An adaptive sparse grid method for elliptic PDEs with stochastic coefficients

Jocelyne Erhel, Zoubida Mghazli, Mestapha Oumouni

An adaptive sparse grid method for elliptic PDEs with stochastic coefficients

Jocelyne Erhel, Zoubida Mghazli, Mestapha Oumouni

Abstract

The stochastic collocation method based on the anisotropic sparse grid has become a significant tool to solve partial differential equations with stochastic inputs. The aim is to seek a vector weight and a convenient level for the method. The classical approach uses a posteriori approach on the solution, yielding to an additional prohibitive cost with a large stochastic dimension.

In this work, we discuss an adaptive approach of this method to calculate the statistics of the solution. It is based on an adaptive approximation of the *inverse* diffusion parameter. We construct an efficient error indicator which is an upper bound of the error on the solution. In the case of unbounded variables, we use an appropriate error estimation to compute suitable weights of the method. Numerical examples are presented to confirm the efficiency of the approach and showing that the cost is considerably reduced without loss of accuracy.

key words: Elliptic PDEs with random coefficients, Stochastic collocation method, Anisotropic sparse grid, Adaptation method.

1 Introduction

The Monte Carlo method [9] is the most classical approach used to compute statistical quantities of interest depending on the solution of partial differential equation with stochastic inputs. It consists of solving M deterministic problems, where M is the number of independent and identically distributed simulations (iid) of the parameters. The main disadvantage of this approach is its slow convergence given by the order $O(\frac{1}{\sqrt{M}})$, hence the method generates usually an exorbitant computational effort.

Recently, spectral methods have been developed such as the stochastic finite element and stochastic collocation methods. They offer a robust tool for

solving problems of stochastic PDE [12, 3, 2]. They approach the response of the model as a stochastic function by a polynomial interpolation in the stochastic space and provide an exponentially convergent approximation when the solution of the problem is a smooth function with respect to the stochastic variable. However, they require solving a large deterministic problem and this limiting their application sometimes to the problem with small stochastic dimension.

Those methods suffer from the curse of dimensionality, since the computational cost increases exponentially with a stochastic dimension and sometimes those methods are impracticable for very large dimension. Therefore, when the stochastic dimension is moderately large, it is interesting to minimize the cost and the effort of the resolution. This can be done by carefully selecting the stochastic degrees of freedom.

In this work, we discuss the anisotropic sparse grid method with adaptive approach inspired from the works [3, 16, 8]. In [8], we have introduced the Kahrunen-Loève expansion of the inverse of the diffusion parameter where we have used the random variables basis of this expansion to compute the average solution of one dimensional elliptic problem. Motivated by this work, we use this inverse of the diffusion to construct an error indicator to carefully choosing the parameter of the collocation method which are the level and the anisotropic weights. More precisely, we give an appropriate expansion of the indicator on the Lagrange polynomials basis, then we use this basis to construct the solution of the stochastic problem. We prove that this indicator satisfies an upper bound of the error. This process requires only a sequence of interpolation problems unlike the traditional approach which requires a large sequence of deterministic problems.

This paper is organized as follows. In the first section, we introduce the mathematical problem setting with a finite noise assumption. In section 2, we recall the stochastic collocation method based on the full tensor product and sparse grid interpolation. In Section 3, we introduce an indicator which will be used to determine the stochastic degrees of freedom of the collocation method. We prove that this indicator gives an upper bound of the error on the solution. Finally, in section 4, we present numerical examples to illustrate the efficiency of the proposed approach and compare with the traditional approach.

2   Problem formulation

Let $(\Omega, \mathcal{F}, d\mathbb{P})$ be a complete probability space, where Ω is the space of elementary event, $\mathcal{F} \subset 2^{\Omega}$ is the σ-algebra of events and $\mathbb{P}$ is the probability measure. Also, we consider a bounded domain $D \subset \mathbb{R}^d$, with a smooth boundary ∂D. The parameters k and f are two random fields on $\Omega \times D$. We focus on the following linear elliptic boundary value problem: find a stochastic

function, $u : \Omega \times D \longrightarrow \mathbb{R}$ such that the following equation holds (a.s) in Ω:

$$(\mathcal{P}_s) \begin{cases} -\operatorname{div}(k\nabla u) = f \ \text{ in } \Omega \times D, \\ u \mid_{\partial D} = 0. \end{cases} \tag{1}$$

We make the following assumptions on the random parameters k and f to keep the well-posedness of the problem $(\mathcal{P}_s)$.

Assumptions 1:

- **H1**: f belongs in $L^p(\Omega) \otimes L^2(D)$, for all $p \in [1, \infty[$.

- **H2**: There exist k_{min}, k_{max} such that, for each $\omega \in \Omega$ (a.s) we have

  $0 < k_{min}(\omega) \leq k(\omega, .) \leq k_{max}(\omega)$ and $k_{min}^{-1} \in L^p(\Omega)$.

Let $H_0^1(D)$ be the subspace of $H^1(D)$ consisting of the functions with vanishing trace on ∂D, then we define the space $L^p(\Omega) \otimes H_0^1(D)$,

$$L^p(\Omega) \otimes H_0^1(D) := \left\{ v : \Omega \longrightarrow H_0^1(D); \int_\Omega \|v\|_{H_0^1(D)}^p d\mathbb{P} < \infty \right\}.$$

To introduce the stochastic discretization, we first state some finite dimensional assumption on the parameters k and f (see [3, 16, 17]).

2.1 Finite dimensional noise

In many practical problems, random coefficients are parameterized by a finite number of uncorrelated random variables, sometimes independent, as in the case of a truncated Karhunen-Loève or Fourier expansion [14, 15, 2]. This motivates us to assume that k and f are parameterized by N random variables $\{Y_n\}_{n=1}^N$. The number N is called the stochastic dimension.

Assumptions 2 (Finite dimensional noise)
The stochastic fields k and f used in the computations are parameterized by a random vector $\mathbf{Y} = (Y_1, ..., Y_N)$ with N components such as $k(\omega, x) = k(\mathbf{Y}(\omega), x)$ and $f(\omega, x) = f(\mathbf{Y}(\omega), x)$.

We denote $\Gamma_n := Y_n(\Omega)$ the image of the set Ω by Y_n for each $n = 1, \ldots, N$ and Γ the image $\mathbf{Y}(\Omega)$. We assume that the components of $\mathbf{Y}$ are independent and ρ_n is the density function of each Y_n. Then $\mathbf{Y}$ has a joint probability density $\varrho(y) = \prod_{n=1}^N \rho_n(y_n), \forall y \in \Gamma$ and $\Gamma := \prod_{n=1}^N \Gamma_n$.

Exemple1 (Karhunen-Loève (K-L) expansion)

The Karhunen-Loòve expansion [14, 15] allows to expand each random field $k \in L^2(\Omega) \otimes L^2(D)$ in the series expansion given by the product of deterministic functions and a random variables:

$$k(\omega, x) = \mathbb{E}[k](x) + \sum_{n=1}^{\infty} \sqrt{\lambda_n} \phi_n(x) Y_n(\omega), \tag{2}$$

where $\{\lambda_n\}_{n=1}^{\infty}$ et $\{\phi_n\}_{n=1}^{\infty}$ are respectively eigenvalues and eigenfunctions solution of the following eigenvalue problem:

$$\int_D \text{cov}[k](x, z) \phi(z) dz = \lambda \phi(x).$$

Where $\text{cov}[k](x, z) := \mathbb{E}[k(., x)k(., z)] - \mathbb{E}[k(., x)]\mathbb{E}[k(., z)]$ and $\{Y_n\}_{n=1}^{\infty}$ are uncorrelated random variables given by $Y_n = \dfrac{1}{\sqrt{\lambda_n}} \int_D (k(., x) - \mathbb{E}[k(x)]) \phi_n(x) dx$. We define the truncated (K-L) expansion k_N by,

$$k_N(\omega, x) = \mathbb{E}[k](x) + \sum_{n=1}^{N} \sqrt{\lambda_n} \phi_n(x) Y_n(\omega). \tag{3}$$

Exemple2 (A nonlinear coefficient)

In some cases, to ensure some properties for data, as the positivity, the parameter k is given by a nonlinear transformation of a Gaussian field G like that $k = g(G)$, where $g > 0$ is non negative and smooth. A widely used example is a log-normal coefficient $k(\omega, x) = e^{G_N}$, where G_N is the (K-L) truncation of the Gaussian field G.

