eta_2, \ldots, \eta_N\} \subset L^2(D). \tag{19}$$

Although one could use the space X^h given in (6), the eigenvalue problem typically has other discretization requirements than the spatial part of the SPDE, so that a separate discretization is warranted. The Galerkin condition

$$(Cu, v) = \lambda(u, v) \quad \forall v \in Y^h, \quad (\cdot, \cdot) = (\cdot, \cdot)_{L^2(D)},$$

is then equivalent to the generalized matrix eigenvalue problem

$$\boldsymbol{Cu} = \lambda \boldsymbol{Mu} \tag{20}$$

with the symmetric and positive semidefinite matrix $\boldsymbol{C}$ and the symmetric positive definite mass matrix $\boldsymbol{M}$ given by

$$[\boldsymbol{C}]_{i,j} = (C\eta_j, \eta_i), \quad [\boldsymbol{M}]_{i,j} = (\eta_j, \eta_i), \quad i, j = 1, 2, \ldots, N.$$

Since the number M of approximate eigenpairs required for the truncated KL expansion is typically much smaller than the dimension N, Krylov subspace methods for eigenvalue problems [32,34] seem promising. Krylov subspace methods require matrix vector products with C and linear system solves with M. For finite element approximations M can often be made (approximately) diagonal, making these system solves inexpensive. The matrix C, however, is in general dense, and the generation, storage and matrix-vector multiplication with this matrix cannot be performed inexpensively in the usual manner. In our work we have used the so-called hierarchical matrix technique (see [6] and the references therein) for these tasks, which, for integral operators with sufficiently smooth kernels, is able to perform them in $O(N \log N)$ operations.

While several authors have proposed using Krylov-subspace algorithms based on the implicitly restarted Arnoldi process (cf. [23]) for the covariance eigenvalue problems, we have found the Lanczos-based thick-restart method (TRL) [44] to be more efficient in this context. Both approaches compute Ritz approximations of eigenpairs with respect to Krylov subspaces of fixed dimension, usually somewhat larger than the number M of desired eigenpairs, which are successively improved by generating a new Krylov space with a cleverly chosen initial vector. The TRL method takes advantage of the symmetry of the problem, resulting in shorter update and restart formulas as well as a posteriori error bounds for the eigenpairs.

4.2 Solving the Galerkin system

The complete SFEM Galerkin Eq. (13) has an $N_\xi \times N_\xi$ block coefficient matrix consisting of blocks of size $N_x \times N_x$. If the deterministic part of the problem itself already contains many degrees of freedom (N_x large), then—even for a moderate value of the stochastic dimension N_ξ—the expense of solving the system can be extremely high unless certain structural properties are taken advantage of.

The problem simplifies considerably if the coefficient function in the differential operator of (4) is deterministic and only source and/or boundary data are random. This case is sometimes called that of a *stochastic right hand side* [9]. In this case the coefficient matrix $\boldsymbol{A}$ in (15) or (17) contains only the factors with index zero, and, if the stochastic basis functions are chosen orthonormal, there results a block diagonal matrix with constant blocks. In [12] a stochastic right hand side problem is treated for an acoustic scattering application with random boundary data. It is shown there how a source field expanded in a KL series with $M+1$ terms permits reducing the global Galerkin problem to a linear system of size $N_x \times N_x$ with $M+1$ right hand sides, and how block Krylov solvers may be applied to solve this multiple right hand side system efficiently. An alternative approach for the stochastic right hand side problem is given in [36].

For the *stochastic left hand side problem*, i.e. when the differential operator has random coefficient functions, one may attack the problem by solving the full coupled block system. The results of iterative solution approaches such as the conjugate gradient method and preconditioning based on hierarchical basis decompositions for coefficient matrices of the type in Sect. 3.2.1 are given in [17] and [33]. For the stochastically linear case described in Sect. 3.2.1, using double-orthogonal polynomials reduces the system to block diagonal form, resulting in N_ξ uncoupled linear systems of size $N_x \times N_x$, i.e. one deterministic problem for each stochastic degree of freedom. This brings the effort for solving the Galerkin system close to the Monte Carlo method. Detailed comparison of the SFEM using double orthogonal polynomials with Monte Carlo finite element calculations are given in [5].

