

IMPLEMENTATION OF OPTIMAL GALERKIN AND COLLOCATION APPROXIMATIONS OF PDES WITH RANDOM COEFFICIENTS^{*,**}

J. BECK¹, F. NOBILE², L. TAMELLINI² AND R. TEMPONE¹

Abstract. In this work we first focus on the Stochastic Galerkin approximation of the solution u of an elliptic stochastic PDE. We rely on sharp estimates for the decay of the coefficients of the spectral expansion of u on orthogonal polynomials to build a sequence of polynomial subspaces that features better convergence properties compared to standard polynomial subspaces such as Total Degree or Tensor Product.

We consider then the Stochastic Collocation method, and use the previous estimates to introduce a new effective class of Sparse Grids, based on the idea of selecting a priori the most profitable hierarchical surpluses, that, again, features better convergence properties compared to standard Smolyak or tensor product grids.

Key words: Uncertainty Quantification, PDEs with random data, elliptic equations, multivariate polynomial approximation, Best M -Terms approximation, Stochastic Galerkin methods, Smolyak approximation, Sparse grids, Stochastic Collocation methods.

AMS Subject Classification: 41A10, 65C20, 65N12, 65N35

INTRODUCTION

Many works have been recently devoted to the analysis and the improvement of the Stochastic Galerkin and Collocation techniques for Uncertainty Quantification for PDEs with random input data. These methods are promising since they can exploit the possible regularity of the solution with respect to the stochastic parameters to achieve much faster convergence than sampling methods like Monte Carlo.

Stochastic Galerkin and Collocation methods can be classified as *parametric* techniques, since both expand u , the solution of the PDE of interest, as a summation over suitable deterministic basis functions in probability space, typically polynomials or piecewise polynomials. Stochastic Galerkin is a projection technique over a set of orthogonal polynomials with respect to the probability measure at hand (see e.g. [1, 12, 15, 20]), while Stochastic Collocation is a sum of Lagrangian interpolants over the probability space (see e.g. [2, 10, 22]).

* The authors would like to recognize the support of the PECOS center at ICES, University of Texas at Austin (Project Number 024550, Center for Predictive Computational Science). Support from the VR project "Effektiva numeriska metoder för stokastiska differentialekvationer med tillämpningar" and King Abdullah University of Science and Technology (KAUST) is also acknowledged.

** The second and third authors have been supported by the Italian grant FIRB-IDEAS "Advanced numerical techniques for uncertainty quantification in engineering and life science problems"

¹ Applied Mathematics and Computational Science, KAUST, Saudi Arabia, e-mail: joakim.back.09@ucl.ac.uk, raul.tempone@kaust.edu.sa

² MOX, Department of Mathematics "F. Brioschi", Politecnico di Milano, Italy, e-mail: fabio.nobile@polimi.it, lorenzo.tamellini@mail.polimi.it

The comparison between performances of these deterministic methods is a matter of study (see e.g. [3]). However, both suffer the so-called ‘‘Curse of Dimensionality’’: using naive projections/interpolations over tensor product polynomials spaces/tensor grids leads to computational costs that grow exponentially fast with the number of input random variables. In such a case, careful construction of approximation spaces/sparse grids is needed in order to retain accuracy while keeping computational work acceptably low.

In a Stochastic Galerkin setting this requirement can be translated to the implementation of algorithms able to compute what is known as ‘‘Best M -Terms approximation’’. In other words, the method should be able to establish a-priori the set of the M most fruitful multivariate polynomials in the spectral approximation of the solution u , and to compute only those terms.

Important contributions in the study of the Best M -Terms approximation have been given by Cohen, DeVore and Schwab: estimates on the decay of the coefficients of the spectral expansion of u have been proved e.g. in [6, 8, 9]. In this work we will reformulate and slightly generalize the result given in [9, Corollary 6.1], and show on few numerical examples that the sequence of polynomial subspaces built upon those estimates (‘‘TD with Factorial Correction’’ sets, TD-FC in the following) performs better than classical choices such as Total Degree or Tensor Product in terms of error versus the dimension of the polynomial space.

In a Stochastic Collocation setting, the construction of an optimal grid can be recast into a classical knapsack problem relying on the notion of profit of each hierarchical surplus composing the sparse grid, as introduced e.g. in [7, 13] for approximation of H_{mix}^r functions. The ‘‘Best M -Terms’’ grid is then the one built with the set of the M most profitable hierarchical surpluses. In this work we provide a heuristic a-priori estimate of the profit of each hierarchical surplus, and use it to build a quasi optimal sparse grid. The estimates of the profit are in turn based on the estimates of the decay of the spectral expansion of u . Numerical investigations show that these new grids perform better than standard Smolyak grids as well as grids constructed with the dimension adaptive approach developed in [11, 14].

The paper is organized as follows. Section 1 defines the elliptic model problem of interest and gives general regularity results of the solution u . In Section 2 we first address the general procedure that leads to the Stochastic Galerkin approximation of u ; next we state the estimate for the decay of the spectral approximation of u and explain how to build practically the TD-FC polynomial subspaces that stem from it. In Section 2.2 we consider some simple numerical tests where we can build explicitly the Best M -Terms approximation, and we compare it with the TD-FC and with some standard choices of polynomial subspaces. Finally, the construction of the approximated optimal sparse grids and their numerical testing are presented in Section 3.

1. PROBLEM SETTING

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

$$\begin{cases} -\operatorname{div}(a(\mathbf{x}, \omega) \nabla u(\mathbf{x}, \omega)) = f(\mathbf{x}) & \mathbf{x} \in D, \\ u(\mathbf{x}, \omega) = 0 & \mathbf{x} \in \partial D. \end{cases} \quad (1)$$

where the operators div and ∇ imply differentiation with respect to the physical coordinate \mathbf{x} only. We make the following assumptions on the random diffusion coefficient:

Assumption 1.1. $a(\mathbf{x}, \omega)$ is strictly positive and bounded with probability 1, i.e. there exist $a_{min} > 0$ and $a_{max} < \infty$ such that $P(a_{min} \leq a(\mathbf{x}, \omega) \leq a_{max}, \forall \mathbf{x} \in \overline{D}) = 1$.

Assumption 1.2. $a(\mathbf{x}, \omega)$ is parametrized by a set of N independent and identically distributed uniform random variables in $(-1, 1)$, $\mathbf{y}(\omega) = [y_1(\omega), \dots, y_N(\omega)]^T : \Omega \rightarrow \mathbb{R}^N$.

Observe that Assumption 1.2 is not that restrictive. Indeed one could assume that a is parametrized by a non uniform random vector \mathbf{z} and introduce a smooth non linear map $\mathbf{y} = \Theta(\mathbf{z})$ that transforms the original variables into i.i.d. uniform variables.

