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

F. Nobile* R. Tempone† C. G. Webster‡

June 22, 2006

Abstract

This work proposes and analyzes a sparse grid stochastic collocation method for solving elliptic partial differential equations with random coefficients and forcing terms (input data of the model). This method can be viewed as an extension of the Stochastic Collocation method proposed in [Babuška-Nobile-Tempone, Technical report, MOX, Dipartimento di Matematica, 2005] which consists of a Galerkin approximation in space and a collocation at the zeros of suitable tensor product orthogonal polynomials in probability space and naturally leads to the solution of uncoupled deterministic problems as in the Monte Carlo method. The full tensor product spaces suffer from the *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 moderately large, this work proposes the use of sparse tensor product spaces utilizing either Clenshaw-Curtis or Gaussian interpolants. For both situations this work provides rigorous convergence analysis of the fully discrete problem and demonstrates: (sub)-exponential convergence of the "probability error" in the asymptotic regime and algebraic convergence of the "probability error" in the pre-asymptotic regime, with respect to the total number of collocation points. The problem setting in which this procedure is recommended as well as suggestions for future enhancements to the method are discussed. Numerical examples exemplify the theoretical results and show the effectiveness of the method.

Key words: Collocation techniques, stochastic PDEs, finite elements, uncertainty quantification, sparse grids, Smolyak algorithm, multivariate polynomial interpolation.

AMS subject classification: 65N30, 65N35, 65N12, 65N15, 65C20

Introduction

Mathematical modeling and computer simulations are nowadays widely used tools to predict the behavior of physical and engineering problems. Whenever a particular application is considered, the mathematical models need to be equipped with input data, such as coefficients, forcing terms, boundary conditions, geometry, etc.

However, in many applications, such input data may be affected by a relatively large amount of uncertainty. This can be due to an intrinsic variability in the physical system as, for instance,

*MOX, Dipartimento di Matematica, Politecnico di Milano, Italy.

†School of Computational Science and Department of Mathematics, Florida State University, Tallahassee FL 32306-4120.

‡School of Computational Science and Department of Mathematics, Florida State University, Tallahassee FL 32306-4120.

in the mechanical properties of many bio-materials, polymeric fluids, or composite materials, the action of wind or seismic vibrations on civil structures, etc.

In other situations, uncertainty may come from our difficulty in characterizing accurately the physical system under investigation as in the study of groundwater flows, where the subsurface properties such as porosity and permeability in an aquifer have to be extrapolated from measurements taken only in few spatial locations.

Such uncertainties can be included in the mathematical model adopting a probabilistic setting, provided enough information is available for a complete statistical characterization of the physical system. In this framework, the input data are modeled as *random variables*, like in the case where the input coefficients are piecewise constant and random over fixed subdomains, or more generally, as *random fields* with a given spatial (or temporal) correlation structure.

Therefore, the goal of the mathematical and computational analysis becomes the prediction of statistical moments (mean value, variance, covariance, etc.) or even the whole probability distribution of some responses of the system (quantities of physical interest), given the probability distribution of the input random data.

A random field can often be expanded as an infinite combination of random variables by means, for instance of the so called Karhunen-Loève [23] or Polynomial Chaos (PC) expansions [34, 38]. Although it is properly described only by means of an infinite number of random variables, whenever the realizations are slowly varying in space, with a correlation length comparable to the size of the domain, only a few terms in the above mentioned expansion are typically needed to describe the random field with sufficient accuracy. Therefore, for this type of application, it is reasonable to limit the analysis to just a few random variables in the expansion (see e.g. [2, 16]).

In this work we focus on elliptic partial differential equations whose coefficients and forcing terms are described by a finite dimensional random vector (*finite dimensional noise assumption*, cf. Section 1.1), either because the problem itself can be described by a finite number of random variables or because the input coefficients are modeled as truncated random fields.

The most popular approach to solve mathematical problems in a probabilistic setting is the Monte Carlo method (see e.g. [15] and references therein). The Monte Carlo method is easy to implement and allows one to reuse available deterministic codes. Yet, the convergence rate is typically very slow, although with a mild dependence on the number on sampled random variables.

In the last few years, other approaches have been proposed, which in certain situations feature a much faster convergence rate. We mention, among others, the Spectral Galerkin method [4, 5, 16, 19, 22, 25, 29, 37], Stochastic Collocation [3, 24, 31, 36], perturbation methods or Neumann expansions [1, 17, 20, 32, 35].

For certain classes of problems, the solution may have a very regular dependence on the input random variables. For instance, it was shown in [3] and [4] that the solution of a linear elliptic PDE with diffusivity coefficient and/or forcing term described as truncated expansions of random fields is analytic in the input random variables. In such situations, Spectral Galerkin or Stochastic Collocation methods based on orthogonal tensor product polynomials feature a very fast convergence rate.

In particular, our earlier work [3] proposed a Stochastic Collocation/Finite Element method based on standard finite element approximations in space and a collocation on a tensor grid built upon the zeros of orthogonal polynomials with respect to the joint probability density function of the input random variables. It was shown that for an elliptic PDE the error converges exponentially fast with respect to the number of points employed in each probability direction.

The Stochastic Collocation can be easily implemented and leads naturally to the solution of uncoupled deterministic problems as for the Monte Carlo method, even in presence of input data which depend nonlinearly on the driving random variables. It can also treat efficiently the case

of non independent random variables with the introduction of an auxiliary density and handle for instance cases with lognormal diffusivity coefficient, which is not bounded in $\Omega \times D$ but that has bounded realizations.

When the number of input random variables is small, Stochastic Collocation is a very effective numerical tool.

On the other hand, approximations based on tensor product spaces suffer from the *curse of dimensionality* since the number of collocation points in a tensor grid grows exponentially fast in the number of input random variables.

If the number of random variables is moderately large, one should rather consider sparse tensor product spaces as first proposed by Smolyak [30] and further investigated by e.g. [6,16,18,36], which will be the primary focus of this paper. The use of sparse grids allows one to reduce dramatically the number of collocation points, while preserving a high level of accuracy.

Motivated by the above, this work analyzes a sparse grid Stochastic Collocation method for solving elliptic partial differential equations whose coefficients or forcing terms are described through a finite number of random variables. The sparse tensor product grids are built upon either Clenshaw-Curtis or Gaussian abscissas. For both situations this work provides a rigorous convergence analysis of the fully discrete problem and demonstrates (sub)-exponential convergence of the "probability error" in the asymptotic regime and algebraic convergence of the "probability error" in the pre-asymptotic regime, with respect to the total number of collocation points used in the sparse grid. This work also addresses the case where the input random variables come from suitably truncated expansions of random fields and discusses how the size of the sparse grid can be algebraically related to the number of random variables retained in the expansion in order to have a discretization error of the same order as that of the error due to the truncation of the input random fields.

The problem setting in which the sparse grid Stochastic Collocation method is recommended as well as suggestions for future enhancements to the method are discussed.

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 focus on applications to linear elliptic PDEs with random input data. In Section 3 we provide an overview of various collocation techniques and describe the sparse approximation method to be considered as well as the different interpolation techniques to be employed. In Section 4 we provide a complete error analysis of the method considered. Finally, in Section 5 we present some numerical results showing the effectiveness of the proposed method.

1 Problem setting

We begin by focusing our attention on an elliptic operator $\mathcal{L}$, linear or nonlinear, on a domain $D \subset \mathbb{R}^d$, which depends on some coefficients $a(\omega, x)$ with $x \in D$, $\omega \in \Omega$ and $(\Omega, \mathcal{F}, P)$ 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. Similarly the forcing term $f = f(\omega, x)$ can be assumed random as well.

Consider the stochastic 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:

$$\mathcal{L}(a)(u) = f \quad \text{in } D \tag{1.1}$$

equipped with suitable boundary conditions. Before introducing some assumptions we denote by $W(D)$ a Banach space of functions $v : D \to \mathbb{R}$ and define, for $q \in [1, \infty]$, the stochastic Banach

spaces

$$L_P^q(\Omega; W(D)) = \left\{ v : \Omega \to W(D) \mid v \text{ is strongly measurable and } \int_\Omega \|v(\omega,\cdot)\|_{W(D)}^q dP(\omega) < +\infty \right\}$$

and

$$L_P^\infty(\Omega; W(D)) = \left\{ v : \Omega \to W(D) \mid v \text{ is strongly measurable and } P - \operatorname*{ess\,sup}_{\omega\in\Omega} \|v(\omega,\cdot)\|_{W(D)}^2 < +\infty \right\}.$$

Of particular interest is the space $L_P^2(\Omega; W(D))$, consisting of Banach valued functions that have finite second moments.

We will now make the following assumptions:

A_1) the solution to (1.1) has realizations in the Banach space $W(D)$, i.e. $u(\cdot,\omega) \in W(D)$ almost surely and $\forall \omega \in \Omega$

$$\|u(\cdot,\omega)\|_{W(D)} \le C \|f(\cdot,\omega)\|_{W^*(D)}$$

where we denote $W^*(D)$ to be the dual space of $W(D)$, and C is a constant independent of the realization $\omega \in \Omega$.

A_2) the forcing term $f \in L_P^2(\Omega; W^*(D))$ is such that the solution u is unique and bounded in $L_P^2(\Omega; W(D))$.

Here we give two example problems that are posed in this setting:

Example 1.1 *The linear problem*

$$\begin{cases} -\nabla \cdot (a(\omega,\cdot)\nabla u(\omega,\cdot)) = f(\omega,\cdot) \text{ in } \Omega \times D, \\ u(\omega,\cdot) = 0 \text{ on } \Omega \times \partial D, \end{cases} \tag{1.2}$$

with $a(\omega,\cdot)$ uniformly bounded and coercive, i.e.

$$\text{there exists } a_{min}, a_{max} \in (0,+\infty) \text{ such that } P(\omega \in \Omega : \ a(\omega,x) \in [a_{min}, a_{max}] \, \forall x \in \overline{D}) = 1$$

and $f(\omega,\cdot)$ square integrable with respect to P, satisfies assumptions A_1 and A_2 with $W(D) = H_0^1(D)$ (see [3]).

Example 1.2 *Similarly, for $k \in \mathbb{N}^+$, the nonlinear problem*

$$\begin{cases} -\nabla \cdot (a(\omega,\cdot)\nabla u(\omega,\cdot)) + u(\omega,\cdot)|u(\omega,\cdot)|^k = f(\omega,\cdot) \quad \text{in } \Omega \times D, \\ u(\omega,\cdot) = 0 \quad \text{on } \Omega \times \partial D, \end{cases} \tag{1.3}$$

with $a(\omega,\cdot)$ uniformly bounded and coercive and $f(\omega,\cdot)$ square integrable with respect to P, satisfies assumptions A_1 and A_2 with $W(D) = H_0^1(D) \cap L^{k+2}(D)$ (see [26]).

1.1 On Finite Dimensional Noise

In some applications, the coefficient a and the forcing term f appearing in (1.1) can be described by a random vector $[Y_1, \ldots, Y_N] : \Omega \to \mathbb{R}^N$, as in the following examples. In such cases, we will emphasize such dependence by writing a_N and f_N.

Example 1.3 (Piecewise constant random fields) *Let us consider again problem* (1.2) *where the physical domain D is the union of non-overlapping subdomains D_i, $i = 1, \ldots, N$. We consider a diffusion coefficient that is piecewise constant and random on each subdomain, i.e.*

$$a_N(\omega, x) = a_{min} + \sum_{i=1}^{N} \sigma_i \, Y_i(\omega) 1_{D_i}(x).$$

Here 1_{D_i} is the indicator function of the set D_i, σ_i, a_{min} are positive constants, and the random variables Y_i are nonnegative with unit variance.

In other applications the coefficients and forcing terms in (1.1) may have other type of spatial variation that is amenable to describe by an expansion. Depending on the decay of such expansion and the desired accuracy in our computations we may retain just the first N terms.

