

GALERKIN FINITE ELEMENT APPROXIMATIONS OF
STOCHASTIC ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS*IVO BABUŠKA[†], RAÚL TEMPONE[†], AND GEORGIOS E. ZOURARIS[‡]

Abstract. We describe and analyze two numerical methods for a linear elliptic problem with stochastic coefficients and homogeneous Dirichlet boundary conditions. Here the aim of the computations is to approximate statistical moments of the solution, and, in particular, we give a priori error estimates for the computation of the expected value of the solution. The first method generates independent identically distributed approximations of the solution by sampling the coefficients of the equation and using a standard Galerkin finite element variational formulation. The Monte Carlo method then uses these approximations to compute corresponding sample averages. The second method is based on a finite dimensional approximation of the stochastic coefficients, turning the original stochastic problem into a deterministic parametric elliptic problem. A Galerkin finite element method, of either the h - or p -version, then approximates the corresponding deterministic solution, yielding approximations of the desired statistics. We present a priori error estimates and include a comparison of the computational work required by each numerical approximation to achieve a given accuracy. This comparison suggests intuitive conditions for an optimal selection of the numerical approximation.

Key words. stochastic elliptic equation, perturbation estimates, Karhunen–Loëve expansion, finite elements, Monte Carlo method, $k \times h$ -version, $p \times h$ -version, expected value, error estimates

AMS subject classifications. 65N30, 65N15, 65C05, 65C20

DOI. 10.1137/S0036142902418680

1. Introduction. Due to the great development in computational resources and scientific computing techniques, more mathematical models can be solved efficiently. Ideally, this artillery could be used to solve many classical partial differential equations, the mathematical models we shall focus on here, to high accuracy. However, in many cases, the information available to solve a given problem is far from complete and is in general very limited. This is the case when solving a partial differential equation whose coefficients depend on material properties that are known to some accuracy. The same may occur with its boundary conditions and even with the geometry of its domain; see, for example, the works [5, 4]. Naturally, since the current engineering trends are going toward more reliance on computational predictions, the need for assessing the level of accuracy in the results grows accordingly. More than ever, the goal then becomes to represent and propagate uncertainties from the available data to the desired result through our partial differential equation. By uncertainty we mean either intrinsic variability of physical quantities or simply lack of knowledge about some physical behavior; cf. [38]. If variability is interpreted as randomness,

*Received by the editors November 26, 2002; accepted for publication (in revised form) September 26, 2003; published electronically June 4, 2004.

<http://www.siam.org/journals/sinum/42-2/41868.html>

[†]Institute for Computational and Engineering Sciences (ICES), University of Texas at Austin, Austin, TX 78759 (babuska@ices.utexas.edu, rtempone@ices.utexas.edu). The first author was partially supported by the National Science Foundation (grant DMS-9802367) and the Office of Naval Research (grant N00014-99-1-0724). The second author was partially supported by the Swedish Council for Engineering Science grant 222-148, National Science Foundation grant DMS-9802367, UdeLaR and UdeM in Uruguay, and the European network HYKE (contract HPRN-CT-2002-00282).

[‡]Department of Mathematics, University of the Aegean, GR-832 00 Karlovassi, Samos, Greece, and Institute of Applied and Computational Mathematics, FO.R.T.H., GR-711 10 Heraklion, Crete, Greece (zouraris@aegean.gr). This author was partially supported by the European network HYKE (contract HPRN-CT-2002-00282).

then naturally we can apply probability theory. To be fruitful, probability theory requires considerable empirical information about the random quantities in question, usually in the form of probability distributions or their statistical moments. On the other hand, if the only available information comes in the form of some bounds for the uncertain variables, the description and analysis of uncertainty may be based on other methods, such as convexity methods; cf. [8, 18]. This approach is closely related to the so-called worst case scenario.

This work addresses elliptic partial differential equations with stochastic coefficients, with applications to physical phenomena, e.g., random vibrations, seismic activity, oil reservoir management, and composite materials; see [2, 17, 19, 22, 27, 28, 30, 39, 43] and the references therein. Solving a stochastic partial differential equation entails finding the joint probability distribution of the solution, which is a hard problem. In practice we shall usually be satisfied with much less, namely, the computation of some moments, e.g., the expected value of the solution, or some probability related to a given event, e.g., the probability of some eventual failure; cf. [26, 34]. Although the results presented in this paper can be generalized to linear elliptic stochastic partial differential equations we now focus our study on the standard model problem, a second order linear elliptic equation with homogeneous Dirichlet boundary conditions.

Let D be a convex bounded polygonal domain in \mathbb{R}^d and (Ω, \mathcal{F}, P) be a complete probability space. Here Ω is the set of outcomes, $\mathcal{F} \subset 2^\Omega$ is the σ -algebra of events, and $P : \mathcal{F} \rightarrow [0, 1]$ is a probability measure. Consider the following stochastic linear elliptic boundary value problem: find a stochastic function, $u : \Omega \times \bar{D} \rightarrow \mathbb{R}$, such that P -a.e. in Ω , or, in other words, almost surely (a.s.), the following equation holds:

$$(1.1) \quad \begin{aligned} -\nabla \cdot (a(\omega, \cdot) \nabla u(\omega, \cdot)) &= f(\omega, \cdot) \quad \text{on } D, \\ u(\omega, \cdot) &= 0 \quad \text{on } \partial D. \end{aligned}$$

Here $a, f : \Omega \times D \rightarrow \mathbb{R}$ are stochastic functions with continuous and bounded covariance functions. If we denote by $B(D)$ the Borel σ -algebra generated by the open subsets of D , then a, f are assumed measurable with the σ -algebra $(\mathcal{F} \otimes B(D))$. In what follows we shall assume that a is bounded and uniformly coercive, i.e.,

$$(1.2) \quad \exists a_{\min}, a_{\max} \in (0, +\infty) : P(\omega \in \Omega : a(\omega, x) \in [a_{\min}, a_{\max}] \quad \forall x \in \bar{D}) = 1.$$

To ensure regularity of the solution u we assume also that a has a uniformly bounded and continuous first derivative; i.e., there exists a real deterministic constant C such that

$$(1.3) \quad P(\omega \in \Omega : a(\omega, \cdot) \in C^1(\bar{D}) \text{ and } \max_{\bar{D}} |\nabla_x a(\omega, \cdot)| < C) = 1$$

and that the right-hand side in (1.1) satisfies

$$(1.4) \quad \int_{\Omega} \int_D f^2(\omega, x) dx \, dP(\omega) < +\infty, \quad \text{which implies} \quad \int_D f^2(\omega, x) dx < +\infty \text{ a.s.}$$

Depending on the structure of the noise that drives an elliptic partial stochastic differential equation, there are different numerical approximations. For example, when the size of the noise is relatively small, a Neumann expansion around the mean value of the elliptic operator in (1.1) is a popular approach. It requires only the solution of standard deterministic partial differential equations, the number of them being equal to the number of terms in the expansion. Equivalently, a Taylor expansion of the solution around its mean value with respect to the noise yields the same result. Similarly, the work [30] uses formal Taylor expansions up to second order of the solution but does not study their convergence properties. Recently, the work [3] proposed a perturbation method with successive approximations. It also proves that uniform coercivity of the diffusion is sufficient for the convergence of the perturbation method.

When only the load f is stochastic, it is also possible to derive deterministic equations for the statistical moments of the solution. This case was analyzed in [1, 32] and more recently in the work [40], where a new method to solve these equations with optimal complexity is presented.

On the other hand, the works by Deb [14], Deb, Babuška, and Oden [15], Ghanem and Red-Horse [21], and Ghanem and Spanos [22] address the general case where all the coefficients are stochastic. Both approaches transform the original stochastic problem into a deterministic one with a large dimensional parameter, and they differ in the choice of the approximating functional spaces. The works [14, 15] use finite elements to approximate the noise dependence of the solution, while [21, 22] use a formal expansion in terms of Hermite polynomials. The approximation error in the approach [14, 15] can then be bounded in terms of deterministic quantities, as described in this work. After finishing this paper the authors became aware of the work [9], which developed a related error analysis for elliptic stochastic differential equations. The work [9] gives approximation error estimates for functionals of the solution, while our work focuses on error estimates for the strong approximation of the statistical moments of the solution. Besides, we use the Karhunen–Loève expansion and characterize the regularity of the solution, yielding, e.g., exponential rates of convergence; cf. section 6. On the other hand, the analysis in [9] uses the regularity of the computed functional together with estimates in negative spaces for the approximation error in the solution of the stochastic partial differential equation. This negative estimate can in principle accommodate rough solutions; however, they require H^2 spatial regularity, an assumption that is not clearly fulfilled by rough solutions.

Monte Carlo methods are both general and simple to code, and they are naturally suited for parallelization. They generate a set of independent identically distributed (iid) approximations of the solution by sampling the coefficients of the equation, using a spatial discretization of the partial differential equation, e.g., by a Galerkin finite element method. Then using these approximations we can compute corresponding sample averages of the desired statistics. Monte Carlo methods have a rate of convergence that may be considered slow, but its computational work grows only like a polynomial with respect to the number of random variables present in the problem. It is worth mentioning that in particular cases their convergence can be accelerated by variance reduction techniques; see [29]. The convergence rate of the Monte Carlo method is interpreted in the probability sense, and a practical estimate of its error needs an a posteriori estimate of the variance of the sampled random variable, which in turn requires an a priori bound on higher statistical moments; cf. the Berry-Essen theorem in [16]. Besides this, if the probability density of a random variable is smooth,

the convergence rate of the Monte Carlo method for the approximation of its expected value can be improved; cf. [35, 45]. Quasi-Monte Carlo methods (see [12, 41, 42]) offer a way to get a better convergence rate than the Monte Carlo method, although this advantage seems to deteriorate in general when the number of random variables present in the problem becomes large.

Another way to provide a notion of stochastic partial differential equations is based on the Wick product and the Wiener chaos expansion; see [27] and [46]. This approach yields solutions in Kondratiev spaces of stochastic distributions that are based on a different interpretation of (1.1); the solutions proposed in [27] and [46] are not the same as those arising from (2.1). The choice between (2.1) and [27] is a modeling decision, based on the physical situation under study. For example, with the Wick product we have $E[a \diamond u] = E[a]E[u]$, regardless of the correlation between a and u , whereas this is in general not true with the usual product. A numerical approximation for Wick stochastic linear elliptic partial differential equations is studied in [44], yielding a priori convergence rates.

This work studies the case of stochastic linear elliptic partial differential equations with random diffusion and load coefficients, stating and proving conditions for existence and uniqueness of solutions; for example, to obtain a meaningful numerical solution for (1.1) its diffusion coefficient should be uniformly coercive. Besides, it compares a Monte Carlo Galerkin method with the stochastic Galerkin finite element method introduced in [14] and introduces a related p -version, developing a priori error estimates for each case. A priori error estimates are useful to characterize the convergence, and ultimately they provide information to compare the number of operations required by numerical methods. The conclusion for now is that if the noise is described by a small number of random parameters or if the accuracy requirement is sufficiently strict, then a stochastic Galerkin method is preferred; otherwise, a Monte Carlo Galerkin method still seems to be the best choice; see section 8. It is worth mentioning that the development of numerical methods for stochastic differential equations is still very much ongoing, and better numerical methods are expected to appear.

2. Theoretical aspects of the continuous problem.

2.1. Notation and function spaces. Let d be a positive integer and D be an open, connected, bounded, and convex subset of \mathbb{R}^d with polygonal boundary ∂D . Denote the volume of D by $|D| \equiv \int_D 1 dx$. For a nonnegative integer s and $1 \leq p \leq +\infty$, let $W^{s,q}(D)$ be the Sobolev space of functions having generalized derivatives up to order s in the space $L^q(D)$. Using the standard multi-index notation, $\alpha = (\alpha_1, \dots, \alpha_d)$ is a d -tuple of nonnegative integers, and the length of α is given by $|\alpha| = \sum_{i=1}^d \alpha_i$. The standard Sobolev norm of $v \in W^{s,q}(D)$ will be denoted by $\|v\|_{W^{s,q}(D)}$, $1 \leq q \leq +\infty$. Whenever $q = 2$, we shall use the notation $H^s(D)$ instead of $W^{s,2}(D)$. As usual, the function space $H_0^1(D)$ is the subspace of $H^1(D)$ consisting of functions which vanish at the boundary of D in the sense of trace, equipped with the norm $\|v\|_{H_0^1(D)} = \{\int_D |\nabla v|^2 dx\}^{1/2}$. Whenever $s = 0$ we shall keep the notation with $L^q(D)$ instead of $W^{0,q}(D)$. For the sake of generality, sometimes we shall let H be a Hilbert space with inner product $(\cdot, \cdot)_H$. In that case we shall also denote the dual space of H , H' , that contains linear bounded functionals, $\mathcal{L} : H \rightarrow \mathbb{R}$, and is endowed with the operator norm $\|\mathcal{L}\|_{H'} = \sup_{0 \neq v \in H} \frac{\mathcal{L}(v)}{\|v\|_H}$.