We denote by $\Gamma_n = (-1, 1)$ the image set of the random variable y_n , and $\Gamma = \Gamma_1 \times \dots \times \Gamma_N$. After Assumption 1.2 the random vector \mathbf{y} has a joint probability density function $\rho : \Gamma \rightarrow \mathbb{R}_+$ that factorizes as $\rho(\mathbf{y}) = \prod_{n=1}^N \rho_n(y_n)$, $\forall \mathbf{y} \in \Gamma$, with $\rho_n = \frac{1}{2}$. Moreover, the solution u of (1) depends on the single realization $\omega \in \Omega$ only through the value taken by the random vector \mathbf{y} . We can therefore replace the probability space (Ω, \mathcal{F}, P) with $(\Gamma, B(\Gamma), \rho(\mathbf{y})d\mathbf{y})$, where $B(\Gamma)$ denotes the Borel σ -algebra on Γ and $\rho(\mathbf{y})d\mathbf{y}$ is the distribution measure of the vector \mathbf{y} . We denote with $L_\rho^2(\Gamma)$ the space of square integrable functions on Γ with respect to the measure $\frac{1}{2^N}d\mathbf{y}$.

The assumption of independence of the random variables is very convenient for the development of the techniques proposed below, since they rely on tensor polynomial approximations. However, such assumption is not essential and could be removed whenever the density ρ does not factorize, by introducing an auxiliary density $\hat{\rho} = \frac{1}{2^N}$ as suggested in [2]. The price to pay in the convergence estimate is then a constant factor proportional to $\|\rho/\hat{\rho}\|_{L^\infty(\Omega)}$.

In the rest of the paper we will use the following notation: given a multi-index $\mathbf{i} \in \mathbb{N}^N$ and a vector $\mathbf{r} \in \mathbb{R}^N$, we define $|\mathbf{i}| = \sum_{n=1}^N i_n$, $\mathbf{i}! = \prod_{n=1}^N (i_n!)$ and $\mathbf{r}^{\mathbf{i}} = \prod_{n=1}^N r_n^{i_n}$. We can now state a regularity assumption on $a(\mathbf{x}, \mathbf{y})$:

Assumption 1.3. $a(\mathbf{x}, \mathbf{y})$ is infinitely many times differentiable with respect to \mathbf{y} and $\exists \mathbf{r} \in \mathbb{R}_+^N$, $\mathbf{r} = [r_1, \dots, r_N]$ independent of \mathbf{y} and \mathbf{i} s.t.

$$\left\| \frac{\partial_{\mathbf{i}} a}{a}(\mathbf{y}) \right\|_{L^\infty(D)} \leq \mathbf{r}^{\mathbf{i}} \quad \forall \mathbf{y} \in \Gamma,$$

where \mathbf{i} is a multi-index in \mathbb{N}^N , $\partial_{\mathbf{i}} a = \frac{\partial^{i_1 + \dots + i_N} a}{\partial y_1^{i_1} \dots \partial y_N^{i_N}}$.

Remark 1.4. A common situation of interest is when $a(\mathbf{x}, \omega)$ is an infinitely dimensional random field, suitably expanded in series (e.g. by a Karhunen-Loève or Fourier expansion) either as a *linear* expansion of the form $a = a_0 + \sum_{n=1}^{\infty} b_n(\mathbf{x})y_n$ or an *exponential* expansion of the form $a = a_0 + \exp(\sum_{n=1}^{\infty} b_n(\mathbf{x})y_n)$, with $b_n \in L^\infty(D)$ in both cases. Then the infinite series is truncated up to N terms, with N large enough to take into account a sufficient amount of the total variability. Both expansions comply with the previous assumption with $r_n = \|b_n\|_{L^\infty(D)}/a_{min}$ and $r_n = \|b_n\|_{L^\infty(D)}$, respectively.

Finally, we denote by $V = H_0^1(D)$, the space of square integrable functions in D with square integrable distributional derivatives and with zero trace on the boundary, equipped with the gradient norm $\|v\|_V = \|\nabla v\|_{L^2(D)}$, $\forall v \in V$. Its dual space will be denoted by V' . We are now in the position to write a weak formulation of problem (1):

Weak Formulation. Find $u \in V \otimes L_\rho^2(\Gamma)$ such that $\forall v \in V \otimes L_\rho^2(\Gamma)$

$$\int_\Gamma \int_D a(\mathbf{x}, \mathbf{y}) \nabla u(\mathbf{x}, \mathbf{y}) \cdot \nabla v(\mathbf{x}, \mathbf{y}) \rho(\mathbf{y}) d\mathbf{x} d\mathbf{y} = \int_\Gamma \int_D f(\mathbf{x}) v(\mathbf{x}, \mathbf{y}) \rho(\mathbf{y}) d\mathbf{x} d\mathbf{y}. \quad (2)$$

Under Assumption 1.1, a straightforward application of the Lax-Milgram lemma yields that there exists a unique solution to problem (2) for any $f \in V'$. Moreover, the following estimate holds:

$$\|u\|_{V \otimes L_\rho^2(\Gamma)} \leq \frac{\|f\|_{V'}}{a_{min}}.$$

Concerning the regularity of the solution with respect to \mathbf{y} , under Assumptions 1.1 - 1.3 the following result is proved in [5].

$$\textbf{Regularity result:} \quad \|\partial_{\mathbf{i}} u(\mathbf{y})\|_V \leq C_0 |\mathbf{i}|! \bar{\mathbf{r}}^{\mathbf{i}} \quad \forall \mathbf{y} \in \Gamma$$

which implies that u is analytic in every $\mathbf{y} \in \Gamma$. Here $C_0 = \frac{\|f\|_{V'}}{\alpha_{min}}$ and $\tilde{\mathbf{r}} = \frac{3}{2}\mathbf{r}$, with \mathbf{r} as in Assumption 1.3. A similar result is given in [9] for the special case $a = a_0 + \sum_{n=1}^N b_n(x)y_n$.

2. STOCHASTIC GALERKIN METHOD

We now seek an approximation of the solution u with respect to \mathbf{y} by global polynomials. As anticipated in the introduction, we remark that the choice of the polynomial space is critical when the number N of input random variables is large, since the number of stochastic degrees of freedom might grow very quickly with N , even exponentially when isotropic tensor product polynomial spaces are used (see Table 1). This effect is known as the *curse of dimensionality*.

