Efficient Solvers for a Linear Stochastic Galerkin Mixed Formulation of Diffusion Problems with Random Data

Ernst, O.G. and Powell, C.E and Silvester, D.J. and Ullmann, E.

2009

MIMS EPrint: **2007.126**

ISSN 1749-9097

SIAM J. Sci. Comput.
Vol. 31, No. 2, pp. 1424–1447

EFFICIENT SOLVERS FOR A LINEAR STOCHASTIC GALERKIN MIXED FORMULATION OF DIFFUSION PROBLEMS WITH RANDOM DATA[*]

O.G. ERNST[†], C.E. POWELL[‡], D.J. SILVESTER[‡], and E. ULLMANN[†]

Abstract. We introduce a stochastic Galerkin mixed formulation of the steady-state diffusion equation and focus on the efficient iterative solution of the saddle-point systems obtained by combining standard finite element discretizations with two distinct types of stochastic basis functions. So-called mean-based preconditioners, based on fast solvers for scalar diffusion problems, are introduced for use with the minimum residual method. We derive eigenvalue bounds for the preconditioned system matrices and report on the efficiency of the chosen preconditioning schemes with respect to all the discretization parameters.

Key words. stochastic Galerkin method, finite elements, mixed approximation, preconditioning, multigrid

AMS subject classifications. 35R60, 65C20, 65F10, 65N22, 65N30

DOI. 10.1137/070705817

1. Introduction.

In the last few years, interest in stochastic finite element methods (SFEMs) for solving partial differential equations (PDEs) with uncertain data has risen sharply. There currently exists a large body of literature on the stochastic Galerkin formulation of the standard (primal) formulation of the steady-state diffusion equation in which the coefficient is a random field rather than a (deterministic) function. Using SFEMs in the context of solving mixed variational problems, however, is still a relatively new and unexplored field. Mixed formulations pervade in applications with rapidly varying material coefficients (e.g., when modelling groundwater flow or semiconductor devices) and are the motivation for this work.

Our starting point is the following deterministic boundary value problem written as a system of first-order PDEs along with boundary conditions:

$$\begin{aligned} T^{-1}\boldsymbol{q} + \nabla u &= 0, & & \\ \nabla\cdot \boldsymbol{q} &= f & &\text{in } D \subset \mathbb{R}^2, \\ u &= g & &\text{on } \partial D_D \neq \emptyset, \\ \boldsymbol{n}\cdot\boldsymbol{q} &= 0 & &\text{on } \partial D_N = \partial D \backslash \partial D_D. \end{aligned} \tag{1.1}$$

In the context of groundwater flow modelling, (1.1) consists of Darcy's law, coupled with a mass conservation constraint and provides a simplified model for single-phase flow in a saturated porous medium (see, for example, [25, 10]). It is also the so-called

[*] Received by the editors October 19, 2007; accepted for publication (in revised form) September 16, 2008; published electronically January 21, 2009. This research was supported by grant 1279 from the British Council and grant D/06/12721 from the DAAD, as part of the British-German Academic Research Collaboration Program.

http://www.siam.org/journals/sisc/31-2/70581.html

[†] Institut für Numerische Mathematik und Optimierung, Technische Universität Bergakademie Freiberg, D-09596, Freiberg, Germany (ernst@math.tu-freiberg.de, ullmann@math.tu-freiberg.de).

[‡] School of Mathematics, University of Manchester, Manchester, M139PL, United Kingdom (c.powell@manchester.ac.uk, d.silvester@manchester.ac.uk).

mixed formulation of the steady-state diffusion problem

$$\begin{aligned} -\nabla\cdot(T\nabla u) &= f, && \text{in } D\subset\mathbb{R}^2,\\ u &= g && \text{on } \partial D_D \neq \emptyset,\\ \boldsymbol{n}\cdot T\nabla u &= 0 && \text{on } \partial D_N = \partial D\backslash\partial D_D. \end{aligned} \tag{1.2}$$

In this setting the variables u and $\boldsymbol{q} = -T\nabla u$ denote the hydraulic head and volumetric flux, respectively. T is a strictly positive scalar function which is assumed to be known at every point in space. Discretizing (1.1) via mixed finite element techniques allows the simultaneous approximation of the scalar and vector unknowns and is favored over the solution of (1.2) in the presence of rough coefficients when the flux $\boldsymbol{q}$ is the variable of primary interest.

In many applications only limited information about the diffusion coefficient T is actually available. A stochastic approach for modelling this data uncertainty is to consider T to be a random field $T = T(\boldsymbol{x}, \omega)$, i.e., a random function with index variable $\boldsymbol{x} \in \overline{D}$, with respect to a probability space $(\Omega, \mathfrak{A}, P)$, where Ω denotes the abstract set of elementary events, $\mathfrak{A}$ is a σ-algebra on Ω, and P is a probability measure. If $T(\boldsymbol{x}, \omega)$ is bounded and strictly positive, that is, if

$$0 < T_1 \leq T(\boldsymbol{x}, \omega) \leq T_2 < \infty \quad \text{a.e. in } D\times\Omega, \tag{1.3}$$

then (1.1) and (1.2) are well-posed. For a fixed spatial location $\boldsymbol{x} \in D$, $T = T(\omega)$ is a random variable, whilst for a fixed realization $\omega \in \Omega$, $T = T(\boldsymbol{x})$ is a bounded function in $\boldsymbol{x}$ only. As a consequence, the two solution components $(\boldsymbol{q}, u)$ of (1.1) will themselves be random fields. We thus consider the problem of finding two random fields $\boldsymbol{q} = \boldsymbol{q}(\boldsymbol{x}, \omega)$ and $u = u(\boldsymbol{x}, \omega)$ such that, P-almost surely,

$$\begin{aligned} T^{-1}(\boldsymbol{x},\omega)\boldsymbol{q}(\boldsymbol{x},\omega) + \nabla u(\boldsymbol{x},\omega) &= 0,\\ \nabla\cdot\boldsymbol{q}(\boldsymbol{x},\omega) &= f(\boldsymbol{x}) && \text{in } D\times\Omega,\\ u(\boldsymbol{x},\omega) &= g(\boldsymbol{x}) && \text{on } \partial D_D\times\Omega,\\ \boldsymbol{n}\cdot\boldsymbol{q}(\boldsymbol{x},\omega) &= 0 && \text{on } \partial D_N\times\Omega. \end{aligned} \tag{1.4}$$

In this paper, we prescribe only the second-order statistics of the reciprocal field T^{-1}, namely, its mean and covariance functions. We will also make the simplifying assumption that T^{-1} possesses a finite separated expansion of the form

$$T^{-1}(\boldsymbol{x},\omega) = t_0(\boldsymbol{x}) + \sum_{m=1}^{M} t_m(\boldsymbol{x})\xi_m(\omega) \tag{1.5}$$

in terms of $t_0(\boldsymbol{x})$, the expected value of T^{-1} at the point $\boldsymbol{x}$, M specified functions t_m, and M independent random variables ξ_m, each having zero mean and unit variance. Since the dependence of T^{-1} on these random variables is *linear*, we refer to (1.4)–(1.5) as the *stochastically linear formulation.* A popular method for constructing such a linear representation is a truncated Karhunen–Loève expansion. For groundwater flow, see [6], a more realistic model is to assume a lognormal distribution for the permeability field—leading to a *stochastically nonlinear formulation*, in which the dependence on each ξ_m in (1.5) is nonlinear. This latter case will be the focus of a subsequent paper, which builds on the theoretical results and the solver methodology established herein. Note that the source term f and boundary data g can also be treated as random fields in a straightforward manner, but we shall not consider these cases in the present work.

In the next section we extend the usual SFEM framework developed for the stochastic version of the primal problem (1.2) to the mixed problem (1.4). In contrast to traditional Monte Carlo methods, SFEMs discretize the probabilistic dimension of the stochastic PDE directly. If a standard orthonormal basis is used for the stochastic component, we are required to solve a single structured but extremely large saddle-point system. Alternatively, the application of a certain so-called doubly orthogonal stochastic basis requires the solution of multiple decoupled saddle-point systems, each with the dimension of the chosen spatial basis. Details are given in section 3. A comparison of the efficiency of stochastic Galerkin methods and Monte Carlo methods in computing moments of solutions can be found in [2]. It is clear from studies such as this, that if stochastic Galerkin methods are to be competitive with popular sampling techniques, such as Monte Carlo methods and stochastic collocation methods [1, 30], which require multiple solves with small deterministic system matrices, then we need fast and robust linear algebra techniques for solving stochastic Galerkin systems that have optimal complexity. Many authors have studied this for positive definite problems (e.g., [20, 21]). Here we tackle an indefinite problem. Preconditioners based on the mean value of the reciprocal field T^{-1} are constructed and discussed in section 4. In addition, eigenvalue bounds that establish the efficacy of our preconditioning approach are derived. An attractive feature is that the building block of our mean-based preconditioning is a scalar diffusion solve based on an algebraic multigrid V-cycle (see [24, 28]). Numerical experiments that show the efficiency of our methodology are discussed in section 5.

2. Stochastic Galerkin formulation.

To define our SFEM, based on Galerkin approximation of (1.4), we first recall the standard variational formulation of (1.1). Following the usual framework for deterministic mixed approximation as given in [5, 10, 23], we set

$$\begin{gathered} X := \boldsymbol{H}_0(\mathrm{div}; D) \times L^2(D), \qquad \text{with} \\ \boldsymbol{H}_0(\mathrm{div}; D) := \left\{ \boldsymbol{r} \in L^2(D)^2 : \nabla\cdot \boldsymbol{r} \in L^2(D), \boldsymbol{n}\cdot\boldsymbol{r}|_{\partial D_N} = 0 \right\} \end{gathered}$$

and seek $(\boldsymbol{q}, u) \in X$ such that

$$\begin{aligned} a(\boldsymbol{q},\boldsymbol{r}) + b(\boldsymbol{r},u) &= -(g, \boldsymbol{n}\cdot\boldsymbol{r})_{\partial D_D} \quad && \forall \boldsymbol{r} \in \boldsymbol{H}_0(\mathrm{div}; D), \\ b(\boldsymbol{q}, v) &= -(f, v) && \forall v \in L^2(D), \end{aligned} \tag{2.1}$$

with bilinear forms $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ defined by

$$\begin{aligned} a(\boldsymbol{q},\boldsymbol{r}) := \int_D T^{-1}\boldsymbol{q}\cdot\boldsymbol{r}\,d\boldsymbol{x}, \qquad b(\boldsymbol{r}, v) \\ := -\int_D v(\nabla\cdot \boldsymbol{r})\,d\boldsymbol{x}, \qquad \boldsymbol{q},\boldsymbol{r} \in \boldsymbol{H}_0(\mathrm{div}; D),\ v \in L^2(D). \end{aligned}$$

To obtain the stochastic formulation of (2.1), we introduce the space $L^2_P(\Omega)$ of all random variables on the probability space $(\Omega, \mathfrak{A}, P)$ with finite variance and we assume that the input random field satisfies $T^{-1}(\boldsymbol{x},\cdot) \in L^2_P(\Omega)\ \forall \boldsymbol{x} \in D$. Moreover, we let $\langle\xi\rangle$ denote the expectation of a random variable $\xi = \xi(\omega) \in L^2_P(\Omega)$. Finally, introducing the tensor product space

$$X \otimes L^2_P(\Omega) = \boldsymbol{V} \times W, \qquad \boldsymbol{V} := \boldsymbol{H}_0(\mathrm{div}; D) \otimes L^2_P(\Omega), \quad W := L^2(D) \otimes L^2_P(\Omega),$$

we arrive at the stochastic variational problem of determining a pair of random fields $\boldsymbol{q} \in \boldsymbol{V}$ and $u \in W$ such that

$$\begin{aligned} \langle a(\boldsymbol{q}, \boldsymbol{r})\rangle + \langle b(\boldsymbol{r}, u)\rangle &= -\left\langle (g, \boldsymbol{n} \cdot \boldsymbol{r})_{\partial D_D}\right\rangle \quad && \forall \boldsymbol{r} \in \boldsymbol{V}, \\ \langle b(\boldsymbol{q}, v)\rangle &= -\left\langle (f, v)\right\rangle && \forall v \in W. \end{aligned} \tag{2.2}$$

The well-posedness of (2.2) can be established using the general framework for the analysis of saddle-point problems given in [5, Chapter II]. To this end, note first that under assumption (1.3), both $\langle a(\cdot, \cdot)\rangle$ and $\langle b(\cdot, \cdot)\rangle$ are continuous bilinear forms on $\boldsymbol{V} \times \boldsymbol{V}$ and $\boldsymbol{V} \times W$, respectively, with respect to the norms

$$\|\boldsymbol{r}\|_{\boldsymbol{V}} := \left\langle \|\boldsymbol{r}\|^2_{\boldsymbol{H}(\mathrm{div};D)} \right\rangle^{1/2}, \quad \boldsymbol{r} \in V \quad \text{ and } \quad \|v\|_W := \left\langle \|v\|^2_{L^2(D)} \right\rangle^{1/2}, \quad v \in W,$$

where, as usual, $\|\boldsymbol{r}\|^2_{\boldsymbol{H}(\mathrm{div};D)} = \int_D \boldsymbol{r}\cdot\boldsymbol{r} + (\nabla\cdot\boldsymbol{r})^2 \, d\boldsymbol{x}$. Next, we introduce the null-space

$$\boldsymbol{V}_0 := \{\boldsymbol{r} \in \boldsymbol{V} : \langle b(\boldsymbol{r}, w)\rangle = 0 \; \forall w \in W\}$$

associated with $\langle b(\cdot, \cdot)\rangle$ and note that $\langle (\nabla \cdot \boldsymbol{r}, \nabla \cdot \boldsymbol{r})\rangle = 0$ if $\boldsymbol{r} \in \boldsymbol{V}_0$. Using (1.3) we deduce that, for all $\boldsymbol{r} \in \boldsymbol{V}_0$,

$$\langle a(\boldsymbol{r}, \boldsymbol{r})\rangle = \left\langle \left(T^{-1}\boldsymbol{r}, \boldsymbol{r}\right)\right\rangle \geq T_1^{-1} \left\langle (\boldsymbol{r}, \boldsymbol{r})\right\rangle = T_1^{-1}\|\boldsymbol{r}\|^2_{\boldsymbol{V}} \tag{2.3}$$

and hence that $\langle a(\cdot, \cdot)\rangle$ is coercive on $\boldsymbol{V}_0$. Finally, to verify the inf-sup stability condition in [5], we need to establish an intermediate result.