2.2 Strong formulation

The stochastic functions k and f are parameterized by the vector $\mathbf{Y}$ as in Assumption 2, then by Doob-Dynkin's lemma, the solution u of $(\mathcal{P}_s)$ can be also parameterized by $\mathbf{Y}$ such that $u(\omega, x) = u(\mathbf{Y}(\omega), x)$. Therefore, we introduce the strong deterministic problem obtained by projecting $(\mathcal{P}_s)$ in $H_0^1(D)$ for all y in Γ (a.e):

$$\begin{cases} \text{Find } u(y, .) \in H_0^1(D) \text{ such that} \\ \displaystyle\int_D k(y, .) \nabla u(y, x) . \nabla v(x) dx = \int_D f(y, x) v(x) dx \;\; \forall v \in H_0^1(D). \end{cases} \tag{4}$$

for all y in Γ (a.e), by Lax-Milgram theorem, the problem (4) is well posed and $u(y, .)$ satisfies:

$$\|u(y, .)\|_{H_0^1(D)} \leq C_D k_{min}^{-1}(y) \|f(y, .)\|_{L^2(D)}, \tag{5}$$

4

where C_D is the Poincaré constant. Under Assumption ($\mathbf{H1}, \mathbf{H2}$), the solution u belongs in $L_\varrho^p(\Gamma) \otimes H_0^1(D)$ for all $p \geq 1$, since:

$$\mathbb{E}[\|u\|_{H_0^1(D)}^p] \leq C_D^p \mathbb{E}[k_{min}^{-p}\|f\|_{L^2(D)}^p] \leq C_D^p \sqrt{\mathbb{E}[(k_{min}^{-1})^{2p}]\mathbb{E}[\|f\|_{L^2(D)}^{2p}]}.$$

3   Stochastic collocation method

The stochastic collocation method has the same principle as Monte Carlo method. It gives realizations of u by solving (4) at deterministic knots $\{y_{\mathbf{j}}\}_{\mathbf{j}}$ of Γ and construct an approximation $u_{h,p}$ of u in the tensor space $W_{p,h} := \mathcal{P}_p(\Gamma) \otimes V_h(D)$. Where $V_h(D)$ is the standard finite element space and $\mathcal{P}_p(\Gamma)$ is the span of tensor product polynomials with degree at most $p = (p_1, ..., p_N)$. The method is fully parallelized and provide exponential convergence when k and f are smooth with respect to y. We introduce the semi-discrete approximation u_h obtained by projecting (4) into $V_h(D)$ for (a.s) $y \in \Gamma$:

$$\begin{cases} \text{Find } u_h(y,.) \in V_h(D) \text{ such that} \\ \int_D k(y,.)\nabla u_h(y,x).\nabla v_h(x)dx = \int_D f(y,x)v_h(x)dx, \quad \forall v_h \in V_h(D). \end{cases} \quad (6)$$

We solve (6) at suitable knots $\{y_{\mathbf{j}}\}_{\mathbf{j}}$ of Γ, then we obtain the family $\{u_h(y_{\mathbf{j}},.)\}_{\mathbf{j}}$ to construct an approximation $u_{h,p}$ by multi-dimensional polynomial interpolation . It can be done either by a full tensor product rule of one-dimensional interpolation or on the sparse grid given by the Smolyak algorithm [18, 19].

3.1   Full tensor product interpolation

The full tensor product interpolation is given by a product of one-dimensional interpolation, where a high order is used in each dimension. For $n = 1, ..., N$, let $i_n \in \mathbb{N}^+$ a level of one-dimensional interpolation, and $\mathcal{X}^{i_n} := \{y_{n,1}, ..., y_{n,m(i_n)}\}$ is a set of $m(i_n)$ collocation points in Γ_n.

For $v \in \mathcal{C}^0(\Gamma_n; V_h(D))$, we introduce a sequence of one-dimensional Lagrange interpolation operator $\mathcal{U}^{i_n} : \mathcal{C}^0(\Gamma_n; V_h(D)) \longrightarrow \mathcal{P}_{p_n}(\Gamma_n) \otimes V_h(D)$ defined by:

$$\mathcal{U}^{i_n}(v)(z,x) = \sum_{j=1}^{m(i_n)} v(y_{n,j}, x)\ell_{n,j}^{i_n}(z), \quad (7)$$

where $\ell_j^{i_n}(z) = \prod_{r=1, r \neq j}^{m(i_n)} \frac{z - y_{n,r}}{y_{n,j} - y_{n,r}}$ is the Lagrange polynomial with degree $p_n = m(i_n) - 1$. We define the full tensor product interpolation on $\mathcal{C}^0(\Gamma; V_h(D))$ by:

$$\mathcal{A}_{\mathbf{i}}v(y,x) := \mathcal{U}^{i_1} \otimes \ldots \otimes \mathcal{U}^{i_N}v(y,x) = \sum_{\mathbf{j} \leq \mathbf{m(i)}} v(y_{\mathbf{j}}, x)L_{\mathbf{j}}^{\mathbf{i}}(y), \quad (8)$$

where $p = (p_1, \ldots, p_N)$, $\mathbf{m(i)} = (m(i_1), \ldots, m(i_N))$ the tensor space $\mathcal{P}_p(\Gamma) := \otimes_{n=1}^N \mathcal{P}_{p_n}(\Gamma_n)$ and we construct the knots in Γ by $y_{\mathbf{j}} = (y_{1,j_1}, \ldots, y_{n,j_n}, \ldots, y_{N,j_N})$ where each $y_{n,j_n} \in \mathcal{X}^{i_n}$. For each $\mathbf{i}$, $\mathbf{j} := (j_1, \ldots, j_N)$ and $y = (y_1, \ldots, y_n) \in \Gamma$, we set $L_{\mathbf{j}}^{\mathbf{i}}(y) = \otimes_{n=1}^N \ell_{n,j_n}^{i_n}(y_n)$ the multidimensional Lagrange polynomial.

3.2 Standard sparse grid method

The Standard sparse grid method is based on the Smolyak algorithm [18, 19] which provides an effective way to approach multivariate functions by a polynomial interpolation. It is given by linear combinations of product formulas (8) where in each dimension, a small order of interpolation is used, this provides a significant reduction of the full interpolation complexity and the curse of dimensionality is reduced.

For each $n = 1, \ldots, N$, we consider the one-dimensional operator $\mathcal{U}^{i_n}$ associated to the knots $\{y_{n,j}\}_{j=1}^{m(i_n)}$ as given in (7), and let $\Delta_{i_n} = \mathcal{U}^{i_n} - \mathcal{U}^{i_n-1}$ be an incremental operator with $\mathcal{U}^0 = 0$, for $w \in \mathbb{N}$, we consider the set of multi-level defined by $X(w, N) := \left\{ \mathbf{i} \in \mathbb{N}^N, \mathbf{i} \geq \mathbf{1}, \ |\mathbf{i}| := \sum_{n=1}^N i_n \leq w + N \right\}$. The Smolyak interpolation operator $\mathcal{A}(w, N)$ with level w is given by the formula:

$$\mathcal{A}(w, N)v = \sum_{\mathbf{i} \in X(w,N)} (\triangle_{i_1} \otimes \ldots \otimes \triangle_{i_N})v. \tag{9}$$

An equivalent formula (see [19]) which is more practical than (9) is given by:

$$\mathcal{A}(w, N)v = \sum_{w+1 \leq |\mathbf{i}| \leq w+N} (-1)^{w+N-|\mathbf{i}|} \binom{N-1}{w+N-|\mathbf{i}|} \mathcal{U}^{i_1} \otimes \ldots \otimes \mathcal{U}^{i_N} v. \tag{10}$$

To compute the interpolation $\mathcal{A}(w, N)v$ given in (10), we need to evaluate v at knots of the "sparse grid" $\mathcal{H}(w, N) = \bigcup_{w+1 \leq |\mathbf{i}| \leq w+N} \{\mathcal{X}^{i_1} \times \ldots \times \mathcal{X}^{i_N}\}$. In the case when the knots are nested, the Smolyak interpolant becomes more interesting, since, a significant reduction of the effort can be gained. Several nested points can be used according to their degree of accuracy and their growing number. For example: Newton-Cotes, Clenshaw-Curtis, Gauss-Patterson [7, 10, 16, 6].

3.3 The anisotropic sparse grid method

In the standard sparse grid method, all of the directions are equitably treated globally with the same order, since any permutation of an admissible multi-levels in $X(w, N)$ is also admissible. However, when the dimensions have different behaviors, the solution is highly anisotropic and the convergence rate can be deteriorated when we use this isotropic method. Thus, an anisotropy interpolation is needed to improve the convergence.