Example 1.4 (Karhunen-Loève expansion) *We recall that any second order random field $g(\omega, x)$, with continuous covariance function $cov[g] : \overline{D \times D} \to \mathbb{R}$, can be represented as an infinite sum of random variables, by means, for instance, of a Karhunen-Loève expansion [23]. For mutually uncorrelated real random variables $\{Y_i(\omega)\}_{i=1}^{\infty}$ with zero mean and unit variance, i.e. $E[Y_i] = 0$ and $E[Y_i Y_j] = \delta_{ij}$ for $i, j \in \mathbb{N}_+$ we let*

$$g(\omega, x) = E[g](x) + \sum_{i=1}^{\infty} \sqrt{\lambda_i} \, b_i(x) \, Y_i(\omega)$$

where $\{\lambda_i\}_{i=1}^{\infty}$ is a sequence of non-negative decreasing eigenvalues and $\{b_i\}_{i=1}^{\infty}$ the corresponding sequence of orthonormal eigenfunctions satisfying

$$T_g b_i = \lambda_i b_i, \quad (b_i, b_j)_{L^2(D)} = \delta_{ij} \;\; \textit{for} \;\; i, j \in \mathbb{N}_+.$$

The compact and self-adjoint operator $T_g : L^2(D) \to L^2(D)$ is defined by

$$T_g v(\cdot) = \int_D cov[g](x, \cdot) \, v(x) \, dx \quad \forall v \in L^2(D).$$

The truncated Karhunen-Loève expansion g_N, of the stochastic function g, is defined by

$$g_N(\omega, x) = E[g](x) + \sum_{i=1}^{N} \sqrt{\lambda_i} \, b_i(x) \, Y_i(\omega) \quad \forall N \in \mathbb{N}_+.$$

The infinite random variables are uniquely determined by

$$Y_i(\omega) = \frac{1}{\sqrt{\lambda_i}} \int_D (g(\omega, x) - E[g](x)) \, b_i(x) \, dx.$$

Then by Mercer's theorem (cf [28, p. 245]), it follows that

$$\lim_{N \to \infty} \left\{ \sup_D E\left[(g - g_N)^2 \right] \right\} = \lim_{N \to \infty} \left\{ \sup_D \left(\sum_{i=N+1}^{\infty} \lambda_i b_i^2 \right) \right\} = 0.$$

Observe that the N random variables in (1.4)*, describing the random data, are then weighted differently due to the decay of the eigen-pairs of the Karhunen-Loève expansion. The decay of eigenvalues and eigenvectors has been investigated e.g. in the works [16] and [32].*

The above examples motivate us to consider problems whose coefficients are described by finitely many random variables. Thus, we will seek a random field $u_N : \Omega \times \overline{D} \to \mathbb{R}$, such that a.s., the following equation holds:

$$\mathcal{L}(a_N)(u_N) = f_N \quad \text{in } D, \tag{1.4}$$

We assume that equation (1.4) admits a unique solution $u_N \in L^2_P(\Omega; W(D))$. We then have, by the Doob–Dynkin's lemma (cf. [27]), that the solution u_N of the stochastic elliptic boundary value problem (1.4) can be described by $u_N = u_N(\omega, x) = u_N(Y_1(\omega), \ldots, Y_N(\omega), x)$. We underline that the coefficients a_N and f_N in (1.4) may be an exact representation of the input data as in Example 1.3 or a suitable truncation of the input data as in Example 1.4. In the latter case, the solution u_N will also be an approximation of the exact solution u in (1.1) and the truncation error $u - u_N$ has to be properly estimated, see section 4.2.

Remark 1.5 (Nonlinear coefficients) *In certain cases, one may need to ensure qualitative properties on the coefficients a_N and f_N and may be worth while to describe them as nonlinear functions of Y. For instance, in Example 1.1 one is required to enforce positiveness on the coefficient $a_N(\omega, x)$, say $a_N(\omega, x) \geq a_{min}$ for all $x \in D$, a.s. in Ω. Then a better choice is to expand $\log(a_N - a_{min})$. The following standard transformation guarantees that the diffusivity coefficient is bounded away from zero almost surely*

$$\log(a_N - a_{min})(\omega, x) = b_0(x) + \sum_{1 \leq n \leq N} \sqrt{\lambda_n} b_n(x) Y_n(\omega), \tag{1.5}$$

i.e. one performs a Karhunen-Loève expansion for $\log(a_N - a_{min})$, assuming that $a_N > a_{min}$ almost surely. On the other hand, the right hand side of (1.4) can be represented as a truncated Karhunen-Loève expansion

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

Remark 1.6 *It is usual to have f_N and a_N independent, because the forcing terms and the parameters in the operator $\mathcal{L}$ are seldom related. In such a situation we have $a_N(Y(\omega), x) = a_N(Y_a(\omega), x)$ and $f_N(Y(\omega), x) = f_N(Y_f(\omega), x)$, with $Y = [Y_a, Y_f]$ and the vectors Y_a, Y_f are independent.*

For this work we denote $\Gamma_n \equiv Y_n(\Omega)$ the image of Y_n, where we assume $Y_n(\omega)$ to be bounded. Without loss of generality we can assume $\Gamma_n = [-1, 1]$. We also let $\Gamma^N = \prod_{n=1}^N \Gamma_n$ and assume that the random variables $[Y_1, Y_2, \ldots, Y_n]$ have a joint probability density function $\rho : \Gamma^N \to \mathbb{R}_+$, with $\rho \in L^\infty(\Gamma^N)$. Thus, the goal is to approximate the function $u_N = u_N(y, x)$, for any $y \in \Gamma^N$ and $x \in \overline{D}$. (see [3], [4])

Remark 1.7 (Unbounded Random Variables) *By using a similar approach to the work [3] we can easily deal with unbounded random variables, such as Gaussian or exponential ones. For the sake of simplicity in the presentation we focus our study on bounded random variables only.*

1.2 Regularity

Before discussing various collocation techniques and going through the convergence analysis of such methods, we need to state some regularity assumptions on the data of the problem and consequent regularity results for the exact solution u_N. We will perform a one-dimensional analysis in each direction y_n, $n = 1, \ldots, N$. 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^*. We require the solution to problem (1.1) to satisfy the following estimate:

Assumption 1.8 *For each $y_n \in \Gamma_n$, there exists $\tau_n > 0$ such that the function $u_N(y_n, y_n^*, x)$ as a function of y_n, $u_N : \Gamma_n \to C^0(\Gamma_n^*; W(D))$ admits an analytic extension $u(z, y_n^*, x)$, $z \in \mathbb{C}$, in the region of the complex plane*

$$\Sigma(\Gamma_n; \tau_n) \equiv \{z \in \mathbb{C}, \quad dist(z, \Gamma_n) \leq \tau_n\}. \tag{1.6}$$

Moreover, $\forall z \in \Sigma(\Gamma_n; \tau_n)$,

$$\|u_N(z)\|_{C^0(\Gamma_n^*; W(D))} \leq \lambda \tag{1.7}$$

with λ a constant independent of n.

The previous assumption should be verified for each particular application. In particular, this has implications on the allowed regularity of the input data, e.g. coefficients, loads, etc., of the stochastic PDE under study. In the next section we recall some theoretical results, which were proved in [3, Section 3], for the linear problem introduced in Example 1.1.

2 Applications to linear elliptic PDEs with random input data

In this section we give more details concerning the linear problem described in Example 1.1. Problem (1.2) can be written in a weak form as: find $u \in L_P^2(\Omega; H_0^1(D))$ such that

$$\int_D E[a\nabla u \cdot \nabla v]\, dx = \int_D E[fv]\, dx \quad \forall\, v \in L_P^2(\Omega; H_0^1(D)). \tag{2.1}$$

A straightforward application of the Lax-Milgram theorem allows one to state the well posedness of problem (2.1). Moreover, the following a priori estimates hold

$$\|u\|_{H_0^1(D)} \leq \frac{C_P}{a_{min}} \|f(\omega, \cdot)\|_{L^2(D)} \quad \text{a.s.} \tag{2.2}$$

and

$$\|u\|_{L_P^2(\Omega; H_0^1(D))} \leq \frac{C_P}{a_{min}} \left(\int_D E[f^2]\, dx \right)^{1/2}, \tag{2.3}$$

where C_P denotes the constant appearing in the Poincaré inequality:

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

Once we have the input random fields described by a finite set of random variables, i.e. $a(\omega, x) = a_N(Y_1(\omega), \ldots, Y_N(\omega), x)$, and similarly for $f(\omega, x)$, the "finite dimensional" version of the stochastic variational formulation (2.1) has a "deterministic" equivalent which is the following: find $u_N \in L_\rho^2(\Gamma^N; H_0^1(D))$ such that

$$\int_{\Gamma^N} \rho\, (a_N \nabla u_N, \nabla v)_{L^2(D)}\, dy = \int_{\Gamma^N} \rho\, (f_N, v)_{L^2(D)}\, dy, \quad \forall\, v \in L_\rho^2(\Gamma^N; H_0^1(D)). \tag{2.4}$$

Observe that in this work the gradient notation, ∇, always means differentiation with respect to $x \in D$ only, unless otherwise stated. The stochastic boundary value problem (2.1) 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_N as a function $u_N : \Gamma^N \to H_0^1(D)$ and we use the notation $u_N(y)$ whenever we want to highlight the dependence on

the parameter y. We use similar notations for the coefficient a_N and the forcing term f_N. Then, it can be shown that problem (2.1) is equivalent to

$$\int_D a_N(y)\nabla u_N(y)\cdot\nabla\phi\,dx = \int_D f_N(y)\phi\,dx, \quad \forall\phi\in H_0^1(D), \;\; \rho\text{-a.e. in } \Gamma^N. \tag{2.5}$$

For our convenience, we will suppose that the coefficient a_N and the forcing term f_N admit a smooth extension on the ρ-zero measure sets. Then, equation (2.5) can be extended a.e. in Γ^N with respect to the Lebesgue measure (instead of the measure ρdy).

It has been proved in [3] that problem (2.5) satisfies the analyticity result stated in Assumption 1.8. For instance, if we take the diffusivity coefficient as in Example 1.3 and a deterministic load the size of the analyticity region is given by

$$\tau_n = \frac{a_{min}}{4\sigma_n}. \tag{2.6}$$

On the other hand, if we take the diffusivity coefficient as a truncated expansion like in Remark 1.5, then the analyticity region $\Sigma(\Gamma_n;\tau_n)$ is given by

$$\tau_n = \frac{1}{4\sqrt{\lambda_n}\|b_n\|_{L^\infty(D)}} \tag{2.7}$$

Observe that, in the latter case, as $\sqrt{\lambda_n}\|b_n\|_{L^\infty(D)}\to 0$ for a regular enough covariance function (see [16]) the analyticity region increases as n increases. This fact introduces, naturally, an anisotropic behavior with respect to the "direction" n. This effect will not be exploited in the numerical methods proposed in the next sections but is the subject of ongoing research.

3 Collocation techniques

We seek a numerical approximation to the exact solution of (1.4) in a suitable finite dimensional subspace. To describe such a subspace properly, we introduce some standard approximation subspaces, namely:

- $W_h(D)\subset W(D)$ 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 $h>0$. We suppose that W_h has the following deterministic approximability property: for a given function $\varphi\in W(D)$,

$$\min_{v\in W_h(D)}\|\varphi - v\|_{W(D)} \le C(s;\varphi)\,h^s, \tag{3.1}$$

  where s is a positive integer determined by the smoothness of φ and the degree of the approximating finite element subspace and $C(s;\varphi)$ is independent of h.

  Example 3.1 *Let D be a convex polygonal domain and $W(D)=H_0^1(D)$. For piecewise linear finite element subspaces we have*

$$\min_{v\in W_h(D)}\|\varphi - v\|_{H_0^1(D)} \le c\,h\,\|\varphi\|_{H^2(D)}.$$

  That is, $s=1$ and $C(s;\varphi)=\|\varphi\|_{H^2(D)}$, see for example [8].

We will also assume that there exists a finite element operator $\pi_h : W(D) \to W_h(D)$ with the optimality condition

$$\|\varphi - \pi_h\varphi\|_{W(D)} \leq C_\pi \min_{v \in W_h(D)} \|\varphi - v\|_{W(D)}, \quad \forall \varphi \in W(D), \tag{3.2}$$

where the constant C_π is independent of the mesh size h.

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

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

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

Stochastic collocation entails the sampling of approximate values $\pi_h u_N(y_k) = u_h^N(y_k) \in W_h(D)$, to the solution u_N of (1.4) on a suitable set of abscissas $y_k \in \Gamma^N$.

Example 3.2 *If we examine the linear PDE for example, then we introduce the semi-discrete approximation $u_h^N : \Gamma^N \to W_h(D)$, obtained by projecting equation (2.5) onto the subspace $W_h(D)$, for each $y \in \Gamma^N$, i.e.*

$$\int_D a_N(y) \nabla u_h^N(y) \cdot \nabla \phi_h \, dx = \int_D f_N(y) \phi_h \, dx, \quad \forall \phi_h \in W_h(D), \;\; \textit{for a.e. } y \in \Gamma^N. \tag{3.3}$$

Notice that the finite element functions $u_h^N(y)$ satisfy the optimality condition (3.2), for all $y \in \Gamma^N$.

Then the construction of a fully discrete approximation, $u_{h,p}^N \in C^0(\Gamma^N; W_h(D))$, is based on a suitable interpolation of the sampled values. That is

$$u_{h,p}^N(y, \cdot) = \sum_k u_h^N(y_k, \cdot) l_k^p(y), \tag{3.4}$$

where, for instance, the functions l_k^p can be taken as the Lagrange polynomials (see Section 3.1 and 3.2).