Several choices of polynomial spaces that mitigate this phenomenon have been proposed in the literature, see e.g. [3]. In this work we consider a general multivariate space: let $w \in \mathbb{N}$ be an integer index denoting the level of approximation and $\mathbf{p} = (p_1, \dots, p_N)$ a multi-index. Let $\Lambda(w)$ be a sequence of increasing index sets such that $\Lambda(0) = \{(0, \dots, 0)\}$ and $\Lambda(w) \subseteq \Lambda(w+1) \subset \mathbb{N}^N$, for $w \geq 0$. We introduce the multivariate $\rho(\mathbf{y})d\mathbf{y}$ -orthonormal Legendre polynomials $\mathcal{L}_{\mathbf{p}}(\mathbf{y}) = \prod_{n=1}^N L_{p_n}(y_n)$, where $L_{p_n}(t)$ are the monodimensional Legendre polynomials of degree p_n , and consider the multivariate polynomial subspace of $L^2_\rho(\Gamma)$ built as

$$\mathbb{P}_{\Lambda(w)}(\Gamma) = \text{span} \{ \mathcal{L}_{\mathbf{p}}(\mathbf{y}) \text{ with } \mathbf{p} \in \Lambda(w) \}.$$

We then seek an approximation $u_w \in V \otimes \mathbb{P}_{\Lambda(w)}(\Gamma)$. Table 1 shows common choices for $\Lambda(w)$.

	index set $\Lambda(w)$	Dimension $ \Lambda(w) $
Tensor product space (TP)	$\{\mathbf{p} \in \mathbb{N}^N : \max_{n=1, \dots, N} p_n \leq w\}$	$(w+1)^N$
Total degree space (TD)	$\{\mathbf{p} \in \mathbb{N}^N : \sum_{n=1}^N p_n \leq w\}$	$\binom{N+w}{N}$
Hyperbolic cross space (HC)	$\{\mathbf{p} \in \mathbb{N}^N : \prod_{n=1}^N (p_n + 1) \leq w + 1\}$	$\leq (w+1) (\log(e(w+1)))^{N-1}$

TABLE 1. Common polynomial spaces. The result for HC is only an upper bound.

For further details on these spaces, see [3] and references therein. In [3] we have also considered anisotropic versions of these spaces as in Table 2, where $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N) \in \mathbb{R}_+^N$ is a vector of positive weights and $\alpha_{min} = \min_n \alpha_n$.

Tensor product space (TP)	$\Lambda(w) = \{\mathbf{p} \in \mathbb{N}^N : \max_{n=1, \dots, N} \alpha_n p_n \leq \alpha_{min} w\}$
Total degree space (TD)	$\Lambda(w) = \{\mathbf{p} \in \mathbb{N}^N : \sum_{n=1}^N \alpha_n p_n \leq \alpha_{min} w\}$
Hyperbolic cross space (HC)	$\Lambda(w) = \{\mathbf{p} \in \mathbb{N}^N : \prod_{n=1}^N (p_n + 1)^{\frac{\alpha_n}{\alpha_{min}}} \leq w + 1\}$

TABLE 2. Anisotropic version of polynomial spaces

We can interpret these weights as a measure of the importance of each random variable y_n on the solution: the smaller the weight, the higher degree we allow in the corresponding variable. The Stochastic Galerkin (SG) approximation then consists in restricting the weak formulation (2) to the subspace $V \otimes \mathbb{P}_{\Lambda(w)}(\Gamma)$ and reads:

Galerkin Formulation. Find $u_w \in V \otimes \mathbb{P}_{\Lambda(w)}(\Gamma)$ such that $\forall v_w \in V \otimes \mathbb{P}_{\Lambda(w)}(\Gamma)$

$$\int_{\Gamma} \int_D a(\mathbf{x}, \mathbf{y}) \nabla u_w(\mathbf{x}, \mathbf{y}) \cdot \nabla v_w(\mathbf{x}, \mathbf{y}) \rho(\mathbf{y}) d\mathbf{x} d\mathbf{y} = \int_{\Gamma} \int_D f(\mathbf{x}) v_w(\mathbf{x}, \mathbf{y}) \rho(\mathbf{y}) d\mathbf{x} d\mathbf{y}. \quad (3)$$

We recall that the choice of the family of polynomials depends on the probability measure considered over Γ : for instance we could use Hermite polynomials for Gaussian measures and Laguerre polynomials for Exponential measures (see e.g. [23] for the general Askey scheme).

Now let $\phi(\mathbf{x})$ be a basis function for the physical space V . Inserting $v_w = \phi(\mathbf{x})\mathcal{L}_{\mathbf{q}}(\mathbf{y})$ with $\mathbf{q} \in \Lambda(w)$ as test function in the weak formulation (3) will result in a set of equations in weak form for the coefficients $u_{\mathbf{p}}(\mathbf{x})$ that will be generally coupled due to the presence in the equation (3) of the term $a(\mathbf{x}, \mathbf{y})\mathcal{L}_{\mathbf{p}}(\mathbf{y})\mathcal{L}_{\mathbf{q}}(\mathbf{y})$. See e.g. [3, 17, 18] for further details on space discretization and on the numerical solution of the system of equations.

2.1. Optimal choice of polynomial spaces

The question that naturally arises in the context of Galerkin approximation concerns the best choice of the polynomial space to be used, to get maximum accuracy for a given dimension M of the space (Best M -Terms approximation). Let us assume the solution $u(\mathbf{x}, \mathbf{y})$ is known and consider its expansion on Legendre polynomials,

$$u(\mathbf{x}, \mathbf{y}) = \sum_{\mathbf{p} \in \mathbb{N}^N} u_{\mathbf{p}}(\mathbf{x})\mathcal{L}_{\mathbf{p}}(\mathbf{y}), \quad u_{\mathbf{p}}(\mathbf{x}) = \int_{\Gamma} u(\mathbf{x}, \mathbf{y})\mathcal{L}_{\mathbf{p}}(\mathbf{y})\rho(\mathbf{y})d\mathbf{y}.$$

We look for an index set $\mathcal{S} \subset \mathbb{N}^N$ with cardinality M that minimizes the error

$$\|u - \sum_{\mathbf{p} \in \mathcal{S}} u_{\mathbf{p}}\mathcal{L}_{\mathbf{p}}\|_{V \otimes L^2(\Gamma)}^2 = \sum_{\mathbf{p} \notin \mathcal{S}} \|u_{\mathbf{p}}\|_V^2$$

where the equivalence is a consequence of Parseval's equality. The obvious solution to this problem is the set \mathcal{S} that contains the M coefficients $u_{\mathbf{p}}$ with largest norm. This solution of course is not constructive; what we need are sharp estimates of the decay of the coefficients $\|u_{\mathbf{p}}\|_V$, based only on computable quantities, to be used in the approximation of the set \mathcal{S} .

Seminal works in this direction are [6, 8, 9], where estimates of the decay of the Legendre coefficients are provided. We consider here a slight generalization of the result in [9, Corollary 6.1] and show numerically that the polynomial sets, hereafter called TD-FC ("TD with Factorial Correction"), built on these modified estimates behave closely to the real Best M -Terms approximation.