We follow closely [16], to define the anisotropic sparse grid formula. We consider a suitable weights $\alpha_1, \ldots, \alpha_N$, for each dimension, $\underline{\alpha} = \min_{1 \leq, n \leq N} \alpha_n$ and the generalized sets of multi-levels: $X_\alpha(w, N) := \left\{ \mathbf{i} \in \mathbb{N}^N, \mathbf{i} \geq \mathbf{1}, \ \sum_{n=1}^N (i_n - 1)\alpha_n \leq w\underline{\alpha} \right\}$ and

6

$$Y_\alpha(w,N) = \left\{ \mathbf{i} \geq \mathbf{1} : w\underline{\alpha} < \sum_{n=1}^N \alpha_n i_n \leq w\underline{\alpha} + |\alpha| \right\} \text{ with } |\alpha| = \sum_{n=1}^N \alpha_n. \text{ Then the}$$

anisotropic sparse grid formula is given by:

$$\mathcal{A}_\alpha(w,N)v = \sum_{\mathbf{i}\in X_\alpha(w,N)} (\triangle_{i_1} \otimes \ldots \otimes \triangle_{i_N})v, \tag{11}$$

An equivalent formula of (11) is given by the product rule [11]:

$$\mathcal{A}_\alpha(w,N)v = \sum_{\mathbf{i}\in Y_\alpha(w,N)} c_\alpha(\mathbf{i})\,\mathcal{U}^{i_1} \otimes \ldots \otimes \mathcal{U}^{i_N}v, \tag{12}$$

where the combining coefficients $c_\alpha(\mathbf{i}) = \sum_{\mathbf{j}\in\{0,1\}^N, \mathbf{i}+\mathbf{j}\in X_\alpha(w,N)} (-1)^{|\mathbf{j}|}.$

The computation of $\mathcal{A}_\alpha(w,N)u$ requires to evaluate v on the sparse grid:

$$\mathcal{H}_\alpha(w,N) = \bigcup_{\mathbf{i}\in Y_\alpha(w,N)} \{\mathcal{X}^{i_1} \times \ldots \times \mathcal{X}^{i_N}\}. \tag{13}$$

The anisotropic weights $\{\alpha_n\}_{n=1}^N$ allows to approximate adaptively u. In the next section we discuss a posteriori strategy to compute them with a reduced cost.

4   Error estimation and adaptivity in the stochastic space

The computational cost increases exponentially according to the stochastic dimension, sometimes, it can be dramatic. Thus, it is interesting to seek an appropriate way to reduce the global computation cost by carefully selecting suitable stochastic degrees of freedom and without loss of accuracy by an efficient strategy of adaptation.

4.1   Motivation

In the traditional approaches, when we attempt to approximate u by the full tensor product, adaptivity in the stochastic space consists of to start with a free degree of interpolation and increase the degree until reaching the desired accuracy, as described in [2, 3]. This approach seems very expensive, since in each iteration, we solve a set of deterministic PDE problems, which increases with N. The same strategy is used to determine the weights $\{\alpha\}_{n=1}^N$ and the level w of the formula $\mathcal{A}_\alpha(w,N)u$.

In what follows, we introduce an indicator λ to determine with a reduce cost, the suitable stochastic degree of freedom to compute u. The problem $(\mathcal{P}_s)$ is equivalent to the following system:

$$\begin{cases} -\nabla u = k^{-1}p \\ div(p) = f. \end{cases}$$

The first equation shows that each stochastic dimension of u and $k^{-1}p$ have the same behavior. From the second equation we expect that p and f have also the same

behavior with respect to the stochastic variable. Therefore, we propose to combine those parameters to obtain an indicator $\lambda = k^{-1}f$, which allows to identify with a small effort the behavior of different directions of u.

In [8], for one-dimensional elliptic problem, we have computed the Karhunen-Loève expansion for k^{-1}, and use its random variables as a basis to build a projection of u which gives a good approximation of the expectation $\mathbb{E}[u]$. In this work, we propose to use a similar methodology to approximate u in $L^2(\Gamma)$. We expand the indicator λ on the multi-dimensional Lagrange polynomials basis defined in (8) or (10, 12), then we use this basis to approximate the solution u.

The following lemma concerns an estimate of the error $\|u - \mathcal{A}_\mathbf{i}u\|_{L^2(\Gamma;H_0^1(D))}$ on u when we approximate it by the full tensor interpolation (8). This is the main result given in [3] showing that each direction contributes to the global error with a decay rate depending on the size of the region where u is analytic.

Lemma 4.1 *([3]) Let u be the solution of $(\mathcal{P}_s)$, $\mathcal{A}_\mathbf{i}u$ its full interpolation defined in (8). We suppose that, in each direction, u admits an analytic extension in the region $\Sigma_n(\Gamma_n; \tau_n) := \{z \in \mathbb{C}, \quad dist(z, \Gamma_n) \leq \tau_n\}$, with the size $\tau_n > 0$. Then, there exists $C_F > 0$ is independent of each $p_n = m(i_n) - 1$, such that:*

$$\|u - \mathcal{A}_\mathbf{i}u\|_{L_\varrho^2(\Gamma)\otimes H_0^1(D)} \leq C_F \sum_{n=1}^N \beta_n(p_n)e^{-\nu_n p_n^{\theta_n}} \tag{14}$$

- *if Γ_n is bounded , $\theta_n = \beta_n = 1$ and $\nu_n = log\left(\dfrac{2\tau_n}{|\Gamma_n|}(1 + \sqrt{1 + \dfrac{|\Gamma_n|^2}{4\tau_n^2}})\right)$.*

- *if Γ_n is unbounded , $\theta_n = 1/2, \beta_n = O(\sqrt{p_n})$ and $\nu_n = \tau_n/2$.*

The next proposition shows that the error $\|\lambda - \mathcal{A}_\mathbf{i}\lambda\|_{L_\varrho^2(\Gamma)\otimes L^2(D)}$ is an upper bound of the error $\|u - \mathcal{A}_\mathbf{i}u\|_{L_\varrho^2(\Gamma)\otimes H_0^1(D)}$.

Proposition 4.1 *Let u be the solution of $(\mathcal{P}_s)$, $\lambda = k^{-1}f$ and $\mathcal{A}_\mathbf{i}u$ and $\mathcal{A}_\mathbf{i}\lambda$ their approximations by the formula (8). Then we have the following estimate:*

$$\|u - \mathcal{A}_\mathbf{i}u\|_{L_\varrho^2(\Gamma;H_0^1(D))} \leq C\|\lambda - \mathcal{A}_\mathbf{i}\lambda\|_{L_\varrho^2(\Gamma,L^2(D))}.$$

Where the positive constant C depends on u and λ.

The estimate given in Lemma 4.1 shows that the interpolation $\mathcal{A}_\mathbf{i}$ converges roughly like $O(e^{-\nu p})$ where $\nu = \min_n \nu_n$ and p the order used in its associated dimension. In the other word, there is $\tilde{p} > 0$ such that:

$C_1 e^{-\nu p} \leq \|u - \mathcal{A}_\mathbf{i}u\|_{L_\varrho^2(\Gamma;H_0^1(D))} \leq C_2 e^{-\nu p}$ for each $p > \tilde{p}$.

Thus, to compare $\|u - \mathcal{A}_\mathbf{i}u\|_{L_\varrho^2(\Gamma;H_0^1(D))}$ with $\|\lambda - \mathcal{A}_\mathbf{i}\lambda\|_{L_\varrho^2(\Gamma;H_0^1(D))}$ it is enough to compare τ_n of u with that of λ for each $n = 1, \ldots, N$. To do this, we adopt the following notations: for each $n = 1, \ldots, N$, we put $\Gamma_n^* = \prod_{j\neq n}^N \Gamma_j$, $y = (y_n, y_n^*)$ where $y_n \in \Gamma_n$ and $y_n^* \in \Gamma_n^*$. We also suppose that f is polynomial with respect to y.

