A stochastic collocation method for elliptic partial differential equations with random input data

I. Babuška [a], F. Nobile[b], R. Tempone[c]

[a] ICES, The University of Texas at Austin, USA

[b] MOX, Dipartimento di Matematica, Politecnico di Milano, ITALY

[c] School of Computational Sciences and Department of Mathematics, Florida State University at Tallahassee, USA

Abstract

In this paper we propose and analyze a Stochastic-Collocation method to solve elliptic Partial Differential Equations with random coefficients and forcing terms (input data of the model). The input data are assumed to depend on a finite number of random variables. The method consists in a Galerkin approximation in space and a collocation in the zeros of suitable tensor product orthogonal polynomials (Gauss points) in the probability space and naturally leads to the solution of uncoupled deterministic problems as in the Monte Carlo approach. It can be seen as a generalization of the Stochastic Galerkin method proposed in [Babuška -Tempone-Zouraris, SIAM J. Num. Anal. 42(2004)] and allows one to treat easily a wider range of situations, such as: input data that depend non-linearly on the random variables, diffusivity coefficients with unbounded second moments , random variables that are correlated or have unbounded support. We provide a rigorous convergence analysis and demonstrate exponential convergence of the "probability error" with respect of the number of Gauss points in each direction in the probability space, under some regularity assumptions on the random input data. Numerical examples show the effectiveness of the method.

Key words: Collocation method, stochastic PDEs, finite elements, uncertainty quantification, exponential convergence.

AMS subject classification: 65N35, 65N15, 65C20

Introduction

Thanks to the fast growing computer power, numerical simulations are used every day more to produce predictions of the behavior of complex physical or engineering systems. Some sources of errors arising in computer simulations can be controlled and reduced, by now, using sophisticated techniques such as

a posteriori error estimation [1, 3, 37], mesh adaptivity and the more recent *modeling error* analysis [30, 31, 10]. All this has increased the accuracy of numerical predictions as well as our confidence in them.

Yet, many engineering applications are affected by a relatively large amount of uncertainty in the input data such as model coefficients, forcing terms, boundary conditions, geometry, etc. In this case, to obtain a reliable numerical prediction, one has to include uncertainty quantification due to the uncertainty in the input data.

Uncertainty can be described in several ways, depending on the amount of information available; among others we mention: the worst case scenario analysis and fuzzy set theory, evidence theory, probabilistic setting, etc (see [6, 24] and the references therein).In this paper we focus on elliptic partial differential equations with a probabilistic description of the uncertainty in the input data. The model problem has the form

$$\mathcal{L}(a)u = f \quad \text{in } D \tag{1}$$

where $\mathcal{L}$ is an elliptic operator in a domain $D \subset \mathbb{R}^d$, which depends on some coefficients $a(x, \omega)$, with $x \in D$, $\omega \in \Omega$, and Ω indicates the set of possible outcomes. Similarly, the forcing term $f = f(x, \omega)$ can be assumed random as well.

We will focus on the case where the probability space has a low dimensionality, that means, the stochastic problem depends only on a relatively small number of random variables.

This can be the case if, for instance, the mathematical model depends on few parameters, which can be taken as random variables with a given joint probability distribution. To make an example we might think at the deformation of an elastic homogeneous material in which the Young's modulus and the Poisson's ratio (parameters that characterize the material properties), are random variables, either independent or not.

In other situations, the input data may vary randomly from one point of the physical domain D to another and their uncertainty should rather be described in terms of random fields with a given covariance structure (i.e. each point of the domain is a random variable and the correlation between two distinct points in the domain is known and non zero, in general; this case is sometimes referred to as *colored noise*).

Examples of this situation are, for instance, the deformation of inhomogeneous materials such as wood, foams, or bio-materials such arteries, bones, etc.; groundwater flow problems where the permeability in each layer of sediments (rocks, sand, etc) should not be assumed constant; the action of wind (direction and point intensity) on structures; etc.

A possible way to describe such random fields consists in using a Karhunen-Loève [27] or a Polynomial Chaos (PC) expansion [38, 42]. The former represents the random field as a linear combination of an infinite number of uncorrelated random variables, while the latter uses polynomial expansions in terms of independent random variables. Both expansions exist provided the random field

$a : \Omega \to V$, as a mapping from the probability space into a functional space V, has bounded second moment. Other non-linear expansions can be considered as well (see e.g. [22] for a technique to express a stationary random field with given covariance structure and marginal distribution as a function of (infinite) independent random variables; non-linear transformations have been used also in [29, 39]). The use of non-polynomial expansions may be advantageous in some situations: for instance, in groundwater flow problems, the permeability coefficient within each layer of sediments can feature huge variability, which is often expressed in a logarithm scale. In this case, one might want to use a Karhunen-Loève (or Polynomial Chaos) expansion for the logarithm of the permeability, instead of the permeability field itself. This leads to an exponential dependence of the permeability on the random variables and the resulting random field might even have unbounded second moments. An advantage of such a non-linear expansion is that it guarantees a positive permeability almost surely (a condition which is difficult to enforce, instead, with a standard truncated Karhunen-Loève or PC expansion).

Although such random fields are properly described only by means of an infinite number of random variables, whenever the input data vary slowly in space, with a correlation length comparable to the size of the domain, only few terms in the above mentioned expansions are typically enough to describe the random field with sufficient accuracy. Therefore, for this type of applications it is reasonable to limit the analysis to just a few number of random variables in the expansion (see e.g. [2]).

This argument is also strengthened by the fact that the amount of measured data at one's disposal to identify the input data as random fields is in general very limited and barely sufficient to identify the first few random variables in the expansion.

Conversely, situations in which the random fields are highly oscillatory with a short correlation length, as in the case of materials with a random microstructure, do not fall in this category and will not be considered in the present work. The interested reader should refer, instead, to the wide literature in homogenization and multiscale analysis (see e.g. [16] and references therein).

To solve numerically the stochastic partial differential equation (1), a relatively new numerical technique, which has gained much attention in the last few years, is the so called *Spectral Galerkin* approximation (see e.g. [21]) . It employs standard approximations in space (finite elements, finite volumes, spectral or h-p finite elements, etc.) and polynomial approximation in the probability domain, either by full polynomial spaces [41, 29, 20], tensor product polynomial spaces [4, 18] or piecewise polynomial spaces [4, 26].

The use of tensor product spaces is particularly attractive in the case of a small number of random variables, since it allows naturally the use of anisotropic spaces where the polynomial degree is chosen differently in each direction in probability. Moreover, whenever the random fields are expanded in a truncated Karhunen-Loève expansion and the underlying random variables are assumed independent, a particular choice of the basis for the tensor product space (as proposed in [4, 5]), leads to the solution of uncoupled deterministic problems

as in a Monte Carlo simulation. In this case, exponential convergence of the "probability error" has been proved in [4].

On the other hand, tensor product spaces suffer from the so called *curse of dimensionality* since the dimension of the approximating space grows exponentially fast in the number of random variables. If the number of random variables is moderate or large, one should rather consider full polynomial spaces or sparse tensor product spaces [7, 18, 40]. We will not address this issue in this paper.

The extension of the Spectral Galerkin method to cases in which the input data depend non-linearly on the random variables and possibly have unbounded second moments is not straightforward and, in any case, would lead to fully coupled systems of equations, which demand for highly efficient parallel solvers.

In this work we propose a Collocation method which consists in collocating problem (1) in the zeros of tensor product orthogonal polynomials with respect to the joint probability density ρ of the random variables, should they be independent, or any other auxiliary density $\hat{\rho}$ corresponding to independent random variables, as long as the ratio $\rho/\hat{\rho}$ is bounded. Stochastic collocation has already been applied in a variety of problems and it is the subject of ongoing research, see among others [35, 28] and the recent work [40] which the authors became aware of upon completion of this work.

As it will be pointed out in the paper, this method offers several advantages:

- it naturally leads to uncoupled deterministic problems also in case of input data which depend non-linearly on the random variables;
- treats efficiently the case of non independent random variables with the introduction of the auxiliary density $\hat{\rho}$;
- can easily deal with random variables with unbounded support, such as Gaussian or exponential ones.
- deals with no difficulty with a diffusivity coefficient a with unbounded second moment.

The main result of the paper is given in Theorem 1, Section 4, where we prove that the Collocation method preserves the same accuracy as the Spectral Galerkin approach and achieves exponential convergence in all the above mentioned cases, provided the input data are infinitely differentiable with respect to the random variables, under very mild assumptions on the growth of their derivatives in the probability directions, as it is the case for standard expansions of random fields.

The Collocation method can also be seen as a Pseudo Spectral method (see e.g. [33] and [19] for unbounded domains), i.e. a Spectral Galerkin approximation with the use of suitable Gaussian quadrature formulas. We will also show that in some particular cases, where such Gaussian quadratures are exact, it actually coincides with the Spectral Galerkin method based on tensor product spaces.

The outline of the paper is the following: in Section 1 we introduce the mathematical problem and the main notation used throughout. In Section 2

we describe the Collocation method. In Section 3 we provide some regularity results on the solution of the stochastic partial differential equation. In particular, we show that the solution is analytic with respect to the random variables, provided that the input data, as functions of the random variables, have infinite derivatives which do not grow too fast. In Section 4 we give a complete convergence result for the Collocation method and prove exponential convergence. Finally, in Section 5 we present some numerical results showing the effectiveness of the proposed method.

1 Problem setting

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

$$\begin{aligned} -\nabla \cdot (a(\omega,\cdot)\nabla u(\omega,\cdot)) &= f(\omega,\cdot) \quad \text{on} \quad D, \\ u(\omega,\cdot) &= 0 \quad \text{on} \quad \partial D. \end{aligned} \tag{2}$$

We will make the following assumptions:

A_1) $a(\omega,\cdot)$ is uniformly bounded from below, i.e.

$$\text{there exist } a_{min} > 0 \text{ s.t.} \quad P\big(\, \omega \in \Omega : \; a(\omega,x) > a_{min} \; \forall x \in \overline{D} \,\big) = 1$$

A_2) $f(\omega,\cdot)$ is square integrable w.r.t. P, i.e. $\int_D E[f^2]\, dx < \infty$

Moreover, we introduce the following Hilbert spaces:

- $V_P = L^2_P(\Omega) \otimes H^1_0(D)$ equipped with the norm $\|v\|_P^2 = \int_D E\big[|\nabla v|^2\big]\, dx$
- $V_{P,a} \equiv \{v \in V_P : \int_D E\big[a|\nabla v|^2\big]\, dx < \infty\}$, equipped with the norm $\|v\|_{P,a} = \int_D E\big[a|\nabla v|^2\big]\, dx$.

Observe that under the above assumptions, the space $V_{P,a}$ is continuously embedded in V_P and

$$\|v\|_P \le \frac{1}{\sqrt{a_{min}}} \|v\|_{P,a}.$$

Problem (2) can be written in a weak form as

$$\text{find } u \in V_{P,a} \text{ s.t.} \quad \int_D E[a\nabla u \cdot \nabla v]\, dx = \int_D E[fv]\, dx \quad \forall\, v \in V_{P,a}. \tag{3}$$

A straightforward application of the Lax-Milgram theorem allows one to state the well posedness of problem (3); precisely

Lemma 1 *Under assumptions A_1 and A_2, problem* (3) *admits a unique solution $u \in V_{P,a}$, which satisfies the estimate*

$$\|u\|_P \leq \frac{C_P}{a_{min}} \left(\int_D E[f^2]\, dx \right)^{\frac{1}{2}}. \tag{4}$$

In the previous Lemma we have used the Poincaré inequality

$$\|w\|_{L^2(D)} \leq C_P \|\nabla w\|_{L^2(D)} \qquad \forall w \in H_0^1(D).$$

1.0.1 Weaker Assumptions on the random coefficients

It is possible to relax the assumptions A_1 and A_2 substantially and still guarantee the existence and uniqueness of the solution u to problem (3). In particular, if the lower bound for the coefficient a is no longer a constant but a random variable, i.e. $a(x,\omega) \geq a_{min}(\omega) > 0$ a.s. a.e. on D, we have the following estimate for the moments of the solution:

Lemma 2 (Moments estimates) *Let $p, q \geq 0$ with $1/p + 1/q = 1$. Take a positive integer k. Then if $f \in L_P^{kp}(\Omega; L^2(D))$ and $1/a_{min} \in L_P^{kq}(\Omega)$ we have that $u \in L_P^k(\Omega; H_0^1(D))$.*

Proof. Since

$$\|u(\cdot,\omega)\|_{H_0^1(D)} \leq C_P \frac{\|f(\cdot,\omega)\|_{L^2(D)}}{a_{min}(\omega)} \quad a.s.$$

the result is a direct application of Hölder's inequality:

$$\begin{aligned}\int_\Omega \|u(\cdot,\omega)\|_{H_0^1(D)}^k dP(\omega) &\leq C_P^k \int_\Omega \left(\frac{\|f(\cdot,\omega)\|_{L^2(D)}}{a_{min}(\omega)} \right)^k dP(\omega) \\ &\leq C_P^k \left(\int_\Omega \|f(\cdot,\omega)\|_{L^2(D)}^{kp} dP(\omega) \right)^{1/p} \left(\int_\Omega \left(\frac{1}{a_{min}(\omega)} \right)^{qk} dP(\omega) \right)^{1/q}\end{aligned}$$

□

Example 1 (Lognormal diffusion coefficient) *As an application of the previous lemma, we can conclude the well posedness of* (3)*. with a lognormal diffusion coefficient. For instance, let*

$$a(x,\omega) = \exp\left(\sum_{n=1}^N b_n(x) Y_n(\omega) \right),\ Y_n \sim N(0,1)\ iid.$$

Use the lower bound

$$a_{min}(\omega) = \exp\left(-\sum_{n=1}^N \|b_n\|_{L^\infty(D)} |Y_n(\omega)| \right)$$

and then for $k, q < \infty$

$$\|1/a_{min}\|^{kq}_{L^{kq}_P(\Omega)} = \int_\Omega \left(\frac{1}{a_{min}(\omega)}\right)^{qk} dP(\omega)$$
$$= \int_{\mathbb{R}^N} \exp\left(kq \sum_{n=1}^{N} \|b_n\|_{L^\infty(D)}|z_n|\right) \exp\left(-\frac{1}{2}\sum_{n=1}^{N} z_n^2\right) dz_1 \dots dz_N < \infty. \tag{5}$$

Now let $\epsilon > 0$. *Then by Lemma 2 the assumption* $f \in L_P^{k(1+\epsilon)}(\Omega; L^2(D))$ *together with* (5) *imply* $u \in L^k_P(\Omega; H^1_0(D))$.