LEMMA 2.1. *For all $w \in W$, there exists a unique $\boldsymbol{v} \in \boldsymbol{V}$ and a constant C such that*

$$\|\boldsymbol{v}\|_{\boldsymbol{V}} \leq C \, \|w\|_W. \tag{2.4}$$

Proof. Given $w \in W$, there exists a unique $s \in H^1(D) \otimes L^2(\Omega)$, which is the solution to the stochastic right-hand side problem

$$-\nabla\cdot \nabla s = w \qquad \text{in } D \times \Omega, \tag{2.5a}$$

$$s = 0 \qquad \text{on } \partial D_D \times \Omega, \tag{2.5b}$$

$$\boldsymbol{n} \cdot \nabla s = 0 \qquad \text{on } \partial D_N \times \Omega \tag{2.5c}$$

(see [7] or [2] and deterministic analysis in [5, p. 136]) and which satisfies

$$\left\langle \|s\|^2_{H^1(D)} \right\rangle \leq C \left\langle \|w\|^2_{L^2(D)} \right\rangle, \tag{2.6}$$

with some constant C depending only on D. Setting $\boldsymbol{v} := -\nabla s$, we note that $\boldsymbol{v} \in L^2(D)^2 \otimes L^2_P(\Omega)$ since $s \in H^1(D) \otimes L^2_P(\Omega)$. Moreover, $\nabla\cdot \boldsymbol{v} = w \in W$ because of (2.5a) and $\boldsymbol{n} \cdot \boldsymbol{v} = 0$ on $\partial D_N \times \Omega$ from (2.5c), and therefore, $\boldsymbol{v} \in \boldsymbol{V}$. Now, using (2.6) gives

$$\|\boldsymbol{v}\|^2_V = \left\langle \|\nabla s\|^2_{L^2(D)} + \|w\|^2_{L^2(D)} \right\rangle \leq C \left\langle \|w\|^2_{L^2(D)} \right\rangle = C \, \|w\|^2_W,$$

yielding (2.4). □

For any $w \in W$ with $\boldsymbol{v} \in \boldsymbol{V}$ as in Lemma 2.1, it now follows that

$$\sup_{r \in \boldsymbol{V}} \frac{\langle b(\boldsymbol{r}, w)\rangle}{\|\boldsymbol{r}\|_{\boldsymbol{V}}} \geq \frac{\langle b(\boldsymbol{v}, w)\rangle}{\|\boldsymbol{v}\|_{\boldsymbol{V}}} = \frac{\langle (\nabla\cdot \boldsymbol{v}, w)\rangle}{\|\boldsymbol{v}\|_{\boldsymbol{V}}} = \frac{\langle (w, w)\rangle}{\|\boldsymbol{v}\|_{\boldsymbol{V}}} = \frac{\|w\|^2_W}{\|\boldsymbol{v}\|_{\boldsymbol{V}}} \geq \frac{1}{C}\|w\|_W. \tag{2.7}$$

Results (2.7) and (2.3) ensure that a solution to (2.2) exists and is unique.

2.1. Finite-dimensional noise. Following a by now well-established approach for the discretization of stochastic boundary value problems [2, 3, 11, 17, 18, 27], we make the assumption that the input random field $T^{-1}(\boldsymbol{x},\omega)$ can be represented as a function of a finite number $M \in \mathbb{N}$ of independent random variables $\xi_1,\dots,\xi_M \in L^2_P(\Omega)$, which is often referred to as finite-dimensional noise. Although such a functional dependence can take on many forms (see, e.g., [17, 18]), in this paper we focus on the truncated Karhunen–Loève (KL) expansion [15, 27]

$$T^{-1}(\boldsymbol{x},\omega) = t_0(\boldsymbol{x}) + \sigma \sum_{m=1}^{M} \sqrt{\lambda_m}\, t_m(\boldsymbol{x})\xi_m(\omega). \tag{2.8}$$

In (2.8), the random variables are uncorrelated and have zero mean and unit variance, $t_0(\boldsymbol{x}) = \langle T^{-1}(\boldsymbol{x},\cdot)\rangle$ is the expected value of the random field at the point $\boldsymbol{x} \in D$, and $\{(\lambda_m, t_m)\}_{m=1}^{M}$ are the leading eigenpairs of the integral operator

$$C : L^2(D) \to L^2(D), \qquad (Cu)(\boldsymbol{x}) = \int_D u(\boldsymbol{y})c(\boldsymbol{x},\boldsymbol{y})\,d\boldsymbol{y}, \quad u \in L^2(D),$$

whose kernel function c is given by

$$c(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{\sigma^2}\left\langle \left(T^{-1}(\boldsymbol{x},\cdot) - t_0(\boldsymbol{x})\right)\left(T^{-1}(\boldsymbol{y},\cdot) - t_0(\boldsymbol{y})\right)\right\rangle, \qquad \boldsymbol{x},\boldsymbol{y} \in D.$$

The parameter σ is a scalar measure of the fluctuation of T^{-1} around its mean value $t_0(\boldsymbol{x})$. If the variance of T^{-1} is constant on D, then it is equal to σ^2 and in this case, the kernel function c is simply the correlation function associated with T^{-1}. If the kernel function is continuous, then the self-adjoint nonnegative-definite operator C is compact and the eigenvalues, assumed in decreasing order, are nonnegative and decay to zero, with the decay rate depending on the smoothness of c. Assuming further that $t_0(\boldsymbol{x}) \equiv \mu$ and $c(\boldsymbol{x},\boldsymbol{x}) \equiv 1$, which is the case, e.g., if the field is homogeneous, then with the eigenfunctions normalized such that $\|t_m\|_{L^2(D)} = 1$, there holds $\sum_{m=1}^{\infty} \lambda_m = |D|$, and the truncation index M can be chosen such that the truncated KL expansion retains a given amount of the field's total variance $\sigma^2 \int_D c(\boldsymbol{x},\boldsymbol{x})\,d\boldsymbol{x}$.

In geostatistics it is common to assume a given correlation structure, and we mention the three popular choices:

$$c(\boldsymbol{x},\boldsymbol{y}) = \exp\left(-\frac{|x_1 - y_1|}{\tau_1} - \frac{|x_2 - y_2|}{\tau_2}\right), \tag{2.9a}$$

$$c(\boldsymbol{x},\boldsymbol{y}) = \exp\left(-\frac{r}{\tau}\right), \tag{2.9b}$$

$$c(\boldsymbol{x},\boldsymbol{y}) = \frac{r}{\tau}K_1\left(\frac{r}{\tau}\right), \tag{2.9c}$$

where r is the Euclidean distance between $\boldsymbol{x}$ and $\boldsymbol{y}$, τ_1, τ_2, and τ are correlation lengths, and K_1 denotes the modified Bessel function of second kind and order one. Many authors (e.g., [12]) use (2.9a) because explicit formulae for the eigenvalues and eigenfunctions exist. For the other choices, the eigenproblem has to be solved numerically. See [8] for further details.

To obtain well-posed Galerkin discretizations of the stochastic boundary value problem (1.4), one could assume that two-sided bounds as in (1.3) hold also for the truncated KL expansion (2.8). For a continuous covariance function, the KL expansion converges only in $L^\infty(D) \otimes L^2_P(\Omega)$, in the sense that

$$\sup_{\boldsymbol{x}\in D} \left\langle \left(T^{-1} - T_M^{-1}\right)^2 \right\rangle \to 0 \text{ as } M \to \infty,$$

and such bounds will, in general, not hold without further assumptions. In [11] additional regularity conditions on the covariance function are shown to assure uniform convergence on $D \times \Omega$, which together with (1.3), yield a similar two-sided bound for (2.8) for a sufficiently large truncation index M. An alternative approach proposed in [18], which only requires (1.3) for the full (nontruncated) random field is to observe that coercivity of the continuous problem implies that of the discrete problem obtained by Galerkin projection onto finite-dimensional subspaces. The orthogonal polynomials we shall use to construct finite element subspaces will, by orthogonality, yield the same Galerkin matrices for the full KL expansion as for its truncation after M terms if the stochastic finite element space is based on these M random variables. Therefore, uniform coercivity follows from (1.3) in this case.

Although the random variables occurring in the KL expansion of a random field are, in general, only uncorrelated, we shall make the stronger assumption that they are independent. These two properties are equivalent for Gaussian random fields. For Gaussian fields, however, boundedness assumption (1.3) fails to hold. One may achieve (1.3) by assuming that the random variables in (2.8) are independent with given distributions, i.e., by introducing independence as a modelling assumption. A simple choice is, e.g., M independent uniformly distributed random variables on the interval $[-\sqrt{3}, \sqrt{3}]$, which have mean zero and unit variance.

Having restricted the variability of the input data, and hence the solution $(\boldsymbol{q}, u)$ of (2.2), to an M-dimensional random vector $\boldsymbol{\xi} = (\xi_1, \dots, \xi_m)$, we may, in view of the Doob–Dynkin lemma, introduce $\boldsymbol{\xi}$ as a new independent random variable in place of ω and write $T^{-1}(\boldsymbol{x}, \boldsymbol{\xi})$, $\boldsymbol{q}(\boldsymbol{x}, \boldsymbol{\xi})$, and $u(\boldsymbol{x}, \boldsymbol{\xi})$. Moreover, setting $\Gamma_m := \xi_m(\Omega)$, $m = 1, \dots, M$, we denote by $\Gamma := \Gamma_1 \times \cdots \times \Gamma_M$ the range of the random vector $\boldsymbol{\xi}$. If, furthermore, each random variable ξ_m possesses the density function $\rho_m : \Gamma_m \to \mathbb{R}_0^+$, we may replace $L^2_P(\Omega)$ by the weighted L^2-space $L^2_\rho(\Gamma)$, where the weight function

$$\rho(\boldsymbol{\xi}) := \rho(\xi_1) \cdots \rho_M(\xi_M) \tag{2.10}$$

is the joint density function of the independent random variables $\xi_1, \dots, \xi_M$. The variational spaces in (2.2) thus become $\boldsymbol{V} = \boldsymbol{H}_0(\mathrm{div}; D) \otimes L^2_\rho(\Gamma)$ and $W = L^2(D) \otimes L^2_\rho(\Gamma)$, with norms

$$\|\boldsymbol{r}\|_{\boldsymbol{V}} = \left(\int_\Gamma \|\boldsymbol{r}\|^2_{\boldsymbol{H}(\mathrm{div};D)}\, \rho(\boldsymbol{\xi}) d\boldsymbol{\xi}\right)^{1/2}, \ \boldsymbol{r} \in \boldsymbol{V},$$

$$\|v\|_W = \left(\int_\Gamma \|v\|^2_{L^2(D)}\, \rho(\boldsymbol{\xi}) d\boldsymbol{\xi}\right)^{1/2}, \ v \in W.$$

2.2. Galerkin approximation. The restriction of the stochastic variability to finite-dimensional noise reduces the stochastic saddle-point problem (2.2) to a deterministic saddle-point problem with the M-dimensional parameter $\boldsymbol{\xi}$. The Galerkin discretization is then obtained in the usual way by restricting trial and test functions in (2.2) to suitable finite-dimensional subspaces of the tensor product spaces $\boldsymbol{V}$ and W, constructed by selecting finite-dimensional subspaces of the component spaces $\boldsymbol{H}_0(\mathrm{div}; D)$, $L^2(D)$, and $L^2_\rho(\Gamma)$. When choosing subspaces, we need to ensure that the discrete analogue of (2.7) holds. To this end, the first two subspaces must be chosen in a compatible way to ensure inf-sup stability of the discrete deterministic saddle-point problem. $L^2_\rho(\Gamma)$ may be discretized independently.