Using double orthogonal polynomials to decouple the system is also attractive for implementation on a parallel computer (see, e.g. [22]), since the uncoupled subproblems are "embarrassingly parallel." Alternatively, one may take advantage of the fact that these linear systems are sometimes closely related, particularly when the stochastic variation is small. Krylov subspace methods which allow one to exploit this fact are discussed in Sect. 5.
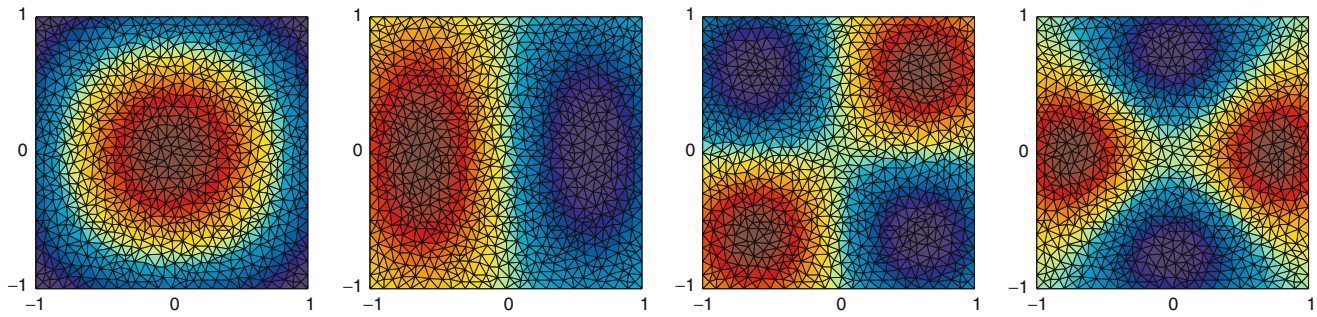
Fig. 5 Approximate eigenfunctions of indices 1,2,4 and 6 of the Bessel covariance operator on the domain $D = [-1, 1]^2$

5 Numerical examples

In this section we present examples of calculations for the covariance eigenproblem and the solution of the Galerkin equations. All calculations were performed in MATLAB version 7 SP 2 on a 1.6 GHz Apple G5 computer with 1.5 GB RAM. The eigenvalue calculations employed the mesh generator of the finite element package FEMLAB as well as Matlab MEX-files based on a preliminary version of the HLIB package for performing hierarchical matrix computations.

5.1 The covariance eigenproblem

We consider the eigenproblem for the covariance operator (18) with Bessel kernel

$$\text{Cov}(\boldsymbol{x}, \boldsymbol{y}) = \frac{r}{a} K_1\left(\frac{r}{a}\right), \quad r = \|\boldsymbol{x} - \boldsymbol{y}\|, \quad a > 0, \quad (21)$$

on the square domain $D = [-1, 1]^2$ with the function space Y^h (19) consisting of piecewise constant functions with respect to a (nonuniform) triangulation of D. The eigenvalues are calculated with the TRL algorithm with fast matrix-vector products computed using the hierarchical matrix approximation. The mass matrix M is diagonal for piecewise constant elements, hence the linear system solve required at every Lanczos step poses no difficulty. Figure 5 shows the approximations to the eigenfunctions with indices $1, 2, 4$ and 6 obtained for $a = 0.2$ and with a mesh consisting of $1, 712$ triangles. Some indication of the performance is given in Table 1, which shows the results of the eigenvalue calculation for variations of the correlation length, mesh size and dimension of the Krylov space. In all cases the parameters used for the hierarchical matrix approximation were an interpolation degree of 4 for the low rank approximation of admissible blocks, an admissibility threshold of one and a minimal block size of 35 for fully calculated block (see [6] for an explanation of these parameters).

Table 1 Performance of the TRL eigenvalue calculations using hierarchical matrix approximation for fast matrix-vector products

m	2,668 DOF	7,142 DOF	14,160 DOF	28,326 DOF
$a = 20,$	$\lambda_{n_e+1}/\lambda_1 < 10^{-3}$	$(n_e = 4)$		
6	18 (3.6)	15 (8.1)	15 (17)	15 (38)
10	3 (1.2)	3 (3.7)	3 (8.5)	3 (19)
12	3 (1.4)	2 (3.6)	2 (8.3)	2 (17)
$a = 2,$	$\lambda_{n_e+1}/\lambda_1 < 10^{-2}$	$(n_e = 8)$		
10	14 (3.1)	14 (8.3)	16 (22)	16 (45)
16	4 (2.1)	4 (6.5)	4 (15)	4 (33)
20	3 (2.2)	3 (6.9)	3 (16)	3 (36)
$a = 0.2,$	$\lambda_{n_e+1}/\lambda_1 < 10^{-2}$	$(n_e = 87)$		
110	5 (18)	5 (84)	5 (154)	5 (296)
120	4 (19)	4 (74)	4 (160)	4 (306)
130	3 (19)	3 (83)	3 (157)	3 (302)