Proof of the Proposition 4.1

1) First, we consider the case where k is linear according to y: $k = b_0 + \sum_{n=1}^{N} b_n y_n$. Since f is polynomial on y, it is enough to compare the analyticity region of u with that of k^{-1}. The function $u(., y_n^*, x)$ as a function of the variable y_n, can be extended analytically in the ball $\{|z - y_n| < \tau_n(u)\}$ and we have by d'Alembert's rule:

$$\frac{1}{\tau_n(u)} = \limsup_{m \longrightarrow \infty} \frac{\frac{\|\sqrt{k}\partial_{y_n}^m \nabla u\|_{L^2(D)}}{m!}}{\frac{\|\sqrt{k}\partial_{y_n}^{m-1}\nabla u\|_{L^2(D)}}{(m-1)!}}.$$

Since, we can write: $k^{-1}(z, y_n^*, x) = \dfrac{k^{-1}(y_n, y_n^*, x)}{1 + \dfrac{(z - y_n)b_n(x)}{k(y_n, y_n^*, x)}}$, then $k^{-1}(., y_n^*, x)$ is analytically extended in the region $\{|z - y_n| < \tau_n(k^{-1})\}$, such that:

$$\frac{1}{\tau_n(k^{-1})} = \|\frac{b_n}{k}\|_{\infty}.$$

On the other hand, by deriving m-time the problem (4) with respect to y_n we obtain:

$$\partial_{y_n}^m \int_D k\nabla u \nabla v dx = \int_D \sum_{l=0}^{m} \binom{m}{l} \partial_{y_n}^l k \partial_{y_n}^{m-l} \nabla u \nabla v dx = \partial_{y_n}^m \int_D fv dx \quad \forall v \in H_0^1(D),$$

for $l \geq 2$ we have $\partial_{y_n}^l k = 0$ and $\partial_{y_n} k = b_n$, then:

$$\int_D k\frac{\partial_{y_n}^m \nabla u}{m!}\nabla v dx = -\int_D b_n \frac{\partial_{y_n}^{m-1}\nabla u}{(m-1)!}\nabla v dx + \int_D \frac{\partial_{y_n}^m f}{m!}v dx. \tag{15}$$

Taking $v = \partial_{y_n}^m \nabla u$ in (15) and by Cauchy-Schwartz inequality we obtain:

$$\frac{\|\sqrt{k}\partial_{y_n}^m \nabla u\|_{L^2(D)}}{m!} \leq \|\frac{b_n}{k}\|_{\infty}\frac{\|\sqrt{k}\partial_{y_n}^{m-1}\nabla u\|_{L^2(D)}}{(m-1)!} + \frac{C_D}{\sqrt{k_{min}(y)}}\frac{\|\partial_{y_n}^m f\|_{L^2(D)}}{m!}.$$

Since f is polynomial on y, we have $\limsup_{m \longrightarrow \infty} \frac{\|\partial_{y_n}^m f\|_{L^2(D)}}{m!} = 0$, thus

$$\limsup_{m \longrightarrow \infty} \frac{\frac{\|\sqrt{k}\partial_{y_n}^m \nabla u\|_{L^2(D)}}{m!}}{\frac{\|\sqrt{k}\partial_{y_n}^{m-1}\nabla u\|_{L^2(D)}}{(m-1)!}} \leq \|\frac{b_n}{k}\|_{\infty}.$$

Then we deduce that $\tau_n(u) \geq \tau_n(k^{-1})$. Therefore, there exists $C > 0$ such that:

$$||u - \mathcal{A}_{\mathbf{i}}u||_{L_\varrho^2(\Gamma; H_0^1(D))} \leq C||\lambda - \mathcal{A}_{\mathbf{i}}\lambda||_{L_\varrho^2(\Gamma; L^2(D))}.$$

2) We suppose now that k is non linear according to y as $k(y, x) = g(b_0(x) + \sum_{n=1}^{N} b_n(x)y_n)$, where g is an analytic and strictly positive function. Observe that in this case, k satisfies the assumption ensuring the analyticity of u (see [3]). Let

9

$\{\mathcal{V}_l\}_l$ a discretization of the set Γ, where for each element $\mathcal{V}_l$ with the size h_l, the following estimate holds:

$$k(y, x) \approx \sum_{n=1}^{N} a_n(y^l, x) y_n \ \ \forall y \in \mathcal{V}_l.$$

Where y^l is the center of $\mathcal{V}_l$. Let $\mathcal{A}_{\mathbf{i}}^l u$ be the full interpolation for the restriction of u in each element $\mathcal{V}_l$, then by the previous case the following estimate holds:

$$||u - \mathcal{A}_{\mathbf{i}}^l u||_{L_\varrho^2(\mathcal{V}_l; H_0^1(D))} \leq \tilde{c} ||\lambda - \mathcal{A}_{\mathbf{i}}^l \lambda||_{L_\varrho^2(\mathcal{V}_l; L^2(D))}. \tag{16}$$

Furthermore, for all y in $\mathcal{V}_l$, if we set $\psi_l(y, \cdot) = u(y, \cdot) - \mathcal{A}_{\mathbf{i}} u(y, \cdot)$, we have:
$\mathcal{A}_{\mathbf{i}}^l \psi_l = \mathcal{A}_{\mathbf{i}}^l u - \mathcal{A}_{\mathbf{i}}^l(\mathcal{A}_{\mathbf{i}} u) = \mathcal{A}_{\mathbf{i}}^l u - \mathcal{A}_{\mathbf{i}} u$, since $\mathcal{A}_{\mathbf{i}} u$ is polynomial. According to the classical h_p convergence (see [2, 5]), with a small size h_l, the following estimation holds,

$$\|\psi_l - \mathcal{A}_{\mathbf{i}}^l \psi_l\|_{L_\varrho^2(\mathcal{V}_l; H_0^1(D))} \approx \hat{c} h_l^p \|\psi_l\|_{L_\varrho^2(\mathcal{V}_l; H_0^1(D))} \tag{17}$$

where p is the smallest order used in the directions of $\mathcal{A}_{\mathbf{i}}^l$. Therefore, we deduce

$$||u - \mathcal{A}_{\mathbf{i}}^l u||_{L_\varrho^2(\mathcal{V}_l; H_0^1(D))} \approx \hat{c} h_l^p ||u - \mathcal{A}_{\mathbf{i}} u||_{L_\varrho^2(\mathcal{V}_l; H_0^1(D))}. \tag{18}$$

Also by the same argument, we have:

$$||\lambda - \mathcal{A}_{\mathbf{i}}^l \lambda||_{L_\varrho^2(\mathcal{V}_l; H_0^1(D))} \approx \hat{c} h_l^p ||\lambda - \mathcal{A}_{\mathbf{i}} \lambda||_{L_\varrho^2(\mathcal{V}_l; L^2(D))}. \tag{19}$$

Combining (18) and (19) with (16) we obtain in each $\mathcal{V}_l$:

$$||u - \mathcal{A}_{\mathbf{i}} u||_{L_\varrho^2(\mathcal{V}_l; H_0^1(D))} \leq \tilde{C} ||\lambda - \mathcal{A}_{\mathbf{i}} \lambda||_{L_\varrho^2(\mathcal{V}_l; L^2(D))}.$$

Then the estimation hold on Γ. ∎

Remark 4.1 *The result of the proposition is still true with Smolyak interpolation $\mathcal{A}_\alpha(N, w)$, since the accuracy of the formula $\mathcal{A}_\alpha(N, w)$ depends on the size of the analyticity region of the function. Moreover, when k is linear with respect to y, the constant C can decrease with total number of collocation points used by $\mathcal{A}_\alpha(w, N)$.*

Remark 4.2 *(Choice of the reference points) When we attempt to determine a suitable polynomial degree p_n according to direction n of $\mathcal{A}_{\mathbf{i}} \lambda$, we set the other directions fixed at reference point y_n^* and we seek the degree p_n such that two successive approximations are close. The degree $m(i_n)$ may depend on y_n^*, so it must be carefully chosen. For a random field discussed in Section 2.1, the reference point $y_n^* = 0$ is adequate, since there is no interaction between directions.*

In the next section, as an application of Proposition 4.1, we use the indicator $\lambda = k^{-1} f$ to determine an appropriate level w and a weight vector α to compute $\mathcal{A}_\alpha(w, N) u$.

4.2 Adaptive sparse grid

The main ingredients of the collocation method on the anisotropic sparse grid is the weight vector α and the level w. The vector α takes into account the importance between the dimensions. The level w is the total order of the interpolation, it controls the cardinality of the knots of the grid. Then, when we determine suitable weights and a convenient level w with a small cost, the method becomes more efficient.

4.2.1 Determination of the anisotropic weights

The main idea proposed in [17] consists of choosing each anisotropic weight α_n as the exponential rate of the convergence according the direction n i.e $\alpha_n = \nu_n$, where ν_n is the rate given in Lemma 4.1. The global error committed by the interpolation $\mathcal{A}_{\mathbf{i}}u$ can be split into N parts, $\|u - \mathcal{A}_{\mathbf{i}}u\|_{L^2_\varrho(\Gamma, H^1_0(D))} \leq \sum_{n=1}^N \varepsilon_n$.