Since stochastic functions intrinsically have different structure with respect to ω and with respect to x , the analysis of numerical approximations requires tensor

spaces. Let H_1, H_2 be Hilbert spaces. The tensor space $H_1 \otimes H_2$ is the completion of formal sums $u(y, x) = \sum_{i=1, \dots, n} v_i(y)w_i(x)$, $\{v_i\} \subset H_1$, $\{w_i\} \subset H_2$, with respect to the inner product $(u, \hat{u})_{H_1 \otimes H_2} = \sum_{i,j} (v_i, \hat{v}_j)_{H_1} (w_i, \hat{w}_j)_{H_2}$. For example, let us consider two domains, $y \in \Gamma$, $x \in D$, and the tensor space $L^2(\Gamma) \otimes H^1(D)$, with tensor inner product

$$(u, \hat{u})_{L^2(\Gamma) \otimes H^1(D)} = \int_{\Gamma} \left(\int_D u(y, x) \hat{u}(y, x) dx \right) dy + \int_{\Gamma} \left(\int_D \nabla_x u(y, x) \cdot \nabla_x \hat{u}(y, x) dx \right) dy.$$

Thus, if $u \in L^2(\Gamma) \otimes H^k(D)$, then $u(y, \cdot) \in H^k(D)$ a.e. on Γ and $u(\cdot, x) \in L^2(\Gamma)$ a.e. on D . Moreover, we have the isomorphism $L^2(\Gamma) \otimes H^k(D) \simeq L^2(\Gamma; H^k(D)) \simeq H^k(D; L^2(\Gamma))$ with the definitions

$$\begin{aligned} L^2(\Gamma; H^k(D)) \\ = \left\{ v : \Gamma \times D \rightarrow \mathbb{R} \mid v \text{ is strongly measurable and } \int_{\Gamma} \|v(y, \cdot)\|_{H^k(D)}^2 dy < +\infty \right\}, \end{aligned}$$

$$\begin{aligned} H^k(D; L^2(\Gamma)) \\ = \left\{ v : \Gamma \times D \rightarrow \mathbb{R} \mid v \text{ is strongly measurable and } \forall |\alpha| \leq k \exists \partial^{\alpha} v \in L^2(\Gamma) \otimes L^2(D) \right. \\ \left. \text{with } \int_{\Gamma} \int_D \partial^{\alpha} v(y, x) \varphi(y, x) dx dy = (-1)^{|\alpha|} \right. \\ \left. \int_{\Gamma} \int_D v(y, x) \partial^{\alpha} \varphi(y, x) dx dy \quad \forall \varphi \in C_0^{\infty}(\Gamma \times D) \right\}. \end{aligned}$$

Similar constructions can be done for the tensor product of Banach spaces, although the norm for the tensor space used to obtain the completion of the formal sums has to be defined explicitly on each case. Here the Banach space $C(\Gamma; H)$ comprises all continuous functions $u : \Gamma \rightarrow H$ with the norm $\|u\|_{C(\Gamma; H)} \equiv \sup_{y \in \Gamma} \|u(y)\|_H$. Similar definitions apply to the spaces $C^k(\Gamma; H)$, $k = 1, \dots$; cf. [20, p. 285].

Let Y be an \mathbb{R}^N -valued random variable in (Ω, \mathcal{F}, P) . If $Y \in L_P^1(\Omega)$, we denote its expected value by $E[Y] = \int_{\Omega} Y(\omega) dP(\omega) = \int_{\mathbb{R}^N} y d\mu_Y(y)$, where μ_Y is the distribution measure for Y , defined for the Borel sets $\tilde{b} \in B(\mathbb{R}^N)$ by $\mu_Y(\tilde{b}) \equiv P(Y^{-1}(\tilde{b}))$. If μ_Y is absolutely continuous with respect to the Lebesgue measure, then there exists a density function $\rho_Y : \mathbb{R} \rightarrow [0, +\infty)$ such that $E[Y] = \int_{\mathbb{R}^N} y \rho_Y(y) dy$. Analogously, whenever $Y_i \in L_P^2(\Omega)$ for $i = 1, \dots, d$, the covariance matrix of Y , $Cov[Y] \in \mathbb{R}^{d \times d}$, is defined by $Cov[Y](i, j) = Cov(Y_i, Y_j) = E[(Y_i - E[Y_i])(Y_j - E[Y_j])]$, $i, j = 1, \dots, d$. Besides this, whenever $u(\omega, x)$ is a stochastic process the positive semidefinite function $Cov[u](x_1, x_2) = Cov[u(x_1), u(x_2)] = Cov[u(x_2), u(x_1)]$ is the covariance function of the stochastic process u .

To introduce the notion of stochastic Sobolev spaces we first recall the definition of stochastic weak derivatives. Let $v \in L_P^2(\Omega) \otimes L^2(D)$; then the α stochastic weak derivative of v , $w = \partial_{\alpha} v \in L_P^2(\Omega) \otimes L^2(D)$, satisfies $\int_D v(\omega, x) \partial^{\alpha} \phi(x) dx = (-1)^{|\alpha|} \int_D w(\omega, x) \phi(x) dx \forall \phi \in C_0^{\infty}(D)$ a.s.

We shall work with stochastic Sobolev spaces $\widetilde{W}^{s,q}(D) = L_P^q(\Omega, W^{s,q}(D))$ containing stochastic functions, $v : \Omega \times D \rightarrow \mathbb{R}$, that are measurable with respect to the

product σ -algebra $\mathcal{F} \otimes B(D)$ and equipped with the averaged norms $\|v\|_{\tilde{W}^{s,q}(D)} = (E[\|v\|_{W^{s,q}(D)}^q])^{1/q} = (E[\sum_{|\alpha| \leq s} \int_D |\partial^\alpha v|^q dx])^{1/q}$, $1 \leq q < +\infty$, and $\|v\|_{\tilde{W}^{s,\infty}(D)} = \max_{|\alpha| \leq s} (\text{ess sup}_{\Omega \times D} |\partial^\alpha v|)$. Observe that if $v \in \tilde{W}^{s,q}(D)$, then $v(\omega, \cdot) \in W^{s,q}(D)$ a.s. and $\partial^\alpha v(\cdot, x) \in L_P^q(\Omega)$ a.e. on D for $|\alpha| \leq s$. Whenever $q = 2$, the above space is a Hilbert space, i.e., $\tilde{W}^{s,2}(D) = \tilde{H}^s(D) \simeq L_P^2(\Omega) \otimes H^s(D)$.

2.2. Existence and uniqueness for the solution of a linear stochastic elliptic problem. Let us consider the tensor product Hilbert space $H = \tilde{H}_0^1(D) \simeq L_P^2(\Omega; H_0^1(D))$ endowed with the inner product $(v, u)_H \equiv E[\int_D \nabla v \cdot \nabla u dx]$.

Define the bilinear form, $\mathcal{B} : H \times H \rightarrow \mathbb{R}$, by $\mathcal{B}(v, w) \equiv E[\int_D a \nabla v \cdot \nabla w dx] \forall v, w \in H$. The standard assumption (1.2) yields both the continuity and the coercivity of \mathcal{B} ; i.e., $|\mathcal{B}(v, w)| \leq a_{\max} \|v\|_H \|w\|_H \forall v, w \in H$, and $a_{\min} \|v\|_H^2 \leq \mathcal{B}(v, v) \forall v \in H$. A direct application of the Lax–Milgram lemma (cf. [11]) implies the existence and uniqueness for the solution to the following variational formulation: find $u \in H$ such that

$$(2.1) \quad \mathcal{B}(u, v) = \mathcal{L}(v) \quad \forall v \in H.$$

Here $\mathcal{L}(v) \equiv E[\int_D f v dx] \forall v \in H$ defines a bounded linear functional since the random field f satisfies (1.4). Since the domain D is convex and bounded and assumptions (1.2), (1.3) on the diffusion a hold, the theory of elliptic regularity (cf. [20]) implies that the solution of (1.1) satisfies $u(\omega, \cdot) \in H^2(D) \cap H_0^1(D)$ a.s. Moreover, standard arguments from measure theory show that the solution to (2.1) also solves (1.1). The formulation (2.1), together with assumption (2.1) on finite dimensional noise, gives the basis for the stochastic Galerkin finite element method (SGFEM) introduced in sections 5 and 6, while formulation (1.1) is the basis for the Monte Carlo Galerkin finite element method (MCGFEM), discussed in section 4.

2.3. Continuity with respect to the coefficients a and f . Since the coefficients a and f are not known exactly, in the next proposition we consider a perturbed weak formulation and estimate the size of the corresponding perturbation in the solution. The proof uses standard estimates and is included in [6].

PROPOSITION 2.1. *Let $(H, (\cdot, \cdot)_H)$ be a Hilbert space. Consider two symmetric bilinear forms $\mathcal{B}, \hat{\mathcal{B}} : H \times H \rightarrow \mathbb{R}$ that are H -coercive and bounded; i.e., there exist real constants $0 < a_{\min} \leq a_{\max}$ such that*

$$(2.2) \quad a_{\min} \|v\|_H^2 \leq \min\{\mathcal{B}(v, v), \hat{\mathcal{B}}(v, v)\} \quad \forall v \in H$$

and

$$(2.3) \quad \max\{|\mathcal{B}(v, w)|, |\hat{\mathcal{B}}(v, w)|\} \leq a_{\max} \|v\|_H \|w\|_H \quad \forall v, w \in H.$$

Consider two bounded linear functionals, $\mathcal{L}, \hat{\mathcal{L}} \in H'$, and let $u, \hat{u} \in H$ be the solutions of the problems

$$\begin{aligned} \mathcal{B}(u, v) &= \mathcal{L}(v) \quad \forall v \in H, \\ \hat{\mathcal{B}}(\hat{u}, v) &= \hat{\mathcal{L}}(v) \quad \forall v \in H. \end{aligned}$$

If, in addition, we know that there exist Banach spaces, V_1, V_2 , and positive constants, C, γ' , such that $u \in V_2 \subseteq H \subset V_1$, $\|\cdot\|_{V_1} \leq C \|\cdot\|_H$, and

$$(2.4) \quad |(\hat{\mathcal{B}} - \mathcal{B})(w, v)| \leq \gamma' \|w\|_{V_1} \|v\|_{V_2} \quad \forall w \in H, v \in V_2,$$

then

$$(2.5) \quad \|u - \hat{u}\|_H \leq \frac{1}{a_{\min}} (\|\mathcal{L} - \hat{\mathcal{L}}\|_{H'} + C\gamma' \|u\|_{V_2}).$$

Next we consider a perturbation of (2.1). A direct application of Proposition 2.1 yields the following estimate.

COROLLARY 2.1. *Let $1 < p < +\infty$ with $1/p + 1/q = 1$. Consider the Hilbert space $H = \tilde{H}_0^1(D)$ and perturbed coefficients, \hat{a}, \hat{f} , satisfying $0 < a_{\min} \leq \hat{a} \leq a_{\max} < \infty$, $(P \otimes dx)$ a.e. on $D \times \Omega$, $\hat{f} \in \tilde{L}^2(D)$. Let u and \hat{u} solve $E[\int_D \hat{a} \nabla \hat{u} \cdot \nabla v dx] = E[\int_D \hat{f} v dx]$ $\forall v \in H$, $E[\int_D a \nabla u \cdot \nabla v dx] = E[\int_D f v dx] \forall v \in H$. Besides this, assume that the solution u belongs to the stochastic Sobolev space $\tilde{W}^{1,2q}(D)$. Then*

$$\|u - \hat{u}\|_{\tilde{H}_0^1(D)} \leq \frac{1}{a_{\min}} (C_D \|\hat{f} - f\|_{\tilde{L}^2(D)} + \|a - \hat{a}\|_{L^{2p}(D)} \|u\|_{\tilde{W}^{1,2q}(D)}),$$

with $C_D > 0$ being the Poincaré constant for the domain D ; i.e., $\|v\|_{L^2(D)} \leq C_D \|v\|_{H_0^1(D)} \forall v \in H_0^1(D)$.