Thus, denoting these subspaces in terms of suitable bases

$$\begin{aligned}\boldsymbol{\Phi}_h &:= \operatorname{span}\{\boldsymbol{\varphi}_i : i = 1, \dots, N_{\boldsymbol{q}}\} \subset \boldsymbol{H}_0(\operatorname{div}; D),\\ \Phi_h &:= \operatorname{span}\{\phi_i : i = 1, \dots, N_u\} \subset L^2(D),\\ \Psi_p &:= \operatorname{span}\{\psi_i : i = 1, \dots, N_{\boldsymbol{\xi}}\} \subset L^2_\rho(\Gamma),\end{aligned}$$

in which the subscripts h and p refer to discretization parameters, we arrive at

$$\begin{aligned}\boldsymbol{V}_{h,p} &= \Big\{\boldsymbol{r}(\boldsymbol{x}, \boldsymbol{\xi}) \in \operatorname{span}\{\boldsymbol{\varphi}_i(\boldsymbol{x})\psi_j(\boldsymbol{\xi}) : i = 1, \dots, N_{\boldsymbol{q}}; j = 1, \dots, N_{\boldsymbol{\xi}}\}\Big\} = \boldsymbol{\Phi}_h \otimes \Psi_p,\\ W_{h,p} &= \Big\{w(\boldsymbol{x}, \boldsymbol{\xi}) \in \operatorname{span}\{\phi_i(\boldsymbol{x})\psi_j(\boldsymbol{\xi}) : i = 1, \dots, N_u; j = 1, \dots, N_{\boldsymbol{\xi}}\}\Big\} = \Phi_h \otimes \Psi_p,\end{aligned}$$

resulting in a total number of degrees of freedom $\dim(\boldsymbol{V}_{h,p} \times W_{h,p}) = N_{\boldsymbol{q}} N_{\boldsymbol{\xi}} + N_u N_{\boldsymbol{\xi}} = N_{\boldsymbol{x}} N_{\boldsymbol{\xi}}$, where $N_{\boldsymbol{x}} := N_{\boldsymbol{q}} + N_u$ denotes the total number of deterministic degrees of freedom. We thus arrive at the discrete version of problem (2.2) and seek $\boldsymbol{q}_{h,p} \in \boldsymbol{V}_{h,p}$ and $u_{h,p} \in W_{h,p}$ such that

$$\begin{aligned}\langle a(\boldsymbol{q}_{h,p}, \boldsymbol{r})\rangle + \langle b(\boldsymbol{r}, u_{h,p})\rangle &= -\langle (g, \boldsymbol{n} \cdot \boldsymbol{r})_{\partial D_D}\rangle \quad &\forall \boldsymbol{r} \in \boldsymbol{V}_{h,p},\\ \langle b(\boldsymbol{q}_{h,p}, w)\rangle &= -\langle (f, w)\rangle \quad &\forall w \in W_{h,p}.\end{aligned} \tag{2.11}$$

For the subspaces $\boldsymbol{\Phi}_h$ and Φ_h, we will use the the lowest-order Raviart–Thomas mixed approximation (see [23]) based on a partition $\mathcal{T}_h$ of the spatial domain D into triangles or rectangles of maximal diameter $h > 0$. More precisely, given a partition $\mathcal{T}_h$ of D into triangles, we set

$$\boldsymbol{\Phi}_h := \left\{\boldsymbol{q} \in \boldsymbol{H}_0(\operatorname{div}; D) : \boldsymbol{q}|_K \in \mathcal{P}_0(K)^2 + \boldsymbol{x}\mathcal{P}_0(K) \ \forall K \in \mathcal{T}_h\right\},$$

where $\mathcal{P}_0(K)$ denotes the space of constant functions on element K. For rectangular partitions, the corresponding Raviart–Thomas space is

$$\boldsymbol{\Phi}_h := \{\boldsymbol{q} \in \boldsymbol{H}_0(\operatorname{div}; D) : \boldsymbol{q}|_K \in \mathcal{Q}_{1,0}(K) \times \mathcal{Q}_{0,1}(K) \ \forall K \in \mathcal{T}_h\},$$

where $\mathcal{Q}_{j,k}$ denotes polynomials of degree j in the first spatial variable and k in the second. In both cases this amounts to constructing a vector field that is piecewise linear in each component and which has a continuous normal component across the edges of the elements of $\mathcal{T}_h$.

As subspaces Ψ_p of $L^2_\rho(\Gamma)$, we employ global M-variate polynomials on Γ. The degree p of these polynomials can be chosen in a variety of ways, with implications for the resulting number of degrees of freedom as well as the structure of the linear system to be solved. Using *tensor product polynomials*, i.e., polynomials of degree at most p separately in each of the M variables, results in $\dim \Psi_p = (p + 1)^M$, an exponential growth of the number of degrees of freedom with M. The major advantage of tensor product polynomials (which are discussed in [2, 3, 8, 14]) is that this space possesses a basis with respect to which the global Galerkin system associated with (2.11) is block-diagonal. It, therefore, decouples into $N_{\boldsymbol{\xi}}$ systems of dimension $N_{\boldsymbol{x}}$. Recently, there have been attempts to reduce the large dimension N_ξ of Ψ_p while retaining the block diagonal structure of the global system matrix. Investigations based on exploiting regularity of the solutions that involve adaptively choosing different polynomial degrees $p_1, \dots, p_M$ in each of the M variables are presented in [3, 11, 16]. Stochastic collocation methods, in which the number of stochastic degrees of freedom can be even further reduced using the techniques of sparse grids and Smolyak quadrature (cf. [30, 1]), are also becoming popular. However, performing stochastic collocation on the mixed problem, with a particular choice of collocation points, leads to the same set of decoupled saddle-point systems encountered in section 3.2. (See Remark 3.1.) The mean-based preconditioner proposed in section 4.4 is suitable for stochastic collocation systems under the same conditions presented below for decoupled stochastic Galerkin systems.

An alternative to tensor product polynomials, that leads to only polynomial growth in the number of stochastic degrees of freedom, is to employ *complete polynomials*, i.e., polynomials in M variables of *total* degree p. In this case, we obtain $\dim \Psi_p = \binom{M+p}{p}$. As shown in [9], there is no basis of this space for which the stochastic degrees of freedom decouple, and therefore, a global system involving all $N_{\boldsymbol{x}} N_{\boldsymbol{\xi}}$ degrees of freedom must be solved. This is often perceived as a serious drawback. Our results in section 4 demonstrate, however, that preconditioning makes the solution of such a coupled system feasible computationally.

3. Matrix properties.

In this section we examine the structure of the linear system of equations associated with stochastic Galerkin equations (2.11).

3.1. Kronecker product representation.

Inserting representation (2.8) of the input random field T^{-1} and the trial functions

$$\boldsymbol{q}_{h,p}(\boldsymbol{x}, \boldsymbol{\xi}) = \sum_{\ell=1}^{N_{\boldsymbol{\xi}}} \sum_{j=1}^{N_q} q_{j,\ell}\, \boldsymbol{\varphi}_j(\boldsymbol{x}) \psi_\ell(\boldsymbol{\xi}), \qquad u_{h,p}(\boldsymbol{x}, \boldsymbol{\xi}) = \sum_{\ell=1}^{N_{\boldsymbol{\xi}}} \sum_{j=1}^{N_u} u_{j,\ell}\, \phi_j(\boldsymbol{x}) \psi_\ell(\boldsymbol{\xi}) \tag{3.1}$$

as well as the basis of test functions $\boldsymbol{r}(\boldsymbol{x}, \boldsymbol{\xi}) = \boldsymbol{\varphi}_i(\boldsymbol{x}) \psi_k(\boldsymbol{\xi})$, $i = 1, \dots, N_{\boldsymbol{q}}; k = 1, \dots, N_{\boldsymbol{\xi}}$, and $v(\boldsymbol{x}, \boldsymbol{\xi}) = \phi_i(\boldsymbol{x}) \psi_k(\boldsymbol{\xi})$, $i = 1, \dots, N_u$, $k = 1, \dots, N_{\boldsymbol{\xi}}$ into stochastic Galerkin equations (2.11) results in the matrix saddle-point problem

$$\begin{bmatrix} \hat{A} & \hat{B}^\top \\ \hat{B} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{q} \\ \boldsymbol{u} \end{bmatrix} = \begin{bmatrix} \boldsymbol{g} \\ \boldsymbol{f} \end{bmatrix}, \tag{3.2}$$

in which the solution vector consists of the two block vectors

$$\boldsymbol{q} = \begin{bmatrix} \boldsymbol{q}_1 \\ \vdots \\ \boldsymbol{q}_{N_{\boldsymbol{\xi}}} \end{bmatrix} \in \mathbb{R}^{N_q N_{\boldsymbol{\xi}}}, \qquad \boldsymbol{u} = \begin{bmatrix} \boldsymbol{u}_1 \\ \vdots \\ \boldsymbol{u}_{N_{\boldsymbol{\xi}}} \end{bmatrix} \in \mathbb{R}^{N_u N_{\boldsymbol{\xi}}}, \tag{3.3}$$

of which each block, in turn, is the size of the corresponding deterministic block, i.e, for $k = 1, 2, \dots, N_{\boldsymbol{\xi}}$, we have

$$[\boldsymbol{q}_\ell]_j = q_{j,\ell}, \quad j = 1, \dots, N_{\boldsymbol{q}}, \qquad [\boldsymbol{u}_\ell]_j = u_{j,\ell}, \quad j = 1, \dots, N_u, \qquad \ell = 1, 2, \dots, N_{\boldsymbol{\xi}}.$$

An analogous representation holds for the two blocks $\boldsymbol{f}$ and $\boldsymbol{g}$ on the right-hand side of (3.2), each comprising the $N_{\boldsymbol{\xi}}$ subblocks

$$[\boldsymbol{g}_k]_i = -\langle \psi_k \rangle (g, \boldsymbol{n}\cdot\boldsymbol{\varphi}_i)_{\partial D_D}, \quad i = 1, \dots, N_{\boldsymbol{q}}, \qquad [\boldsymbol{f}_k]_i = -\langle \psi_k \rangle (f, \phi_i), \quad i = 1, \dots, N_u. \tag{3.4}$$

The block matrices $\hat{A}$ and $\hat{B}$ in (3.2) are, using the double-indexing for rows and columns introduced in (3.1), given by

$$\begin{aligned}[\hat{A}]_{(i,k),(j,\ell)} &= \left\langle \left(T^{-1}\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_i\right) \psi_\ell \psi_k \right\rangle \\ &= (t_0 \boldsymbol{\varphi}_j, \boldsymbol{\varphi}_i) \left\langle \psi_\ell \psi_k \right\rangle + \sigma \sum_{m=1}^{M} \sqrt{\lambda_m} (t_m \boldsymbol{\varphi}_j, \boldsymbol{\varphi}_i) \left\langle \xi_m \psi_\ell \psi_k \right\rangle , \\ & \qquad i,j = 1,\dots,N_{\boldsymbol{q}},\ k,\ell = 1,\dots,N_{\boldsymbol{\xi}}\end{aligned} \tag{3.5a}$$

and

$$[\hat{B}]_{(i,k),(j,\ell)} = - \left\langle (\nabla\cdot \boldsymbol{\varphi}_i, \phi_j) \psi_\ell \psi_k \right\rangle = -(\nabla\cdot \boldsymbol{\varphi}_i, \phi_j) \left\langle \psi_\ell \psi_k \right\rangle , \quad i = 1,\dots,N_{\boldsymbol{q}},\ j = 1,\dots,N_u,\ k,\ell = 1,\dots,N_{\boldsymbol{\xi}}. \tag{3.5b}$$

The bilinear structure implicit in (3.5), due to the fact that the integrals with respect to $\boldsymbol{\xi}$ and $\boldsymbol{x}$ can be separated, which in turn is a consequence of the separation of these two variables in expansion (2.8), allows these matrices to be expressed as sums of Kronecker products

$$\hat{A} = G_0 \otimes A_0 + \sum_{m=1}^{M} G_m \otimes A_m, \qquad \hat{B} = G_0 \otimes B, \tag{3.6}$$

the factors of which are given by

$$A_0 \in \mathbb{R}^{N_q \times N_q}, \quad [A_0]_{i,j} = (t_0 \boldsymbol{\varphi}_j, \boldsymbol{\varphi}_i), \qquad i,j = 1,\dots,N_{\boldsymbol{q}}, \tag{3.7a}$$

$$A_m \in \mathbb{R}^{N_q \times N_q}, \quad [A_m]_{i,j} = \sigma\sqrt{\lambda_m} (t_m \boldsymbol{\varphi}_j, \boldsymbol{\varphi}_i), \quad i,j = 1,\dots,N_{\boldsymbol{q}},\ m = 1,\dots,M, \tag{3.7b}$$

$$B \in \mathbb{R}^{N_u \times N_q}, \quad [B]_{i,j} = -(\nabla\cdot \boldsymbol{\varphi}_i, \phi_j), \qquad i = 1,\dots,N_u,\ j = 1,\dots,N_{\boldsymbol{q}}, \tag{3.7c}$$

$$G_0 \in \mathbb{R}^{N_{\boldsymbol{\xi}} \times N_{\boldsymbol{\xi}}}, \quad [G_0]_{k,\ell} = \left\langle \psi_\ell \psi_k \right\rangle , \qquad k,\ell = 1,\dots,N_{\boldsymbol{\xi}}, \tag{3.7d}$$

$$G_m \in \mathbb{R}^{N_{\boldsymbol{\xi}} \times N_{\boldsymbol{\xi}}}, \quad [G_m]_{k,\ell} = \left\langle \xi_m \psi_\ell \psi_k \right\rangle , \qquad k,\ell = 1,\dots,N_{\boldsymbol{\xi}},\ m = 1,\dots,M. \tag{3.7e}$$

Note that the matrices A_0 and A_m can be viewed as the $(1,1)$-blocks of the Galerkin discretization of the associated deterministic mixed problem (1.1) with a material parameter characterized by $T^{-1} = t_0$ and $T^{-1} = \sigma\sqrt{\lambda_m} t_m$, respectively. The matrix B is exactly the $(2,1)$-block of the deterministic problem, since the input random field does not occur in the bilinear form $b(\cdot,\cdot)$. The first term in $\hat{A}$ in (3.6) as well as the matrix $\hat{B}$ represent the discretization of the *mean problem*, i.e., the deterministic problem obtained by replacing the input random field T^{-1} with its expectation $\langle T^{-1} \rangle$.