Under Assumptions 1.1 - 1.3 it is possible to prove that the following estimate holds for the Legendre coefficients:

$$\|u_{\mathbf{p}}\|_V \leq C_0 e^{-\sum_n g_n p_n} \frac{|\mathbf{p}|!}{\mathbf{p}!} \quad (4)$$

with $g_n = -\log(r_n/(\sqrt{3}\log 2))$ and r_n as in Assumption 1.3. For a proof of (4) see [5]. Again, a similar result is given in [9] for the special case $a = a_0 + \sum_{n=1}^N b_n(x)y_n$.

We define now the sequence of TD-FC sets, with increasing approximating accuracy, by selecting all multi-indices \mathbf{p} for which the *estimated decay* of the corresponding Legendre coefficient, given in (4), is above a fixed threshold ϵ . This, in turn, corresponds to selecting those indices \mathbf{p} such that

$$e^{-\sum_{n=1}^N g_n p_n} \frac{|\mathbf{p}|!}{\mathbf{p}!} \geq \epsilon \quad \iff \quad -\sum_{n=1}^N g_n p_n + \log \frac{|\mathbf{p}|!}{\mathbf{p}!} \geq \log \epsilon.$$

The TD-FC sets are then defined as

$$\Lambda(w) = \left\{ \mathbf{p} \in \mathbb{N}^N : \sum_{n=1}^N g_n p_n - \log \frac{|\mathbf{p}|!}{\mathbf{p}!} \leq w \right\} \quad (5)$$

with $w \in \mathbb{N}^+ = \lceil -\log \epsilon \rceil$. The quantities $g_n = -\log(r_n/(\sqrt{3}\log 2))$ appearing in (5) can be estimated *a-priori* (see Remark 1.4). Note however that none of the Assumptions 1.1-1.3 nor Remark 1.4 prescribe bounds on the magnitude of r_n , so that the rates g_n are not guaranteed to have always the same sign. This in turn implies that estimate (4) does not always predict that $\|u_{\mathbf{p}}\|_V \rightarrow 0$ when $|\mathbf{p}| \rightarrow \infty$. On the one hand this is clearly

unsatisfactory; on the other hand it is found that bound (4) works particularly well if the rates g_n are estimated numerically rather than *a-priori*, as the following numerical results will show. To estimate numerically g_n , one increases the polynomial degree in one random variable at a time while keeping degree zero in all the others variables and estimate numerically the exponential rate of convergence. Observe that in such monivariate analyses the factorial term does not appear so the expected convergence rate is precisely $\sim e^{-g_n p_n}$.

Remark 2.1. Observe that $\Lambda(w)$ actually depends on N but one can extend this also to the case where \mathbf{p} is a sequence of natural numbers (“infinite dimensions multi-indices”) with only a finite number of non zero terms, provided $g_n \rightarrow +\infty$ as $n \rightarrow \infty$. This is an alternative way to work with random fields, without truncating them a priori to a certain level (see e.g. [8, 16]).

2.2. Numerical Tests

In this section we show the performance of the TD-FC sets (5) compared to the isotropic and anisotropic versions of TD sets defined in Tables 1 and 2 as well as the Best M -Terms approximation. We consider the following elliptic problem in one physical dimension

$$\begin{cases} -(a(x, \mathbf{y})u(x, \mathbf{y}))' = 1 & x \in D = (0, 1), \mathbf{y} \in \Gamma \\ u(0, \mathbf{y}) = u(1, \mathbf{y}) = 0, & \mathbf{y} \in \Gamma \end{cases} \quad (6)$$

with different choices of diffusion coefficient $a(x, \mathbf{y})$, for which Assumptions 1.1 - 1.3 hold. We focus on a linear functional $\psi : V \rightarrow \mathbb{R}$, so that $\psi(u)$ is a scalar random variable, function of \mathbf{y} only. In our examples, ψ is defined as $\psi(v) = v(\frac{1}{2})$.

To obtain the Best M -Terms approximation we compute explicitly all the Legendre coefficients of $\psi(u)$ in a sufficiently large index set \mathbb{U} , evaluating the integrals $\psi_{\mathbf{p}} = \int_{\Gamma} \psi(u) L_{\mathbf{p}}(\mathbf{y}) \rho(\mathbf{y}) d\mathbf{y}$ with a high-level sparse grid. We order then the coefficients in decreasing order, according to their V norm and take the partial sums of the reordered sequence as the Best M -Terms approximation. The rates \mathbf{g} used to build the TD-FC space, as well as the anisotropic TD space (with $\alpha_n = g_n$), are computed numerically as explained in the previous Section.

Test 1: diffusion coefficient depending on 2 random variables

The first case we consider has two random variables (y_1, y_2) and a diffusion coefficient $a(x, \mathbf{y}) = 1 + 0.1xy_1 + 0.5x^2y_2$; results are shown in Figure 1.

Figure 1(a) shows the Legendre coefficients ordered in lexicographic order, giving this peculiar sawtooth shape. The first tooth corresponds to multi-indices of the form $[0, k]$, the second one to $[1, k]$ and so on. We have also added to the plot the estimate (4) of the magnitude of the Legendre coefficients, which leads to the TD-FC sets (5), as well as the estimate $|\psi_{\mathbf{p}}| \leq C_0 e^{-\sum_n g_n p_n}$, where the factorial terms have been dropped, which leads to the anisotropic TD spaces. The plot suggests that estimate (4) is quite sharp, whereas the estimate corresponding to the TD space underestimates considerably the Legendre coefficients. This highlights the importance of the factorial term in (4). We expect, therefore, that the TD-FC approximation performs better than the aniso-TD one. Moreover, we point out the non intuitive fact that the Legendre coefficients $\psi_{\mathbf{p}}$ are not strictly decreasing in absolute value in the lexicographic order. As an example, $|\psi_{[5\ 0]}| < |\psi_{[5\ 1]}|$, and the same holds for all teeth but the first few.

Figure 1(b) shows convergence plots for the error in L_{ρ}^2 -norm squared for the various polynomial spaces used versus the dimension of the polynomial space. As the TD-FC sequence is the only sequence that captures correctly the non decreasing behaviour of the Legendre coefficients in lexicographic order, the convergence of the TD-FC sequence in Figure 1(b) is the closest to the Best M -Terms approximation, even though the anisotropic TD spaces give good results as well. We also point out the poor performance of the standard isotropic TD space compared to both the anisotropic TD and the TD-FC spaces: this confirms the importance of using anisotropic spaces to reduce computational cost.

