

CONVERGENCE OF ADAPTIVE STOCHASTIC COLLOCATION WITH FINITE ELEMENTS

MICHAEL FEISCHL AND ANDREA SCAGLIONI

ABSTRACT. We consider an elliptic partial differential equation with a random diffusion parameter discretized by a stochastic collocation method in the parameter domain and a finite element method in the spatial domain. We prove for the first time convergence of a stochastic collocation algorithm which adaptively enriches the parameter space as well as refines the finite element meshes.

1. INTRODUCTION

Partial differential equations with random data are a ubiquitous tool in the modeling of real life phenomena such as structural vibrations [21], groundwater flow [31], and composite material behavior [1]. The efficient approximation of solutions of those equations is a challenging problem as it requires the approximation of high-dimensional functions in a parameter domains as well as low-dimensional but in general non-regular functions in the spatial domain. While effective ways to generate the random data have been studied in [23, 34], we focus on the numerical approximation of the resulting solution of the PDE.

To that end, we consider an adaptive stochastic collocation algorithm for a random diffusion problem proposed in [32] and extend it to include spatial mesh refinement for a finite element method. We give the first proof of convergence of the adaptive algorithm to the exact solution and even derive some convergence rates as well as optimality statements. The main difficulty to overcome is the interplay of parametric enrichment and finite element refinement to ensure overall convergence.

Stochastic collocation is a so-called non-intrusive method, which has the big advantage that it does not require new solver algorithms, but reuses deterministic solvers only. Roughly speaking, the exact solution depends on a parametric variable (the random input) and a spatial variable. While the spatial dependence is resolved by standard finite element approximation, the parametric dependence is discretized by collocation. For each collocation point, we only need to solve a deterministic problem and therefore can reuse well tested finite element codes.

Problems of this kind have been considered in many prior works. See, e.g., [37] for the (apparent) first appearance of the term *stochastic collocation* in the field of computational fluid dynamics. The authors combine a polynomial chaos expansion with a collocation solver. This is generalized and formalized in [48], where different strategies and collocation points are discussed, and error analysis is provided. In [2], rigorous error analysis is developed for the discretization of a random PDE with a combination of standard finite elements and Gaussian collocation points (and also sparse grid points). In [39], dimension dependent anisotropy of the solution is addressed by use of anisotropic sparse grids and in [5], the optimal choice of the sparse grid parameters (the multi-index set) is discussed.

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 258734477 – SFB 1173 and the Austrian Science Fund (FWF) the SFB Taming complexity in partial differential systems (grant SFB F65)

Institute of Analysis and Scientific Computing TU Wien, Wiedner Hauptstraße 8-10, 1040 Vienna.

More recent works, which technically do not use stochastic collocation but still deal with similar mathematical techniques and problems, include [15, 12] for quasi-Monte Carlo sampling approaches, [14, 16] for multi-level methods, and [13] for a multi-index method. Those methods have in common that they do not recover the full probability distribution of the exact solution but only compute a certain quantity of interest of it, e.g., the expectation or higher moments.

Another important and mathematically well-understood branch of methods for PDEs with random parameters are so-called stochastic Galerkin methods. Belonging to the class of intrusive methods, they employ a Galerkin method in both the spatial and the parametric domain. Algorithms have been proposed in [29, 24, 28, 38, 49].

The question of whether intrusive or non-intrusive algorithms are more advantageous is a fundamental one and depends on the specific problem setting and goals of the user. An overview is given in [30].

When dealing with PDEs with random data, adaptivity comes into mind for two reasons: First, spatial adaptivity is necessary to resolve singularities originating from geometric features (e.g., concave corners) and from irregular coefficients induced by the random input. Uniform meshes suffer from drastic reduction of convergence rate in the presence of such singularities, see, e.g., [7] for an exhaustive overview on h -adaptive methods.

Second, parametric adaptivity is necessary to resolve anisotropies in the random coefficient. The random input is often parameterized on high-dimensional parameter domains, and usually not all directions of that domain are equally important. Therefore, a straightforward tensor approximation approach would suffer dramatically from the curse of dimensionality. Here, an adaptive approach can outperform uniform methods significantly, see [10, 11] for an overview.

Adaptive approximation of high-dimensional parameter domains has a long history. A starting point is often the work by Kolmogorov [35], in which the decomposition of a high-dimensional function into a sum of lower dimensional terms is discussed. Similar ideas have been pursued in statistics, we mention, e.g., so-called additive models [33], *Multivariate adaptive regression splines* [25], or the *ANOVA* decomposition [46, 50]. One of the first dimension adaptive algorithms for sparse grids can be found in [27]. The method uses an error estimator that, in practice, agrees very well with the actual error but does not constitute a rigorous upper bound (it may underestimate the error by an arbitrarily large factor in exotic cases).

For PDEs with random data, adaptive stochastic Galerkin algorithms have been investigated in [6, 20] with convergence and even optimality proofs. Even low-rank tensor formats have been used in [22, 17] to speed up the computation of stochastic Galerkin matrix and [17] shows that in certain cases, the low-rank approximation of the full tensor product approximation can be stored and manipulated faster than some adaptive sparse approximations.

An adaptive sparse grid collocation algorithm based on a reliable error estimator was proposed in [32]. The work uses a sparse grid interpolation operator to discretize the parametric domain and proposes an error estimator which consists of a parametric estimator as well as a finite element estimator.

A couple of recent works deal with similar approaches. A non-adaptive but true multi-level collocation method is proposed in [45]. The multi-level aspect allows the method to treat high-fidelity finite element approximations with low-fidelity parametric approximation and vice versa. This further reduces the impact of the curse of dimensionality and results in a very efficient algorithm. Such an approach could be built on top of the method proposed in the present manuscript in order to reduce the cost. An adaptive version of the multi-level algorithm was recently proposed in [36]. While the algorithm is adaptive in both the parameter and spatial domain, the authors use an error estimator from [27] that is not an upper bound for the error and hence can not prove convergence

without extra assumptions. Finally, the very recent but independent work [19] analyzes the same parameter adaptive algorithm as in this work and proves convergence. However, they do not consider spatial adaptive refinement of the finite element meshes. As we show in numerical experiments in Section 4, spatial adaptivity clearly improves overall performance and hence should be included in any adaptive algorithm (however, a careful choice of the adaptive strategy is important as minor variations can lead to significant performance differences).

The remainder of this work is organized as follows: We present the model problem in Section 1.1 and describe the adaptive algorithm in Section 1.3. In Section 2, we prove convergence of the adaptive algorithm for the pure parameter enrichment problem (i.e., the problem considered in [32]), and Section 3 proves the convergence of the full adaptive algorithm including spatial adaptivity (with one adaptive mesh per collocation point or one global adaptive mesh). Section 4 presents some numerical experiments and in the final section, we draw some conclusions.

1.1. Problem statement. Consider an integer $d \geq 2$ and an open bounded domain $D \subset \mathbb{R}^d$ with Lipschitz continuous boundary ∂D . Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a complete probability space. Let $Y_n : \Omega \rightarrow \mathbb{R}$ be independent random variables with ranges $\Gamma_n := Y_n(\Omega)$ and densities $\rho_n : \Gamma_n \rightarrow \mathbb{R}_{\geq 0}$ for all $n \in 1, \dots, N$. In the present work, we assume that the ranges Γ_n are *bounded* subsets of \mathbb{R} . Let $\Gamma := \bigotimes_{n=1}^N \Gamma_n \subset \mathbb{R}^N$ and $\rho := \bigotimes_{n=1}^N \rho_n$. The triple $(\Gamma, \mathcal{B}(\Gamma), \rho(\mathbf{y})d\mathbf{y})$ ($\mathcal{B}(\Gamma)$ the Borel σ -algebra on Γ) is a probability space. Consider $f \in L^2(D)$ and $a : \Gamma \times D \rightarrow \mathbb{R}$ with the following properties:

- uniform boundedness

$$\exists a_{\min}, a_{\max} \in \mathbb{R}_{>0} : a_{\min} \leq a(\mathbf{y}, x) \leq a_{\max} \quad \rho\text{-a.e. } \mathbf{y} \in \Gamma, \forall x \in D$$

- affine dependence on $\mathbf{y} \in \Gamma$

$$\forall n \in 0, \dots, N \exists a_n : D \rightarrow \mathbb{R} : a(\mathbf{y}, x) = a_0(x) + \sum_{n=1}^N a_n(x)y_n$$

- regularity in space $\nabla a(\mathbf{y}, \cdot)|_T \in L^\infty(T)$ for all elements T of a coarse *initial* mesh $\mathcal{T}_{\text{init}}$ of D .

We consider the parametric weak formulation of the Poisson problem: Find $u : \Gamma \rightarrow V$ such that

$$(1) \quad \int_D a(x, \mathbf{y}) \nabla u(x, \mathbf{y}) \cdot \nabla v(x) dx = \int_D f(x) v(x) dx \quad \forall v \in V, \rho\text{-a.e. } \mathbf{y} \in \Gamma.$$

Here, V denotes the Sobolev space $H_0^1(D)$ with the norm $\|v\|_V := \|\nabla v\|_{L^2(D)}$.

Due to uniform ellipticity of the problem the exact solution is unique and (see also, e.g., [2, Lemma 3.1]) there exists $\tau \subset \mathbb{R}_{>0}^N$ such that $u : \Gamma \rightarrow V$ can be extended to a bounded holomorphic function on the set

$$(2) \quad \Sigma(\Gamma, \tau) := \{z \in \mathbb{C}^N : \text{dist}(z_n, \Gamma_n) \leq \tau_n \forall n = 1, \dots, N\}.$$

1.2. The sparse grid stochastic collocation interpolant. We aim at building a discretization of the solution u of (1) in the space

$$(3) \quad \mathbb{P}(\Gamma, W) \cong \mathbb{P}(\Gamma) \otimes W,$$

where $\mathbb{P}(\Gamma)$ is a finite-dimensional polynomial space on Γ and W is a finite-dimensional subspace of V . In order to do so, we fix a set \mathcal{H} of distinct collocation points in Γ and denote by $\{L_{\mathbf{y}}\}_{\mathbf{y} \in \mathcal{H}}$ the related set of Lagrange basis functions (i.e. the unique set of polynomials of minimal degree over Γ such that $L_{\mathbf{z}}(\mathbf{y}) = \delta_{\mathbf{y}, \mathbf{z}}$ for any $\mathbf{y}, \mathbf{z} \in \mathcal{H}$). By $\mathbb{P}(\Gamma)$, we denote the polynomial space spanned by $\{L_{\mathbf{y}}\}_{\mathbf{y} \in \mathcal{H}}$. For any $\mathbf{y} \in \mathcal{H}$, we consider $\mathcal{T}_{\mathbf{y}}$, a shape-regular triangulation on D depending on \mathbf{y} , and $V_{\mathbf{y}} := S_0^1(\mathcal{T}_{\mathbf{y}})$, the classical finite elements space of piecewise-linear functions over $\mathcal{T}_{\mathbf{y}}$ with

zero boundary conditions. We denote by $U_{\mathbf{y}} \in V_{\mathbf{y}}$ the finite element solution of the problem for the parameter \mathbf{y} , i.e.,

$$(4a) \quad \int_D a(x, \mathbf{y}) \nabla U_{\mathbf{y}}(x) \cdot \nabla v(x) dx = \int_D f(x) v(x) dx \quad \forall v \in V_{\mathbf{y}}.$$

Finally, the discretization of u takes the form

$$(4b) \quad u_{\mathcal{H}}(x, \mathbf{z}) = \sum_{\mathbf{y} \in \mathcal{H}} U_{\mathbf{y}}(x) L_{\mathbf{y}}(\mathbf{z}).$$

The number of degrees of freedom of $u_{\mathcal{H}}$ is $\sum_{\mathbf{y} \in \mathcal{H}} \dim(V_{\mathbf{y}})$. The space W from (3) will be the smallest subspace of V that contains each of the finite element spaces $\{V_{\mathbf{y}}\}_{\mathbf{y} \in \mathcal{H}}$. The set of collocation nodes and the polynomial space are defined following the sparse grid construction, which we now describe briefly. We start by considering a family of 1D nodes, i.e. a set $\mathcal{Y}^n := \left\{ y_j^{(n)} \right\}_{j=1}^n \subset \mathbb{R}$ defined for any positive integer n . We require the family of \mathcal{Y}^n to be nested, i.e. $\mathcal{Y}^n \subset \mathcal{Y}^{n+1}$ for any $n \in \mathbb{N}$. The particular number of the quadrature nodes used in the algorithm is encoded in the function $m(\cdot): \mathbb{N} \rightarrow \mathbb{N}$. Finally, let $I \subset \mathbb{N}^N$ be a *downward-closed* multi-index set, i.e.,

$$\forall \mathbf{i} \in I, \quad \mathbf{i} - \mathbf{e}_n \in I \quad \forall n = 1, \dots, N \text{ such that } i_n > 1.$$

with \mathbf{e}_n the n -th unit vector in \mathbb{N}^N . The *sparse grid interpolant* of a function $v \in C^0(\Gamma, V)$ is

$$(5) \quad S_I[v](\mathbf{y}) := \sum_{\mathbf{i} \in I} \Delta^{m(\mathbf{i})}(v)(\mathbf{y}),$$

where the *hierarchical surplus* operator is defined as $\Delta^{m(\mathbf{i})} := \bigotimes_{n=1}^N \Delta^{m(i_n)}$, the *detail operator* is defined as $\Delta^{m(i_n)} := \mathcal{U}_n^{m(i_n)} - \mathcal{U}_n^{m(i_n-1)}$ and $\mathcal{U}_n^{m(i_n)}: C^0(\Gamma_n) \rightarrow \mathbb{P}_{m(i_n)-1}(\Gamma_n)$ is the Lagrange interpolant with respect to the nodes $\mathcal{Y}^{m(i_n)} \subset \Gamma_n$. Finally, we set $\mathcal{U}_n^0 \equiv 0$ for all $n \in 1, \dots, N$.