An equivalent representation of (3.2) is obtained by permuting the blocks of unknowns $\boldsymbol{q}_\ell$ and $\boldsymbol{u}_\ell$ in (3.3) such that corresponding pairs $\boldsymbol{q}_\ell$ and $\boldsymbol{u}_\ell$ are adjacent. In this case, the coefficient matrix of (3.2) becomes

$$G_0 \otimes C_0 + \sum_{m=1}^{M} G_m \otimes C_m, \tag{3.8}$$

with matrices $C_0, \dots, C_M$ of dimension $N_{\boldsymbol{q}} + N_u$ given by

$$C_0 := \begin{bmatrix} A_0 & B^\top \\ B & 0 \end{bmatrix}, \qquad C_m := \begin{bmatrix} A_m & 0 \\ 0 & 0 \end{bmatrix}, \quad m = 1, \dots, M. \tag{3.9}$$

Here, C_0 is the saddle-point matrix associated with the mean problem, and C_m may be viewed as the contributions of the stochastic fluctuations. The structure of the Galerkin matrices G_0 and G_m will depend on the basis chosen for the space Ψ_p used to discretize the parameter space $L^2_\rho(\Gamma)$. We examine two such choices below.

3.2. Choice of stochastic basis. As discussed in section 2.2, the subspace $\Psi_p \subset L^2_\rho(\Gamma)$ consists of polynomials of degree p in the M variables $\xi_1, \dots, \xi_M$, and we distinguish

$$\Psi_p = \operatorname{span}\{\boldsymbol{\xi}^{\boldsymbol{\alpha}} : 0 \le \alpha_m \le p,\ m = 1, \dots, M\} \quad \text{(tensor product polynomials)}, \tag{3.10}$$

$$\Psi_p = \operatorname{span}\{\boldsymbol{\xi}^{\boldsymbol{\alpha}} : |\boldsymbol{\alpha}| \le p\} \quad \text{(complete polynomials)}, \tag{3.11}$$

where we have introduced the multi-index $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_M) \in \mathbb{N}_0^M$, the notation $\boldsymbol{\xi}^{\boldsymbol{\alpha}} = \xi_1^{\alpha_1} \xi_2^{\alpha_2} \cdots \xi_M^{\alpha_M}$ for the monomials in M variables, as well as $|\boldsymbol{\alpha}| = \alpha_1 + \cdots + \alpha_M$. In either case, a basis of the space Ψ_p can be constructed from products of univariate polynomials

$$\psi_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \psi^{(1)}_{\alpha_1}(\xi_1)\, \psi^{(2)}_{\alpha_2}(\xi_2) \cdots \psi^{(M)}_{\alpha_M}(\xi_M),$$

where each $\psi^{(m)}_j$, $0 \le j \le p$, is a fixed polynomial of exact degree j in the variable ξ_m, $1 \le m \le M$. Given two such polynomials $\psi_{\boldsymbol{\alpha}}$ and $\psi_{\boldsymbol{\beta}}$, since the joint density $\rho(\boldsymbol{\xi})$ in (2.10) separates, the integrations in $\langle \psi_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\beta}} \rangle$ with respect to each variable ξ_m are independent, and we obtain

$$\langle \psi_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\beta}} \rangle = \prod_{m=1}^{M} \left\langle \psi^{(m)}_{\alpha_m} \psi^{(m)}_{\beta_m} \right\rangle,$$

revealing that an orthonormal basis of Ψ_p is obtained by choosing each of the M sets of univariate polynomials $\{\psi^{(m)}_j\}_{j=0}^p$ to be the polynomials on the interval Γ_m that are orthonormal with respect to the weight function ρ_m. In this case, the matrix G_0, which is the Grammian matrix of the basis $\{\psi_{\boldsymbol{\alpha}}\}$ with respect to the inner product $(\psi_{\boldsymbol{\alpha}}, \psi_{\boldsymbol{\beta}})_{L^2_\rho(\Gamma)} := \langle \psi_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\beta}} \rangle$, is simply the identity matrix. For the space of tensor product polynomials, it is shown in [2, 3] that it is possible to construct a basis of Ψ_p whose elements, in addition to being orthonormal, also satisfy

$$\langle \xi_m \psi_{\boldsymbol{\alpha}} \psi_{\boldsymbol{\beta}} \rangle = \nu^{(m)}_{\alpha_m} \prod_{n=1}^{M} \delta_{\alpha_n, \beta_n}, \tag{3.12}$$

in which $\delta_{j,k}$ denotes the Kronecker delta. The explicit construction of this basis, sometimes referred to as *doubly orthogonal polynomials*, requires the solution of M dense generalized eigenvalue problems of size $p+1$. This calculation can be performed a priori, since it depends only on the distribution of the random variables chosen in (2.8). In [9] it was shown that this construction can also be done by solving M standard tridiagonal eigenvalue problems of size $p+1$. When a doubly orthogonal polynomial basis is used, (3.12) means that the matrices G_m are all diagonal. The tensor product form of the basis also means that each G_m takes the form

$$G_m = I \otimes \cdots \otimes I \otimes D^{(m)} \otimes I \otimes \cdots \otimes I, \qquad m = 1, \dots, M,$$

where I denotes the identity matrix of dimension $p+1$ and $D^{(m)}$ is the diagonal matrix $D^{(m)} = \mathrm{diag}(\nu_0^{(m)}, \dots, \nu_p^{(m)})$. In this case, determining the solution of the stochastic Galerkin problem (2.11), i.e., solving linear system (3.2), entails the solution of $N_{\boldsymbol{\xi}}$ deterministic saddle-point problems. More precisely, after permuting the $N_{\boldsymbol{\xi}}$ blocks of unknowns as in (3.8), the saddle-point problem with multi-index $\boldsymbol{\alpha}$ is given by

$$(3.13) \qquad \begin{bmatrix} A^{(\ell(\boldsymbol{\alpha}))} & B^\top \\ B & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{q}_{\ell(\boldsymbol{\alpha})} \\ \boldsymbol{u}_{\ell(\boldsymbol{\alpha})} \end{bmatrix} = \begin{bmatrix} \boldsymbol{g}_{\ell(\boldsymbol{\alpha})} \\ \boldsymbol{f}_{\ell(\boldsymbol{\alpha})} \end{bmatrix}, \qquad A^{(\ell(\boldsymbol{\alpha}))} := A_0 + \sum_{m=1}^{M} \nu_{\alpha_m}^{(m)} A_m,$$

where $\ell(\boldsymbol{\alpha})$ denotes the scalar index assigned to the multi-index $\boldsymbol{\alpha}$ in some enumeration of the $(p+1)^M$ multi-indices.

Remark 3.1. The saddle-point matrices in (3.13) are identical to the deterministic saddle-point matrices that arise when the random variables in diffusion coefficient (2.8) are sampled at the points $\{\xi_1^\ell, \dots \xi_M^\ell\} = \{\nu_{\alpha_1}^{(1)}, \dots, \nu_{\alpha_M}^{(M)}\}, \ell = 1 : (p+1)^M$ and a standard mixed finite element method is used to discretize each resulting variational problem of the form (2.1). Hence, the preconditioning strategy described below can be applied in a straightforward way to the saddle-point systems arising in traditional sampling methods.

4. Iterative solution.

In this section we address the solution of the linear system of equations (3.2) by preconditioned Krylov subspace iteration. As the coefficient matrix in (3.2) is symmetric and indefinite, a suitable Krylov subspace method is MINRES iteration [19], which minimizes the Euclidean norm of the residual at every step. Ideally, we would like the iterative solver to be robust with respect to the many parameters in the problem, i.e., h, p, M, σ, and t_0, and this necessitates an efficient preconditioning scheme. For simplicity, in the analysis presented below, we assume that the input random field T^{-1} is *homogeneous*, so that the mean $t_0(\boldsymbol{x})$ is constant, denoted μ, and so is the variance, denoted σ^2.

4.1. Mean-based preconditioning.

The preconditioning approach that we adopt is based on the mean value t_0 of the input random field T^{-1}. This leads to practical computations. If the fluctuations represented by the terms $\sigma\sqrt{\lambda_m} t_m(\boldsymbol{x})\xi_m$ in (2.8) are small relative to t_0, then it is to be expected that an efficient preconditioner for the mean problem obtained for zero fluctuations will be effective also for the full stochastic problem, and we refer to such a preconditioner as a *mean-based preconditioner*. Note that, when using an orthonormal stochastic basis, the coefficient matrix associated with the mean problem can be written as $I \otimes C_0$, with C_0 defined in (3.9). This is the first term in the sum of Kronecker products in (3.8) that represents the global matrix. Basing a preconditioner for a stochastic Galerkin matrix on the first term in its Kronecker product sum has been studied for primal formulation (1.2) in [13, 20, 21]. In that case, mean-based preconditioning is exactly block-Jacobi preconditioning. Here, we shall derive a preconditioning scheme for the stochastic Galerkin matrix based on a block-diagonal preconditioner for the deterministic saddle-point system C_0.

4.2. Deterministic preconditioner.

In this section we summarize earlier work [22] on preconditioning saddle-point problems with coefficient matrices

$$(4.1) \qquad C := \begin{bmatrix} A & B^\top \\ B & 0 \end{bmatrix},$$

where $A \in \mathbb{R}^{n\times n}$ is symmetric positive definite and $B \in \mathbb{R}^{m\times n}$, $n \geq m$, has full rank. Our approach uses block-diagonal preconditioners of the form

$$P = \begin{bmatrix} D & 0 \\ 0 & V \end{bmatrix}, \tag{4.2}$$

where D is a diagonal matrix with positive entries approximating A and V is a symmetric positive definite approximation to the matrix $S_D := BD^{-1}B^\top$, which may be viewed as a sparse approximation of the Schur complement $S := BA^{-1}B^\top$. As both blocks of P are symmetric positive definite matrices, it is possible to use P as a preconditioner for MINRES iteration. Bounds for the linear convergence rate of preconditioned MINRES may be obtained from inclusion intervals for the negative and positive components of the spectrum of the preconditioned matrix $P^{-1}C$ or equivalently, that of the symmetric matrix $P^{-1/2}CP^{-1/2}$.

THEOREM 4.1. *Let $\alpha_{\min}$ and $\alpha_{\max}$ denote the extremal eigenvalues of $D^{-1}A$, and let θ and Θ be two real constants such that*

$$0 < \theta^2 \leq \frac{\boldsymbol{v}^\top S_D \boldsymbol{v}}{\boldsymbol{v}^\top V \boldsymbol{v}} \leq \Theta^2 \qquad \forall \boldsymbol{v} \in \mathbb{R}^m \setminus \{0\}.$$

Then the eigenvalues of the preconditioned matrix $P^{-1}C$, with P as in (4.2) *and C as in* (4.1) *lie in the union of the intervals*

$$\left[\frac{1}{2}\left(\alpha_{\min} - \sqrt{\alpha_{\min}^2 + 4\Theta^2}\right), \frac{1}{2}\left(\alpha_{\max} - \sqrt{\alpha_{\max}^2 + 4\theta^2}\right)\right] \\ \cup \left[\alpha_{\min}, \frac{1}{2}\left(\alpha_{\max} + \sqrt{\alpha_{\max}^2 + 4\Theta^2}\right)\right].$$

Proof. See [22, Corollaries 3.3 and 3.4]. □

Such preconditioners are known to be very effective when applied to discretizations of (1.1). In particular, in [22] we derive eigenvalue inclusion bounds which are independent of the spatial mesh size h and robust with respect to the coefficient function T^{-1} when D is chosen as the diagonal of A and the action of V^{-1} on a vector is defined as one V-cycle of algebraic multigrid (AMG) applied to a linear system with coefficient matrix S_D. Using lowest-order Raviart–Thomas mixed approximation, the constants θ^2 and Θ^2 are bounded independently of h because the matrix S_D is equivalent to a finite difference approximation of the operator $\nabla \cdot (T\nabla)$ and, crucially, is an M-matrix.