The table gives the number of restart cycles of length m that were necessary to compute the dominant n_e eigenvalues of the covariance operator discretized by piecewise constant elements with various numbers of degrees of freedom (DOF). The numbers in parentheses denote the elapsed time in seconds for the eigenvalue calculation

The table shows the number of restart cycles necessary to calculate the dominant n_e eigenpairs of the discretized covariance eigenvalue problem (20) as well as the elapsed time. The number of eigenvalues n_e to compute was determined by the condition that the magnitude of the eigenvalue with index $n_e + 1$ be at most 1% of that of the largest eigenvalue. In the example of strongest correlation ($a = 20$) we lowered the threshold to 0.1% in order to compute more than a trivial number of eigenpairs. We observe that the TRL algorithm is very robust with respect to the resolution of the integral oparator once the dimension m of the Krylov space is set sufficiently large, in these experiments a small number in addition to the number n_e of desired eigenpairs. We further observe that the solution time for the eigenvalue problem increases linearly with the size of the matrix, which is somewhat better than the asymptotic complexity of $O(N \log N)$ of the hierarchical matrix method.

5.2 Solving the Galerkin system

In this section we present the results of some numerical experiments solving the family of N_ξ linear systems resulting from the block-diagonalization of the global Galerkin system (13) using Krylov solvers designed to reuse information over a sequence of linear systems. We consider the following diffusion-type model problem from [45] posed on the square domain $D = [-1, 1]^2$:

$$-\nabla \cdot \big(\kappa(\boldsymbol{x}, \omega) \nabla u(\boldsymbol{x}, \omega)\big) = f(\boldsymbol{x}, \omega), \quad \boldsymbol{x} \in D, \tag{22a}$$

$$u(\boldsymbol{x}, \omega) = 1, \quad \boldsymbol{x} \in \{-1\} \times [-1, 1], \tag{22b}$$

$$u(\boldsymbol{x}, \omega) = 0, \quad \boldsymbol{x} \in [-1, 1] \times \{-1\}, \tag{22c}$$

$$\frac{\partial u}{\partial \boldsymbol{n}}(\boldsymbol{x}, \omega) = 0, \quad \boldsymbol{x} \in \{1\} \times [-1, 1] \cup [-1, 1] \times \{1\}. \tag{22d}$$

The random fields κ and f are given by their finite KL expansions

$$\kappa(\boldsymbol{x}, \omega) = \kappa_0(\boldsymbol{x}) + \alpha_\kappa \sum_{m=1}^{M} \sqrt{\lambda_m^{(\kappa)}} \kappa_m(\boldsymbol{x}) \xi_m(\omega),$$

$$f(\boldsymbol{x}, \omega) = f_0(\boldsymbol{x}) + \alpha_f \sum_{m=1}^{M} \sqrt{\lambda_m^{(f)}} f_m(\boldsymbol{x}) \xi_m(\omega), \tag{23}$$

with $\{\xi_m\}_{m=1}^{M}$ uncorrelated centered random variables of unit variance, both expansions resulting from a Bessel covariance function (21) with $a = 20$ and mean values $\kappa_0(\boldsymbol{x}) \equiv 1$, $f_0(\boldsymbol{x}) \equiv 0$. The factors α_κ and α_f are chosen such that the resulting fields have a prescribed (constant) standard deviations of, unless specified otherwise, $\sigma_\kappa = \sigma_f = 0.4$. Moreover, we assume the random variables $\{\xi_m\}_{m=1}^{M}$ are fully cross-correlated, i.e. that the random variables occurring in the random fields κ and f are the same. In the experiments that follow we consider both uniformly and normally distributed random variables ξ_m. The stochastic space is spanned by tensor product polynomials of degree p_ξ in M random variables, and we set $p_\xi = 3$ unless stated otherwise. We use double orthogonal polynomials for the stochastic basis functions, so that the Galerkin system is block diagonal.