Proof. Take $V_1 = \tilde{H}_0^1(D)$ and $V_2 = \tilde{W}^{1,2q}(D)$. In order to apply (2.5) it is enough to bound the difference of bilinear forms

$$\begin{aligned} & \left| E \left[\int_D (a - \hat{a}) \nabla u \cdot \nabla v dx \right] \right| \\ & \leq \left(E \left[\int_D (a - \hat{a})^2 |\nabla u|^2 dx \right] \right)^{1/2} \left(E \left[\int_D |\nabla v|^2 dx \right] \right)^{1/2} \\ & \leq \left(E \left[\int_D (a - \hat{a})^{2p} dx \right] \right)^{1/2p} \left(E \left[\int_D |\nabla u|^{2q} dx \right] \right)^{1/2q} \left(E \left[\int_D |\nabla v|^2 dx \right] \right)^{1/2}. \quad \square \end{aligned}$$

2.4. Karhunen–Loèeve expansions and finite dimensional noise. Here we recall the Karhunen–Loèeve expansion, a suitable tool for the approximation of stochastic processes. Consider a stochastic process a with continuous covariance function $Cov[a] : \overline{D} \times \overline{D} \rightarrow \mathbb{R}$. Besides this, let $\{(\lambda_n, b_n)\}_{n=1}^{\infty}$ denote the sequence of eigenpairs associated with the compact self-adjoint operator that maps

$$f \in L^2(D) \mapsto \int_D Cov[a](x, \cdot) f(x) dx \in L^2(D).$$

Its nonnegative eigenvalues, $\sqrt{\int_{D \times D} (Cov[a](x_1, x_2))^2 dx_1 dx_2} \geq \lambda_1 \geq \lambda_2 \geq \dots \geq 0$, satisfy $\sum_{n=1}^{+\infty} \lambda_n = \int_D Var[a](x) dx$. The corresponding eigenfunctions are orthonormal, i.e., $\int_D b_i(x) b_j(x) dx = \delta_{ij}$. The truncated Karhunen–Loèeve expansion of the stochastic process a (cf. [33]) is

$$a_N(\omega, x) = E[a](x) + \sum_{n=1}^N \sqrt{\lambda_n} b_n(x) Y_n(\omega),$$

where the real random variables, $\{Y_n\}_{n=1}^{\infty}$, are mutually uncorrelated and have mean zero and unit variance. These random variables are uniquely determined by $Y_n(\omega) = \frac{1}{\sqrt{\lambda_n}} \int_D (a(\omega, x) - E[a](x)) b_n(x) dx$ for $\lambda_n > 0$. Then, by Mercer's theorem (cf. [37, p. 245]), we have

$$\sup_{x \in D} E[(a - a_N)^2](x) = \sup_{x \in D} \sum_{n=N+1}^{+\infty} \lambda_n b_n^2(x) \rightarrow 0 \text{ as } N \rightarrow \infty.$$

If, in addition,

- the images $Y_n(\Omega)$, $n = 1, \dots$, are uniformly bounded in \mathbb{R} ,
- the eigenfunctions b_n are smooth, which is the case when the covariance function is smooth,
- and the eigenpairs have at least the decay $\sqrt{\lambda_n} \|b_n\|_{L^\infty(D)} = \mathcal{O}(\frac{1}{1+n^s})$ for some $s > 1$,

then $\|a - a_N\|_{\tilde{L}^\infty(D)} \rightarrow 0$. Notice that for larger values of the decay exponent s we can also obtain the convergence of higher spatial derivatives of a_N in $\tilde{L}^\infty(D)$. The last two conditions can be readily verified once the covariance function of a is known. However, observe that it is also necessary to verify the uniform coercivity of a_N , which depends on the probability distributions of Y_n , $n = 1, \dots$.

In many problems the source of the randomness can be approximated using just a small number of mutually uncorrelated, sometimes mutually independent, random variables. Take, for example, the case of a truncated Karhunen–Loève expansion described previously.

Assumption 2.1 (finite dimensional noise). Whenever we apply some numerical method to solve (1.1) we assume that the coefficients used in the computations, $a, f : \Omega \times D \rightarrow \mathbb{R}$, are finite Karhunen–Loève expansions; i.e., $a(\omega, x) = E[a](x) + \sum_{n=1}^N \sqrt{\lambda_n} b_n(x) Y_n(\omega)$ and $f(\omega, x) = E[f](x) + \sum_{n=1}^N \sqrt{\lambda_n} \hat{b}_n(x) Y_n(\omega)$, where $\{Y_n\}_{n=1}^N$ are real random variables with mean value zero and unit variance, are uncorrelated, and have images, $\Gamma_n \equiv Y_n(\Omega)$, that are bounded intervals in \mathbb{R} for $n = 1, \dots, N$. Moreover, we assume that each Y_n has a density function $\rho_n : \Gamma_n \rightarrow \mathbb{R}^+$ for $n = 1, \dots, N$.

In what follows we use the notation $\rho(y) \forall y \in \Gamma$ for the joint probability density of (Y_1, \dots, Y_N) and $\Gamma \equiv \prod_{n=1}^N \Gamma_n \subset \mathbb{R}^N$ for the support of such probability density.

After making Assumption 2.1, we have by the Doob–Dynkin lemma (cf. [36]) that u , the solution corresponding to the stochastic partial differential equation (1.1), can be described by just a finite number of random variables, i.e., $u(\omega, x) = u(Y_1(\omega), \dots, Y_N(\omega), x)$. The number N has to be large enough so that the approximation error is sufficiently small. Now the goal is to approximate the function $u(y, x)$. In addition, the stochastic variational formulation (2.1) has a deterministic equivalent in the following: find $u \in L_\rho^2(\Gamma) \otimes H_0^1(D)$ such that

$$(2.6) \quad \begin{aligned} & \int_{\Gamma} \rho(y) \int_D a(y, x) \nabla u(y, x) \cdot \nabla v(y, x) dx dy \\ &= \int_{\Gamma} \rho(y) \int_D f(y, x) v(y, x) dx dy \quad \forall v \in L_\rho^2(\Gamma) \otimes H_0^1(D). \end{aligned}$$

In this work the gradient notation, ∇ , always means differentiation with respect to $x \in D$ only, unless otherwise stated. The corresponding strong formulation for (2.6) is an elliptic partial differential equation with an N -dimensional parameter, i.e.,

$$(2.7) \quad \begin{aligned} & -\nabla \cdot (a(y, x) \nabla u(y, x)) = f(y, x) \quad \forall (y, x) \in \Gamma \times D, \\ & u(y, x) = 0 \quad \forall (y, x) \in \Gamma \times \partial D. \end{aligned}$$

Making Assumption 2.1 is a crucial step, turning the original stochastic elliptic equation (1.1) into a deterministic parametric elliptic one and allowing the use of finite element and finite difference techniques to approximate the solution of the resulting deterministic problem.

Truncation of the outcomes set, Γ . For the sake of efficiency, it may be useful to compute the solution of (2.7) in a subdomain with strictly positive probability, $\Gamma_0 \subset \Gamma$. Besides, we assume the probability density of Y to be strictly positive in Γ_0 . In that case, we approximate the function

$$E[u(Y, \cdot) 1_{\{Y \in \Gamma_0\}}] = E[u(Y, \cdot) | Y \in \Gamma_0] P(Y \in \Gamma_0)$$

instead of the original $E[u]$. If \bar{u} is an approximation of u in Γ_0 , then we have the splitting

$$(2.8) \quad \begin{aligned} & \|E[u(Y, \cdot)] - E[\bar{u}(Y, \cdot) 1_{\{Y \in \Gamma_0\}}]\| \\ & \leq \|E[u(Y, \cdot)] - E[u(Y, \cdot) 1_{\{Y \in \Gamma_0\}}]\| + \|E[u(Y, \cdot) - \bar{u}(Y, \cdot) | Y \in \Gamma_0]\| P(Y \in \Gamma_0). \end{aligned}$$

Property 2.1 below gives a simple estimate for the first error contribution, which is related to the truncation of Γ . The second error contribution in (2.8) is the discretization error, and it will be analyzed for each numerical approximation separately; see sections 4, 5, and 6. In those sections we shall simplify the notation by writing $\Gamma = \Gamma_0$ and work with the corresponding conditional probability space.

PROPERTY 2.1. *Let u be the solution of the problem (2.7); then there exists a constant C such that*

$$(2.9) \quad \|E[u(Y, \cdot)] - E[u(Y, \cdot) 1_{\{Y \in \Gamma\}}]\|_{H_0^1(D)} \leq C \sqrt{P(Y \notin \Gamma_0)} \|f\|_{L_\rho^2(\Gamma \setminus \Gamma_0) \otimes L^2(D)}.$$

3. The finite element spaces. In this section, we define tensor product finite element spaces on the set $\Gamma \times D$, which we will use to construct approximations of the solution of the parametric boundary value problem (2.7), stating their approximation properties.

3.1. Finite element spaces on the spatial set $D \subset \mathbb{R}^d$: h -version. Consider a family of finite element approximation spaces, $X_h^d \subset H_0^1(D)$, consisting of piecewise linear continuous functions on conforming triangulations (of simplices), T_h^d , of the convex polyhedral domain, $D \subset \mathbb{R}^d$, with a maximum mesh spacing parameter $h > 0$. We will always assume that the triangulations are nondegenerate (sometimes also called regular); cf. [11, p. 106]. Then (cf. [11, 13]) the finite element spaces X_h^d satisfy a standard approximation estimate, namely, that for all $v \in H^2(D) \cap H_0^1(D)$

$$(3.1) \quad \min_{\chi \in X_h^d} \|v - \chi\|_{H_0^1(D)} \leq C h \|v\|_{H^2(D)},$$

where $C > 0$ is a constant independent of v and h .

3.2. Tensor product finite element spaces on the outcomes set $\Gamma \subset \mathbb{R}^N$: k -version. Let $\Gamma = \prod_{n=1}^N \Gamma_n$ be as in subsection 2.4. Consider a partition of Γ consisting of a finite number of disjoint \mathbb{R}^N -boxes, $\gamma = \prod_{n=1}^N (a_n^\gamma, b_n^\gamma)$, with $(a_n^\gamma, b_n^\gamma) \subset \Gamma_n$ for $n = 1, \dots, N$. The mesh spacing parameters, $k_n > 0$, are defined by $k_n \equiv \max_\gamma |b_n^\gamma - a_n^\gamma|$ for $1 \leq n \leq N$. For every nonnegative integer multi-index $q = (q_1, \dots, q_N)$ consider the finite element approximation space of (discontinuous) piecewise polynomials with degree at most q_n on each direction y_n , $Y_k^q \subset L^2(\Gamma)$. Thus, if $\varphi \in Y_k^q$, its restriction to each of the partition boxes satisfies $\varphi|_\gamma \in \text{span}(\prod_{n=1}^N y_n^{\alpha_n} : \alpha_n \in \mathbb{N} \text{ and } \alpha_n \leq q_n, n = 1, \dots, N)$. It is easy to verify

that the finite element spaces Y_k^q have the following approximation property: for all $v \in H^{q+1}(\Gamma)$,

$$(3.2) \quad \min_{\varphi \in Y_k^q} \|v - \varphi\|_{L^2(\Gamma)} \leq \sum_{n=1}^N \left(\frac{k_n}{2} \right)^{q_n+1} \frac{\|\partial_{y_n}^{q_n+1} v\|_{L^2(\Gamma)}}{(q_n+1)!}.$$

3.3. Tensor product finite element spaces on $\Gamma \times D$: $k \times h$ -version. Here we will discuss some approximation properties of the following tensor product finite element spaces:

(3.3)

$$Y_k^q \otimes X_h^d \equiv \{ \psi = \psi(y, x) \in L^2(\Gamma \times D) : \psi \in \text{span}(\varphi(y) \chi(x) : \varphi \in Y_k^q, \chi \in X_h^d) \}$$

with X_h^d and Y_k^q as in subsections 3.1 and 3.2.

