

## A POSTERIORI ERROR ESTIMATION FOR THE STOCHASTIC COLLOCATION FINITE ELEMENT METHOD\*

DIANE GUIGNARD<sup>†</sup> AND FABIO NOBILE<sup>†</sup>

**Abstract.** In this work, we consider an elliptic partial differential equation (PDE) with a random coefficient solved with the stochastic collocation finite element method (SC-FEM). The random diffusion coefficient is assumed to depend in an affine way on independent random variables. We derive a residual-based a posteriori error estimate that is constituted of two parts controlling the SC error and the FE error, respectively. The SC error estimator is then used to drive an adaptive sparse grid algorithm. Several numerical examples are given to illustrate the efficiency of the error estimator and the performance of the adaptive algorithm.

**Key words.** PDEs with random input, finite element method, stochastic collocation method, a posteriori error estimation, adaptive algorithm

**AMS subject classifications.** 65N15, 65C20, 65N30, 65N35

**DOI.** 10.1137/17M1155454

**1. Introduction.** Partial differential equations (PDEs) are the mathematical formulation of many physical and engineering phenomena. For such problems, the input data are often affected by uncertainty due to either a lack of knowledge or an inherent variability of the system. Probability theory offers a possible way to describe the uncertainties, characterizing the uncertain input data with random variables or random fields and yielding PDEs with random inputs.

The development of efficient methods to tackle the numerical approximation of such problems has thus been of great interest and has attracted the attention of many scientists over the past decades. In this work, we will consider the stochastic collocation (SC) method [1, 2, 3] for the stochastic approximation, and the finite element method (FEM) for the physical space discretization. As sampling methods of Monte-Carlo type or quasi- and multilevel Monte-Carlo types [4, 5, 6, 7], and contrary to *intrusive* methods such as stochastic Galerkin [8, 9], the SC method requires only the solution of decoupled deterministic problems and thus allows the re-use of deterministic solvers. Moreover, exploiting the possible regularity of the solution with respect to the random parameters, the SC method has the advantage of a potentially much faster convergence rate than the Monte-Carlo method. It is also suitable for large uncertainties, contrary to perturbation-type methods as considered in our previous works [10, 11].

Whenever a numerical method is used to approximate the solution of the problem under consideration, an error analysis should be performed to estimate the numerical error thus introduced. The derivation of a priori error estimates for the stochastic collocation finite element method (SC-FEM) is done, e.g., in [1, 12, 13], but to our knowledge no a posteriori error estimate for the whole solution in suitable norms has been derived. It is of great importance to have a posteriori error estimators at our disposal since such estimators are the foundation of many adaptive strategies which aim at reaching a numerical solution with prescribed accuracy while keeping

---

\*Received by the editors November 6, 2017; accepted for publication (in revised form) August 13, 2018; published electronically October 23, 2018.

<http://www.siam.org/journals/sinum/56-5/M115545.html>

<sup>†</sup>SB-MATH, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland (diane.guignard@epfl.ch, fabio.nobile@epfl.ch).

the computational cost as low as possible. Here, the numerical solution is affected by two sources of error, namely the SC error and the FE error, and the estimator should not only provide an upper bound of the error but also furnish an estimation of the contribution of each error component to the total error, so that it can be used for balancing errors in an adaptive algorithm. We mention that recently, a posteriori error estimates for a specific quantity of interest, usually referred to as *goal-oriented* error estimates, have been developed; see, for instance, [14].

The main drawback of the SC method is that it suffers from the so-called *curse of dimensionality* when tensor grids are used; namely, the performance of the method deteriorates as the number of random variables increases. A remedy is then to exploit the possible anisotropy of the solution, in the sense that the different random variables might not have the same influence on the solution. Examples of works in this direction are the anisotropic sparse grid method proposed in [15] and the quasi-optimal sparse grid method introduced in [12]. In the latter, the adaptive algorithm is based on a priori error estimates whose constants are numerically tuned during the process, yielding what the authors called an a priori/a posteriori strategy. A proof of convergence has been obtained in [16] for the pure a priori algorithm. An a posteriori sparse grid adaptive algorithm was first proposed in [17] and then used, for instance, in [13, 18, 19, 20, 21, 22]. In [20], the adaptive process is driven by profit indicators obtained by solving additional PDEs. The method is applicable to a wide range of problems, including, for instance, the case of unbounded random variables or non-nested grids, and can be combined with a Monte-Carlo sampling, using a control variate technique, to handle rough random fields [23]. However, the error indicators proposed so far are heuristic and do not provide a certified control of the error.

We mention that adaptive strategies have also been investigated for the case when a different method is used for the stochastic space approximation. For instance, in [24] the solution is approximated via a Taylor series, and an adaptive algorithm is proposed with a proof of its convergence. In [25, 26], where the random PDEs are solved with the stochastic Galerkin FEM, the convergence is proved when the adaptation is performed in both physical and stochastic spaces. In this case, the extension of the results obtained for the adaptive finite element method (AFEM) in [27] is feasible and heavily depends on the so-called Galerkin orthogonality property. So far, at least to our knowledge, there is no proof of convergence for adaptive SC methods.

The main goal of this paper is to derive an a posteriori error estimate that controls both the FE error and the SC error. We consider an elliptic diffusion problem with random coefficient that depends in an affine way on a finite number of independent random variables. Moreover, we restrict our attention to the case where the source term is deterministic and the SC scheme is interpolatory. The error estimate we obtain is residual-based, provides an upper bound of the total error, is localizable, and hence is suitable for adaptive algorithms. We use then the SC error estimator to drive an adaptive sparse grid algorithm in which the collocation points are iteratively selected based on a criterion that uses the error estimator. It is important to mention that so far, we have no proof of convergence of the adaptive algorithm proposed here. Moreover, this algorithm is suitable only for random spaces of moderate dimension (or if the anisotropy in the problem is significant). An alternative procedure should be used for high-dimensional problems, adapting, for instance, the *dimension adaptive* strategy proposed in [20] to our context. Finally, since physical mesh refinements are rather standard, we focus here only on adaptive strategies in the stochastic dimension. The next step would be to propose an adaptive strategy with refinements in both the physical and random spaces, combining, for instance, the algorithm proposed here

for the adaptive selection of the collocation points with a standard AFEM for the physical mesh refinement.

The outline of the paper is the following. In section 2 we give the statement of the problem, namely an elliptic diffusion PDE with random coefficient. In section 3 we present the SC-FEM that we use to solve this problem approximately. Section 4 is devoted to the a posteriori error analysis, more precisely, to the derivation of a residual-based a posteriori error estimate that controls the two error components. In section 5 we give a possible strategy to adaptively construct the sparse grid using the stochastic error estimator. In section 6 we perform several numerical experiments to test the efficiency of the error estimator and the performance of the proposed adaptive strategy. Finally, section 7 contains some insight into how to deal with high-dimensional problems, and section 8 presents some conclusions.

**2. Problem statement.** Let  $D \subset \mathbb{R}^d$  be an open bounded domain with Lipschitz continuous boundary  $\partial D$ , and let  $(\Omega, \mathcal{F}, P)$  be a complete probability space. We seek  $u : \bar{D} \times \Omega \rightarrow \mathbb{R}$  that solves  $P$ -almost everywhere in  $\Omega$ , or in other words almost surely (a.s.), the problem

$$(1) \quad \begin{cases} -\nabla \cdot (a(\cdot, \omega) \nabla u(\cdot, \omega)) &= f(\cdot) \quad \text{in } D, \\ u(\cdot, \omega) &= 0 \quad \text{on } \partial D, \end{cases}$$

with deterministic forcing term  $f \in L^2(D)$  and random field  $a$  on  $(\Omega, \mathcal{F}, P)$  over  $W^{1,\infty}(D)$ . Moreover, we make the following assumptions on the random diffusion coefficient  $a$ :

$$(2) \quad \exists a_{\min}, a_{\max} : P(\omega \in \Omega : 0 < a_{\min} \leq a(\mathbf{x}, \omega) \leq a_{\max} < \infty \forall \mathbf{x} \in D) = 1$$

and

$$(3) \quad a(\mathbf{x}, \omega) = a_0(\mathbf{x}) + \sum_{n=1}^N a_n(\mathbf{x}) Y_n(\omega),$$

where  $(Y_n)_{n=1}^N$  are real-valued independent random variables. Thanks to the Doob-Dynkin lemma, the solution  $u$  depends on the same random variables as the diffusion coefficient  $a$ ; i.e., we have  $u(\mathbf{x}, \omega) = u(\mathbf{x}, Y_1(\omega), \dots, Y_N(\omega))$ . Let us introduce  $\Gamma = \Gamma_1 \times \dots \times \Gamma_N$  with  $\Gamma_n = Y_n(\Omega)$  for  $n = 1, \dots, N$ . Moreover, let  $\rho : \Gamma \rightarrow \mathbb{R}_+$  be the joint probability density function of the random vector  $\mathbf{Y} = (Y_1, \dots, Y_N)$ , which factorizes as  $\rho(\mathbf{y}) = \prod_{n=1}^N \rho_n(y_n)$  for all  $\mathbf{y} = (y_1, \dots, y_N) \in \Gamma$ . We can then replace the probability space  $(\Omega, \mathcal{F}, P)$  by  $(\Gamma, B(\Gamma), \rho(\mathbf{y}) d\mathbf{y})$ , where  $B(\Gamma)$  denotes the Borel  $\sigma$ -algebra defined on  $\Gamma$ , and  $\rho(\mathbf{y}) d\mathbf{y}$  denotes the probability measure of  $\mathbf{Y}$ . Due to assumption (2) on  $a$ , the energy and the  $H_0^1$  norms are equivalent; i.e., there exist  $0 < c_{\min} \leq c_{\max}$  such that

$$(4) \quad c_{\min} \|\nabla v(\mathbf{y})\|_{L^2(D)} \leq \|a^{\frac{1}{2}}(\mathbf{y}) \nabla v(\mathbf{y})\|_{L^2(D)} \leq c_{\max} \|\nabla v(\mathbf{y})\|_{L^2(D)} \quad \rho\text{-a.e. in } \Gamma$$

for any  $v(\mathbf{y}) \in H_0^1(D)$ . In particular, (4) holds with  $c_{\min} = \sqrt{a_{\min}}$  and  $c_{\max} = \sqrt{a_{\max}}$ . Finally, for a given Banach space  $V$  with norm  $\|\cdot\|_V$ , and for  $p \in [1, \infty]$ , we define the Bochner space

$$L_p^p(\Gamma; V) := \{v : \Gamma \rightarrow V \mid v \text{ is strongly measurable and } \|v\|_{L_p^p(\Gamma; V)} < \infty\}$$

with

$$\|v\|_{L^p_\rho(\Gamma;V)} := \begin{cases} \left( \int_\Gamma \|v(\mathbf{y})\|_V^p \rho(\mathbf{y}) d\mathbf{y} \right)^{\frac{1}{p}} & \text{if } p < \infty, \\ \rho\text{-ess sup}_{\mathbf{y} \in \Gamma} \|v(\mathbf{y})\|_V & \text{if } p = \infty. \end{cases}$$

The (parametric, pointwise) weak formulation of problem (1) reads as follows: find  $u : \Gamma \rightarrow V$  such that

$$(5) \quad \int_D a(\mathbf{x}, \mathbf{y}) \nabla u(\mathbf{x}, \mathbf{y}) \cdot \nabla v(\mathbf{x}) d\mathbf{x} = \int_D f(\mathbf{x}) v(\mathbf{x}) d\mathbf{x} \quad \forall v \in V, \rho\text{-a.e. in } \Gamma,$$

where  $V := H_0^1(D)$  is the usual Sobolev space that we endow with the gradient norm  $\|v\|_V := \|\nabla v\|_{L^2(D)}$ . By a straightforward application of the Lax–Milgram lemma, assumption (2) ensures the well-posedness of problem (5), namely that there exists a unique solution  $u \in L^2_\rho(\Gamma; V)$  which satisfies the a priori estimate

$$\|u\|_{L^2_\rho(\Gamma;V)} \leq \frac{C_P}{a_{\min}} \|f\|_{L^2(D)}.$$

In particular, we have  $u \in L^p_\rho(\Gamma; V)$  for any  $p \in [1, \infty]$ . Moreover, it has been shown (see, for instance, [1]) that the parametric solution  $u$  of problem (5) is analytic with respect to each parameter  $y_n \in \Gamma_n$ ,  $n = 1, \dots, N$ . Finally, we mention that imposing  $a(\cdot, \omega) \in L^\infty(D)$  is enough for the well-posedness of the problem. We assume  $W^{1,\infty}(D)$  regularity for ease of derivation of our a posteriori error estimate, more precisely, for the part controlling the FE error; see (17) below.