1.1 Finite Dimensional Noise Assumption

In many problems the source of randomness can be approximated using just a small number of uncorrelated, sometimes independent, random variables; take for example the case of a truncated Karhunen-Loève expansion [4]. This motivates us to assume that

Assumption 1 (finite dimensional noise) *The coefficients used in the computations have the form:*

$$a(\omega, x) = a(Y_1(\omega), \dots, Y_N(\omega), x) \text{ and } f(\omega, x) = f(Y_1(\omega), \dots, Y_N(\omega), x) \text{ on } \Omega \times \overline{D},$$

where $N \in \mathbb{N}_+$, $\{Y_n\}_{n=1}^N$ *are real valued random variables with mean value zero and unit variance.*

We will denote with $\Gamma_n \equiv Y_n(\Omega)$ the image of Y_n, $\Gamma = \prod_{n=1}^N \Gamma_n$ and we will assume that the random variables $[Y_1, Y_2, \dots, Y_n]$ have a joint probability *density function* $\rho : \Gamma \to \mathbb{R}_+$, with $\rho \in L^\infty(\Gamma)$.

Example 2 *The following standard transformation guarantees that the diffusivity coefficient is bounded away from zero almost surely*

$$\log(a - a_{min})(\omega, x) = b_0(x) + \sum_{1 \le n \le N} \sqrt{\lambda_n} b_n(x) Y_n(\omega), \tag{6}$$

i.e. one performs a Karhunen-Loève expansion for $\log(a - a_{min})$, *assuming that* $a > a_{min}$ *almost surely. On the other hand, the right hand side of* (2) *can be represented as a truncated Karhunen-Loève expansion*

$$f(\omega, x) = c_0(x) + \sum_{1 \le n \le N} \sqrt{\mu_n} c_n(x) Y_n(\omega).$$

Remark 1 *It is usual to have* f *and* a *to be independent, because the loads and the material properties are seldom related. In such a situation we have* $a(Y(\omega), x) = a(Y_a(\omega), x)$ *and* $f(Y(\omega), x) = f(Y_f(\omega), x)$, *with* $Y = [Y_a, Y_f]$ *and the vectors* Y_a, Y_f *independent.*

After making Assumption 1, we have by Doob–Dynkin's lemma (cf. [32]), that the solution u of the stochastic elliptic boundary value problem (3) can be described by just a finite number of random variables, i.e. $u(\omega, x) = u(Y_1(\omega), \ldots, Y_N(\omega), x)$. Thus, the goal is to approximate the function $u = u(y, x)$, where $y \in \Gamma$ and $x \in \overline{D}$. Observe that the stochastic variational formulation (3) has a "deterministic" equivalent which is the following: find $u \in V_{\rho,a}$ such that

$$\int_\Gamma \rho \, (a\nabla u, \nabla v)_{L^2(D)} \, dy = \int_\Gamma \rho \, (f, v)_{L^2(D)} \, dy, \quad \forall \, v \in V_{\rho,a} \tag{7}$$

noting that here and later in this work the gradient notation, ∇, always means differentiation with respect to $x \in D$ only, unless otherwise stated. The space $V_{\rho,a}$ is the analogue of $V_{P,a}$ with $(\Omega, \mathcal{F}, P)$ replaced with $(\Gamma, \mathcal{B}^N, \rho dy)$. The stochastic boundary value problem (2) now becomes a deterministic Dirichlet boundary value problem for an elliptic partial differential equation with an $N-$dimensional parameter. For convenience, we consider the solution u as a function $u : \Gamma \to H_0^1(D)$ and we use the notation $u(y)$ whenever we want to highlight the dependence on the parameter y. We use similar notations for the coefficient a and the forcing term f. Then, it can be shown that problem (2) is equivalent to

$$\int_D a(y)\nabla u(y) \cdot \nabla \phi \, dx = \int_D f(y)\phi \, dx, \quad \forall \phi \in H_0^1(D), \;\; \rho\text{-a.e. in } \Gamma. \tag{8}$$

For our convenience, we will suppose that the coefficient a and the forcing term f admit a smooth extension on the ρ-zero measure sets. Then, equation (8) can be extended a.e. in Γ with respect to the Lebesgue measure (instead of the measure ρ).

Hence, making Assumption 1 is a crucial step, turning the original stochastic elliptic equation 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 (cf. [25, 13]).

Remark 2 *Strictly speaking, equation* (8) *will hold only for those values of* $y \in \Gamma$ *for which the coefficient* $a(y)$ *is finite. In this paper we will assume that* $a(y)$ *may go to infinity only at the boundary of the parameter domain* Γ.

2 Collocation method

We seek a numerical approximation to the exact solution of (7) in a finite dimensional subspace $V_{p,h}$ based on a tensor product, $V_{p,h} = \mathcal{P}_p(\Gamma) \otimes H_h(D)$, where

- $H_h(D) \subset H_0^1(D)$ is a standard finite element space of dimension N_h, which contains continuous piecewise polynomials defined on regular triangulations $\mathcal{T}_h$ that have a maximum mesh spacing parameter $h > 0$.

- $\mathcal{P}_p(\Gamma) \subset L^2_\rho(\Gamma)$ is the span of tensor product polynomials with degree at most $p = (p_1, \ldots, p_N)$ i.e. $\mathcal{P}_p(\Gamma) = \bigotimes_{n=1}^N \mathcal{P}_{p_n}(\Gamma_n)$, with

$$\mathcal{P}_{p_n}(\Gamma_n) = \text{span}(y_n^m,\, m = 0, \ldots, p_n), \quad n = 1, \ldots, N.$$

  Hence the dimension of $\mathcal{P}_p$ is $N_p = \prod_{n=1}^N (p_n + 1)$.

We first introduce the semi-discrete approximation $u_h : \Gamma \to H_h(D)$, obtained by projecting equation (8) onto the subspace $H_h(D)$, for each $y \in \Gamma$, i.e.

$$\int_D a(y)\nabla u_h(y) \cdot \nabla \phi_h \, dx = \int_D f(y)\phi_h \, dx, \quad \forall \phi_h \in H_h(D), \text{ for a.e. } y \in \Gamma. \quad (9)$$

The next step consists in collocating equation (9) on the zeros of orthogonal polynomials and build the discrete solution $u_{h,p} \in \mathcal{P}_p(\Gamma) \otimes H_h(D)$ by interpolating in y the collocated solutions.

To this end, we first introduce an auxiliary probability density function $\hat{\rho} : \Gamma \to \mathbb{R}^+$ that can be seen as the joint probability of N independent random variables, i.e. it factorizes as

$$\hat{\rho}(y) = \prod_{n=1}^N \hat{\rho}_n(y_n), \;\; \forall y \in \Gamma, \qquad \text{and is such that} \quad \left\| \frac{\rho}{\hat{\rho}} \right\|_{L^\infty(\Gamma)} < \infty. \qquad (10)$$

For each dimension $n = 1, \ldots, N$ let y_{n,k_n}, $1 \le k_n \le p_n + 1$ be the $p_n + 1$ roots of the orthogonal polynomial q_{p_n+1} with respect to the weight $\hat{\rho}_n$, which satisfies then $\int_{\Gamma_n} q_{p_n+1}(y) v(y) \hat{\rho}_n(y) dy = 0, \quad \forall v \in \mathcal{P}_{p_n}(\Gamma_n)$.

Standard choices for $\hat{\rho}$, such as constant, Gaussian, etc., lead to well known roots of the polynomial q_{p_n+1}, which are tabulated to full accuracy and do not need to be computed.

To any vector of indexes $[k_1, \ldots, k_N]$ we associate the global index

$$k = k_1 + p_1(k_2 - 1) + p_1 p_2 (k_3 - 1) + \ldots$$

and we denote by y_k the point $y_k = [y_{1,k_1}, y_{2,k_2}, \ldots, y_{N,k_N}] \in \Gamma$. We also introduce, for each $n = 1, 2, \ldots, N$, the Lagrange basis $\{l_{n,j}\}_{j=1}^{p_n+1}$ of the space $\mathcal{P}_{p_n}$:

$$l_{n,j} \in \mathcal{P}_{p_n}(\Gamma_n); \qquad l_{n,j}(y_{n,k}) = \delta_{jk}, \quad j, k = 1, \ldots, p_n + 1$$

where δ_{jk} is the Kronecker symbol, and we set $l_k(y) = \prod_{n=1}^N l_{n,k_n}(y_n)$. Hence, the final approximation is given by

$$u_{h,p}(y, x) = \sum_{k=1}^{N_p} u_h(y_k, x) l_k(y)$$

where $u_h(y_k, x)$ is the solution of problem (9) for $y = y_k$.

Equivalently, if we introduce the Lagrange interpolant operator $\mathcal{I}_p : C^0(\Gamma; H_0^1(D)) \to \mathcal{P}_p(\Gamma) \otimes H_0^1(D)$, such that

$$\mathcal{I}_p v(y) = \sum_{n=1}^{N} v(y_k) l_k(y), \qquad \forall v \in C^0(\Gamma; H_0^1(D))$$

then we have simply $u_{h,p} = \mathcal{I}_p u_h$.

Finally, for any continuous function $g : \Gamma \to \mathbb{R}$ we introduce the Gauss quadrature formula $E_{\hat{\rho}}^p[g]$ approximating the integral $\int_\Gamma g(y)\hat{\rho}(y)\,dy$ as

$$E_{\hat{\rho}}^p[g] = \sum_{k=1}^{N_p} \omega_k g(y_k), \quad \omega_k = \prod_{n=1}^{N} \omega_{k_n}, \quad \omega_{k_n} = \int_{\Gamma_n} l_{k_n}^2(y)\hat{\rho}_n(y)\,dy. \tag{11}$$

This can be used to approximate the mean value or the variance of u as

$$\bar{u}_h \in H_h(D), \qquad \bar{u}_h(x) = E_{\hat{\rho}}^p\left[\frac{\rho}{\hat{\rho}} u_h(x)\right]$$
$$\mathrm{var}_h(u_h) \in L^1(D), \qquad \mathrm{var}_h(u_h)(x) = E_{\hat{\rho}}^p\left[\frac{\rho}{\hat{\rho}}\left(u_h(x) - \bar{u}_h(x)\right)^2\right]$$

as long as $\rho/\hat{\rho}$ is a smooth function. Otherwise, $\bar{u}_h$ and $\mathrm{var}_h(u_h)$ should be computed with a suitable quadrature formula which takes into account eventual discontinuities or singularities of $\rho/\hat{\rho}$.

2.1 Collocation versus Spectral Galerkin approximation

An approach, alternative to the Collocation method introduced so far, consists in approximating problem (7) with a Spectral Galerkin method: *find* $u_{h,p}^G \in \mathcal{P}_p(\Gamma) \otimes H_h(D)$ *such that*

$$\int_\Gamma \rho\,(a\nabla u_{h,p}^G, \nabla v)_{L^2(D)}\,dy = \int_\Gamma \rho\,(f, v)_{L^2(D)}\,dy, \quad \forall\, v \in \mathcal{P}_p(\Gamma) \otimes H_h(D). \tag{12}$$

This approach has been considered by several authors ([4, 13, 18, 41, 21, 29]). Observe that, in general, problem (12) leads to a fully coupled system of linear equations, whose dimension is $N_h \times N_p$ and demands for highly efficient strategies and parallel computations for its numerical solution [15]. Conversely, the Collocation method only requires the solutions of N_p uncoupled linear systems of dimension N_h, and is fully parallelizable.

In [4, 5] a particular choice of basis functions (named *double orthogonal polynomials*) for the space $P_p(\Gamma)$ is proposed . This choice allows to decouple the system in the special case where the diffusivity coefficient and the forcing term are multi-linear combinations of the random variables $Y_n(\omega)$ (as it is the case if one performs a truncated linear Karhunen-Loève expansion) and the random variables are independent, i.e. $\rho(y) = \prod_{n=1}^{N} \rho_n(y_n)$. The proposed

basis is then obtained by solving the following eigenvalue problems, for each $n = 1, \dots, N$,

$$\int_{\Gamma_n} z\psi_{kn}(z)v(z)\rho_n(z)\,dz = c_{kn}\int_{\Gamma_n} \psi_{kn}(z)v(z)\rho_n(z)\,dz, \qquad k = 1, \dots, p_n + 1.$$

The eigenvectors ψ_{kn} are normalized so as to satisfy the property

$$\int_{\Gamma_n} \psi_{kn}(z)\psi_{jn}(z)\rho_n(z)\,dz = \delta_{kj}, \qquad \int_{\Gamma_n} z\psi_{kn}(z)\psi_{jn}(z)\rho_n(z)\,dz = c_{kn}\delta_{kj}.$$

See [4, 5] for further details on the double orthogonal basis.