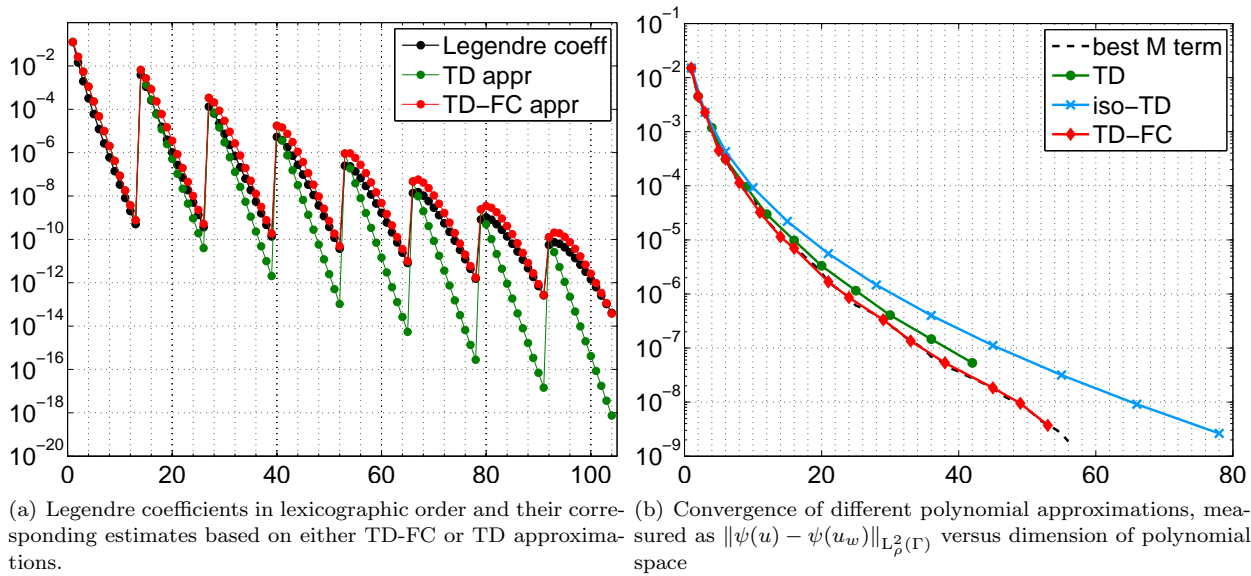


FIGURE 1. Results for $a(x, \mathbf{y}) = 1 + 0.1xy_1 + 0.5x^2y_2$. Here we have $g \simeq (2.96, 1.57)$, $\mathbb{U} = \text{TP}(12)$, Legendre coefficients computed with a standard Smolyak sparse grid of level 9, with Gauss-Legendre abscissae.

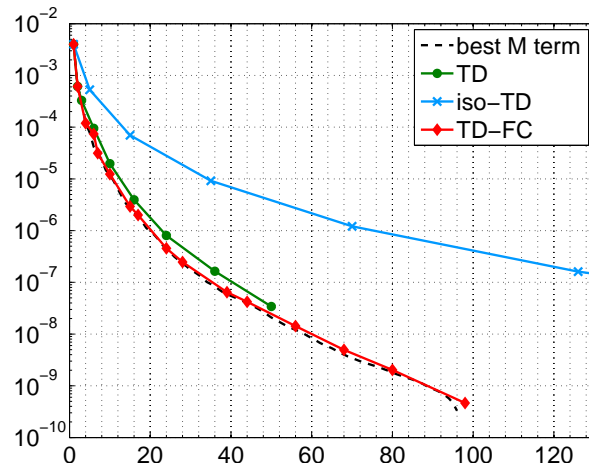


FIGURE 2. Results for $a(x, \mathbf{y}) = 4 + y_1 + 0.2 \sin(\pi x)y_2 + 0.04 \sin(2\pi x)y_3 + 0.008 \sin(3\pi x)y_4$. Here we have $g \simeq (2.035, 4.11, 5.73, 7.05)$, $\mathbb{U} = \text{TD}(9)$. The convergence is measured by $\|\psi(u) - \psi(u_w)\|_{L^2(\Gamma)}$ versus the dimension of the polynomial space.

Test 2: diffusion coefficient depending on 4 random variables

We now consider $a(x, \mathbf{y}) = 4 + y_1 + 0.2 \sin(\pi x)y_2 + 0.04 \sin(2\pi x)y_3 + 0.008 \sin(3\pi x)y_4$, and look at the functional $\psi(v) = v(0.7)$ (the functional $\psi(v) = v(1/2)$ is not suited for analysis in this case as, by symmetry, many of the Legendre coefficients are zero). Figure 2 shows the results, and again we see that the TD-FC approximation is the best performing, with anisotropic TD closely following and isotropic TD far worse.

3. STOCHASTIC COLLOCATION

The Stochastic Collocation (SC) Finite Element method consists in collocating problem (1) in a set of points $\{\mathbf{y}_j \in \Gamma, j = 1, \dots, M_w\}$, i.e. computing the corresponding solutions $u(\cdot, \mathbf{y}_j)$ and building a global polynomial approximation u_w , not necessarily interpolatory, upon those evaluations: $u_w(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{M_w} u(\mathbf{x}, \mathbf{y}_j) \tilde{\psi}_j(\mathbf{y})$ for suitable multivariate polynomials $\{\tilde{\psi}_j\}_{j=1}^{M_w}$.

Building the set of evaluation points $\{\mathbf{y}_j\}$ as a cartesian product of monodimensional grids becomes quickly unfeasible since the computational cost grows exponentially fast with the number of stochastic dimensions needed. We consider instead the so-called *sparse grid* procedure, originally introduced by Smolyak in [19] for high dimensional quadrature purposes; see also [4, 7] for polynomial interpolation. In the following we briefly review and generalize this construction.

For each direction y_n we introduce a sequence of one dimensional polynomial interpolant operators of increasing order: $\mathcal{U}_n^{m(i)} : C^0(\Gamma_n) \rightarrow \mathbb{P}_{m(i)-1}(\Gamma_n)$. Here $i \geq 1$ denotes the level of approximation and $m(i)$ the number of collocation points used to build the interpolation at level i . As a consequence, $\mathcal{U}_n^{m(i)}[q] = q$ if q is a polynomial of degree up to $m(i) - 1$. We require the function m to satisfy the following assumptions: $m(0) = 0$, $m(1) = 1$ and $m(i) < m(i + 1)$ for $i \geq 1$. In addition, let $\mathcal{U}_n^0 = 0, \forall q \in C^0(\Gamma_n)$.