**3. Stochastic collocation finite element method.** In this section, we briefly present the SC-FEM for numerically solving PDEs with random input data, following closely [16] and focusing on the model problem (1). We also refer the reader to [1, 3] for a complete discussion on this method. The idea is to proceed in two steps: first, a semidiscretization of problem (5) using the FEM for the physical space approximation, and then the application of a collocation method for the stochastic space approximation using global polynomials in  $\mathbf{y}$ . We thus seek an approximate solution in a space  $\mathbb{P}(\Gamma) \otimes V_h$ , with  $\mathbb{P}(\Gamma) \subset L^2_\rho(\Gamma)$  a polynomial space on  $\Gamma$ , and  $V_h$  a FE subspace of  $V$ .

More precisely, for any  $h > 0$ , let  $\mathcal{T}_h$  be a regular triangulation of  $D$  with elements  $T$  of diameter  $h_T \leq h$ . We assume that there exists a constant  $c > 0$  satisfying

$$(6) \quad \frac{h_T}{\rho_T} \leq c \quad \forall T \in \mathcal{T}_h, \quad \forall h > 0,$$

where  $\rho_T = \sup\{\text{diam}(B) : B \text{ is a ball contained in } T\}$ . Let  $V_h \subset V$ , with  $\dim(V_h) = N_h$ , be the space of continuous, piecewise linear FE functions associated to  $\mathcal{T}_h$  that vanish on  $\partial D$ . The semidiscretized problem is therefore given by the following: find  $u_h : \Gamma \rightarrow V_h$  such that

$$(7) \quad \int_D a(\mathbf{x}, \mathbf{y}) \nabla u_h(\mathbf{x}, \mathbf{y}) \cdot \nabla v_h(\mathbf{x}) d\mathbf{x} = \int_D f(\mathbf{x}) v_h(\mathbf{x}) d\mathbf{x} \quad \forall v_h \in V_h, \rho\text{-a.e. in } \Gamma.$$

Problem (7) is then further discretized by considering a set  $\{\mathbf{y}_1, \dots, \mathbf{y}_{N_c}\}$  of  $N_c$  collocation points in  $\Gamma$  and building the global polynomial approximation

$$(8) \quad u_{h,N_c}(\mathbf{y}) := \sum_{k=1}^{N_c} u_h(\mathbf{y}_k) L_k(\mathbf{y})$$

for appropriate multivariate (for instance, Lagrange) polynomials  $L_k$ , where  $u_h(\mathbf{y}_k)$  is the solution of problem (7) with  $\mathbf{y} = \mathbf{y}_k$ . A possible choice for the collocation points  $\mathbf{y}_k \in \Gamma$  is to take the Cartesian product of certain abscissas in each direction. However, using such a tensor grid would rapidly become computationally too expensive due to the *curse of dimensionality*: the number of nodes increases exponentially with  $N$ . To alleviate this drawback, the idea is to use a so-called sparse grid, first introduced by Smolyak in [28]. Let us define

$$(9) \quad \mathcal{U}_n^{m(i_n)} : C^0(\Gamma_n) \rightarrow \mathbb{P}_{m(i_n)-1}(\Gamma_n),$$

a sequence of univariate polynomial interpolant operators along each direction  $\Gamma_n$  for  $n = 1, \dots, N$ , using abscissas  $\{\xi_j^{n,i_n}\}_{j=1}^{m(i_n)}$ . Here,  $m(i_n)$  denotes the number of collocation points used to build the interpolant of level  $i_n$ , and  $\mathbb{P}_q(\Gamma_n)$  is the space of polynomials in  $y_n$  of degree at most  $q$ . The function  $m$  should satisfy  $m(0) = 0$ ,  $m(1) = 1$ , and  $m(i) < m(i+1)$  for any  $i \geq 1$ . Moreover, let  $I \subset \mathbb{N}_+^N$  be a multi-index set, where  $\mathbb{N}_+ = \{1, 2, \dots\}$  denotes the positive integers. In what follows, the only restriction on  $I$  will be that it is a downward closed set (a.k.a. lower set), i.e., it satisfies

$$(10) \quad \forall \mathbf{i} \in I, \quad \mathbf{i} - \mathbf{e}_j \in I \quad \forall j = 1, \dots, N \text{ such that } i_j > 1.$$

This condition is necessary to get good approximation properties; see, for instance, [17]. Setting  $\mathcal{U}_n^0 = 0$  for  $n = 1, \dots, N$ , we define then the sparse grid interpolant  $S_I$  by

$$(11) \quad u_{h,I}(\mathbf{y}) := S_I[u_h](\mathbf{y}) := \sum_{\mathbf{i} \in I} \Delta^{\mathbf{m}(\mathbf{i})}(u_h)(\mathbf{y}),$$

where

$$\Delta^{\mathbf{m}(\mathbf{i})} := \bigotimes_{n=1}^N \Delta_n^{m(i_n)} := \bigotimes_{n=1}^N \left( \mathcal{U}_n^{m(i_n)} - \mathcal{U}_n^{m(i_n-1)} \right)$$

and  $\mathbf{m}(\mathbf{i}) = (m(i_1), \dots, m(i_N))$ . The operators  $\Delta_n^{m(i_n)}$  and  $\Delta^{\mathbf{m}(\mathbf{i})}$  are often referred to as *difference* (or *detail*) and *hierarchical surplus* operators, respectively. In what follows, we assume that

$$(12) \quad u_h(\mathbf{y}) = \sum_{\mathbf{i} \in \mathbb{N}_+^N} \Delta^{\mathbf{m}(\mathbf{i})}(u_h)(\mathbf{y}) \quad \rho\text{-a.e. in } \Gamma,$$

where the series converges absolutely in  $V$ , which holds if  $u$  is sufficiently smooth in  $\mathbf{y}$  and if the operators  $\mathcal{U}_n^{m(i_n)}$  in (9) are such that  $\bigotimes_{n=1}^N \mathcal{U}_n^{m(i_n)} u \rightarrow u$  in  $V$  as  $\mathbf{i} \rightarrow \infty$ . Finally, we mention that the operator  $S_I$  in (11) can be equivalently written as a linear combination of tensor grid interpolations (see, for instance, [29]),

$$(13) \quad S_I[u_h](\mathbf{y}) = \sum_{\mathbf{i} \in I} c_{\mathbf{i}} \bigotimes_{n=1}^N \mathcal{U}_n^{m(i_n)}(u_h)(\mathbf{y}), \quad c_{\mathbf{i}} := \sum_{\substack{\mathbf{j} \in \{0,1\}^N \\ (\mathbf{i}+\mathbf{j}) \in I}} (-1)^{|\mathbf{j}|},$$

with  $|\mathbf{j}| = \sum_{n=1}^N j_n$  for  $\mathbf{j} = (j_1, \dots, j_N)$ . Notice that many of the coefficients  $c_{\mathbf{i}}$  are actually zero: for instance, for  $\mathbf{i} \in I$ ,  $c_{\mathbf{i}} = 0$  if  $(\mathbf{i} + \mathbf{j}) \in I$  for all  $\mathbf{j} \in \{0, 1\}^N$ . We then

call *sparse grid* the set of  $N_c$  collocation points needed by (13) to compute  $S_I[u_h]$ . To summarize, the sparse grid interpolant  $S_I$  is characterized by the multi-index set  $I$ , the function  $m$  defining the number of collocation points on each level, and the type of univariate nodes. Note that  $I$  must contain the multi-index  $\mathbf{1}$ , which allows us to approximate constant functions.

Our error estimate will only be valid in the case when  $S_I$  is interpolatory, i.e., it satisfies  $S_I[f](\mathbf{y}_k) = f(\mathbf{y}_k)$  for  $k = 1, \dots, N_c$ , where  $\{\mathbf{y}_1, \dots, \mathbf{y}_{N_c}\}$  are the collocation points in the sparse grid underlying the multi-index set  $I$  and function  $m$ . Notice that such a property requires the use of nested sequences of univariate nodes  $\{\xi_j^{n, i_n-1}\} \subset \{\xi_j^{n, i_n}\}$ ; see, for instance, [30, p. 277]. Finally, we introduce the notion of margin  $\mathcal{M}_I$  and reduced margin, defined, respectively, by

$$\begin{aligned}\mathcal{M}_I &:= \{\mathbf{i} \in \mathbb{N}_+^N \setminus I : \mathbf{i} - \mathbf{e}_n \in I \text{ for some } n \in \{1, \dots, N\}\}, \\ \mathcal{R}_I &:= \{\mathbf{i} \in \mathcal{M}_I : \mathbf{i} - \mathbf{e}_n \in I \text{ for all } n = 1, \dots, N \text{ with } i_n > 1\}.\end{aligned}$$

**4. Residual-based a posteriori error estimate.** We will now derive an a posteriori error estimate for the error  $u - S_I[u_h]$ , which consists of two parts controlling the FE error and the SC error, respectively. We first give two results that we will use in the derivation of the error estimate.

**PROPOSITION 4.1.** *Let  $S_I$  be the operator defined in (11). Then for any  $f, g \in C^0(\Gamma)$  we have*

$$S_I[fg] = S_I[fS_I[g]].$$

*Proof.* Since  $S_I$  is assumed to be interpolatory, we have  $S_I[g](\mathbf{y}_k) = g(\mathbf{y}_k)$  for all  $k = 1, \dots, N_c$ . By the definition of  $S_I$ , we get then

$$\begin{aligned}S_I[fS_I[g]](\mathbf{y}) &= \sum_{k=1}^{N_c} (fS_I[g])(\mathbf{y}_k) L_k(\mathbf{y}) = \sum_{k=1}^{N_c} f(\mathbf{y}_k) S_I[g](\mathbf{y}_k) L_k(\mathbf{y}) \\ &= \sum_{k=1}^{N_c} f(\mathbf{y}_k) g(\mathbf{y}_k) L_k(\mathbf{y}) = S_I[fg](\mathbf{y})\end{aligned}$$

for any  $\mathbf{y} \in \Gamma$ . □

For any downward closed multi-index set  $I$ , let us define the polynomial space  $\mathbb{P}_I$  by

$$(14) \quad \mathbb{P}_I := \sum_{\mathbf{i} \in I} \mathbb{P}_{\mathbf{m}(\mathbf{i})-1} \quad \text{with} \quad \mathbb{P}_{\mathbf{m}(\mathbf{i})-1} := \mathbb{P}_{m(i_1)-1} \otimes \dots \otimes \mathbb{P}_{m(i_N)-1}.$$

We have the following approximation properties that will be crucial in the derivation of our error estimate.

**PROPOSITION 4.2.** *Let  $S_I$  be the operator defined in (11). Then*

1.  $S_I[f] \in \mathbb{P}_I \quad \forall f \in C^0(\Gamma)$ ;
2.  $S_I$  is exact on  $\mathbb{P}_I$ , i.e.,  $S_I[f] = f \quad \forall f \in \mathbb{P}_I$ .