4.3. Preconditioning the stochastic Galerkin system. In the following, we shall construct a preconditioner to the stochastic Galerkin equations (3.2) based on the deterministic preconditioner in section 4.2. Specifically, we set $D := D_0 := \mathrm{diag}(A_0)$, with A_0 the so-called mean mass matrix defined in (3.7), and, as in [22], we define V such that its inverse is effected by the action of a single AMG V-cycle applied to the sparse matrix $S_{D_0} = BD_0^{-1}B^\top$, with B defined in (3.7). Thus, writing the coupled system matrix from (3.2) in the form

$$\hat{C} := \begin{bmatrix} \hat{A} & \hat{B}^\top \\ \hat{B} & 0 \end{bmatrix} = \begin{bmatrix} G_0 \otimes A_0 + \sum_{m=1}^M G_m \otimes A_m, & G_0 \otimes B^\top \\ G_0 \otimes B & 0 \end{bmatrix}$$

and noting that $G_0 = I$ when using an orthonormal stochastic basis, our preconditioner is of the form

$$\hat{P} := \begin{bmatrix} \hat{D}_0 & 0 \\ 0 & \hat{V} \end{bmatrix} := \begin{bmatrix} I \otimes D_0 & 0 \\ 0 & I \otimes V \end{bmatrix}. \tag{4.3}$$

Our choice of V ensures the existence of spectral equivalence bounds θ and Θ independent of h such that

$$0 < \theta^2 \leq \frac{\boldsymbol{v}^\top B D_0^{-1} B^\top \boldsymbol{v}}{\boldsymbol{v}^\top V \boldsymbol{v}} \leq \Theta^2 \qquad \forall \boldsymbol{v} \in \mathbb{R}^{N_u} \setminus \{0\}. \tag{4.4}$$

Using elementary properties of Kronecker products, we deduce that

$$\frac{\hat{\boldsymbol{v}}^\top \hat{B}\hat{D}_0^{-1}\hat{B}^\top \hat{\boldsymbol{v}}}{\hat{\boldsymbol{v}}^\top \hat{V} \hat{\boldsymbol{v}}} = \frac{\hat{\boldsymbol{v}}^\top (I \otimes B)(I \otimes D_0)^{-1}(I \otimes B)^\top \hat{\boldsymbol{v}}}{\hat{\boldsymbol{v}}(I \otimes V)\hat{\boldsymbol{v}}} = \frac{\hat{\boldsymbol{v}}^\top (I \otimes B D_0^{-1} B)^\top \hat{\boldsymbol{v}}}{\hat{\boldsymbol{v}}(I \otimes V)\hat{\boldsymbol{v}}}$$

for all nonzero $\hat{\boldsymbol{v}} \in \mathbb{R}^{N_u N_{\boldsymbol{\xi}}}$, showing that spectral equivalence bounds (4.4) also hold for $\hat{B}\hat{D}_0^{-1}\hat{B}^\top$ and $\hat{V}$. Applying Theorem 4.1 now immediately yields spectral inclusion bounds for stochastic Galerkin system (3.2) preconditioned by (4.3).

Corollary 4.2. *Let $\hat{\alpha}_{\min}$ and $\hat{\alpha}_{\max}$ denote the extremal eigenvalues of $\hat{D}_0^{-1}\hat{A}$, and let θ and Θ be the constants in (4.4). Then the eigenvalues of the preconditioned matrix $\hat{P}^{-1}\hat{C}$ lie in the union of the intervals*

$$\left[\frac{1}{2}\left(\hat{\alpha}_{\min} - \sqrt{\hat{\alpha}_{\min}^2 + 4\Theta^2}\right), \frac{1}{2}\left(\hat{\alpha}_{\max} - \sqrt{\hat{\alpha}_{\max}^2 + 4\theta^2}\right)\right]$$
$$\cup \left[\hat{\alpha}_{\min}, \frac{1}{2}\left(\hat{\alpha}_{\max} + \sqrt{\hat{\alpha}_{\max}^2 + 4\Theta^2}\right)\right].$$

The limits of the spectral inclusions in Corollary 4.2 are solely determined by the eigenvalues of

$$\hat{D}_0^{-1}\hat{A} = I \otimes D_0^{-1}A_0 + \sum_{m=1}^{M} G_m \otimes D_0^{-1}A_m. \tag{4.5}$$

We bound the eigenvalues of the Kronecker product factors separately in the following two lemmas.

Lemma 4.3. *Assume that square or right-angled triangular lowest-order Raviart–Thomas mixed approximation is used for spatial discretization, and let $D_0 = \mathrm{diag}(A_0)$ and $\{A_m\}_{m=1}^M$ be as defined in (3.7). If the individual eigenfunction t_m in (2.8) is not strictly positive, then*

$$-\frac{3\sigma}{2\mu}\sqrt{\lambda_m}\,\|t_m\|_{L^\infty(D)} \leq \frac{\boldsymbol{r}^\top A_m \boldsymbol{r}}{\boldsymbol{r}^\top D_0 \boldsymbol{r}} \leq \frac{3\sigma}{2\mu}\sqrt{\lambda_m}\,\|t_m\|_{L^\infty(D)} \quad \forall \boldsymbol{r} \in \mathbb{R}^{N_q} \setminus \{0\},$$

where

$$t_m^{\min} := \inf_{\boldsymbol{x} \in D} t_m(\boldsymbol{x}), \qquad \text{and} \qquad t_m^{\max} := \|t_m\|_{L^\infty(D)}. \tag{4.6}$$

Alternatively, if t_m is uniformly positive, then

$$0 < \frac{\sigma}{2\mu}\sqrt{\lambda_m}\,t_m^{\min} \leq \frac{\boldsymbol{r}^\top A_m \boldsymbol{r}}{\boldsymbol{r}^\top D_0 \boldsymbol{r}} \leq \frac{3\sigma}{2\mu}\sqrt{\lambda_m}\,t_m^{\max} \qquad \forall\, \boldsymbol{r} \in \mathbb{R}^{N_q} \setminus \{0\}.$$

Proof. Given any $\boldsymbol{q} \in \mathbb{R}^{N_q} \setminus \{0\}$, we may define $\boldsymbol{r} \in \Phi_h$ by $\boldsymbol{r}(\boldsymbol{x}) = \sum q_i \boldsymbol{\varphi}_i(\boldsymbol{x})$. If $t_m(\boldsymbol{x}) \geq 0$ on D, then

$$\boldsymbol{q}^\top A_m \boldsymbol{q} = \sigma\sqrt{\lambda_m}\int_D t_m(\boldsymbol{x})\boldsymbol{r}\cdot\boldsymbol{r}\,d\boldsymbol{x} \leq \frac{\sigma}{\mu}t_m^{\max}\sqrt{\lambda_m}\int_D \mu\,\boldsymbol{r}\cdot\boldsymbol{r}\,d\boldsymbol{x} = \frac{\sigma}{\mu}t_m^{\max}\sqrt{\lambda_m}\;\boldsymbol{q}^\top A_0 \boldsymbol{q},$$

$$\boldsymbol{q}^\top A_m \boldsymbol{q} = \sigma\sqrt{\lambda_m}\int_D t_m(\boldsymbol{x})\boldsymbol{r}\cdot\boldsymbol{r}\,d\boldsymbol{x} \geq \frac{\sigma}{\mu}t_m^{\min}\sqrt{\lambda_m}\int_D \mu\,\boldsymbol{r}\cdot\boldsymbol{r}\,d\boldsymbol{x} = \frac{\sigma}{\mu}t_m^{\min}\sqrt{\lambda_m}\;\boldsymbol{q}^\top A_0 \boldsymbol{q},$$

where $t_m^{\min}$ and $t_m^{\max}$ are as defined in (4.6). Dividing through by $\boldsymbol{q}^\top A_0 \boldsymbol{q} > 0$ gives

$$0 < \frac{\sigma}{\mu} t_m^{\min} \sqrt{\lambda_m} \leq \frac{\boldsymbol{q}^\top A_m \boldsymbol{q}}{\boldsymbol{q}^\top A_0 \boldsymbol{q}} \leq \frac{\sigma}{\mu} t_m^{\max} \sqrt{\lambda_m} \qquad \forall \boldsymbol{q} \in \mathbb{R}^{N_q} \setminus \{0\}. \tag{4.7}$$

If t_m also takes on negative values in D, we have

$$\left| \boldsymbol{q}^\top A_m \boldsymbol{q} \right| = \left| \sigma \sqrt{\lambda_m} \int_D \frac{t_m(\boldsymbol{x})}{\mu} \mu \, \boldsymbol{r} \cdot \boldsymbol{r} \, d\boldsymbol{x} \right| \leq \frac{\sigma}{\mu} \|t_m\|_{L^\infty(D)} \sqrt{\lambda_m} \; \boldsymbol{q}^\top A_0 \boldsymbol{q}, \tag{4.8}$$

leading to

$$-\frac{\sigma}{\mu} \sqrt{\lambda_m} \|t_m\|_{L^\infty(D)} \;\leq\; \frac{\boldsymbol{q}^\top A_m \boldsymbol{q}}{\boldsymbol{q}^\top A_0 \boldsymbol{q}} \;\leq\; \frac{\sigma}{\mu} \sqrt{\lambda_m} \|t_m\|_{L^\infty(D)} \quad \forall\, \boldsymbol{q} \in \mathbb{R}^{N_q} \setminus \{0\}. \tag{4.9}$$

Now, let $A_0^\square$ denote the element matrix associated with the mean matrix A_0. Using uniform square elements as a specific example, we have

$$A_0^\square = \frac{h^2}{6} \begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}, \qquad \left(\operatorname{diag}(A_0^\square \right)^{-1} A_0^\square = \frac{1}{2} \begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}, \tag{4.10}$$

and so

$$\frac{1}{2} \leq \frac{\boldsymbol{q}^\top A_0^\square \boldsymbol{q}}{\boldsymbol{q}^\top \operatorname{diag}\left(A_0^\square\right) \boldsymbol{q}} \leq \frac{3}{2} \qquad \forall \boldsymbol{q} \in \mathbb{R}^4 \setminus \{0\} \text{ and all elements } \square. \tag{4.11}$$

Using a standard result from [29], we thus arrive at

$$\frac{1}{2} \leq \frac{\boldsymbol{q}^\top A_0 \boldsymbol{q}}{\boldsymbol{q}^\top \operatorname{diag}(A_0) \boldsymbol{q}} \leq \frac{3}{2} \quad \forall\, \boldsymbol{q} \in \mathbb{R}^{N_q} \setminus \{0\}. \tag{4.12}$$

Combining (4.12) with (4.7) and (4.9) gives the desired result. □

Remark 4.4. Lemma 4.3 is readily extended to cover Raviart–Thomas mixed approximation on general meshes. For equilateral triangles, the constants in (4.11) are $\frac{1}{3}$ and $\frac{2}{3}$.

The following result gives us a handle on the eigenvalues of the stochastic Galerkin matrices G_m appearing in (4.5).

LEMMA 4.5. *Assume that Ψ_p consists of either complete or tensor product multivariate polynomials of degree p. The eigenvalues of each of the G_m are zeros of the set of univariate polynomials of degree $p+1$ or less that are orthogonal with respect to the weight function ρ_m. In particular, if ρ_m has bounded support, the eigenvalues are uniformly bounded with respect to p.*

Proof. See Lemma 3.1 in [21] and [9]. □

Note that if the random variables ξ_m are Gaussian, then the support of the associated density function is unbounded and the extremal eigenvalues of G_m are bounded by the extremal roots of the univariate Hermite polynomial of degree $p+1$. These grow like $O(\sqrt{p})$ as $p \to \infty$.

Combining Lemma 4.3 with Lemma 4.5 gives us a bound on the eigenvalues of $\hat{D}_0^{-1} \hat{A}$ in (4.5).

LEMMA 4.6. *Assume that square or right-angled triangular lowest-order Raviart–Thomas mixed approximation is used for the spatial discretization. If the random* *variables in* (2.8) *range over a real interval symmetric about zero, then the eigenvalues of* $\hat{D}_0^{-1}\hat{A}$ *lie in the bounded interval*

$$\left[\frac{1}{2} - c_p\tau, \frac{3}{2} + c_p\tau\right], \qquad \textit{where} \quad \tau = \frac{3\sigma}{2\mu}\sum_{m=1}^{M}\sqrt{\lambda_m}\,\|t_m\|_{L^\infty(D)}, \tag{4.13}$$

where σ *and* μ *are the standard deviation and mean of the input random field* T^{-1}, $\{(\lambda_m, t_m)\}$ *are the eigenpairs of the correlation function, and* $c_p > 0$ *is a constant possibly depending on* p.