This formulation can be used to compute the mean value or variance of u, as:

$$E[u](x) \approx \overline{u}_h^N \equiv \sum_k u_h^N(y_k, x) \int_{\Gamma^N} l_k^p(y) \rho(y) dy$$

and

$$\text{Var}[u](x) \approx \sum_k \left(u_h^N(y_k, x) - \overline{u}_h^N\right)^2 \int_{\Gamma^N} l_k^p(y) \rho(y) dy.$$

Several choices are possible for the interpolation points. We will discuss two of them, namely Clenshaw-Curtis and Gaussian in Sections 3.2.1 and 3.2.2 respectively. Regardless of the choice of interpolating knots, the interpolation can be constructed by using either full tensor product polynomials, see Section 3.1, or the space of sparse polynomials, see Section 3.2.

3.1 Full tensor product interpolation

In this section we briefly recall interpolation based on Lagrange polynomials. Let $i \in \mathbb{N}_+$ and $\{y_1^i, \ldots, y_{m_i}^i\} \subset [-1,1]$ be a sequence of abscissas for Lagrange interpolation on $[-1,1]$.

For $u \in C^0(\Gamma^1; W(D))$ and $N = 1$ we introduce a sequence of one-dimensional Lagrange interpolation operators $\mathscr{U}^i : C^0(\Gamma^1; W(D)) \to V_{m_i}(\Gamma^1; W(D))$

$$\mathscr{U}^i(u)(y) = \sum_{j=1}^{m_i} u(y_j^i) \cdot l_j^i(y), \quad \forall u \in C^0(\Gamma^1; W(D)), \tag{3.5}$$

where $l_j^i \in \mathcal{P}_{m_i-1}(\Gamma^1)$ are Lagrange polynomials of degree $p_i = m_i - 1$ and

$$V_m(\Gamma^1; W(D)) = \left\{ v \in C^0(\Gamma^1; W(D)) \,:\, v(y,x) = \sum_{k=1}^{m} \widetilde{v}_k(x) l_k(y),\ \{\widetilde{v}_k\}_{k=1}^m \in W(D) \right\}.$$

Here of course we have, for $i \in \mathbb{N}_+$,

$$l_j^i(y) = \prod_{\substack{k=1 \\ k \neq j}}^{m_i} \frac{(y - y_k^i)}{(y_j^i - y_k^i)}$$

and formula (3.5) reproduces exactly all polynomials of degree less than m_i. Now, in the multivariate case $N > 1$, for each $u \in C^0(\Gamma^N; W(D))$ and the multi-index $\mathbf{i} = (i_1, \ldots, i_N) \in \mathbb{N}_+^N$ we define the full tensor product interpolation formulas

$$\mathscr{I}_{\mathbf{i}}^N u(y) = \left(\mathscr{U}^{i_1} \otimes \cdots \otimes \mathscr{U}^{i_N}\right)(u)(y) = \sum_{j_1=1}^{m_{i_1}} \cdots \sum_{j_N=1}^{m_{i_N}} u\left(y_{j_1}^{i_1}, \ldots, y_{j_N}^{i_N}\right) \cdot \left(l_{j_1}^{i_1} \otimes \cdots \otimes l_{j_N}^{i_N}\right). \tag{3.6}$$

Clearly, the above product needs $(m_{i_1} \cdots m_{i_N})$ function values, sampled on a grid. These formulas will also be used as the building blocks for the Smolyak method, described next.

3.2 The Smolyak method

Here we follow closely the work [7] and describe the Smolyak *i*sotropic formulas $\mathscr{A}(q,N)$. The Smolyak formulas are just linear combinations of product formulas (3.6) with the following key properties: only products with a relatively small number of knots are used and the linear combination is chosen in such a way that an interpolation property for $N = 1$ is preserved for $N > 1$. With $\mathscr{U}^0 = 0$ define

$$\Delta^i = \mathscr{U}^i - \mathscr{U}^{i-1} \tag{3.7}$$

for $i \in \mathbb{N}_+$. Moreover, we put $|\mathbf{i}| = i_1 + \cdots + i_N$ for $\mathbf{i} = (i_1, i_2, \ldots, i_N) \in \mathbb{N}_+^N$. Then the Smolyak algorithm is given by

$$\mathscr{A}(q,N) = \sum_{|\mathbf{i}| \leq q} \left(\Delta^{i_1} \otimes \cdots \otimes \Delta^{i_N}\right) \tag{3.8}$$

for integers $q \geq N$. Equivalently, formula (3.8) can be written as (see [33])

$$\mathscr{A}(q,N) = \sum_{q-N+1 \leq |\mathbf{i}| \leq q} (-1)^{q-|\mathbf{i}|} \binom{N-1}{q-|\mathbf{i}|} \cdot \left(\mathscr{U}^{i_1} \otimes \cdots \otimes \mathscr{U}^{i_N}\right). \tag{3.9}$$

To compute $\mathscr{A}(q,N)(u)$, one only needs to know function values on the "sparse grid"

$$\mathscr{H}(q,N) = \bigcup_{q-N+1\leq|\mathbf{i}|\leq q} \left(\vartheta^{i_1}\times\cdots\times\vartheta^{i_N}\right) \tag{3.10}$$

where $\vartheta^i = \left\{y_1^i,\ldots,y_{m_i}^i\right\}\subset[-1,1]$ denotes the set of points used by $\mathscr{U}^i$. Note that the Smolyak algorithm, as presented, is isotropic and we will later discuss possible improvements that can be made to further reduce the number of points used to compute $\mathscr{U}^i$.

3.2.1 Clenshaw-Curtis Formulas

We first suggest to use the Smolyak algorithm based on polynomial interpolation at the extrema of Chebyshev polynomials. For any choice of $m_i > 1$ these knots are given by

$$y_j^i = -\cos\left(\frac{\pi(j-1)}{m_i-1}\right), \quad j=1,\ldots,m_i. \tag{3.11}$$

In addition, we define $y_1^i = 0$ if $m_i = 1$. It remains to specify the numbers m_i of knots that are used in formulas $\mathscr{U}^i$. In order to obtain nested sets of points, i.e., $\vartheta^i\subset\vartheta^{i+1}$ and thereby $\mathscr{H}(q,N)\subset\mathscr{H}(q+1,N)$, we choose

$$m_1 = 1 \text{ and } m_i = 2^{i-1}+1, \text{ for } i>1. \tag{3.12}$$

For such a choice of m_i we arrive at Clenshaw-Curtis formulas, see [10]. It is important to choose $m_1 = 1$ if we are interested in optimal approximation in relatively large N, because in all other cases the number of points used by $\mathscr{A}(q,N)$ increases too fast with N.

A variant of the Clenshaw-Curtis formulas are the Filippi formulas in which the abscissas at the boundary of the interval are omitted [18]. In either case the degree m_i-1 of exactness is obtained.

3.2.2 Gaussian formulas

We also propose to apply the Smolyak formulas based on polynomial interpolation at the zeros of the orthogonal polynomials with respect to a weight ρ. This naturally leads to the Gauss formulas that have a maximum degree of exactness of $2m_i-1$. However, these Gauss-Legendre formulas are in general not nested. Regardless, as in the Clenshaw-Curtis case, we choose

$$m_1 = 1 \text{ and } m_i = 2^{i-1}+1, \text{ for } i>1.$$

The natural choice of the weight ρ should be the probability density of the random variables $Y_i(\omega)$ for all i. Yet, in the general multivariate case, if the random variables Y_i are not independent, the density ρ does not factorize, i.e.

$$\rho(y_1,\ldots,y_n)\neq\prod_{n=1}^{N}\rho_n(y_n).$$

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

$$\hat{\rho}(y_1,\ldots,y_n)=\prod_{n=1}^{N}\hat{\rho}_n(y_n),\ \ \forall y\in\Gamma^N, \qquad \text{and is such that} \quad \left\|\frac{\rho}{\hat{\rho}}\right\|_{L^\infty(\Gamma^N)}<\infty. \tag{3.13}$$

For each dimension $n=1,\ldots,N$ let the m_n Gaussian abscissas be the roots of the m_n degree polynomial that is $\hat{\rho}_n$-orthogonal to all polynomials of degree m_n-1 on the interval $[-1,1]$.

4 Error analysis

Collocation methods can be used to approximate the solution $u_N \in C^0(\Gamma^N; W(D))$ using finitely many function values. By Assumption 1.8, u_N admits an analytic extension. Further, each function value will be computed by means of a finite element technique. We define the numerical approximation $u_{h,p}^N = \mathscr{A}(q, N)\pi_h u_N$. Our aim is to give a priori estimates for the total error

$$\epsilon = u - u_{h,p}^N = u - \mathscr{A}(q, N)\pi_h u_N$$

where the operator $\mathscr{A}(q, N)$ is described by (3.8) and π_h is the finite element projection operator described by (3.2). We will investigate the error

$$\|u - \mathscr{A}(q, N)\pi_h u_N\| \leq \underbrace{\|u - u_N\|}_{(I)} + \underbrace{\|u_N - \pi_h u_N\|}_{(II)} + \underbrace{\|\pi_h(u_N - \mathscr{A}(q, N)u_N)\|}_{(III)} \tag{4.1}$$

evaluated in the natural norm $L_P^2(\Omega; W(D))$. Since the error functions in (II) and (III) are finite dimensional the natural norm is equivalent to $L_\rho^2(\Gamma^N; W(D))$. By controlling the error in this natural norm we also control the error in the expected value of the solution, for example:

$$\left\|E[u - u_{h,p}^N]\right\|_{W(D)} \leq E\left[\left\|u - u_{h,p}^N\right\|_{W(D)}\right] \leq \left\|u - u_{h,p}^N\right\|_{L_P^2(\Omega; W(D))}.$$

The quantity (I) controls the truncation error for the case where the input data a_N and f_N are suitable truncations of random fields. This contribution to the total error will be considered in Section 4.2. The quantity (I) is otherwise zero if the representation of a_N and f_N is exact, as in Example 1.3. The second term (II) controls the convergence with respect to h, i.e. the finite element error, which will be dictated by standard approximability properties of the finite element space $W_h(D)$, given by (3.1), and the regularity in space of the solution u (see e.g. [8,9]). Specifically,

$$\|u_N - \pi_h u_N\|_{L_\rho^2(\Gamma^N; W(D))} \leq C_\pi h^s \left(\int_{\Gamma^N} C(s; u)^2 \rho(y)\, dy\right)^{1/2}$$

by the finite element approximability property (3.1).

The full tensor product convergence results are given by [3, Theorem 1] and therefore, we will only concern ourselves with the convergence results when implementing the Smolyak algorithm described in Section 3.2. Namely, our primary concern will be to analyze the interpolation error (III)

$$\|\pi_h(u_N - \mathscr{A}(q, N)u_N)\|_{L_\rho^2(\Gamma^N; W(D))} \leq C_\pi \|u_N - \mathscr{A}(q, N)u_N\|_{L_\rho^2(\Gamma^N; W(D))}$$

where C_π is defined by the finite element optimality condition (3.2). Hence, in the next sections we estimate the interpolation error

$$\|u_N - \mathscr{A}(q, N)u_N\|_{L_\rho^2(\Gamma^N; W(D))},$$

for both the Clenshaw-Curtis and Gaussian versions of the Smolyak algorithm.

4.1 Analysis of the interpolation error

There are techniques to get error bounds for Smolyak's algorithm for $N > 1$ from those for the case $N = 1$. Therefore, we first address the case $N = 1$. Let us first recall the best approximation error for a function $v : \Gamma^1 \to W(D)$ which admits an analytic extension in the region $\Sigma(\Gamma^1; \tau) = \{z \in$

$\mathbb{C},\ \ \mathrm{dist}(z,\Gamma) < \tau\}$ of the complex plane, for some $\tau > 0$. We will still denote the extension by v; in this case, τ represents the distance between $\Gamma^1 \subset \mathbb{R}$ and the nearest singularity of $v(z)$ in the complex plane. Since we assume $\Gamma^1 = [-1, 1]$ and hence bounded, we present the following result, whose proof can be found in [3, Lemma 7] and which is an immediate extension of the result given in [12, Chapter 7, Section 8]:

Lemma 4.1 *Given a function $v \in C^0(\Gamma^1; W(D))$ which admits an analytic extension in the region of the complex plane $\Sigma(\Gamma^1; \tau) = \{z \in \mathbb{C},\ \ dist(z, \Gamma^1) \leq \tau\}$ for some $\tau > 0$, there holds*

$$E_{m_i} \equiv \min_{w \in V_{m_i}} \|v - w\|_{C^0(\Gamma^1; W(D))} \leq \frac{2}{\varrho - 1} e^{-m_i \log(\varrho)} \max_{z \in \Sigma(\Gamma^1;\tau)} \|v(z)\|_{W(D)}$$

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

Remark 4.2 (Approximation with unbounded random variables) *A related result with weighted norms holds for unbounded random variables whose probability density decays as the Gaussian density at infinity (see [3]).*

In the multidimensional case, the size of the analyticity region will depend, in general, on the direction n and it will be denoted by τ_n, as in (2.7). The same holds for the decay coefficient ϱ_n. In what follows, we set

$$\varrho \equiv \min_n \varrho_n. \tag{4.2}$$

As stated in Section 3.2, the Smolyak construction treats all directions equally and is therefore an isotropic algorithm. Moreover, the convergence analysis presented in Sections 4.1.1 and 4.1.2 does not exploit possible anisotropic behaviors of problem (1.1). Therefore, we can expect a slower convergence rate for such problems that exhibit strong anisotropic effects. See Section 5 where we explore numerically the consequences of introducing an anisotropy into the model problem described by Example 1.1.