*Proof.* See Proposition 1 in [31]. □

Finally, we introduce the (generalized) jump of a function  $\varphi$  across an edge ( $d = 2$ ) or a face ( $d = 3$ )  $e$  in the direction  $\mathbf{n}_e$  orthogonal to  $e$  by

$$[\varphi]_{\mathbf{n}_e}(\mathbf{x}) := \begin{cases} \lim_{t \rightarrow 0^+} (\varphi(\mathbf{x} + t\mathbf{n}_e) - \varphi(\mathbf{x} - t\mathbf{n}_e)) & \text{if } e \not\subset \partial D, \\ 0 & \text{if } e \subset \partial D. \end{cases}$$

We can now state our residual-based a posteriori error estimate.

**PROPOSITION 4.3.** *Let  $u$  and  $u_h$  be the solutions of (5) and (7), respectively, and let  $S_I[u_h]$  be the sparse grid approximation of  $u_h$  computed using the multi-index set  $I$ . There exists a constant  $C > 0$  depending only on the mesh aspect ratio  $c$  such that for any  $p \in [1, \infty]$ , we have*

$$(15) \quad \|u - S_I[u_h]\|_{L^p_\rho(\Gamma; V)} \leq \frac{1}{c_{min}^2} [C\eta_{FE} + \zeta_{SC}],$$

where

$$(16) \quad \eta_{FE} := \sum_{k=1}^{N_c} \eta_k \|L_k\|_{L^p_\rho(\Gamma)}, \quad \eta_k := \left( \sum_{T \in \mathcal{T}_h} \eta_{k,T}^2 \right)^{\frac{1}{2}},$$

with

$$(17) \quad \eta_{k,T} := h_T^2 \|f + \nabla \cdot (a(\mathbf{y}_k) \nabla u_h(\mathbf{y}_k))\|_{L^2(T)}^2 + \sum_{e \subset \partial T} h_e \left\| \frac{1}{2} [a(\mathbf{y}_k) \nabla u_h(\mathbf{y}_k) \cdot \mathbf{n}_e]_{\mathbf{n}_e} \right\|_{L^2(e)}^2$$

and

$$(18) \quad \zeta_{SC} := \sum_{\mathbf{i} \in \mathcal{M}_I} \zeta_{\mathbf{i}}, \quad \zeta_{\mathbf{i}} := \|\Delta^{\mathbf{m}(\mathbf{i})} (a \nabla S_I[u_h])\|_{L^p_\rho(\Gamma; L^2(D))}.$$

*Proof.* In what follows, all equations hold  $\rho$ -a.e. in  $\Gamma$  without specifically mentioning it. Moreover, the dependence of each function on variables will not necessarily be indicated unless ambiguity arises. For any  $v \in V$ , we have

$$(19) \quad \begin{aligned} \int_D a \nabla (u - S_I[u_h]) \cdot \nabla v &= \int_D f v - \int_D a \nabla S_I[u_h] \cdot \nabla v \\ &= S_I \left[ \underbrace{\int_D f v - \int_D a \nabla u_h \cdot \nabla v}_{=: A_1} \right] \\ &\quad + \underbrace{S_I \left[ \int_D a \nabla u_h \cdot \nabla v \right] - \int_D a \nabla S_I[u_h] \cdot \nabla v}_{=: A_2}. \end{aligned}$$

For the second equality, we used the fact that  $f$  is deterministic and thus  $S_I[f] = f$  for any multi-index set  $I$ . We analyze the terms  $A_1$  and  $A_2$  separately. For the first term, thanks to the Galerkin orthogonality we have

$$(20) \quad \begin{aligned} A_1 &= \sum_{k=1}^{N_c} \left[ \int_D f v - \int_D a(\mathbf{y}_k) \nabla u_h(\mathbf{y}_k) \cdot \nabla v \right] L_k(\mathbf{y}) \\ &= \sum_{k=1}^{N_c} \left[ \int_D f(v - v_h) - \int_D a(\mathbf{y}_k) \nabla u_h(\mathbf{y}_k) \cdot \nabla (v - v_h) \right] L_k(\mathbf{y}) \end{aligned}$$

for any  $v_h \in V_h$ . We take  $v_h = I_h v$  as the Clément interpolant of  $v$  for which we have the following interpolation error bounds [32]:

$$(21) \quad \|v - I_h v\|_{L^2(T)} \leq Ch_T \|\nabla v\|_{L^2(N(T))} \quad \text{and} \quad \|v - I_h v\|_{L^2(e)} \leq Ch_e^{\frac{1}{2}} \|\nabla v\|_{L^2(N(T_e))}$$

for any element  $T$  and any edge or face  $e$ . Here, for an internal edge or face  $e$ ,  $T_e$  is the union of the two elements sharing  $e$ . Moreover,  $N(T)$  denotes the patch of elements associated to  $T$ , i.e., all  $K \in \mathcal{T}_h$  with  $\bar{K} \cap \bar{T} \neq \emptyset$  (the definition of  $N(T_e)$  is analogous). After splitting the integral in (20) over each element  $T$  and integrating by parts, we obtain

$$(22) \quad A_1 \leq C \sum_{k=1}^{N_c} |L_k(\mathbf{y})| \eta_k \|\nabla v\|_{L^2(D)},$$

with  $\eta_k$  defined in (16). Notice that this term  $\eta_k$  is *deterministic*; namely, it does not depend on  $\mathbf{y}$ . It controls the FE error made when solving approximately the problem for the collocation point  $\mathbf{y}_k$ .

We now bound the second term  $A_2$ . We first notice that, thanks to Proposition 4.1, we have  $S_I[a \nabla u_h] = S_I[a \nabla S_I[u_h]]$  since  $S_I$  is assumed to be interpolatory. Therefore, using relation (12) we get

$$\begin{aligned} A_2 &= \int_D (S_I[a \nabla S_I[u_h]] - a \nabla S_I[u_h]) \cdot \nabla v = - \int_D \sum_{\mathbf{i} \notin I} \Delta^{\mathbf{m}(\mathbf{i})}(a \nabla S_I[u_h]) \cdot \nabla v \\ &= - \int_D \sum_{\mathbf{i} \in \mathcal{M}_I} \Delta^{\mathbf{m}(\mathbf{i})}(a \nabla S_I[u_h]) \cdot \nabla v \\ (23) \quad &\leq \left\| \sum_{\mathbf{i} \in \mathcal{M}_I} \Delta^{\mathbf{m}(\mathbf{i})}(a \nabla S_I[u_h]) \right\|_{L^2(D)} \|\nabla v\|_{L^2(D)}. \end{aligned}$$

We used the fact that  $a$  depends in an affine way on the random variables (see (2)) to restrict the summation over the multi-indices of the margin  $\mathcal{M}_I$  of  $I$ . Indeed, by Proposition 4.2 we have  $S_I[u_h] \in \mathbb{P}_I$ , with  $\mathbb{P}_I$  defined in (14), and by assumption,

$$a \in \mathbb{P}_0 + \sum_{n=1}^N \mathbb{P}_{\mathbf{e}_n}, \quad \text{with } \mathbb{P}_{\mathbf{e}_n} = \mathbb{P}_0 \otimes \cdots \otimes \mathbb{P}_0 \otimes \underbrace{\mathbb{P}_1}_{nth \text{ index}} \otimes \mathbb{P}_0 \cdots \otimes \mathbb{P}_0.$$

Therefore, we have  $a \nabla S_I[u_h] \in \sum_{n=1}^N \sum_{\mathbf{i} \in I} \mathbb{P}_{\mathbf{m}(\mathbf{i}) - \mathbf{1} + \mathbf{e}_n} \subset \mathbb{P}_{I \cup \mathcal{M}_I}$ , and thus

$$(24) \quad \Delta^{\mathbf{m}(\mathbf{i})}(a \nabla S_I[u_h]) = 0 \quad \forall \mathbf{i} \notin I \cup \mathcal{M}_I,$$

using again Proposition 4.2, namely, that  $S_{I \cup \mathcal{M}_I}$  is exact on  $\mathbb{P}_{I \cup \mathcal{M}_I}$ . Thanks to the norm equivalence (4), taking then  $v = u(\mathbf{y}) - S_I[u_h](\mathbf{y})$  in (19) and using the bounds (22) and (23) for the terms  $A_1$  and  $A_2$ , respectively, yields

$$(25) \quad \|\nabla(u(\mathbf{y}) - S_I[u_h](\mathbf{y}))\|_{L^2(D)} \leq \frac{1}{c_{min}^2} \left( C \sum_{k=1}^{N_c} |L_k(\mathbf{y})| \eta_k + \left\| \sum_{\mathbf{i} \in \mathcal{M}_I} \Delta^{\mathbf{m}(\mathbf{i})}(a \nabla S_I[u_h])(\mathbf{y}) \right\|_{L^2(D)} \right).$$

To conclude the proof, it only remains to take the  $L_\rho^p(\Gamma)$  norm on both sides of the last inequality and to use the triangle inequality for the norm  $L_\rho^p(\Gamma; L^2(D))$  to take the sum over the multi-indices  $\mathbf{i} \in \mathcal{M}_I$  outside the norm.  $\square$



Notice that in this proof, we have relied on the fact that  $S_I$  is interpolatory and that  $a$  depends in an affine way on the random variables. The latter allows us to restrict the summation over all the multi-indices outside  $I$  in the bound of  $A_2$  to the multi-indices belonging to the margin  $\mathcal{M}_I$ ; see (23). Moreover, it is worth mentioning that (25) yields a pointwise (in  $\mathbf{y}$ ) error estimate.

*Remark 4.4.* The spatial error estimate  $\eta_{FE}$  in (16) depends on  $\|L_k(\mathbf{y})\|_{L^p_\rho(\Gamma)}$ ,  $k = 1, \dots, N_c$ , i.e., on the stability constant of the operator  $S_I$ . These quantities can be bounded using the Lebesgue constant for  $S_I$ , whose growth depends on the choice of the function  $m$  and the family of interpolation points used by  $\mathcal{U}_n^{m(i)}$ ,  $n = 1, \dots, N$ . For instance, when using a doubling rule for  $m$  as in [30], defined by  $m(1) = 1$  and  $m(i) = 2^{i-1} + 1$  if  $i > 1$ , and Clenshaw–Curtis (CC) nodes, the Lebesgue constant associated with the operator  $S_I$  can be bounded by  $|I|^2$  [33]. As an alternative, we could bound the term  $A_1$  in (20) as

$$\begin{aligned} A_1 &= \sum_{T \in \mathcal{T}_h} \left[ \int_T \sum_{k=1}^{N_c} L_k(\mathbf{y})(f + \nabla \cdot (a(\mathbf{y}_k) \nabla u_h(\mathbf{y}_k)))(v - v_h) \right. \\ &\quad \left. + \frac{1}{2} \sum_{e \subset \partial T} \int_e \sum_{k=1}^{N_c} L_k(\mathbf{y})[a(\mathbf{y}_k) \nabla u_h(\mathbf{y}_k) \cdot \mathbf{n}_e]_{\mathbf{n}_e}(v - v_h) \right] \\ &\leq C \left( \sum_{T \in \mathcal{T}_h} \eta_T^2 \right)^{\frac{1}{2}} \|\nabla v\|_{L^2(D)}, \end{aligned}$$

with

$$\begin{aligned} \eta_T(\mathbf{y})^2 &:= h_T^2 \left\| \sum_{k=1}^{N_c} L_k(\mathbf{y})(f + \nabla \cdot (a(\mathbf{y}_k) \nabla u_h(\mathbf{y}_k))) \right\|_{L^2(T)}^2 \\ (26) \quad &+ \sum_{e \subset \partial T} h_e \left\| \frac{1}{2} \sum_{k=1}^{N_c} L_k(\mathbf{y})[a(\mathbf{y}_k) \nabla u_h(\mathbf{y}_k) \cdot \mathbf{n}_e]_{\mathbf{n}_e} \right\|_{L^2(e)}^2. \end{aligned}$$

Since  $(\sum_{T \in \mathcal{T}_h} \eta_T^2)^{\frac{1}{2}} \leq \sum_{T \in \mathcal{T}_h} \eta_T$ , we can then replace (15) by

$$(27) \quad \|u - S_I[u_h]\|_{L^p_\rho(\Gamma; V)} \leq \frac{1}{c_{min}^2} \left[ C \sum_{T \in \mathcal{T}_h} \|\eta_T\|_{L^p_\rho(\Gamma)} + \zeta_{SC} \right].$$

Mesh refinement, using the error estimate of Proposition 4.3 or the one proposed here, would lead to different adaptive strategies. The estimator in (16) gives an estimation of the spatial error for each collocation point that is further localized on each element  $T \in \mathcal{T}_h$ . Indeed, the estimator  $\eta_{k,T}$  in (17) is an estimator of the FE error for the element  $T$  and the collocation point  $\mathbf{y}_k$ . Therefore, different spatial meshes could be considered for different collocation points. On the contrary, the estimator in (26) gives an estimation of the spatial error for each element  $T \in \mathcal{T}_h$  and contains the contribution of all the collocation points. In this case, the same spatial mesh would then be used for all the collocation points.