Each part ε_n is the error contribution in the direction n which is estimated by:

$$\varepsilon_n \approx c_n e^{-\nu_n p_n}, \qquad \text{when } \Gamma_n \text{ is bounded ,} \qquad (20)$$

$$\varepsilon_n \approx O(\sqrt{p_n}) e^{-\nu_n \sqrt{p_n}}, \qquad \text{when } \Gamma_n \text{ is unbounded ,} \qquad (21)$$

When we attempt to solve highly anisotropic problem, we obtain numerically (Section 5) a non optimal vector weight with (21), since the resulting sparse grid is quite dense. The anisotropic property of the weight can is reduced because the square $\sqrt{\cdot}$ in (21). In order to remedy to this situation when Γ is unbounded, we propose to use an estimate error as in (20). We suppose that the density ϱ decreases like a gaussian kernel, and each density ρ_n of Y_n satisfies the following estimate:

$$\rho_n(z) \leq C_n e^{-\delta z^2}, \quad C_n > 0 \text{ et } \delta > 0, \quad \forall z \in \mathbb{R}. \qquad (22)$$

We denote by $H_p(t) \in \mathcal{P}_p(\mathbb{R})$ the normalized Hermite polynomials,

$$H_p(t) = \frac{1}{\sqrt{p! 2^p \pi^{\frac{1}{2}}}} (-1)^p e^{-t^2} \frac{d^p}{dt^p}\left(e^{-t^2}\right).$$

We recall that they form a complete orthonormal basis of the space $L^2_\mu(\mathbb{R})$, where $\mu(t) = e^{-t^2}$. We note also $h_p(t) = \sqrt{\mu(t)} H_p(t)$, the associated Hermite functions. The following proposition gives an estimate of the error committed by $\mathcal{U}^m v$ for an entire function v satisfying (22). The proof is based on an estimate of the Fourier-Hermite coefficients (see[4]) for an entire function decreasing as a gaussian kernel.

Proposition 4.2 *We consider κ in $L^2_\rho(\mathbb{R}; L^2(D))$, $\mathcal{U}^i(\kappa)$ its Lagrange interpolation defined by $m(i)$ Gauss points. We suppose that ρ satisfies (22), for each $t \in \mathbb{R}$, $\kappa(t,.)$ have an entire extension in the complex plan, and there is $\nu > 0$ such that*

$$\|e^{-\nu|z|}\kappa(z)\|_{L^2(D)} \leq C_\nu < \infty, \quad \forall z \in \mathbb{C}. \qquad (23)$$

Then, there is a decay $g > 0$ and a constant C does not depend on $m(i)$ such that:

$$\|\kappa(t,.) - \mathcal{U}^i(\kappa)(t,.)\|_{L^2_\rho(\mathbb{R}; L^2(D))} \leq C e^{-gm(i)}. \qquad (24)$$

PROOF: We consider the function $\tilde{\kappa}(t,x) = \kappa(w(t), x)$ where $w(t) = \dfrac{\sqrt{2}t}{\delta}$ and the expansion of $\tilde{\kappa}$ in Hermite polynomials basis:

$$\tilde{\kappa}(t,x) = \sum_{p=0}^{\infty} \kappa_p(x) H_p(t) \quad \text{where} \quad \kappa_p(x) = \int_{\mathbb{R}} \tilde{\kappa}(t,x) H_p(t) e^{-t^2} dt. \qquad (25)$$

We set $F(t,x) := \tilde{\kappa}(t,x) e^{-\frac{t^2}{2}}$, its Fourier-Hermite expansion is given by:

$$F(t,x) = \sum_{p=0}^{\infty} F_p(x) h_p(t) \quad \text{where} \quad F_p(x) = \int_{\mathbb{R}} F(t,x) h_p(t) dt, \qquad (26)$$

The Hermite coefficients F_p in (26) are the same as in (25). Indeed,

$$F_p(x) = \int_{\mathbb{R}} F(t,x) h_p(t) dt = \int_{\mathbb{R}} \tilde{\kappa}(t,x) e^{-t^2} H_p(t) dt = \kappa_p.$$

The function κ satisfies (23), then F decreases as a gaussian kernel, since

$$\|\tilde{\kappa}(t,x)\| e^{-t^2/2} \le C_\nu e^{\sqrt{2}\frac{\nu}{\delta}t} e^{-t^2/2} \le C_\nu e^{-\frac{1}{2}(t-\sqrt{2}\frac{\nu}{\delta})^2} e^{\frac{\nu^2}{\delta^2}} \le d e^{-qt^2},$$

where d and q are positive, finite, and depending on ν and δ. Clearly F is an entire function and decrease as a gaussian kernel, therefore (see [4]) the Fourier-Hermite coefficients F_p decreases geometrically as:

$$F_p \le K e^{-gp} \quad \forall p \in \mathbb{N},$$

where the constant $K > 0$ does not depend on p and the decay $g > 0$. The following error can be bounded (see [3], Lemma 7) as:

$$\|\kappa(t,.) - \mathcal{U}^{m(i)}(\kappa)(t,.)\|_{L^2_\rho(\mathbb{R}; L^2(D))} \le \min_{v \in \mathcal{P}_{m(i)} \otimes L^2(D)} \max_{t \in \mathbb{R}} \left| \|\kappa - v\|_{L^2(D)} e^{-\frac{(t\delta)^2}{4}} \right|. \qquad (27)$$

Taking $v = \Pi_{m(i)}\kappa$ in (27) the truncated Hermite expansion of κ up the order $m(i)-1$, and let $\tilde{v} = \Pi_{m(i)}\tilde{\kappa}$ that of $\tilde{\kappa}$. We deduce:

$$
\begin{aligned}
\|\kappa(t,.) - \mathcal{U}^{m(i)}(\kappa)(t,.)\|_{L^2_\rho(\mathbb{R}; L^2(D))} &\le \max_{z \in \mathbb{R}} \left| \|\kappa - \Pi_{m(i)}(\kappa)\|_{L^2(D)} e^{-\frac{(z\delta)^2}{4}} \right| \\
&\le \max_{t \in \mathbb{R}} \left| \|\tilde{\kappa} - \Pi_{m(i)}(\tilde{\kappa})\|_{L^2(D)} e^{-\frac{t^2}{2}} \right| \\
&\le \sum_{p=m(i)}^{\infty} \|\kappa_p\|_{L^2(D)} \max_{t \in \mathbb{R}} \left| H_p(t) e^{-\frac{t^2}{2}} \right| \\
&\le C_1 \sum_{p=m(i)}^{\infty} e^{-gp} \max_{t \in \mathbb{R}} |h_p(t)|.
\end{aligned}
$$

Knowing that $\max_{t \in \mathbb{R}} |h_p(t)| \le 1$ ((see e.g. [4]), we get,
$\|\kappa(t,.) - \mathcal{U}^{m(i)}(\kappa)(t,.)\|_{L^2_\rho(\mathbb{R}; L^2(D))} \le C_1 \sum_{p=m(i)}^{\infty} e^{-gp} \le C_2 e^{-gm(i)}.$ ∎

Remark 4.3 *When $k = e^G$ follows a log-normal law, the indicator λ is an entire function. Thus, in the case of unbounded variables, we use the estimate (24) to compute numerically the rate convergence in each direction of λ. Numerical tests have shown that using estimate (24) yields a fast convergence of the formula $\mathcal{A}_\alpha(w, N)$.*

In order to determine the anisotropic weight, we use the bound given by the Proposition 4.2 for both bounded and unbounded set Γ, we use the algorithm given in [16] for the indicator λ instead of the function u. The bound given in (20) and (24) is only true foe the Gaussian knots, then to obtain suitable weights , we use those knots with the growth rule of the order $m(i_n) = i_n$ for each dimension of λ.

Algorithm to determine α_n

- We fix $y_n^* = 0$, as a reference point, i.e $y_j = 0$ for all $j \neq n$.

- Let $m(\tilde{i}_n) = \tilde{i}_n$ be an integer such that $\tilde{\lambda}(y_n, 0) := \mathcal{U}^{\tilde{i}_n} \lambda(y_n, 0)$ is considered as a reference approximation according to the direction n.

- For each $p = 1, \ldots, \tilde{i}_n - 1$, we compute the error: $\mathrm{err}_n(p) = \|\tilde{\lambda}(y_n, 0) - \mathcal{U}^p \lambda\|$.

- We use a linear least squares algorithm to estimate α_n and c_n solutions of the following minimization problem:

$$\min_{\alpha_n, c_n} \| \log(err_n) - log(c_n) + P\alpha_n \|^2,$$

    where err_n is the vector with components of the error on λ and P is the vector with components the number of collocation points.

We observe that with this algorithm the computation cost is clearly reduced. Indeed to determine each α_n with the indicator λ, we compute $P_n = \frac{\tilde{m}_n(\tilde{m}_n+1)}{2}$ interpolation problems, while when we use u, we must solve P_n of a large linear systems. The algorithm which compute the weights of u is given in [17], by the following steps:

- We fix $y_n^* = 0$, as a reference point, i.e $y_j = 0$ for all $j \neq n$.