Proof. The $N_{\boldsymbol{\xi}} N_{\boldsymbol{q}}$ eigenvalues$\{\hat{\alpha}_j\}$ we are seeking satisfy

$$\left(I \otimes D_0^{-1}A_0 + \sum_{m=1}^{M} G_m \otimes D_0^{-1}A_m\right)\boldsymbol{q} = \hat{\alpha}\boldsymbol{q}.$$

Using (4.12) and elementary properties of the Kronecker product, notice first that

$$\frac{1}{2} \le \frac{\boldsymbol{q}^\top (I \otimes A_0)\,\boldsymbol{q}}{\boldsymbol{q}^\top (I \otimes D_0)\,\boldsymbol{q}} \le \frac{3}{2} \quad \forall\, \boldsymbol{q} \in \mathbb{R}^{N_{\boldsymbol{q}} N_{\boldsymbol{\xi}}} \setminus \{0\}\,. \tag{4.14}$$

If the random variables in (2.8) vary over a (bounded or unbounded) symmetric interval, then their densities must have a support symmetric to zero and therefore, by Lemma 4.5, the eigenvalues of G_m, $m = 1, \ldots, M$, belong to a symmetric interval $[-c_p, c_p]$. By Lemma 4.3, the eigenvalues of each matrix $D_0^{-1}A_m$ belong to

$$\left[\frac{\sigma}{2\mu}\sqrt{\lambda_m}t_m^{\min}, \frac{3\sigma}{2\mu}\sqrt{\lambda_m}t_m^{\max}\right] \quad \text{or} \quad \left[-\frac{3\sigma}{2\mu}\sqrt{\lambda_m}\|t_m\|_{L^\infty(D)}, \frac{3\sigma}{2\mu}\sqrt{\lambda_m}\|t_m\|_{L^\infty(D)}\right]$$

depending on the positivity of the eigenfunction t_m. Denoting the minimum and maximum eigenvalues of $G_m \otimes D_0^{-1}A_m$ by $\gamma_m^{\min}$ and $\gamma_m^{\max}$, we have, in both cases,

$$\hat{\alpha}_{\min} \ge \frac{1}{2} + \sum_{m=1}^{M}\gamma_m^{\min} \ge \frac{1}{2} - c_p\tau, \qquad \hat{\alpha}_{\max} \le \frac{3}{2} + \sum_{m=1}^{M}\gamma_m^{\max} \le \frac{3}{2} + c_p\tau,$$

where τ is as defined in (4.13). □

Corollary 4.2 and Lemma 4.6 tell us that the convergence of preconditioned MINRES is independent of h but is likely to deteriorate when the ratio $\sigma\mu^{-1}$ is increased. Convergence is independent of p if bounded random variables are used. However, if Gaussian random variables are used, then $\hat{A}$ and $\hat{D}_0^{-1}\hat{A}$ become indefinite as $p \to \infty$. Recall that the problem is not well-posed in that case. Finally, we note that the boundedness assumption on KL expansion (2.8) means that τ in (4.13) converges to a finite limit as $M \to \infty$.

The cost of applying the preconditioner $\hat{P}$ in each MINRES iteration amounts to one solve with a diagonal matrix of dimension $N_\xi N_{\boldsymbol{q}} \times N_\xi N_{\boldsymbol{q}}$ and N_ξ multigrid V-cycles, with the $N_u \times N_u$ matrix S_{D_0}, where $N_\xi = \binom{M+p}{p}$ or $N_{\boldsymbol{\xi}} = (p+1)^M$ depending on whether Ψ_p consists of complete polynomials or tensor product polynomials of degree p, respectively. Since the cost of performing one AMG V-cycle grows linearly in N_u, (unlike traditional factorization methods), we have a computationally optimal preconditioner. The set-up of the AMG preconditioner only has to be performed once on a deterministic matrix, so it is a relatively trivial component of the overall computational cost.

4.4. Preconditioning the decoupled system. The derivation of the eigenvalue inclusion intervals for the preconditioned stochastic Galerkin problem given in Lemma 4.6 assumes only that an orthonormal set of stochastic basis functions are used. In particular, the eigenvalue bounds also hold if Ψ_p is chosen as in (3.10). However, in that case, a doubly orthogonal basis, characterized by (3.12), exists for which the (suitably reordered) stochastic Galerkin system decouples into the $N_{\boldsymbol{\xi}}$ saddle-point problems (3.13). It is prefereable then to solve these systems separately, and we can derive somewhat sharper bounds in this case by applying the analysis of the preceding section to each of the decoupled systems in turn.

The simplest approach to take is to solve the uncoupled linear systems in serial. For computational efficiency, we would like to use the same preconditioner for each system so as to minimize the set-up cost, and the strategy we advocate here is use the mean-based preconditioner

$$P = \begin{bmatrix} D_0 & 0 \\ 0 & V \end{bmatrix}, \tag{4.15}$$

where, as in the previous section, $D_0 = \operatorname{diag}(A_0)$ and V^{-1} represents a V-cycle of AMG applied to the deterministic matrix $S = BD_0^{-1}B^\top$. If (4.4) holds and each $A^{(\ell)} := A^{(\ell(\boldsymbol{\alpha}))}$, $\ell = 1, \ldots, N_{\boldsymbol{\xi}}$, is positive definite, we have the following result analogous to Corollary 4.2 (see [22] and [26]).

LEMMA 4.7. *Let* $0 < \alpha_{\min}^{(\ell)} \leq \alpha_{\max}^{(\ell)}$ *denote the extremal eigenvalues of* $D_0^{-1}A^{(\ell)}$. *The eigenvalues of*

$$P^{-1}\begin{bmatrix} A^{(\ell)} & B^\top \\ B & 0 \end{bmatrix}, \qquad \ell = 1, \ldots, N_{\boldsymbol{\xi}} = (p+1)^M$$

lie in the union of the intervals

$$\left[\frac{1}{2}\left(\alpha_{\min}^{(\ell)} - \sqrt{{\alpha_{\min}^{(\ell)}}^2 + 4\Theta^2}\right)\frac{1}{2}\left(\alpha_{\max}^{(\ell)} - \sqrt{{\alpha_{\max}^{(\ell)}}^2 + 4\theta^2}\right)\right]$$
$$\cup \left[\alpha_{\min}^{(\ell)}, \frac{1}{2}\left(\alpha_{\max}^{(\ell)} + \sqrt{{\alpha_{\max}^{(\ell)}}^2 + 4\Theta^2}\right)\right],$$

where θ^2 *and* Θ^2 *are the constants appearing in* (4.4).

We demonstrate the tightness of the above bounds using the following example.

Example 4.1. Consider decoupled system (3.13) arising from SFEM discretization of (1.4) on $D = [0,1] \times [0,1]$, with $f = 0$, $\partial D_D = \{0,1\} \times [0,1]$, and $\partial D_N = \partial D \backslash \partial D_D$. We select covariance function (2.9a) with $\tau_1 = \tau_2 = 1$, use uniform random variables in (2.8), and use doubly orthogonal Legendre polynomials for the basis of Ψ_p. For the spatial discretization, we use square Raviart–Thomas elements with $h^{-1} = 16$. We set $p = 2$ and $M = 2$ so that we have $N_{\boldsymbol{\xi}} = 9$ decoupled systems.

In Table 4.1 we present data corresponding to the specific case of $\mu = 1$ and $\sigma = 0.1$ so that the signal/noise ratio is 10. For each of the nine uncoupled systems, we list the number of preconditioned MINRES iterations required to reach a specified tolerance, together with a comparison of the extremal eigenvalues with the bounds given in Lemma 4.7.

Table 4.2 gives the values $\alpha_{\min}^{(\ell)}$, $\alpha_{\max}^{(\ell)}$, θ^2, and Θ^2 that we used to compute the bounds in Table 4.1. Note that the values θ^2 and Θ^2 are independent of all the

Table 4.1
Computed and estimated extremal eigenvalues of preconditioned saddle-point matrices.

Iters	ℓ	Bounds	Computed eigenvalues
38	1	$[-0.8087, -0.4842] \cup [0.4278, 1.9542]$	$[-0.8068, -0.4993] \cup [0.4315, 1.8972]$
39	2	$[-0.7886, -0.4580] \cup [0.4796, 2.0632]$	$[-0.7876, -0.4716] \cup [0.4817, 2.0076]$
39	3	$[-0.7710, -0.4324] \cup [0.5260, 2.1828]$	$[-0.7694, -0.4460] \cup [0.5314, 2.1252]$
38	4	$[-0.8013, -0.4987] \cup [0.4467, 1.8985]$	$[-0.7999, -0.5039] \cup [0.4521, 1.8594]$
38	5	$[-0.7808, -0.4728] \cup [0.5000, 2.0000]$	$[-0.7806, -0.4755] \cup [0.5048, 1.9726]$
39	6	$[-0.7646, -0.4433] \cup [0.5432, 2.1305]$	$[-0.7624, -0.4492] \cup [0.5562, 2.0929]$
39	7	$[-0.8087, -0.4842] \cup [0.4278, 1.9542]$	$[-0.8070, -0.4987] \cup [0.4315, 1.8996]$
39	8	$[-0.7886, -0.4580] \cup [0.4796, 2.0632]$	$[-0.7879, -0.4711] \cup [0.4817, 2.0089]$
39	9	$[-0.7710, -0.4324] \cup [0.5260, 2.1828]$	$[-0.7697, -0.4455] \cup [0.5314, 2.1257]$

Table 4.2
Maximal eigenvalues illustrating the efficiency of diagonal scaling and multigrid.

ℓ	$\alpha_{\min}^{(\ell)}$	$\alpha_{\max}^{(\ell)}$	θ^2	Θ^2
1	0.4278	1.4425	0.9328	1.0000
2	0.4796	1.5785	–	–
3	0.5260	1.7247	–	–
4	0.4467	1.3717	–	–
5	0.5000	1.5000	–	–
6	0.5432	1.6612	–	–
7	0.4278	1.4425	–	–
8	0.4796	1.5785	–	–
9	0.5260	1.7247	–	–

statistical parameters since μ is constant, so the only factor influencing the iteration counts from system to system is the efficiency of the diagonal scaling. We can get a tight theoretical handle on this. Notice that in the above example, $\alpha_{\min}^{(\ell)}$ is always a perturbation from 0.5 and $\alpha_{\max}^{(\ell)}$ is a perturbation of 1.5. This is entirely predictable in view of (4.12). Before analyzing the dependence of $\alpha_{\min}^{(\ell)}$ and $\alpha_{\max}^{(\ell)}$ on the parameters M, p, h and μ, and σ, we first present a sufficient condition for all the matrices $A^{(\ell)}$ in (3.13) to be positive definite.

Lemma 4.8. *If the random variables in* (2.8) *are bounded on the interval* $[-\gamma, \gamma]$ *and if*

$$\frac{\mu}{\sigma} > \gamma \sum_{m=1}^{M} \sqrt{\lambda_m}\, \|t_m\|_{L^\infty(D)}, \tag{4.16}$$

then each matrix $A^{(\ell)}$ *occurring in the sequence of saddle-point systems* (3.13) *is positive definite.*

Proof. If the random variables in (2.8) each vary on the interval $[-\gamma, \gamma]$, then each coefficient $\nu_m^{(\ell)} := \nu_{\alpha_m}^{(m)}$ from (3.13) lies in this interval. Hence, for each ℓ, we have

$$\mu + \sigma \sum_{m=1}^{M} \nu_m^{(\ell)} \sqrt{\lambda_m} t_m(\boldsymbol{x}) \geq \mu - \sigma\gamma \sum_{m=1}^{M} \sqrt{\lambda_m} \|t_m\|_{L^\infty(D)} =: \kappa^{(\ell)} \qquad \forall \boldsymbol{x} \in D.$$

Now associating with $\boldsymbol{q} \in \mathbb{R}^{N_q} \setminus \{0\}$ the function $\boldsymbol{r} = \sum q_i \boldsymbol{\varphi}_i \in \Phi_h$, we obtain

$$\boldsymbol{q}^\top A^{(\ell)} \boldsymbol{q} = \left(\left(\mu + \sigma \sum_{m=1}^{M} \nu_m^{(\ell)} \sqrt{\lambda_m} t_m \right) \boldsymbol{r}, \boldsymbol{r} \right) \geq \kappa^{(\ell)} (\boldsymbol{r}, \boldsymbol{r}) \geq \kappa^{(\ell)} c_* \boldsymbol{q}^\top \boldsymbol{q}, \tag{4.17}$$

where $c_* > 0$ denotes the minimum eigenvalue of the mass matrix with unit coefficients, represented by the bilinear form $(\boldsymbol{r}, \boldsymbol{r})$. The result follows if $\kappa^{(\ell)} > 0$, which is assured if (4.16) holds. □

We now return to Lemma 4.7 and assess the efficiency of the diagonal scaling $D_0^{-1} A^{(\ell)}$.