Example 4.3 *For the linear problem described in Section 2 it was shown in the work [3] that for a multi-index $\mathbf{p} = (p_1, \ldots, p_N)$, a tensor product polynomial interpolation on Gaussian abscissas achieves exponential convergence in each direction Y_n and the error can be bounded as*

$$\|u_N - I^N_{\mathbf{p}} u_N\|_{L^2_\rho(\Gamma^N; W(D))} \leq C \sum_{n=1}^{N} {\varrho_n}^{-p_N}. \tag{4.3}$$

The constant C in (4.3) is independent of N and, using (2.7), we have

$$\begin{aligned} \varrho_n &= \frac{2\tau_n}{|\Gamma_n|} + \sqrt{1 + \frac{4{\tau_n}^2}{|\Gamma_n|^2}} \\ &\geq 1 + \frac{2\tau_n}{|\Gamma_n|}. \end{aligned} \tag{4.4}$$

where τ_n can be estimated e.g. as in (2.6) and (2.7).

4.1.1 Clenshaw-Curtis interpolation estimates

In this section we develop error estimates for interpolating functions $u \in C^0(\Gamma^N; W(D))$ that admit an analytic extension as described by Assumption 1.8 using the Smolyak formulations based on the choice (3.11) and (3.12) described in Section 3.2.1. We remind the reader that in the global estimate (4.1) we need to bound the interpolation error (III) in the $L^2_\rho(\Gamma^N; W(D))$ norm. Yet, this norm is always bounded by the $L^\infty(\Gamma^N; W(D))$. Namely, for all $v \in L^\infty(\Gamma^N; W(D))$ we have

$$\|v\|_{L^2_\rho(\Gamma^N;W(D))} \leq \|v\|_{L^\infty(\Gamma^N;W(D))}.$$

In our notation the norm $\|\cdot\|_{\infty,N}$ is shorthand for $\|\cdot\|_{L^\infty(\Gamma^N;W(D))}$ and will be used henceforth. We also define $I_N : \Gamma^N \to \Gamma^N$ as the identity operator on an N-dimensional space.

We begin by letting E_m be the error of the best approximation to functions $u \in C^0(\Gamma^1; W(D))$ by functions $w \in V_m$. Similarly to [7], since $\mathscr{U}^i$ is exact on V_{m_i-1} we can apply the general formula

$$\left\|u - \mathscr{U}^i(u)\right\|_{\infty,1} \leq E_{m_i-1}(u) \cdot (1 + \Lambda_{m_i}) \tag{4.5}$$

where Λ_m is the Lebesgue constant for our choice (3.11). It is known that

$$\Lambda_m \leq \frac{2}{\pi}\log(m-1) + 1 \tag{4.6}$$

for $m \geq 2$, see [13].

Using Lemma 4.1, the best approximation to functions $u \in C^0(\Gamma^1; W(D))$ that admit an analytic extension as described by Assumption 1.8 is bounded by:

$$E_{m_i}(u) \leq C\,\varrho^{-m_i} \tag{4.7}$$

where C is a constant dependent on τ defined in Lemma 4.1. Hence (4.5)-(4.7) implies

$$\left\|(I_1 - \mathscr{U}^i)(u)\right\|_{\infty,1} \leq C\,\log(m_i)\varrho^{-m_i} \leq C\,i\varrho^{-2^i},$$

$$\left\|(\Delta^i)(u)\right\|_{\infty,1} = \left\|(\mathscr{U}^i - \mathscr{U}^{i-1})(u)\right\|_{\infty,1} \leq \left\|(I_1 - \mathscr{U}^i)(u)\right\|_{\infty,1} + \left\|(I_1 - \mathscr{U}^{i-1})(u)\right\|_{\infty,1} \leq E\,i\varrho^{-2^{i-1}}$$

for all $i \in \mathbb{N}_+$ with positive constants C and E depending on u but not on i.

Lemma 4.4 *For functions $u \in L^2_\rho(\Gamma^N; W(D))$ that admit an analytic extension as described by Assumption 1.8 we obtain*

$$\|(I_N - \mathscr{A}(q,N))\,(u)\|_{L^2_\rho(\Gamma^N;W(D))} \leq CF^N\,\Psi(q,N)\varrho^{-\frac{p(q,N)}{2}} \tag{4.8}$$

where

$$p(q,N) := \begin{cases} N\,2^{q/N}, & \text{if } q > N\chi, \\ (q-N)\log(2)\cdot 2^\chi, & \text{otherwise,} \end{cases}, \tag{4.9}$$

$$\Psi(q,N) := \begin{cases} 1 & \text{if } N = 1 \\ \min\left\{q^{2N-1}, q^3\,e^{q^2}\right\} & \text{otherwise,} \end{cases} \tag{4.10}$$

and $\chi = \left(\frac{1+\log(2)}{\log(2)}\right)$.

Proof. First we define, for $j \geq 1$ and $s \geq d$, the two functions

$$f(s,j) := \frac{1}{2}\left(j2^{s/j} + 2^{q-N+j+2-s}\right)$$

and

$$g(s,j) := \sum_{i\in\mathbb{N}_+,|\mathbf{i}|=s} \prod_{n=1}^{j} i_n.$$

We begin by claiming that

$$\|(I_N - \mathscr{A}(q,N))(u)\|_{\infty,N} \leq C\sum_{j=1}^{N-1} E^j \sum_{s=j}^{q-N+j} g(s,j)\,(q-N+j+1-s)\,\varrho^{-f(s,j)} + \|(I_1 - \mathscr{A}(q-N+1,1))(u)\|_{\infty,N} \quad (4.11)$$

where, for the trivial case, we get

$$\|(I_1 - \mathscr{A}(q-N+1,1))(u)\|_{\infty,N} = \left\|(I_1 - \mathscr{U}^{q-N+1})(u)\right\|_{\infty,N} \leq C(q-N+1)\varrho^{-2^{q-N+1}} \leq Cq\varrho^{-2^{q-N+1}}.$$

This error estimate is computed inductively. For $N > 1$ we use recursively,

$$\begin{aligned} I_{N+1} - \mathscr{A}(q+1,N+1) &= I_{N+1} - \sum_{|\mathbf{i}|\leq q}\left(\bigotimes_{n=1}^{N}\Delta^{i_n}\otimes\mathscr{U}^{q+1-|\mathbf{i}|}\right) \\ &= \sum_{|\mathbf{i}|\leq q}\left(\bigotimes_{n=1}^{N}\Delta^{i_n}\otimes\left(I_1 - \mathscr{U}^{q+1-|\mathbf{i}|}\right)\right) + (I_N - \mathscr{A}(q,N))\otimes I_1. \end{aligned}$$

Furthermore,

$$\begin{aligned} &\left\|\sum_{|\mathbf{i}|\leq q}\left(\bigotimes_{n=1}^{N}(\Delta^{i_n})(u)\otimes\left(I_1 - \mathscr{U}^{q+1-|\mathbf{i}|}\right)(u)\right)\right\|_{\infty,N} \\ &\qquad \leq \sum_{|\mathbf{i}|\leq q}\prod_{n=1}^{N}\left\|(\Delta^{i_n})(u)\right\|_{\infty,N}\left\|(I_1 - \mathscr{U}^{q+1-|\mathbf{i}|})(u)\right\|_{\infty,N} \\ &\qquad \leq CE^N\sum_{|\mathbf{i}|\leq q}\left(\prod_{n=1}^{N} i_n\right)\varrho^{-\sum_{n=1}^{N}2^{i_n-1}}(q+1-|\mathbf{i}|)\,\varrho^{-2^{q+1-|\mathbf{i}|}} \\ &\qquad \leq CE^N\sum_{|\mathbf{i}|\leq q}\left(\prod_{n=1}^{N} i_n\right)(q+1-|\mathbf{i}|)\,\varrho^{-\frac{1}{2}\left(N2^{|\mathbf{i}|/N}+2^{q+2-|\mathbf{i}|}\right)} \\ &\qquad \leq CE^N\sum_{s=N}^{q} g(s,N)\,(q+1-s)\,\varrho^{-f(s,N)} \end{aligned}$$

where we have used the convexity estimate

$$\varrho^{-\sum_{n=1}^{N}2^{i_n}} \leq \varrho^{-N2^{|\mathbf{i}|/N}}.$$

Then, by the inductive assumption (4.11),

$$\begin{aligned}\left\|\left(\left(I_N-\mathscr{A}(q,N)\right)\otimes I_1\right)(u)\right\|_{\infty,N} &\le C\sum_{j=1}^{N-1}E^j\sum_{s=j}^{q-N+j}g(s,j)\,(q-N+j+1-s)\,\varrho^{-f(s,j)}\\ &\quad+\left\|(I_1-\mathscr{A}(q-N+1,1))(u)\right\|_{\infty,N}.\end{aligned}$$

Therefore,

$$\begin{aligned}\left\|(I_{N+1}-\mathscr{A}(q+1,N+1))(\,u)\right\|_{\infty,N} &\le C\sum_{j=1}^{N}E^j\sum_{s=j}^{q-N+j}g(s,j)\,(q-N+j+1-s)\,\varrho^{-f(s,j)}\\ &\quad+\left\|(I_1-\mathscr{A}(q-N+1,1))(u)\right\|_{\infty,N}.\end{aligned}$$

and (4.11) is proved. Set $F=\max\{1,E\}$ to obtain

$$\begin{aligned}\left\|\left(I_N-\mathscr{A}(q,N)\right)(u)\right\|_{\infty,N} &\le C\sum_{j=1}^{N-1}\left(\max\{1,E\}\right)^N\sum_{s=j}^{q-N+j}g(s,j)\,(q-N+j+1-s)\,\varrho^{-f(s,j)}\\ &\quad+\left\|(I_1-\mathscr{A}(q-N+1,1))(u)\right\|_{\infty,N}\\ &\le CF^N\sum_{j=1}^{N-1}\sum_{s=j}^{q-N+j}g(s,j)\,(q-N+j+1-s)\,\varrho^{-f(s,j)}\\ &\quad+\left\|(I_1-\mathscr{A}(q-N+1,1))(u)\right\|_{\infty,N}.\end{aligned}\tag{4.12}$$

We now turn our attention to finding a maximum for $\varrho^{-f(s,j)}$ on the set $\{(s,j)\,:\,j\le s\le q-N+j\ \text{ and } 1\le j\le N-1\}$. Clearly

$$\frac{\partial f}{\partial s}=\left(2^{s/j}\;-\;2^{q-N+j+2-s}\right)\log(2)=0$$

implies that $s=s(j)=j+\frac{j(q-N+1)}{j+1}$, which satisfies for any $j\in\mathbb{N}_+$

$$j\le s(j)\le q-N+1+j.$$

Hence,

$$\begin{aligned}\max_{j\le s\le q-N+j}\varrho^{-f(j,s)} &=\varrho^{-f(j,s(j))}\\ &\le\varrho^{-h(j)}\end{aligned}$$

where $h(j)=(j+1)2^{(q-N+1)/(j+1)}$. Then we get

$$\frac{\mathrm{d}h}{\mathrm{d}j}=2^{(q-N+1)/(j+1)}\left(1-\frac{(q-N+1)\log(2)}{j+1}\right)=0$$

which yields $j=(q-N+1)\log(2)-1$. For q sufficiently large, the minimum of $h(j)$ falls outside the interval $[1,N-1]$ and the function $h(j)$ is decreasing on this interval. Therefore, there are two cases to consider. The first being the situation when $q>N\left(\frac{1+\log(2)}{\log(2)}\right)=N\chi$ and the second when $N\le q\le N\chi$. In either case

$$\max_{1\le j\le N-1}\varrho^{-h(j)}=\varrho^{-p(q,N)},$$

hence

$$\max_{\substack{1\le j\le N-1\\ j\le s\le q-N+j}} \varrho^{-f(j,s)} \le \varrho^{-p(q,N)}.$$