**5. Adaptive algorithm.** The error estimator deduced from Proposition 4.3 can be used to adaptively refine the mesh and increase the multi-index set. Such an

adaptive strategy aims at reaching a given accuracy of the (FE and SC) error with computational cost as low as possible. Since the theory for mesh adaptation, often referred to as the AFEM, is well developed and studied, we will focus on the SC error. More precisely, we will consider an adaptive construction of the multi-index set  $I$  proceeding similarly to that originally proposed in [17] and further used, for instance, in [13, 20].

We give below a possible adaptive strategy which uses the error estimators  $\zeta_{\mathbf{i}}$  given in (18) to drive the process, with the requirement that the multi-index set  $I$  must remain downward closed during the adaptation. Basically, at each iteration we select the multi-index in the margin  $\mathcal{M}_I$  of the current set  $I$  that has the largest profit, with the latter being defined as follows: for any  $\mathbf{i} \in \mathcal{M}_I$ , we define

$$(28) \quad P_{\mathbf{i}} := \frac{\sum_{\mathbf{j} \in A_{\mathbf{i}}} \zeta_{\mathbf{j}}}{\sum_{\mathbf{j} \in A_{\mathbf{i}}} W_{\mathbf{j}}},$$

where  $A_{\mathbf{i}} = J_{\mathbf{i}} \setminus I$  and  $J_{\mathbf{i}}$  is the downward closed set of minimal cardinality containing  $I \cup \{\mathbf{i}\}$ ; i.e.,  $A_{\mathbf{i}}$  is the set containing  $\mathbf{i}$  plus all the multi-indices  $\mathbf{j} \in \mathcal{M}_I$  that must also be included in  $I$  if  $\mathbf{i}$  is added to  $I$  so that the set remains downward closed. Moreover, we have denoted by  $W_{\mathbf{i}}$  the *work* contribution of the multi-index  $\mathbf{i}$ , which can be defined as in [20] by

$$(29) \quad W_{\mathbf{i}} := \Pi_{n=1}^N (m(i_n) - m(i_n - 1)).$$

In the case of nested sets of points, as considered here, it corresponds to the number of new points in  $\Gamma$  introduced if  $\mathbf{i}$  is added to  $I$ . We could also choose to set  $W_{\mathbf{i}} = 1$  if we want to drive the adaptation based only on the error estimators. Finally, notice that for any  $\mathbf{i} \in \mathcal{R}_I$ , since  $I \cup \{\mathbf{i}\}$  is always downward closed, we have  $A_{\mathbf{i}} = \{\mathbf{i}\}$ , and the profit is simply given by  $P_{\mathbf{i}} = \frac{\zeta_{\mathbf{i}}}{W_{\mathbf{i}}}$ .

We can now introduce the adaptive algorithm we consider in this work.

---

**Algorithm 1** Adaptive algorithm (stochastic space adaptation).

---

**Require:**  $Tol > 0$

**Ensure:** multi-index set  $I$  such that  $\zeta_{SC} \leq Tol$

```

1:  $I = \{\mathbf{1}\}$ ,  $u_{h,I} = S_I[u_h]$ ,  $\zeta_{SC} = \zeta_{\mathbf{1}}$ 
2: while  $\zeta_{SC} > Tol$  do
3:    $\mathbf{i}^* = \operatorname{argmax}_{\mathbf{i} \in \mathcal{M}_I} P_{\mathbf{i}}$            select the most profitable multi-index
4:    $I \leftarrow I \cup A_{\mathbf{i}^*}$                  update the multi-index set
5:    $u_{h,I} = S_I[u_h]$                    compute the new sparse grid approximation
6:    $\zeta_{SC} = \sum_{\mathbf{i} \in \mathcal{M}_I} \zeta_{\mathbf{i}}$          compute the error estimator (18)
7: end while
```

---

*Remark 5.1.* Algorithm 1 is one possible adaptive strategy. In particular, we choose to select only one multi-index at each iteration; see line 3. Another possibility would be to allow the selection of several multi-indices, for instance, to satisfy a Dörfler-type criterion. Moreover, the selection of the most profitable element is made on the full margin in Algorithm 1. To reduce the computational cost, we could alternatively drive the adaptive process only by the profit of the elements of the reduced margin. In such a case, we do not need to compute  $\zeta_{\mathbf{i}}$  for each  $\mathbf{i} \in \mathcal{M}_I \setminus \mathcal{R}_I$ . However, the global error estimator  $\zeta_{SC}$  would then no longer be available, and another stopping criterion must be used.

**6. Numerical results.** We consider here numerical examples to test the efficiency of the SC error estimator derived in Proposition 4.3 and, in particular, to test the performance of Algorithm 1. In all what follows, the FE error is not accounted for. Moreover, we consider the case  $p = \infty$ , and we thus consider the error and estimator defined, respectively, by

$$(30) \quad \|\nabla(u_h - S_I[u_h])\|_{L_\rho^\infty(\Gamma; L^2(D))} \quad \text{and} \quad \sum_{\mathbf{i} \in \mathcal{M}_I} \|\Delta^{\mathbf{m}(\mathbf{i})}(a \nabla S_I[u_h])\|_{L_\rho^\infty(\Gamma; L^2(D))}.$$

In order to compute the  $L_\rho^\infty(\Gamma)$  norm approximately, we use a set  $\Theta \subset \Gamma$  of finite cardinality; that is, we use the approximation

$$\|g\|_{L_\rho^\infty(\Gamma)} \approx \max_{\mathbf{y} \in \Theta} |g(\mathbf{y})|$$

for any  $g \in C^0(\Gamma)$ . In what follows, we set  $\Theta$  to be constituted of 500 points randomly sampled according to the distribution  $\rho$ . Finally, instead of taking the a priori bound  $\sqrt{a_{\min}}$  for  $c_{\min}$ , which may be overly conservative, we estimate it by

$$c_{\min} := \min_{\mathbf{y} \in \Xi} \frac{\|a^{\frac{1}{2}}(\mathbf{y}) \nabla v(\mathbf{y})\|_{L^2(D)}}{\|\nabla v(\mathbf{y})\|_{L^2(D)}},$$

with  $\Xi \subset \Gamma$  a random set of small cardinality and  $v = u_h$  or  $v = u_h - S_I[u_h]$  using any (small) set  $I$ . In all the numerical examples below, we have observed that  $c_{\min} \approx 1$ , i.e., the energy and the  $H_0^1$  norms of the error are comparable, which is the reason why we decided to drop this constant and consider the estimator defined in (30).

*Remark 6.1.* In practice, we first compute (approximately) the error estimator  $\zeta_{\mathbf{i}}$  defined in (18) for each  $\mathbf{i} \in \mathcal{R}_I$ , the reduced margin of  $I$ , using

$$(31) \quad \zeta_{\mathbf{i}} \approx \max_{\mathbf{y} \in \Theta} \|S_{G \cup \{\mathbf{i}\}}[a \nabla S_I[u_h]](\mathbf{y}) - S_G[a \nabla S_I[u_h]](\mathbf{y})\|_{L^2(D)},$$

with  $G = I$ . We then proceed *layer by layer*, namely, we compute the estimator for each  $\mathbf{i} \in \mathcal{M}_I$  that belongs to the reduced margin of  $I \cup \mathcal{R}_I$  using (31) with  $G = I \cup \mathcal{R}_I$ , and so on until the full margin  $\mathcal{M}_I$  is covered (which requires a finite number of iterations). A key point is that no additional PDE solve is needed throughout this process.

Before performing sparse grid adaptation, we will test the efficiency of the SC estimator considering different approximation spaces chosen a priori. We will consider both cases  $m(i) = i$  and

$$(32) \quad m(i) = \begin{cases} 0 & \text{if } i = 0, \\ 1 & \text{if } i = 1, \\ 2^{i-1} + 1 & \text{if } i > 1. \end{cases}$$

Since we need nested sequences of points, we use Leja points for the linear case  $m(i) = i$  and CC points if  $m$  is defined by (32). We recall that for a generic compact set  $X$  and a given initial point  $y^0 \in X$ , the (standard) Leja points are defined recursively by [34]

$$y^k = \operatorname{argmax}_{y \in X} \prod_{j=1}^{k-1} (y - y^j), \quad k = 1, 2, \dots$$

In what follows, when using Leja points on an interval  $\Gamma_i = [a_i, b_i] \subset \mathbb{R}$ , we will set the initial point to the endpoint  $b_i$ . To test the efficiency of the estimator, we will consider an arbitrary (downward closed) multi-index set  $I$  or, for a given level of approximation  $w$ , the classical approximation spaces [31] given in Table 1.

TABLE 1  
*Approximation spaces for testing the efficiency of the SC error estimator.*

Approximation space	$m$	$I$	Points
Tensor product (TP)	$m(i) = i$	$I(w) = \{\mathbf{i} \in \mathbb{N}_+^N : \max_n (i_n - 1) \leq w\}$	Leja
Total degree (TD)	$m(i) = i$	$I(w) = \{\mathbf{i} \in \mathbb{N}_+^N : \sum_n (i_n - 1) \leq w\}$	Leja
Hyperbolic cross (HC)	$m(i) = i$	$I(w) = \{\mathbf{i} \in \mathbb{N}_+^N : \Pi_n (i_n) \leq w + 1\}$	Leja
Smolyak (SM)	$m$ in (32)	$I(w) = \{\mathbf{i} \in \mathbb{N}_+^N : \sum_n (i_n - 1) \leq w\}$	CC

**6.1. First example.** We start with the analysis of an inclusion problem, first with  $N = 2$  inclusions and then with  $N = 8$  inclusions; see [31]. The physical domain is the unit square  $D = (0, 1)^2$  in which we identify the subdomains  $F$  and  $C_n$ ,  $n = 1, \dots, N$ , as depicted in Figure 1(left) for the case  $N = 2$  and in Figure 7(left) for  $N = 8$ . The square subdomain  $F$  has a side length of 0.2, while the radius of each circular subdomain  $C_n$  is equal to 0.13. We set the forcing term to  $f(\mathbf{x}) = 100\chi_F(\mathbf{x})$  and define the random diffusion coefficient by

$$(33) \quad a(\mathbf{x}, \mathbf{Y}(\omega)) = a_0(\mathbf{x}) + \sum_{n=1}^N \gamma_n \chi_n(\mathbf{x}) Y_n(\omega) \quad \text{with } a_0 = 1,$$

where  $\chi_F$  and  $\chi_n$ ,  $n = 1, \dots, N$ , denote the indicator function of each subdomain. The parameters  $\gamma_n$ ,  $n = 1, \dots, N$ , are used to introduce anisotropy in the problem, assigning more importance to one or another direction  $y_n$ .