The polynomial space $\mathbb{P}(\Gamma)$ introduced in (3) corresponds to

$$\mathbb{P}_I(\Gamma) := \sum_{\mathbf{i} \in I} \mathbb{P}_{m(\mathbf{i})-1}(\Gamma) \quad \text{where} \quad \mathbb{P}_{m(\mathbf{i})-1}(\Gamma) := \bigotimes_{n=1}^N \mathbb{P}_{m(i_n)-1}(\Gamma_n).$$

The sparse grid stochastic collocation interpolant can be written as a linear combination of tensor product Lagrange interpolants (see, for instance, [47]):

$$(6) \quad S_I[u](\mathbf{y}) = \sum_{\mathbf{i} \in I} c_{\mathbf{i}} \bigotimes_{n=1}^N \mathcal{U}_n^{m(i_n)}(u)(\mathbf{y}), \quad c_{\mathbf{i}} := \sum_{\substack{j \in \{0,1\}^N \\ \mathbf{i}+j \in I}} (-1)^{|j|_1}.$$

The set of collocation points \mathcal{H} in (6) and also in (4) is referred to as *sparse grid* and we will also denote it by \mathcal{H}_I in order to make the dependence on I explicit. The *nestedness* of the family of 1D nodes \mathcal{Y}^n makes $S_I[\cdot]$ interpolatory in the collocation nodes (see [4, proposition 6])

$$S_I[u](\mathbf{y}) = u(\mathbf{y}) \quad \forall \mathbf{y} \in \mathcal{H}_I.$$

Due to this fact, (4) can be rewritten as

$$(7) \quad u_{\mathcal{H}}(x, \mathbf{z}) = u_I(x, \mathbf{z}) = S_I[u](x, \mathbf{z}) = \sum_{\mathbf{y} \in \mathcal{H}_I} U_{\mathbf{y}}(x) L_{\mathbf{y}}(\mathbf{z}) \quad x \in D, \mathbf{z} \in \Gamma.$$

The nestedness is satisfied, e.g., by choosing *Clenshaw-Curtis (CC)* nodes to construct the sparse grid, i.e.

$$y_j^{(m)} := -\cos \frac{\pi(j-1)}{m-1} \quad \forall j = 1, \dots, m,$$

with the *doubling rule*

$$(8) \quad m(i) := \begin{cases} 0 & i = 0, \\ 1 & i = 1, \\ 2^{i-1} + 1 & i > 1. \end{cases}$$

We will stick with this particular choice for the remainder of this work for simplicity, remark however that other choices are possible (see, e.g., [32]). The essential properties of \mathcal{Y}^n used in the proofs below are nestedness $\mathcal{Y}^n \subseteq \mathcal{Y}^{n+1}$ and the fact that the Lebesgue constants of the associated interpolation operators grow sub-exponentially.

The requirement on the multi-index set I to be downward-closed is needed to ensure that the sum (5) is actually telescopic.

Since u is analytic in \mathbf{y} , we may consider the expansion (see again [32])

$$(9) \quad u(\mathbf{y}) = \sum_{\mathbf{i} \in \mathbb{N}^N} \Delta^{m(\mathbf{i})} u(\mathbf{y}) \quad \text{a.e. } \mathbf{y} \in \Gamma$$

converging absolutely in V . As it will be central in the following discussion, we recall the definition of the *margin of a multi-index set I* :

$$\mathcal{M}_I := \{ \mathbf{i} \in \mathbb{N}^N : \mathbf{i} - \mathbf{e}_n \in I \text{ for some } n \in 1, \dots, N \text{ such that } i_n > 1 \}.$$

1.3. The adaptive stochastic collocation finite element algorithm. The adaptive algorithm employs the error estimator proposed in [32, Proposition 4.3]. We recall that u denotes the analytic solution of the problem (1) while the discrete solution is $S_I[U] = \sum_{\mathbf{y} \in \mathcal{H}_I} U_{\mathbf{y}} L_{\mathbf{y}}$. By $U : \Gamma \rightarrow W$, we denote a function that takes the value $U_{\mathbf{y}}$ on the collocation point $\mathbf{y} \in \mathcal{H}_I$ (sometimes we will also use the notation $U(\mathbf{y}) = U_{\mathbf{y}}$).

The total estimator is composed of a *parametric estimator*

$$(10) \quad \zeta_{SC,I} := \sum_{\mathbf{i} \in \mathcal{M}_I} \zeta_{\mathbf{i},I}, \quad \zeta_{\mathbf{i},I} := \left\| \Delta^{m(\mathbf{i})} (a \nabla S_I[U]) \right\|_{L^\infty(\Gamma, L^2(D))}$$

(the gradient ∇ here acts exclusively on the space variable $x \in D$) as well as a *finite element estimator*

$$(11) \quad \eta_{FE,I} := \sum_{\mathbf{y} \in \mathcal{H}_I} \eta_{\mathbf{y}} \|L_{\mathbf{y}}\|_{L^\infty(\Gamma)}, \quad \eta_{\mathbf{y}} := \left(\sum_{T \in \mathcal{T}_{\mathbf{y}}} \eta_{\mathbf{y},T}^2 \right)^{\frac{1}{2}},$$

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

where $[\cdot]_{\mathbf{n}_e}$ denotes the jump over the edge (face) in normal direction \mathbf{n}_e . The combination of both yields a reliable upper bound, i.e.,

$$(12) \quad \|u - S_I[U]\|_{L^\infty(\Gamma, V)} \leq \frac{1}{a_{\min}} (C \eta_{FE,I} + \zeta_{SC,I}),$$

where $a_{\min} > 0$ appears in the equivalence relation between $H_0^1(D)$ and energy norm

$$a_{\min}^{1/2} \|v\|_{H_0^1(D)} \leq \left\| a(\mathbf{y})^{\frac{1}{2}} \nabla v \right\|_{L^2(D)} \leq a_{\max}^{1/2} \|v\|_{H_0^1(D)} \quad \text{a.e. } \mathbf{y} \in \Gamma \text{ for all } v \in H_0^1(D)$$

and $C > 0$ depends only on the shape regularity of $\mathcal{T}_{\text{init}}$. We consider the following adaptive algorithm.

Algorithm 1 $u_\epsilon \leftarrow \text{SCFE}(\epsilon, \theta, \alpha, \mathcal{T}_{\text{init}})$

```

1:  $I_{-1} := \emptyset$ 
2:  $I_0 := \{\mathbf{1}\}$ 
3: compute finite element solution  $U_{0,\mathbf{y}}$  on  $\mathcal{T}_{\text{init}}$  for all  $\mathbf{y} \in \mathcal{H}_{I_0}$ 
4: for  $\ell = 0, 1, 2, \dots$  do
5:    $U_\ell \leftarrow \text{Refine\_FE\_spaces}(I_\ell, U_\ell, \alpha, \theta)$ 
6:   compute parametric estimators  $(\zeta_{i,I_\ell})_{i \in \mathcal{M}_{I_\ell}}, \zeta_{SC,I_\ell}$ 
7:   compute finite element estimator  $\eta_{\text{FE},I_\ell}$ 
8:   if  $a_{\min}^{-1}(\zeta_{SC,I_\ell} + C\eta_{\text{FE},I_\ell}) < \epsilon$  then
9:     return  $u_\epsilon \leftarrow S_{I_\ell}[U_\ell]$ 
10:  end if
11:   $(U_{\ell+1}, I_{\ell+1}) \leftarrow \text{Refine\_parameter\_space}(I_\ell, U_\ell, (\zeta_{i,I_\ell})_{i \in \mathcal{M}_{I_\ell}}, \mathcal{T}_{\text{init}})$ 
12: end for

```

The algorithm consists of alternating between enriching the polynomial space \mathbb{P}_I in Line 11 (Alg. 1) and refining the finite element spaces corresponding to each collocation point *independently from each other* in Line 5 (Alg. 1). The intuitive idea behind this choice is the following: In order for the parameter enrichment routine to make a meaningful choice, the finite element solution in the collocation points has to be "close enough" to the exact solution. The algorithm terminates when the a-posteriori estimator falls below a given tolerance $\epsilon > 0$ in Line 8 (Alg. 1). The reliable upper bound (12) guarantees that the error of the discrete solution is also bounded by ϵ .

The sub-routine `Refine_FE_spaces` reads:

Algorithm 2 $U \leftarrow \text{Refine_FE_spaces}(I, U, \alpha, \theta)$

```

1: compute finite element estimator  $(\eta_{\mathbf{y}})_{\mathbf{y} \in \mathcal{H}_I}, \eta_{\text{FE},I}$ 
2: compute parametric estimator  $\zeta_{SC,I}$ 
3:  $\text{Tol} := \alpha \frac{1}{\sum_{i \in \mathcal{M}_I} \prod_{n=1}^N i_n} \zeta_{SC,I}$ 
4: while  $\eta_{\text{FE},I} > \text{Tol}$  do
5:   find minimal  $\mathcal{K} \subseteq \bigsqcup_{\mathbf{y} \in \mathcal{H}} \mathcal{T}_{\mathbf{y}} := \bigcup_{\mathbf{y} \in \mathcal{H}} \bigcup_{T_{\mathbf{y}} \in \mathcal{T}_{\mathbf{y}}} (\mathbf{y}, T_{\mathbf{y}})$  such that
      
$$\sum_{(\mathbf{y}, T_{\mathbf{y}}) \in \mathcal{K}} \eta_{\mathbf{y}, T_{\mathbf{y}}}^2 \|L_{\mathbf{y}}\|_{L^\infty(\Gamma)} \geq \theta \eta_{\text{FE},I}^2$$

6:   for  $\mathbf{y} \in \mathcal{H}$  do
7:     refine  $\mathcal{T}_{\mathbf{y}}$  with  $\mathcal{K}_{\mathbf{y}} := \{T \in \mathcal{T}_{\mathbf{y}} : (\mathbf{y}, T) \in \mathcal{K}\}$  as marked elements
8:   end for
9:   compute  $U_{\mathbf{y}}$  over  $\mathcal{T}_{\mathbf{y}}$ 
10:  compute finite element estimator  $(\eta_{\mathbf{y}})_{\mathbf{y} \in \mathcal{H}_I}, \eta_{\text{FE},I}$ 
11:  compute parametric estimator  $\zeta_{SC,I}$ 
12:   $\text{Tol} \leftarrow \alpha \frac{1}{\sum_{i \in \mathcal{M}_I} \prod_{n=1}^N i_n} \zeta_{SC,I}$ 
13: end while

```

The aim of this sub-routine is to refine the finite element solutions in the collocation points until the finite element estimator falls below the tolerance defined in Line 3 (Alg. 2). We use newest-vertex-bisection with mesh closure for mesh refinement (see, e.g., [44] for further details). Observe that, since the tolerance depends on the parametric estimator $\zeta_{SC,I}$, which in turn depends

on the discrete solution, the tolerance needs to be re-computed at every finite-element refinement. This is necessary as one parametric refinement might uncover new features of the solution which need to be resolved in the finite-element refinement step. In practical computations, this rarely happens after the first few iterations of the algorithm. Note also that the linear convergence result in Proposition 3.2 shows that starting from the initial mesh does not significantly increase the computational cost (at worst case, it contributes logarithmically). It might be possible to prove convergence without the scaling $1/\sum_{\mathbf{i} \in \mathcal{M}_I} \prod_{n=1}^N i_n$ of the tolerance in Line 3 of Algorithm 1 (the experiments in Section 4 indicate that it is not necessary) however with the present techniques we didn't find a way to do that.

In Section 3 we will prove that the sub-routine terminates (i.e. that the finite element estimator eventually falls below the tolerance) and that the choice of tolerance made in Line 3 (Alg. 2) is a sufficient condition for convergence.

Finally, the sub-routine `Refine_parameter_space` reads as follows:

Algorithm 3 $(U', I') \leftarrow \text{Refine_parameter_space}(I, U, (\zeta_{\mathbf{i}, I})_{\mathbf{i} \in \mathcal{M}_I}, \mathcal{T}_{\text{init}})$

- 1: $\mathbf{i} := \arg \max_{\mathbf{i} \in \mathcal{M}_I} \mathcal{P}_{\mathbf{i}, I}$
 - 2: $I' := I \cup A_{\mathbf{i}, I}$
 - 3: $U' \leftarrow$ update U by computing finite element solution $U_{\mathbf{y}}$ on $\mathcal{T}_{\text{init}}$ for all $\mathbf{y} \in \mathcal{H}_{I'} \setminus \mathcal{H}_I$
-

The aim here is to enrich the polynomial space \mathbb{P}_I as done in [32, Algorithm 1]. At each iteration, the algorithm enlarges the multi-index set I by adding multi-indices from the margin of I depending on the values of the pointwise error estimators $(\zeta_{\mathbf{i}, I})_{\mathbf{i} \in I}$. More precisely, in Line 1 (Alg. 3) we select a *profit maximizer*, i.e. a multi index in the margin that maximizes a given *profit function* $\mathcal{P}_{\mathbf{i}, I}$ (see below for some examples):

$$(13) \quad \mathbf{i} = \arg \max_{\mathbf{i} \in \mathcal{M}_I} \mathcal{P}_{\mathbf{i}, I}$$

(in case more than one multi-index maximizes the profit, we pick the one that comes first in the lexicographic ordering).

Then, in Line 2 (Alg. 3) I is enlarged by adding $A_{\mathbf{i}, I}$, the smallest subset of \mathcal{M}_I containing \mathbf{i} such that $I \cup A_{\mathbf{i}, I}$ is downward-closed. Finally, in Line 3 (Alg. 3) we compute the finite element solution over the default mesh $\mathcal{T}_{\text{init}}$ corresponding to each new collocation point, while preserving the old ones.