In conclusion we have, for $q \ge N$

$$\begin{aligned}\left\|\left(I_N - \mathscr{A}(q,N)\right)(u)\right\|_{\infty,N} &\le CF^N \varrho^{-p(q,N)} \sum_{j=1}^{N-1} \sum_{s=j}^{q-N+j} g(s,j)(q-N+j+1-s) \\ &= CF^N \varrho^{-p(q,N)}\,\kappa + \left\|\left(I_1 - \mathscr{A}(q-N+1,1)\right)(u)\right\|_{\infty,N}\end{aligned} \tag{4.13}$$

and

$$\begin{aligned}\kappa &= \sum_{j=1}^{N-1} \sum_{s=j}^{q-N+j} g(s,j)(q-N+j+1-s) \\ &= \sum_{j=1}^{N-1} \sum_{s=j}^{q-N+j} (q-N+j+1-s) \sum_{i\in\mathbb{N}_+^j,\,|\mathbf{i}|=s} \left(\prod_{n=1}^{j} i_n\right) \\ &\le \sum_{j=1}^{N-1} \sum_{s=j}^{q-N+j} (q-N+j+1-s) \sum_{i\in\mathbb{N}_+^j,\,|\mathbf{i}|=s} \left(\frac{|\mathbf{i}|}{j}\right)^j \\ &= \sum_{j=1}^{N-1} \sum_{s=j}^{q-N+j} (q-N+j+1-s) \left(\frac{s}{j}\right)^j \binom{s-1}{j-1} \\ &\le (q-N+1) \sum_{j=1}^{N-1} \sum_{s=j}^{q-N+j} \left(\frac{q-N+j}{j}\right)^j \frac{(s-1)^{j-1}}{(j-1)!} \\ &\le (q-N+1)(q-N) \sum_{j=1}^{N-1} (q-1)^j \frac{(q-N+j)^{j-1}}{(j-1)!} \\ &\le (q-N+1)(q-N) \sum_{j=1}^{N-1} \frac{(q-1)^{2j-1}}{(j-1)!} \\ &= (q-N+1)(q-N)(q-1) \sum_{j=0}^{N-2} \frac{(q-1)^{2j}}{j!}.\end{aligned} \tag{4.14}$$

Since the sum $\sum_{j=0}^{N-2} \frac{(q-1)^{2j}}{j!}$ can be bounded by $e^{(q-1)^2}$ or by $\frac{(q-1)^{2N-2}}{q(q-2)}$, then in either case $\kappa \le \Psi(q,N)$. Finally, from (4.12) and using (4.13) and (4.14) we conclude that

$$\left\|\left(I_N - \mathscr{A}(q,N)\right)(u)\right\|_{\infty,N} \le CF^N\,\Psi(q,N)\varrho^{-\frac{p(q,N)}{2}} + \varrho^{-2^{q-N+1}}. \tag{4.15}$$

We also observe that $\Psi(q,N) \ge 1$ and by straightforward calculations

$$2^{q-N+1} \ge \frac{p(q,N)}{2}, \quad \forall N \ge 1, q \ge N,$$

to conclude that

$$\varrho^{-2^{q-N+1}} \le \Psi(q,N)\varrho^{-\frac{p(q,N)}{2}}$$

and this completes the proof. □

Now we relate the number of collocation points $n = n(q,N) = \#\mathscr{H}(q,N)$ to the level q of the Smolyak algorithm. We state the result in the following lemma.

Lemma 4.5 *Using the Smolyak interpolant described by* (3.8) *where the abscissas are the Clenshaw-Curtis knots, described in Section 3.2.1, the total number of points required at level q satisfies the following bounds:*

$$2^{q-N+1} \leq n \leq \frac{2^q q^N}{(N-1)!}. \tag{4.16}$$

Proof. The proof follows immediately but will be shown for completeness. By using formula (3.8) and exploiting the nested structure of the Clenshaw-Curtis abscissas the number of points $n = n(q,N) = \#\mathscr{H}(q,N)$ can be counted in the following way:

$$n = \sum_{|\mathbf{i}|\leq q} \prod_{n=1}^{N} r(i_n), \text{ where } r(i) := \begin{cases} 1 & \text{if } i = 1 \\ 2 & \text{if } i = 2 \\ 2^{i-2} & \text{if } i > 2 \end{cases}. \tag{4.17}$$

If we take $i_1 = i_2 = \ldots = i_{N-1} = 1$ then to satisfy the constraint $|\mathbf{i}| \leq q$ we required $i_N \leq q - N + 1$. Then we get

$$\begin{aligned} 2^{q-N-1} \leq n = \sum_{|\mathbf{i}|\leq q} \prod_{n=1}^{N} r(i_n) \leq \sum_{|\mathbf{i}|\leq q} 2^{|\mathbf{i}|} \leq \sum_{j=N}^{q} \sum_{|\mathbf{i}|=j} 2^j &= \sum_{j=N}^{q} 2^j \binom{j-1}{N-1} \\ &\leq \sum_{j=N}^{q} 2^q \frac{(q-1)^{N-1}}{(N-1)!} \\ &\leq \frac{2^q q^N}{(N-1)!} \end{aligned}$$

which completes the proof. □

The next Theorem relates the error bound (4.8) to the number of collocation points $n = n(q,N) = \#\mathscr{H}(q,N)$, described by Lemma 4.5.

Theorem 4.6 *Assume the conditions of Lemma 4.4 and Lemma 4.5, and define the function*

$$\gamma(n,N) = \log_2(n) + N + 1,$$

then for $N \leq q < N\chi$

$$\left\| \left(I_N - \mathscr{A}(q,N)\right)(u) \right\|_{L^2_\rho(\Gamma^N;W(D))} \leq C F^N \, \Psi(\gamma(n,N),N) \left(\frac{(2\gamma(n,N))^N}{n\,(N-1)!} \right)^{\frac{\Theta}{2}\log(\varrho)} \tag{4.18}$$

and for $q \geq N\chi$

$$\left\| \left(I_N - \mathscr{A}(q,N)\right)(u) \right\|_{L^2_\rho(\Gamma^N;W(D))} \leq C F^N \, \Psi(\gamma(n,N),N) \, \varrho^{-\frac{N}{2}[(N-1)!]^{1/N}\frac{n^{1/N}}{\gamma(n,N)}}. \tag{4.19}$$

where $\Theta = 2^\chi$ and $n = n(q,N)$ is the number of knots that are used by $\mathscr{A}(q,N)$ and Ψ was defined in (4.10).

Proof. Recall that the error bound will be separated into two estimates depending on the domain of definition of (4.9). First for $N \leq q < N\chi$ and using (4.16) we arrive at

$$\begin{aligned} \log(n) + \log((N-1)!) &\leq q\log(2) + N\log(q) \\ &\leq q\log(2) + N\log\gamma(n,N) \\ &= (q-N)\log(2) + N\log(2\gamma(n,N)). \end{aligned}$$

Hence,

$$q - N \geq \frac{\log(n) + \log((N-1)!) - N\log(2\gamma(n,N))}{\log(2)}$$

and using (4.9) implies that

$$\begin{aligned} p(q,N) &\geq (\log(n) + \log((N-1)!) - N\log(2\gamma(n,N))) \cdot \Theta \\ &= \Theta \log\left(\frac{n\,(N-1)!}{(2\gamma(n,N))^N}\right). \end{aligned}$$

Therefore, using (4.16) we deduce that

$$\begin{aligned} \|(I_N - \mathscr{A}(q,N))\,(u)\|_{\infty,N} &\leq CF^N\,\Psi(\gamma(n,N),N)\,\varrho^{-\Theta\,\log\left(\frac{(N-1)!+n}{(2\gamma(n,N))^N}\right)} \\ &= CF^N\,\Psi(\gamma(n,N),N)\,\left(e^{-\log(\varrho)\,\log\left(\frac{(N-1)!+n}{(2\gamma(n,N))^N}\right)}\right)^{\Theta} \\ &= CF^N\,\Psi(\gamma(n,N),N)\left(\frac{(2\gamma(n,N))^N}{n(N-1)!}\right)^{\Theta\cdot\log(\varrho)} \end{aligned} \tag{4.20}$$

and we recover (4.18).

On the other hand, for $q \geq N\chi$ and using (4.16) we find that

$$\left(\frac{2^q q^N}{(N-1)!}\right)^{1/N} \geq n^{1/N}$$

which implies that

$$2^{q/N} \geq \frac{(n(N-1)!)^{1/N}}{\gamma(n,N)}$$

and

$$p(q,N) \geq N\,[(N-1)!]^{1/N}\,\frac{n^{1/N}}{\gamma(n,N)}.$$

Therefore, again with (4.16) we conclude that

$$\|(I_N - \mathscr{A}(q,N))\,(u)\|_{\infty,N} \leq CF^N\,\Psi(\gamma(n,N),N)\varrho^{-\frac{N}{2}[(N-1)!]^{1/N}\frac{n^{1/N}}{\gamma(n,N)}}. \tag{4.21}$$

and we recover (4.19). □

4.1.2 Gaussian interpolation estimates

Similarly to the previous section we now develop error estimates for interpolating functions $u \in C^0(\Gamma^N; W(D))$ that admit an analytic extension as described by Assumption 1.8 using the Smolyak formulations based on Gaussian abscissas described in Section 3.2.2. As before, we remind the reader that in the global estimate (4.1) we need to bound the interpolation error (III) in the norm $L^2_\rho(\Gamma^N; W(D))$. Yet, the Gaussian points defined in Section 3.2.2 are constructed for the more appropriate density $\hat{\rho} = \prod_{n=1}^N \hat{\rho}_n$ and we have

$$\|v\|_{L^2_\rho(\Gamma^N;W(D))} \leq \left\|\frac{\rho}{\hat{\rho}}\right\|_{L^\infty(\Gamma^N)} \cdot \|v\|_{L^2_{\hat{\rho}}(\Gamma^N;W(D))} \quad \text{for all } v \in C^0(\Gamma^N; W(D)).$$

In what follows we will use the shorthand notation $\|\cdot\|_{\hat{\rho},N}$ for $\|\cdot\|_{L^2_{\hat{\rho}}(\Gamma^N;W(D))}$. Utilizing the work of Erdös and Turán [14] we present the following lemma:

Lemma 4.7 *For every function $u \in C^0(\Gamma^1; W(D))$ the interpolation error satisfies*

$$\|u - \mathscr{U}^i(u)\|_{\hat{\rho},1} \leq 2\sqrt{C_{\hat{\rho}}} \inf_{w \in V_{m_i}} \|u - w\|_{\infty,1}.$$

where $C_{\hat{\rho}} = \displaystyle\int_{\Gamma^1} \hat{\rho}(y)\, dy$.

Proof. We have, indeed, for any $v \in V_{m_i}$

$$\begin{aligned} \|u - \mathscr{U}^i(u)\|^2_{\hat{\rho},1} &= \left\|u - v + (v - \mathscr{U}^i(u))\right\|^2_{\hat{\rho},1} \\ &= \left\|u - v + \mathscr{U}^i(v - u)\right\|^2_{\hat{\rho},1} \\ &\leq 2\left(\|u - v\|^2_{\hat{\rho},1} + \left\|\mathscr{U}^i(u - v)\right\|^2_{\hat{\rho},1}\right) \end{aligned} \tag{4.22}$$

where we observe that $\forall v \in V_{m_i}$, it holds $\mathscr{U}^i(v) = v$. Then it is easy to see that

$$\begin{aligned} \|u - v\|^2_{\hat{\rho},1} &\leq \int_{\Gamma^1} \hat{\rho}(y)\, |(u - v)(y)|^2\, dy \\ &\leq \|u - v\|^2_{\infty,1} \int_{\Gamma^1} \hat{\rho}(y)\, dy = C_{\hat{\rho}}\, \|u - v\|^2_{\infty,1} \end{aligned}$$

and

$$\begin{aligned} \left\|\mathscr{U}^i(u - v)\right\|^2_{\hat{\rho},1} &= \left\|\sum_{j=1}^{m_i} (u - v)(y^i_j) l^i_j(y)\right\|^2_{\hat{\rho},1} \\ &\leq \sum_{j,j'=1}^{m_i} \left|(u - v)(y^i_j)\right| \left|(u - v)(y^i_{j'})\right| \int_{\Gamma^1} \hat{\rho}(y) l^i_j(y) l^i_{j'}(y)\, dy \\ &\leq \|u - v\|^2_{\infty,1} \sum_{j=1}^{m_i} \int_{\Gamma^1} \hat{\rho}(y)(l^i_j(y))^2\, dy = \left(\int_{\Gamma^1} \hat{\rho}(y) dy\right) \|u - v\|^2_{\infty,1} \end{aligned}$$

where we exploit the orthogonality of the Lagrange polynomial basis. Then from (4.22) we conclude that

$$\|u - \mathscr{U}^i(u)\|^2_{\hat{\rho},1} \leq 4 \left(\int_{\Gamma^1} \hat{\rho}(y) dy\right) \|u - v\|^2_{\infty,1}$$

and the result follows directly. □

Similar to Section 4.1.1 we let E_m be the error of the best approximation to functions $u \in C^0(\Gamma^1; W(D))$ that admit an analytic extension as described by Assumption 1.8 by functions $w \in V_m$. Then, from Lemma 4.7 we begin with

$$\left\|u - \mathscr{U}^i(u)\right\|_{\hat{\rho},1} \leq 2\sqrt{\left(\int_{\Gamma^1} \hat{\rho}(y) dy\right)}\, E_{m_i - 1}(u). \tag{4.23}$$