For later use we recall the definition of the standard L^2 -projection operators $\Pi_k^q : L^2(\Gamma) \rightarrow Y_k^q$ by

$$(3.4) \quad (\Pi_k^q w - w, \varphi)_{L^2(\Gamma)} = 0 \quad \forall \varphi \in Y_k^q, \quad \forall w \in L^2(\Gamma)$$

and the H_0^1 -projection operator $\mathcal{R}_h : H_0^1(D) \rightarrow X_h^d$ by

$$(3.5) \quad (\nabla(\mathcal{R}_h v - v), \nabla \chi)_{L^2(D)} = 0 \quad \forall \chi \in X_h^d, \quad \forall v \in H_0^1(D).$$

Estimates (3.1) and (3.2) imply

$$(3.6) \quad \begin{aligned} \|v - \mathcal{R}_h v\|_{H_0^1(D)} &\leq C h \|v\|_{H^2(D)}, \\ \|w - \Pi_k^q w\|_{L^2(\Gamma)} &\leq \sum_{n=1}^N \left(\frac{k_n}{2} \right)^{q_n+1} \frac{\|\partial_{y_n}^{q_n+1} w\|_{L^2(\Gamma)}}{(q_n+1)!} \end{aligned}$$

for all $v \in H^2(D) \cap H_0^1(D)$ and $w \in H^{q+1}(\Gamma)$. We now state an approximation property for the tensor product finite element spaces defined in (3.3) which is a direct implication of the approximation properties of the spaces Y_k^q and X_h^d .

PROPOSITION 3.1. *There exists a constant $C > 0$ independent of h , N , q , and k such that*

$$(3.7) \quad \begin{aligned} \inf_{\psi \in Y_k^q \otimes X_h^d} \|v - \psi\|_{L^2(\Gamma; H_0^1(D))} \\ \leq C \left\{ h \|v\|_{L^2(\Gamma; H^2(D))} + \sum_{n=1}^N \left(\frac{k_n}{2} \right)^{q_n+1} \frac{\|\partial_{y_n}^{q_n+1} v\|_{L^2(\Gamma; H_0^1(D))}}{(q_n+1)!} \right\} \end{aligned}$$

for all $v \in C^{q+1}(\Gamma; H^2(D) \cap H_0^1(D))$.

Proof. Since $\Pi_k^q(\mathcal{R}_h v) \in Y_k^q \otimes X_h^d$, using (3.6) we obtain

$$(3.8) \quad \begin{aligned} \inf_{\psi \in Y_k^q \otimes X_h^d} \|v - \psi\|_{L^2(\Gamma; H_0^1(D))} &\leq \|v - \Pi_k^q(\mathcal{R}_h v)\|_{L^2(\Gamma; H_0^1(D))} \\ &\leq \|v - \mathcal{R}_h v\|_{L^2(\Gamma; H_0^1(D))} \\ &\quad + \|\mathcal{R}_h v - \Pi_k^q(\mathcal{R}_h v)\|_{L^2(\Gamma; H_0^1(D))} \\ &\leq C h \|v\|_{L^2(\Gamma; H^2(D))} \\ &\quad + \|\mathcal{R}_h v - \Pi_k^q(\mathcal{R}_h v)\|_{L^2(\Gamma; H_0^1(D))}. \end{aligned}$$

Applying the estimate (3.2) and using the boundedness of \mathcal{R}_h in $H_0^1(D)$ yield

$$\begin{aligned} \|\mathcal{R}_h v - \Pi_k^q(\mathcal{R}_h v)\|_{L^2(\Gamma; H_0^1(D))} &\leq \|v - \Pi_k^q v\|_{L^2(\Gamma; H_0^1(D))} \\ &\leq \sum_{n=1}^N \left(\frac{k_n}{2}\right)^{q_n+1} \frac{\|\partial_{y_n}^{q_n+1} v\|_{L^2(\Gamma; H_0^1(D))}}{(q_n+1)!}. \end{aligned}$$

The estimate (3.7) follows, combining (3.8) with the last estimate. \square

3.4. Tensor product finite element spaces on $\Gamma \times D$: $p \times h$ -version. This approximation space is in fact a particular case of the $k \times h$ -version with no k partition of Γ , i.e., $k_n = |\Gamma_n|$, $n = 1, \dots, N$. Instead, only the polynomial degree is increased. Here the multi-index $p = (p_1, \dots, p_N)$ plays the role of the q from section 3.3, and we use the tensor finite element space $Z^p = \bigotimes_{n=1}^N Z_n^{p_n}$, where the one dimensional global polynomial subspaces, $Z_n^{p_n}$, are defined by $Z_n^{p_n} = \{v : \Gamma_n \rightarrow \mathbb{R} : v \in \text{span}(y^s, s = 0, \dots, p_n)\}$, $n = 1, \dots, N$.

4. Monte Carlo Galerkin finite element method. In this section we describe the use of the standard Monte Carlo Galerkin finite element method (MCGFEM) to construct approximations of the expected value of the solution.

Formulation of the MCGFEM.

- Give a number of realizations, M , a piecewise linear finite element space on D , X_h^d , as defined in subsection 3.1.
- For each $j = 1, \dots, M$, sample iid realizations of the diffusion $a(\omega_j, \cdot)$ and the load $f(\omega_j, \cdot)$, based on realizations of $\{Y_n\}_{n=1}^N$, and find a corresponding approximation $u_h(\omega_j, \cdot) \in X_h^d$ such that

$$(4.1) \quad (a(\omega_j, \cdot) \nabla u_h(\omega_j, \cdot), \nabla \chi)_{L^2(D)} = (f(\omega_j, \cdot), \chi)_{L^2(D)} \quad \forall \chi \in X_h^d.$$

- Finally, use the sample average $\frac{1}{M} \sum_{j=1}^M u_h(\omega_j, \cdot)$ to approximate $E[u]$. \square

Here we consider only the case where X_h^d is the same for all realizations; i.e., the spatial triangulation is deterministic. The computational error naturally separates into two parts:

(4.2)

$$E[u] - \frac{1}{M} \sum_{j=1}^M u_h(\omega_j, \cdot) = (E[u] - E[u_h]) + \left(E[u_h] - \frac{1}{M} \sum_{j=1}^M u_h(\omega_j, \cdot) \right) \equiv \mathcal{E}_h + \mathcal{E}_S.$$

The size of the spatial triangulation controls the space discretization error \mathcal{E}_h , while the number of realizations, M of u_h , controls the statistical error \mathcal{E}_S .

To study the behavior of the statistical error, let us first consider the random variable $\|\mathcal{E}_S\|_{H_0^1(D)}$ which, due to the independence of the realizations $u_h(\omega_j, \cdot)$, $j = 1, \dots, M$, satisfies the estimate

$$(4.3) \quad M E[\|\mathcal{E}_S\|_{H_0^1(D)}^2] \leq \|u_h\|_{H_0^1(D)}^2 \leq \left(\frac{C_D}{a_{\min}} \right)^2 \|f\|_{L^2(D)}^2,$$

and a similar result also holds in $L^2(D)$. Then, thanks to (4.3) we have that, for either $H = L^2(D)$ or $H = H_0^1(D)$, the statistical error $\|\mathcal{E}_S\|_H$ tends a.s. to zero as we increase the number of realizations; i.e., we have the following.

PROPOSITION 4.1. *Suppose that there exists a constant $C > 0$ independent of M and h such that the statistical error in H norm satisfies*

$$(4.4) \quad M E[\|\mathcal{E}_S\|_H^2] \leq C \quad \forall M, h.$$

Then, taking the number of realizations, M_k , increasingly from the set $\{2^k : k \in \mathbb{N}\}$, we have, for any $\alpha \in (0, 1/2)$ and any choice of mesh size h , $\lim_{M_k \rightarrow \infty} M_k^\alpha \|\mathcal{E}_S\|_H = 0$ a.s.

Proof. Let $\epsilon > 0$. Then (4.4) and Markov's inequality give

$$P((M_k)^\alpha \|\mathcal{E}_S\|_H > \epsilon) \leq \frac{E[(M_k)^{2\alpha} \|\mathcal{E}_S\|_H^2]}{\epsilon^2} \leq \frac{C}{\epsilon^2 (M_k)^{1-2\alpha}}.$$

Thus, for $\alpha \in (0, 1/2)$ we have

$$\sum_{k=1}^{\infty} P(M_k^\alpha \|\mathcal{E}_S\|_H > \epsilon) \leq \frac{C}{\epsilon^2} \sum_{k=1}^{\infty} \frac{1}{M_k^{1-2\alpha}} \leq \frac{C}{\epsilon^2} \sum_{k=1}^{\infty} \frac{1}{(2^{1-2\alpha})^k} < \infty,$$

which, together with the Borel–Cantelli lemma, finishes the proof. \square

Under the same assumptions as in Proposition 4.1 we have that for any given $\epsilon > 0$ there exists a constant $C > 0$ independent of ϵ, M , and h such that

$$(4.5) \quad P\left(\|\mathcal{E}_S\|_H > \frac{\epsilon}{\sqrt{M}}\right) \leq \frac{C}{\epsilon^2}.$$

Thus, within a given confidence level we have the usual convergence rate for the Monte Carlo method, which is independent of the mesh size h . Next we present error estimates for the space discretization error, namely, we have the following.

PROPOSITION 4.2 (spatial discretization error estimates). *There holds*

$$\begin{aligned} h\|u - u_h\|_{H_0^1(D)} + \|u - u_h\|_{L^2(D)} &\leq C h^2 \|f\|_{L^2(D)} \quad \text{a.s.}, \\ h\|E[u] - E[u_h]\|_{H_0^1(D)} + \|E[u] - E[u_h]\|_{L^2(D)} &\leq C h^2 E[\|f\|_{L^2(D)}^2]^{1/2}. \end{aligned}$$

The results from Proposition 4.2 and estimate (4.5) will be used in section 8 to compare the MCGFEM with other discretizations for (1.1).

5. Stochastic Galerkin finite element method: $k \times h$ -version. This section defines and analyzes the $k \times h$ -version of the stochastic Galerkin finite element method ($k \times h$ -SGFEM) which, via a Galerkin variational formulation, yields approximations, $u_{kh} \in Y_k^q \otimes X_h^d$, of the solution u of the parametric elliptic boundary value problem (2.7).

Formulation of the $k \times h$ -SGFEM. Denote by $q = (q_1, \dots, q_N) \in \mathbb{N}^N$ a multi-index, and let Γ be a bounded box in \mathbb{R}^N . The $k \times h$ -SGFEM approximation is the tensor product, $u_{kh} \in Y_k^q \otimes X_h^d$, such that

(5.1)

$$(u_{kh}, \psi)_E \equiv \int_{\Gamma} \rho (a \nabla u_{kh}, \nabla \psi)_{L^2(D)} dy = \int_{\Gamma} \rho (f, \psi)_{L^2(D)} dy \quad \forall \psi \in Y_k^q \otimes X_h^d.$$

Recall that $\rho : \Gamma \rightarrow (0, +\infty)$ is the density function of the vector-valued random variable $Y : \Omega \rightarrow \Gamma \subset \mathbb{R}^N$. Hence, the assumption (1.2) on the random function $a(\omega, x) \equiv a(Y(\omega), x)$ reads

$$(5.2) \quad a(y, x) \in [a_{\min}, a_{\max}] \quad \forall (y, x) \in \overline{\Gamma \times D}.$$

Although the analysis can be generalized [6], we now focus on the practical case where both a and f are truncated Karhunen–Loëve expansions. Later, section 7 discusses how to compute efficiently u_{kh} , the solution of (5.1), by a double orthogonal polynomials technique. By Lemma 4.1 in [31], the solution u of (2.7) satisfies $u \in C^\infty(\Gamma; H^2(D) \cap H_0^1(D))$. Use (5.2) and (2.7) to obtain

$$(5.3) \quad \|u(y, \cdot)\|_{H_0^1(D)} \leq \frac{C_D}{a_{\min}} \|f(y, \cdot)\|_{L^2(D)} \quad \forall y \in \Gamma,$$

where C_D is the constant of the Poincaré–Friedrichs inequality on D . Also, elliptic regularity yields

$$(5.4) \quad \|u(y, \cdot)\|_{H^2(D)} \leq C_{0,B} \|f(y, \cdot)\|_{L^2(D)} \quad \forall y \in \Gamma,$$

where $C_{0,B}$ is a constant which depends on D and $\|a\|_{L^\infty(\Gamma; W^{1,\infty}(D))}$. Finally, take derivatives with respect to y_n in (2.7), proceed as in the derivation of (5.3), and follow an inductive argument arriving at

$$(5.5) \quad \frac{\|\partial_{y_n}^{q_n+1} u(y, \cdot)\|_{H_0^1(D)}}{(q_n + 1)!} \leq (r_n)^{q_n} \frac{C_D}{a_{\min}} (\|\partial_{y_n} f(y, \cdot)\|_{L^2(D)} + r_n \|f(y, \cdot)\|_{L^2(D)}), \quad q_n \geq 0,$$

with $r_n \equiv \sqrt{\lambda_n} \|\frac{b_n}{a}\|_{L^\infty(\Gamma \times D)}$, and $n = 1, \dots, N$. As a consequence of (3.7), (5.4), and (5.5) we have an a priori error estimate for the $k \times h$ -SGFEM in the energy norm.