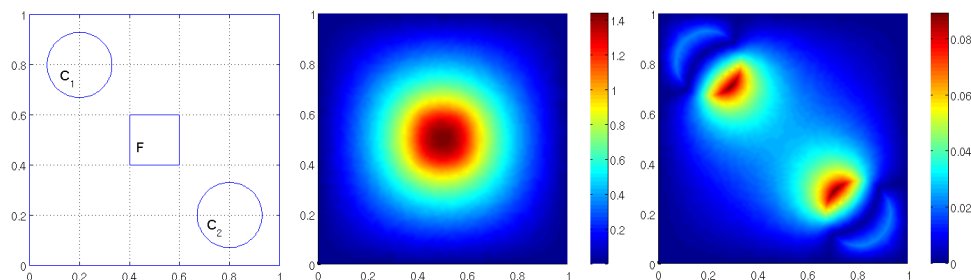


FIG. 1. *Geometry of the problem (left), expected value (middle), and standard deviation (right) of the solution for the case  $\gamma_1 = \gamma_2 = 1$ .*

**Case  $N = 2$ .** We start with  $N = 2$  and take  $Y_n \sim \mathcal{U}[-0.99, 0.99]$  for  $n = 1, 2$ . The FE mesh we are using consists of 4961 vertices and 9696 triangles, with minimal and maximal diameter  $h_T$  of about  $7.367\text{e-}3$  and  $2.854\text{e-}2$ , respectively. The mean and the standard deviation of the solution is given in Figure 1 for the isotropic case  $\gamma_1 = \gamma_2 = 1$ .

In Figure 2 we give the error and the estimator with respect to the number of points in the sparse grids for the four types of approximation spaces defined in Table

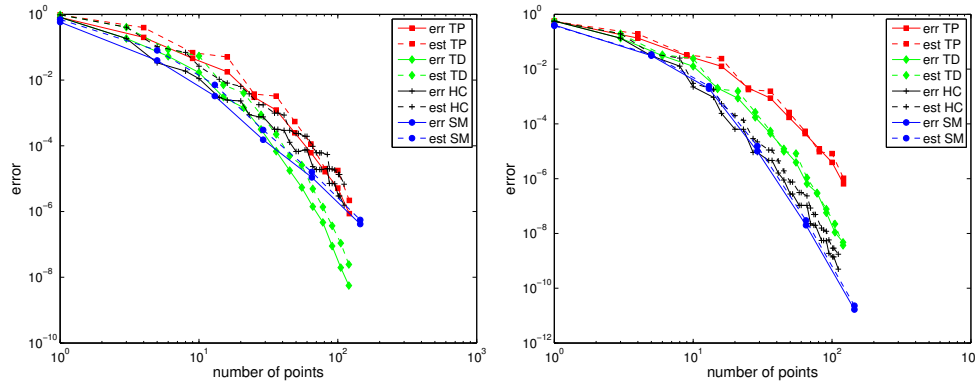


FIG. 2. Error and estimator with respect to the number of points in logarithmic scale for the four approximation spaces given in Table 1. Isotropic case  $\gamma_1 = \gamma_2 = 1$  (left), and anisotropic case  $\gamma_1 = 1$  and  $\gamma_2 = 0.1$  (right).

1. We consider the isotropic case  $\gamma_1 = \gamma_2 = 1$  but also an anisotropic one, namely,  $\gamma_1 = 1$  and  $\gamma_2 = 0.1$ . The maximum level of approximation  $w$  is set to 10 for TP, 14 for TD, 29 for HC, and 5 for SM, which correspond to a sparse grid of 121, 120, 111, and 145 points in  $\Gamma$ , respectively.

We can see that the estimator provides a good control of the error for all the considered approximation spaces and for both the isotropic and the anisotropic cases. This is also the case when an arbitrary multi-index set is considered. Indeed, let us take, for instance,  $I_1 = [(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (2, 1), (2, 2), (2, 3), (3, 1)]$ , which is a priori not a good set for the considered values of  $\gamma_1$  and  $\gamma_2$  as it uses more points for  $y_2$  rather than  $y_1$ , and  $I_2 = [(1, 1), (1, 2), (2, 1), (2, 2), (3, 1), (4, 1), (5, 1)]$ . The results we obtain for the two cases  $m(i) = i$  with Leja points and  $m$  in (32) with CC points are presented in Table 2. Finally, we mention that we observe similar behavior for all the numerical examples presented below.

TABLE 2

Number of points, error, and estimator for the given multi-index sets  $I_1$  and  $I_2$ . Case iso:  $\gamma_1 = \gamma_2 = 1$ . Case aniso:  $\gamma_1 = 1$  and  $\gamma_2 = 0.1$ .

		$m(i) = i$ and Leja points			$m$ in (32) and CC points		
		# pts	error	estimator	# pts	error	estimator
iso	$I_1$	9	3.5977e-2	4.8307e-2	29	2.2348e-3	3.7151e-3
	$I_2$	7	1.4035e-1	2.1013e-1	23	3.1768e-2	3.5094e-2
aniso	$I_1$	9	3.0888e-2	3.3697e-2	29	1.9359e-3	2.3869e-3
	$I_2$	7	2.4784e-3	4.5417e-3	23	8.3280e-5	1.3008e-4

We now consider the adaptive strategy proposed in Algorithm 1. From now on, we restrict our attention to CC nodes and  $m$  defined in (32). We start with the isotropic case  $\gamma_1 = \gamma_2 = 1$ . We set the tolerance to  $Tol = 10^{-6}$ . The evolution of the multi-index set  $I$  during the adaptive process is presented in Figure 3. The multi-index in green denotes the selected element at the current iteration of Algorithm 1, i.e., the one with the highest profit, before it is added to  $I$ .

We can detect the isotropy of the problem by the *symmetrical* construction of the multi-index set. For instance, at iteration 11 the point  $(4, 2)$  is added, while  $(2, 4)$  is selected at the next iteration. Moreover, we see that the estimator provides a good

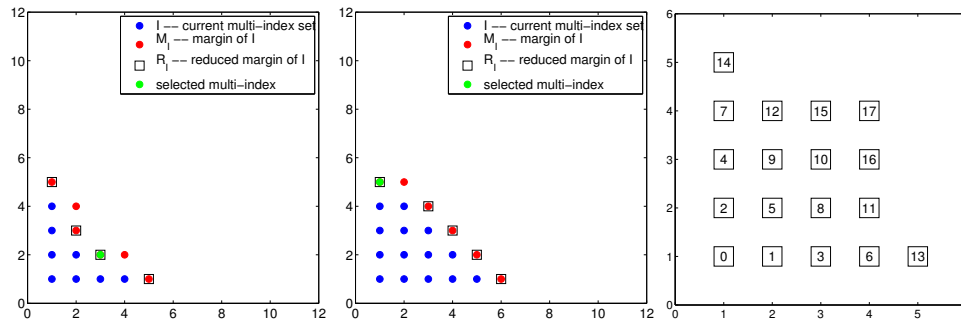


FIG. 3. Evolution of  $I$  during the adaptive process for the case  $\gamma_1 = \gamma_2 = 1$ . From left to right: iterations 8 and 14 and order of selection of the multi-indices. (See online version for color.)

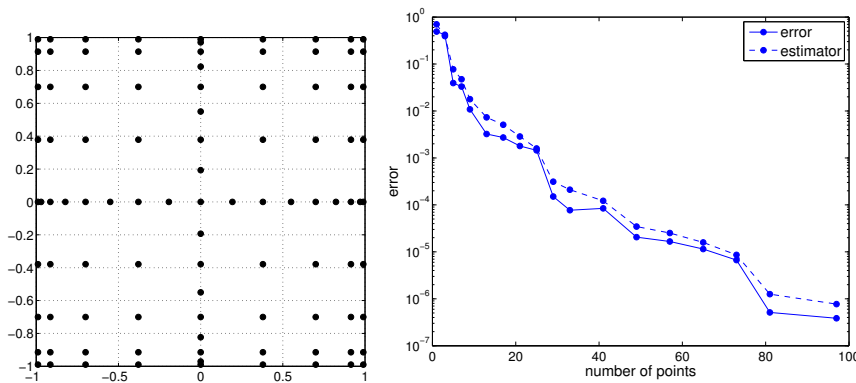


FIG. 4. Final sparse grid (left) and error and estimator with respect to the number of points in semilogarithmic scale (right) for the case  $\gamma_1 = \gamma_2 = 1$ .

control of the error as shown in Figure 4, where the final sparse grid is also given. It has been obtained after 17 iterations, yielding a grid of 97 points and an error and an estimator of about  $3.8464\text{e-}7$  and  $7.6980\text{e-}7$ , respectively. The error in the energy norm at this final stage, namely  $\|a^{\frac{1}{2}} \nabla e\|_{L^\infty_\rho(\Gamma; L^2(D))}$ , is about  $3.0020\text{e-}7$  and thus close to the error in the  $H^1$  seminorm. Finally, we mention that the highest profit of the elements of the margin of this final stage is about  $2.3220\text{e-}8$  and is achieved at  $(2, 5)$ , which belongs to the reduced margin.

We now set different values for  $\gamma_1$  and  $\gamma_2$  in (33) to see whether the adaptive algorithm is able to capture the anisotropy of the problem. We thus set  $\gamma_1 = 1$  and  $\gamma_2 = 0.1$ . In Figure 5 we present the set  $I$  at various steps of the adaptive construction. As expected, the algorithm clearly identifies a preferred direction, namely, the horizontal direction which corresponds to  $y_1$ .

The final sparse grid for a tolerance of  $\text{ToI} = 10^{-6}$  in Algorithm 1 is given in Figure 6 and has been reached in 10 iterations. In this case, there are 41 points in the sparse grid, the error and estimator are  $6.9851\text{e-}8$  and  $1.2506\text{e-}7$ , respectively, and the maximal profit among the elements of the margin is about  $2.0030\text{e-}8$  at  $(3, 3)$ , which belongs to the reduced margin. Finally, the error in the  $H^1$  seminorm is comparable to the error in the energy norm, which is about  $6.3569\text{e-}8$ .

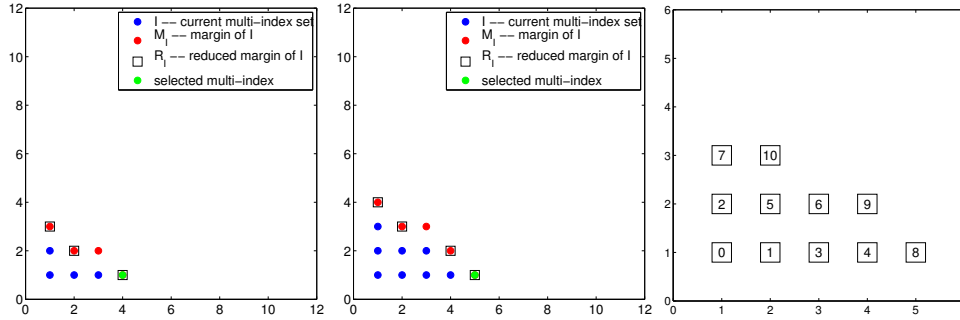


FIG. 5. Evolution of the multi-index set  $I$  during the adaptive process for the case  $\gamma_1 = 1$  and  $\gamma_2 = 0.1$ . From left to right: iterations 4 and 8 and order of selection of the multi-indices.