Next we introduce the difference operators $\Delta_n^{m(i)} = \mathcal{U}_n^{m(i)} - \mathcal{U}_n^{m(i-1)}$, an integer value $w \geq 0$ and a sequence of index sets $\mathcal{I}(w)$ such that $\mathcal{I}(w) \subset \mathcal{I}(w + 1)$ and $\mathcal{I}(0) = \{(1, 1, \dots, 1)\}$. We define the sparse grid approximation of $u(\mathbf{y}) : \Gamma \rightarrow \mathbb{V}$ at level w as

$$u_w(\mathbf{y}) = \mathcal{S}_{\mathcal{I}(w)}^m[u](\mathbf{y}) = \sum_{\mathbf{i} \in \mathbb{N}_+^N : \mathbf{i} \in \mathcal{I}(w)} \bigotimes_{n=1}^N \Delta_n^{m(i_n)}[u](\mathbf{y}). \quad (7)$$

The set of all evaluation points needed is called *sparse grid* and denoted by $\mathcal{H}_{\mathcal{I}(w)}^m \subset \Gamma$. To fully characterize the sparse approximation operator $\mathcal{S}_{\mathcal{I}(w)}^m$ one has to provide the sequence of sets $\mathcal{I}(w)$, the relation between the level i and the number of points $m(i)$ in the corresponding one dimensional polynomial interpolation formula $\mathcal{U}^{m(i)}$, and the family of points to be used at each level, e.g. Clenshaw-Curtis or Gauss abscissae (see e.g. [21]).

Remark 3.1. As pointed out in [11], the sparse approximation is well defined only if the sum (7) is actually a telescopic sum. This is not ensured by any arbitrary \mathcal{I} , and we have to pose some additional constraints on \mathcal{I} . Following [11] we say that a set \mathcal{I} is *admissible* if for all $\mathbf{i} \in \mathcal{I}$

$$\mathbf{i} - \mathbf{e}_j \in \mathcal{I} \text{ for } 1 \leq j \leq N, i_j > 1. \quad (8)$$

We refer to this property as *admissibility condition*, or *ADM* in short. Given a set \mathcal{I} we will denote by \mathcal{I}^{ADM} the smallest set such that $\mathcal{I} \subset \mathcal{I}^{ADM}$ and \mathcal{I}^{ADM} is admissible.

In what follows we will consider Clenshaw-Curtis abscissae and the “doubling” rule $m(i) = db(i) = 2^{i-1} + 1$, which leads to nested grids. The classical Smolyak sparse grid (SM) uses $\mathcal{I}(w) = \{\mathbf{i} \in \mathbb{N}_+^N : |\mathbf{i} - \mathbf{1}| \leq w\}$, which clearly satisfies the admissibility condition (8). A quasi optimal choice of $\mathcal{I}(w)$ will be discussed in the next Section.

3.1. Quasi-optimal sparse grids

We now aim at constructing the quasi-optimal sparse grid for Stochastic collocation method, i.e. at choosing the best sequence of sets of indices. We will rely on the estimate (4) on the decay of the Legendre expansion of u .

To this end, let us define the error associated to a sparse grid as $E(\mathcal{S}) = \left\| u - \mathcal{S}_{h,w}^m[u] \right\|_{\mathbb{V} \otimes L_\rho^2(\Gamma)}$, and the work $W(\mathcal{S})$ as the number of evaluations needed, i.e. $W(\mathcal{S}) = |\mathcal{H}_{\mathcal{I}(w)}^m|$. Next we define the error and work contribution

of a multi-index \mathbf{i} . Let \mathcal{I} be any set of indices such that $\mathbf{i} \notin \mathcal{I}$ and $\{\mathcal{I} \cup \mathbf{i}\}$ is admissible. Then the error contribution of \mathbf{i} is $\Delta E(\mathbf{i}) = \left\| \mathcal{S}_{\{\mathcal{I} \cup \mathbf{i}\}}^m[u] - \mathcal{S}_{\mathcal{I}}^m[u] \right\|_{V \otimes L_p^2(\Gamma)}$ and the work contribution is $\Delta W(\mathbf{i}) = |W(\mathcal{S}_{\{\mathcal{I} \cup \mathbf{i}\}}^m) - W(\mathcal{S}_{\mathcal{I}}^m)|$.

Following [7, 11] we can define the profit of an index \mathbf{i} as

$$P(\mathbf{i}) = \frac{\Delta E(\mathbf{i})}{\Delta W(\mathbf{i})}$$

and define the optimal sparse approximation operator \mathcal{S}^* as the one using the set of most profitable indices, i.e. $\mathcal{I}^*(\epsilon) = \{\mathbf{i} \in \mathbb{N}_+^N : P(\mathbf{i}) \geq \epsilon\}$.

To build the set \mathcal{I} we rely on estimates for both $\Delta E(\mathbf{i})$ and $\Delta W(\mathbf{i})$. Since using Clenshaw-Curtis abscissae and doubling rule $db(\cdot)$ we get nested grids, we can compute *exactly* $\Delta W(\mathbf{i})$ as

$$\Delta W(\mathbf{i}) = \prod_{n=1}^N (db(i_n) - db(i_n - 1)). \quad (9)$$

On the other hand an estimate of the error contribution $\Delta E(\mathbf{i})$ requires some additional effort. We conjecture that the decay of $\Delta E(\mathbf{i})$ is related to the decay of the Legendre coefficients, through the Lebesgue constant $\mathbb{L}(m(\mathbf{i}))$ associated to $\bigotimes_{n=1}^N \mathcal{U}_n^{m(i_n)}$:

$$\Delta E(\mathbf{i})[u] \lesssim \|u_{m(\mathbf{i}-1)}\|_V \mathbb{L}(m(\mathbf{i})), \quad (10)$$

where $a \lesssim b$ means that there exists a constant c independent of \mathbf{i} such that $a \leq cb$, $u_{m(\mathbf{i}-1)}$ is the Legendre coefficient associated to the multi-index $m(\mathbf{i}-1)$, and for Clenshaw-Curtis abscissae with doubling relation the Lebesgue constant is

$$\mathbb{L}(db(\mathbf{i})) \leq \prod_{n=1}^N \left(\frac{2}{\pi} \log(db(i_n) + 1) + 1 \right).$$

Figure 3 shows the decay of the error $\Delta E(\mathbf{i})$ and the corresponding $m(\mathbf{i}-1)$ -th Legendre coefficients for the quantity $|\psi(u)|$, where $u(x, \mathbf{y})$ solves problem (6) with $a(x, \mathbf{y}) = 1 + 0.1xy_1 + 0.5x^2y_2$, as in Section 2.2, Test 1. These results, as well as the ones presented in the next Section, confirm that estimate (10) is accurate enough for our purposes.