We analyze two possible choices of profit:

- Workless profit:

$$(14) \quad \mathcal{P}_{\mathbf{i}, I} := \sum_{\mathbf{j} \in A_{\mathbf{i}, I}} \zeta_{\mathbf{j}, I}$$

- Profit with work:

$$(15) \quad \mathcal{P}_{\mathbf{i}, I} := \frac{\sum_{\mathbf{j} \in A_{\mathbf{i}, I}} \zeta_{\mathbf{j}, I}}{\sum_{\mathbf{j} \in A_{\mathbf{i}, I}} W_{\mathbf{j}}}$$

where the *work* is defined as $W_{\mathbf{j}} := \prod_{n=1}^N (m(j_n) - m(j_{n-1}))$.

2. CONVERGENCE OF THE PARAMETRIC ENRICHMENT ALGORITHM

We examine the convergence properties of a simplified version of Algorithm 1, also discussed in [32]. In the present case, we suppose to be able to sample the function $u : \Gamma \rightarrow V$ for any fixed parameter $\mathbf{y} \in \Gamma$. Thus, a discrete solution is given by the sparse grid interpolant $S_I[u] \in \mathbb{P}_I(\Gamma, V)$,

for a downward-close multi-index set $I \subset \mathbb{N}^N$. Moreover, the a-posteriori estimator, which in the fully discrete setting was the sum of parametric estimator (10) and finite element estimator (11), simplifies to $\zeta_{SC,I} := \sum_{i \in \mathcal{M}_I} \zeta_{i,I}$ (no additional term accounting for the finite element discretization) where the pointwise estimator is

$$\zeta_{i,I} := \left\| \Delta^{m(i)}(a \nabla S_I[u]) \right\|_{L^\infty(\Gamma, L^2(D))}.$$

In this setting, the reliable upper bound (12) simplifies to: $\|u - S_I[u]\|_{L^\infty(\Gamma, V)} \lesssim \zeta_{SC,I}$. Workless-profit and profit with work are defined analogously to (14) and (15) respectively. This simplified version of the algorithm reads:

Algorithm 4 $u_\epsilon \leftarrow SC(\epsilon)$

```

1:  $I := \{\mathbf{1}\}$ 
2:  $u_\epsilon := S_I[u]$ 
3: compute  $\zeta_{SC,I}$ 
4: while  $\zeta_{SC,I} \geq \epsilon$  do
5:    $i := \arg \max_{i \in \mathcal{M}_I} \mathcal{P}_{i,I}$ 
6:    $I \leftarrow I \cup A_{i,I}$ 
7:    $u_\epsilon \leftarrow S_I[u]$ 
8:   compute new a-posteriori estimator  $\zeta_{SC,I}$ 
9: end while

```

2.1. Preliminary results.

2.1.1. *Stability and convergence of the hierarchical surplus $\Delta^{m(i)}$.* In this section we recall basic results on the hierarchical surplus operator $\Delta^{m(i)}$ (see for instance [40]). The analysis is carried out in the $L^\infty(\Gamma, V)$ norm as it is the most "stringent" among the $L^p(\Gamma, V)$ norms for $p \in [1, \infty]$. We note that all the arguments below work with other choices of p .

We will first state 1D results (corresponding to the case $N = 1$). For $i \in \mathbb{N}$, the Lebesgue constant $\lambda_{m(i)}$ of the interpolant $\mathcal{U}^{m(i)}$ satisfies the relation

$$(16) \quad \left\| \mathcal{U}^{m(i)} v \right\|_{L^\infty(\Gamma, V)} \leq \lambda_{m(i)} \|v\|_{L^\infty(\Gamma, V)} \quad \forall v \in C^0(\Gamma, V).$$

Moreover, since CC nodes and the *doubling rule* (8) are used, it can be estimated as (see [18])

$$(17) \quad \lambda_{m(i)} \leq 2i.$$

Therefore, the relation (16) can be rewritten explicitly with respect to i as

$$(18) \quad \left\| \mathcal{U}^{m(i)} v \right\|_{L^\infty(\Gamma, V)} \lesssim i \|v\|_{L^\infty(\Gamma, V)} \quad \forall v \in C^0(\Gamma, V).$$

The estimate (18) can be used to derive a stability estimate for the detail operator

$$\left\| \left(\mathcal{U}^{m(i)} - \mathcal{U}^{m(i-1)} \right) v \right\|_{L^\infty(\Gamma, V)} \lesssim i \|v\|_{L^\infty(\Gamma, V)}.$$

Moving to the general case $N \in \mathbb{N}$, we can now exploit the tensor product structure of $\Gamma \subset \mathbb{R}^N$ to obtain a stability estimate for the hierarchical surplus operator

$$(19) \quad \left\| \Delta^{m(i)} v \right\|_{L^\infty(\Gamma, V)} \lesssim \left(\prod_{n=1}^N i_n \right) \|v\|_{L^\infty(\Gamma, V)}.$$

Since this estimate will be employed several times in the rest of the paper, we denote this bound on the norm of $\Delta^{m(\mathbf{i})}$ by

$$(20) \quad \Lambda_{\mathbf{i}} := \prod_{n=1}^N i_n.$$

We derive another estimate of $\|\Delta^{m(\mathbf{i})}u\|_{L^\infty(\Gamma, V)}$ that relies on the fact that $u : \Gamma \rightarrow V$ is analytic with respect to \mathbf{y} . The tensor product structure of Γ allows us again to start from a 1D results and then generalize to N dimensions. We state a result that relates the best approximation error in $\mathbb{P}_m(\Gamma, V)$ to the size of the domain of the holomorphic extension of u (2).

Lemma 2.1 ([2, Lemma 4.4]). *If $v \in C^0(\Gamma, V)$ and it exists $\tau > 0$ such that v admits an analytic extension to $\Sigma(\Gamma, \tau)$ (defined in (2)), then for $m \in \mathbb{N}$*

$$(21) \quad E_m(v) := \min_{w \in \mathbb{P}_m(\Gamma, V)} \|v - w\|_{L^\infty(\Gamma, V)} \leq \frac{2}{e^\sigma - 1} e^{-\sigma m} \max_{z \in \Sigma(\Gamma, \tau)} \|v(z)\|_V$$

where $\sigma := \log\left(\frac{2\tau}{|\Gamma|} + \sqrt{1 + \frac{4\tau^2}{|\Gamma|^2}}\right) > 0$. □

Since $\mathcal{U}^{m(\mathbf{i})}$ is exact on $\mathbb{P}_{m(\mathbf{i})-1}(\Gamma, V)$, its error can be expressed as (see [4])

$$\left\|u - \mathcal{U}^{m(\mathbf{i})}u\right\|_{L^\infty(\Gamma, V)} \leq (1 + \lambda_{m(\mathbf{i})}) E_{m(\mathbf{i})-1}(u).$$

Remembering (17) and the previous lemma, the error estimate for $\mathcal{U}^{m(\mathbf{i})}$ can be simplified as

$$\left\|u - \mathcal{U}^{m(\mathbf{i})}u\right\|_{L^\infty(\Gamma, V)} \lesssim i e^{-\sigma m(\mathbf{i})} \max_{z \in \Sigma(\Gamma, \tau)} \|u(z)\|_V.$$

This estimate can be applied to the detail operator after a triangle inequality to obtain

$$(22) \quad \left\|\Delta^{m(\mathbf{i})}u\right\|_{L^\infty(\Gamma, V)} \lesssim i e^{-\sigma m(\mathbf{i}-1)} \max_{z \in \Sigma(\Gamma, \tau)} \|u(z)\|_V.$$

Applying (22) to the multidimensional case (by considering one component at a time) leads to the following estimate:

Lemma 2.2. *For $\mathbf{i} \in \mathbb{N}^N$, the hierarchical surplus of an analytic function $u : \Gamma \rightarrow V$ satisfies*

$$(23) \quad \left\|\Delta^{m(\mathbf{i})}(u)\right\|_{L^\infty(\Gamma, V)} \lesssim \Lambda_{\mathbf{i}} e^{-\sigma |m(\mathbf{i}-1)|_1}.$$

where

$$\sigma := \min_{n \in \{1, \dots, N\}} \sigma_n, \quad \sigma_n := \log\left(\frac{2\tau_n}{|\Gamma_n|} + \sqrt{1 + \frac{4\tau_n^2}{|\Gamma_n|^2}}\right),$$

where the hidden constant depends on u .

2.1.2. *A simplified formula for $\zeta_{\mathbf{i}, I}$.* In the present section we highlight elementary facts on the zeros of $\Delta^{m(\mathbf{j})}u$ and the kernel of $\Delta^{m(\mathbf{j})}$. These facts are combined to show that the operator $\Delta^{m(\mathbf{i})}(a\nabla\Delta^{m(\mathbf{j})})$ is identically zero unless the multi-indices $\mathbf{i}, \mathbf{j} \in \mathbb{N}^N$ are “close to each other” (see also [32, Proposition 4.3] for partial results in this direction).

We will denote by $\mathcal{R}_{\mathbf{i}} \subset \mathbb{N}^N$ the axis-aligned rectangle with opposite vertices $\mathbf{1}$ and \mathbf{i} :

$$(24) \quad \mathcal{R}_{\mathbf{i}} := \{\mathbf{j} \in \mathbb{N}^N : j_n \leq i_n \ \forall n \in \{1, \dots, N\}\}.$$

Theorem 2.3. Given, $\mathbf{i}, \mathbf{j} \in \mathbb{N}^N$, if one of the following two conditions

$$(25) \quad \exists n \in 1, \dots, N : i_n < j_n$$

or

$$(26) \quad \forall n \in 1, \dots, N : \mathbf{j} + \mathbf{e}_n < \mathbf{i},$$

is satisfied, then

$$(27) \quad \Delta^{m(\mathbf{i})} \left(a \nabla \Delta^{m(\mathbf{j})} u \right) \equiv 0 \quad \forall u \in C^0(\Gamma, V).$$

Proof. Fix $\mathbf{y} \in \mathcal{Y}^{m(\mathbf{i})}$. By (25) and the nestedness of CC nodes, it exists $n \in 1, \dots, N$ such that $y_n \in \mathcal{Y}^{m(j_n-1)}$. This implies that $\Delta^{m(\mathbf{j})} u(\mathbf{y}) = 0$, as y_n is an interpolation point for both $\mathcal{U}_n^{m(j_n)}$ and $\mathcal{U}_n^{m(j_n-1)}$. Thus, recalling that ∇ acts on the space variable x only,

$$(28) \quad a(\mathbf{y}) \nabla \Delta^{m(\mathbf{j})} u(\mathbf{y}) = 0 \quad \forall \mathbf{y} \in \mathcal{Y}^{m(\mathbf{i})}.$$

Next, observe that a hierarchical surplus can be written as a linear combination of Lagrange interpolants

$$\Delta^{m(\mathbf{i})} = \sum_{\alpha \in \{0,1\}^N} (-1)^{|\alpha|} \mathcal{U}^{m(\mathbf{i}-\alpha)}.$$

By the nestedness of CC nodes, (28) implies that $a \nabla \Delta^{m(\mathbf{j})} u$ is in the kernel of each of the interpolants $\mathcal{U}^{m(\mathbf{i}-\alpha)}$, $\alpha \in \{0,1\}^N \setminus \{0\}$, which in turn implies (27). To show that (25) also implies (27), first observe that

$$(29) \quad a \nabla \Delta^{m(\mathbf{j})} u \in \sum_{n=1}^N \mathbb{P}_{m(\mathbf{j})-\mathbf{1}+\mathbf{e}_n} = \mathbb{P}_{\{\mathbf{j}\} \cup \mathcal{M}_{\{\mathbf{j}\}}} \subseteq \mathcal{P}_{\mathcal{R}_i \setminus \{\mathbf{i}\}},$$

where the last inclusion is due to assumption (26). Next, observe that a hierarchical surplus can be written as a difference of sparse grid interpolants: $\Delta^{m(\mathbf{i})} = S_{\mathcal{R}_i} - S_{\mathcal{R}_i \setminus \{\mathbf{i}\}}$. This implies that $\mathbb{P}_{\mathcal{R}_i \setminus \{\mathbf{i}\}}$ is a subset of the kernel of $\Delta^{m(\mathbf{i})}$, as both $S_{\mathcal{R}_i}$ and $S_{\mathcal{R}_i \setminus \{\mathbf{i}\}}$ are exact on this space. Together with (29), this concludes the proof. \square

Remark 2.4. The previous theorem can be used to simplify the computation of $\zeta_{\mathbf{i},I}$. Consider a multi-index set $I \subset \mathbb{N}^N$ and $\mathbf{i} \in \mathcal{M}_I$. Define

$$J_{\mathbf{i},I} := \{\mathbf{j} \in I : \exists n \in 1, \dots, N : \mathbf{j} = \mathbf{i} - \mathbf{e}_n\}.$$

Then, thanks to the previous theorem:

$$\Delta^{m(\mathbf{i})} (a \nabla S_I [u]) = \Delta^{m(\mathbf{i})} \left(a \nabla \sum_{\mathbf{j} \in I} \Delta^{m(\mathbf{j})} u \right) = \Delta^{m(\mathbf{i})} \left(a \nabla \sum_{\mathbf{j} \in J_{\mathbf{i},I}} \Delta^{m(\mathbf{j})} u \right),$$

so

$$(30) \quad \zeta_{\mathbf{i},I} = \left\| \Delta^{m(\mathbf{i})} \left(a \nabla \sum_{\mathbf{j} \in J_{\mathbf{i},I}} \Delta^{m(\mathbf{j})} u \right) \right\|_{L^\infty(\Gamma, L^2(D))}.$$

See Figure 1 for a graphical representation.

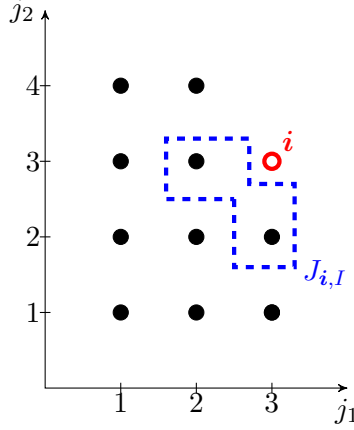


Figure 1. Graphical representation of the simplified computation of $\zeta_{i,I}$ from (30). We consider $N = 2$ parameters, each point in the plot corresponds to an element $\mathbf{j} = (j_1, j_2) \in \mathbb{N}^2$, where on the x-axis we represent j_1 and on the y-axis j_2 . Filled dots represent I , the red hollow one is $i \in \mathcal{M}_I$. The blue dashed line encircles the multi-indices in $J_{i,I}$, i.e. the only relevant ones in I for the computation of $\zeta_{i,I}$, as explained in Remark 2.4.