- Let $m(\tilde{i}_n) = \tilde{i}_n$ be an integer such that $\tilde{u}_h(y_n, 0) := \mathcal{U}^{\tilde{i}_n} u_h(y_n, 0)$ is considered as a reference approximation of u_h according to the direction n, where u_h is the finite element approximation of u.

- For each $p = 1, \ldots, \tilde{i}_n - 1$, we compute the error: $\mathrm{err}_n(p) = \|\tilde{u}_h(y_n, 0) - \mathcal{U}^p u_h\|$.

- We use a linear least squares algorithm to estimate α_n and c_n solutions of the following minimization problem:

$$\min_{\alpha_n, c_n} \| \log(err_n) - log(c_n) + P\alpha_n \|^2,$$

    where err_n is the vector with components of the error on u_h and P is the vector with components the number of collocation points.

Remark 4.4 *In the computation of the vector α, we use the growth rule $m(i_n) = i_n$ to compute the interpolation $\mathcal{U}^{i_n}$ since we need many results on the error to obtain a good estimate of α using a linear squares method.*

In practice, to accelerate the convergence of the Formula $\mathcal{A}_\alpha(w, N)$, we use an exponential growth rule $m(i_n) = 2^{i_n-1}$ or at least, the linear growth rule $m(i_n) = 2i_n - 1$. Thus, most of the computation realized to determine the vector α for u can not be reused to compute $\mathcal{A}_\alpha(w, N)u$, since only the knots of the grid $\mathcal{H}_\alpha(w, N)$ are involved in the construction of $\mathcal{A}_\alpha(w, N)u$. In particular, if the grid $\mathcal{H}_\alpha(w, N)$ is given by Gaussian knots, which are non nested points, almost the computation realized to determine the vector α for u is not reuse to compute $\mathcal{A}_\alpha(w, N)u$.

4.2.2 Determination of the level w for Smolyak formula

In the classical approach, when we attempt to compute $\mathcal{A}_\alpha(w, N)u$, the appropriate level w satisfies $\|\mathcal{A}_\alpha(w, N)u - \mathcal{A}_\alpha(w + 1, N)u\| \approx \varepsilon$, where ε is a desirable tolerance. The cost requires by $\mathcal{A}_\alpha(w+1, N)u$ is generally greater than the combined cost of all previous levels, it is given by a large number of deterministic problems.

Non-linear coefficient

Motivated by Proposition 4.1, we use the indicator λ to determine the convenient level w for the formula $\mathcal{A}_\alpha(w, N)$, since the error $\|\lambda - \mathcal{A}_\alpha(w, N)\lambda\|$ is an upper bound of the error $\|u - \mathcal{A}_\alpha(w, N)u\|$. The strategy consists of choosing the level $w \in \mathbb{N}$ such that for a given desirable tolerance $\varepsilon > 0$, the interpolation $\mathcal{A}_\alpha(w, N)\lambda$ satisfies:

$$\|\lambda - \mathcal{A}_\alpha(w, N)\lambda\| \leq \frac{\varepsilon}{C}, \tag{28}$$

where the constant C is given as in Proposition 4.1. It is shown in [17] that the error committed by $\mathcal{A}_\alpha(w, N)u$ satisfies the following bound:

$$\|u - \mathcal{A}_\alpha(w, N)u\| \approx C_1\|u\|M^{-\nu_1}, \tag{29}$$

where the decay rate ν_1 depends on the analyticity of u. We have a similar bound for the error committed by $\mathcal{A}_\alpha(w, N)\lambda$.

$$\|\lambda - \mathcal{A}_\alpha(w, N)\lambda\| \approx C_1\|\lambda\|M^{-\nu_2}. \tag{30}$$

Therefore, if we assume that $\nu_1 \approx \nu_2$ in the case where k is non linear with respect to y, we can estimate the constant C as follows:

$$C \approx \frac{\|u\|}{\|\lambda\|} \approx \frac{\|\mathcal{A}_\alpha(0, N)u\|}{\|\mathcal{A}_\alpha(0, N)\lambda\|}. \tag{31}$$

The indicator λ allows us to determine a suitable w to compute $\mathcal{A}_\alpha(w, N)u$. By this strategy we avoid great computations, since $\mathcal{A}_\alpha(w, N)\lambda$ requires only a sequence of the interpolation problems and we avoid computing $\mathcal{A}_\alpha(w + 1, N)u$ given by a sequence of PDE problems.

Linear coefficient

When the parameter k is linear with respect to y, the constant C given in Proposition 4.1 decrease with M_w, the number of the collocation points with level w, because the size of the analyticity region of u is large than that of λ. The decay rates ν_1 and ν_2 as in (29,30) satisfies $\nu_1 \geq \nu_2$ and depend on α_{min} the smallest anisotropic weight vector α [17]. In this case we propose to estimate the constant C for each w as follows:

$$C = C_w \approx \frac{\|\mathcal{A}_\alpha(0,N)u\|}{\|\mathcal{A}_\alpha(0,N)\lambda\|} M_w^{-\mu}.$$

The decay rate $\mu > 0$ is computed numerically by the following way, we estimate the weight vector α for λ and consider $\tilde{n}$ such that $\alpha_{\tilde{n}} = \min_{1 \leq n \leq N} \alpha$, we compute only the weight $\tilde{\alpha}_{\tilde{n}}$ for u according to direction $n = \tilde{n}$ and then we set $\mu = \alpha_{\tilde{n}} - \tilde{\alpha}_{\tilde{n}}$.

Using the indicator λ to determine a suitable level w of the formula $\mathcal{A}_\alpha(w,N)$, a significant reduction of the computational effort is kept. By this approach, the computation of the weight of anisotropy and the interpolation $\mathcal{A}_\alpha(w+1,N)$ required only an evaluation of λ on the collocation knots and the sparse grid $\mathcal{H}_\alpha(w+1,N)$, instead of solving a set of deterministic PDE problems.

5 Numerical Examples

This Section illustrates the convergence and effectiveness of the proposed adaptive approach to the stochastic linear elliptic problem, as described in Section 4. We consider two cases of the parameter k, non-linear and linear with respect to y. The computational results presented here are realized by Matlab software.

$$\begin{cases} -\mathrm{div}(k\nabla u) = f & \text{in } \Omega \times D, \\ u\,|_{\partial D} = 0, \end{cases} \tag{32}$$

with $D :=]0;1[^2$ and f is deterministic $f(\omega, x) = \cos(x_1)\sin(x_2), \ \ x = (x_1, x_2) \in D$.

5.1 Non-linear case

In the first example, we consider k_N which is non linear with respect to the stochastic variable and with one-dimensional spatial dependence:

$$k_N(\omega, x) = 0.5 + \exp\left(Y_1(\omega)\left(\frac{\sqrt{\pi l_c}}{2}\right)^{1/2} + \sum_{n=2}^{N} \zeta_n \varphi_n(x) Y_n(\omega)\right), \tag{33}$$

the reals $(\zeta_n)_{n=1}^{N}$ are strictly positive and given by:

$$\zeta_n := \left(\sqrt{\pi}l_c\right)^{1/2} \exp\left(\frac{-(\lfloor \frac{n}{2} \rfloor \pi l_c)^2}{8}\right), \ \ \text{if } n > 1$$

the functions $\{\varphi_n\}_{n=1}^{N}$ are defined by:

$$\varphi_n(x) := \begin{cases} \sin\left(\frac{\lfloor \frac{n}{2} \rfloor \pi x}{2}\right), & \text{if } n \text{ even,} \\ \cos\left(\frac{\lfloor \frac{n}{2} \rfloor \pi x}{2}\right), & \text{if } n \text{ odd.} \end{cases}$$

15

Where l_c is the correlation length and $\{Y_n\}_{n=1}^N$ are independent gaussian random variables with mean zero and unit variance. The parameter k_N is a truncation of a log-normal field $k = e^G$, where G is a Gaussian field with a Gaussian covariance function [16, 17]. The order N is strongly related to the size l_c, such that a large value of l_c corresponds to a small value of N, then the problem is strongly anisotropic and vice-versa [2, 16, 17].

To check that the proposed approach is effective for the anisotropic Smolyak method, we consider the case where $l_c = \frac{1}{2}$ with $N = 11$ and $l_c = \frac{1}{5}$ with $N = 41$. The finite element space of the spatial discretization is the span of piecewise linear functions over a uniform triangulation of D with $nd = 2712$ unknowns. The collocation points are the roots of the Hermite polynomials and $\Gamma = \mathbb{R}^d$.