Starting from (9) and (10), we can estimate the profit of each index, and estimate the sequence $\mathcal{S}_{\mathcal{I}^*(\epsilon)}$ of quasi-optimal grids with

$$\mathcal{I}^*(\epsilon) = \left\{ \mathbf{i} \in \mathbb{N}_+^N : \frac{C_0 \exp \left(- \sum_{n=1}^N db(i_n - 1) g_n \right) \frac{|db(\mathbf{i}-1)|!}{db(\mathbf{i}-1)!} \mathbb{L}(db(\mathbf{i}))}{\prod_{n=1}^N (db(i_n) - db(i_n - 1))} \geq \epsilon \right\}^{ADM} \quad (11)$$

with $\epsilon > 0 \in \mathbb{R}$. Equivalently, for $w = 0, 1, \dots$ we can define the sequence of sets

$$\mathcal{I}^*(w) = \left\{ \mathbf{i} \in \mathbb{N}_+^N : \sum_{i=n}^N db(i_n - 1) g_n - \log \frac{|db(\mathbf{i}-1)|!}{db(\mathbf{i}-1)!} - \sum_{n=1}^N \log \frac{\frac{2}{\pi} \log(db(i_n) + 1) + 1}{db(i_n) - db(i_n - 1)} \leq w \right\}^{ADM} \quad (12)$$

that will be used in (7) to build the quasi optimal sparse grids. In analogy with the quasi optimal sets in Section 2.1 we will refer to these grids as ‘‘Error-Work’’, EW in short.

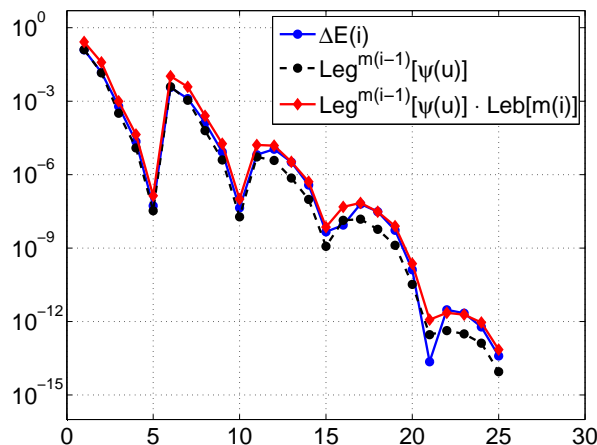


FIGURE 3. Numerical comparison between $\Delta E(\mathbf{i})$, $|\psi(u)_{m(\mathbf{i}-1)}|$ and estimate (10) for $\psi(u)$ as in Section 2.2 Test 1. Both the $\Delta E(\mathbf{i})$ for $\mathbf{i} \in TP(4)$ and the corresponding Legendre coefficients $|\psi(u)_{m(\mathbf{i}-1)}|$ have been computed with a standard sparse grid SM(10).

Remark 3.2. Observe that in definitions (11) and (12) the set has to be made admissible, as the condition on the multi-indices given inside the brackets might not satisfy *ADM*. This simply implies that if at level w an index \mathbf{j} is added, all indices $\{\mathbf{i} \in \mathbb{N}^N : i_1 \leq j_1, i_2 \leq j_2, \dots, i_N \leq j_N\}$ have to be added as well, if not already present in the set.

3.2. Numerical tests on sparse grids

In this Section we consider again problem (6) with the diffusion coefficients $a(x, \mathbf{y})$ as in Section 2.2 and use it to test the performance of the EW grids derived above, comparing them with the classical SM grid and the Best M -Terms approximation. The decay coefficients g_n in (12) are estimated numerically as in Section 2.2.

To approximate the Best M -Terms we again consider a sufficiently large set \mathbb{U} of multi-indices and for each of them we compute $\Delta W(\mathbf{i})$, $\Delta E(\mathbf{i})$ and their profit $P(\mathbf{i})$. Next, we sort the multi-indices according to $P(\mathbf{i})$, modify the sequence to fulfil the *ADM* condition (8) and compute the sparse grids according to this sequence.

We remark that the procedure just described only leads to an approximation of the Best M -Terms solution. Indeed, on the one hand replacing the total error $E(\mathcal{S})$ with the sum $\sum_{\mathbf{i}} \Delta E(\mathbf{i})$ provides only an upper bound that could be pessimistic because of possible cancellations, since the details $\Delta^{m(\mathbf{i})}[u]$ are not mutually orthogonal, in general. On the other hand the fact that the most profitable index may be not admissible suggests that the solution cannot be found using a greedy algorithm.

We also compare our results with the dimension adaptive algorithm proposed in [11], in the implementation of [14], available at <http://www.ians.uni-stuttgart.de/spinterp>. This is an adaptive algorithm that given a sparse grid $\mathcal{S}_{\mathcal{I}}$ explores all neighbour multi-indices and adds to \mathcal{I} the most profitable ones. The algorithm implemented in [14] has a tunable parameter $\tilde{\omega}$ that allows one to move continuously from the classical Smolyak formula ($\tilde{\omega} = 0$) to the fully adaptive algorithm ($\tilde{\omega} = 1$). Following [14], in the present work we have set $\tilde{\omega} = 0.9$, that numerically has been proved to be a good performing choice. The cost of this algorithm is the *total* number of evaluations needed, including also those necessary to explore all neighbours, to find the most profitable multi-index.

Figure 4 shows the convergence of the quantity $\|\psi(u) - \psi(u_w)\|_{L^2_\rho(\Gamma)}$ versus the number of grid points, for the different sparse grids considered. The L^2_ρ -norm has been computed with a high level isotropic Smolyak grid. The EW grid is the best performing, even compared to the version of the a-posteriori dimension adaptive algorithm implemented in [14], and the closest to the Best M -Terms grids sequence.

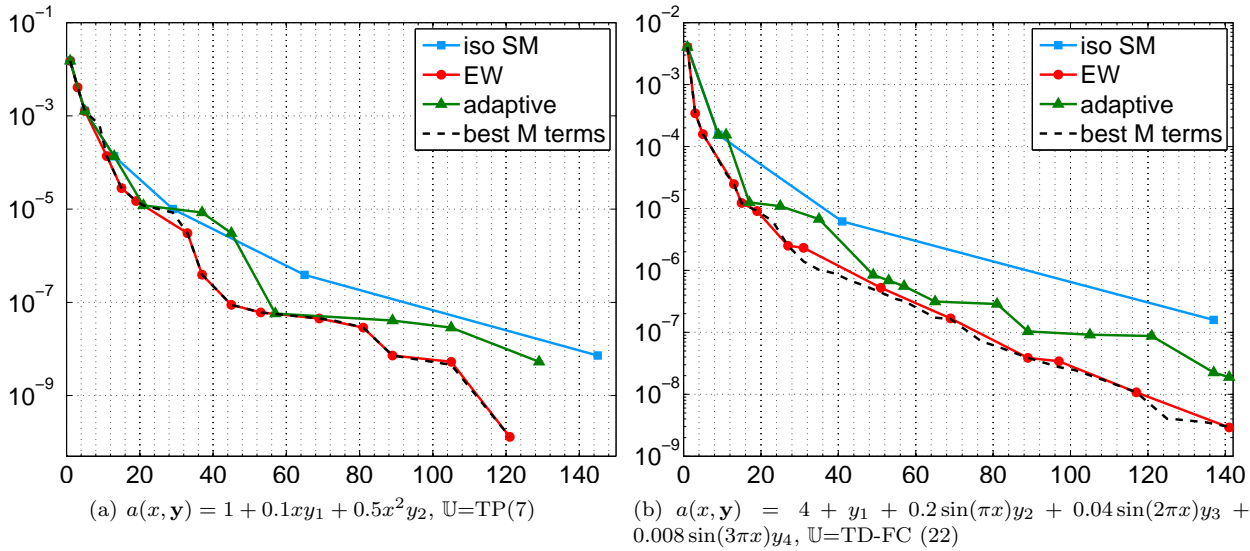


FIGURE 4. Results for EW sparse grids compared with Best M -Terms, isotropic Smolyak and dimension adaptive algorithm. Convergence is measured in terms of $\|\psi(u) - \psi(u_w)\|_{L^2(\Gamma)}$ versus the number of evaluations (grid points).