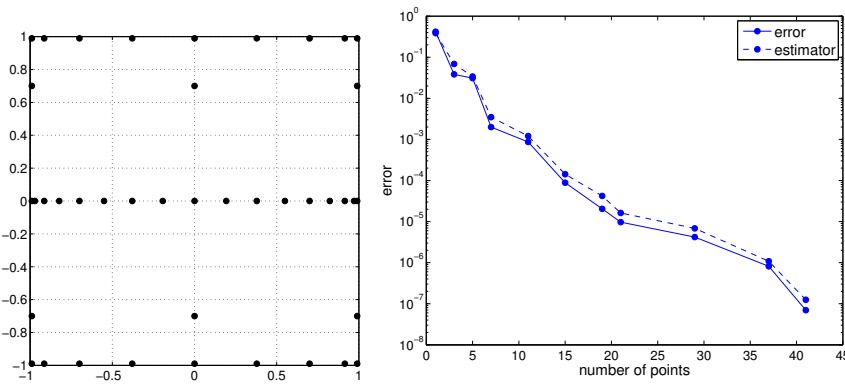


FIG. 6. Final sparse grid (left) and error and estimator with respect to the number of points in semi-logarithmic scale (right) for the case  $\gamma_1 = 1$  and  $\gamma_2 = 0.1$ .

**Case  $N = 8$ .** To conclude on this inclusion problem, we consider the case  $N = 8$  as in [31] and choose  $Y_n \sim \mathcal{U}[-0.99, 0.2]$  for  $n = 1, \dots, 8$  in (33). The geometry is given in Figure 7(left), where the value of the coefficients  $\gamma_n$ ,  $n = 1, \dots, 8$ , is also given. The FE mesh we are using contains 3805 vertices and 7416 triangles with minimal diameter and maximal diameter  $h_T$  of about  $1.0041e-2$  and  $3.1153e-2$ , respectively. For this case, we set the tolerance to  $Tol = 10^{-3}$  in Algorithm 1.

In Figure 7(right), we give the error and the estimator with respect to the number of points in the grid. At the final stage, obtained in 79 iterations, the sparse grid contains 363 points, and the error and estimator are about  $1.0852e-4$  and  $9.9014e-4$ , respectively. Moreover, the error in the energy norm is about  $8.7246e-5$ . Finally, the maximum profit among the elements of the margin is about  $5.4553e-6$  and is achieved at  $(1, 1, 1, 2, 1, 2, 2, 1)$ .

In this case, the estimator still provides a reasonable control of the error, even though it is less efficient than for the case  $N = 2$ . We see several possible explanations for this behavior and give a nonexhaustive list below. First, we have not been able to prove that the error estimator provides a lower bound for the error. The difficulties arise from, among other things, the lack of *Galerkin orthogonality* and the use of the triangle inequality to localize the estimator on each multi-index of the margin. Moreover, we are not taking into account the error due to the approximation of the  $L_\rho^\infty(\Gamma)$  norm, and further investigation should be made in this direction, namely, trying to

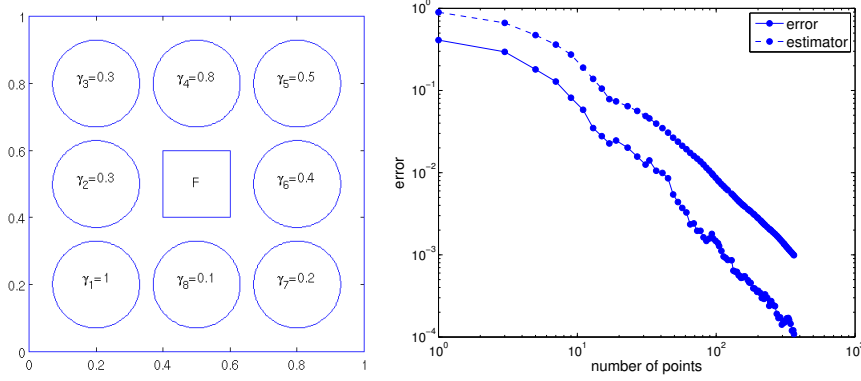


FIG. 7. Geometry of the problem for  $N = 8$  with indication of the coefficients  $\gamma_n$ ,  $n = 1, \dots, 8$ , (left) and error and estimator with respect to the number of points in logarithmic scale (right).

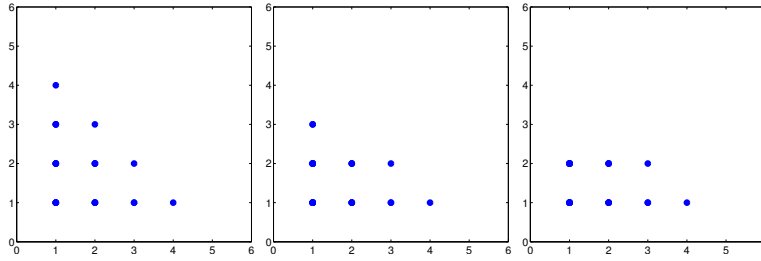


FIG. 8. Projection of the multi-index set  $I$ , obtained for  $Tol = 10^{-3}$  in Algorithm 1, on  $(y_1, y_4)$  (left),  $(y_1, y_5)$  (middle), and  $(y_1, y_8)$  (right).

quantify this additional error and perform additional tests with other training sets  $\Theta$ . The size of the training set could also be adapted with respect to the number of points in the sparse grid and not be fixed once and for all as considered here.

The projection of the obtained multi-index set  $I$  over two directions, namely,  $y_1$  and  $y_4$ ,  $y_1$  and  $y_5$ , and  $y_1$  and  $y_8$ , is presented in Figure 8. These results are consistent with the choice we made for the value of the coefficients  $\gamma_n$ ,  $n = 1, 2, \dots, 8$ ; see Figure 7(left).

**6.2. Second example.** As a second numerical experiment, we consider problem (1) with  $D = (0, 1)^2$ ,  $f(\mathbf{x}) = 32(x_1(1 - x_1) + x_2(1 - x_2))$ , and

$$(34) \quad a(\mathbf{x}, \mathbf{Y}(\omega)) = 1 + \sum_{n=1}^N \frac{\cos(2\pi n x_1) + \cos(2\pi n x_2)}{(\pi n)^2} Y_n(\omega) \quad \text{with} \quad Y_n \sim \mathcal{U}[-\sqrt{3}, \sqrt{3}]$$

for  $\mathbf{x} = (x_1, x_2) \in D$ . We use a spatial mesh consisting of 2673 vertices and 5184 triangles with minimum diameter and maximum diameter  $h_T$  of about 0.01 and 0.04, respectively. Finally, we consider the two cases  $N = 3$  and  $N = 5$  and set the tolerance to  $Tol = 10^{-6}$  in Algorithm 1.

The results for the case  $N = 3$  are given in Figure 9. We plot the error and the estimator with respect to the number of collocation points. We also give the projection of the final multi-index set  $I$  over two directions, namely,  $y_1$  and  $y_3$ . For this final state, obtained in 27 iterations, the error and the estimator are about  $4.3746\text{e-}7$  and  $9.2363\text{e-}7$ , respectively, and the grid contains 141 points. The error in the energy



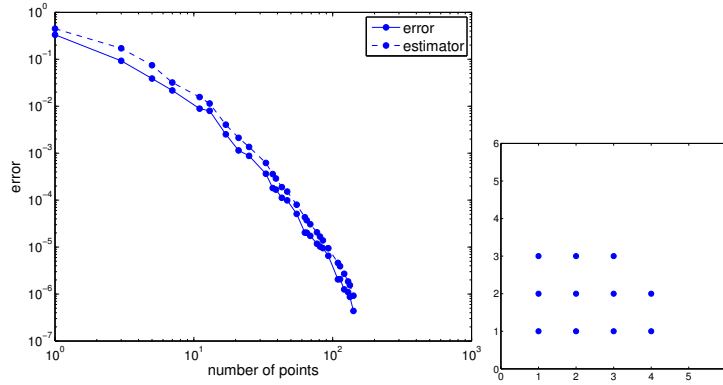


FIG. 9. Error and estimator with respect to the number of points in logarithmic scale (left) and projection of the final multi-index set on  $(y_1, y_3)$  (right) for the case  $N = 3$ .

norm is about  $3.5904e-7$ . Finally, we mention that the multi-index added in the last iteration to the final set  $I$  is  $(4, 3, 1)$  and that the maximum profit among the elements of  $\mathcal{M}_I$  is about  $3.0550e-8$  and is reached at  $(3, 2, 3)$ , which belongs to  $\mathcal{R}_I$ .

Figure 10 contains the results for the case  $N = 5$ . The final multi-index set  $I$  is projected on  $y_1$  and  $y_3$  and on  $y_1$  and  $y_5$ . The final grid has 973 points, for an error and estimator of about  $1.7666e-7$  and  $9.9454e-7$ , respectively, and was reached in 110 iterations. The error in the energy norm at this final stage is about  $1.4942e-7$ . Finally, the last multi-index added to the set is  $(4, 2, 1, 2, 2)$ , and the maximum profit among the elements of the margin of the final set is about  $3.9748e-9$  at  $(4, 3, 1, 2, 1) \in \mathcal{R}_I$ .

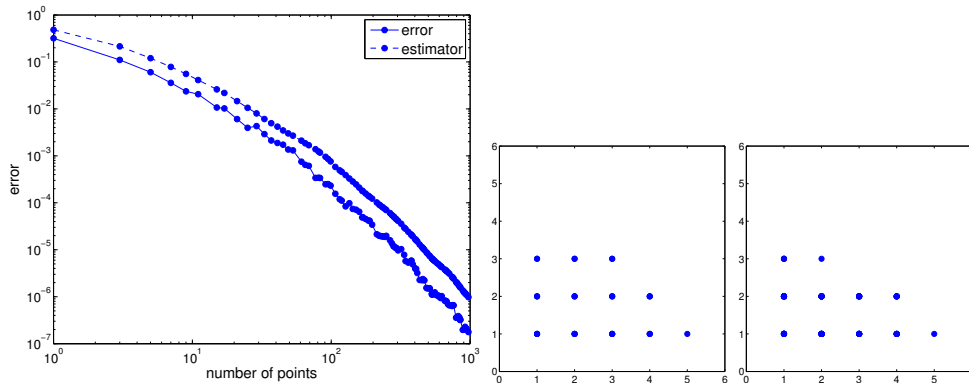


FIG. 10. Error and estimator with respect to the number of points in logarithmic scale (left) and projection of the final multi-index set on  $(y_1, y_3)$  (middle) and on  $(y_1, y_5)$  (right) for the case  $N = 5$ .

In both cases  $N = 3$  and  $N = 5$ , the error estimator provides a good control of the error, the overestimation being slightly bigger for  $N = 5$  than for  $N = 3$ . Moreover, due to the decay of the  $a_n$  in  $n^{-2}$ , the random variables  $Y_n$  should have less and less influence as  $n$  increases. The adaptive algorithm is able to capture this feature, as seen, for instance, when projecting the obtained multi-index set over two different directions. From this experiment, together with the numerical results obtained for the inclusion problems, we see that the efficiency of the stochastic error estimator

seems to be linked to the number of random variables. Further investigation should be made in this direction to determine whether this is indeed the case or if the reason lies elsewhere, for instance, in the error due to the approximation of the  $L^\infty_\rho(\Gamma)$  norm.

*Remark 6.2.* For all the numerical examples given above, the selected multi-index at each iteration of Algorithm 1 belongs to  $\mathcal{R}_I$ . In what follows, we consider a 1D example for which the optimal set is not downward closed, as observed in [12]. The goal is then to see whether our adaptive algorithm captures this feature.

**6.3. 1D numerical example.** We consider problem (1) with  $D = (0, 1)$  the unit interval,  $f(x) = 1$ , and  $a(x, \mathbf{Y}(\omega)) = 1 + 0.1Y_1(\omega) + 0.5Y_2(\omega)$ , where  $Y_n \sim \mathcal{U}[-1, 1]$  for  $n = 1, 2$ .