2.1.3. *A priori estimates for estimators and index sets.*

Proposition 2.5. *Given $u : \Gamma \rightarrow V$ analytic, a multi-index set $I \subset \mathbb{N}^N$ and $i \in \mathcal{M}_I$, the pointwise error estimator can be bounded as*

$$\zeta_{i,I} \lesssim N \Lambda_i^2 e^{-\sigma|m(\mathbf{i}-\mathbf{1})|_1},$$

where Λ_i is defined in (20).

Proof. Observe that $S_I[u]$ is analytic but not uniformly with respect to I , so one cannot apply directly the convergence result for the hierarchical surplus. Recalling Remark 2.4, we can simplify the expression of $\zeta_{i,I}$ as

$$\begin{aligned} \zeta_{i,I} &= \left\| \Delta^{m(\mathbf{i})} (a \nabla S_I[u]) \right\|_{L^\infty(\Gamma, L^2(D))} \\ &= \left\| \Delta^{m(\mathbf{i})} \left(a \nabla \sum_{\substack{n \in \{1, \dots, N\} \\ \mathbf{i} - \mathbf{e}_n \in I}} \Delta^{m(\mathbf{i} - \mathbf{e}_n)} u \right) \right\|_{L^\infty(\Gamma, L^2(D))}. \end{aligned}$$

Applying the stability of $\Delta^{m(\mathbf{i})}$, boundedness of a , and the triangle inequality, we obtain

$$\zeta_{i,I} \lesssim \Lambda_i \sum_{\substack{n \in \{1, \dots, N\} \\ \mathbf{i} - \mathbf{e}_n \in I}} \left\| \Delta^{m(\mathbf{i} - \mathbf{e}_n)} \nabla u \right\|_{L^\infty(\Gamma, L^2(D))}.$$

Observe finally that, since u is analytic, we can apply Lemma 2.2 to obtain

$$\zeta_{i,I} \leq \Lambda_i \sum_{\substack{n \in \{1, \dots, N\} \\ \mathbf{i} - \mathbf{e}_n \in I}} \Lambda_{\mathbf{i} - \mathbf{e}_n} e^{-\sigma|m(\mathbf{i} - \mathbf{e}_n - \mathbf{1})|_1} \lesssim N \Lambda_i^2 e^{-\sigma|m(\mathbf{i} - \mathbf{1})|_1}.$$

□

Remark 2.6. A direct consequence of the previous proposition is the uniform boundedness of the sequence of a-posteriori estimators $(\zeta_{SC, I_\ell})_\ell$. Indeed, we have the following bound independently of the iteration number ℓ

$$\zeta_{SC, I_\ell} = \sum_{\mathbf{i} \in \mathcal{M}_{I_\ell}} \zeta_{\mathbf{i}, I_\ell} \lesssim N \sum_{\mathbf{i} \in \mathcal{M}_{I_\ell}} \Lambda_{\mathbf{i}} e^{-\sigma |m(\mathbf{i}-1)|_1} \leq N \sum_{\mathbf{i} \in \mathbb{N}^N} \Lambda_{\mathbf{i}} e^{-\sigma |m(\mathbf{i}-1)|_1} < \infty.$$

Lemma 2.7. The profit maximizer $\mathbf{i}_\ell \in \mathbb{N}^N$ at iteration ℓ of Algorithm 3 satisfies

$$\Lambda_{\mathbf{i}_\ell} = \prod_{n=1}^N \langle \mathbf{i}_\ell, \mathbf{e}_n \rangle \leq \left(1 + \frac{\ell}{N}\right)^N.$$

Moreover, there holds

$$(31a) \quad \#A_{\mathbf{i}_\ell, I_\ell} \leq \left(1 + \frac{\ell}{N}\right)^N$$

as well as

$$(31b) \quad \#\mathcal{M}_{I_\ell} \leq N \left(1 + (\ell - 1) \left(1 + \frac{\ell - 1}{N}\right)^N\right).$$

Proof. First observe that due to the arithmetic-geometric inequality,

$$\prod_{n=1}^N j_n \leq \left(\frac{\sum_{n=1}^N j_n}{N}\right)^N = \left(\frac{|\mathbf{j}|_1}{N}\right)^N \quad \forall \mathbf{j} \in \mathbb{R}^N.$$

Then, it can be easily proved by induction that $|\mathbf{i}_\ell|_1 = N + \ell$. To prove (31), first observe that $A_{\mathbf{i}} = \mathcal{R}_{\mathbf{i}} \setminus I$, where $\mathcal{R}_{\mathbf{i}}$ is the axis-aligned rectangle in \mathbb{N}^N as defined in (24). Thus, $\#A_{\mathbf{i}_\ell, I_\ell} \leq \#\mathcal{R}_{\mathbf{i}_\ell, I_\ell} = \Lambda_{\mathbf{i}_\ell}$ and due to the previous lemma we obtained the desired bound. As for the second estimate, first observe that $\#\mathcal{M}_{I_\ell} \leq N \#I_\ell$. Then, an estimate on $\#I_\ell$ comes from the partition $I_\ell = \{\mathbf{1}\} \cup \bigcup_{m=1}^{\ell-1} A_{\mathbf{i}_m}$ and the estimate on $\#A_{\mathbf{i}_\ell}$. \square

2.1.4. Remarks on the algorithm driven by workless profit. In this section, we point out some elementary facts on the behavior of the algorithm when the workless profit defined in (14) is used. Inspired by [9], we give the following definition:

Definition 2.8. Given a downward closed multi-index set $I \subset \mathbb{N}^N$, $\mathbf{i} \in \mathcal{M}_I$ is maximal in \mathcal{M}_I if and only if

$$\forall \mathbf{j} \in \mathcal{M}_I \setminus \{\mathbf{i}\}, \exists n \in 1, \dots, N : i_n > j_n.$$

The set of maximal points in \mathcal{M}_I is denoted by μ_I .

Example 2.9. If $\mathbf{i} \in \mathbb{N}^N$ and $I = \mathcal{R}_{\mathbf{i}}$ is an axis-aligned rectangle as defined in (24), then

$$\mu_I = \{\mathbf{i} + \mathbf{e}_n, n \in 1, \dots, N\}.$$

Lemma 2.10. For the workless profit (14), the selected point \mathbf{i}_ℓ is maximal in \mathcal{M}_{I_ℓ}

$$(32) \quad \mathbf{i}_\ell \in \mu_{I_\ell}.$$

Therefore, I_ℓ is an axis-aligned rectangle in \mathbb{N}^N , i.e.

$$(33) \quad I_\ell = \mathcal{R}_{\mathbf{i}_{\ell-1}}.$$

Proof. We prove (32) by contradiction. If \mathbf{i}_ℓ is not maximal, there exists $\mathbf{j} \in \mathcal{M}_{I_\ell} \setminus \{\mathbf{i}_\ell\}$ such that for all $n \in 1, \dots, N$ $\langle \mathbf{i}_\ell, \mathbf{e}_n \rangle \leq j_n$, which implies $\mathbf{i}_\ell \in \mathcal{R}_{\mathbf{j}}$. Thus, $\mathbf{i}_\ell \in A_{\mathbf{j}, I_\ell} = \mathcal{R}_{\mathbf{j}} \setminus I_\ell$ and by definition of the workless profit, we have the contradiction $\mathcal{P}_{\mathbf{i}_\ell, I_\ell} < \mathcal{P}_{\mathbf{j}, I_\ell}$.

The second fact (33) can be proved by induction. For $\ell = 1$, $I_1 = \mathcal{R}_{\mathbf{1}} = \{\mathbf{1}\}$. Assume that for fixed $\ell \in \mathbb{N}$, $I_\ell = \mathcal{R}_{\mathbf{i}_{\ell-1}}$. With (32) and Example 2.9, we know

$$\mathbf{i}_\ell \in \mu_{I_\ell} = \mu_{\mathcal{R}_{\mathbf{i}_{\ell-1}}} = \{\mathbf{i}_{\ell-1} + \mathbf{e}_n, n \in 1, \dots, N\}.$$

Thus $I_{\ell+1} = I_\ell \cup A_{\mathbf{i}_\ell, I_\ell} = \mathcal{R}_{\mathbf{i}_\ell}$. □

To summarize, the use of the workless profit (14) implies that, for all $\ell > 0$,

- it exists a *unique* number $n(\ell) \in 1, \dots, N$ such that

$$(34) \quad \mathbf{i}_{\ell+1} = \mathbf{i}_\ell + \mathbf{e}_{n(\ell)}.$$

- as a consequence, the norm of \mathbf{i}_ℓ is given by:

$$(35) \quad |\mathbf{i}_{\ell+1}|_1 = |\mathbf{i}_\ell|_1 + 1 = N + \ell.$$

- I_ℓ is a rectangle:

$$(36) \quad I_{\ell+1} = \mathcal{R}_{\mathbf{i}_\ell}.$$

Therefore, the sparse grid stochastic collocation interpolant is actually a full tensor product Lagrange interpolant:

$$S_{I_{\ell+1}} = \bigotimes_{n=1}^N \mathcal{U}_n^{m(\langle \mathbf{i}_\ell, \mathbf{e}_n \rangle)}.$$

- the multi-indices added at iteration ℓ are

$$(37) \quad A_{\mathbf{i}_\ell, I_\ell} = I_{\ell+1} \setminus I_\ell = \{\mathbf{j} \in \mathcal{R}_{\mathbf{i}_\ell} : j_{n(\ell)} = \langle \mathbf{i}_\ell, \mathbf{e}_{n(\ell)} \rangle\}.$$

In other words, the evolution of the approximation space is determined by the sequence of integers $(n(\ell))_\ell$. This allows us to simplify the notation as follows

$$\begin{aligned} A_{n, I_\ell} &:= A_{\mathbf{i}_{\ell-1} + \mathbf{e}_n, I_\ell} \\ \mathcal{P}_{n, I_\ell} &:= \sum_{\mathbf{j} \in A_{n, I_\ell}} \zeta_{\mathbf{j}, I_\ell} \end{aligned}$$

Let us moreover denote the maximal n -th dimension of I_ℓ as

$$(38) \quad r_{n, \ell} := \max_{\mathbf{j} \in I_\ell} j_n.$$

See Figure 2 for a graphical representation.

The estimate for the pointwise error estimator from Proposition 2.5 can be improved as follows. First observe that, due to (34) and (36),

$$\mathcal{J}_{\mathbf{i}, I_\ell} = \{\mathbf{j} \in I_\ell : \exists n \in 1, \dots, N : \mathbf{j} = \mathbf{i} - \mathbf{e}_n\} = \{\mathbf{i} - \mathbf{e}_{n(\ell)}\}.$$

Thus, $\#\mathcal{J}_{\mathbf{i}, I_\ell} = 1$ and we may reduce the dependence on N by

$$(39) \quad \zeta_{\mathbf{i}, I_\ell} \lesssim \Lambda_{\mathbf{i}}^2 e^{-\sigma|m(\mathbf{i}-1)|}.$$

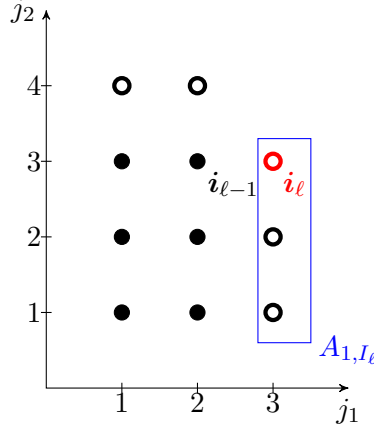


Figure 2. Example of approximation parameters at a generic step ℓ of the algorithm when the workless profit (14) is used. We consider $N = 2$ parameters, each point in the plot corresponds to an element $\mathbf{j} = (j_1, j_2) \in \mathbb{N}^2$, where on the x-axis we represent j_1 and on the y-axis j_2 . Filled dots represent I_ℓ , hollow ones \mathcal{M}_{I_ℓ} . The multi-index selected by the algorithm at current step, \mathbf{i}_ℓ , is in red (so in this case $n(\ell) = 1$). The blue rectangle encircles multi-indices in $A_{n(\ell), I_\ell}$.

2.2. Convergence of the parametric estimator. In the following two lemmata, we prove that Algorithm 4 driven by workless profit and profit with work respectively forces the maximum profit over the margin to zero.