Lemma 4.9. *Assume that square or right-angled triangular lowest-order Raviart–Thomas mixed approximation is used for spatial discretization and that piecewise constant approximation is used for the eigenfunctions t_m. If $A^{(\ell)}$ is positive definite, then the constants $\alpha_{\min}^{(\ell)}$ and $\alpha_{\max}^{(\ell)}$ occurring in the eigenvalue bounds in Lemma 4.7 satisfy*

$$\frac{1}{2}\left(1 + \frac{\sigma m^{(\ell)}}{\mu}\right) \le \alpha_{\min}^{(\ell)}, \qquad \alpha_{\max}^{(\ell)} \le \frac{3}{2}\left(1 + \frac{\sigma M^{(\ell)}}{\mu}\right), \tag{4.18}$$

where μ and σ are the mean and standard deviation of the field T^{-1} and

(4.19)

$$m^{(\ell)} := \inf_{\boldsymbol{x} \in D} \sum_{m=1}^{M} h_m^{(\ell)}(\boldsymbol{x}), \quad M^{(\ell)} := \sup_{\boldsymbol{x} \in D} \sum_{m=1}^{M} h_m^{(\ell)}(\boldsymbol{x}), \quad \textit{with } h_m^{(\ell)}(\boldsymbol{x}) := \nu_m^{(\ell)} \sqrt{\lambda_m} t_m(\boldsymbol{x}).$$

Proof. Each $A^{(\ell)}$ is a weighted mass matrix, hence it suffices to consider the diagonally scaled element matrices (see [29]). Let $A_0^{\square}$ and $A_\ell^{\square}$ denote the element mass matrices associated with A_0 and $A^{(\ell)}$, respectively. Using piecewise constant approximation for the eigenfunctions, we have

$$A_\ell^{\square} = \left(1 + \frac{\sigma}{\mu} \sum_{m=1}^{M} \nu_m^{(\ell)} \sqrt{\lambda_m} t_m^{\square}\right) A_0^{\square}, \tag{4.20}$$

where $t_m^{\square}$ is the value of the mth eigenfunction t_m in the element under consideration. Hence,

$$\frac{\boldsymbol{v}^\top A_\ell^{\square} \boldsymbol{q}}{\boldsymbol{q}^\top \operatorname{diag}\left(A_0^{\square}\right) \boldsymbol{q}} = \left(1 + \frac{\sigma}{\mu} \sum_{m=1}^{M} \nu_m^{(\ell)} \sqrt{\lambda_m} t_m^{\square}\right) \frac{\boldsymbol{q}^\top A_0^{\square} \boldsymbol{q}}{\boldsymbol{q}^\top \operatorname{diag}\left(A_0^{\square}\right) \boldsymbol{q}}. \tag{4.21}$$

Using square elements as a specific example, (4.11) holds and so using the standard result from [29],

$$\alpha_{\min}^{(\ell)} \ge \min_{\square} \frac{\boldsymbol{q}^\top A_\ell^{\square} \boldsymbol{q}}{\boldsymbol{q}^\top \operatorname{diag}\left(A_0^{\square}\right) \boldsymbol{q}} \ge \frac{1}{2} \min_{\square} \left(1 + \frac{\sigma}{\mu} \sum_{m=1}^{M} \nu_m^{(\ell)} \sqrt{\lambda_m} t_m^{\square}\right),$$
$$\alpha_{\max}^{(\ell)} \le \max_{\square} \frac{\boldsymbol{q}^\top A_\ell^{\square} \boldsymbol{q}}{\boldsymbol{q}^\top \operatorname{diag}\left(A_0^{\square}\right) \boldsymbol{q}} \le \frac{3}{2} \max_{\square} \left(1 + \frac{\sigma}{\mu} \sum_{m=1}^{M} \nu_m^{(\ell)} \sqrt{\lambda_m} t_m^{\square}\right),$$

and the result follows. □

Example 4.2. To illustrate the sharpness of the bounds (4.18), consider Example 4.1 but now with $\mu = 1$, $\sigma = 0.3$, so that the signal/noise ratio is smaller than previously and take $p = 3$ and $M = 4$.

In Table 4.3, we list the computed extremal eigenvalues of $D_0^{-1} A^{(\ell)}$ for the first few systems in (3.13) together with the bounds from Lemma 4.9. Results are presented for uniform random variables.

Table 4.3

Computed and estimated values of $\alpha_{\min}^{(\ell)}$ *and* $\alpha_{\max}^{(\ell)}$ *(Uniform random variables).*

iters	ℓ	$\alpha_{\min}^{(\ell)}$	$\alpha_{\max}^{(\ell)}$	$\frac{1}{2}\left(1+\frac{\sigma m^{(\ell)}}{\mu}\right)$	$\frac{3}{2}\left(1+\frac{\sigma M^{(\ell)}}{\mu}\right)$
50	1	0.1794	1.7029	0.1586	1.7624
44	2	0.2848	1.9524	0.2612	2.0064
45	3	0.4131	2.2936	0.3897	2.3452
47	4	0.5033	2.5637	0.4833	2.6197
45	5	0.1973	1.5571	0.1746	1.6104

We observe that our bounds are tight. If Gaussian random variables are used, however, $A^{(1)}$ and $A^{(5)}$ are not positive definite, and so the bounds are not valid.[1] Note also that since the ratio $\frac{\sigma}{\mu}$ is larger compared to that in Table 4.1, the MINRES iteration counts are slightly higher.

Lemmas 4.7 and 4.9 tell us that the convergence of preconditioned MINRES for *all* systems in (3.13) is independent of h but will deteriorate when the ratio $\frac{\sigma}{\mu}$ is large. Convergence, for the ℓth system, is ultimately determined by the constants $m^{(\ell)}$ and $M^{(\ell)}$ in (4.19), and these can vary a great deal from system to system. Lemma 4.5 tells us that the set of coefficients $\{\nu_1^{(\ell)}, \dots \nu_M^{(\ell)}\}$ that determine each $m^{(\ell)}$ and $M^{(\ell)}$ are, again, just roots of orthogonal polynomials in $\boldsymbol{\xi}$ (see [9]). These are bounded with respect to p if bounded random variables are used.

In the decoupled approach, the cost of applying P *to each system*, in each MINRES iteration, now amounts to one solve with a diagonal matrix of dimension $N_{\boldsymbol{q}} \times N_{\boldsymbol{q}}$ and one multigrid V-cycle with the $N_u \times N_u$ matrix S_{D_0}. However, there are $N_{\boldsymbol{\xi}}$ systems to solve where, here, $N_{\boldsymbol{\xi}} = (p+1)^M$. Again, set-up is a trivial cost, and we have a computationally optimal preconditioner for each system. However, $N_{\boldsymbol{\xi}}$ grows unacceptably large for increasing M and p. For example, if $M = 6$ and $p = 4$, the dimension of the complete polynomial space is 210 compared to $N_{\boldsymbol{\xi}} = 15{,}625$ for the tensor product polynomial space. Comparing the results in Lemma 4.6 and Corollary 4.2 with those in Lemma 4.9 and Lemma 4.7, we see that the number of preconditioned MINRES iterations required to solve the large coupled system (when complete polynomials are employed) is likely to deteriorate at the same rate, with respect to p and the ratio $\frac{\sigma}{\mu}$, as the highest number of iterations required to solve any of the small decoupled systems associated with the tensor product polynomials. If the decoupled systems are simply solved in serial and if $(p+1)^M$ is significantly larger than $\binom{M+p}{p}$, then the coupled approach is almost certainly cheaper overall.

5. Numerical results.

We now present numerical results for two test problems, employing both the set of tensor product polynomials *and* the corresponding set of complete polynomials for the stochastic solution space Ψ_p. In the first case, we solve a sequence of $(p+1)^M$ decoupled saddle-point systems of dimension $(N_{\boldsymbol{q}} + N_u)$. In the second case, we solve a single large saddle-point system of dimension $\binom{M+p}{p}(N_{\boldsymbol{q}} + N_u)$. We employ uniform random variables for the stochastic input and construct the stochastic bases using Legendre polynomials.

To compare the methods, we record the number of MINRES iterations required to reduce the Euclidean norm of the preconditioned relative residual error to 10^{-8} when zero initial guesses are prescribed. In addition, we list set-up and solve times (in

[1] The discrete problem is not well-posed in this case (since (1.3) does not hold), and so the solution is meaningless.

TABLE 5.1
Example 5.1: *Numerical results obtained with tensor product polynomials.*

	p	1	2	3	4
	N_ξ	64	729	4,096	15,625
	$N_\xi(N_{\boldsymbol{q}} + N_u)$	786,432	8,957,952	50,331,648	192,000,000
$\frac{\sigma}{\mu} = 0.1$	avg. iters.	40	40	40	40
	max. iters.	42	43	43	44
	N_V	2,568	29,317	165,223	630,825
	set-up time	0.56s	0.46s	0.46s	0.50s
	avg. solve time	0.40s	0.41s	0.46s	0.47s
$\frac{\sigma}{\mu} = 0.2$	avg. iters.	43	43	43	43
	max. iters.	46	49	50	51
	N_V	2,723	31,287	177,484	678,832
	set-up time	0.57s	0.49s	0.54s	0.46s
	avg. solve time	0.45s	0.51s	0.52s	0.51s
$\frac{\sigma}{\mu} = 0.3$	avg. iters.	46	47	47	48
	max. iters.	50	57	63	66
	N_V	2,954	34,028	193,785	743,509
	set-up time	0.61s	0.54s	0.55s	0.48s
	avg. solve time	0.49s	0.52s	0.56s	0.57s

seconds) and the total number N_V of black-box AMG V-cycles performed on a system with the coefficient matrix S_{D_0}. The AMG code we use is a MATLAB version of the code HSL_MI20 [4]. N_V, the total number of diffusion solves, is the basic work unit and can be used to compare the costs of the two approaches. All reported experiments were performed in serial on a modest single processor Linux machine with 2 GB RAM and on a more powerful two-processor dual-core Linux machine with 16 GB RAM. The timings reported below were obtained using the second machine.

Remark 5.1. The dominant components of our Krylov subspace solver methodology are the matrix multiply of the coefficient matrix and the action of the inverse of the preconditioner which must be done once per iteration. Using either tensor product or complete polynomials, both these components are completely straightforward to parallelize over the number of stochastic degrees of freedom.

Example 5.1. Consider system (3.2) arising from the SFEM discretization of (1.4) with $D = [0,1] \times [0,1]$, $f = 0$, $\partial D_D = \{0,1\} \times [0,1]$, and $\partial D_N = \partial D \backslash \partial D_D$. We set $\boldsymbol{n} \cdot \boldsymbol{q} = 0$ on the horizontal boundaries, $u = 1$ on $\{0\} \times [0,1]$, and $u = 0$ on $\{1\} \times [0,1]$. We select covariance function (2.9c) (which is discussed in [31]) with $\tau = 1$ and the constant mean $\langle T^{-1} \rangle = \mu = 1$. With this choice of τ, $M = 6$ random variables are required to capture 98% of the variance of the input random field. For the spatial discretization, we select square elements on a uniform grid with $h^{-1} = 64$, yielding $N_{\boldsymbol{q}} = 8{,}192$ and $N_u = 4{,}096$.

Iteration counts and timings obtained with varying p and $\frac{\sigma}{\mu}$ are listed in Tables 5.1 and 5.2. As expected, the iteration counts deteriorate for increasing $\frac{\sigma}{\mu}$. The means and variances of $u_{h,p}$ and those of the x and y components of $\boldsymbol{q}_{h,p}$, for the particular case $p = 3$ and $\frac{\sigma}{\mu} = 0.2$, are plotted in Figures 5.1 and 5.2. In the decoupled case, we solve 4,096 saddle-point systems of dimension 12,288. Each system requires on average 43 preconditioned MINRES iterations, corresponding to a total of 177,484 multigrid V-cycles. In the coupled case, we solve one saddle-point system of dimension 1,032,192. A total of 59 preconditioned MINRES iterations are required, corresponding to only 4,956 multigrid V-cycles. Thus, solving (2.11) using (3.11) rather than (3.10) requires approximately one thirty-sixth of the number of fast diffusion solves. Although tensor product space (3.10) is richer than the space of complete polynomials

Table 5.2
Example 5.1*: Numerical results obtained with complete polynomials.*

	p	1	2	3	4	5
	N_ξ	7	28	84	210	462
	$N_\xi(N_q+N_u)$	86,016	344,064	1,032,192	2,580,480	5,677,056
$\frac{\sigma}{\mu}=0.1$	total iters.	43	45	46	48	48
	N_V	301	1,260	3,864	10,080	22,176
	set-up time	0.56s	0.47s	0.47s	0.47s	0.47s
	total solve time	3.27s	14.0s	45.35s	119.01s	262.04s
$\frac{\sigma}{\mu}=0.2$	total iters.	49	55	59	62	63
	N_V	343	1,540	4,956	13,020	29,106
	set-up time	0.54s	0.47s	0.49s	0.47s	0.47s
	total solve time	3.79s	17.18s	58.51s	154.82s	379.01s
$\frac{\sigma}{\mu}=0.3$	total iters.	55	66	74	80	86
	N_V	385	1,848	6,216	16,800	39,732
	set-up time	0.47s	0.48s	0.48s	0.47s	0.48s
	total solve time	4.08s	20.66s	72.97s	199.75s	486.74s

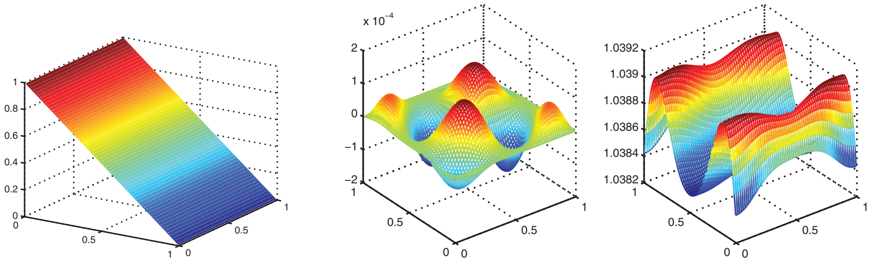

Fig. 5.1. *Example* 5.1*: Computed means* $\langle u_{h,p}\rangle$*,* $\left\langle \boldsymbol{q}^y_{h,p}\right\rangle$*, and* $\left\langle \boldsymbol{q}^x_{h,p}\right\rangle$ *(left to right) for* $p=3$*,* $\frac{\sigma}{\mu}=0.2$.