We aim at analyzing, now, the analogies between the Collocation and the Spectral Galerkin methods. The Collocation method can be seen as a *Pseudo-Spectral* Galerkin method (see e.g. [33]) where the integrals over Γ in (12) are replaced by the quadrature formula (11): *find* $u_{h,p} \in \mathcal{P}_p(\Gamma) \otimes H_h(D)$ *such that*

$$E^p_{\hat\rho}\left[\frac{\rho}{\hat\rho}\,(a\nabla u_{h,p}, \nabla v)_{L^2(D)}\right] = E^p_{\hat\rho}\left[\frac{\rho}{\hat\rho}\,(f, v)_{L^2(D)}\right], \quad \forall\, v \in \mathcal{P}_p(\Gamma)\otimes H_h(D). \tag{13}$$

Indeed, by choosing in (13), the test functions of the form $v(y, x) = l_k(y)\phi(x)$, where $\phi(x) \in H_h(D)$ and $l_k(y)$ is the Lagrange basis function associated to the knot y_k, $k = 1, \dots, N_p$, one is led to solve a sequence of uncoupled problems of the form (9) collocated in the points y_k, which, ultimately, gives the same solution as the Collocation method.

In the particular case where the diffusivity coefficient and the forcing term are multi-linear combinations of the random variables $Y_n(\omega)$, and the random variables are independent, it turns out that the quadrature formula is exact if one chooses $\hat\rho = \rho$. In this case, the solution obtained by the Collocation method actually coincides with the Spectral Galerkin one. This can be seen easily observing that, with the above assumptions, the integrand in (12), i.e. $(a\nabla u_{h,p} \cdot \nabla v)$ is a polynomial at most of degree $2p_n + 1$ in the variable y_n and the Gauss quadrature formula is exact for polynomials up to degree $2p_n + 1$ integrated against the weight ρ.

The Collocation method is a natural generalization of the Spectral Galerkin approach, and has the following advantages:

- decouples the system of linear equations in Y also in the case where the diffusivity coefficient a and the forcing term f are non linear functions of the random variables Y_n;
- treats efficiently the case of non independent random variables with the introduction of the auxiliary measure $\hat\rho$;
- can easily deal with random variables with unbounded support (see Theorem 1 in Section 4).

As it will be shown in Section 4, the Collocation method preserves the same accuracy as the Spectral Galerkin approach and achieves exponential convergence if the coefficient a and forcing term f are infinitely differentiable with respect to the random variables Y_n, under very mild requirements on the growth of their derivatives in Y.

As a final remark, we show that the double orthogonal polynomials proposed in [4] coincide with the Lagrange basis $l_k(y)$ and the eigenvalues c_{kn} are nothing but the Gauss knots of integration.

Lemma 3 *Let $\Gamma \subset \mathbb{R}$, $\rho : \Gamma \to \mathbb{R}$ a positive weight and $\{\psi_k\}_{k=1}^{p+1}$ the set of double orthogonal polynomials of degree p satisfying*

$$\int_\Gamma \psi_k(y)\psi_j(y)\rho(y)\,dy = \delta_{kj}, \qquad \int_\Gamma y\psi_k(y)\psi_j(y)\rho(y)\,dy = c_k\delta_{kj}.$$

Then, the eigenvalues c_k are the nodes of the Gaussian quadrature formula associated to the weight ρ and the eigenfunctions ψ_k are, up to multiplicative factors, the corresponding Lagrange polynomials build on the nodes c_k.

Proof. We have, for $k = 1, \ldots, p+1$,

$$\int_\Gamma (y - c_k)\psi_k(y)v(y)\rho(y)dy = 0, \quad \forall v \in P_p(\Gamma).$$

Take $v = \prod_{\substack{j=1 \\ j\neq k}}^{p+1}(y - c_j) \in P_p(\Gamma)$ in the above and let $w = \prod_{j=1}^{p+1}(y - c_j)$. Then

$$\int_\Gamma w(y)\psi_k(y)\rho(y)dy = 0, \quad \forall k = 1, \ldots, p+1.$$

Since $\{\psi_k\}_{k=1}^{p+1}$ defines a basis of the space $\mathcal{P}_p(\Gamma)$, the previous relation implies that w is ρ-orthogonal to $P_p(\Gamma)$. Besides, the functions $(y - c_k)\psi_k$ are also orthogonal to the same subspace: this yields, due to the one dimensional nature of the orthogonal complement of $P_p(\Gamma)$ over $P_{p+1}(\Gamma)$,

$$(y - c_k)\psi_k = \alpha_k w = \alpha_k \prod_{j=1}^{p+1}(y - c_j), \quad k = 1, \ldots, p+1$$

which gives

$$\psi_k = \alpha_k \prod_{\substack{j=1 \\ j\neq k}}^{p+1}(y - c_j), \quad k = 1, \ldots, p+1$$

i.e. the double orthogonal polynomials ψ_k are collinear to Lagrange interpolants at the nodes c_j. Moreover, the eigenvalues c_j are the roots of the polynomial $w \in P_{p+1}(\Gamma)$, which is ρ-orthogonal to $P_p(\Gamma)$ and therefore they coincide with the nodes of the Gaussian quadrature formula associated with the weight ρ. □

3 Regularity results

Before going through the convergence analysis of the method, we need to state some regularity assumptions on the data of the problem and consequent regularity results for the exact solution u and the semi-discrete solution u_h.

In what follows we will need some restrictive assumptions on f and ρ. In particular, we will assume f to be a continuous function in y, whose growth at infinity, whenever the domain Γ is unbounded, is at most exponential. At the same time we will assume that ρ behaves as a Gaussian weight at infinity, and so does the auxiliary density $\hat{\rho}$, in light of assumption (10).

Other types of growth of f at infinity and corresponding decay of the probability density ρ, for instance of exponential type, could be considered as well. Yet, we will limit the analysis to the aforementioned case.

To make precise these assumptions, we introduce a weight $\sigma(y) = \prod_{n=1}^{N} \sigma_n(y_n) \leq 1$ where

$$\sigma_n(y_n) = \begin{cases} 1 & \text{if } \Gamma_n \text{ is bounded} \\ e^{-\alpha_n |y_n|}, \text{ for some } \alpha_n > 0 & \text{if } \Gamma_n \text{ is unbounded} \end{cases} \tag{14}$$

and the functional space

$$C^0_\sigma(\Gamma; V) \equiv \{v : \Gamma \to V, \;\; v \text{ continuous in } y, \;\; \max_{y \in \Gamma} \|\sigma(y) v(y)\|_V < +\infty\}$$

where V is a Banach space of functions defined in D.

Assumption 2 (growth at infinity) *In what follows we will assume*

a) $f \in C^0_\sigma(\Gamma; L^2(D))$

b) the joint probability density ρ *satisfies*

$$\rho(y) \leq C_\rho \, e^{-\sum_{n=1}^{N} (\delta_n y_n)^2}, \qquad \forall y \in \Gamma, \tag{15}$$

for some constant $C_\rho > 0$ *and* δ_n *strictly positive if* Γ_n *is unbounded and zero otherwise.*

The parameter δ_n in (15) gives a scale for the decay of ρ at infinity and provides an estimate of the dispersion of the random variable Y_n. On the other hand, the parameter α_n in (14) controls the growth of the forcing term f at infinity.

Remark 3 (growth of f) *The convergence result given in Theorem 1, in the next section, extends to a wider class of functions* f*. For instance we could take* $f \in C^0_\sigma(\Gamma; L^2(D))$ *with* $\sigma = e^{-\sum_{n=1}^{N} (\delta_n y_n)^2/8}$*. Yet, the class given in* (14) *is already large enough for most practical applications (see Example 2).*

We can now choose any suitable auxiliary density $\hat{\rho}(y) = \prod_{n=1}^{N} \hat{\rho}_n(y_n)$ that satisfies, for each $n = 1, \dots, N$

$$C_{min}^n e^{-(\delta_n y_n)^2} \leq \hat{\rho}_n(y_n) < C_{max}^n e^{-(\delta_n y_n)^2}, \qquad \forall y_n \in \Gamma_n \tag{16}$$

for some positive constants C_{min}^n and C_{max}^n that do not depend on y_n.

Observe that this choice satisfies the requirement given in (10), i.e. $\|\rho/\hat{\rho}\|_{L^\infty(\Gamma)} \leq C_\rho/C_{min}$ with $C_{min} = \prod_{n=1}^{N} C_{min}^n$.

Under the above assumptions, the following inclusions hold true

$$C_\sigma^0(\Gamma; V) \subset L_{\hat{\rho}}^2(\Gamma; V) \subset L_\rho^2(\Gamma; V)$$

with continuous embedding. Indeed, on one hand we have

$$\|v\|_{L_\rho^2(\Gamma;V)} \leq \left\| \frac{\rho}{\hat{\rho}} \right\|_{L^\infty(\Gamma)}^{\frac{1}{2}} \|v\|_{L_{\hat{\rho}}^2(\Gamma;V)} \leq \sqrt{\frac{C_\rho}{C_{min}}} \|v\|_{L_{\hat{\rho}}^2(\Gamma;V)}.$$

On the other hand,

$$\|v\|_{L_{\hat{\rho}}^2(\Gamma;V)}^2 = \int_\Gamma \hat{\rho}(y) \|v(y)\|_V^2 \, dy \leq \|v\|_{C_\sigma^0(\Gamma;V)}^2 \int_\Gamma \frac{\hat{\rho}(y)}{\sigma^2(y)} \, dy \leq \prod_{n=1}^{N} M_n \|v\|_{C_\sigma^0(\Gamma;V)}^2$$

with $M_n = \int_{\Gamma_n} \hat{\rho}_n/\sigma_n^2$. Now, for Γ_n bounded, $M_n \leq C_{max}^n |\Gamma_n|$, whereas if Γ_n is unbounded

$$M_n = \int_{\Gamma_n} \left(e^{-\frac{(\delta_n y)^2}{2} + 2\alpha_n |y|} \right) e^{\frac{(\delta_n y)^2}{2}} \hat{\rho}_n(y) \, dy \leq C_{max}^n \sqrt{\frac{2\pi}{\delta_n}} e^{2(\alpha_n/\delta_n)^2}.$$

The first result we need is the following

Lemma 4 *If $f \in C_\sigma^0(\Gamma; L^2(D))$ and $a \in C_{loc}^0(\Gamma; L^\infty(D))$, uniformly bounded away from zero, then the solution to problem (8) satisfies $u \in C_\sigma^0(\Gamma; H_0^1(D))$.*

The proof of this Lemma is immediate. The next result concerns the analyticity of the solution u whenever the diffusivity coefficient a and the forcing term f are infinitely differentiable w.r.t. y, under mild assumptions on the growth of their derivatives in y. We will perform a one-dimensional analysis in each direction y_n, $n = 1, \dots, N$. For this, we introduce the following notation: $\Gamma_n^* = \prod_{\substack{j=1 \\ j \neq n}}^{N} \Gamma_j$, y_n^* will denote an arbitrary element of Γ_n^*. Similarly, we set $\hat{\rho}_n^* = \prod_{\substack{j=1 \\ j \neq n}}^{N} \hat{\rho}_j$ and $\sigma_n^* = \prod_{\substack{j=1 \\ j \neq n}}^{N} \sigma_j$.

Lemma 5 *Under the assumption that, for every $y = (y_n, y_n^*) \in \Gamma$, there exists $\gamma_n < +\infty$ such that*

$$\left\| \frac{\partial_{y_n}^k a(y)}{a(y)} \right\|_{L^\infty(D)} \leq \gamma_n^k k! \qquad \textit{and} \qquad \frac{\|\partial_{y_n}^k f(y)\|_{L^2(D)}}{1 + \|f(y)\|_{L^2(D)}} \leq \gamma_n^k k!, \tag{17}$$

the solution $u(y_n, y_n^*, x)$ *as a function of* y_n, $u : \Gamma_n \to C^0_{\sigma_n^*}(\Gamma_n^*; H_0^1(D))$ *admits an analytic extension* $u(z, y_n^*, x)$, $z \in \mathbb{C}$, *in the region of the complex plane*

$$\Sigma(\Gamma_n; \tau_n) \equiv \{z \in \mathbb{C}, \ \ \mathrm{dist}(z, \Gamma_n) \leq \tau_n\}. \tag{18}$$

with $0 < \tau_n < 1/(2\gamma_n)$. *Moreover,* $\forall z \in \Sigma(\Gamma_n; \tau_n)$,

$$\|\sigma_n(\mathrm{Re}\, z)\, u(z)\|_{C^0_{\sigma_n^*}(\Gamma_n^*; H_0^1(D))} \leq \frac{C_P e^{\alpha_n \tau_n}}{a_{min}(1 - 2\tau_n\gamma_n)}(2\|f\|_{C^0_\sigma(\Gamma; H_0^1(D))} + 1) \tag{19}$$

with the constant C_p *as in* (4).