Again, from Lemma 4.1 the best approximation is bounded by :

$$E_{m_i}(u) \leq C\, \varrho^{-m_i} \tag{4.24}$$

where C is a constant dependent on τ defined in Lemma 4.1. Hence (4.23) and (4.24) imply

$$\left\|(I_1 - \mathscr{U}^i)(u)\right\|_{\hat{\rho},1} \leq \widetilde{C}\, \varrho^{-2^i},$$

$$\begin{aligned}\left\|(\Delta^i)(u)\right\|_{\hat{\rho},1} &= \left\|(\mathscr{U}^i - \mathscr{U}^{i-1})(u)\right\|_{\hat{\rho},1} \\ &\leq \left\|(I_1 - \mathscr{U}^i)(u)\right\|_{\hat{\rho},1} + \left\|(I_1 - \mathscr{U}^{i-1})(u)\right\|_{\hat{\rho},1} \\ &\leq \widetilde{E}\,\varrho^{-2^{i-1}}\end{aligned}$$

for all $i \in \mathbb{N}_+$ with positive constants $\widetilde{C}$ and $\widetilde{E}$ depending on u but not on i. We then present the following lemma and theorem whose proofs follow, with minor changes, those given in Lemma 4.4 and Theorem 4.6 respectively.

Lemma 4.8 *For functions $u \in L^2_\rho(\Gamma^N; W(D))$ that admit an analytic extension as described by Assumption 1.8 we obtain*

$$\left\|\left(I_N - \mathscr{A}(q,N)\right)(u)\right\|_{L^2_\rho(\Gamma^N;W(D))} \leq \|\rho/\hat{\rho}\|_{L^\infty(\Gamma^N)}\,\widetilde{C}\widetilde{F}^N\,\widetilde{\Psi}(q,N)\varrho^{-\frac{p(q,N)}{2}} \tag{4.25}$$

where

$$p(q,N) := \begin{cases} N\,2^{q/N}, & \text{if } q > N\chi, \\ (q-N)\log(2)\cdot 2^{\chi}, & \text{otherwise} \end{cases}, \tag{4.26}$$

$$\widetilde{\Psi}(q,N) := \begin{cases} 1 & \text{if } N = 1 \\ \min\left\{q^{N-2}, qe^q\right\} & \text{otherwise} \end{cases} \tag{4.27}$$

and $\chi = \left(\frac{1+\log(2)}{\log(2)}\right)$.

Now we relate the number of collocation points $n = n(q,N) = \#\mathscr{H}(q,N)$ to the level q of the Smolyak algorithm. We state the result in the following lemma:

Lemma 4.9 *Using the Smolyak interpolant described by* (3.9) *where the abscissas are the Gaussian knots described in Section 3.2.2, the total number of points required at level q satisfies the following bounds:*

$$2^{q-N} \leq n \leq \frac{2^q q^N}{(N-1)!}. \tag{4.28}$$

Proof. The proof follows immediately but will be shown for completeness. By using formula (3.9), where we collocate using the Gaussian abscissas the number of points $n = n(q,N) = \#\mathscr{H}(q,N)$, can be counted in the following way:

$$n = \sum_{|\mathbf{i}|\leq q}\prod_{n=1}^{N}\widetilde{r}(i_n), \text{ where } 2^{i-1} \leq \widetilde{r}(i) := \begin{cases} 1 & \text{if } i = 1 \\ 2^{i-1}+1 & \text{if } i \geq 2 \end{cases}. \tag{4.29}$$

If we take $i_1 = i_2 = \ldots = i_{N-1} = 1$, then to satisfy the constraint $|\mathbf{i}| \leq q$ we required $i_N \leq q - N + 1$. Then we get

$$\begin{aligned} 2^{q-N} \leq 2^{q-N} + 1 \leq n = \sum_{|\mathbf{i}|\leq q}\prod_{n=1}^{N}\widetilde{r}(i_n) \leq \sum_{|\mathbf{i}|\leq q} 2^{|\mathbf{i}|} &\leq \sum_{j=N}^{q}\sum_{|\mathbf{i}|=j} 2^j = \sum_{j=N}^{q} 2^j \binom{j-1}{N-1} \\ &\leq \sum_{j=N}^{q} 2^q\,\frac{(q-1)^{N-1}}{(N-1)!} \\ &\leq \frac{2^q q^N}{(N-1)!} \end{aligned}$$

which completes the proof. □

Finally, the next Theorem relates the error bound (4.25) to the number of collocation points $n = n(q,N) = \#\mathscr{H}(q,N)$, described by Lemma 4.9.

Theorem 4.10 *Assume the conditions of Lemma 4.8 and 4.9, and define the function*

$$\widetilde{\gamma}(n,N) = \log_2(n) + N,$$

then for $N \leq q < N\chi$ *there holds*

$$\|(I_N - \mathscr{A}(q,N))(u)\|_{L^2_\rho(\Gamma^N;W(D))} \leq \widehat{C}\widetilde{F}^N \widetilde{\Psi}(\widetilde{\gamma}(n,N),N) \left(\frac{(2\widetilde{\gamma}(n,N))^N}{n(N-1)!}\right)^{\frac{\Theta}{2}\log(\varrho)} \tag{4.30}$$

and for $q \geq N\chi$

$$\|(I_N - \mathscr{A}(q,N))(u)\|_{L^2_\rho(\Gamma^N;W(D))} \leq \widehat{C}\widetilde{F}^N \widetilde{\Psi}(\widetilde{\gamma}(n,N),N) \varrho^{-\frac{N}{2}[(N-1)!]^{1/N}\frac{n^{1/N}}{\widetilde{\gamma}(n,N)}}, \tag{4.31}$$

where $\Theta = 2^\chi$, $\widehat{C} = \widetilde{C}\,\|\rho/\hat{\rho}\|_{L^\infty(\Gamma^N)}$ *and* $n = n(q,N)$ *is the number of knots that are used by* $\mathscr{A}(q,N)$ *and* $\widetilde{\Psi}$ *was defined in* (4.27).

4.2 Influence of truncation errors

In this Section we consider the case where the coefficients a_N and f_N from (1.4) are suitably truncated random fields. In this case the truncation error $u - u_N$ is nonzero and contributes to the total error. Such contribution should be considered as well as the relationship between this error and the discretization error.

To this end, if we take the level q to be dimension dependent, i.e. $q = \alpha N$ where $\alpha \geq \chi$ is some constant, then we can estimate the total error $\|u - \mathscr{A}(q,N)(u_N)\|_{L^2_P(\Omega;W(D))}$ in terms of N only. Consider first the case of Gaussian abscissas described in Section 3.2.2. The following theorem holds:

Theorem 4.11 *Let* $q = \alpha N$ *such that* $\alpha \geq \chi$ *and* $\zeta(N)$ *is a monotonic decreasing function of* N *such that* $\zeta(N) \to 0$ *as* $N \to \infty$. *Further define* $\beta(\alpha) = \alpha + \log(\widetilde{F}) - \log(\varrho)\,2^{\alpha-1}$, $\widetilde{\alpha}$ *the solution to* $\beta(\widetilde{\alpha}) = 0$ *and* $\alpha > \max\{\chi, \widetilde{\alpha}\}$. *Under the assumptions of Lemma 4.8 and Theorem 4.10 and the further assumption that*

$$\|u - u_N\|_{L^2_P(\Omega;W(D))} \leq \zeta(N)$$

where $u \in L^2_P(\Omega;W(D))$ *and* $u_N \in L^2_\rho(\Gamma^N;W(D))$, *we get*

$$\|u - \mathscr{A}(q,N)(u_N)\|_{L^2_P(\Omega;W(D))} \leq \zeta(N) + \alpha\widehat{C}N\,e^{\beta(\alpha)N} \tag{4.32}$$

Proof. We begin by writing the total error when approximating $u \in C^0(\Omega;W(D))$ by its N-dimensional interpolant $\mathscr{A}(q,N)(u_N)$. That is, we want to understand

$$\begin{aligned}\|u - \mathscr{A}(q,N)(u_N)\|_{L^2_P(\Omega;W(D))} &\leq \|u - u_N\|_{L^2_P(\Omega;W(D))} + \|(I_N - \mathscr{A}(q,N)(u_N))\|_{L^2_P(\Omega;W(D))} \\ &= \underbrace{\|u - u_N\|_{L^2_P(\Omega;W(D))}}_{\text{(I)}} + \underbrace{\|(I_N - \mathscr{A}(q,N)(u_N))\|_{L^2_\rho(\Gamma^N;W(D))}}_{\text{(II)}}\end{aligned} \tag{4.33}$$

By the assumption, the first term (I) is bounded by $\zeta(N)$ for all N and from Lemma 4.8, and by the assumption $q = \alpha N \geq \chi N$, the second term (II) can be bounded by

$$\begin{aligned}
\|(I_N - \mathscr{A}(q,N))\,(u_N)\|_{L^2_\rho(\Gamma^N;W(D))} &\leq \widehat{C}\widetilde{F}^N q\, e^q \varrho^{-\frac{N}{2}\,2^{q/N}} \\
&\leq \widehat{C}\widetilde{F}^N \alpha N\, e^{\alpha N} \varrho^{-N\,2^{\alpha-1}} \\
&\leq \alpha\widehat{C}\widetilde{F}^N N\, e^{\alpha N - \log(\varrho) N\,2^{\alpha-1}} \\
&= \alpha\widehat{C} N\, e^{N(\alpha + \log(\widetilde{F}) - \log(\varrho)\,2^{\alpha-1})} \\
&= \alpha\widehat{C} N\, e^{\beta(\alpha) N}
\end{aligned} \tag{4.34}$$

where $\beta(\alpha) = \alpha + \log(\widetilde{F}) - \log(\varrho)\,2^\alpha$; such that, for sufficiently large α, $\beta(\alpha)$ is negative. With $\alpha > \max\{\chi, \widetilde{\alpha}\}$ equation (4.34) becomes

$$\|(I_N - \mathscr{A}(q,N))\,(u_N)\|_{L^2_\rho(\Gamma^N;W(D))} \leq \alpha\widehat{C} N\, e^{\beta(\alpha) N},$$

which substituted into the total error (4.33) yields

$$\begin{aligned}
\|u - \mathscr{A}(q,N)(u)\|_{L^2_P(\Omega;W(D))} &\leq \|u - u_N\|_{L^2_P(\Omega;W(D))} + \|(I_N - \mathscr{A}(q,N)(u_N))\|_{L^2_\rho(\Gamma^N;W(D))} \\
&\leq \zeta(N) + \alpha\widehat{C} N\, e^{\beta(\alpha) N}
\end{aligned}$$

as required by (4.32). □

We want to understand the cases where (II) is negligible when compared with (I). In Theorem 4.11 we assume that the truncation error $\|u - u_N\|_{L^2_P(\Omega;W(D))}$ is bounded by $\zeta(N)$ for all N. The function $\zeta(N)$ is typically related to the decay of the eigenvalues if one truncates the noise with a Karhunen-Loève expansion (see [16]). For example, if

$$\|u - u_N\|_{L^2_P(\Omega;W(D))} \leq \theta N^{-r}, \text{ for } r > 0,$$

for some constant θ, then

$$\|u - \mathscr{A}(q,N)(u)\|_{L^2_P(\Omega;W(D))} \leq \underbrace{\theta N^{-r}}_{\text{(I)}} + \underbrace{\alpha\widehat{C} N\, e^{\beta(\alpha) N}}_{\text{(II)}}.$$

In such a situation the Smolyak error (II) is asymptotically negligible with respect to the truncation (I) as $N \to \infty$. Therefore, the isotropic Smolyak algorithm is an efficient interpolation scheme to choose in computational experiments. On the other hand, if

$$\|u - u_N\|_{L^2_P(\Omega;W(D))} \leq \theta e^{-\gamma N} \text{ where } \gamma > \beta(\alpha)$$

then

$$\|u - \mathscr{A}(q,N)(u)\|_{L^2_P(\Omega;W(D))} \leq \underbrace{\theta e^{-\gamma N}}_{\text{(I)}} + \underbrace{\alpha\widehat{C} N\, e^{\beta(\alpha) N}}_{\text{(II)}},$$

which implies that the truncation error (I) is dominated by the Smolyak error (II). In this case the Smolyak algorithm is an inadequate interpolation scheme and improvements to this algorithm must be investigated. We recommend the development of an anisotropic version of the Smolyak algorithm to facilitate faster convergence of such problems.