Proposition 2.11. *If the workless profit (14) is used, then*

$$\lim_{\ell \rightarrow \infty} \mathcal{P}_{n(\ell), I_\ell} = 0.$$

Proof. For fixed $n \in 1, \dots, N$, we estimate each pointwise error estimator appearing in \mathcal{P}_{n, I_ℓ} by (39) and the fact that for any \mathbf{i} in A_{n, I_ℓ} , $i_n = r_{n, \ell} + 1$.

$$\begin{aligned} \mathcal{P}_{n, I_\ell} &= \sum_{\mathbf{j} \in A_{n, I_\ell}} \zeta_{\mathbf{j}, I_\ell} \lesssim \sum_{\mathbf{i} \in A_{n, I_\ell}} \Lambda_{\mathbf{i}}^2 e^{-\sigma |m(\mathbf{i}-1)|_1} \\ &= \sum_{\mathbf{i} \in A_{n, I_\ell}} \prod_{k=1}^N \left(i_k^2 e^{-\sigma |m(i_k-1)|} \right) \\ &\leq (r_{n, \ell} + 1)^2 e^{-\frac{\sigma}{2} m(r_{n, \ell})} \sum_{\mathbf{i} \in A_{n, I_\ell}} \left(i_n^2 e^{-\frac{\sigma}{2} m(i_n+1)} \prod_{k=1, k \neq n}^N \left(i_k^2 e^{-\sigma m(i_k-1)} \right) \right) \\ &\leq (r_{n, \ell} + 1)^2 e^{-\frac{\sigma}{2} m(r_{n, \ell})} \sum_{\mathbf{i} \in A_{n, I_\ell}} \Lambda_{\mathbf{i}}^2 e^{-\frac{\sigma}{2} |m(\mathbf{i}-1)|_1}. \end{aligned}$$

The last factor is uniformly bounded with respect to ℓ (but this bound depends on the number of dimensions N)

$$\sum_{\mathbf{i} \in A_{n, I_\ell}} \Lambda_{\mathbf{i}}^2 e^{-\frac{\sigma}{2} |m(\mathbf{i}-1)|_1} \leq \sum_{\mathbf{i} \in \mathbb{N}^N} \Lambda_{\mathbf{i}}^2 e^{-\frac{\sigma}{2} |m(\mathbf{i}-1)|_1} < \infty.$$

We are left with:

$$\mathcal{P}_{n, I_\ell} \lesssim (r_{n, \ell} + 1)^2 e^{-\frac{\sigma}{2} m(r_{n, \ell})}.$$

The proof is completed by observing that $\lim_{\ell \rightarrow \infty} r_{n(\ell), \ell} = \infty$. □

For the profit with work, we can even show convergence to zero of the profit without using the analyticity assumption on u . This is not relevant for the problem at hand, as the analyticity follows immediately, but may be relevant for more complicated and less regular random coefficients.

Proposition 2.12. *There holds $\lim_{\ell \rightarrow \infty} \mathcal{P}_{\mathbf{i}_\ell, I_\ell} = 0$.*

Proof. As in the proof of Proposition 2.5, but without using any analyticity of u , we obtain with (19) that

$$\zeta_{\mathbf{i}, I} \lesssim \Lambda_{\mathbf{i}}^2 N \|\nabla u\|_{L^\infty(\Gamma, L^2(D))}.$$

We observe that the doubling rule (8) implies

$$(40) \quad 2^{|\mathbf{i}|_1 - 2N} \leq W_{\mathbf{i}} \leq 2^{|\mathbf{i}|_1 - N}.$$

Thus, the profit can be estimated as:

$$\mathcal{P}_{\mathbf{i}_\ell, I_\ell} \lesssim \frac{\sum_{\mathbf{j} \in A_{\mathbf{i}_\ell, I_\ell}} \zeta_{\mathbf{j}, I_\ell}}{\sum_{\mathbf{j} \in A_{\mathbf{i}_\ell, I_\ell}} W_{\mathbf{j}}} \lesssim \frac{\#A_{\mathbf{i}_\ell, I_\ell} \Lambda_{\mathbf{i}_\ell}^2 N}{W_{\mathbf{i}_\ell}} \leq N(1 + \ell/N)^N \Lambda_{\mathbf{i}_\ell}^2 2^{2N - |\mathbf{i}_\ell|_1}.$$

Since $2^{|\mathbf{i}_\ell|_1}$ grows much faster than $\Lambda_{\mathbf{i}_\ell}^2 = \prod_{n=1}^N i_{\ell, n}^2$, we conclude the proof. \square

The following result shows that, if a multi-index $\mathbf{i} \in \mathbb{N}^N$ stays in the margin indefinitely, then it's pointwise estimator vanishes. This result is valid for both workless profit and profit with work.

Proposition 2.13. *Let $\widehat{\mathbf{i}} \in \mathbb{N}^N$ and suppose the index remains in the margin indefinitely, i.e.,*

$$\exists \ell_0 \in \mathbb{N} : \forall \ell \geq \ell_0, \widehat{\mathbf{i}} \in \mathcal{M}_{I_\ell}.$$

Then, the pointwise error estimator corresponding to $\widehat{\mathbf{i}}$ vanishes

$$\lim_{\ell \rightarrow \infty} \zeta_{\widehat{\mathbf{i}}, I_\ell} = 0.$$

Proof. Let $\widehat{\mathbf{i}} \in \mathbb{N}^N$ such that $\widehat{\mathbf{i}} \in \mathcal{M}_{I_\ell}$ for all $\ell > \ell_0$. Thus, $\widehat{\mathbf{i}} \neq \mathbf{i}_\ell$ for any $\ell > \ell_0$, which means that

$$\mathcal{P}_{\widehat{\mathbf{i}}, I_\ell} \leq \mathcal{P}_{\mathbf{i}_\ell, I_\ell} \quad \forall \ell > \ell_0.$$

In case the profit with work (15) is used, since $\lim_{\ell \rightarrow \infty} \mathcal{P}_{\mathbf{i}_\ell, I_\ell} = 0$ as proved in Proposition 2.12, we have that $\lim_{\ell \rightarrow \infty} \mathcal{P}_{\widehat{\mathbf{i}}, I_\ell} = 0$ (otherwise $\widehat{\mathbf{i}}$ would be selected at some point). Moreover, since $\sum_{\mathbf{j} \in A_{\widehat{\mathbf{i}}, I_\ell}} W_{\mathbf{j}}$ (i.e. the denominator in the profit $\mathcal{P}_{\widehat{\mathbf{i}}, I_\ell}$) is eventually constant with respect to ℓ , we have that $\lim_{\ell \rightarrow \infty} \sum_{\mathbf{j} \in A_{\widehat{\mathbf{i}}, I_\ell}} \zeta_{\mathbf{j}, I_\ell} = 0$, and in particular we obtain the statement. The same holds if the profit without work (14) is employed, as in Proposition 2.11 we have proved that also in this case $\lim_{\ell \rightarrow \infty} \mathcal{P}_{\mathbf{i}_\ell, I_\ell} = 0$. \square

Remark 2.14. *Recall the simplified formula (30) for $\zeta_{\widehat{\mathbf{i}}, I_\ell}$ with $J_{\widehat{\mathbf{i}}, I_\ell} := \{\widehat{\mathbf{i}} - \mathbf{e}_n : n \in 1, \dots, N\}$.*

Observe that $\left(J_{\widehat{\mathbf{i}}, I_\ell}\right)_\ell$ is eventually constant, i.e. it exists $\ell_2 > \ell_0$ (as defined in the previous proposition) such that for all $\ell > \ell_2$ $J_{\widehat{\mathbf{i}}, I_\ell} = J_{\widehat{\mathbf{i}}, I_{\ell_2}}$. Thus, $\left(\zeta_{\widehat{\mathbf{i}}, I_\ell}\right)_\ell$ is also eventually constant. Therefore, $\left(\zeta_{\widehat{\mathbf{i}}, I_\ell}\right)_\ell$ does not only vanish in the limit, but is actually eventually zero:

$$\forall \ell > \ell_2, \zeta_{\widehat{\mathbf{i}}, I_\ell} = 0.$$

We can finally prove the convergence of the parameter-enrichment algorithm with a technique inspired by [6, Proposition 10].

Theorem 2.15 (Convergence of the parameter-enrichment algorithm). *The adaptive stochastic collocation Algorithm 4 driven by either workless profit or profit with work, leads to a vanishing sequence of a-posteriori error estimators, thus also leading to a convergent sequence of discrete solutions*

$$\lim_{\ell \rightarrow \infty} \zeta_{SC, I_\ell} = 0 = \lim_{\ell \rightarrow \infty} \|u - S_{I_\ell}[u]\|_{L^\infty(\Gamma, V)}$$

Proof. The a-posteriori error estimator at step $\ell \in \mathbb{N}$ can be written as

$$\zeta_{SC, I_\ell} = \sum_{\mathbf{i} \in \mathbb{N}^N} \zeta_{\mathbf{i}, I_\ell} \mathbb{1}_{\mathcal{M}_{I_\ell}}(\mathbf{i}),$$

where $\mathbb{1}_{\mathcal{M}_{I_\ell}}$ is the indicator function of the margin \mathcal{M}_{I_ℓ} . In order to prove that the sequence vanishes by dominated convergence, it is sufficient to prove that (i) for any $\mathbf{i} \in \mathbb{N}^N$, $\lim_{\ell \rightarrow \infty} \zeta_{\mathbf{i}, I_\ell} \mathbb{1}_{\mathcal{M}_{I_\ell}} = 0$ and (ii) that the sequence $(\zeta_{SC, I_\ell})_\ell$ is bounded. The uniform boundedness (ii) was proved in Remark 2.6. As for (i), observe that at least one of the following cases applies:

- \mathbf{i} is eventually added to I_ℓ , thus $\mathbb{1}_{\mathcal{M}_{I_\ell}}(\mathbf{i})$ is eventually zero;
- \mathbf{i} is never added to the margin (for all $\ell \in \mathbb{N}$, $\mathbf{i} \in \mathbb{N}^N \setminus \mathcal{M}_{I_\ell}$), thus $\zeta_{\mathbf{i}, I_\ell}$ is constantly zero;
- it exists $\bar{\ell} \in \mathbb{N}$ such that for any $\ell \geq \bar{\ell}$, $\mathbf{i} \in \mathcal{M}_{I_\ell}$. In this case, due to Proposition 2.13, $\lim_{\ell \rightarrow \infty} \zeta_{\mathbf{i}, I_\ell} = 0$.

This concludes the proof. \square

2.3. Convergence of the parametric error. In the present section we denote by $\mathcal{L}(L^\infty(\Gamma, V))$ the space of linear bounded operators $T : L^\infty(\Gamma, V) \rightarrow L^\infty(\Gamma, V)$. It is well known that this is a Banach space when equipped with the usual operator norm

$$\|T\|_{\mathcal{L}(L^\infty(\Gamma, V))} := \sup_{u \in L^\infty(\Gamma, V), u \neq 0} \frac{\|Tu\|_{L^\infty(\Gamma, V)}}{\|u\|_{L^\infty(\Gamma, V)}}.$$

We have the following monotonicity property of the approximation error of $S_I[\cdot]$ with respect to I :

Lemma 2.16. *Let $u \in C^0(\Gamma, V)$ and $I, J \subset \mathbb{N}^N$ downward-closed multi-index sets such that $J \subset I$. Then*

$$\|u - S_I[u]\|_{L^\infty(\Gamma, V)} \leq \left(1 + \|S_I\|_{\mathcal{L}(L^\infty(\Gamma, V))}\right) \|u - S_J[u]\|_{L^\infty(\Gamma, V)}.$$

Proof. With the identity operator $\mathbf{1}$ on $C^0(\Gamma, V)$, observe that

$$u - S_I[u] = (\mathbf{1} - S_I)u = (\mathbf{1} - S_I)(\mathbf{1} - S_J)u$$

since $J \subset I$ implies $S_I[S_J[u]] = S_J[u]$. The triangle inequality concludes the proof. \square

In the present section we provide error estimates for S_{I_ℓ} with respect to the number of iterations ℓ . We consider both the possible definitions of profit (14) and (15).

Remark 2.17. *The quantity $\|S_{I_\ell}\|_{\mathcal{L}(L^\infty(\Gamma, V))}$ from Lemma 2.16 satisfies*

- *Workless profit:* $I_\ell = \mathcal{R}_{\mathbf{i}_{\ell-1}}$, i.e. S_{I_ℓ} is actually a tensor-product Lagrange interpolant (see Section 2.1.4). Therefore, we can estimate

$$(41) \quad \|S_{I_\ell}\|_{\mathcal{L}(L^\infty(\Gamma, V))} = \left\| \bigotimes_{n=1}^N \mathcal{U}_n^{m(\langle \mathbf{i}_{\ell-1}, \mathbf{e}_n \rangle)} \right\|_{\mathcal{L}(L^\infty(\Gamma, V))} \leq \prod_{n=1}^N \langle \mathbf{i}_{\ell-1}, \mathbf{e}_n \rangle \leq \left(1 + \frac{\ell-1}{N}\right)^N,$$

where in the first inequality we used the stability bound for the Lagrange interpolant (18) and Lemma 2.7 for the second inequality.

- *Profit with work: Partitioning I_ℓ with the sequence $(A_{i_m, I_m})_{m=1}^{\ell-1}$ and using Lemma 2.7.*

$$(42) \quad \begin{aligned} \|S_{I_\ell}\|_{\mathcal{L}(L^\infty(\Gamma, V))} &\leq \sum_{i \in I_\ell} \left\| \Delta^{m(i)} \right\|_{\mathcal{L}(L^\infty(\Gamma, V))} \leq \sum_{m=1}^{\ell-1} \#A_{i_m, I_m} \Lambda_{i_m} \\ &\leq (\ell-1) \left(1 + \frac{\ell-1}{N} \right)^{2N}. \end{aligned}$$

We finally prove the parametric error estimates, first with workless profit, then with profit with work.