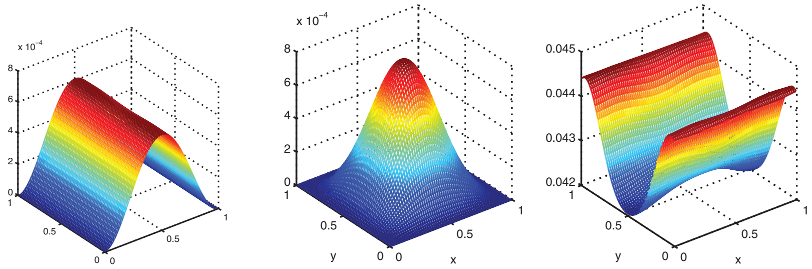

Fig. 5.2. *Example* 5.1*: Computed variances* $Var(u_{h,p})$*,* $Var(\boldsymbol{q}^y_{h,p})$*, and* $Var(\boldsymbol{q}^x_{h,p})$ *(left to right) for* $p=3$*,* $\frac{\sigma}{\mu}=0.2$.

(3.11), the solutions obtained in each case are observed to be qualitatively the same. Using (3.10), the maximum recorded values of the variances of $u_{h,p}, q^x_{h,p}$, and $q^y_{h,p}$ are $6.1332\times10^{-4}, 0.0445$, and 7.9302×10^{-4}, respectively. Using (3.11), we obtain the corresponding values $6.1329\times10^{-4}, 0.0444$, and 7.9113×10^{-4}.

Example 5.2. Consider the same test problem as above but now with a piecewise constant mean. Let $D_1=[0,0.5]\times[0,0.5]$, $D_2=[0.5,1]\times[0,0.5]$, $D_3=[0.5,1]\times[0.5,1]$, and $D_4=[0,0.5]\times[0.5,1]$ and set

$$\langle T^{-1}\rangle=\mu=\begin{cases}1 & \text{in } D_1 \text{ and } D_4,\\ 10^3 & \text{in } D_2 \text{ and } D_3.\end{cases}$$

Table 5.3
Example 5.2*: Numerical results obtained with tensor product polynomials.*

p	1	2	3	4
N_ξ	64	729	4,096	15,625
$N_\xi(N_{\boldsymbol{q}} + N_u)$	583,872	6,650,667	37,367,808	142,546,875
avg. iters.	43	44	44	44
max. iters.	46	49	51	53
N_V	2,773	31,824	179,974	688,085

Table 5.4
Example 5.2*: Numerical results obtained with complete polynomials.*

p	1	2	3	4
N_ξ	7	28	84	210
$N_\xi(N_{\boldsymbol{q}} + N_u)$	63,861	255,444	766,332	1,915,830
total iters.	48	55	58	62
N_V	336	1,540	4,872	13,020

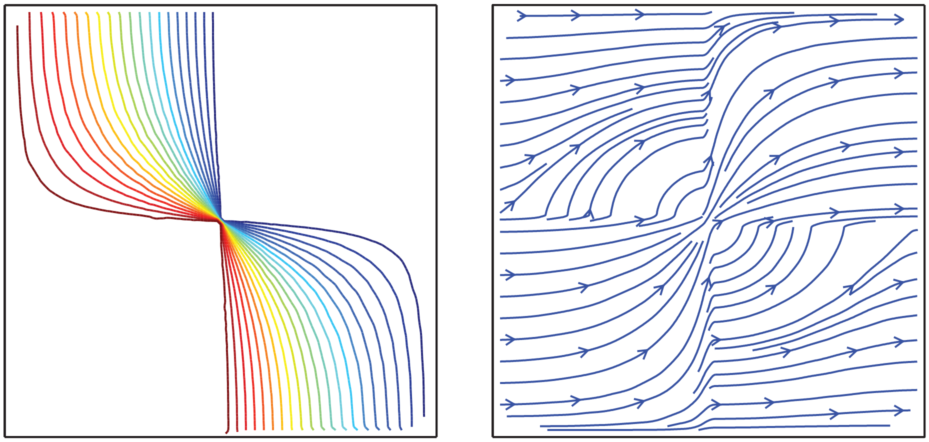

Fig. 5.3. *Example* 5.2*: Contours of* $\langle u_{h,p}\rangle$ *(left) and streamlines of* $\langle \boldsymbol{q}_{h,p}\rangle$ *(right),* $p = 3$.

Note that T takes small values in D_2 and D_3 so most "flow" occurs in D_1 and D_4. We choose correlation function (2.9c) with $\tau = 1$. In addition, we fix $\sigma = 0.2$ and vary p. For the spatial discretization, we select a locally adapted mesh of triangular elements yielding $N_{\boldsymbol{q}} = 5{,}474$ and $N_u = 3{,}649$.

minres iteration counts obtained for the decoupled and coupled systems, with varying p are listed in Tables 5.3 and 5.4, respectively. The contours of the expected head $\langle u_{h,p}\rangle$ and the streamlines of the expected flow-field $\langle \boldsymbol{q}_{h,p}\rangle$ for the case $p = 3$ are plotted in Figure 5.3. Again, the solutions obtained are qualitatively the same using either (3.10) or (3.11) for the stochastic solution space with a fixed p. When $p = 3$, choosing Ψ_p as in (3.10) requires the solution of 4,096 saddle-point systems, and our preconditioning strategy requires a total of 179,974 fast diffusion solves. In contrast, using complete polynomials to construct Ψ_3 requires the solution of one saddle-point system of dimension 766,332, and our mean-based solver requires only 4,872 multigrid V-cycles with the deterministic matrix S_{D_0}.

6. Conclusions. In this study we have developed a *mean-based* preconditioner for linear algebra systems that arise from a stochastic Galerkin mixed formulation of the steady-state diffusion equation with random data. If stochastic Galerkin methods are to be competitive with traditional deterministic methodologies based on sampling techniques, then we need fast and robust linear algebra techniques to solve the large indefinite systems that arise. Our approach uses a black-box algebraic multigrid on the spatial component of the problem, and we have demonstrated that this gives an effective way of solving the extremely large coupled and decoupled systems that arise when the fluctuations in the data are not too large relative to their mean value. We intend to extend our methodology to cover stochastically nonlinear formulations of diffusion problems in future publications.

REFERENCES

[1] I. Babuška, F. Nobile, and R. Tempone, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM J. Numer. Anal., 45 (2007), pp. 1005–1034.

[2] I. Babuška, R. Tempone, and G. E. Zouraris, *Galerkin finite element approximations of stochastic elliptic partial differential equations*, SIAM J. Numer. Anal., 42 (2004), pp. 800–825.

[3] I. Babuška, R. Tempone, and 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), pp. 1251–1294.

[4] J. Boyle, M. Mihajlović, and J. Scott, *HSL_MI*20: *An Efficient AMG Preconditioner*, Technical report RAL–TR–2007–021, STFC Rutherford Appleton Laboratory, Didcot, UK, 2007, http://epubs.cclrc.ac.uk/bitstream/1961/bmsRALTR2007021.pdf.

[5] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.

[6] G. Dagan, *Stochastic modeling of groundwater flow by unconditional and conditional probabilities* 1. *Conditional simulation and the direct problem*, Water Resources Res., 18 (1982), pp. 813–833.

[7] M. K. Deb, I. M. Babuška, and J. T. Oden, *Solution of stochastic partial differential equations using Galerkin finite element techniques*, Comput. Methods Appl. Mech. Engrg., 190 (2001), pp. 6359–6372.

[8] M. Eiermann, O. G. Ernst, and E. Ullmann, *Computational aspects of the stochastic finite element method*, Comput. Vis. Sci., 10 (2007), pp. 3–15.

[9] O. G. Ernst and E. Ullmann, *On Stochastic Galerkin Matrices*, Fakultät für Mathematik und Informatik, Technische Universität Bergakademie Freiberg, Germany, 2008–03, preprint.

[10] R. E. Ewing and M. F. Wheeler, *Computational aspects of mixed finite element methods*, in Numerical Methods for Scientific Computing, R. Stepleman, ed., North-Holland, Amsterdam, 1983, pp. 163–172.

[11] P. Frauenfelder, C. Schwab, and R. A. Todor, *Finite elements for elliptic problems with stochastic coefficients*, Comput. Methods Appl. Mech. Engrg., 194 (2005), pp. 205–228.

[12] R. Ghanem and P. D. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, New York, 1991.

[13] R. G. Ghanem and R. M. Kruger, *Numerical solution of spectral stochastic finite element systems*, Comput. Methods Appl. Mech. Engrg., 129 (1996), pp. 289–303.

[14] C. Jin, X.-C. Cai, and C. Li, *Parallel domain decomposition methods for stochastic elliptic equations*, SIAM J. Sci. Comput., 29 (2007), pp. 2096–2114.

[15] M. Loève, *Probability Theory*, Vol. II, Springer-Verlag, New York, 1977.

[16] L. Mathelin and O. P. L. Maître, *Dual-based a posteriori error estimate for stochastic finite element methods*, Comm. App. Math. Comput. Sci., 2 (2007), pp. 83–115.

[17] H. G. Matthies and C. Bucher, *Finite elements for stochastic media problems*, Comput. Methods Appl. Mech. Engrg., 168 (1999), pp. 3–17.

[18] H. G. Matthies and A. Keese, *Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations*, Comput. Methods Appl. Mech. Engrg., 194 (2005), pp. 1295–1331.

[19] C. C. Paige and M. A. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629.

[20] M. F. Pellissetti and R. G. Ghanem, *Iterative solution of systems of linear equations arising in the context of stochastic finite elements*, Adv. Engrg. Software, 31 (2000), pp. 607–616.

[21] C. E. Powell and H. C. Elman, *Block-diagonal preconditioning for spectral stochastic finite element systems*, IMA J. Numer. Anal., to appear, 2008; also available online from http://dx.doi.org/10.1093/imanum/drn014.

[22] C. E. Powell and D. Silvester, *Optimal preconditioning for Raviart-Thomas mixed formulation of second-order elliptic problems*, SIAM J. Matrix Anal. Appl., 25 (2003), pp. 718–738.

[23] P.-A. Raviart and J. M. Thomas, *A mixed finite element method for second order elliptic problems*, in Mathematical Aspects of the Finite Element Method, Lect. Notes Math. 606, I. Galligani and E. Magenes, eds., Springer-Verlag, New York, 1977, pp. 292–315.

[24] J. W. Ruge and K. Stüben, *Efficient solution of finite difference and finite element equations by algebraic multigrid (AMG)*, in Multigrid Methods for Integral and Differential Equations, IMA Monogr. Ser. 3, D. J. Paddon and H. Holstein, eds., Clarendon Press, Oxford, 1985, pp. 169–212.

[25] T. F. Russell and M. F. Wheeler, *Finite element and finite difference methods for continuous flows in porous media*, in The Mathematics of Reservoir Simulation, R. E. Ewing, ed., SIAM, Philadelphia, 1983, pp. 35–106.

[26] T. Rusten and R. Winther, *A preconditioned iterative method for saddlepoint problems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 887–904.

[27] C. Schwab and R.-A. Todor, *Karhunen-Loève approximation of random fields by generalized fast multipole methods*, J. Comput. Phys., 217 (2006), pp. 100–122.

[28] U. Trottenberg, C. Oosterlee, and A. Schüller, *Multigrid*, Academic Press, London, 2001.

[29] A. J. Wathen, *Realistic eigenvalue bounds for the Galerkin mass matrix*, IMA J. Numer. Anal., 7 (1987), pp. 449–457.

[30] D. Xiu and J. S. Hesthaven, *High-order collocation methods for differential equations with random inputs*, SIAM J. Sci. Comput., 27 (2005), pp. 1118–1139.

[31] D. Xiu and G. E. Karniadakis, *Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos*, Comput. Methods Appl. Mech. Engrg., 191 (2002), pp. 4927–4948.