For the FE mesh, we consider a uniform partition of the unit interval with mesh size  $h = 2^{-12}$ ; that is, we discretize  $[0, 1]$  taking the nodes  $x_i = ih$  with  $i = 0, \dots, 2^{12}$ . In Figure 11 we give three different examples for which the selected multi-index belongs to  $\mathcal{M}_I \setminus \mathcal{R}_I$ . For such a multi-index, more than one element is added to  $I$  because of the constraint that  $I$  remains downward closed during the adaptive process. If we set the tolerance to  $Tol = 10^{-8}$  in Algorithm 1, the adaptive process stops after 16 iterations and the sparse grid contains 153 points in  $\Gamma$ . Moreover, the corresponding error and estimator are about  $2.4882\text{e-}9$  and  $4.2305\text{e-}9$ , respectively, while the error in the energy norm is about  $2.0389\text{e-}9$ .

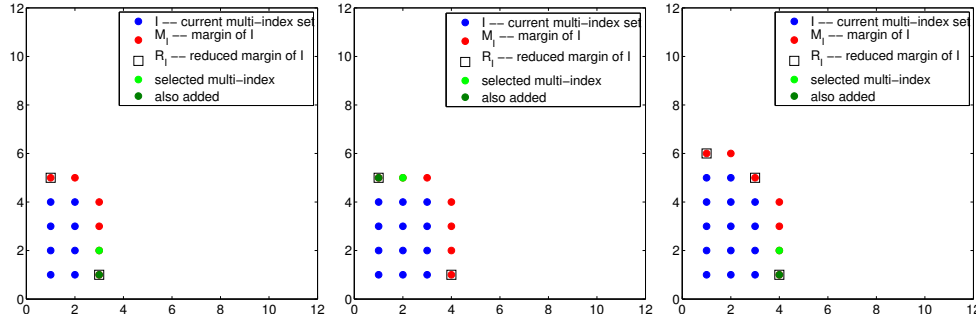


FIG. 11. Three examples for which the selected multi-index belongs to  $\mathcal{M}_I \setminus \mathcal{R}_I$  which correspond to iterations 8 (left), 11 (middle), and 12 (right).

**6.4. Comparison with a heuristic error indicator.** Finally, we compare our error estimator with the heuristic error indicator

$$(35) \quad \Delta_{\mathcal{M}_I} := \sum_{\mathbf{i} \in \mathcal{M}_I} \Delta_{\mathbf{i}} := \sum_{\mathbf{i} \in \mathcal{M}_I} \|\Delta^{\mathbf{m}(\mathbf{i})}(u_h)\|_{L^p_\rho(\Gamma; V)}$$

that is based on the relation  $\|u_h - S_I[u_h]\|_{L^p_\rho(\Gamma; V)} \leq \sum_{\mathbf{i} \notin I} \|\Delta^{\mathbf{m}(\mathbf{i})}(u_h)\|_{L^p_\rho(\Gamma; V)}$ . Notice that we use the FE approximation  $u_h$  of  $u$  since the FE error is not taken into account here. This error indicator is used, for instance, in [17, 20], although its definition is restricted to the reduced margin. We can then consider the adaptive strategy of Algorithm 1 with  $\zeta_{\mathbf{i}}$  replaced by  $\Delta_{\mathbf{i}}$  and the appropriate modification in the definition (29) of the profit. Considering the case  $p = \infty$  as above, the computation of  $\Delta_{\mathbf{i}}$  can be done similarly to the procedure described in Remark 6.1. We emphasize that the computation of  $\Delta_{\mathbf{i}}$  requires the solution of additional PDEs; namely, we need to solve (7) for each new collocation point between the sparse grid for  $S_{G \cup \{\mathbf{i}\}}$  and  $S_G$ . In Figure 12 we give the evolution of *estimator*, the residual-based error estimator (30),

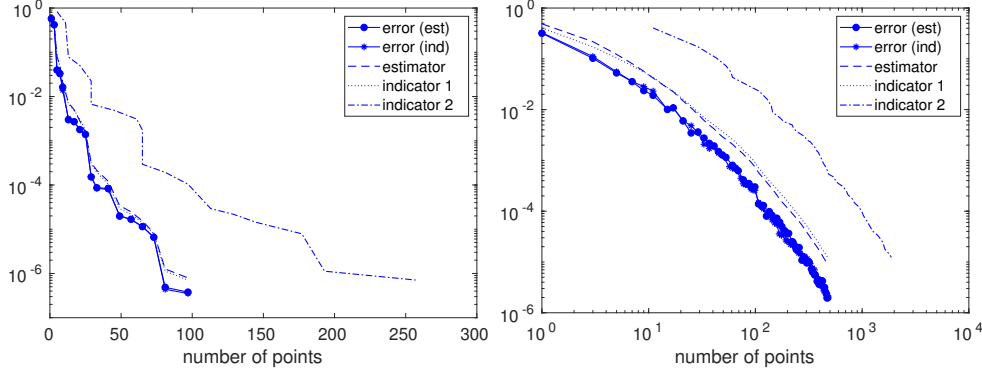


FIG. 12. Adaptive algorithm based on the residual-based estimator and the heuristic error indicator. Case (33) with  $N = 2$  and  $\gamma_1 = \gamma_2 = 1$  in semilogarithmic scale (left), and first 70 iterations for the case (34) with  $N = 5$  in logarithmic scale (right).

and *indicator*, the hierarchical error indicator (35), throughout the adaptive process, each setup being run independently. The error obtained for each case, denoted by *error (est)* and *error (ind)*, respectively, is also provided. The results are plotted with respect to not only the cardinality of the current sparse grid but also the number of collocation points for  $S_{I \cup M_I}$  for the error indicator (see *indicator 2*), i.e., the number of PDE solves needed to compute it. Even if the error indicator is heuristic, it seems to provide a good control of the error, at least in the two cases considered here. Indeed, the error estimator and the error indicator are similar and produce comparable adaptive sparse grids (the same grid for the inclusion problem). Therefore, the *true* error is also similar. However, the error indicator is much more costly to compute, and the higher the dimension, the higher the cost, since additional PDE solves are needed. This last observation is slightly tempered by the fact that these additional solves are not wasted since the approximation  $S_{I \cup M_I}[u_h]$  can be returned at the end of the adaptive process. The error estimator has the advantage of providing a certified control of the error and is cheaper to compute. In its version provided here, it presents the drawback of requiring interpolatory sparse grid approximation, and, contrary to the error indicator (35), it is problem dependent.

**7. Dimension adaptive algorithm.** We provide here some details on how we could proceed to deal with the case where the number  $N$  of random variables in the system is large, possibly infinite. For such a problem, the cardinality of the margin of the multi-index set  $I$  becomes large, and the computation of the error estimator  $\zeta_{SC}$  is no longer feasible. The idea is then to activate initially only a finite (small) number of directions  $y_n$ —the most important ones—as proposed in [20]. The error committed by neglecting some directions has then to be appropriately estimated. A first step in this direction is proposed below. Let us rewrite the diffusion coefficient  $a$  defined in (3) as  $a_N$  to highlight its dependence on the  $N$  random variables  $Y_n$ ,  $n = 1, \dots, N$ . Similarly, we write  $u_N$  for the solution of the diffusion problem with diffusion coefficient  $a_N$ . For  $1 \leq M \leq N$ , let  $u_M$  be the solution of the diffusion problem with coefficient  $a_M(x, \mathbf{y}) = a_0(x) + \sum_{n=1}^M a_n(x)y_n$ . The goal is to estimate the error  $u_N - S_{I_M}[u_{M,h}]$ , where  $u_{M,h}$  is the FE approximation of  $u_M$  and  $S_{I_M}$  is the sparse grid interpolant based on  $I_M \subset \mathbb{N}_+^M$ , i.e., with  $M$  *active variables*. This is the content of the following proposition.

PROPOSITION 7.1. *There exists a constant  $C > 0$  depending only on the mesh aspect ratio  $c$  such that for any  $p \in [1, \infty]$ , we have*

$$(36) \quad \|u_N - S_{IM}[u_{M,h}]\|_{L^p_\rho(\Gamma;V)} \leq \frac{1}{c_{min}^2} [C\eta_{FE,M} + \zeta_{SC,M} + \theta],$$

with  $\eta_{FE,M}$  and  $\zeta_{SC,M}$  defined as in (16)–(18) upon replacing  $a$ ,  $S_I$ , and  $u_h$  by  $a_M$ ,  $S_{IM}$ , and  $u_{M,h}$ , respectively, and

$$(37) \quad \theta := \|(a_N - a_M)\nabla S_{IM}[u_{M,h}]\|_{L^p_\rho(\Gamma;L^2(D))}.$$

*Proof.* We can easily show that for any  $v \in H_0^1(D)$  and a.s. in  $\Omega$ , we have

$$(38) \quad \int_D a_N \nabla(u_N - S_{IM}[u_{M,h}]) \cdot \nabla v = A_1 + A_2 + A_3,$$

with

$$(39) \quad A_1 := S_{IM} \left[ \int_D f v - \int_D a_M \nabla u_{M,h} \cdot \nabla v \right],$$

$$(40) \quad A_2 := - \int_D \sum_{i \in \mathcal{M}_{IM}} \Delta^{m(i)} (a_M \nabla S_{IM}[u_{M,h}]) \cdot \nabla v,$$

$$(41) \quad A_3 := - \int_D (a_N - a_M) \nabla S_{IM}[u_{M,h}] \cdot \nabla v.$$

Indeed, we have

$$\begin{aligned} \int_D a_N \nabla(u_N - S_{IM}[u_{M,h}]) \cdot \nabla v &= \int_D f v - \int_D a_N \nabla S_{IM}[u_{M,h}] \cdot \nabla v \\ &= \int_D f v - \int_D a_M \nabla S_{IM}[u_{M,h}] \cdot \nabla v - \int_D (a_N - a_M) \nabla S_{IM}[u_{M,h}] \cdot \nabla v, \end{aligned}$$

and the first two terms of the right-hand side can be split into  $A_1$  and  $A_2$ , defined in (39) and (40), respectively, proceeding exactly as in the proof of Proposition 4.3. Then, these two terms  $A_1$  and  $A_2$ , which correspond to the errors due to the FEM and the SC method, respectively, can be estimated by proceeding exactly as in section 4. For the term  $A_3$ , which corresponds to the error due to neglecting some directions, we simply apply the Cauchy–Schwarz inequality. Plugging these estimations into (38) with  $v = u_N - S_{IM}[u_{M,h}]$ , we finally obtain (36) using the norm equivalence (4) and taking the  $L^p_\rho(\Gamma)$  norm on both sides.  $\square$

The term  $\theta$  defined in (37), which controls the neglected directions, can be further estimated. Considering the case  $p = 2$  and denoting by  $\sigma_n^2$  the variance of  $Y_n$ ,  $n = M+1, \dots, N$ , we get

$$\theta^2 = \sum_{n=M+1}^N \theta_n^2 \quad \text{with} \quad \theta_n^2 := \sigma_n^2 \|a_n \nabla S_{IM}[u_{M,h}]\|_{L^2_\rho(\Gamma;L^2(D))}^2,$$

where we have used the independence of the random variables and the fact that  $S_{IM}[u_{M,h}]$  depends only on  $y_n$ ,  $n = 1, \dots, M$ . Since the computation of  $\theta$  might still be prohibitive if  $N \gg M$ , we can further decompose it into two parts as

$$\theta^2 = \sum_{n=M+1}^{M+Q} \theta_n^2 + \sum_{n=M+Q+1}^N \theta_n^2 \leq \theta_{Q,M}^2 + \theta_{tail}^2,$$

with

(42)

$$\theta_{Q,M}^2 := \sum_{n=M+1}^{M+Q} \theta_n^2 \quad \text{and} \quad \theta_{tail}^2 := \|\nabla S_{I_M}[u_{M,h}]\|_{L_\rho^2(\Gamma; L^2(D))}^2 \sum_{n=M+Q+1}^N \sigma_n^2 \|a_n\|_{L^\infty(D)}^2$$

for some  $1 \leq Q \leq N - M$ , and compute  $\theta_n$  only for the first  $Q$  terms, i.e., for  $n = M + 1, \dots, M + Q$ . For the choice of  $Q$ , we can, for instance, fix it to some prescribed value. Another possibility is to choose  $Q$  such that  $\theta_{Q,M}^2 \approx \theta_{tail}^2$ , i.e., such that the two terms are balanced. In the spirit of Algorithm 1, a possible adaptive strategy for high-dimensional problems is provided in Algorithm 2, where the same notation as in section 5 is used and  $\theta_{Q,M}$  is defined in (42).