Proof. In every point $y \in \Gamma$, the k-th derivative of u w.r.t y_n satisfies the problem

$$B(y; \partial^k_{y_n} u, v) = \sum_{l=1}^{k} \binom{k}{l} \partial^l_{y_n} B(y; \partial^{k-l}_{y_n} u, v) + (\partial^k_{y_n} f, v), \quad \forall v \in H_0^1(D),$$

where B is the parametric bilinear form $B(y; u, v) = \int_D a(y)\nabla u \cdot \nabla v\, dx$. Hence

$$\|\sqrt{a(y)}\nabla \partial^k_{y_n} u\|_{L^2(D)} \leq \sum_{l=1}^{k} \binom{k}{l} \left\| \frac{\partial^l_{y_n} a(y)}{a(y)} \right\|_{L^\infty(D)} \|\sqrt{a(y)}\nabla \partial^{k-l}_{y_n} u\|_{L^2(D)} + \frac{C_p}{\sqrt{a_{min}}} \|\partial^k_{y_n} f\|_{L^2(D)}.$$

Setting $R_k = \|\sqrt{a(y)}\nabla \partial^k_{y_n} u\|_{L^2(D)}/k!$ and using the bounds on the derivatives of a and f, we obtain the recursive inequality

$$R_k \leq \sum_{l=1}^{k} \gamma_n^l R_{k-l} + \frac{C_p}{\sqrt{a_{min}}} \gamma_n^k (1 + \|f\|_{L^2(D)}).$$

The generic term R_k admits the bound

$$R_k \leq (2\gamma_n)^k R_0 + \frac{C_p}{\sqrt{a_{min}}}(1 + \|f\|_{L^2(D)})\gamma_n^k \sum_{l=0}^{k-1} 2^l.$$

Observing, now that $R_0 = \|\sqrt{a(y)}\nabla u(y)\|_{L^2(D)} \leq \frac{C_p}{\sqrt{a_{min}}}\|f(y)\|_{L^2(D)}$ and

$$\frac{\|\nabla \partial^k_{y_n} u\|_{L^2(D)}}{k!} \leq \frac{R_k}{\sqrt{a_{min}}},$$

we get the final estimate on the growth of the derivatives of u

$$\frac{\|\nabla \partial^k_{y_n} u(y)\|_{L^2(D)}}{k!} \leq \frac{C_p}{a_{min}}(2\|f(y)\|_{L^2(D)} + 1)(2\gamma_n)^k.$$

We now define for every $y_n \in \Gamma_n$ the power series $u : \mathbb{C} \to C^0_{\sigma_n^*}(\Gamma_n^*, H_0^1(D))$ as

$$u(z, y_n^*, x) = \sum_{k=0}^{\infty} \frac{(z - y_n)^k}{k!} \partial^k_{y_n} u(y_n, y_n^*, x).$$

Hence,

$$\sigma_n(y_n)\|u(z)\|_{C^0_{\sigma_n^*}(\Gamma_n^*,H_0^1(D))} \le \sum_{k=0}^{\infty} \frac{|z-y_n|^k}{k!}\sigma_n(y_n)\|\partial_{y_n}^k u(y_n)\|_{C^0_{\sigma_n^*}(\Gamma_n^*;H_0^1(D))}$$

$$\le \frac{C_P}{a_{min}} \max_{y_n\in\Gamma_n} \left\{\sigma_n(y_n)\left(2\|f(y_n)\|_{C^0_{\sigma_n^*}(\Gamma_n^*;L^2(D))}+1\right)\right\} \sum_{k=0}^{\infty} (|z-y_n|2\gamma_n)^k$$

$$\le \frac{C_P}{a_{min}}(2\|f\|_{C^0_\sigma(\Gamma;L^2(D))}+1)\sum_{k=0}^{\infty}(|z-y_n|2\gamma_n)^k$$

where we have exploited the fact that $\sigma_n(y_n) \le 1$ for all $y_n \in \Gamma_n$; the series converges for all $z \in \mathbb{C}$ such that $|z-y_n| \le \tau_n < 1/(2\gamma_n)$. Moreover, in the ball $|z-y_n| \le \tau_n$, we have, by virtue of (14), $\sigma_n(\mathrm{Re}\, z) \le e^{\alpha_n\tau_n}\sigma_n(y_n)$ and then

$$\sigma_n(\mathrm{Re}\, z)\|u(z)\|_{C^0_{\sigma_n^*}(\Gamma_n^*,H_0^1(D))} \le \frac{C_P e^{\alpha_n\tau_n}}{a_{min}(1-2\tau_n\gamma_n)}(2\|f\|_{C^0_\sigma(\Gamma;L^2(D))}+1)$$

The power series converges for every $y_n \in \Gamma_n$, hence, by a continuation argument, the function u can be extended analytically on the whole region $\Sigma(\Gamma_n;\tau_n)$ and estimate (19) follows. □

Example 3 *Let us consider the case where the diffusivity coefficient a is expanded in a linear truncated Karhunen-Loève expansion*

$$a(\omega,x) = b_0(x) + \sum_{n=1}^{N}\sqrt{\lambda_n}b_n(x)Y_n(\omega),$$

provided that such expansion guarantees $a(\omega,x) \ge a_{min}$ for almost every $\omega \in \Omega$ and $x \in D$ [34]. In this case we have

$$\left\|\frac{\partial_{y_n}^k a}{a}\right\|_{L^\infty(\Gamma\times D)} \le \begin{cases}\sqrt{\lambda_n}\|b_n\|_{L^\infty(D)}/a_{min} & \text{for } k=1\\ 0 & \text{for } k>1\end{cases}$$

and we can safely take $\gamma_n = \sqrt{\lambda_n}\|b_n\|_{L^\infty(D)}/a_{min}$ in (17).

If we consider, instead, a truncated exponential expansion

$$a(\omega,x) = a_{min} + e^{b_0(x)+\sum_{n=1}^{N}\sqrt{\lambda_n}b_n(x)Y_n(\omega)}$$

we have

$$\left\|\frac{\partial_{y_n}^k a}{a}\right\|_{L^\infty(\Gamma\times D)} \le \left(\sqrt{\lambda_n}\|b_n\|_{L^\infty(D)}\right)^k$$

and we can take $\gamma_n = \sqrt{\lambda_n}\|b_n\|_{L^\infty(D)}$. Hence, both choices fulfill the assumption in Lemma 5.

Example 4 *Similarly to the previous case, let us consider a forcing term f of the form*

$$f(\omega, x) = c_0(x) + \sum_{n=1}^{N} c_n(x) Y_n(\omega)$$

where the random variables Y_n are Gaussian (either independent or not), and the functions $c_n(x)$ are square integrable for any $n = 1, \dots, N$. Then, the function f belongs to the space $C^0_\sigma(\Gamma; L^2(D))$, with weight σ defined in (14)*, for any choice of the exponent coefficients $\alpha_n > 0$.*

Moreover,

$$\frac{\|\partial^k_{y_n} f(y)\|_{L^2(D)}}{1 + \|f(y)\|_{L^2(D)}} \leq \begin{cases} \|c_n\|_{L^2(D)} & \text{for } k = 1 \\ 0 & \text{for } k > 1 \end{cases}$$

and we can safely take $\gamma_n = \|c_n\|_{L^2(D)}$ in (17)*. Hence, such a forcing term satisfies the assumptions of Lemma 5. In this case, though, the solution u is linear with respect to the random variables Y_n (hence, clearly analytic) and our theory is not needed.*

Observe that the regularity results are valid also for the semidiscrete solution u_h, as well.

4 Convergence analysis

Our aim is to give a priori estimates for the total error $\epsilon = u - u_{h,p}$ in the natural norm $L^2_\rho(\Gamma) \otimes H^1_0(D)$. The next Theorem states the sought convergence result, and the rest of the section will be devoted to its proof. In particular, we will prove that the error decays (sub)exponentially fast w.r.t. p under the regularity assumptions made in Section 3. The convergence w.r.t. h will be dictated by standard approximability properties of the finite element space $H_h(D)$ and the regularity in space of the solution u (see e.g. [12, 11]).

Theorem 1 *Under the assumptions of Lemmas 4 and 5, there exist positive constants r_n, $n = 1, \dots, N$, and C, independent of h and p, such that*

$$\begin{aligned} \|u - u_{h,p}\|_{L^2_\rho \otimes H^1_0} \leq & \frac{1}{\sqrt{a_{min}}} \inf_{v \in L^2_\rho \otimes H_h} \left(\int_{\Gamma \times D} \rho a |\nabla(u - v)|^2 \right)^{\frac{1}{2}} \\ & + C \sum_{n=1}^{N} \beta_n(p_n) \exp\{-r_n \, p_n^{\theta_n}\} \end{aligned} \tag{20}$$

where

- *if Γ_n is bounded* $\quad \begin{cases} \theta_n = \beta_n = 1 \\ r_n = \log\left[\frac{2\tau_n}{|\Gamma_n|} \left(1 + \sqrt{1 + \frac{|\Gamma_n|^2}{4\tau_n^2}} \right) \right] \end{cases}$
- *if Γ_n is unbounded* $\quad \begin{cases} \theta_n = 1/2, \quad \beta_n = O(\sqrt{p_n}) \\ r_n = \tau_n \delta_n \end{cases}$

τ_n is the distance between Γ_n and the nearest singularity in the complex plane, as defined in Lemma 5, and δ_n is defined as in (15).

The first term on the right hand side of (20) concerns the space approximability of u in the subspace $H_h(D)$ and is controlled by the mesh size h. The actual rate of convergence will depend on the regularity in space of $a(y)$ and $f(y)$ for each $y \in \Gamma$ as well as on the smoothness on the domain D. Observe that an h or $h-p$ adaptive strategy to reduce the error in space is not precluded by this approach.

The exponential rate of convergence in the Y direction depends on the constants r_n, which on their turn, are related to distances τ_n to the nearest singularity in the complex plane. In Examples 3 and 4 we have estimated these constants in the case where the random fields a and f are represented by either a linear or exponential truncated Karhunen-Loève expansion. Hence, a full characterization of the convergence rate is available in these cases.

Observe that in Theorem 1 it is not necessary to assume the finiteness of the second moment of the coefficient a. Before proving the theorem, we recall

some known results of approximation theory for a function f defined on a one dimensional domain (bounded or unbounded) with values in a Banach space V, $f : \Gamma \subset \mathbb{R} \to V$. As in Section 2, let $\rho : \Gamma \to \mathbb{R}^+$ be a positive weight which satisfies, for all $y \in \Gamma$, $\rho(y) \leq C_M e^{-(\delta y)^2}$ for some $C_M > 0$ and δ strictly positive is Γ is unbounded and zero otherwise; $y_k \in \Gamma$, $k = 1, \ldots, p+1$ the set of zeros of the polynomial of degree p orthogonal to the space $\mathcal{P}_{p-1}$ with respect to the weight ρ; σ an extra positive weight such that $\sigma(y) \geq C_m e^{-(\delta y)^2/4}$ for some $C_m > 0$. With this choice, the embedding $C^0_\sigma(\Gamma; V) \subset L^2_\rho(\Gamma; V)$ is continuous. Observe that the condition on σ is satisfied both by a Gaussian weight $\sigma = e^{-(\mu y)^2}$ with $\mu \leq \delta/2$ and by an exponential weight $\sigma = e^{-\alpha|y|}$ for any $\alpha \geq 0$. Finally, we denote by $\mathcal{I}_p$ the Lagrange interpolant operator: $\mathcal{I}_p v(y) = \sum_{k=1}^{p+1} v(y_k) l_k(y)$, for every continuous function v and by $\omega_k = \int_\Gamma l_k^2(y)\rho(y)\, dy$ the weights of the Gaussian quadrature formula built upon $\mathcal{I}_p$.

The following two lemmas are a slight generalization of a classical result by Erdös and Turán [17].

Lemma 6 *The operator $\mathcal{I}_p : C^0_\sigma(\Gamma; V) \to L^2_\rho(\Gamma; V)$ is continuous.*

Proof. We have, indeed, that for any $v \in C^0_\sigma(\Gamma; V)$

$$\|\mathcal{I}_p v\|^2_{L^2_\rho(\Gamma;V)} = \int_\Gamma \| \sum_{k=1}^{p+1} v(y_k) l_k(y) \|^2_V \rho(y)\, dy \leq \int_\Gamma \left(\sum_{k=1}^{p+1} \|v(y_k)\|_V l_k(y) \right)^2 \rho(y)\, dy.$$

Thanks to the orthogonality property $\int_\Gamma l_j(y)l_k(y)\rho(y)\,dy = \delta_{jk}$, we have

$$\begin{aligned}\|\mathcal{I}_p v\|^2_{L^2_\rho(\Gamma;V)} &\le \int_\Gamma \sum_{k=1}^{p+1} \|v(y_k)\|^2_V l_k^2(y)\rho(y)\,dy \\ &\le \max_{k=1,\dots,p+1} \|v(y_k)\|^2_V \sigma^2(y_k) \sum_{k=1}^{p+1} \int_\Gamma \frac{l_k^2(y)\rho(y)}{\sigma^2(y_k)}\,dy \\ &\le \|v\|^2_{C^0_\sigma(\Gamma;V)} \sum_{k=1}^{p+1} \frac{\omega_k}{\sigma^2(y_k)}.\end{aligned}$$

In the case of Γ bounded, we have $\sigma \ge C_m$ and $\sum_{k=1}^{p+1}\omega_k = 1$ for any p and the result follows immediately. For Γ unbounded, since $\rho(y) \le C_M e^{-(\delta y)^2}$, all the even moments $c_{2m} = \int_\Gamma y^{2m}\rho(y)\,dy$ are bounded, up to a constant, by the moments of the Gaussian density $e^{-(\delta y)^2}$. Therefore, using a result from Uspensky '28 [36], it follows

$$\sum_{k=1}^{p+1} \frac{\omega_k}{\sigma^2(y_k)} \overset{p\to\infty}{\longrightarrow} \int_\Gamma \frac{\rho(y)}{\sigma^2(y)}\,dy \le \frac{C_M}{C_m^2}\sqrt{\frac{2\pi}{\delta}}$$

and we conclude that

$$\|\mathcal{I}_p v\|_{L^2_\rho(\Gamma;V)} \le C_1 \|v\|_{C^0_\sigma(\Gamma;V)}$$

□

Lemma 7 *For every function $v \in C^0_\sigma(\Gamma;V)$ the interpolation error satisfies*

$$\|v - \mathcal{I}_p v\|_{L^2_\rho(\Gamma;V)} \le C_2 \inf_{w\in\mathcal{P}_p(\Gamma)\otimes V} \|v - w\|_{C^0_\sigma(\Gamma;V)}.$$

with a constant C_2 independent of p.