We will approximate the solution u by the interpolation $\mathcal{A}_\alpha(w, N)u$, then compute $\mathbb{E}[\mathcal{A}_\alpha(w, N)u]$ and $\text{var}[\mathcal{A}(w, N)u]$ as an approximation of the mean value and the variance of the function u respectively.

First, we determine the suitable weight vector α for the indicator $\lambda = k^{-1}f$ with a posteriori approach then we seek a convenient level w such that the interpolation $\mathcal{A}_\alpha(w, N)\lambda$ satisfies a given threshold for $\mathbb{E}[\lambda]$. Therefore, we use this vector α and the level w to compute $\mathbb{E}[\mathcal{A}_\alpha(w, N)u]$ and $\text{var}[\mathcal{A}(w, N)u]$.

To determine the vector α in this case of unbounded set Γ, we use the bound (24) given in Proposition 4.2 instead of the bound (21) proposed in [3, 16].

We denote α and $\tilde{\alpha}$ the weight vectors obtained by a posteriori information with the bound (24), respectively for λ and u. We also denote by $\hat{\alpha}$ the weight vector for λ obtained using the bound (21).

Table 1 shows the first ten values of vectors α, $\tilde{\alpha}$ and $\hat{\alpha}$. We remark that the components of the vector α for the indicator λ are closely comparable with those of the vector $\tilde{\alpha}$ for u. However, α is completely different with $\hat{\alpha}$, since they are obtained with different one-dimensional bound.

l_c	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}	α_{11}
1/2	2.57	2.26	2.27	4.31	4.31	8.18	8.18	12.50	12.51	20.58	20.59
1/5	3.13	2.55	2.55	2.88	2.88	3.4	3.4	4.34	4.35	5.92	5.92

l_c	$\tilde{\alpha}_1$	$\tilde{\alpha}_2$	$\tilde{\alpha}_3$	$\tilde{\alpha}_4$	$\tilde{\alpha}_5$	$\tilde{\alpha}_6$	$\tilde{\alpha}_7$	$\tilde{\alpha}_8$	$\tilde{\alpha}_9$	$\tilde{\alpha}_{10}$	$\tilde{\alpha}_{11}$
1/2	2.57	2.36	2.09	3.58	3.51	8.31	7.04	11.73	10.8	21.48	21.77
1/5	3.13	2.88	2.50	2.33	2.16	2.56	2.62	3.55	3.27	4.22	4.18

l_c	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$	$\hat{\alpha}_5$	$\hat{\alpha}_6$	$\hat{\alpha}_7$	$\hat{\alpha}_8$	$\hat{\alpha}_9$	$\hat{\alpha}_{10}$	$\hat{\alpha}_{11}$
1/2	9.47	7.91	7.92	15.14	15.14	21.87	21.87	30.16	30.17	42.96	42.95
1/5	10.98	9.25	9.27	10.24	10.24	12.03	12.04	15.34	15.34	16.17	16.17

Table 1: The first eleventh values of the weights obtained by a posteriori approach.

Figure 1 gives the cardinality of the grid $\mathcal{H}_\alpha(w, N)$ and $\mathcal{H}_{\hat{\alpha}}(w, N)$. We observe that the weights obtained by the error bound given in Proposition 4.2 are more convenient than those obtained using the error bound $O(e^{-\sqrt{m}g})$, since we obtained

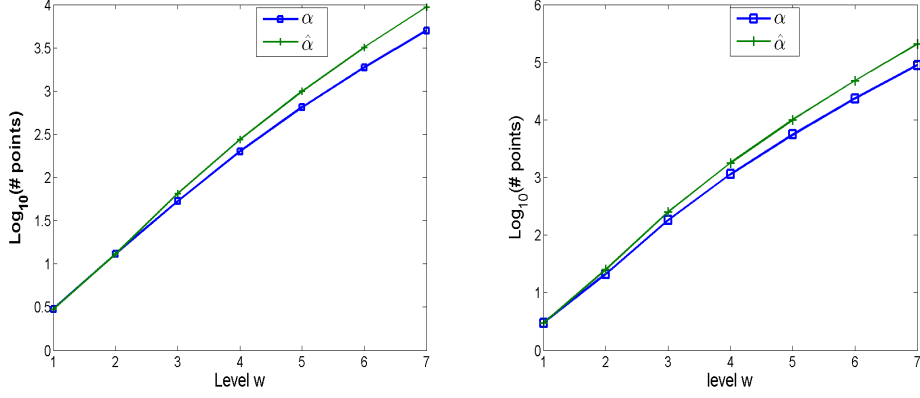$$\#\mathcal{H}_\alpha(w, N) \ll \#\mathcal{H}_{\hat{\alpha}}(w, N).$$

Figure 1: $\#\mathcal{H}_{\hat{\alpha}}(w, N)$ versus $\#\mathcal{H}_\alpha(w, N)$, (left $l_c = 0.5$, right $l_c = 0.2$).

In Figure (2), we compare $\#\mathcal{H}_\alpha(w, N)$ versus $\#\mathcal{H}_{\tilde{\alpha}}(w, N)$, we observe that they are closely comparable, since $\mathcal{H}_\alpha(w, N) \approx \mathcal{H}_{\tilde{\alpha}}(w, N)$.
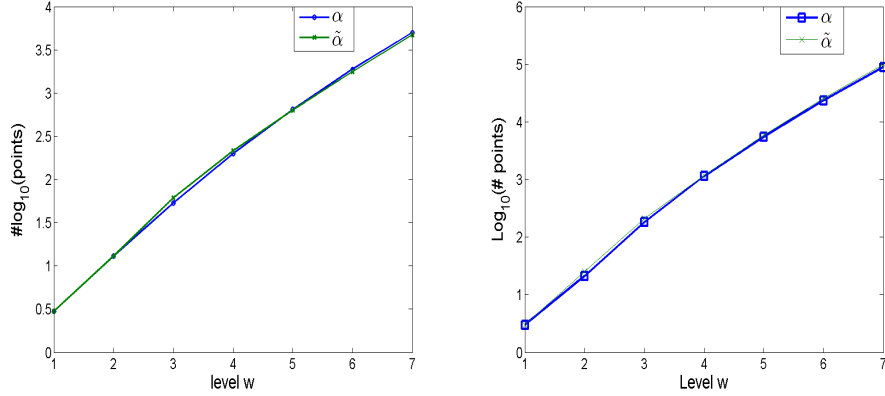

Figure 2: $\#\mathcal{H}_\alpha(w, N)$ versus $\#\mathcal{H}_{\tilde{\alpha}}(w, N)$, (left $l_c = 0.5$, right $l_c = 0.2$).

Proposition 4.1 shows that the error $C\|\lambda - \mathcal{A}_\alpha(w, N)\lambda\|$ is an upper bound of the error $\|u - \mathcal{A}_\alpha(w, N)u\|$, thus we use this error to determine a convenient level w satisfying $\|u - \mathcal{A}_\alpha(w, N)u\| \approx tol$. The strategy consists of seeking w such that $C\|\lambda - \mathcal{A}_\alpha(w, N)\lambda\| \approx tol$, where the constant C is given in (31).

Figure 3 shows that the errors $\|u - \mathcal{A}_\alpha(w, N)u\|$ and $C\|\lambda - \mathcal{A}_\alpha(w, N)\lambda\|$ have the same behavior. We have taken $\mathcal{A}_\alpha(9, N)$ and $\mathcal{A}_\alpha(7, N)$ as a reference solution for both functions u and λ, respectively when $l_c = \frac{1}{2}$ and $l_c = \frac{1}{5}$.
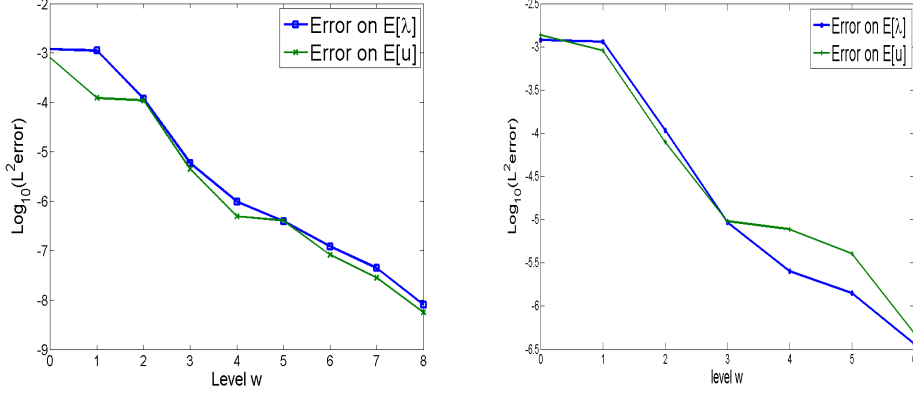
Figure 3: Error $C\|\mathbb{E}[\lambda_{ref} - \mathcal{A}_\alpha(w, N)\lambda]\|$ versus error on $\mathbb{E}[u]$ (left $l_c = 0.5$, right $l_c = 0.2$).