Remark 4.12 *In the situation in which the Clenshaw-Curtis abscissas are used, the term (II) in* (4.33) *can be bounded as*

$$\begin{aligned}\|(I_N - \mathscr{A}(q,N)(u_N))\|_{L^2_\rho(\Gamma^N;W(D))} &\leq C\, F^N q^2 e^{q^2}\, \varrho^{-\frac{N}{2}2^{q/N}} \\ &= C q^2 e^{N\log(F)+q^2-\frac{N}{2}\log\varrho 2^{q/N}}.\end{aligned}$$

In the presence of the term q^2, global convergence can only be achieved if one takes $q = N^{1+\alpha}$ with $\alpha > 0$.

5 Numerical Examples

This section illustrates the convergence of the sparse collocation method for the stochastic linear elliptic problem in two spatial dimensions, as described in Section 2. The computational results are in accordance with the convergence rate predicted by the theory. We will also use this problem to compare the convergence of the isotropic Smolyak method with that of the anisotropic full tensor product method described by [3] using the adaptive algorithm described in the work [5, Section 9].

The problem is to solve

$$\begin{cases} -\nabla\cdot(a_N(\omega,\cdot)\nabla u(\omega,\cdot)) = f_N(\omega,\cdot) \quad \text{in } D\times\Omega, \\ u(\omega,\cdot) = 0 \quad \text{on } \partial D\times\Omega. \end{cases} \tag{5.1}$$

with $D = \left\{\mathbf{x} = (x,z)\in\mathbb{R}^2 : 0\leq x,z\leq 1\right\}$. For this numerical example we take a deterministic load $f_N(\omega,x,z) = \cos(x)\sin(z)$ and construct the random diffusion coefficient $a_N(\omega,x)$ with one-dimensional spatial dependence as

$$\log(a_N(\omega,\tilde{x}) - 0.5) = 1 + \sigma\sum_{n=1}^{N}\left(\frac{\sqrt{\pi}L}{2}\right)^{1/2} e^{\frac{-(n-1)^2\pi^2L^2}{8}}\cos((n-1)\tilde{x})\,\mathrm{Y}_n(\omega) \tag{5.2}$$

where $\tilde{x}\in[0,2\pi]$, $\sigma = 1$. Therefore, for $x\in[0,L_p]$ we simply shift coordinates such that

$$\tilde{x} = \frac{2\pi x}{L_p} \text{ and } L = \frac{L_c}{L_p}$$

where $L_p = 1$ is the length of the x spatial direction. The parameter L_c appearing in (5.2) dictates the decay of the terms in the expansion and is related to a "physical correlation length". Small values of L_c will be related with slow decay in (5.2).

The random variables $\{Y_n(\omega)\}_{n=1}^{\infty}$ are independent, have zero mean and unit variance, i.e. $E[Y_n] = 0$ and $\mathbb{E}[Y_nY_m] = \delta_{nm}$ for $n,m\in\mathbb{N}_+$, and are taken uniform in the interval $[-\sqrt{3},\sqrt{3}]$. Expansion (5.2) is related to a Karhunen-Loève expansion of a one-dimensional random field with stationary covariance

$$\mathrm{cov}[\log(a_N - 0.5)](x_1,x_2) = \sigma^2\exp\left(\frac{-(x_1-x_2)^2}{L_c^2}\right).$$

To formulate the constant ϱ defined by (4.2) for the problem (5.1) we investigate a lower bound for ϱ_n. That is (see (4.4))

$$\begin{aligned}\varrho_n &\geq 1 + \frac{1}{4\sqrt{3}}\left(\frac{2}{\sqrt{\pi}L}\right)^{1/2} e^{\frac{(n-1)^2\pi^2L^2}{8}} \\ &= 1 + \frac{1}{2}\left(\frac{1}{6\sqrt{\pi}L}\right)^{1/2} e^{\frac{(n-1)^2\pi^2L^2}{8}}\end{aligned}$$

and therefore

$$\varrho = \min_n \varrho_n \geq 1 + \sqrt{\frac{1}{24\sqrt{\pi}L}}. \tag{5.3}$$

To illustrate the behavior of the sparse collocation method constructed from either Clenshaw-Curtis or Gaussian abscissas we assume the random variables Y_n are bounded with uniform densities. The corresponding collocation points are then sparse cartesian products determined by the roots of either Chebyshev or Legendre polynomials using the Smolyak method described in Section 3.2. Recall from Section 3.2.1 that the Clenshaw-Curtis abscissas are nested and therefore, in practice, we exploit this fact and construct the Smolyak interpolant using formula (3.8). Therefore, the number of points $n = n(q, N) = \#\mathscr{H}(q, N)$ can be counted as in formula (4.17). On the other hand, the Gaussian abscissas described in Section 3.2.2 are not nested, and to reduce the number of points necessary to build the Smolyak interpolant one utilizes the variant of (3.8), given by (3.9). In doing this, we can count the number of points n used by the Smolyak interpolant as in formula (4.29).

The finite element space for the spatial discretization is the span of continuous functions that are piecewise polynomials with degree two over a uniform triangulation of D with 1089 unknowns.

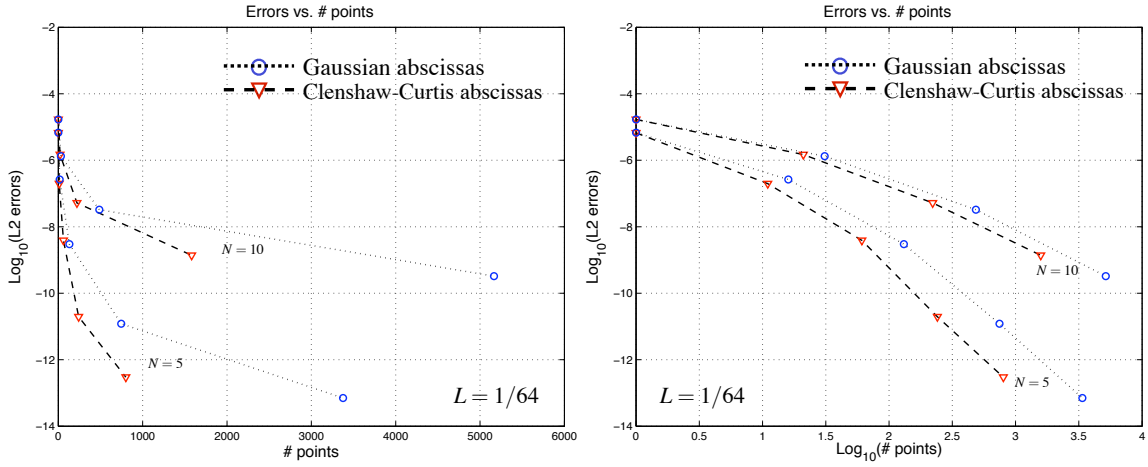

Figure 1: The rate of convergence of the Smolyak algorithm for a given correlation length $L_c = 1/64$ using both the Gaussian and Clenshaw-Curtis abscissas. For the values $N = 5$ and $N = 10$ in (5.2) we plot: on the left, $\log(\epsilon)$ versus the number of collocation points and on the right, $\log(\epsilon)$ versus the logarithm of the number of collocation points.

Observe, in general, that the collocation method only requires the solution of uncoupled deterministic problems over the set of collocation points, even in the presence of a diffusivity coefficient which depends nonlinearly on the random variables as in (5.2). This is a significant advantage that the collocation method offers compared to the classical Stochastic-Galerkin finite element method as considered in [4] or [16, 19, 25, 37]. To study the convergence of the isotropic Smolyak algorithm we consider a problem with a fixed dimension N and investigate the behavior when the level q of the interpolation in the Smolyak algorithm is increased linearly. The computational results for the $L^2(D)$ approximation error to the expected value, $E[u]$, are shown in Figure 1. Here we consider two cases, namely $N = 5$ and $N = 10$ for the finite sum (5.2). To estimate the computational error in the q-th level we approximate $E[\epsilon] \approx E[\mathscr{A}(q, N)\pi_h u_N - \mathscr{A}(q+1, N)\pi_h u_N]$ using either Gaussian

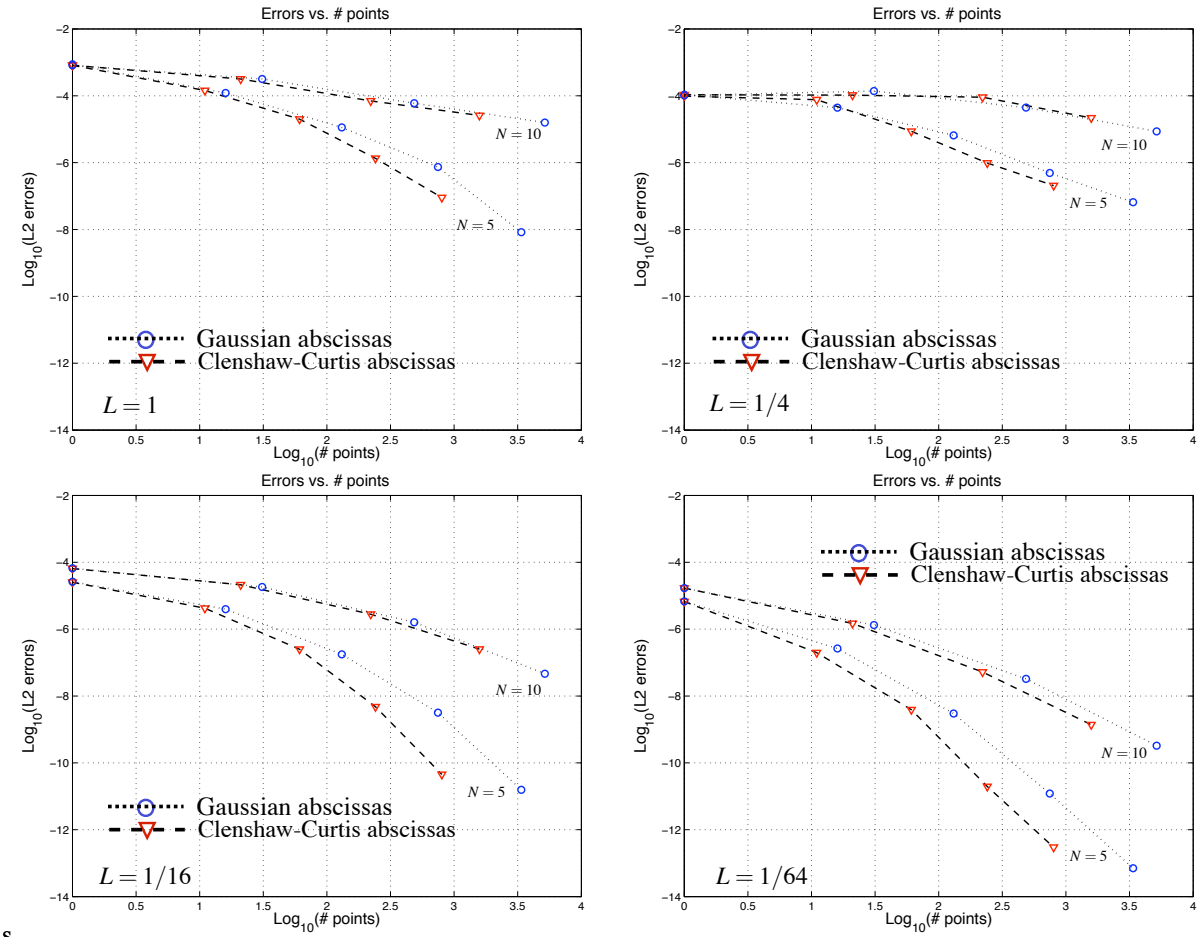

Figure 2: The convergence of the Smolyak algorithm in $N = 5$ and $N = 10$ dimensions for correlation lengths $L_c = 1, 1/4, 1/16$ and $1/64$, using both Gaussian and Clenshaw-Curtis abscissas.

or Clenshaw-Curtis abscissas. The results reveal, as expected, that for a small non-degenerate correlation length, i.e. $L_c = 1/64$, the error decreases (sub)-exponentially, as the level q increases. We also observe that the convergence rate is dimension dependent and slightly deteriorates as N increases.