4. CONCLUSIONS

In this work we have proposed a new sequence of polynomial subspaces (TD-FC spaces in short) to be used in the solution of elliptic stochastic PDEs with Galerkin method, based on sharp estimates of the decay of the Legendre coefficients.

The performances of TD-FC spaces have been assessed through some simple test cases. Here we have compared TD-FC with some standard choices of polynomial spaces and with the Best M -Terms approximation of the solution, that can be explicitly built for the examples considered. Results show that the TD-FC spaces perform better than the anisotropic standard ones, and are close to the Best M -Terms approximation. However, the standard spaces may still have reasonable performances, if used in an appropriate anisotropic framework.

Using the estimate for the decay of the Legendre coefficients we have also defined a new class of sparse grids to be used in the context of Stochastic Collocation, relying on the concept of profit of each multi-index in the sparse grid. Again numerical tests suggest that these new sparse grids outperform the classical Smolyak construction, as well as the a-posteriori dimension adaptive algorithm as implemented in [14]. The reason for this apparent success is that our algorithm picks up the hierarchical surpluses based purely on a priori estimates and inexpensive y -one dimensional auxiliary problems. These estimates turn out to be quite sharp, and do not have any extra cost to explore neighbor points as the algorithm in [14] does.

The new polynomial spaces and sparse grids proposed here are valid in the case of analytic dependence of the solution on the random variables. We point out, however, that the general strategy outlined in Sections 2.1 and 3.1 on how to build optimal polynomial spaces / sparse grids is applicable to any problem and any kind of underlying random variables. Of course, this strategy requires a sharp estimate of the decay of the coefficients of the spectral expansion of the solution on a orthonormal hierarchical basis (not necessarily polynomial). This step is highly problem dependent and should be analyzed carefully in each situation, as we did here for a linear elliptic PDE with a stochastic coefficient dependent on uniformly distributed random variables.

REFERENCES

- [1] 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, 2004.
- [2] I. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM J. Numer. Anal.*, 45(3):1005–1034, 2007.
- [3] J. Bäck, F. Nobile, L. Tamellini, and R. Tempone. Stochastic spectral Galerkin and collocation methods for PDEs with random coefficients: a numerical comparison. In J.S. Hesthaven and E.M. Ronquist, editors, *Spectral and High Order Methods for Partial Differential Equations*, volume 76 of *Lecture Notes in Computational Science and Engineering*, pages 43–62. Springer, 2011. Selected papers from the ICOSAHOM '09 conference, June 22–26, Trondheim, Norway.
- [4] V. Barthelmann, E. Novak, and K. Ritter. High dimensional polynomial interpolation on sparse grids. *Adv. Comput. Math.*, 12(4):273–288, 2000.
- [5] J. Beck, F. Nobile, L. Tamellini, and R. Tempone. On the optimal polynomial approximation of stochastic PDEs by Galerkin and collocation methods. MOX-Report 2011-23, MOX - Department of Mathematics, Politecnico di Milano, 2011. To appear on *Mathematical Models and Methods in Applied Sciences*.
- [6] M. Bieri, R. Andreev, and C. Schwab. Sparse tensor discretization of elliptic sPDEs. SAM-Report 2009-07, Seminar für Angewandte Mathematik, ETH, Zurich, 2009.
- [7] H.J. Bungartz and M. Griebel. Sparse grids. *Acta Numer.*, 13:147–269, 2004.
- [8] A. Cohen, R. DeVore, and C. Schwab. Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs. SAM-Report 2010-03, Seminar für Angewandte Mathematik, ETH, Zurich, 2010.
- [9] A. Cohen, R. DeVore, and C. Schwab. Convergence rates of best n -term Galerkin approximations for a class of elliptic sPDEs. *Foundations of Computational Mathematics*, 10:615–646, 2010. 10.1007/s10208-010-9072-2.
- [10] B. Ganapathysubramanian and N. Zabaras. Sparse grid collocation schemes for stochastic natural convection problems. *Journal of Computational Physics*, 225(1):652–685, 2007.
- [11] T. Gerstner and M. Griebel. Dimension-adaptive tensor-product quadrature. *Computing*, 71(1):65–87, 2003.
- [12] R. G. Ghanem and P. D. Spanos. *Stochastic Finite Elements: a Spectral Approach*. Springer-Verlag, New York, 1991.
- [13] M. Griebel and S. Knapek. Optimized general sparse grid approximation spaces for operator equations. *Math. Comp.*, 78(268):2223–2257, 2009.
- [14] A. Klimke. *Uncertainty modeling using fuzzy arithmetic and sparse grids*. PhD thesis, Universität Stuttgart, Shaker Verlag, Aachen, 2006.
- [15] 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.
- [16] F. Nobile, R. Tempone, and C.G. Webster. An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. Numer. Anal.*, 46(5):2411–2442, 2008.
- [17] M.F. Pellissetti and R.G. Ghanem. Iterative solution of systems of linear equations arising in the context of stochastic finite elements. *Adv. Eng. Software*, 31:607–616, 2000.
- [18] C.E. Powell and H.C. Elman. Block-diagonal preconditioning for spectral stochastic finite-element systems. *IMA J. Numer. Anal.*, 29(2):350–375, 2009.
- [19] S.A. Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Dokl. Akad. Nauk SSSR*, 4:240–243, 1963.
- [20] R. A. Todor and C. Schwab. Convergence rates for sparse chaos approximations of elliptic problems with stochastic coefficients. *IMA J Numer Anal*, 27(2):232–261, 2007.
- [21] Lloyd N. Trefethen. Is Gauss quadrature better than Clenshaw-Curtis? *SIAM Rev.*, 50(1):67–87, 2008.
- [22] D. Xiu and J.S. Hesthaven. High-order collocation methods for differential equations with random inputs. *SIAM J. Sci. Comput.*, 27(3):1118–1139, 2005.
- [23] D. Xiu and G.E. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24(2):619–644, 2002.