---

**Algorithm 2** Dimension adaptive algorithm (stochastic space adaptation).

---

**Require:**  $Tol > 0$ ,  $M_{init}$

**Ensure:** multi-index set  $I_M \subset \mathbb{N}_+^M$  such that  $\zeta_{SC,M} + \theta \leq Tol$

- 1:  $M = M_{init}$ ,  $I_M = \{\mathbf{1}\} \subset \mathbb{N}_+^M$ ,  $u_{h,I_M} = S_{I_M}[u_{M,h}]$
  - 2: compute  $\zeta_{SC,M}$  and  $\theta_{Q,M}$  with  $Q = \min\{\hat{Q} : \theta_{\hat{Q},M} \geq \theta_{tail}\}$
  - 3: **while**  $\zeta_{SC,M} + \sqrt{2}\theta_{Q,M} > Tol$  **do**
  - 4:   **if**  $\zeta_{SC,M} \geq \sqrt{2}\theta_{Q,M}$  **then**
  - 5:      $\mathbf{i}^* = \operatorname{argmax}_{\mathbf{i} \in \mathcal{M}_{I_M}} P_{\mathbf{i}}$
  - 6:      $I_M \leftarrow I_M \cup A_{\mathbf{i}^*}$
  - 7:      $u_{h,I_M} = S_{I_M}[u_{M,h}]$
  - 8:   **else**
  - 9:      $m^* = \operatorname{argmax}_{M+1 \leq n \leq M+Q} \theta_n^2$
  - 10:    extend  $I_M$  and  $\mathcal{M}_{I_M}$  by adding the direction  $y_{m^*}$  (does not affect  $u_{h,I_M}$ )
  - 11:     $M \leftarrow M + 1$  (and relabeling of the neglected directions if needed)
  - 12:   **end if**
  - 13:   compute  $\zeta_{SC,M}$  and  $\theta_{Q,M}$  with  $Q = \min\{\hat{Q} : \theta_{\hat{Q},M} \geq \theta_{tail}\}$
  - 14: **end while**
- 

**8. Conclusions.** In this work we have derived a residual-based a posteriori error estimate for the SC-FEM, focusing on an elliptic model problem with a random diffusion coefficient. Our error estimate is valid under the assumptions that the diffusion coefficient depends affinely on the random variables and that the sparse grid approximation is interpolatory, which requires the use of nested points. The error estimate, which provides an upper bound of the total error, is constituted of two parts accounting for the FE error and the SC error, respectively. We have then used the SC estimator to drive an adaptive strategy in which the multi-index set characterizing the sparse grid is constructed step by step. We assign a profit to each element of the margin of the set and, at each iteration, select the most profitable one to enter the set. We have provided several numerical examples of moderate dimension to illustrate the theoretical findings. More precisely, we have compared the error and the estimator for various given multi-index sets and have then tested the efficiency of the proposed adaptive algorithm. This algorithm, which uses the SC estimator to drive the adaptive process, is one possible strategy. Several other versions could be considered as well, for instance, by selecting more than one multi-index at each iteration using a so-called Dörfler or maximum marking strategy with, ideally, a proof of convergence. In the case of high-dimensional problems, that is, when the coefficient depends on

(possibly infinitely) many random variables, the computation of the profit of each element of the margin becomes prohibitive, and an alternative should be used. We have given some insight in this direction and proposed a possible adaptive strategy; however, further investigations, including numerical experiments, should be done.

## REFERENCES

- [1] 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 (2007), pp. 1005–1034, <https://doi.org/10.1137/050645142>.
- [2] F. NOBILE, R. TEMPONE, AND C. G. WEBSTER, *A sparse grid stochastic collocation method for partial differential equations with random input data*, SIAM J. Numer. Anal., 46 (2008), pp. 2309–2345, <https://doi.org/10.1137/060663660>.
- [3] D. XIU AND J. S. HESTHAVEN, *High-order collocation methods for differential equations with random inputs*, SIAM J. Sci. Comput., 27 (2005), pp. 1118–1139, <https://doi.org/10.1137/040615201>.
- [4] G. S. FISHMAN, *Monte Carlo: Concepts, Algorithms, and Applications*, Springer Ser. Oper. Res. Financ. Eng., Springer, New York, 1996.
- [5] J. DICK AND F. PILLICHSHAMMER, *Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration*, Cambridge University Press, Cambridge, UK, 2010.
- [6] J. DICK, F. Y. KUO, AND I. H. SLOAN, *High-dimensional integration: The quasi-Monte Carlo way*, Acta Numer., 22 (2013), pp. 133–288.
- [7] M. B. GILES, *Multi-level Monte Carlo path simulation*, Oper. Res., 56 (2008), pp. 607–617.
- [8] R. G. GHANEM AND P. D. SPANOS, *Stochastic Finite Elements: A Spectral Approach*, Springer, New York, 1991.
- [9] O. P. LE MAÎTRE AND O. M. KNIO, *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*, Sci. Comput., Springer Netherlands, 2010.
- [10] D. GUIGNARD, F. NOBILE, AND M. PICASSO, *A posteriori error estimation for elliptic partial differential equations with small uncertainties*, Numer. Methods Partial Differential Equations, 32 (2016), pp. 175–212.
- [11] D. GUIGNARD, F. NOBILE, AND M. PICASSO, *A posteriori error estimation for the steady Navier-Stokes equations in random domains*, Comput. Methods Appl. Mech. Engrg., 313 (2017), pp. 483–511.
- [12] J. BECK, F. NOBILE, L. TAMELLINI, AND R. TEMPONE, *On the optimal polynomial approximation of stochastic PDEs by Galerkin and collocation methods*, Math. Models Methods Appl. Sci., 22 (2012), 1250023.
- [13] A. CHKIFA, A. COHEN, AND C. SCHWAB, *High-dimensional adaptive sparse polynomial interpolation and applications to parametric PDEs*, Found. Comput. Math., (2013), pp. 1–33.
- [14] R. C. ALMEIDA AND J. T. ODEN, *Solution verification, goal-oriented adaptive methods for stochastic advection–diffusion problems*, Comput. Methods Appl. Mech. Engrg., 199 (2010), pp. 2472–2486.
- [15] 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 (2008), pp. 2411–2442, <https://doi.org/10.1137/070680540>.
- [16] F. NOBILE, L. TAMELLINI, AND R. TEMPONE, *Convergence of quasi-optimal sparse grid approximation of Hilbert-valued functions: Application to random elliptic PDEs*, Numer. Math., 134 (2016), pp. 343–388.
- [17] T. GERSTNER AND M. GRIEBEL, *Dimension-adaptive tensor-product quadrature*, Computing, 71 (2003), pp. 65–87.
- [18] H. J. BUNGARTZ AND M. GRIEBEL, *Sparse grids*, Acta Numer., 13 (2004), pp. 147–269.
- [19] A. KLIMKE, *Uncertainty Modeling Using Fuzzy Arithmetic and Sparse Grids*, Ph.D. thesis, Universität Stuttgart, Stuttgart, Germany, 2006.
- [20] F. NOBILE, L. TAMELLINI, F. TESEI, AND R. TEMPONE, *An adaptive sparse grid algorithm for elliptic PDEs with lognormal diffusion coefficient*, in Sparse Grids and Applications—Stuttgart 2014, J. Garcke and D. Pflüger, eds., Springer International Publishing, 2016, pp. 191–220.
- [21] M. GRIEBEL AND S. KNAPEK, *Optimized general sparse grid approximation spaces for operator equations*, Math. Comp., 78 (2009), pp. 2223–2257.
- [22] C. SCHILLINGS AND C. SCHWAB, *Sparse, adaptive Smolyak quadratures for Bayesian inverse problems*, Inverse Probl., 29 (2013), 065011.

- [23] F. NOBILE AND F. TESEI, *A multi level Monte Carlo method with control variate for elliptic PDEs with log-normal coefficients*, Stoch. Partial Differ. Equ. Anal. Comput., 3 (2015), pp. 398–444.
- [24] A. CHKIFA, A. COHEN, R. DEVORE, AND C. SCHWAB, *Sparse adaptive Taylor approximation algorithms for parametric and stochastic elliptic PDEs*, ESAIM Math. Model. Numer. Anal., 47 (2013), pp. 253–280.
- [25] M. EIGEL, C. J. GITTELSON, C. SCHWAB, AND E. ZANDER, *Adaptive stochastic Galerkin FEM*, Comput. Methods Appl. Mech. Engrg., 270 (2014), pp. 247–269.
- [26] M. EIGEL, C. J. GITTELSON, C. SCHWAB, AND E. ZANDER, *A convergent adaptive stochastic Galerkin finite element method with quasi-optimal spatial meshes*, ESAIM Math. Model. Numer. Anal., 49 (2015), pp. 1367–1398.
- [27] J. M. CASCON, C. KREUZER, R. H. NOCHETTO, AND K. G. SIEBERT, *Quasi-optimal convergence rate for an adaptive finite element method*, SIAM J. Numer. Anal., 46 (2008), pp. 2524–2550, <https://doi.org/10.1137/07069047X>.
- [28] S. A. SMOLYAK, *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Soviet Math. Dokl., 4 (1963), pp. 240–243. [Russian original in Dokl. Akad. Nauk SSSR, 148 (1963), pp. 1042–1045.]
- [29] G. W. WASILKOWSKI AND H. WOZNAKOWSKI, *Explicit cost bounds of algorithms for multivariate tensor product problems*, J. Complexity, 11 (1995), pp. 1–56.
- [30] V. BARTHELMANN, E. NOVAK, AND K. RITTER, *High dimensional polynomial interpolation on sparse grids*, Adv. Comput. Math., 12 (2000), pp. 273–288.
- [31] J. BÄCK, F. NOBILE, L. TAMELLINI, AND R. TEMPONE, *Stochastic spectral Galerkin and collocation methods for PDEs with random coefficients: A numerical comparison*, in Spectral and High Order Methods for Partial Differential Equations: Selected Papers from the ICOSAHOM '09 Conference (June 22–26, 2009, Trondheim, Norway), S. J. Hesthaven and M. E. Rønquist, eds., Lect. Notes Comput. Sci. Eng. 76, Springer, Berlin, 2011, pp. 43–62.
- [32] P. CLÉMENT, *Approximation by finite element functions using local regularization*, RAIRO Anal. Numér., 9 (1975), pp. 77–84.
- [33] A. CHKIFA, *Méthodes polynomiales parcimonieuses en grande dimension. Application aux EDP paramétriques*, Ph.D. thesis, Laboratoire Jacques Louis Lions, Paris, 2014.
- [34] F. LEJA, *Sur certaines suites liées aux ensembles plans et leur application à la représentation conforme*, Ann. Polon. Math., 4 (1957), pp. 8–13.