The spatial discretization uses piecewise tensor product polynomials of degree p_x on $n \cdot n$ axis-parallel rectangular elements with function values at Gauss–Lobatto nodes as the degrees of freedom (GLL spectral elements). To resolve the singularity at the corner $(-1, -1)$ we use a graded rectangular mesh resulting from the tensor product of the nodes $\{-1, -1 + 2\rho^{n-1}, \ldots, -1 + 2\rho, 1\}$ with a grading factor $\rho = 0.15$. In our examples we fix $n = 5$ and $p_x = 13$, resulting in $N_x = 4,225$ spatial degrees of freedom accounting for the essential boundary conditions.

To solve the block diagonal systems efficiently, we employ Krylov subspace methods which solve these block equations in succession and which are designed to reuse search subspaces generated in previous solves. Two such recently introduced methods are known as GCROT and GCRO-DR [10,11,31]. While GCROT selects subspaces based on canonical angles between the search space and its image under the system matrix, GCRO-DR employs subspaces spanned by harmonic Ritz vectors of the system matrix with respect to the search space, i.e. approximate eigenvectors. In all experiments we have used as a preconditioner in these Krylov solvers an incomplete Cholesky decomposition with no fill-in of the mean stiffness matrix $\boldsymbol{K}_0$. The stopping criterion for each solve was a reduction of the (Euclidean) residual norm by a factor of 10^{-8}. The parameters used for GCROT (see [10] for their meanings) were $m = 15$ inner steps, inner selection cutoff of $s = 10$, inner selection parameters $p_1 = 0$, $p_2 = 2$, and outer truncation parameters $k_{\text{thresh}} = 20$, $k_{\text{max}} = 25$ and $k_{\text{new}} = 5$. For GCRO-DR (see [31]) we used $m = 40$ and $k = 25$. These choices results in methods which at any time during the iteration require storage of at most 40 basis vectors and are able to recycle a subspace of dimension up to 25 from one linear system solve to the next.

Table 2 shows the average iteration counts per block equation of the Galerkin system for stochastic polynomial degree $p_\xi = 3$ ($N_\xi = 256$) using various Krylov solvers for both uniformly and normally distributed random variables $\{\xi_m\}$ for variations of the variance of the random fields. We compare the full GMRES method with GCROT with and without subspace recycling as well as GCRO-DR with recycling. We observe that subspace recycling results in considerable savings per system and that, in these examples, the effect is more pronounced for the GCROT method than for GCRO-DR, from which we conclude that subspace angle information across block systems was more useful than spectral information. One also observes that, for both methods, smaller variances lead to larger savings due to subspace recycling. Table 3 gives analogous iteration counts for variation of the number of stochastic degrees of freedom N_ξ, which is seen to have little influence. Finally, in Table 4, we vary the correlation parameter a, adjusting the number of random variables $\{\xi_m\}_{m=1}^{M}$ such that $\lambda_j/\lambda_1 < 0.01$ for all $j > M$. We observe that, although increasing M leads to a strong increase in the number of stochastic degrees of freedom, the average iteration counts increase only modestly.

Table 2 Average iteration counts per linear system for solving the Galerkin system of equations using (full) GMRES, GCROT without recycling as well as both GCROT and GCRO-DR with recycling for $\xi_m \sim U[-\sqrt{3}, \sqrt{3}]$ (top) and $\xi_m \sim N(0, 1)$ (bottom)

$\sigma_\kappa (= \sigma_f)$	GMRES	GCROT	GCROT-re	GCRO-DR-re
Uniform distribution				
0.05	32	34	7	12
0.1	33	34	10	12
0.2	33	34	13	13
0.4	33	34	14	13
0.8	33	34	17	14
Normal distribution				
0.05	33	33	9	12
0.1	33	33	12	13
0.2	33	33	14	13
0.4	33	33	16	13

Table 3 Average iteration counts per linear system for solving the Galerkin system for variations of the number of degrees of freedom in the stochastic space W^h (uniform distribution)

p_ξ	N_ξ	GMRES	GCROT	GCROT-re	GCRO-DR-re
1	16	33	34	16	16
2	81	33	34	15	14
3	256	33	34	14	13
4	625	33	34	14	13
5	1,296	33	34	14	13