Theorem 2.18. *Consider Algorithm 4 with workless profit defined in (14). Denote by I_ℓ the downward-closed multi-index sets chosen by the algorithm at step $\ell > 0$ and by $S_{I_\ell}[u]$ the corresponding sparse grid stochastic collocation approximation of the analytic function $u : \Gamma \rightarrow V$. Then,*

$$(43) \quad \|u - S_{I_\ell}[u]\|_{L^\infty(\Gamma, V)} \lesssim \left(1 + \left(1 + \frac{\ell-1}{N} \right)^N \right) N \ell^2 e^{-\frac{\sigma}{2} m(1 + \frac{\ell}{N})} \quad \forall \ell > 0.$$

Proof. Fix $\ell > 0$. Recall the definition of $r_{n, \ell}$ from (38) and consider the direction $\bar{n} \in \{1, \dots, N\}$ which maximizes $r_{n, \ell}$. With $n(\ell)$ from (34), define

$$\ell' := \max \{ \ell' \in 1, \dots, \ell : n(\ell') = \bar{n} \}$$

and observe that with each iteration, at least one side of the axis aligned rectangle I_ℓ is increased by one, i.e.,

$$(44) \quad r_{n(\ell'), \ell'} = r_{\bar{n}, \ell} \geq 1 + \frac{\ell}{N}.$$

Applying estimate (41) from the previous remark, we can bound

$$\|u - S_{I_\ell}[u]\|_{L^\infty(\Gamma, V)} \leq \left(1 + \left(1 + \frac{\ell-1}{N} \right)^N \right) \|u - S_{I_{\ell'}}[u]\|_{L^\infty(\Gamma, V)}.$$

Now, apply the reliability of the error estimator proved in [32, Proposition 4.3] to obtain

$$\|u - S_{I_{\ell'}}[u]\|_{L^\infty(\Gamma, V)} \lesssim \sum_{i \in \mathcal{M}_{I_{\ell'}}} \zeta_{i, I_{\ell'}}.$$

Recalling the definition of $A_{n, I_{\ell'}}$ and $\mathcal{P}_{n, I_{\ell'}}$ for $n \in 1, \dots, N$ given in Section 2.1.4, we have

$$\sum_{i \in \mathcal{M}_{I_{\ell'}}} \zeta_{i, I_{\ell'}} = \sum_{n=1}^N \sum_{i \in A_{n, I_{\ell'}}} \zeta_{i, I_{\ell'}} = \sum_{n=1}^N \mathcal{P}_{n, I_{\ell'}} \leq N \mathcal{P}_{n(\ell'), I_{\ell'}}.$$

The profit $\mathcal{P}_{n(\ell'), I_{\ell'}}$ can now be bounded as a function of $r_{n(\ell'), \ell'}$ as we did in Proposition 2.11

$$\mathcal{P}_{n(\ell'), I_{\ell'}} = \sum_{j \in A_{n(\ell'), I_{\ell'}}} \zeta_{j, I_{\ell'}} \leq \sum_{j \in A_{n(\ell'), I_{\ell'}}} \left(\prod_{k=1}^N j_k \right)^2 e^{-\frac{\sigma}{2} |m(j-1)|} \lesssim r_{n(\ell'), \ell'}^2 e^{-\frac{\sigma}{2} m(r_{n(\ell'), \ell'})},$$

where in the first inequality we have applied the estimate (39) on $\zeta_{j, I_{\ell'}}$ and in the second we have exploited the fact that, for $\mathbf{j} \in A_{n(\ell'), I_{\ell'}}$, $j_{n(\ell')} = r_{n(\ell'), \ell'} + 1$. Recalling that $1 + \frac{\ell}{N} \leq r_{n(\ell'), \ell'} \leq \ell + 1$, we obtain

$$\mathcal{P}_{n(\ell'), I_{\ell'}} \lesssim \ell^2 e^{-\frac{\sigma}{2} m(1 + \frac{\ell}{N})}.$$

□

Let us now prove the analogous result for the algorithm driven by profit with work.

Theorem 2.19. *Consider Algorithm 4 with profit with work defined in (15). Denote by I_ℓ the downward-closed multi-index sets chosen by the algorithm at step $\ell > 0$ and by $S_{I_\ell}[u]$ the corresponding sparse grid stochastic collocation approximation of the analytic function $u : \Gamma \rightarrow V$. Then,*

$$(45) \quad \|u - S_{I_\ell}[u]\|_{L^\infty(\Gamma, V)} \lesssim \ell^5 \left(\frac{\ell}{N}\right)^{4N} 2^{\ell(1-\frac{1}{N})} e^{-\frac{\sigma}{2}m\left(\ell^{\frac{1}{N}}\right)} \quad \forall \ell > 0.$$

Proof. For brevity, we write $\zeta_{\mathbf{i}}$, $A_{\mathbf{i}}$ and $\mathcal{P}_{\mathbf{i}}$ instead of $\zeta_{\mathbf{i}, I}$, $A_{\mathbf{i}, I}$ and $\mathcal{P}_{\mathbf{i}, I}$ respectively. Fix $\ell > 0$ and consider $\bar{r} := \max_{\mathbf{i} \in I_\ell} |\mathbf{i}|_{\ell^\infty}$ and $\bar{n} \in 1, \dots, N$ such that, for some $\mathbf{i} \in I_\ell$, $i_{\bar{n}} = \bar{r}$. Observe that $\#I_\ell \gtrsim \ell$ and hence

$$\bar{r} \geq \ell^{\frac{1}{N}}.$$

Consider the last step ℓ' in which I_ℓ has been extended in direction \bar{n} , i.e.,

$$(46) \quad \ell' := \max \{ \ell' \in 1, \dots, \ell : \langle \mathbf{i}_{\ell'}, \mathbf{e}_{\bar{n}} \rangle = \bar{r} \text{ and } \mathbf{i}_{\ell'} - \mathbf{e}_{\bar{n}} \in I_{\ell'} \}.$$

Applying estimate (42) from Remark 2.17, we can bound

$$(47) \quad \|u - S_{I_\ell}[u]\|_{L^\infty(\Gamma, V)} \leq \left(1 + (\ell - 1) \left(1 + \frac{\ell - 1}{N}\right)^{2N}\right) \|u - S_{I_{\ell'}}[u]\|_{L^\infty(\Gamma, V)}.$$

In [32, Proposition 4.3], the reliability of the error estimator is proved

$$\|u - S_{I_{\ell'}}[u]\|_{L^\infty(\Gamma, V)} \lesssim \sum_{\mathbf{i} \in \mathcal{M}_{I_{\ell'}}} \zeta_{\mathbf{i}}.$$

Recalling the definition of $\mu_{I_{\ell'}}$, the set of maximal elements in $\mathcal{M}_{I_{\ell'}}$ (Definition 2.8), the margin can be represented (but in general not partitioned) as

$$(48) \quad \mathcal{M}_{I_{\ell'}} = \bigcup_{\mathbf{j} \in \mu_{I_{\ell'}}} A_{\mathbf{j}}.$$

Thus, we can estimate

$$\begin{aligned} \sum_{\mathbf{i} \in \mathcal{M}_{I_{\ell'}}} \zeta_{\mathbf{i}} &\leq \sum_{\mathbf{j} \in \mu_{I_{\ell'}}} \sum_{\mathbf{i} \in A_{\mathbf{j}}} \zeta_{\mathbf{i}} = \sum_{\mathbf{j} \in \mu_{I_{\ell'}}} \frac{\sum_{\mathbf{i} \in A_{\mathbf{j}}} \zeta_{\mathbf{i}}}{\sum_{\mathbf{i} \in A_{\mathbf{j}}} W_{\mathbf{i}}} \sum_{\mathbf{i} \in A_{\mathbf{j}}} W_{\mathbf{i}} = \sum_{\mathbf{j} \in \mu_{I_{\ell'}}} \mathcal{P}_{\mathbf{j}} \sum_{\mathbf{i} \in A_{\mathbf{j}}} W_{\mathbf{i}} \\ &\leq \mathcal{P}_{\mathbf{i}_{\ell'}} \sum_{\mathbf{j} \in \mu_{I_{\ell'}}} \sum_{\mathbf{i} \in A_{\mathbf{j}}} W_{\mathbf{i}} = \left(\sum_{\mathbf{i} \in A_{\mathbf{i}_{\ell'}}} \zeta_{\mathbf{i}} \right) \frac{1}{\sum_{\mathbf{i} \in A_{\mathbf{i}_{\ell'}}} W_{\mathbf{i}}} \left(\sum_{\mathbf{j} \in \mu_{I_{\ell'}}} \sum_{\mathbf{i} \in A_{\mathbf{j}}} W_{\mathbf{i}} \right), \end{aligned}$$

where in the second inequality we have used the fact that $\mathcal{P}_{\mathbf{i}_{\ell'}} \geq \mathcal{P}_{\mathbf{j}}$ for any $\mathbf{j} \in \mu_{I_{\ell'}}$. Let us now estimate each of the three factors separately.

- $\sum_{\mathbf{i} \in A_{\mathbf{i}_{\ell'}}} \zeta_{\mathbf{i}}$: As in the proof of Theorem 2.18 (using the estimate from Proposition 2.5 instead of the one in (39)) we obtain with $\ell^{\frac{1}{N}} \leq \bar{r} \leq \ell + 1$ that

$$(49) \quad \sum_{\mathbf{i} \in A_{\mathbf{i}_{\ell'}}} \zeta_{\mathbf{i}} \lesssim N \ell^2 e^{-\frac{\sigma}{2}m\left(\ell^{\frac{1}{N}}\right)}.$$

- $\sum_{\mathbf{i} \in A_{\mathbf{i}_{\ell'}}} W_{\mathbf{i}}$: There holds

$$(50) \quad \sum_{\mathbf{i} \in A_{\mathbf{i}_{\ell'}}} W_{\mathbf{i}} \geq W_{\mathbf{i}_{\ell'}} \geq m(\langle \mathbf{i}_{\ell'}, \mathbf{e}_{\bar{n}} \rangle) - m(\langle \mathbf{i}_{\ell'}, \mathbf{e}_{\bar{n}} \rangle - 1) \geq 2^{\bar{r}-2} \geq 2^{\frac{\ell}{N}-2}$$

- $\sum_{\mathbf{j} \in \mu_{I_{\ell'}}} \sum_{\mathbf{i} \in A_{\mathbf{j}}} W_{\mathbf{i}}$: We observe

$$\sum_{\mathbf{j} \in \mu_{I_{\ell'}}} \sum_{\mathbf{i} \in A_{\mathbf{j}}} W_{\mathbf{i}} = \sum_{\mathbf{i} \in \mathcal{M}_{I_{\ell'}}} \#\{\mathbf{j} \in \mu_{I_{\ell'}} : \mathbf{i} \in A_{\mathbf{j}}\} W_{\mathbf{i}}.$$

Thus, being $\#\{\mathbf{j} \in \mu_{I_{\ell'}} : \mathbf{i} \in A_{\mathbf{j}}\} \leq \#\mathcal{M}_{I_{\ell'}}$, we can estimate

$$(51) \quad \sum_{\mathbf{j} \in \mu_{I_{\ell'}}} \sum_{\mathbf{i} \in A_{\mathbf{j}}} W_{\mathbf{i}} \leq \#\mathcal{M}_{I_{\ell'}} \sum_{\mathbf{i} \in \mathcal{M}_{I_{\ell'}}} W_{\mathbf{i}} \leq (\#\mathcal{M}_{I_{\ell'}})^2 \max_{\mathbf{i} \in \mathcal{M}_{I_{\ell'}}} W_{\mathbf{i}}.$$

An estimate for $\#\mathcal{M}_{I_{\ell'}}$ is given in (31). For the second factor, use the bound on $W_{\mathbf{i}}$ from (40) and the fact that for any $\mathbf{i} \in \mathcal{M}_{I_{\ell}}, |\mathbf{i}|_1 \leq N + \ell$ to obtain:

$$(52) \quad \sum_{\mathbf{j} \in \mu_{I_{\ell'}}} \sum_{\mathbf{i} \in A_{\mathbf{j}}} W_{\mathbf{i}} \leq \left(N + N(\ell - 1) \left(1 + \frac{\ell - 1}{N} \right)^N \right)^2 2^{\ell}.$$

Finally, the statement of the theorem is obtained combining (49), (50) and (52). \square

Remark 2.20. *We note that the convergence rates in Theorems 2.18–2.19 above compare the error to the number of adaptive steps ℓ . This is hard to compare to classical a priori results which bound the error in terms of the number of collocation points (see, e.g., [39, 2]). Due to the adaptive nature of the algorithm we have no knowledge about the shape of I_{ℓ} and hence the number of collocation points $\#\mathcal{H}_{I_{\ell}}$. Additionally, we do not assume any a priori information about the anisotropy of the solution. Hence, the term $\ell^{1/N}$ is the worst-case for a fully isotropic solution. We point out that the observed rate of convergence is much better (see Section 4) and further research is required to explain the performance of the adaptive algorithm.*

3. CONVERGENCE OF THE FULLY DISCRETE ALGORITHM

In order to prove the convergence of Algorithm 1, it is sufficient to prove that

- in Algorithm 2 (the finite element refinement sub-routine) the finite element error eventually falls below the tolerance prescribed in Line 3 (Alg. 2) and iteratively updated in Line 12 (Alg. 2) (proved in Section 3.1)
- that the parametric estimator $\zeta_{SC, I_{\ell}}$ in Algorithm 1 vanishes (proved in Section 3.2).

Indeed, if this is the case, $\eta_{FE, I_{\ell}}$ will vanish with $\zeta_{SC, I_{\ell}}$ because of the definition of the finite element refinement tolerance and the reliability of the estimator will ensure the convergence of the discrete solution to the analytic one.

In the present section, we will write $\zeta_{SC, I}(\cdot), \zeta_{\mathbf{i}, I}(\cdot)$ to denote the dependence on the function explicitly. The same will be done for the finite element estimator $\eta_{FE, I}(\cdot)$. For instance, the parametric estimator from Section 1.3 (defined in (10)) can be written as $\zeta_{SC, I}(U)$, if we denote by U the current discrete finite element solution. In the previous section, in which we assumed to be able to sample the analytic solution, we were dealing with $\zeta_{SC, I}(u)$.

The following lemma will be used in the next sections.