Proof. Let us observe that $\forall w \in \mathcal{P}_p(\Gamma)\otimes V$, it holds $\mathcal{I}_p w = w$. Then,

$$\begin{aligned}\|v - \mathcal{I}_p v\|_{L^2_\rho(\Gamma;V)} &\le \|v - w\|_{L^2_\rho(\Gamma;V)} + \|\mathcal{I}_p(w - v)\|_{L^2_\rho(\Gamma;V)} \\ &\le C_2 \|v - w\|_{C^0_\sigma(\Gamma;V)}.\end{aligned}$$

Since w is arbitrary in the right hand side, the result follows. □

The previous Lemma relates the approximation error $(v - \mathcal{I}_p v)$ in the L^2_ρ-norm with the *best approximation* error in the weighted C^0_σ-norm, for any weight $\sigma(y) \ge C_m e^{-(\delta y)^2/4}$. We analyze now the best approximation error for a function $v : \Gamma \to V$ which admits an analytic extension in the complex plane, in the region $\Sigma(\Gamma;\tau) = \{z \in \mathbb{C},\ \ \mathrm{dist}(z,\Gamma) < \tau\}$ for some $\tau > 0$. We will still denote the extension by v; in this case, τ represents the distance between $\Gamma \subset \mathbb{R}$ and the nearest singularity of $v(z)$ in the complex plane.

We study separately the two cases of Γ bounded and unbounded. We start with the bounded case, in which the extra weight σ is set equal to 1. The following result is an immediate extension of the result given in [14, Chapter 7, Section 8]

Lemma 8 *Given a function $v \in C^0(\Gamma; V)$ which admits an analytic extension in the region of the complex plane $\Sigma(\Gamma;\tau) = \{z \in \mathbb{C},\ \ \mathrm{dist}(z,\Gamma) \leq \tau\}$ for some $\tau > 0$, it holds:*

$$\min_{w \in \mathcal{P}_p \otimes V} \|v - w\|_{C^0(\Gamma;V)} \leq \frac{2}{\varrho - 1} e^{-p \log(\varrho)} \max_{z \in \Sigma(\Gamma;\tau)} \|v(z)\|_V$$

where $1 < \varrho = \dfrac{2\tau}{|\Gamma|} + \sqrt{1 + \dfrac{4\tau^2}{|\Gamma|^2}}$.

Proof. We sketch the proof for completeness. We first make a change of variables, $y(t) = y_0 + \frac{|\Gamma|}{2} t$, where y_0 is the midpoint of Γ. Hence, $y([-1,1]) = \Gamma$. We set $\tilde{v}(t) = v(y(t))$. Clearly, $\tilde{v}$ can be extended analytically in the region of the complex plane $\Sigma([-1,1]; 2\tau/|\Gamma|) \equiv \{z \in \mathbb{C}, \mathrm{dist}(z,[-1,1]) \leq 2\tau/|\Gamma|\}$.

We then introduce the Chebyshev polynomials $C_k(t)$ on $[-1,1]$ and the expansion of $\tilde{v} : [-1,1] \to V$ as

$$\tilde{v}(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k C_k(t), \tag{21}$$

where the Fourier coefficients $a_k \in V$, $k = 0, 1, \ldots$, are defined as

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} \tilde{v}(\cos(t))\, \cos(kt)\, dt.$$

It is well known (see e.g. [14, 9]) that the series (21) converges in any elliptic disc $D_\varrho \subset \mathbb{C}$, with $\varrho > 1$, delimited by the ellipse

$$E_\varrho = \{z = t + is \in \mathbb{C},\ \ t = \frac{\varrho + \varrho^{-1}}{2} \cos\phi,\ \ s = \frac{\varrho - \varrho^{-1}}{2} \sin(\phi),\ \ \phi \in [0, 2\pi)\}$$

in which the function $\tilde{v}$ is analytic. Moreover (see [14] for details) we have

$$\|a_k\|_V \leq 2\varrho^{-k} \max_{z \in D_\varrho} \|\tilde{v}(z)\|_V.$$

If we denote by $\Pi_p v \in \mathcal{P}_p(\Gamma) \otimes V$ the truncated Chebyshev expansion up to the polynomial degree p and we observe that $|C_k(t)| \leq 1$, for all $t \in [-1,1]$, we have

$$\begin{aligned} \min_{w \in \mathcal{P}_p \otimes V} \|v - w\|_{C^0(\Gamma;V)} &\leq \|\tilde{v} - \Pi_p \tilde{v}\|_{C^0([-1,1];V)} \\ &\leq \sum_{k=p+1}^{\infty} \|a_k\|_V \leq \frac{2}{\varrho - 1} \varrho^{-p} \max_{z \in D_\varrho} \|\tilde{v}(z)\|_V. \end{aligned}$$

Finally, we have to link ϱ to the size of the analyticity region of $\tilde{v}$. It is easy to verify that the ellipse given by

$$\varrho = \frac{2\tau}{|\Gamma|}\left(1 + \sqrt{1 + \frac{|\Gamma|^2}{4\tau^2}}\right)$$

is the largest ellipse that can be drawn inside $\Sigma([-1,1]; 2\tau/|\Gamma|)$ and this proves the stated result. □

For the case of unbounded Γ we first recall a result given in [23] and then we state in Lemma 10 a result tuned to our situation.

We denote by $H_n(y) \in P_n(\mathbb{R})$ the normalized Hermite polynomials

$$H_n(y) = \sqrt{\pi^{\frac{1}{2}} 2^n n!}\,(-1)^n e^{y^2} \frac{\partial^n}{\partial y^n}\left(e^{-y^2}\right)$$

and by $h_n(y) = e^{-y^2/2} H_n(y)$ the Hermite functions. We recall that the Hermite polynomials form a complete orthonormal basis of the $L^2(\mathbb{R})$ space with respect to the weight e^{-y^2}, i.e.

$$\int_{\mathbb{R}} H_k(y) H_l(y) e^{-y^2}\, dy = \delta_{kl}.$$

Lemma 9 (Hille, 1940) *Let $f(z)$ be an analytic function in the strip of the complex plane $\Sigma(\mathbb{R};\tau) \equiv \{z = (y + iw) \in \mathbb{C}, -\tau \le w \le \tau\}$. A necessary and sufficient condition in order that the Fourier-Hermite series*

$$\sum_{k=0}^{\infty} f_k h_k(z), \qquad f_k = \int_{\mathbb{R}} f(y) h_k(y)\, dy, \tag{22}$$

shall exist and converge to the sum $f(z)$ in $\Sigma(\mathbb{R};\tau)$ is that to every β, $0 \le \beta < \tau$, there exists a finite positive $C(\beta)$ such that

$$|f(y + iw)| \le C(\beta) e^{-|y|\sqrt{\beta^2 - w^2}}, \quad -\infty < y < \infty, \;\; -\beta \le w \le \beta. \tag{23}$$

Moreover, the following bound for the Fourier coefficients hold

$$|f_n| \le C e^{-\tau\sqrt{2n+1}}. \tag{24}$$

In particular, the previous result tells us that, in order to have exponential decay of the Fourier coefficients f_n, we not only need $f(z)$ to be analytic in $\Sigma(\mathbb{R};\tau)$ but we have to require also that it decays on the real line, for $y \to \infty$, at least as $e^{-\tau|y|}$.

We introduce, now, two weights: the exponential one $\sigma = e^{-\alpha|y|}$, for some $\alpha > 0$ and the Gaussian one $G = e^{-(\delta y)^2/4}$. We recall that Lemma 7 holds for both of them. We will assume that the function v is in the space $C^0_\sigma(\Gamma; V)$, but we will measure the best approximation error in the weaker norm $C^0_G(\Gamma; V)$, with Gaussian weight, so that we can use the result from Hille given in Lemma 9. It holds:

Lemma 10 *Let v be a function in $C^0_\sigma(\mathbb{R};V)$. We suppose that v admits an analytic extension in the strip of the complex plane $\Sigma(\mathbb{R};\tau)=\{z\in\mathbb{C},\ \mathrm{dist}(z,\mathbb{R})\le\tau\}$ for some $\tau>0$, and*

$$\forall z=(y+iw)\in\Sigma(\mathbb{R};\tau),\quad \sigma(y)\|v(z)\|_V\le C_v(\tau).$$

Then, for any $\delta>0$, there exists a constant C, independent of p, and a function $\Theta(p)=O(\sqrt{p})$, such that

$$\min_{w\in\mathcal{P}_p\otimes V}\max_{y\in\mathbb{R}}\left|\|v(y)-w(y)\|_V e^{-\frac{(\delta y)^2}{4}}\right|\le C\Theta(p)e^{-\tau\delta\sqrt{p}}.$$

Proof. We introduce the change of variable $t=\delta\,y/\sqrt{2}$ and we denote by $\tilde{v}(t)=v(y(t))$. Observe that $\tilde{v}\in C^0_{\tilde{\sigma}}(\mathbb{R};V)$ with weight $\tilde{\sigma}=e^{-\sqrt{2}\frac{\alpha}{\delta}|t|}$. We consider the expansion of $\tilde{v}$ in Hermite polynomials

$$\tilde{v}(t)=\sum_{k=0}^{\infty}v_kH_k(t),\quad \text{where } v_k\in V,\quad v_k=\int_{\mathbb{R}}\tilde{v}(t)H_k(t)e^{-t^2}\,dt. \tag{25}$$

We set, now, $f(z)=\tilde{v}(z)e^{-\frac{z^2}{2}}$. Observe that the Hermite expansion of f as defined in (22) has the same Fourier coefficients as the expansion of $\tilde{v}$ defined in (25). Indeed

$$f_k=\int_{\mathbb{R}}f(t)h_k(t)\,dt=\int_{\mathbb{R}}\tilde{v}(t)H_k(t)e^{-t^2}\,dt=v_k.$$

Clearly, $f(z)$ is analytic in the strip $\Sigma(\mathbb{R};\frac{\tau\delta}{\sqrt{2}})$, being the product of analytic functions. Moreover,

$$\|f(y+iw)\|_V=|e^{-\frac{(y+iw)^2}{2}}|\|\tilde{v}(z)\|_V\le e^{-\frac{y^2-w^2}{2}}e^{\sqrt{2}\frac{\alpha}{\delta}|y|}C_v(\tau).$$

Setting

$$C(\beta)=\max_{\substack{-\infty<y<\infty\\ -\beta\le w\le\beta}}\exp\{-\frac{y^2-w^2}{2}+\sqrt{2}\frac{\alpha}{\delta}|y|+|y|\sqrt{\beta^2-w^2}\}$$

which is bounded for all $-\frac{\tau\delta}{\sqrt{2}}\le\beta\le\frac{\tau\delta}{\sqrt{2}}$, the function $f(z)$ satisfies the hypotheses of Lemma 9. Hence the Hermite series converges in $\Sigma(\mathbb{R};\frac{\tau\delta}{\sqrt{2}})$ and the Fourier coefficients v_k behave as in (24). We chose $w\in\mathcal{P}_p\otimes V$ as the truncated Hermite expansion of v, up to degree p: $\tilde{w}(t)=\Pi_p\tilde{v}(t)=\sum_{k=0}^{p}v_kH_k(t)$. We have

$$\begin{aligned}E_p(v)=\min_{w\in\mathcal{P}_p\otimes V}\max_{y\in\mathbb{R}}&\left|\|v(y)-w(y)\|_V e^{-\frac{(\delta y)^2}{4}}\right|\\ &\le\max_{t\in\mathbb{R}}\left|\|\tilde{v}(t)-\Pi_p\tilde{v}(t)\|_V e^{-\frac{t^2}{2}}\right|\le\max_{t\in\mathbb{R}}\|\sum_{k=p+1}^{\infty}v_kh_k(t)\|_V.\end{aligned}$$

It is well known (see e.g. [8]) that the Hermite functions $h_k(t)$ satisfy $|h_k(t)| < 1$ for all $t \in \mathbb{R}$ and all $k = 0, 1, \ldots$. Hence, the previous series can be bound as

$$E_p(v) \leq \sum_{k=p+1}^{\infty} \|v_k\|_V \leq C \sum_{k=p+1}^{\infty} e^{-\frac{\tau\delta}{\sqrt{2}}\sqrt{2k+1}}.$$

Lemma 14 in Appendix provides a bound for such a series and this concludes the proof. □

We are now ready to prove Theorem 1.

Proof. [of Theorem 1] The error naturally splits into $\epsilon = (u - u_h) + (u_h - u_{h,p})$. The first term depends on the space discretization only and can be estimated easily; indeed the function u_h is the orthogonal projection of u onto the subspace $L^2_\rho(\Gamma) \otimes H^1_0(D)$ with respect to the inner product $\int_{\Gamma\times D} \rho a |\nabla \cdot|^2$. Hence

$$\begin{aligned}\|u - u_h\|_{L^2_\rho(\Gamma)\otimes H^1_0(D)} &\leq \frac{1}{\sqrt{a_{min}}} \left(\int_{\Gamma\times D} \rho a |\nabla(u - u_h)|^2\right)^{\frac{1}{2}} \\ &\leq \frac{1}{\sqrt{a_{min}}} \inf_{v\in L^2_\rho(\Gamma)\otimes H_h(D)} \left(\int_{\Gamma\times D} \rho a |\nabla(u - v)|^2\right)^{\frac{1}{2}}\end{aligned}$$

The second term $u_h - u_{h,p}$ is an interpolation error. We recall, indeed, that $u_{h,p} = \mathcal{I}_p u_h$. To lighten the notation, we will drop the subscript h, being understood that we work on the semidiscrete solution. We recall, moreover, that u_h has the same regularity as the exact solution u w.r.t y.