To investigate the performance of the algorithm by varying the correlation length L we also include the cases where $L_c = 1/16$, $L_c = 1/4$ and $L_c = 1$ for both $N = 5$ and $N = 10$, seen in Figure 2. We notice that the larger correlation lengths have negative effects on the rate of convergence. This can be explained by examining ϱ defined by (5.3). From this we see that the coefficient ϱ appearing in the estimates (4.18)-(4.19) and (4.30)-(4.31), is approaching 1 as L becomes large. Hence, the effect of increasing L is a deterioration of the rate of convergence. Therefore, our final interest is to compare our isotropic sparse tensor product method with an anisotropic full tensor product method, proposed in [5].

The anisotropic full tensor product algorithm can be described in the following way: given a tolerance *tol* the method computes a multi-index $p = (p_1, p_2, \ldots, p_N)$, corresponding to the order of the approximating polynomial spaces $\mathcal{P}_p(\Gamma)$. 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. Table 1 and Table 2 show the values of components of the 5-dimensional multi-index p for different values of tol, corresponding to $L_c = 1$ and $L_c = 1/64$ respectively. These tables help give insight into the anisotropic behavior of each particular problem. Observe, in particular, for the case $L_c = 1/64$ the algorithm predicts a multi-index p which is equal in all directions, i.e. an isotropic tensor product space.

A convergence plot for $L_c = 1$ and $L_c = 1/64$ can be constructed by examining each row of Table 1 and Table 2 respectively, and plotting the number of points in the tensor product grid versus the error in expectation. We estimate the error in expectation by $E[e] \approx E[u_{h,p} - u_{h,\tilde{p}}]$, with $\tilde{p} = (p_1 + 1, p_2 + 1, \ldots, p_N + 1)$.

tol	$N = 1$	$N = 2$	$N = 3$	$N = 4$	$N = 5$
1.0e-02	$p_1 = 1$	$p_2 = 1$	$p_3 = 1$	$p_4 = 1$	$p_5 = 1$
1.0e-03	$p_1 = 2$	$p_2 = 1$	$p_3 = 1$	$p_4 = 1$	$p_5 = 1$
1.0e-04	$p_1 = 2$	$p_2 = 2$	$p_3 = 1$	$p_4 = 1$	$p_5 = 1$
1.0e-05	$p_1 = 3$	$p_2 = 2$	$p_3 = 2$	$p_4 = 2$	$p_5 = 1$
1.0e-06	$p_1 = 3$	$p_2 = 3$	$p_3 = 3$	$p_4 = 2$	$p_5 = 2$
1.0e-07	$p_1 = 5$	$p_2 = 4$	$p_3 = 3$	$p_4 = 3$	$p_5 = 2$
1.0e-08	$p_1 = 5$	$p_2 = 5$	$p_3 = 3$	$p_4 = 3$	$p_5 = 3$
1.0e-09	$p_1 = 6$	$p_2 = 6$	$p_3 = 4$	$p_4 = 4$	$p_5 = 3$
1.0e-10	$p_1 = 7$	$p_2 = 7$	$p_3 = 4$	$p_4 = 4$	$p_5 = 3$
1.0e-11	$p_1 = 8$	$p_2 = 8$	$p_3 = 5$	$p_4 = 5$	$p_5 = 4$
1.0e-12	$p_1 = 8$	$p_2 = 8$	$p_3 = 6$	$p_4 = 5$	$p_5 = 4$

Table 1: The five components of the multi-index p used as the input information for the anisotropic full tensor product algorithm when solving problem (5.1) with a correlation length $L_c = 1$.

tol	$N = 1$	$N = 2$	$N = 3$	$N = 4$	$N = 5$
1.0e-03	$p_1 = 1$	$p_2 = 1$	$p_3 = 1$	$p_4 = 1$	$p_5 = 1$
1.0e-06	$p_1 = 2$	$p_2 = 2$	$p_3 = 2$	$p_4 = 2$	$p_5 = 2$
1.0e-09	$p_1 = 3$	$p_2 = 3$	$p_3 = 3$	$p_4 = 3$	$p_5 = 3$
1.0e-12	$p_1 = 4$	$p_2 = 4$	$p_3 = 4$	$p_4 = 4$	$p_5 = 4$

Table 2: The five components of the multi-index p used as the input information for the anisotropic full tensor product algorithm when solving problem (5.1) with a correlation length $L_c = 1/64$.

To study the advantages and/or disadvantages to collocating in a sparse tensor product space as opposed to the anisotropic full tensor product space we show, in Figure 3, the convergence of both methods when solving problem (5.1), using correlation lengths $L_c = 1$ and $L_c = 1/64$ with $N = 5$. Figure 3 reveals that for $L_c = 1/64$ the isotropic Smolyak method converges faster than the anisotropic full tensor product method. This is due to a slower decay of the terms in expansion (5.2) and hence, an almost equal weighting of all 5 random variables. On the contrary, the opposite conclusions can be drawn from the comparison for $L_c = 1$. Since, in this case, the rate of decay of the expansion is faster, the anisotropic full tensor method weighs heavily these important modes and, therefore achieves a faster convergence.

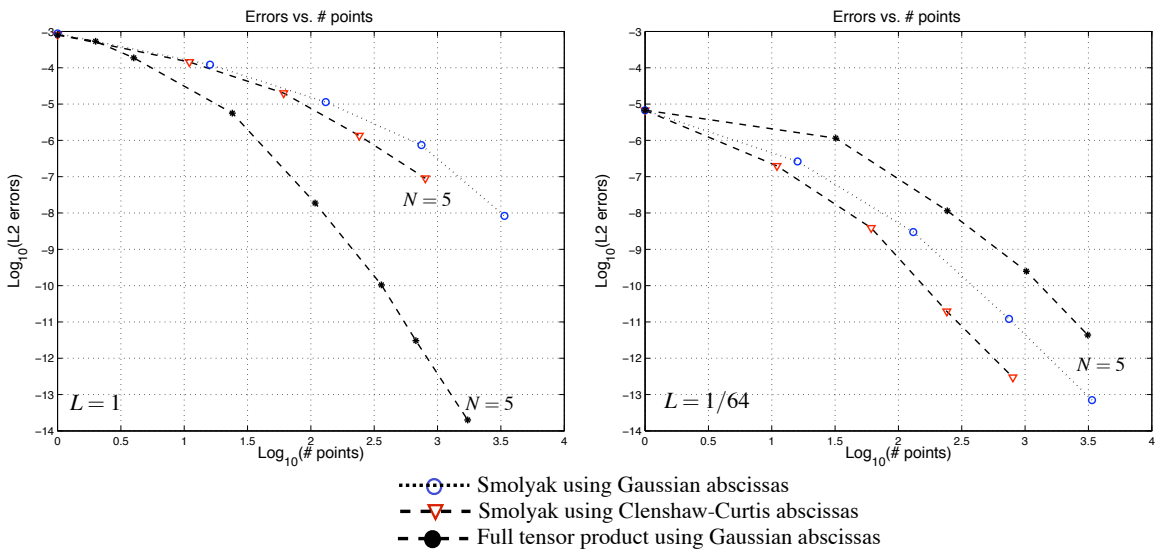

Figure 3: A 5-dimensional comparison of the Smolyak method versus the anisotropic full tensor product algorithm when solving problem (5.1).

6 Conclusions

In this work we proposed and analyzed a sparse grid stochastic collocation method for solving elliptic partial differential equations whose coefficients and forcing terms depend on a finite number of random variables. The sparse grids are constructed from the Smolyak algorithm, utilizing either Clenshaw-Curtis or Gaussian abscissas. The method leads to the solution of uncoupled deterministic problems and, as such, is fully parallelizable like a Monte Carlo method.

This method extends the work proposed in [3] where a stochastic collocation method on tensor product grids was proposed. The use of sparse grids considered in the present work (as opposed to full tensor grids), reduces considerably the *curse of dimensionality* and allows us to treat effectively problems that depend on a moderately large number of random variables, while keeping a high level of accuracy.

Upon assumption that the solution depends analytically on each random variable (which is a reasonable assumption for a certain class of applications, see [3, 4]), we have provided a full convergence analysis and demonstrated (sub)-exponential convergence of the "probability error" in the asymptotic regime and algebraic convergence of the "probability error" in the pre-asymptotic regime, with respect to the total number of collocation points used in the sparse grid.

The main theoretical results are given in Theorem 4.6 and Theorem 4.10 and confirmed numerically by the examples presented in Section 5.

The method is very effective for problems whose input data depend on a moderate number of random variables, which "weigh equally" in the solution. For such an isotropic situation the displayed convergence is faster than standard collocation techniques built upon full tensor product spaces.

On the other hand, the convergence rate deteriorates when we attempt to solve highly anisotropic problems, such as those appearing when the input random variables come e.g. from KL-type truncations of "smooth" random fields. In such cases, a full anisotropic tensor product approximation, as proposed in [3,5], may still be more effective for a small or moderate number of random variables.

Future directions of this research will include the development of an anisotropic version of the Sparse Grid Stochastic Collocation method, which will combine an optimal treatment of the anisotropy of the problem while reducing the *curse of dimensionality* via the use of sparse grids.

Acknowledgments

The first author was partially supported by M.U.R.S.T. Cofin 2005 "Numerical Modeling for Scientific Computing and Advanced Applications". The second author wants to acknowledge the support of UdelaR in Uruguay. The first and second authors were partially supported by the SANDIA project # 523695. The third author was supported by the School of Computation Science (SCS) at Florida State University and would like to thank MOX, Dipartimento di Matematica, Politecnico di Milano, Italy, for hosting a visit to complete this research. Finally, the third author would like to thank Prof. Max Gunzburger and Dr. John Burkardt for their insight, guidance and many helpful discussions.

References

[1] I. Babuška and P. Chatzipantelidis. On solving elliptic stochastic partial differential equations. *Comput. Methods Appl. Mech. Engrg.*, 191(37-38):4093–4122, 2002.

[2] I. M. Babuška, K. M. Liu, and R. 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] I. M. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. Technical report, MOX, Dipartimento di Matematica, 2005.

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

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

[6] V. Barthelmann, E. Novak, and K. Ritter. High dimensional polynomial interpolation on sparse grids. *Adv. Comput. Math.*, 12(4):273–288, 2000. Multivariate polynomial interpolation.

[7] V. Barthelmann, E. Novak, and K. Ritter. High dimensional polynomial interpolation on sparse grids. *Advances in Computational Mathematics*, 12:273–288, 2000.

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

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

[10] C. W. Clenshaw and A. R. Curtis. A method for numerical integration on an automatic computer. *Numer. Math.*, 2:197–205, 1960.

[11] 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(48):6359–6372, 2001.

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

[13] V. K. Dzjadyk and V. V. Ivanov. On asymptotics and estimates for the uniform norms of the lagrange interpolation polynomials corresponding to the chebyshev nodal points. *Analysis Mathematica*, 9(11):85–97, 1983.

[14] 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.

[15] G.S. Fishman. *Monte Carlo*. Springer Series in Operations Research. Springer-Verlag, New York, 1996. Concepts, algorithms, and applications.

[16] P. Frauenfelder, C. Schwab, and R. A. Todor. Finite elements for elliptic problems with stochastic coefficients. *Comput. Methods Appl. Mech. Engrg.*, 194(2-5):205–228, 2005.

[17] A. Gaudagnini and S. Neumann. Nonlocal and localized analysis of conditional mean steady state flow in bounded, randomly nonuniform domains. 1. Theory and computational approach. 2. Computational examples. *Water Resources Research*, 35(10):2999–3039, 1999.

[18] T. Gerstner and M. Griebel. Numerical integration using sparse grids. *Numer. Algorithms*, 18(3-4):209–232, 1998.

[19] R. G. Ghanem and P. D. Spanos. *Stochastic finite elements: a spectral approach*. Springer-Verlag, New York, 1991.

[20] G.E. Karniadakis, C.-H. Su, D. Xiu, D. Lucor, C. Schwab, and R.A. Todor. Generalized polynomial chaos solution for differential equations with random inputs. SAM Report 2005-01, ETH, 2005.

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

[22] 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.

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

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

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

[26] F. Nobile, R. Tempone, and C. Webster. A stochastic collocation method for nonlinear elliptic partial differential equations with random input data. In progress.

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

[28] F. Riesz and B. Sz-Nagy. *Functional Analysis*. Dover, 1990.

[29] L.J. Roman and M. Sarkis. Stochastic galerkin method for elliptic SPDES: a white noise approach. *DCDS-B Journal*, 6(4):941–955, 2006.

[30] S.A. Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Dokl. Akad. Nauk SSSR*, 4:240–243, 1963.

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

[32] R. A. Todor. *Sparse Perturbation Algorithms for Elliptic PDE's with Stochastic Data*. PhD thesis, Dipl. Math. University of Bucharest, 2005.

[33] G. W. Wasilkowski and H. Woźniakowski. Explicit cost bounds of algorithms for multivariate tensor product problems. *Journal of Complexity*, 11:1–56, 1995.

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

[35] 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.

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

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

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