5.2 Linear case

In the second example we consider a diffusion parameter k_N which is linear with respect to the stochastic variable $y \in \Gamma$ where Γ is a bounded set in $\mathbb{R}^N$:

$$k(\omega, x) = \frac{\pi^2}{6} + \sum_{n=1}^{N} \frac{1}{1 + n^2} \cos(n\pi x_1) \sin(n\pi x_2) Y_n(\omega), \ x \in [0, 1]^2.$$

The deterministic load function is given by $f(\omega, x) = 1 + \sin(x_1)x_2 Y_1(\omega)$. The variables $Y_1, \ldots, Y_N$ are independent and uniformly distributed in the box $(-1, 1)$.

We take $N = 51$ and we compute α using a posteriori approach, we consider $\mathcal{A}_\alpha(9, N)$ as a reference solution for both functions u and λ. Table 2 gives the First ten weights for $\lambda = k^{-1}f$ and u.

	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	α_9	α_{10}
α for λ	3.89	5.77	7.19	8.26	9.11	9.87	10.48	11.01	11.48	11.9
$\tilde{\alpha}$ for u	4.24	6.44	8.00	9.32	10.19	10.92	11.53	12.06	12.53	12.96

Table 2: The first ten values of the weights obtained by a posteriori approach.

We can see that the components of α are slightly smaller than the components of $\tilde{\alpha}$. This is consistent with the theory, since when k is linear with respect to y, the analyticity region of u is larger than that of λ (proof of Proposition 4.1).

Figure 5.2 (left) shows that $\#\mathcal{H}_\alpha(w, N)$ is slightly smaller than $\#\mathcal{H}_{\tilde{\alpha}}(w, N)$. This could be explained by the fact that numerically we have $\frac{\alpha}{\min(\alpha)} \geq \frac{\tilde{\alpha}}{\min(\tilde{\alpha})}$.

Figure 5.2 (right) compares the error on $\mathbb{E}[u]$ with $C_w\|\mathbb{E}[\lambda_{ref} - \mathcal{A}_\alpha(w, N)\lambda]\|$, where C_w decreases with $M_w = \#\mathcal{H}_\alpha(w, N)$ the total number of points, we ap-

proximate it by $C_w \approx \frac{\|\mathcal{A}_\alpha(0,N)u\|}{\|\mathcal{A}_\alpha(0,N)\lambda\|} M_w^\mu$, where the decay rate is defined by $\mu = \min(\tilde\alpha) - \min(\alpha) = \tilde\alpha_1 - \alpha_1$.
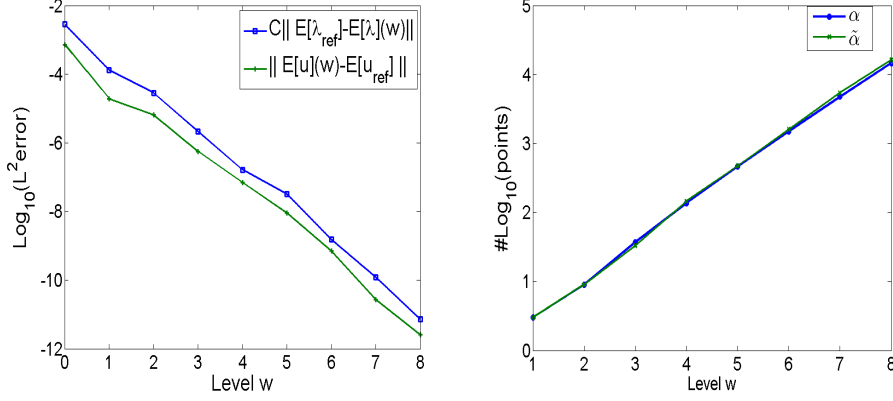

Figure 4: Left: Error $C_w\|\mathbb{E}[\lambda_{ref} - \mathcal{A}_\alpha(w,N)\lambda]\|_2$ versus error on $\mathbb{E}[u]$. Right: Sparse grid cardinality $\#\mathcal{H}_\alpha(w,N)$ versus $\#\mathcal{H}_{\tilde\alpha}(w,N)$.

6 Conclusion

In this work we have discussed one strategy of an adaptive collocation method based on the sparse grid method to approximate an elliptic problem with random inputs and data. The approximation of the problem is given by an anisotropic collocation in the stochastic space and the Finite Element approximation in the physical space. The emphasis has been on to seek the suitable number of knots in each dimension and the level of the method. We have introduced an error indicator based on the inverse of the diffusion coefficient and we have shown that this indicator gives an upper bound of the error on the solution.

The strategy consists of using this indicator to determine a convenient weight vector and level of the anisotropic sparse grid formula. The classical approach is based on a posteriori information on the solution and yields an additional cost, which is given by the resolution of a large sequence linear systems. Here the adaptation approach requires only a set of interpolation problems providing a significant reduction in the overall computation.

The numerical examples considered here are given with two examples of diffusion parameters with bounded and unbounded random variables. The numerical results show that we obtain a closely comparable weight vector to that obtained by classical approach. Also, we have obtained an equivalent error on the solution. The proposed adaptive approach yields substantial computational savings in comparison with the conventional adaptive approach. An immediate next step is to study the impact of the spatial discretization on the approximation quality and then make a posteriori analysis in the physical space.

19

References

[1] I.Babuŝka and P.Chatzipantelidis: *"On solving linear elliptic stochastic partial differential equations"*, Comput.Methods Appl.Mech.Engrg., 191:4093-4122, 2002.

[2] I. M. Babuŝka, 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.

[3] I.M. Babuŝka, F. Nobile, and R. Tempone: *"A stochastic collocation method for elliptic partial differential equations with random input data"*, SIAM J. Numer. Anal, 43(3):1005-1034, 2007

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

[5] P. G.Ciarlet. :*"The Finite Element Method for Elliptic Problems"*. North-Holland, 1978.

[6] J. Burkardt: *"SGMGA, Sparse Grid Mixed Growth Anisotropic Rules"*, http://people.sc.fsu.edu/ jburkardt/cpp-src/sgmga/sgmga.html

[7] P.J. Davis, P. Rabinowitz: *"Methods of Numerical Integration"*, Academic Press, New York, 1975.

[8] J. Erhel, Z. Mghazli, M. Oumouni: *"Calcul de l'espérance de la solution d'une EDP stochastique unidimensionnelle à l'aide d'une base réduite"*, C.R.Acad.Sci.Paris, Ser. I, 349 2011, 861-865.

[9] G. Fishman: *"Monte Carlo: Concepts, Algorithms, and Applications"*, Springer-Verlag, New York, 1996.

[10] T. Gerstner, M. Griebel: *"Numerical integration using sparse grid"*, Numer. Algorithms 18 1998 209-232.

[11] T. Gerstner and M. Griebel:*" Dimension-adaptive tensor-product quadrature"*. Computing, 71:65âĂŞ87, 2003.

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

[13] O.P. Le Maître, O.M. Knio, H.N. Najm, and R.G. Ghanem: *"Uncertainty propagation using Wiener-Haar expansions"*, Journal of Computational Physics, 197(1) :28-57, 2004.

[14] M. Loève: *"Probability Theory I"* I, fourth edition, in: Graduate Texts in Mathematics, vol. 45, Springer-Verlag, New York, 1977.

[15] M. Loève: *"Probability Theory"* II, fourth edition, Graduate Texts in Mathematics, Vol. 46. Springer-Verlag, New York, 1978.

[16] F. Nobile, R. Tempone, and C.Webster: *"A sparse grid stochastic collocation method for elliptic partial differential equations with random input data"*, SIAM. Volume 46 Issue 5, May 2008, pp. 2309-2345.

[17] F. Nobile, R. Tempone, and C.Webster: *"An anisotropic sparse grid stochastic collocation method for elliptic partial differential equations with random input data"*, SIAM J. Numer. Anal., 2008, vol. 46/5, pp. 2411–2442.

[18] S. Smolyak: *"Quadrature and Interpolation Formulas for Tensor Products of Certain Classes of Functions*, Doklady Akademii Nauk SSSR, Volume 4, 1963, pages 240-243.

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