To analyze this term we employ a one-dimensional argument. We first pass from the norm L^2_ρ to $L^2_{\hat\rho}$:

$$\|u - \mathcal{I}_p u\|_{L^2_\rho\otimes H^1_0} \leq \left\|\frac{\rho}{\hat\rho}\right\|^{\frac{1}{2}}_{L^\infty(\Gamma)} \|u - \mathcal{I}_p u\|_{L^2_{\hat\rho}\otimes H^1_0}$$

Here we adopt the same notation as in Section 3, namely we indicate with $\bullet_n$ a quantity relative to the direction y_n and $\bullet^*_n$ the analogous quantity relative to all other directions y_j, $j \neq n$. We focus on the first direction y_1 and define an interpolation operator $\mathcal{I}_1 : C^0_{\sigma_1}(\Gamma_1; L^2_{\hat\rho_1^*} \otimes H^1_0) \to L^2_{\hat\rho_1}(\Gamma_1; L^2_{\hat\rho_1^*} \otimes H^1_0)$,

$$\mathcal{I}_{p_1} v(y_1, y_1^*, x) = \sum_{k=1}^{p_1+1} v(y_{1,k}, y_1^*, x) l_{1,k}(y_1).$$

Then, the global interpolant $\mathcal{I}_p$ can be written as the composition of two interpolation operators $\mathcal{I}_p = \mathcal{I}_1 \circ \mathcal{I}^{(1)}_p$ where $\mathcal{I}^{(1)}_p$ is the interpolation operator in all directions $y_2, y_3, \ldots, y_N$ except y_1: $\mathcal{I}^{(1)}_p : C^0_{\sigma_1^*}(\Gamma_1^*; H^1_0) \to L^2_{\hat\rho_1^*}(\Gamma_1^*; H^1_0)$. We have, then

$$\|u - \mathcal{I}_p u\|_{L^2_{\hat\rho}\times H^1_0} \leq \underbrace{\|u - \mathcal{I}_1 u\|_{L^2_{\hat\rho}\times H^1_0}}_{\text{I}} + \underbrace{\|\mathcal{I}_1(u - \mathcal{I}^{(1)}_p u)\|_{L^2_{\hat\rho}\times H^1_0}}_{\text{II}}$$

Let us bound the first term. We think of u as a function of y_1 with values in a Banach space V, $u \in L^2_{\hat{\rho}_1}(\Gamma_1; V)$, where $V = L^2_{\hat{\rho}_1^*}(\Gamma_1^*) \otimes H^1_0(D)$. Under Assumption 2, in Section 3, and the choice of $\hat{\rho}$ given in (16), the following inclusions hold true

$$C^0_{\sigma_1}(\Gamma_1; V) \subset C^0_{G_1}(\Gamma_1; V) \subset L^2_{\hat{\rho}_1}(\Gamma_1; V)$$

with $\sigma_1 = G_1 = 1$ if Γ_1 is bounded and $\sigma_1 = e^{-\alpha_1|y_1|}$, $G_1 = e^{-\frac{(\delta_1 y_1)^2}{4}}$ if Γ_1 is unbounded. We know also from Lemma 6 that the interpolation operator $\mathcal{I}_1$ is continuous both as an operator from $C^0_{\sigma_1}(\Gamma_1; V)$ with values in $L^2_{\hat{\rho}_1}(\Gamma_1; V)$ and from $C^0_{G_1}(\Gamma_1; V)$ in $L^2_{\hat{\rho}_1}(\Gamma_1; V)$. In particular, we can estimate

$$\mathrm{I} = \|u - \mathcal{I}_1 u\|_{L^2_{\hat{\rho}_1}(\Gamma_1; V)} \leq C_2 \inf_{w \in \mathcal{P}_{p_1} \otimes V} \|u - w\|_{C^0_{G_1}(\Gamma; V)}.$$

To bound the best approximation error in $C^0_{G_1}(\Gamma; V)$, in the case Γ_1 bounded we use Lemma 8 whereas if Γ_1 is unbounded we employ Lemma 10 and the fact that $u \in C^0_{\sigma_1}(\Gamma_1; V)$ (see Lemma 4). In both cases, we need the analyticity result, for the solution u, stated in Lemma 5. Putting everything together, we can say that

$$\mathrm{I} \leq \begin{cases} Ce^{-r_1 p_1}, & \Gamma_1 \text{ bounded} \\ C\Theta(p_1)e^{-r_1\sqrt{p_1}}, & \Gamma_1 \text{ unbounded} \end{cases}$$

the value of r_1 being specified in Lemmas 8 and 10. To bound the term II, we use Lemma 6:

$$\mathrm{II} \leq C_1 \|u - \mathcal{I}^{(1)}_p u\|_{C^0_{\sigma_1}(\Gamma_1; V)}.$$

The term on the right hand side is again an interpolation error. So we have to bound the interpolation error in all the other $N-1$ directions, uniformly with respect to y_1 (in the weighted norm $C^0_{\sigma_1}$). We can proceed iteratively, defining an interpolation $\mathcal{I}_2$, bounding the resulting error in the direction y_2 and so on. □

4.1 Convergence of moments

In some cases one might be interested only in computing the first few moments of the solution, namely $E[u^m]$, $m = 1, 2, \ldots$. We show in the next two lemmas that the error in the first two moments, measured in a suitable spatial norm, is bounded by the *mean square error* $\|u - u_{h,p}\|_{L^2_\rho \otimes H^1_0}$, which, upon Theorem 1, is exponentially convergent with respect to the polynomial degree p employed in the probability directions. In particular, without extra regularity assumptions on the solution u of the problem, we have optimal convergence for the error in the mean value (first moment) measured in $L^2(D)$ or $H^1(D)$ and for the error in the second moment measures in $L^1(D)$.

Lemma 11 (approximation of mean value)

$$\|E[u - u_{h,p}]\|_{V(D)} \leq \|u - u_{h,p}\|_{L^2_\rho(\Gamma) \otimes V(D)}, \text{ with } V(D) = L^2(D) \text{ or } H^1(D).$$

The proof is immediate and omitted. Although the previous estimate implies exponential convergence with respect to p, under the assumptions of Theorem 1, the above estimate is suboptimal and can be improved by a duality argument (see [4] and Remark 5.2 from [5]).

Lemma 12 (approximation of the second moment)

$$\|E[u^2 - u_{h,p}^2]\|_{L^1(D)} \leq C\|u - u_{h,p}\|_{L^2_\rho(\Gamma)\otimes L^2(D)} \|u\|_{C^0_\sigma(\Gamma;H^1(D))}$$

with C independent of the discretization parameters h and p.

Proof. We have

$$\begin{aligned}\|E[u^2 - u_{h,p}^2]\|_{L^1(D)} &\leq \|E[(u - u_{h,p})(u + u_{h,p})]\|_{L^1(D)} \\ &\leq \|u - u_{h,p}\|_{L^2_\rho(\Gamma)\otimes L^2(D)}\|u + u_{h,p}\|_{L^2_\rho(\Gamma)\otimes L^2(D)} \\ &\leq \|u - u_{h,p}\|_{L^2_\rho(\Gamma)\otimes L^2(D)} \left(\|u\|_{L^2_\rho(\Gamma)\otimes L^2(D)} + \|u_{h,p}\|_{L^2_\rho(\Gamma)\otimes L^2(D)}\right).\end{aligned}$$

The term $\|u_{h,p}\|_{L^2_\rho\otimes L^2}$ can be bounded as

$$\|u_{h,p}\|_{L^2_\rho(\Gamma)\otimes L^2(D)} = \|\mathcal{I}_p u_h\|_{L^2_\rho(\Gamma)\otimes L^2(D)} \leq C_1\|u_h\|_{C^0_\sigma(\Gamma;L^2(D))} \leq C\|u\|_{C^0_\sigma(\Gamma;H^1(D))}$$

where we have used the boundedness of the interpolation operator $\mathcal{I}_p$ stated in Lemma 6. The last inequality follows from the fact that the semidiscrete solution u_h is the orthogonal projection of the exact solution u onto the subspace H_h with respect to the *energy* inner product, hence

$$\|\sqrt{a(y)}\nabla u_h(y)\|_{L^2(D)} \leq \|\sqrt{a(y)}\nabla u(y)\|_{L^2(D)}, \quad \forall y \in \Gamma$$

and the energy norm is equivalent to the H^1 norm. □

Similarly, it is possible to estimate the approximation error in the covariance function of the solution u.

On the other hand, to estimate the convergence rate of the error in higher order moments, or of the second moment in higher norms, we need extra regularity assumptions on the solution to ensure proper integrability and then be able to use analyticity.

5 Numerical Examples

This section illustrates the convergence of the collocation method for a stochastic elliptic problem in two dimensions. The computational results are in accordance with the convergence rate predicted by the theory.

The problem to solve is

$$\begin{aligned}-\nabla\cdot(a\nabla u) &= 0, \; on\, \Omega \times D, \\ u &= 0, \; on\, \Omega \times \partial D_D, \\ -a\partial_n u &= 1, \; on\, \Omega \times \partial D_N \\ \partial_n u &= 0, \; on\, \Omega \times (\partial D - (\partial D_D \cup \partial D_N)),\end{aligned}$$

with

$$D = \{(x,z) \in \mathbb{R}^2 : -1.5 \le x \le 0,\ -0.4 \le z \le 0.8\},$$
$$\partial D_D = \{(x,z) \in \mathbb{R}^2 : -1 \le x \le -0.5,\ z = 0.8\},$$
$$\partial D_N = \{(x,z) \in \mathbb{R}^2 :\ -1.5 \le x \le 0,\ z = -0.4\},$$

cf. Figure 1.

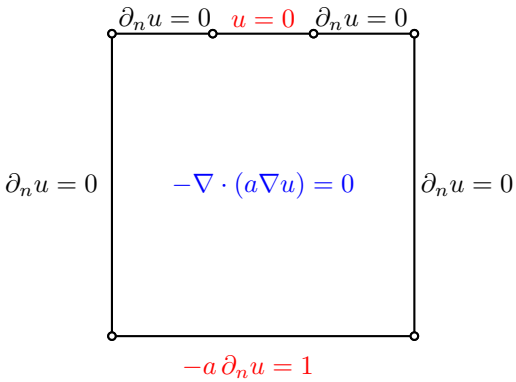

Figure 1: Geometry and boundary conditions for the numerical example.

The random diffusivity coefficient is a nonlinear function of the random vector Y, namely

$$\begin{aligned} a(\omega, x) =& a_{min} + \\ & \exp\Big\{ \left[Y_1(\omega) \cos(\pi z) + Y_3(\omega) \sin(\pi z) \right] e^{-\frac{1}{8}} + \\ & \left[Y_2(\omega) \cos(\pi x) + Y_4(\omega) \sin(\pi x) \right] e^{-\frac{1}{8}} \Big\}. \end{aligned} \tag{26}$$

Here $a_{min} = 1/100$ and the real random variables Y_n, $n = 1, \ldots, 4$ are independent and identically distributed with mean value zero and unit variance. To illustrate on the behavior of the collocation method with either unbounded or bounded random variables Y_n, this section presents two different cases, corresponding to either Gaussian or Uniform densities. The corresponding collocation points are then cartesian products determined by the roots of either Hermite or Legendre polynomials.

Observe that the collocation method only requires the solution of uncoupled deterministic problems in the collocation points, also in presence of a diffusivity coefficient which depends non-linearly on the random variables as in (26). This is a great advantage with respect to the classical Stochastic-Galerkin finite element

method as considered in [4] or [29] (see also the considerations given in Section 2.1). Observe, moreover, how easily the Collocation method can deal with random variables with unbounded support.

Figure 2 shows some realizations of the logarithm of the diffusivity coefficient while Figures 3 and 4 show the mean and variance of the corresponding solutions.

The finite element space for spatial discretization is the span of continuous functions that are piecewise polynomials with degree five over a triangulation with 1178 triangles and 642 vertices, see Figure 5. This triangulation has been adaptively graded to control the singularities at the boundary points $(-1, 0.8)$ and $(-0.5, 0.8)$. These singularities occur where the Dirichlet and Neumann boundaries meet and they essentially behave like $\sqrt{r}$, with r being the distance to the closest singularity point.

To study the convergence of the tensor product collocation method we increase the order p for the approximating polynomial spaces, $\mathcal{P}_p(\Gamma)$, following the adaptive algorithm described on page 1287 of the work [5]. This adaptive algorithm increases the tensor polynomial degree with an anisotropic strategy: it increases the order of approximation in one direction as much as possible before considering the next direction.

The computational results for the $H_0^1(D)$ approximation error in the expected value, $E[u]$, are shown on Figure 6 while those corresponding to the approximation of the second moment, $E[u^2]$, are shown on Figure 7. To estimate the computational error in the i-th direction, corresponding to a multi index $p = (p_1, \ldots, p_i, \ldots, p_N)$, we approximate it by $E[e] \approx E[u_{h,p} - u_{h,\tilde{p}}]$, with $\tilde{p} = (p_1, \ldots, p_i + 1, \ldots, p_N)$. We proceed similarly for the error in the approximation of the second moment.

As expected, the estimated approximation error decreases exponentially fast as the polynomial order increases, for both the computation of $E[u]$ and $E[u^2]$, with either Gaussian or Uniform probability densities.

6 Conclusions

In this work we have proposed a Collocation method for the solution of elliptic partial differential equations with random coefficients and forcing terms. This method has the advantages of: leading to uncoupled deterministic problems also in case of input data which depend non-linearly on the random variables; treating efficiently the case of non independent random variables with the introduction of an auxiliary density $\hat{\rho}$; dealing easily with random variables with unbounded support, such as Gaussian or exponential ones; dealing with no difficulty with a diffusivity coefficient a with unbounded second moment.

We have provided a full convergence analysis and proved exponential convergence "in probability" for a broad range of situations. The theoretical result is given in Theorem 1 and confirmed numerically by the tests presented in Section 5.