PROPOSITION 5.1. *Let u be the solution of the problem (2.7) and $u_{kh} \in Y_k^q \otimes X_h^d$ be the $k \times h$ -SGFEM approximations of u defined in (5.1). If $\rho \in L^\infty(\Gamma)$ and $f \neq 0$, then*

$$(5.6) \quad \frac{\|u - u_{kh}\|_E}{\sqrt{\|a \rho\|_{L^\infty(\Gamma \times D)} \|f\|_{L^2(\Gamma; L^2(D))}}} \leq Ch + \frac{C_D}{a_{\min}} \sum_{n=1}^N \frac{k_n}{2} \left(\frac{k_n r_n}{2} \right)^{q_n} \left(\frac{\|\partial_{y_n} f\|_{L^2(\Gamma; L^2(D))}}{\|f\|_{L^2(\Gamma; L^2(D))}} + r_n \right),$$

where the constant C depends on D and a and is independent of q, k, h , and u .

The next step is to use Proposition 5.1 together with a duality technique to estimate the $H_0^1(D)$ and $L^2(D)$ errors in the approximation of the expected value of $u(Y, \cdot)$.

THEOREM 5.1. *Let u be the solution of the problem (2.7) and $u_{kh} \in Y_k^q \otimes X_h^d$ be the $k \times h$ -SGFEM approximations of u defined in (5.1). If $\rho \in L^\infty(\Gamma)$, then for $\ell = 0, 1$ we have*

$$(5.7) \quad \frac{\|E[u(Y, \cdot)] - E[u_{kh}(Y, \cdot)]\|_{H^\ell(D)}}{\|a \rho\|_{L^\infty(\Gamma \times D)}} \leq C \left(h^{2-\ell} + \sum_{n=1}^N \left(\frac{k_n}{2} \right)^{2-\ell} \left(\frac{k_n r_n}{2} \right)^{(2-\ell)q_n} \right).$$

The constant C depends on D, f , and a , and it is independent of k, q , and h .

Proof. The case $\ell = 1$ follows directly from Proposition 5.1, so we now prove the case $\ell = 0$. Denote $e \equiv u - u_{kh}$ and use the auxiliary function \hat{u} that solves

$$(5.8) \quad \begin{aligned} -\nabla_x \cdot (a(y, \cdot) \nabla \hat{u}(y, \cdot)) &= E[e](\cdot) \quad \text{in } D, \\ \hat{u}(y, \cdot) &= 0 \quad \text{on } \partial D. \end{aligned}$$

Then (5.4) reads

$$(5.9) \quad \|\hat{u}(y, \cdot)\|_{H^2(D)} \leq C_{0,B} \|E[e]\|_{L^2(D)} \quad \forall y \in \Gamma,$$

and, since $E[e]$ is independent of y , the estimate (5.5) for the problem (5.8) reads

$$(5.10) \quad \frac{\|\partial_{y_n}^{q_n+1} \hat{u}(y, \cdot)\|_{H_0^1(D)}}{(q_n+1)!} \leq (r_n)^{q_n+1} \frac{C_D}{a_{min}} \|E[e]\|_{L^2(D)} \quad \forall y \in \Gamma.$$

Now use Galerkin orthogonality, together with (5.8), to obtain

$$\int_{\Gamma} \rho (E[e], e)_{L^2(D)} dy = \int_{\Gamma} \rho (\partial_{y_n}^q \hat{u}, \psi)_{L^2(D)} dy \quad \forall \psi \in Y_k^q \otimes X_h^d,$$

which yields, by the Cauchy–Schwarz inequality,

$$(5.11) \quad \|E[e]\|_{L^2(D)}^2 \leq \tilde{B}_1 \tilde{B}_2,$$

with

$$\tilde{B}_1 \equiv \left(\int_{\Gamma} \rho \|\sqrt{a} \nabla e\|_{L^2(D)}^2 dy \right)^{\frac{1}{2}}$$

and

$$\tilde{B}_2 \equiv \inf_{\psi \in Y_k^q \otimes X_h^d} \left(\int_{\Gamma} \rho \|\sqrt{a} \nabla (\hat{u} - \psi)\|_{L^2(D)}^2 dy \right)^{\frac{1}{2}}.$$

Next observe that \tilde{B}_1 can be bounded using (5.6). Finally, use (3.7), (5.9), and (5.10) to bound \tilde{B}_2 as follows:

$$(5.12) \quad \begin{aligned} \tilde{B}_2 &\leq C \|a \rho\|_{L^\infty(\Gamma \times D)}^{\frac{1}{2}} \left\{ h \|\hat{u}\|_{L^2(\Gamma; H^2(D))} + \sum_{n=1}^N \left(\frac{k_n}{2} \right)^{q_n+1} \frac{\|\partial_{y_n}^{q_n+1} \hat{u}\|_{L^2(\Gamma; H_0^1(D))}}{(q_n+1)!} \right\} \\ &\leq C \|a \rho\|_{L^\infty(\Gamma; L^\infty(D))}^{\frac{1}{2}} \left(h + \frac{C_D}{a_{min}} \sum_{n=1}^N \left(\frac{r_n k_n}{2} \right)^{q_n+1} \right) \|E[e]\|_{L^2(D)}. \end{aligned}$$

Combining (5.11), (5.6), and (5.12), the estimate (5.7) follows. \square

The estimates given in Proposition 5.1 and Theorem 5.1 give the optimal order of convergence with respect to k but are not optimal with respect to q . They can be improved by the analysis given in section 6, yielding exponential convergence with respect to q without the need to decrease k .

6. Stochastic Galerkin finite element method: $p \times h$ -version. The goal of this section is to analyze the $p \times h$ -version of the SGFEM method, which does not refine the set Γ . This method yields an exponential rate of convergence with respect to p , the degree of the polynomials used for approximation; cf. Theorem 6.2. The application of the p -version in the y -direction is motivated by the fact that u is analytic with respect to $y \in \Gamma$; cf. Lemma 6.1. The basic assumption for this section is the following.

Assumption 6.1. Let $\hat{\Gamma}_n \equiv \prod_{1 \leq j \leq N, j \neq n} \Gamma_j$, and let \hat{y}_n be an arbitrary element of $\hat{\Gamma}_n$. Then for each $\hat{y}_n \in \hat{\Gamma}_n$ let $\tilde{a}_n(\hat{y}_n) \equiv \min_{x \in D} \{E[a](x) + \sum_{1 \leq j \leq N, j \neq n} \sqrt{\lambda_j} b_j(x) y_j\}$, and assume a slightly stronger uniform coercivity requirement; i.e., there exists a constant $\nu > 0$, independent of N , such that

$$\tilde{a}_n(\hat{y}_n) - \sqrt{\lambda_n} \|b_n\|_{L^\infty(D)} \max_{y \in \Gamma_n} |y| \geq \nu > 0 \quad \forall \hat{y}_n \in \hat{\Gamma}_n. \quad \square$$

Observe that with the above construction we have $0 < \nu \leq a_{\min}$.

$p \times h$ -version of the SGFEM method. The $p \times h$ -version SGFEM approximation is the tensor product $u_h^p \in Z^p \otimes X_h^d$ (cf. section 3.4) that satisfies

$$(6.1) \quad (u_h^p, \chi)_E \equiv \int_{\Gamma} \rho (a \nabla u_h^p, \nabla \chi)_{L^2(D)} dy = \int_{\Gamma} \rho (f, \chi)_{L^2(D)} dy \quad \forall \chi \in Z^p \otimes X_h^d.$$

6.1. Error estimates. A first step in the analysis of the $p \times h$ -version is to study the energy error, i.e., to consider

$$\begin{aligned} \|u - u_h^p\|_E &\leq \sqrt{\|\rho a\|_{L^\infty(\Gamma \times D)}} \min_{v \in Z^p \otimes X_h^d} \|u - v\|_{L^2(\Gamma) \otimes H_0^1(D)} \\ &\leq \sqrt{\|\rho a\|_{L^\infty(\Gamma \times D)}} \left\{ \min_{v \in Z^p \otimes H_0^1(D)} \|u - v\|_{L^2(\Gamma) \otimes H_0^1(D)} \right. \\ &\quad \left. + \min_{v \in L^2(\Gamma) \otimes X_h^d} \|u - v\|_{L^2(\Gamma) \otimes H_0^1(D)} \right\}. \end{aligned}$$

This bound splits the error into an $L^2(\Gamma)$ approximation error and a standard $H_0^1(D)$ FEM approximation error. The rest of this section studies the first one, since for the second we can apply the results from Proposition 3.1 together with a density argument. The minimizer

$$\|u - u^p\|_{L^2(\Gamma) \otimes H_0^1(D)} = \min_{v \in Z^p \otimes H_0^1(D)} \|u - v\|_{L^2(\Gamma) \otimes H_0^1(D)}$$

is the projection $u^p = (\Pi_1 \dots \Pi_N)u$ with $\Pi_n : L^2(\Gamma) \otimes H_0^1(D) \rightarrow L^2(\Gamma) \otimes H_0^1(D)$ being the natural extension of the L^2 projection $\bar{\Pi}_n : L^2(\Gamma_n) \rightarrow Z_n^{p_n}$, so the difference $u - u^p$ splits into $u - u^p = (1 - \Pi_1)u + \dots + (\Pi_1 \dots \Pi_{N-1})(1 - \Pi_N)u$. In addition, the boundedness of the projections Π_n , $n = 1, \dots, N$, yields

$$(6.2) \quad \|u - u^p\|_{L^2(\Gamma) \otimes H_0^1(D)} \leq \sum_{n=1}^N \|(1 - \Pi_n)u\|_{L^2(\Gamma) \otimes H_0^1(D)}.$$

Without loss of generality we now estimate the first term on the right-hand side of (6.2), since the other terms have a completely similar behavior. Moreover, since

$$\|(1 - \Pi_1)u\|_{L^2(\Gamma) \otimes H_0^1(D)}^2 = \int_{\hat{\Gamma}_1} \left(\int_{\Gamma_1} \|(1 - \Pi_1)u(y_1, \hat{y}_1, \cdot)\|_{H_0^1(D)}^2 dy_1 \right) d\hat{y}_1$$

it is enough to estimate

$$(6.3) \quad (E_1)^2(\hat{y}_1) \equiv \int_{\Gamma_1} \|(1 - \Pi_1)u(y_1, \hat{y}_1, \cdot)\|_{H_0^1(D)}^2 dy_1,$$

and thus our analysis requires only one dimensional arguments in the y -direction. Let $\Gamma_1 = (y_{min}, y_{max})$, and consider the map $\Psi : [-1, 1] \rightarrow H_0^1(D)$ defined by

$$\Psi(t) = u(y_1(t), \hat{y}_1, \cdot) \in H_0^1(D)$$

with the affine transformation, $y_1 : [-1, 1] \rightarrow \Gamma_1$, $y_1(t) \equiv \left(\frac{y_{max}+y_{min}}{2}\right) + \left(\frac{y_{max}-y_{min}}{2}\right) t$. In the upcoming estimate of the quantities $\|d_n\|_{H_0^1(D)}$, to be proved in Lemma 6.2, we need to consider a continuation of Ψ to the complex plane, namely, the following.