Lemma 3.1. *Given a downward-closed multi-index set $I \subset \mathbb{N}^N$, there holds*

$$|\zeta_{SC, I}(u) - \zeta_{SC, I}(U)| \lesssim \left(\sum_{\mathbf{i} \in \mathcal{M}_I} \Lambda_{\mathbf{i}} \right) \eta_{FE, I}(U).$$

Proof. The stability bound (19) for the hierarchical surplus operator implies

$$\begin{aligned} |\zeta_{SC,I}(u) - \zeta_{SC,I}(U)| &\leq \sum_{i \in \mathcal{M}_I} |\zeta_{i,I}(u) - \zeta_{i,I}(U)| \\ &\leq \sum_{i \in \mathcal{M}_I} \left\| \Delta^{m(i)} (a \nabla S_I[u - U]) \right\|_{L^\infty(\Gamma, L^2(D))} \\ &\lesssim \left(\sum_{i \in \mathcal{M}_I} \Lambda_i \right) \|\nabla S_I[u - U]\|_{L^\infty(\Gamma, L^2(D))}. \end{aligned}$$

Now we only need to bound the last factor with the finite element estimator:

$$\begin{aligned} \|\nabla S_I[u - U]\|_{L^\infty(\Gamma, L^2(D))} &\leq \sum_{\mathbf{y} \in \mathcal{H}_I} \|(u(\mathbf{y}) - U_{\mathbf{y}}) L_{\mathbf{y}}\|_{L^\infty(\Gamma, V)} \\ &\leq \sum_{\mathbf{y} \in \mathcal{H}_I} \|\nabla(u(\mathbf{y}) - U_{\mathbf{y}})\|_{L^2(D)} \|L_{\mathbf{y}}\|_{L^\infty(\Gamma)}. \end{aligned}$$

The reliability of the residual-based error estimator in each collocation point \mathbf{y} concludes the proof. \square

3.1. Convergence under h-refinement. The stochastic collocation finite element algorithm (Algorithm 1) delegates to Algorithm 2 the task of refining the finite element solutions corresponding to the collocation points until the finite element a-posteriori estimator defined in (11) falls below a given tolerance. Recall that Algorithm 2 is given a multi-index set I , or equivalently a sparse grid \mathcal{H}_I consisting of N_c collocation points that will *not* change during its execution. Hence, we will drop the index I in the following. Moreover, the index $\ell \in \mathbb{N}$ will denote the current iteration of the adaptive loop starting at Line 4 (Alg. 2) (so $U_{\ell, \mathbf{y}}$ and $\eta_{\ell, \mathbf{y}}$ will denote respectively the finite element solution and finite element estimator on the collocation point $\mathbf{y} \in \mathcal{H}$ at iteration ℓ and with respect to the mesh $\mathcal{T}_{\ell, \mathbf{y}}$).

From the theory of the classical h-adaptive finite element algorithm, we have the following contraction property (see, e.g., [8, 43, 7]) for all $\mathbf{y} \in \mathcal{H}$:

$$(53) \quad \sum_{T \in \mathcal{T}_{\ell+1, \mathbf{y}} \setminus \mathcal{T}_{\ell, \mathbf{y}}} \eta_{\ell+1, \mathbf{y}, T}^2 \leq q \sum_{T \in \mathcal{T}_{\ell, \mathbf{y}} \setminus \mathcal{T}_{\ell+1, \mathbf{y}}} \eta_{\ell, \mathbf{y}, T}^2 + C \|U_{\ell+1, \mathbf{y}} - U_{\ell, \mathbf{y}}\|_V^2$$

as well as

$$(54) \quad \left(\sum_{T \in \mathcal{T}_{\ell+1, \mathbf{y}} \cap \mathcal{T}_{\ell, \mathbf{y}}} \eta_{\ell+1, \mathbf{y}, T}^2 \right)^{1/2} \leq \left(\sum_{T \in \mathcal{T}_{\ell, \mathbf{y}} \cap \mathcal{T}_{\ell+1, \mathbf{y}}} \eta_{\ell, \mathbf{y}, T}^2 \right)^{1/2} + C^{1/2} \|U_{\ell+1, \mathbf{y}} - U_{\ell, \mathbf{y}}\|_V$$

for $0 < q < 1$ and $C > 0$ independent of ℓ but depending on the shape-regularity of the mesh and the regularity assumptions on the coefficient $a(\mathbf{y}, \cdot)$ on $\mathcal{T}_{\text{init}}$. Since we use newest-vertex-bisection for mesh refinement, the shape regularity depends only on $\mathcal{T}_{\text{init}}$.

As in the deterministic setting, Dörfler marking together with (53)–(54) can be used to prove a contraction property of the estimator (see also [6] for a similar argument with a slightly different marking strategy).

Proposition 3.2. *Given an arbitrary downward closed index set $I \subseteq \mathbb{N}^{\mathbb{N}}$, Algorithm 2 satisfies*

$$(55) \quad \sum_{\mathbf{y} \in \mathcal{H}} \eta_{\ell+k, \mathbf{y}}^2 \leq C_{\text{lin}} q_{\text{lin}}^k \sum_{\mathbf{y} \in \mathcal{H}} \eta_{\ell, \mathbf{y}}^2$$

for all $\ell, k \in \mathbb{N}$ and some uniform constants $0 < q_{\text{lin}} < 1$, $C_{\text{lin}} > 0$. In particular, we have:

$$\lim_{\ell \rightarrow \infty} \|S_I[u] - S_I[U_\ell]\|_{L^\infty(\Gamma, V)} = 0 = \lim_{\ell \rightarrow \infty} \eta_{\text{FE}, I}(U_\ell).$$

Proof. We show with (53)–(54) that all $\delta > 0$ satisfy (recall the definition of \sqcup from Line 5 (Alg. 2))

$$\begin{aligned} \sum_{\mathbf{y} \in \mathcal{H}} \eta_{\ell+1, \mathbf{y}}^2 &= \sum_{(\mathbf{y}, T_{\mathbf{y}}) \in \sqcup_{\mathbf{y} \in \mathcal{H}} \overline{\mathcal{T}}_{\ell+1, \mathbf{y}} \setminus \overline{\mathcal{T}}_{\ell, \mathbf{y}}} \eta_{\ell+1, \mathbf{y}, T_{\mathbf{y}}}^2 + \sum_{(\mathbf{y}, T_{\mathbf{y}}) \in \sqcup_{\mathbf{y} \in \mathcal{H}} \overline{\mathcal{T}}_{\ell+1, \mathbf{y}} \cap \overline{\mathcal{T}}_{\ell, \mathbf{y}}} \eta_{\ell+1, \mathbf{y}, T_{\mathbf{y}}}^2 \\ &\leq q \sum_{(\mathbf{y}, T_{\mathbf{y}}) \in \sqcup_{\mathbf{y} \in \mathcal{H}} \overline{\mathcal{T}}_{\ell, \mathbf{y}} \setminus \overline{\mathcal{T}}_{\ell+1, \mathbf{y}}} \eta_{\ell, \mathbf{y}, T_{\mathbf{y}}}^2 + (1 + \delta) \sum_{(\mathbf{y}, T_{\mathbf{y}}) \in \sqcup_{\mathbf{y} \in \mathcal{H}} \overline{\mathcal{T}}_{\ell, \mathbf{y}} \cap \overline{\mathcal{T}}_{\ell+1, \mathbf{y}}} \eta_{\ell, \mathbf{y}, T_{\mathbf{y}}}^2 \\ &\quad + C(2 + \delta^{-1}) \sum_{\mathbf{y} \in \mathcal{H}} \|U_{\ell+1, \mathbf{y}} - U_{\ell, \mathbf{y}}\|_V^2 \\ &\leq (q - 1) \sum_{(\mathbf{y}, T_{\mathbf{y}}) \in \sqcup_{\mathbf{y} \in \mathcal{H}} \overline{\mathcal{T}}_{\ell, \mathbf{y}} \setminus \overline{\mathcal{T}}_{\ell+1, \mathbf{y}}} \eta_{\ell, \mathbf{y}, T_{\mathbf{y}}}^2 + (1 + \delta) \sum_{(\mathbf{y}, T_{\mathbf{y}}) \in \sqcup_{\mathbf{y} \in \mathcal{H}} \overline{\mathcal{T}}_{\ell, \mathbf{y}}} \eta_{\ell+1, \mathbf{y}, T_{\mathbf{y}}}^2 \\ &\quad + C(2 + \delta^{-1}) \sum_{\mathbf{y} \in \mathcal{H}} \|U_{\ell+1, \mathbf{y}} - U_{\ell, \mathbf{y}}\|_V^2. \end{aligned}$$

The Dörfler marking from Algorithm 2 ensures $\mathcal{K} \subseteq \sqcup_{\mathbf{y} \in \mathcal{H}} \overline{\mathcal{T}}_{\ell, \mathbf{y}} \setminus \overline{\mathcal{T}}_{\ell+1, \mathbf{y}}$ and hence

$$(q - 1) \sum_{(\mathbf{y}, T_{\mathbf{y}}) \in \sqcup_{\mathbf{y} \in \mathcal{H}} \overline{\mathcal{T}}_{\ell, \mathbf{y}} \setminus \overline{\mathcal{T}}_{\ell+1, \mathbf{y}}} \eta_{\ell, \mathbf{y}, T_{\mathbf{y}}}^2 \leq \theta(q - 1) \sum_{(\mathbf{y}, T_{\mathbf{y}}) \in \sqcup_{\mathbf{y} \in \mathcal{H}} \overline{\mathcal{T}}_{\ell, \mathbf{y}}} \eta_{\ell, \mathbf{y}, T_{\mathbf{y}}}^2.$$

Altogether, we obtain for $\kappa := 1 + \delta - \theta(1 - q)$ and $\tilde{C} := C(2 + \delta^{-1})$ that

$$\sum_{\mathbf{y} \in \mathcal{H}} \eta_{\ell+1, \mathbf{y}}^2 \leq \kappa \sum_{\mathbf{y} \in \mathcal{H}} \eta_{\ell, \mathbf{y}}^2 + \tilde{C} \sum_{\mathbf{y} \in \mathcal{H}} \|U_{\ell+1, \mathbf{y}} - U_{\ell, \mathbf{y}}\|_V^2.$$

With the Galerkin orthogonality

$$\begin{aligned} &\sum_{\mathbf{y} \in \mathcal{H}} \left\| a(\mathbf{y})^{1/2} \nabla (U_{\ell+1, \mathbf{y}} - U_{\ell, \mathbf{y}}) \right\|_{L^2(D)}^2 \\ &= \sum_{\mathbf{y} \in \mathcal{H}} \left(\left\| a(\mathbf{y})^{1/2} \nabla (u(\mathbf{y}) - U_{\ell, \mathbf{y}}) \right\|_{L^2(D)}^2 - \left\| a(\mathbf{y})^{1/2} \nabla (u(\mathbf{y}) - U_{\ell+1, \mathbf{y}}) \right\|_{L^2(D)}^2 \right) \end{aligned}$$

we may follow [7, Section 4] verbatim in order to prove (55). Since $\#\mathcal{H}$ is fixed, we have $\sum_{\mathbf{y} \in \mathcal{H}} \eta_{\ell, \mathbf{y}}^2 \simeq \eta_{\text{FE}, I}(U_\ell)^2$ and reliability proves $\lim_{\ell \rightarrow \infty} \eta_{\text{FE}, I}(U_\ell) = \lim_{\ell \rightarrow \infty} \|S_I[u] - S_I[U_\ell]\|_{L^\infty(\Gamma, V)} = 0$. This concludes the statement. \square

Remark 3.3. *The previous proposition implies that Algorithm 2 terminates. In particular, the algorithm will eventually satisfy the condition $\eta_{\text{FE}, I}(U_\ell) < \text{Tol}_\ell$, where $\text{Tol}_\ell := \alpha \frac{1}{\sum_{i \in \mathcal{M}_I} \Lambda_i} \zeta_{\text{SC}, I}(U_\ell)$. Indeed, due to Lemma 3.1 we have that, as $(\eta_{\text{FE}, I_\ell}(U_\ell))_\ell$ vanishes, $\zeta_{\text{SC}, I}(U_\ell)$ converges to $\zeta_{\text{SC}, I}(u) \geq \epsilon > 0$, therefore $\lim_{\ell \rightarrow \infty} \text{Tol}_\ell = \alpha \frac{1}{\sum_{i \in \mathcal{M}_I} \Lambda_i} \zeta_{\text{SC}, I}(u) > 0$. Note that the convergence proof uses an ℓ_2 -type estimator instead of an ℓ_1 -type as in $\eta_{\text{FE}, I}$. In this regard, the ℓ_2 -type might seem more natural and we refer to Section 3.3 for further discussion.*

Theorem 3.4. *Given an arbitrary downward closed index set $I \subseteq \mathbb{N}^{\mathbb{N}}$, Algorithm 2 converges with the optimal rate in the following sense: Let \mathbb{T} denote the set of all meshes which can be obtained from $\mathcal{T}_{\text{init}}$ by iterated newest-vertex-bisection with mesh closure. Let $s > 0$ such that*

$$(56) \quad \sup_{N \in \mathbb{N}} \inf_{\substack{\mathcal{T}_{\mathbf{y}} \in \mathbb{T} \\ \sum_{\mathbf{y} \in \mathcal{H}} \#\mathcal{T}_{\mathbf{y}} \leq N}} \left(\sum_{\mathbf{y} \in \mathcal{H}} \|u(\mathbf{y}) - U_{\mathcal{T}_{\mathbf{y}}}\|_V^2 + \|h_{\mathcal{T}_{\mathbf{y}}}(1 - \Pi_{\mathcal{T}_{\mathbf{y}}})f\|_{L^2(D)}^2 \right)^{1/2} N^s < \infty,$$

where $h_{\mathcal{T}}$ denotes the local mesh-size function and $\Pi_{\mathcal{T}}$ is the $L^2(D)$ -orthogonal projection onto \mathcal{T} -elementwise constant functions. Then, there holds

$$\sup_{\ell \in \mathbb{N}} \|S_I[u] - S_I[U_\ell]\|_{L^\infty(\Gamma, V)} \left(\sum_{\mathbf{y} \in \mathcal{H}} \#\mathcal{T}_{\ell, \mathbf{y}} \right)^s < \infty.$$

Proof. First note that standard upper/lower bounds for the residual error estimator together with the regularity assumptions on $a(\mathbf{y}, \cdot)$ show $\eta_{\mathbf{y}} \simeq \sqrt{\|u(\mathbf{y}) - U_{\mathcal{T}_{\mathbf{y}}}\|_V^2 + \|h_{\mathcal{T}_{\mathbf{y}}}(1 - \Pi_{\mathcal{T}_{\mathbf{y}}})f\|_{L^2(D)}^2}$ and hence (56) is equivalent to

$$\sup_{N \in \mathbb{N}} \inf_{\substack{\mathcal{T}_{\mathbf{y}} \in \mathcal{T} \\ \sum_{\mathbf{y} \in \mathcal{H}} \#\mathcal{T}_{\mathbf{y}} \leq N}} \left(\sum_{\mathbf{y} \in \mathcal{H}} \eta_{\mathbf{y}}^2 \right)^{1/2} N^s < \infty.$$