The method is very versatile and very accurate for the class of problems considered (as accurate as the Stochastic Galerkin approach). It leads to the

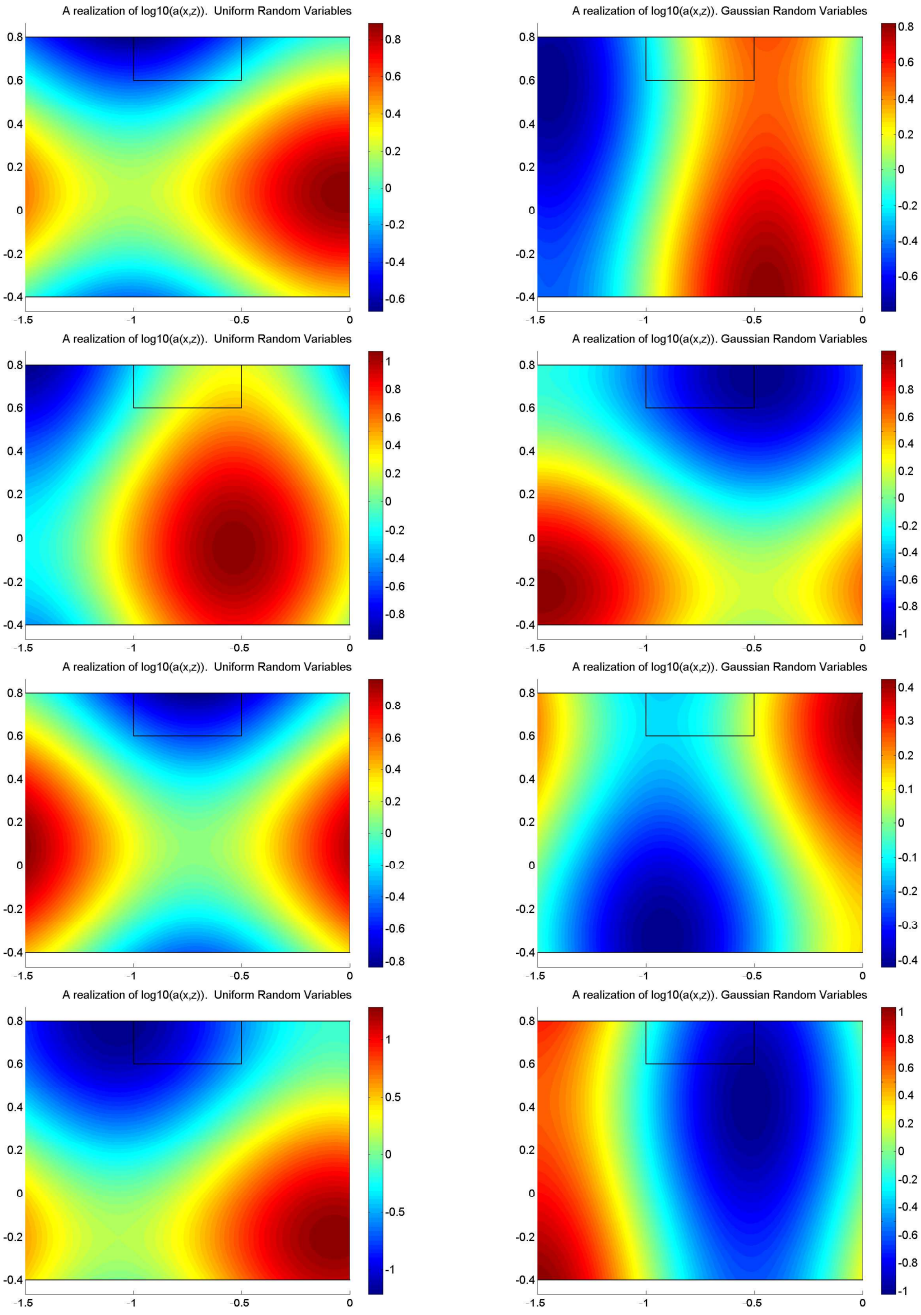

Figure 2: Some realizations of $\log(a)$.

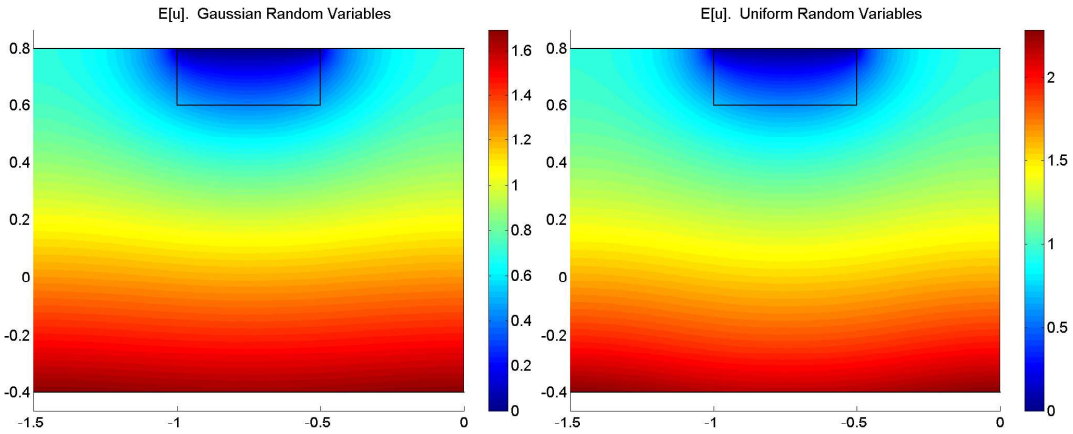

Figure 3: Results for the computation of the expected value for the solution, $E[u]$.

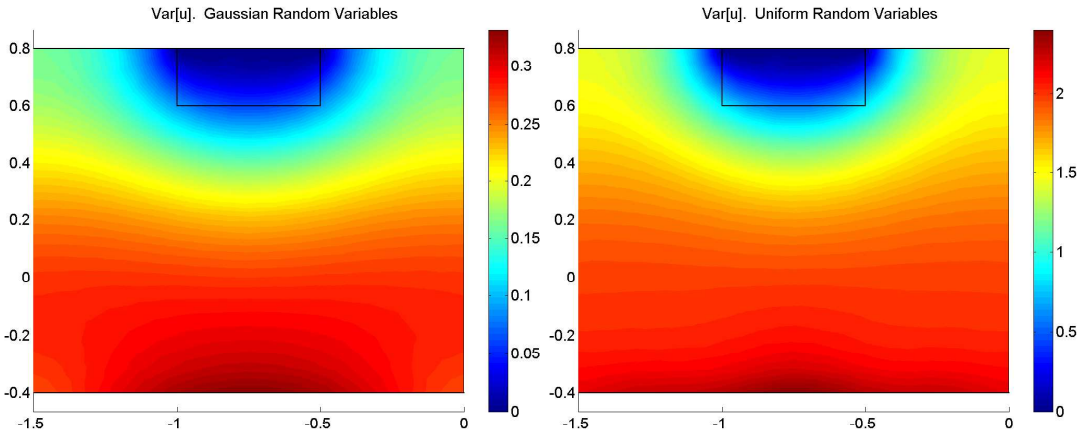

Figure 4: Results for the computation of the variance of the solution, $Var[u]$.

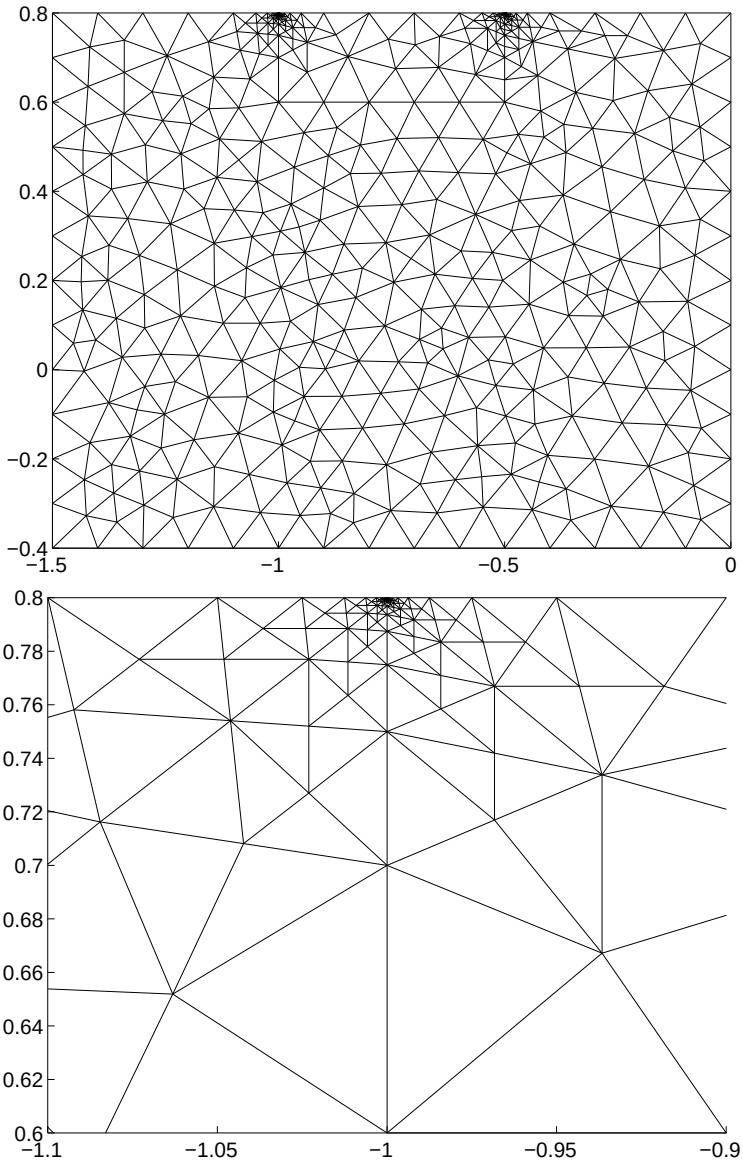

Figure 5: Top: Unstructured grid for the spatial discretization. The corresponding finite element spaces are the span of continuous functions that are piecewise polynomials with degree five. Bottom: Detail of the mesh refinement near the left singularity.

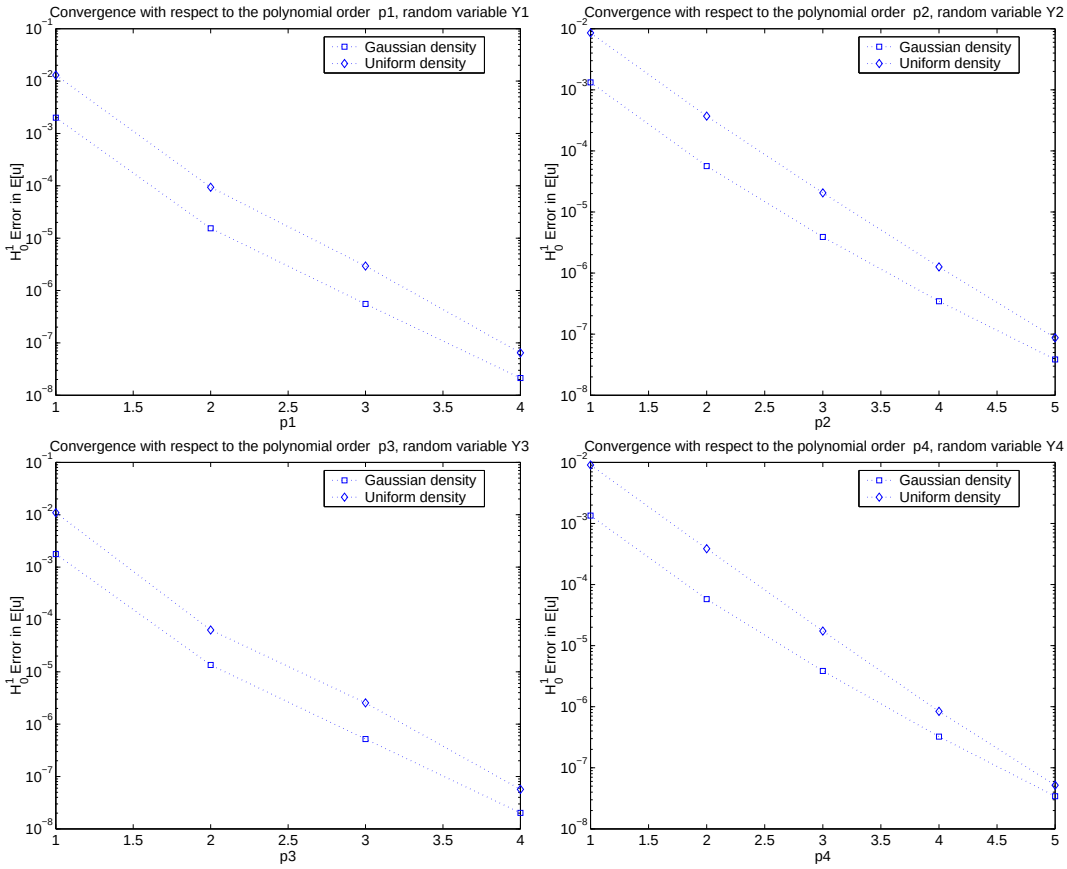

Figure 6: Convergence results for the approximation of the expected value, $E[u]$.