LEMMA 6.1 (complex continuation). *The function $\Psi : [-1, 1] \rightarrow H_0^1(D)$ can be analytically extended to the complex domain.*

Proof. Let $t_0 \in (-1, 1)$. We shall prove that the real function Ψ can be represented as a power series for $|t - t_0| < r_{t_0}$ for some $r_{t_0} > 0$. Since Ψ depends linearly on f , let us assume that $f(y, x) = f(x)$ only, without loss of generality. Let $y(t) = (y_1(t), \hat{y}_1)$, and consider the formal series

$$u_F(t) \equiv \sum_{j=0}^{+\infty} \left(\frac{|\Gamma_1|(t - t_0)}{2} \right)^j u_j,$$

with $u_j \in H_0^1(D)$ satisfying

$$(6.4) \quad \int_D a(y(t_0), \cdot) \nabla u_0 \cdot \nabla v = \int_D f v \quad \forall v \in H_0^1(D)$$

and, for $j \geq 0$,

$$(6.5) \quad \int_D a(y(t_0), \cdot) \nabla u_{j+1} \cdot \nabla v = - \int_D \sqrt{\lambda_1} b_1 \nabla u_j \cdot \nabla v \quad \forall v \in H_0^1(D).$$

This construction implies $\|u_j\|_{H_0^1(D)} \leq (\sqrt{\lambda_1} \|\frac{b_1}{a(y(t_0), \cdot)}\|_{L^\infty(D)})^j \frac{C_D \|f\|_{L^2(D)}}{a_{min}}$, $j \geq 1$, and then

$$\|u_F\|_{H_0^1(D)} \leq \frac{C_D \|f\|_{L^2(D)}}{a_{min}} \frac{1}{1-q} < \infty$$

for $q \equiv \frac{|t-t_0||\Gamma_1|\sqrt{\lambda_1}}{2} \|\frac{b_1}{a(y(t_0), \cdot)}\|_{L^\infty(D)} < 1$. Thus, for any $t_0 \in (-1, 1)$ and $|t - t_0| < r_{t_0} \equiv \frac{2}{|\Gamma_1|\sqrt{\lambda_1} \|\frac{b_1}{a(y(t_0), \cdot)}\|_{L^\infty(D)}}$, the function u_F can be represented as a power series in $t - t_0$. At the same time, we have the equality $u_F(t) = \Psi(t)$ for $t \in (-1, 1)$ since both functions solve the linear elliptic equation

$$\begin{aligned} -\nabla \cdot (a(y(t), x) \nabla u(y(t), x)) &= f(x) \quad \forall x \in D, \\ u(y(t), x) &= 0 \quad \forall x \in \partial D, \end{aligned}$$

which has a unique solution. Then u_F is the analytic continuation of Ψ , and the proof is complete. \square

Remark 6.1. Consider the natural extension of the real variable t to the complex variable η . Observe that $\Psi(\eta)$ from Lemma 6.1 solves

$$(6.6) \quad \begin{aligned} -\nabla \cdot (a(y(\eta), x) \nabla \Psi(\eta, x)) &= f(x) \quad \forall x \in D, \\ \Psi(\eta, x) &= 0 \quad \forall x \in \partial D. \end{aligned}$$

Following [24], we use the Legendre polynomials to prove approximation estimates for the $p \times h$ -version of the SGFEM. Since the Legendre polynomials

$$p_n(t) \equiv \frac{1}{2^n n!} \frac{d^n}{dt^n} ((t^2 - 1)^n), \quad n = 0, 1, \dots,$$

are orthogonal with respect to the $L^2(-1, 1)$ inner product we have the error representation

$$(6.7) \quad (E_1)^2(\hat{y}_1) = \frac{|\Gamma_1|}{2} \sum_{n=p_1+1}^{+\infty} \frac{2}{2n+1} \|d_n\|_{H_0^1(D)}^2$$

with the corresponding Fourier coefficients

$$d_n \equiv \frac{2n+1}{2} \int_{-1}^1 \Psi(t) p_n(t) dt \in H_0^1(D).$$

Therefore, to obtain an estimate for E_1 we shall study the convergence of the tail series in (6.7).

NOTATION 6.1. *For each $\hat{y}_1 \in \hat{\Gamma}_1$, consider the natural extension of the variable t to the complex η and introduce the real function, $\mathcal{A} : \mathbb{C} \rightarrow \mathbb{R}$:*

$$\begin{aligned} \mathcal{A}(\eta) &\equiv \min_{x \in D} \operatorname{Re}\{a(y_1(\eta), \hat{y}_1, x)\} \\ &= \min_{x \in D} [a(0, \hat{y}_1, x) + y_1(\operatorname{Re}\{\eta\}) \sqrt{\lambda_1} b_1(x)], \end{aligned}$$

with $\operatorname{Re}\{\eta\}$ being the real part of $\eta \in \mathbb{C}$. Whenever $\mathcal{A}(\eta) \neq 0$, the extended function Ψ , the solution of (6.6), satisfies the bound

$$(6.8) \quad \|\Psi(\eta)\|_{H_0^1(D)} \leq C_D \frac{\|f(y_1(\eta), \hat{y}_1, \cdot)\|_{L^2(D)}}{\mathcal{A}(\eta)},$$

with C_D being the Poincaré constant for the domain D . Besides this, observe that

$$(6.9) \quad \mathcal{A}(\eta) \geq \tilde{a}_1(\hat{y}_1) - |y_1(\operatorname{Re}\{\eta\})| \sqrt{\lambda_1} \|b_1\|_{L^\infty(D)}.$$

We are now ready to estimate the Fourier coefficients in (6.7).

LEMMA 6.2. *Let $\tau \in (0, 1)$. Under Assumption 6.1 there exists a positive constant $\theta_f(\hat{y}_1, \tau) > 0$ such that*

$$\|d_n\|_{H_0^1(D)} \leq \frac{C_D \theta_f (2n+1)}{\tau \nu 2^n} \int_{-1}^1 \left(\frac{1-t^2}{t+1+\delta} \right)^n dt,$$

with $0 < \delta = \frac{2(1-\tau)\nu}{|\Gamma_1| \sqrt{\lambda_1} \|b_1\|_{L^\infty(D)}}$.

Proof. Consider

$$d_n = \frac{2n+1}{2} \int_{-1}^1 \Psi(t) p_n(t) dt = \frac{(2n+1)(-1)^n}{n! 2^{n+1}} \int_{-1}^1 \frac{d^n}{dt^n} \Psi(t) (1-t^2)^n dt.$$

Use the analytic continuation of the real function Ψ to the complex domain as in Lemma 6.1. The application of Cauchy's formula gives

$$\frac{d^n}{dt^n} \Psi(t) = \frac{n!(-1)^n}{2\pi i} \int_{\gamma_t} \frac{\Psi(\eta)}{(\eta-t)^{n+1}} d\eta,$$

where γ_t is a positively oriented closed circumference with the center at the real point $t \in (-1, 1)$, with radius $R(t)$, and such that all singularities from Ψ are exterior to γ_t . Estimate (6.8) implies

$$\begin{aligned}
 \left\| \frac{d^n}{dt^n} \Psi(t) \right\|_{H_0^1(D)} &\leq \frac{C_D}{2\pi} \frac{n!}{\mathcal{A}(\eta)} \int_{\gamma_t} \frac{\|f(y_1(\eta), \hat{y}_1, \cdot)\|_{L^2(D)}}{\mathcal{A}(\eta)|\eta - t|^{n+1}} |d\eta| \\
 (6.10) \quad &\leq \frac{C_D}{2\pi} \frac{n!}{\mathcal{A}(\eta)} \left(\sup_{\eta \in \gamma_t} \|f(y_1(\eta), \hat{y}_1, \cdot)\|_{L^2(D)} \right) \int_{\gamma_t} \frac{|d\eta|}{\mathcal{A}(\eta)|\eta - t|^{n+1}} \\
 &\leq \frac{C_D}{(R(t))^n} \left(\sup_{\eta \in \gamma_t} \|f(y_1(\eta), \hat{y}_1, \cdot)\|_{L^2(D)} \right) \sup_{\eta \in \gamma_t} \frac{1}{\mathcal{A}(\eta)}.
 \end{aligned}$$

Let

$$\theta_f \equiv \sup_{t \in [-1, 1]} \sup_{\eta \in \gamma_t} \|f(y_1(\eta), \hat{y}_1, \cdot)\|_{L^2(D)};$$

then estimate (6.10) implies

$$(6.11) \quad \|d_n\|_{H_0^1(D)} \leq \frac{(2n+1)C_D\theta_f}{2^{n+1}} \int_{-1}^1 \left(\frac{1}{\inf_{\eta \in \gamma_t} \mathcal{A}(\eta)} \right) \left(\frac{1-t^2}{R(t)} \right)^n dt.$$

Let $\tau \in (0, 1)$. We want to choose $R(t)$ such that

$$(6.12) \quad \inf_{t \in [-1, 1]} \inf_{\eta \in \gamma_t} \mathcal{A}(\eta) \geq \tau \nu.$$

Since Assumption 6.1 holds, then (6.12) is satisfied taking $R(t) = 1 - |t| + \delta$, with $\delta = \frac{2(1-\tau)\nu}{|\Gamma_1| \sqrt{\lambda_1} \|b_1\|_{L^\infty(D)}}$. Finally, the proof concludes by combining (6.11)–(6.12) and the definition of $R(t)$. \square

Now we use a result from [24], namely, that we have the following.

LEMMA 6.3 (integral estimate). *Let $\xi < -1$, and define*

$$r \equiv \frac{1}{|\xi| + \sqrt{\xi^2 - 1}}, \quad 0 < r < 1.$$

Then there holds

$$(-1)^n \int_{-1}^1 \left(\frac{t^2 - 1}{t + \xi} \right)^n dt = (2r)^n 2^{n+1} \frac{n!}{(2n+1)!!} \Phi_{n,0}(r^2),$$

where $\Phi_{n,0}(r^2)$ is the Gauss hypergeometric function. Moreover, we have

$$\Phi_{n,0}(r^2) = \sqrt{1 - r^2} + \mathcal{O}\left(\frac{1}{n^{1/3}}\right)$$

uniformly with respect to $0 < r < 1$.

Finally, we can state the estimate for the size of the series in (6.7).

LEMMA 6.4. *Let $\tau \in (0, 1)$. Under Assumption 6.1 there exists a positive constant $\theta_f > 0$ such that*

$$(E_1)(\hat{y}_1) \leq \frac{C_D \theta_f \sqrt{|\Gamma_1|}}{\tau \tilde{a}_1} \left(\sqrt{1 - r^2} + \mathcal{O}\left(\frac{1}{(p_1)^{1/3}}\right) \right) \sqrt{\pi} \frac{r^{p_1+1}}{\sqrt{1 - r^2}},$$

with $r \equiv \frac{1}{|\xi| + \sqrt{\xi^2 - 1}}, 0 < r < 1$, and $\xi < -(1 + \frac{2(1-\tau)\nu}{|\Gamma_1| \sqrt{\lambda_1} \|b_1\|_{L^\infty(D)}})$.

Proof. Use Lemmas 6.2 and 6.3, together with the asymptotic equivalence $\frac{(2n)!!}{(2n-1)!!} \sim \sqrt{\frac{\pi n}{2}}, n \rightarrow \infty$, yielding

$$\|d_n\|_{H_0^1(D)} \leq \frac{2C_D \theta_f}{\tau \tilde{a}_1} \sqrt{\frac{\pi n}{2}} \left(\sqrt{1 - r^2} + \mathcal{O}\left(\frac{1}{n^{1/3}}\right) \right) r^n.$$

Then use the result to estimate the tail of the series:

$$(E_1)^2(\hat{y}_1) = \frac{|\Gamma_1|}{2} \sum_{n=p_1+1}^{+\infty} \frac{2}{2n+1} \|d_n\|_{H_0^1(D)}^2. \quad \square$$

The main result of this section, namely the exponential convergence with respect to the multi-index p as in [24], follows from the above lemmas; i.e., we have the following.

THEOREM 6.2. *Let $\tau \in (0, 1)$ and u be the solution of (2.6), $u \in L^2(\Gamma) \otimes H_0^1(D)$, which is analytic with respect to y , onto the subspace $Z^p \otimes H_0^1(D)$. Under Assumption 6.1 there exist positive constants, $0 < C, C_f$, such that*

$$(6.13) \quad \begin{aligned} E_p &\equiv \min_{v \in Z^p \otimes H_0^1(D)} \|u - v\|_{L^2(\Gamma) \otimes H_0^1(D)} \\ &\leq \frac{C_D C_f}{\tau \nu} \sqrt{\pi |\Gamma|} \sum_{n=1}^N \left(1 + \frac{1}{\sqrt{1 - r_n^2}} \mathcal{O}\left(\frac{1}{(p_n)^{1/3}}\right) \right) (r_n)^{p_n+1}, \end{aligned}$$

with $0 < r_n \equiv \frac{1}{|\xi_n| + \sqrt{\xi_n^2 - 1}} < 1$, and $\xi_n < -(1 + \frac{2(1-\tau)\nu}{|\Gamma_n| \sqrt{\lambda_n} \|b_n\|_{L^\infty(D)}})$ for $n = 1, \dots, N$.

Similarly, as in the $k \times h$ -version (cf. (5.7)), the p -version has a convergence result for the approximation to the expected value of the solution.

THEOREM 6.3. *With the same assumptions as in Theorem 6.2 and for $\ell = 0, 1$, we have*

$$\|E[u - u_h^\ell]\|_{H^\ell(D)} \leq C \left(h^{2-\ell} + \frac{1}{\tau} \sum_{n=1}^N (r_n)^{(2-\ell)(p_n+1)} \right),$$

with $0 < r_n < 1$ as in Theorem 6.2 and $C > 0$ independent of h , p_n , and r_n .

The proof of the previous theorem uses Theorem 6.2 and is completely similar to the proof of Theorem 5.1.