Table 4 Average iteration counts per linear system for solving the Galerkin system for variations of the correlation length a (uniform distribution)

a	M	N_ξ	GCROT-re	GCRO-DR-re
20	2	16	14	17
5	4	256	15	14
2	6	4,096	17	15

6 Conclusions

We have given a brief overview of how boundary value problems with random data may be solved using the SFEM and have presented some of the variations of this approach. We have further shown how analytical and structural properties of the resulting linear algebra problems may be exploited to make the otherwise extremely costly computational requirements of implementing the SFEM tractable. Presently, work on efficient implementation of the SFEM is in its early stages and in this work we have only treated the simplest case of random fields with linear expansions in independent random variables, and already there, the linear system and eigenvalue computations are daunting. Consideration of random fields with nonlinear expansions, for which the systems do not decouple, are orders of magnitude more challenging. Yet further challenges are posed by nonlinear stochastic equations, time-dependence as well as inverse problems with stochastic data. In summary, the SFEM is a promising technique, but much remains to be done to make particularly the linear algebra calculations required for its implementation tractable.

References

1. Adler, R.J.: The Geometry of Random Fields. Wiley, London (1981)
2. Adler, R.J., Taylor, J.E.: Random Fields and Geometry. Birkhäuser (2005)
3. Babuška, I., Chleboun, J.: Effects of uncertainties in the domain on the solution of Neumann boundary value problems in two spatial dimensions.. Math. Comput. **71**(240), 1339–1370 (2001)
4. Babuška, I., Tempone, R., Zouraris, G.E.: Galerkin finite element approximations of stochastic elliptic partial differential equations.. SIAM J. Numer. Anal. **42**(2), 800–825 (2004)
5. Babuška, I., Tempone, R., Zouraris, G.E.: Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation.. Comput. Meth. Appl. Mech. Eng. **194**(12–16), 1251–1294 (2005)
6. Börm, S., Grasedyck, L., Hackbusch, W.: Introduction to hierarchical matrices with applications.. Eng Anal Bound. Elem. **27**, 405–422 (2003)
7. Christiakos, G.: Random Field Models in Earth Sciences. Academic, San Diego (1992)
8. Deb, M.K.: Solution of stochastic partial differential equations (spdes) using Galerkin method: theory and applications. Ph.D. thesis, The University of Texas, Austin (2000)
9. Deb, M.K., Babuška, I.M., Oden, J.T.: Solution of stochastic partial differential equations using Galerkin finite element techniques.. Comput. Meth. Appl. Mech. Eng. **190**, 6359–6372 (2001)
10. de Sturler, E.: Truncation strategies for optimal Krylov subspace methods.. SIAM J. Numer. Anal. **36**(3), 864–889 (1999)
11. Eiermann, M., Ernst, O.G., Schneider, O.: Analysis of acceleration strategies for restarted minimal residual methods.. J. Comput. Appl. Math. **123**, 261–292 (2000)
12. Elman, H.C., Ernst, O.G., O'Leary, D.P., Stewart, M.: Efficient iterative algorithms for the stochastic finite element method with applications to acoustic scattering.. Comput. Meth. Appl. Mech. Eng. **194**, 1037–1055 (2005)
13. Ernst, O.G., Ullmann, E.: Efficient iterative solution of stochastic finite element equations. (in press)
14. Frauenfelder, P., Schwab, C., Todor, R.A.: Finite elements for elliptic problems with stochastic coefficients.. Comput. Meth. Appl. Mech. Eng. **194**, 205–228 (2005)
15. Ghanem, R., Spanos, P.D.: Stochastic Finite Elements: A Spectral Approach. Springer, Berlin Heidelberg New York (1991)
16. Ghanem, R.G., Brzkala, V.: Stochastic finite element analysis of randomly layered media.. J. Eng. Mech. **129**(3), 289–303 (1996)
17. Ghanem, R.G., Kruger, R.M.: Numerical solution of spectral stochastic finite element systems.. Comput. Meth. Appl. Mech. Eng. **129**, 289–303 (1996)
18. Hida, T., Potthoff, J., Streit, L.: White Noise Analysis—An Infinite Dimensional Calculus. Kluwer, Dordrecht (1993)