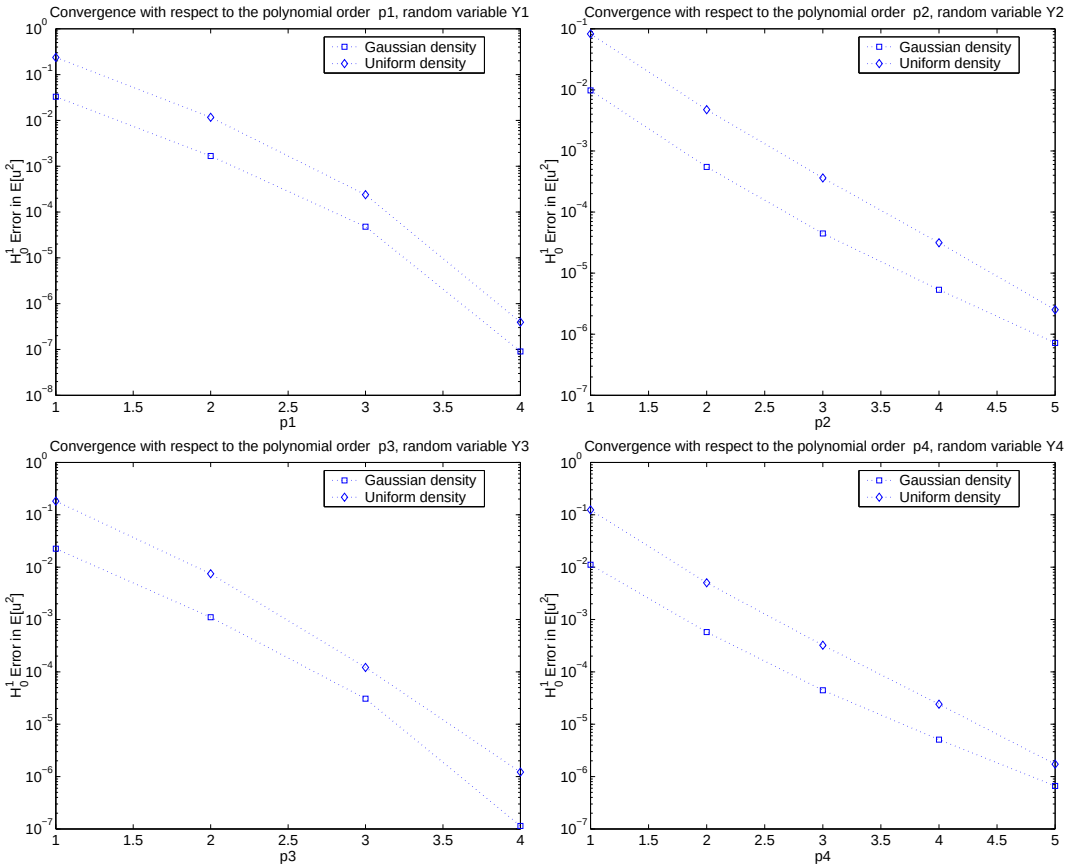

Figure 7: Convergence results for the approximation of the second moment, $E[u^2]$.

solution of uncoupled deterministic problems and, as such, is fully parallelizable like a Monte Carlo method. The extension of the analysis to other classes of linear and non-linear problems is an ongoing research.

The use of tensor product polynomials suffers from the *curse of dimensionality*. Hence, this method is efficient only for a small number of random variables. For a moderate or large dimensionality of the probability space one should rather turn to *sparse tensor product spaces*. This aspect will be investigated in a future work.

Acknowledgments

The first author was partially supported by the Sandia National Lab (Contract Number 268687) and the Office of Naval Research (Grant N00014-99-1-0724). The second author was partially supported by M.U.R.S.T. Cofin 2004 Grant "Metodi Numerici Avanzati per Equazioni alle Derivate Parziali di Interesse Applicativo" and a "J.T. Oden Visiting Faculty Fellowship". The third author was partially supported by the the ICES Postdoctoral Fellowship Program, UdelaR in Uruguay and the European Network HYKE (HYperbolic and Kinetic Equations: Aymptotics, Numerics, Analysis), funded by the EC (Contract HPRN-CT-2002-00282).

Appendix

Lemma 13 *Let $r \in \mathbb{R}^+$, $r < 1$. Then*

- $\displaystyle\sum_{k=0}^{n}(2k+1)r^k = \frac{1}{(1-r)^2}\left\{1+r-r^{n+1}\left[(2n+1)(1-r)+2\right]\right\}$
- $\displaystyle\sum_{k=n+1}^{\infty}(2k+1)r^k = r^{n+1}\frac{(2n+1)(1-r)+2}{(1-r)^2}$

Proof. We use the summation by part formula

$$\sum_{k=0}^{n} f_k g_k = f_n G_n - \sum_{k=0}^{n-1} G_k(f_{k+1}-f_k), \qquad G_k = \sum_{j=0}^{k} g_j.$$

with $f_k = (2k+1)$, $g_k = r^k$ and $G_k = (1-r^{k+1})/(1-r)$. Then

$$\begin{aligned}\sum_{k=0}^{n}(2k+1)r^k &= (2n+1)\frac{1-r^{n+1}}{1-r} - \sum_{k=0}^{n-1} 2\frac{1-r^{k+1}}{1-r} \\ &= (2n+1)\frac{1-r^{n+1}}{1-r} - \frac{2}{1-r}\left[n - r\frac{1-r^n}{1-r}\right] \\ &= \frac{1}{1-r}\left[(2n+1) - (2n+1)r^{n+1} - 2n + 2r\frac{1-r^n}{1-r}\right] \\ &= \frac{1}{1-r}\left\{1 + \frac{2r}{1-r} - r^{n+1}\left[(2n+1) + \frac{2}{1-r}\right]\right\}\end{aligned}$$

which gives the first result. Clearly,

$$\sum_{k=0}^{\infty}(2k+1)r^k = \frac{1+r}{(1-r)^2}.$$

Then, computing the tail series as

$$\sum_{k=n+1}^{\infty}(2k+1)r^k = \sum_{k=0}^{\infty}(2k+1)r^k - \sum_{k=0}^{n}(2k+1)r^k$$

we obtain easily the second result as well. □

Lemma 14 *Let $r \in \mathbb{R}^+$, $r < 1$. Then*

$$\sum_{k=n+1}^{\infty} r^{\sqrt{2k+1}} \leq \left[\frac{2\sqrt{n+1}}{a(1-a)} + O(1)\right] a^{\sqrt{n}}, \qquad a = r^{\sqrt{2}}.$$

Proof. We start bounding

$$\sum_{k=n+1}^{\infty} r^{\sqrt{2k+1}} \leq \sum_{k=n+1}^{\infty} r^{\sqrt{2k}} = \sum_{k=n+1}^{\infty} a^{\sqrt{k}}.$$

Let us observe, now, that

$$\sum_{k=n+1}^{\infty} a^{\sqrt{k}} \leq \sum_{k=[\sqrt{n+1}]}^{\infty}(2k+1)a^k$$

where we have denoted by $[v]$ the integer part of a real number v. Then, using the result from Lemma 13, we have

$$\sum_{k=[\sqrt{n+1}]}^{\infty}(2k+1)a^k \leq a^{[\sqrt{n+1}]}\frac{(2[\sqrt{n+1}]-1)(1-a)+2}{(1-a)^2}.$$

Observing, now, that $\sqrt{n+1} - 1 \leq [\sqrt{n+1}] \leq \sqrt{n+1} + 1$, we obtain

$$\sum_{k=n+1}^{\infty} a^{\sqrt{k}} \leq a^{\sqrt{n+1}} \frac{(2\sqrt{n+1}+1)(1-a)+2}{a(1-a)^2},$$

which leads immediately to the final result. □

References

[1] M. Ainsworth and J-T. Oden. *A posteriori Error Estimation in Finite Element Analysis*. Wiley, 2000.

[2] Ivo Babuška, Kang-Man Liu, and Raúl Tempone. Solving stochastic partial differential equations based on the experimental data. *Math. Models Methods Appl. Sci.*, 13(3):415–444, 2003. Dedicated to Jim Douglas, Jr. on the occasion of his 75th birthday.

[3] Ivo Babuška and Theofanis Strouboulis. *The finite element method and its reliability*. Numerical Mathematics and Scientific Computation. The Clarendon Press Oxford University Press, New York, 2001.

[4] Ivo Babuška, Raúl Tempone, and Georgios E. Zouraris. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM J. Numer. Anal.*, 42(2):800–825 (electronic), 2004.

[5] Ivo Babuška, Raúl Tempone, and Georgios E. Zouraris. Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation. *Comput. Methods Appl. Mech. Engrg.*, 194(12-16):1251–1294, 2005.

[6] I. Babuška, F. Nobile, and R. Tempone. Worst-case scenario analysis for elliptic problems with uncertainty. *Numer. Math.*, 101:185–219, 2005.

[7] Volker Barthelmann, Erich Novak, and Klaus Ritter. High dimensional polynomial interpolation on sparse grids. *Adv. Comput. Math.*, 12(4):273–288, 2000. Multivariate polynomial interpolation.

[8] John P. Boyd. Asymptotic coefficients of Hermite function series. *J. Comput. Phys.*, 54(3):382–410, 1984.

[9] John P. Boyd. *Chebyshev and Fourier spectral methods*. Dover Publications Inc., Mineola, NY, second edition, 2001.

[10] Malte Braack and Alexandre Ern. A posteriori control of modeling errors and discretization errors. *Multiscale Model. Simul.*, 1(2):221–238 (electronic), 2003.

[11] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer–Verlag, New York, 1994.

[12] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland, New York, 1978.

[13] Manas K. Deb, Ivo M. Babuška, and J. Tinsley Oden. Solution of stochastic partial differential equations using Galerkin finite element techniques. *Comput. Methods Appl. Mech. Engrg.*, 190(48):6359–6372, 2001.

[14] Ronald A. DeVore and George G. Lorentz. *Constructive approximation*, volume 303 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1993.

[15] Howard C. Elman, Oliver G. Ernst, Dianne P. O'Leary, and Michael Stewart. Efficient iterative algorithms for the stochastic finite element method with application to acoustic scattering. *Comput. Methods Appl. Mech. Engrg.*, 194(9-11):1037–1055, 2005.

[16] B. Engquist, P. Löstedt, and O. Runborg, editors. *Multiscale Methods in Science and Engineering*, volume 44 of *Lecture Notes in Computetional Science and Engineering*. Springer, 2005.

[17] P. Erdös and P. Turan. On interpolation. I. Quadrature- and mean-convergence in the Lagrange-interpolation. *Ann. of Math. (2)*, 38(1):142–155, 1937.

[18] Philipp Frauenfelder, Christoph Schwab, and Radu Alexandru Todor. Finite elements for elliptic problems with stochastic coefficients. *Computer Methods in Applied Mechanics and Engineering*, 194(2-5):205–228, 2005.

[19] Daniele Funaro and Otared Kavian. Approximation of some diffusion evolution equations in unbounded domains by Hermite functions. *Math. Comp.*, 57(196):597–619, 1991.

[20] Roger Ghanem. Ingredients for a general purpose stochastic finite elements implementation. *Comput. Methods Appl. Mech. Engrg.*, 168(1-4):19–34, 1999.

[21] Roger G. Ghanem and Pol D. Spanos. *Stochastic finite elements: a spectral approach*. Springer-Verlag, New York, 1991.

[22] Mircea Grigoriu. *Stochastic calculus*. Birkhäuser Boston Inc., Boston, MA, 2002. Applications in science and engineering.

[23] Einar Hille. Contributions to the theory of Hermitian series. II. The representation problem. *Trans. Amer. Math. Soc.*, 47:80–94, 1940.

[24] J. Hlaváček, I. Chleboun, and I. Babuška. *Uncertain input data problems and the worst scenario method*. Elsevier, Amsterdam, 2004.

[25] S. Larsen. *Numerical analysis of elliptic partial differential equations with stochastic input data*. PhD thesis, University of Maryland, 1986.

[26] O. P. Le Maître, O. M. Knio, H. N. Najm, and R. G. Ghanem. Uncertainty propagation using Wiener-Haar expansions. *J. Comput. Phys.*, 197(1):28–57, 2004.

[27] Michel Loève. *Probability theory*. Springer-Verlag, New York, fourth edition, 1977. Graduate Texts in Mathematics, Vol. 45 and 46.

[28] Lionel Mathelin, M. Yousuff Hussaini, and Thomas A. Zang. Stochastic approaches to uncertainty quantification in CFD simulations. *Numer. Algorithms*, 38(1-3):209–236, 2005.

[29] Hermann G. Matthies and Andreas Keese. Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations. *Comput. Methods Appl. Mech. Engrg.*, 194(12-16):1295–1331, 2005.

[30] J. T. Oden and S. Prudhomme. Estimation of modeling error in computational mechanics. *Journal of Computational Physics*, 182:496–515, 2002.

[31] J. Tinsley Oden and Kumar S. Vemaganti. Estimation of local modeling error and goal-oriented adaptive modeling of heterogeneous materials. I. Error estimates and adaptive algorithms. *J. Comput. Phys.*, 164(1):22–47, 2000.

[32] Bernt Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003. An introduction with applications.

[33] Alfio Quarteroni and Alberto Valli. *Numerical approximation of partial differential equations*, volume 23 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1994.

[34] C. Schwab and R.A. Todor. Stochastic galerkin fem for the numerical solution of elliptic pdes with stochastic coefficients. in preparation.

[35] M.A. Tatang. *Direct incorporation of uncertainty in chemical and environmental engineering systems*. PhD thesis, MIT, 1995.

[36] J. V. Uspensky. On the convergence of quadrature formulas related to an infinite interval. *Trans. Amer. Math. Soc.*, 30(3):542–559, 1928.

[37] R. Verfürth. *A review of a posteriori error estimation and adaptive mesh refinement techniques*. Wiley-Teubner, 1996.

[38] N. Wiener. The homogeneous chaos. *Amer. J. Math.*, 60:897–936, 1938.

[39] C. L. Winter and D. M. Tartakovsky. Groundwater flow in heterogeneous composite aquifers. *Water Resour. Res.*, 38(8):23.1 (doi:10.1029/2001WR000450), 2002.

[40] D. Xiu and J.S. Hesthaven. High order collocation methods for differential equations with random inputs. submitted to SIAM J. Sci. Comput.

[41] Dongbin Xiu and George Em Karniadakis. Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos. *Comput. Methods Appl. Mech. Engrg.*, 191(43):4927–4948, 2002.

[42] Dongbin Xiu and George Em Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24(2):619–644 (electronic), 2002.