Remark 6.2. Whenever the coefficients a and f are independent the constants θ_f from Lemma 6.2, C_f from Theorem 6.2, and C from Theorem 6.3 do not depend on N .

7. Double orthogonal polynomials. Here we explain the idea of double orthogonal polynomials, used by various authors in different contexts; see, e.g., [47] to compute efficiently the solution of the $k \times h$ -version and the $p \times h$ -version studied in sections 5 and 6, respectively. The idea is to use a special basis to decouple the system in the y -direction, yielding just a number of uncoupled systems, each one with the size and structure of one Monte Carlo realization of (4.1). The double orthogonal polynomials are able to perform the decoupling whenever the random variables in the Karhunen–Loëve expansion of $a, Y_n, n = 1, \dots, N$, are independent.

Without loss of generality we focus on the p -version, i.e., find $u_h^p \in Z^p \otimes X_h^d$ such that

$$(7.1) \quad \begin{aligned} & \int_{\Gamma} \rho(y) (a(y, \cdot) \nabla u_h^p(y, \cdot), \nabla v(y, \cdot))_{L^2(D)} dy \\ &= \int_{\Gamma} \rho(y) (f(y, \cdot), v(y, \cdot))_{L^2(D)} dy \quad \forall v \in Z^p \otimes X_h^d. \end{aligned}$$

Let $\{\psi_j(y)\}$ be a basis of the subspace $Z^p \subset L^2(\Gamma)$ and $\{\varphi_i(x)\}$ be a basis of the subspace $X_h^d \subset H_0^1(D)$. Write the approximate solution as

$$(7.2) \quad u_h^p(y, x) = \sum_{j,i} u_{ij} \psi_j(y) \varphi_i(x)$$

and use test functions $v(y, x) = \psi_k(y) \varphi_\ell(x)$ to find the coefficients u_{ij} . Then (7.1) gives

$$\begin{aligned} & \sum_{j,i} \left(\int_{\Gamma} \rho(y) \psi_k(y) \psi_j(y) (a(y, \cdot) \nabla \varphi_i, \nabla \varphi_\ell)_{L^2(D)} dy \right) u_{ij} \\ &= \int_{\Gamma} \rho(y) \psi_k(y) (f(y, \cdot), \varphi_\ell)_{L^2(D)} dy \quad \forall k, \ell, \end{aligned}$$

which can be rewritten as

$$\sum_{j,i} \left(\int_{\Gamma} \rho(y) \psi_k(y) \psi_j(y) K_{i,\ell}(y) dy \right) u_{ij} = \int_{\Gamma} \rho(y) \psi_k(y) f_\ell(y) dy \quad \forall k, \ell,$$

with $K_{i,\ell}(y) \equiv (a(y, \cdot) \nabla \varphi_i, \nabla \varphi_\ell)_{L^2(D)}$ and $f_\ell(y) \equiv (f(y, \cdot), \varphi_\ell)_{L^2(D)}$. If the diffusion coefficient, a , is a truncated Karhunen–Loèvre expansion, $a(y, x) = E[a](x) + \sum_{n=1}^N b_n(x) y_n$, and by the independence of the $Y_n, n = 1, \dots, N$, we have a corresponding ‘‘Karhunen–Loèvre’’ expression for the stiffness matrix $K_{i,\ell}(y) \equiv \int_D (E[a](x) + \sum_{n=1}^N b_n(x) y_n) \nabla \varphi_i(x) \cdot \nabla \varphi_\ell(x) dx = K_{i,\ell}^0 + \sum_{n=1}^N y_n K_{i,\ell}^n$ with deterministic coefficients $K_{i,\ell}^0 \equiv (E[a] \nabla \varphi_i, \nabla \varphi_\ell)_{L^2(D)}$ and $K_{i,\ell}^n \equiv (b_n \nabla \varphi_i, \nabla \varphi_\ell)_{L^2(D)}$. By the same token we have

$$\begin{aligned} \int_{\Gamma} \rho(y) \psi_k(y) \psi_j(y) K_{i,\ell}(y) dy &= K_{i,\ell}^0 \int_{\Gamma} \rho(y) \psi_k(y) \psi_j(y) dy \\ &+ \sum_{n=1}^N K_{i,\ell}^n \int_{\Gamma} y_n \rho(y) \psi_k(y) \psi_j(y) dy. \end{aligned}$$

Since $\psi_k \in Z^p$, with multi-index $p = (p_1, \dots, p_N)$, it is enough to take it as the product $\psi_k(y) = \prod_{r=1}^N \psi_{kr}(y_r)$, where $\psi_{kr}: \Gamma_r \rightarrow \mathbb{R}$ is a basis function of the subspace

$$Z^{p_r} = \text{span}[1, y, \dots, y^{p_r}] = \text{span}[\psi_{hr}: h = 1, \dots, p_r + 1].$$

Keeping this choice of ψ_k in mind,

$$\begin{aligned} \int_{\Gamma} \rho(y) \psi_k(y) \psi_j(y) K_{i,\ell}(y) dy &= K_{i,\ell}^0 \int_{\Gamma} \prod_{m=1}^N \rho_m(y_m) \psi_{km}(y_m) \psi_{jm}(y_m) dy \\ &+ \sum_{n=1}^N K_{i,\ell}^n \int_{\Gamma} y_n \prod_{m=1}^N \rho_m(y_m) \psi_{km}(y_m) \psi_{jm}(y_m) dy. \end{aligned}$$

Now, for every set Γ_n , $n = 1, \dots, N$, choose the polynomials, $\psi_j(y) = \prod_{n=1}^N \psi_{jn}(y_n)$, to be biorthogonal; i.e., for $n = 1, \dots, N$ they must satisfy

$$(7.3) \quad \begin{aligned} \int_{\Gamma_n} \rho_n(z) \psi_{kn}(z) \psi_{jn}(z) dz &= \delta_{kj}, \\ \int_{\Gamma_n} z \rho_n(z) \psi_{kn}(z) \psi_{jn}(z) dz &= c_{kn} \delta_{kj}. \end{aligned}$$

To find the polynomials ψ_k we have to solve N eigenproblems, each of them with size $(1 + p_n)$. The computational work required by these eigenproblems is negligible with respect to the one required to solve for u_{ij} ; cf. [23, section 8.7.2]. The orthogonality properties (7.3) for ψ_k imply the decoupling $\int_{\Gamma} \rho \psi_k \psi_j dy = \delta_{kj}$, $\int_{\Gamma} y_n \rho \psi_k \psi_j dy = c_{kn} \delta_{kj}$. By means of this decoupling we now conclude that

$$\begin{aligned} &\sum_{j,i} \left(\int_{\Gamma} \rho(y) \psi_k(y) \psi_j(y) K_{i,\ell}(y) dy \right) \\ &= K_{i,\ell}^0 \int_{\Gamma} \rho(y) \psi_k(y) \psi_j(y) dy + \sum_{n=1}^N K_{i,\ell}^n \int_{\Gamma} y_n \rho(y) \psi_k(y) \psi_j(y) dy \\ &= \left(K_{i,\ell}^0 + \sum_{n=1}^N c_{kn} K_{i,\ell}^n \right) \delta_{kj}. \end{aligned}$$

The structure of the linear system that determines u_{ij} now becomes block diagonal, with each block being coercive and with the sparsity structure identical to one deterministic FEM stiffness matrix, i.e.,

$$\begin{bmatrix} \left(K^0 + \sum_{n=1}^N c_{1n} K^n \right) & 0 & \dots & 0 \\ 0 & \left(K^0 + \sum_{n=1}^N c_{2n} K^n \right) & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \left(K^0 + \sum_{n=1}^N c_{Nn} K^n \right) \end{bmatrix}.$$

Observe that as a consequence of the uniform coercivity assumption, each of the diagonal blocks in the system above is symmetric and strictly positive definite. The conclusion is that the computational work to find the coefficients u_{ij} in (7.2) is the same as the one needed to compute $\prod_{i=1}^N (1 + p_i)$ Monte Carlo realizations of u_h defined in (4.1), but the accuracies of these methods may differ. Section 8 studies this issue.

8. Asymptotical efficiency comparisons. In this section we compare the asymptotical numerical complexity for the Monte Carlo Galerkin finite element method (cf. section 4) with the stochastic Galerkin finite element method introduced in sections 5 and 6. The quantity of interest, i.e., the goal of the computation, is the expected value of the solution, $E[u]$, and its approximation is studied in both the $L^2(D)$ and the $H_0^1(D)$ sense. In all the cases, the spatial discretization is done by piecewise linear finite elements on globally quasi-uniform meshes. For the $k \times h$ -SGFEM the Γ partitions are also assumed to be globally quasi-uniform. Besides this, the diffusion function a is assumed to be a truncated Karhunen–Loève expansion with independent random variables Y_n , $n = 1, \dots, N$.

8.1. MCGFEM versus $k \times h$ -SGFEM. Here we consider the computational work to achieve a given accuracy bounded by a positive constant TOL for both the MCGFEM and the $k \times h$ -SGFEM methods. This optimal computational work indicates under which circumstances one method may be best suited. When using the MCGFEM method to approximate the solution of (1.1) in $H_0^1(D)$ the error becomes, applying Proposition 4.2 together with (4.5), that given a confidence level, $0 < c_0 < 1$, there exists a constant $C > 0$ depending only on c_0 such that

$$(8.1) \quad P\left(\left\|E[u] - \frac{1}{M} \sum_{j=1}^M u_h(\cdot; \omega_j)\right\|_{H_0^1(D)} \leq C \left(h + \frac{1}{\sqrt{M}}\right)\right) \geq c_0.$$

Then in the sense of (8.1) we write $E_{MCGFEM} = O(h) + O(1/\sqrt{M})$. The corresponding computational work for the MCGFEM method is $Work_{MCGFEM} = O((1/h^d)^r + 1/h^d)M$, where the parameter $1 \leq r \leq 3$ relates to the computational effort devoted to solve one linear system with n unknowns, $O(n^r)$. From now on we continue the discussion with the optimal $r = 1$ that can be achieved by means of the multigrid method; cf. [10]. Thus, choosing h and M to minimize the computational work for a given desired level of accuracy $TOL > 0$ yields the optimal work

$$(8.2) \quad Work_{MCGFEM}^* = O(TOL^{-(2+d)}).$$

On the other hand, if we apply a $k \times h$ -SGFEM with piecewise polynomials of order q in the y -direction, the computational error in $H_0^1(D)$ norm is (cf. Theorem 5.1)

$$E_{SGFEM} = O(h) + O(k^{q+1}),$$

and the corresponding computational work for the $k \times h$ -version is

$$Work_{SGFEM} = O(h^{-d}(1+q)^N k^{-N}).$$

Here N is the number of terms in the truncated Karhunen–Loève expansion of the coefficients a and f and k is the discretization parameter in the y -direction. Similarly as before, we can compute the optimal work for the $k \times h$ -SGFEM method, yielding

$$Work_{SGFEM}^* = O((1+q)^N TOL^{-\frac{N}{q+1}} TOL^{-d}).$$

Therefore, a $k \times h$ -version SGFEM is likely to be preferred whenever TOL is sufficiently small and $N/2 < 1 + q$; i.e., if the number of terms in the Karhunen–Loève expansion of a is large, then the degree of approximation in the y -direction, q , has to become correspondingly large. We summarize the comparison results in Table 1, where we also include corresponding results from the $p \times h$ -version, to be derived in subsection 8.2. Similarly, if we are interested in controlling the difference $\|E[u] - E[u_h]\|_{L^2(D)}$, the application of (4.2) and Proposition 4.2 for the MCGFEM method and Theorem 5.1 on the convergence of the $k \times h$ -SGFEM method imply the results shown in Table 2. In this case $k \times h$ -SGFEM is likely to be preferred whenever $N/4 < (q+1)$ and TOL is sufficiently small. In addition, the comparison tells us that to be able to be competitive with the Monte Carlo method when the number of relevant terms in the Karhunen–Loève expansion is not so small, an optimal method should have a high order of approximation and should avoid as much as possible the coupling between the different components of the numerical solution to preserve computational efficiency. The approach proposed by Ghanem and Spanos [22] based on

TABLE 1

Approximation of the function $E[u]$ in $H_0^1(D)$. Asymptotical numerical complexity for the MCGFEM and SGFEM methods.