19. Holden, H., ksendal, B., Ube, J., Zhang, T.: Stochastic Partial Differential Equations—A Modeling White Noise Functional Approach. Birkhäuser (1996)
20. Karhunen, K.: Über lineare methoden in der wahrscheinlichkeitsrechnung. . Ann. Acad. Sci. Fenn. Ser. A. I. **37**, 3–79 (1947)
21. Keese, A.: A review of recent developments in the numerical solution of stochastic partial differential equations. Tech. Rep. Informatikbericht Nr.: 2003-6, Universität Braunschweig, Institut für Wissenschaftliches Rechnen (2003)
22. Keese, A.: Numerical solution of systems with stochastic uncertainties—a general purpose framework for stochastic finite elements. Ph.D. thesis, TU Braunschweig, Germany, Fachbereich Mathematik und Informatik (2004)
23. Lehoucq, R.B., Sorensen, D.C., Yang, C.: ARPACK User's Guide: Solution of Large Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods. SIAM, Philadelphia (1998)
24. Liu, J.S.: Monte Carlo. Springer, Berlin Heidelberg New York (2001)
25. Loève, M.: Fonctions aléatoires de second ordre. In: Processus Stochastic et Mouvement Brownien. Gauthier Villars, Paris (1948)
26. Loève, M.: Probability Theory, vol II. Springer, Berlin Heidelberg New York (1977)
27. Matthies, H.G., Brenner, C., Bucher, C.G., Soares, C.G.: Uncertainties in probabilistic numerical analysis of structures and solids—stochastic finite elements.. Struct. Saf. **19**, 283–336 (1997)
28. Matthies, H.G., Bucher, C.: Finite elements for stochastic media problems.. Comput. Meth. Appl. Mech. Eng. **168**, 3–17 (1999)
29. Matthies, H.G., Keese, A.: Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations.. Comput. Meth. Appl. Mech. Eng. **194**, 1295–1331 (2005)
30. Papoulis, A., Pilla, S.U.: Probability, Random Variables and Stochastic Processes, 4th edn. McGraw-Hill, New York (2002)
31. Parks, M., de Sturler, E., Mackey, G., Johnson, D., Maiti, S.: Recycling Krylov subspaces for sequences of linear systems.

Technical Report UIUCDCS-R-2004-2421(CS), University of Illinois (2004)
32. Parlett, B.N.: The Symmetric Eigenvalue Problem. SIAM, Philadelphia (1998)
33. Pellissetti, M.F., Ghanem, R.G.: Iterative solution of systems of linear equations arising in the context of stochastic finite elements.. Adv. Eng. Softw. **31**, 607–616 (2000)
34. Saad, Y.: Numerical Methods for Large Eigenvalue Problems. Halsted Press, New York (1992)
35. Schuëller, G.I.: A state-of-the-art report on computational stochastic mechanics.. Probab. Eng. Mech. **12**(4), 197–321 (1997)
36. Schwab, C., Todor, R.A.: Sparse finite elements for elliptic problems with stochastic data.. Numer. Math. **95**, 707–734 (2003)
37. Sudret, B., Kiureghian, A.D.: Stochastic finite element methods and reliability: a state-of-the-art report. Technical Report on UCB/SEMM-2000/08, Department of Civil and Environmental Engineering, UC Berkeley (2000)
38. Theting, T.G.: Solving Wick–stochastic boundary value problems using a finite element method.. Stoch. Stoch. Rep. **70**(3–4), 241–270 (2000)
39. Trèves, F.: Topological Vector Spaces, Distributions, and Kernels. Academic, New York (1967)
40. Våge, G.: Variational methods for PDEs applied to stochastic partial differential equations. . Math. Scand. **82**, 113–137 (1998)
41. Vanmarcke, E.: Random Fields: Analysis and Synthesis. MIT Press, Cambridge (1983)
42. Whittle, P.: On stationary processes in the plane. Biometrika **41**(434–449) (1954)
43. Wiener, N.: The homogeneous chaos. Am. J. Math. **60**(897–936) (1938)
44. Wu, K., Simon, H.: Thick-restart lanczos method for large symmetric eigenvalue problems.. SIAM J. Matrix Anal. Appl. **22**(2), 602–616 (2000)
45. Xiu, D., Karniadakis, G.E.: Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos.. Comput. Meth. Appl. Mech. Eng. **191**, 4927–4948 (2002)