	MCGFEM	$k \times h$ -version SGFEM	$p \times h$ -version SGFEM
Work	M/h^d	$\frac{(1+q)^N}{h^d k^N}$	$\frac{(1+p)^N}{h^d}$
$H_0^1(D)$ Error	$h + \frac{1}{\sqrt{M}}$	$h + k^{q+1}$	$h + r^{(p+1)}$
$H_0^1(D)$ Work*	$\text{TOL}^{-(2+d)}$	$\text{TOL}^{-\frac{N}{q+1}} \text{TOL}^{-d}$	$(\log_r(\text{TOL}))^N \text{TOL}^{-d}$

TABLE 2

Approximation of the function $E[u]$ in $L^2(D)$. Asymptotical numerical complexity for the MCGFEM and SGFEM methods.

	MCGFEM	$k \times h$ -version SGFEM	$p \times h$ -version SGFEM
Work	M/h^d	$\frac{(1+q)^N}{h^d k^N}$	$\frac{(1+p)^N}{h^d}$
$L^2(D)$ Error	$h^2 + \frac{1}{\sqrt{M}}$	$h^2 + k^{2(q+1)}$	$h^2 + r^{2(p+1)}$
$L^2(D)$ Work*	$\text{TOL}^{-(2+d/2)}$	$(\text{TOL})^{-\frac{N}{2(q+1)}} \text{TOL}^{-d/2}$	$(\log_r(\text{TOL}))^N \text{TOL}^{-d/2}$

orthogonal polynomials has, whenever the approximate diffusion satisfies (1.2), a high order of approximation but introduces coupling between the different components of the numerical solution. The uncoupling can be achieved for linear equations using double orthogonal polynomials; see the description in section 7. With this motivation, section 6 studies the convergence of the $p \times h$ -SGFEM.

8.2. MCGFEM versus $p \times h$ -SGFEM. Here we consider the computational work to achieve a given accuracy for both the $p \times h$ -version of SGFEM defined in (6.1) and the MCGFEM method for the approximation of $E[u]$ defined in section 4; i.e., we are interested in controlling the difference $\|E[u] - E[u_h^p]\|_{L^2(D)}$ or $\|E[u] - \frac{1}{M} \sum_{j=1}^M u_h(\cdot; \omega_j)\|_{L^2(D)}$, respectively. This computational work indicates under which circumstances one method may be better suited than the other. Besides this, let us assume that we use in our computations a piecewise linear FEM space in D . When using the MCGFEM method to approximate the expected value of the solution of (1.1), we have the optimal work required to achieve a given desired level of accuracy $\text{TOL} > 0$ (cf. (8.2)):

$$\text{Work}_{\text{MCGFEM}}^* = O(1/\text{TOL}^{2+\frac{d}{2}}).$$

On the other hand, if we apply a $p \times h$ -version of the SGFEM, with $p_i = p$, $i = 1, \dots, N$, the computational error is (cf. Theorem 6.3)

$$E_{\text{SGFEM}} = O(h^2) + O(r^{2(p+1)}), \quad 0 < r < 1,$$

and the corresponding computational work is (cf. section 7)

$$\text{Work}_{\text{SGFEM}} = O\left(\frac{(1+p)^N}{h^d}\right).$$

Recall that N is the number of terms in the truncated Karhunen–Loève expansion of the coefficients a and f , and k is the discretization parameter in the y -direction. As

before, we can compute the optimal work for the SGFEM method, yielding

$$Work_{SGFEM}^* \leq O((\log_r(\text{TOL}))^N \text{TOL}^{-\frac{d}{2}})$$

and the asymptotical comparison

$$\lim_{\text{TOL} \rightarrow 0} \frac{Work_{SGFEM}^*}{Work_{MCGFEM}^*} = \lim_{\text{TOL} \rightarrow 0} (\log_r(\text{TOL}))^N \text{TOL}^2 = 0.$$

Therefore, for sufficiently strict accuracy requirements, i.e., sufficiently small TOL, in the computation of $E[u]$, SGFEM requires less computational effort than MCGFEM. The work of Bahvalov and its subsequent extensions (cf. [7, 25, 45, 35]) generalize the standard Monte Carlo method, taking advantage of the available integrand's smoothness and yielding a faster order of convergence. The optimal work of such a method is for our case, i.e., the approximation of $E[u]$ in $L^2(D)$, $\mathcal{O}(C(N)\text{TOL}^{-\frac{1}{1/2+q/N}} \text{TOL}^{-d/2})$, where it is assumed that the integrand u has bounded derivatives up to order q with respect to y and the integral is done in the N -dimensional unit cube.

The result on the computational work of the $p \times h$ -version of the SGFEM presented in this work is then related to the case $q = \infty$, since u is analytic with respect to y . This analyticity allows the exponential convergence with respect to p ; cf. Theorem 6.3.

Notice that we discussed only the optimal asymptotical computational work required by both methods, but in practice the constants involved in the asymptotic approximations make these comparisons just indicative and not conclusive. In addition, we have studied only the case where the integrals $\int_{\Gamma_i} \rho_i y^k dy$ can be computed exactly for $k = 0, 1, \dots$, and not considered the more general case where quadrature rules are needed to approximate such integrals. For further information we refer the reader to [6].

Acknowledgment. The authors would like to thank Prof. Anders Szepessy for many valuable discussions regarding this work.

REFERENCES

- [1] I. BABUŠKA, *On randomized solutions of Laplace's equation*, Casopis Pest. Mat., 86 (1961), pp. 269–276.
- [2] I. BABUŠKA, B. ANDERSSON, P. J. SMITH, AND K. LEVIN, *Damage analysis of fiber composites. I. Statistical analysis on fiber scale*, Comput. Methods Appl. Mech. Engrg., 172 (1999), pp. 1–4.
- [3] I. BABUŠKA AND P. CHATZIPANTELIDIS, *On solving linear elliptic stochastic partial differential equations*, Comput. Methods Appl. Mech. Engrg., 191 (2002), pp. 4093–4122.
- [4] I. BABUŠKA AND J. CHLEBOUN, *Effects of uncertainties in the domain on the solution of Neumann boundary value problems in two spatial dimensions*, Math. Comp., 71 (2002), pp. 1339–1370.
- [5] I. BABUŠKA AND J. CHLEBOUN, *Effects of uncertainties in the domain on the solution of Dirichlet boundary value problems*, Numer. Math., 93 (2003), pp. 583–610.
- [6] I. BABUŠKA, R. TEMPONE, AND G. ZOURARIS, *Galerkin Finite Element Methods for Stochastic Elliptic Partial Differential Equations*, Technical report 02-38, TICAM, Austin, TX, 2002.
- [7] N. S. BAHVALOV, *Approximate computation of multiple integrals*, Vestnik Moskov. Univ. Ser. Mat. Meh. Astr. Fiz. Him., 1959 (1959), pp. 3–18.
- [8] Y. BEN-HAIM AND I. ELISHAKOFF, *Convex Models of Uncertainty in Applied Mechanics*, Elsevier, Amsterdam, 1990.
- [9] F. E. BENTH AND J. GJERDE, *Convergence rates for finite element approximations of stochastic partial differential equations*, Stochastics Stochastics Rep., 63 (1998), pp. 313–326.

- [10] J. H. BRAMBLE, *Multigrid Methods*, Longman Scientific and Technical, Harlow, UK, 1993.
- [11] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 1994.
- [12] R. E. CAFLISCH, *Monte Carlo and Quasi-Monte Carlo Methods*, Acta Numer. 7, Cambridge University Press, Cambridge, UK, 1998.
- [13] P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, North-Holland, New York, 1978.
- [14] M.-K. DEB, *Solution of Stochastic Partial Differential Equations (SPDEs) Using Galerkin Method: Theory and Applications*, Ph.D. thesis, University of Texas at Austin, Austin, TX, 2000.
- [15] 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.
- [16] R. DURRETT, *Probability: Theory and Examples*, 2nd ed., Duxbury Press, Belmont, CA, 1964.
- [17] I. ELISHAKOFF, ED., *Whys and Hows in Uncertainty Modelling*, Springer-Verlag, Vienna, 1999.
- [18] I. ELISHAKOFF, Y. K. LIN, AND L. P. ZHU, *Probabilistic and Convex Modelling of Acoustically Excited Structures*, Elsevier, Amsterdam, 1994.
- [19] I. ELISHAKOFF AND Y. REN, *The bird's eye view on finite element method for structures with large stochastic variations*, Comput. Methods Appl. Mech. Engrg., 168 (1999), pp. 51–61.
- [20] L. EVANS, *Partial Differential Equations*, Grad. Stud. Math. 19, AMS, Providence, RI, 1998.
- [21] R. GHANEM AND J. RED-HORSE, *Propagation of probabilistic uncertainty in complex physical systems using a stochastic finite element approach*, Phys. D, 133 (1999), pp. 137–144.
- [22] R. G. GHANEM AND P. D. SPANOS, *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, New York, 1991.
- [23] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, MD, 1996.
- [24] W. GUI AND I. BABUŠKA, *The h , p and $h\cdot p$ versions of the finite element method in 1 dimension. I. The error analysis of the p -version*, Numer. Math., 49 (1986), pp. 577–612.
- [25] S. HABER, *Stochastic quadrature formulas*, Math. Comp., 23 (1969), pp. 751–764.
- [26] E. HAUGEN, *Probabilistic Mechanical Design*, Wiley, New York, 1980.
- [27] H. HOLDEN, B. ØKSENDAL, J. UBØE, AND T. ZHANG, *Stochastic Partial Differential Equations*, Birkhäuser Boston, Boston, MA, 1996.
- [28] J. L. JENSEN, L. W. LAKE, P. W. CORBETT, AND D. LOGAN, *Statistics for Petroleum Engineers and Geoscientists*, Prentice-Hall, Upper Saddle River, NJ, 1997.
- [29] E. JOUINI, J. CVITANIĆ, AND M. MUSIELA, EDs., *Option Pricing, Interest Rates and Risk Management*, Cambridge University Press, Cambridge, UK, 2001.
- [30] M. KLEIBER AND T.-D. HIEN, *The Stochastic Finite Element Method*, Wiley, Chichester, 1992.
- [31] J. E. LAGNESE, *General boundary value problems for differential equations of Sobolev type*, SIAM J. Math. Anal., 3 (1972), pp. 105–119.
- [32] S. LARSEN, *Numerical Analysis of Elliptic Partial Differential Equations with Stochastic Input Data*, Ph.D. thesis, University of Maryland, College Park, MD, 1986.
- [33] P. LÉVY, *Processus Stochastiques et Mouvement Brownien*, 10th ed., Éditions Jacques Gabay, Paris, 1992.
- [34] G. MAYMON, *Some Engineering Applications in Random Vibrations and Random Structures*, Progress in Astronautics and Aeronautics 178, P. Zarchan, ed., American Institute of Aeronautics and Astronautics, Reston, VA, 1998.
- [35] E. NOVAK, *Stochastic properties of quadrature formulas*, Numer. Math., 53 (1988), pp. 609–620.
- [36] B. ØKSENDAL, *Stochastic Differential Equations. An Introduction with Applications*, 5th ed., Springer-Verlag, Berlin, 1998.
- [37] F. RIESZ AND B. SZ.-NAGY, *Functional Analysis*, Dover, New York, 1990.
- [38] P. J. ROACHE, *Verification and Validation in Computational Science and Engineering*, Hermosa, Albuquerque, NM, 1998.
- [39] J. B. ROBERTS AND P.-T. D. SPANOS, *Random Vibration and Statistical Linearization*, Wiley, Chichester, 1990.
- [40] C. SCHWAB AND R.-A. TODOR, *Sparse finite elements for elliptic problems with stochastic loading*, Numer. Math., 95 (2003), pp. 707–734.
- [41] I. M. SOBOL', *A Primer for the Monte Carlo Method*, CRC Press, Boca Raton, FL, 1994.
- [42] I. M. SOBOL', *On quasi-Monte Carlo integrations*, Math. Comput. Simulation, 47 (1998), pp. 103–112.
- [43] J. SÓLNES, *Stochastic Processes and Random Vibrations*, Wiley, New York, 1997.
- [44] T. G. THETING, *Solving Wick-stochastic boundary value problems using a finite element*

method, Stochastics Stochastics Rep., 70 (2000), pp. 241–270.

[45] J. F. TRAUB AND A. G. WERSCHULZ, *Complexity and Information*, Cambridge University Press, Cambridge, UK, 1998.

[46] G. VÅGE, *Variational methods for PDEs applied to stochastic partial differential equations*, Math. Scand., 82 (1998), pp. 113–137.

[47] T. WERDER, K. GERDES, D. SCHÖTZAU, AND C. SCHWAB, *hp-discontinuous Galerkin time stepping for parabolic problems*, Comput. Methods Appl. Mech. Engrg., 190 (2001), pp. 49–50.