With the error norm $\|u - U_\ell\| := \sqrt{\sum_{\mathbf{y} \in \mathcal{H}} \|u(\mathbf{y}) - U_{\mathcal{T}_{\ell, \mathbf{y}}}\|_V^2}$ and (53)–(54), the estimator $\eta_{\text{FE}, I}$ satisfies (A1) and (A2) from [7, Section 3]. From the classical theory of h -adaptivity [8], we immediately obtain discrete reliability (A3) in the sense

$$\|U_{\ell+k} - U_\ell\|^2 = \sum_{\mathbf{y} \in \mathcal{H}} \|U_{\ell+k, \mathbf{y}} - U_{\ell, \mathbf{y}}\|_V^2 \leq C_{\text{drel}} \sum_{\mathbf{y} \in \mathcal{H}} \sum_{T \in \omega(\mathcal{T}_{\ell, \mathbf{y}} \setminus \mathcal{T}_{\ell+k, \mathbf{y}})} \eta_{\ell, \mathbf{y}, T}^2,$$

where $\omega(\cdot)$ denotes the set of elements with non-empty intersection with (\cdot) . With these ingredients and the linear convergence from Proposition 3.2, [7, Proposition 4.12 & Proposition 4.15] show optimal convergence of the error estimator

$$\sup_{\ell \in \mathbb{N}} \sqrt{\sum_{\mathbf{y} \in \mathcal{H}} \eta_{\ell, \mathbf{y}}^2} \left(\sum_{\mathbf{y} \in \mathcal{H}} \#\mathcal{T}_{\ell, \mathbf{y}} \right)^s < \infty.$$

With constants depending only on the size of I , the quantity $\sqrt{\sum_{\mathbf{y} \in \mathcal{H}} \eta_{\ell, \mathbf{y}}^2}$ is equivalent to $\eta_{\text{FE}, I}$ and hence reliability concludes the proof. \square

3.2. Proof of convergence of the fully discrete algorithm. The tolerance for finite element refinement was defined in Algorithm 2 as:

$$(57) \quad \text{Tol} = \text{Tol}(I, \zeta_{i, I}(U), \alpha) := \alpha \frac{1}{\sum_{i \in \mathcal{M}_I} \Lambda_i} \zeta_{SC, I}(U).$$

where $\alpha \in (0, 1)$, Λ_i was defined in (20) and $\zeta_{SC, I}(U)$ is the parametric a-posteriori error estimator. This choice is motivated by the following estimate: For fixed downward closed $I \subset \mathbb{N}^N$, Lemma 3.1 shows

$$\zeta_{SC, I}(U) \leq \zeta_{SC, I}(u) + \left(\sum_{i \in \mathcal{M}_I} \Lambda_i \right) \eta_{\text{FE}, I}(U) \leq \zeta_{SC, I}(u) + \alpha \zeta_{SC, I}(U),$$

and hence

$$(58) \quad \zeta_{SC, I}(U) \leq \frac{1}{1 - \alpha} \zeta_{SC, I}(u).$$

In the context of the adaptive algorithm, this implies that $(\zeta_{SC, I_\ell}(U_\ell))_\ell$ is uniformly bounded since $(\zeta_{SC, I_\ell}(u))_\ell$ is. This last fact was proved in Remark 2.6 using the estimate on the pointwise error estimator from Proposition 2.5.

Lemma 3.5. *Algorithm 1 with either workless profit (and $0 < \alpha < 1$ sufficiently small) or profit with work (and arbitrary $0 < \alpha < 1$) and the tolerance (57) satisfies $\lim_{\ell \rightarrow \infty} \mathcal{P}_{i_\ell, I_\ell} = 0$.*

Proof. We consider the two definitions of profit separately:

$$\text{Profit with work: } \mathcal{P}_{\mathbf{i}, I} := \frac{\sum_{j \in A_{\mathbf{i}, I}} \zeta_{j, I}(U)}{\sum_{j \in A_{\mathbf{i}, I}} W_j}.$$

The uniform boundedness of the parametric a-posteriori error estimator, together with the fact that the work over $A_{\mathbf{i}_\ell, I_\ell}$ diverges, gives

$$\mathcal{P}_{\mathbf{i}_\ell, I_\ell} \leq \frac{\zeta_{SC, I_\ell}(U_\ell)}{\sum_{j \in A_{\mathbf{i}_\ell, I_\ell}} W_j} \lesssim \frac{1}{\sum_{j \in A_{\mathbf{i}_\ell, I_\ell}} W_j} \rightarrow 0.$$

Workless profit: $\mathcal{P}_{\mathbf{i}, I} := \sum_{j \in A_{\mathbf{i}, I}} \zeta_{j, I}(U)$. We recall from (48) that, for the profit-maximizer $\mathbf{i}_\ell \in \mathcal{M}_{I_\ell}$, $\mathcal{P}_{\mathbf{i}_\ell, I_\ell} \geq \frac{1}{N} \zeta_{SC, I_\ell}(U)$. Thus, Lemma 3.1 shows

$$\begin{aligned} \mathcal{P}_{\mathbf{i}_\ell, I_\ell} &\leq \sum_{j \in A_{\mathbf{i}_\ell, I_\ell}} \zeta_{j, I_\ell}(u) + \alpha \frac{\sum_{j \in A_{\mathbf{i}_\ell, I_\ell}} \Lambda_j}{\sum_{j \in \mathcal{M}_{I_\ell}} \Lambda_j} \zeta_{SC, I_\ell}(U_\ell) \\ &\leq \sum_{j \in A_{\mathbf{i}_\ell, I_\ell}} \zeta_{j, I_\ell}(u) + \alpha \frac{\sum_{j \in A_{\mathbf{i}_\ell, I_\ell}} \Lambda_j}{\sum_{j \in \mathcal{M}_{I_\ell}} \Lambda_j} N \mathcal{P}_{\mathbf{i}_\ell, I_\ell} \\ &\leq \sum_{j \in A_{\mathbf{i}_\ell, I_\ell}} \zeta_{j, I_\ell}(u) + \alpha N \mathcal{P}_{\mathbf{i}_\ell, I_\ell}, \end{aligned}$$

so

$$\mathcal{P}_{\mathbf{i}_\ell, I} \leq \frac{1}{1 - \alpha N} \sum_{j \in A_{\mathbf{i}_\ell, I_\ell}} \zeta_{j, I_\ell}(u) \rightarrow 0 \quad \text{as } l \rightarrow \infty.$$

Observe that this introduces the constraint on α with respect to the number of dimensions: $\alpha < N^{-1}$. This constraint can be improved by replacing the crude estimate

$$\frac{\sum_{j \in A_{\mathbf{i}_\ell, I_\ell}} \Lambda_j}{\sum_{j \in \mathcal{M}_{I_\ell}} \Lambda_j} \leq 1,$$

with the better bound

$$\alpha \leq \left(\frac{\max_{n \in \{1, \dots, N\}} \sum_{j \in A_{\mathbf{i}_{\ell-1} + e_n, I_\ell}} \Lambda_j}{\sum_{j \in \mathcal{M}_{I_\ell}} \Lambda_j} N \right)^{-1}.$$

This concludes the proof. \square

We can finally prove that the error estimator vanishes with a technique similar to that used in Theorem 2.15 for the parametric algorithm.

Theorem 3.6. *Algorithm 1 with either workless profit (and $0 < \alpha < 1$ sufficiently small) or profit with work (and arbitrary $0 < \alpha < 1$) and the tolerance (57) satisfies the following: The sequence of parametric a-posteriori error estimators $(\zeta_{SC, I_\ell}(U_\ell))_\ell$ vanishes*

$$\lim_{\ell \rightarrow \infty} \zeta_{SC, I_\ell}(U_\ell) = 0.$$

Thus, also the finite element error estimator vanishes

$$\lim_{\ell \rightarrow \infty} \eta_{FE, I_\ell}(U_\ell) = 0,$$

and the reliability of the a-posteriori error estimator implies error convergence

$$\lim_{\ell \rightarrow \infty} \|u - S_{I_\ell}[U_\ell]\|_{L^\infty(\Gamma, V)} = 0.$$

Proof. The a-posteriori error estimator can be expressed as

$$\zeta_{SC,I_\ell}(U_\ell) = \sum_{\mathbf{i} \in \mathbb{N}^N} \zeta_{\mathbf{i},I_\ell}(U_\ell) \mathbb{1}_{\mathcal{M}_{I_\ell}}(\mathbf{i}).$$

Since the sequence $(\zeta_{SC,I_\ell}(U_\ell))_\ell$ is uniformly bounded (58), it is sufficient to prove that $(\zeta_{\mathbf{i},I_\ell}(U_\ell) \mathbb{1}_{\mathcal{M}_{I_\ell}})_\ell$ vanishes for any fixed $\mathbf{i} \in \mathbb{N}^N$. We can distinguish three cases:

- if \mathbf{i} is eventually added to I_ℓ , then $\mathbb{1}_{\mathcal{M}_{I_\ell}}(\mathbf{i})$ is eventually zero;
- if \mathbf{i} is never added to the margin \mathcal{M}_{I_ℓ} , then $\zeta_{\mathbf{i},I_\ell}(U_\ell)$ is constantly zero;
- finally, if it exists $\bar{\ell} \in \mathbb{N}$ such that for all $\ell > \bar{\ell}$, $\mathbf{i} \in \mathcal{M}_{I_\ell}$, then $\lim_{\ell \rightarrow \infty} \zeta_{\mathbf{i},I_\ell}(U_\ell) = 0$. Indeed, because of Lemma 3.5, $\lim_{\ell \rightarrow \infty} \mathcal{P}_{\mathbf{i},I_\ell} = 0$ (for both workless profit and profit with work), thus $(\zeta_{\mathbf{i},I_\ell}(U_\ell))_\ell$ vanishes as in Proposition 2.13.

This concludes the proof. \square

3.3. Other versions of the finite element estimator. In the previous section we followed [32] to derive the estimator via

$$\begin{aligned} S_I \left[\int_D f v - a \nabla S_I[U] \cdot \nabla v \right] &= \sum_{\mathbf{y} \in \mathcal{H}_I} \left[\int_D f v - a(\mathbf{y}) \nabla S_I[U](\mathbf{y}) \cdot \nabla v \right] L_{\mathbf{y}} \\ &\leq C \sum_{\mathbf{y} \in \mathcal{H}_I} \eta_{\mathbf{y}} |L_{\mathbf{y}}| \|\nabla v\|_{L^2(D)}. \end{aligned}$$

Choosing $v = u - S_I[U]$ and taking the $L^\infty(\Gamma)$ norm leads to the estimator we used above, i.e.,

$$\left\| \sum_{\mathbf{y} \in \mathcal{H}_I} \eta_{\mathbf{y}} |L_{\mathbf{y}}| \right\|_{L^\infty(\Gamma)} \leq \sum_{\mathbf{y} \in \mathcal{H}_I} \eta_{\mathbf{y}} \|L_{\mathbf{y}}\|_{L^\infty(\Gamma)} = \eta_{\text{FE},I}(U).$$

Using the Hölder estimates with other combinations of $(p, q) \in \{(2, 2), (\infty, 1)\}$, we obtain

$$\left\| \sum_{\mathbf{y} \in \mathcal{H}_I} \eta_{\mathbf{y}} |L_{\mathbf{y}}| \right\|_{L^\infty(\Gamma)} \leq \left\| \left(\sum_{\mathbf{y} \in \mathcal{H}_I} \eta_{\mathbf{y}}^2 \right)^{\frac{1}{2}} \left(\sum_{\mathbf{y} \in \mathcal{H}_I} |L_{\mathbf{y}}|^2 \right)^{\frac{1}{2}} \right\|_{L^\infty(\Gamma)} = \eta_{p,I} \Lambda_{q,I},$$

where

$$\eta_{p,I} := \begin{cases} \left(\sum_{\mathbf{y} \in \mathcal{H}_I} \eta_{\mathbf{y}}^2 \right)^{\frac{1}{2}} & p = 2, \\ \max_{\mathbf{y} \in \mathcal{H}_I} \eta_{\mathbf{y}} & p = \infty, \end{cases} \quad \text{and} \quad \Lambda_{q,I} := \begin{cases} \left\| \left(\sum_{\mathbf{y} \in \mathcal{H}_I} |L_{\mathbf{y}}|^2 \right)^{\frac{1}{2}} \right\|_{L^\infty(\Gamma)} & q = 2, \\ \left\| \sum_{\mathbf{y} \in \mathcal{H}_I} |L_{\mathbf{y}}| \right\|_{L^\infty(\Gamma)} & q = 1. \end{cases}$$

The perturbation result from Lemma 3.1 can be analogously modified to obtain:

$$|\zeta_{SC,I}(u) - \zeta_{SC,I}(U)| \lesssim \left(\sum_{\mathbf{i} \in \mathcal{M}_I} \Lambda_{\mathbf{i}} \right) \eta_{p,I} \Lambda_{q,I}$$

From these results, the sufficient condition in (57) for convergence becomes respectively

$$\eta_{p,I}(U) \leq \alpha \left(\Lambda_{q,I} \sum_{\mathbf{i} \in \mathcal{M}_I} \Lambda_{\mathbf{i}} \right)^{-1} \zeta_{SC,I}(U).$$

With these ingredients, all the other results of the previous sections hold for the variants of the finite element estimator discussed above.

3.4. Convergence of a single mesh version of the fully discrete algorithm. We also consider SCFE with the same *adaptively* refined mesh in all collocation points. The idea is that, if the set of singularities of the solution u is small, one single adaptive mesh can resolve all of them simultaneously and thus substantially reduce the computational effort. We employ the following estimator from [32, Remark 4.4] for the finite element part

$$\begin{aligned} \eta_{\text{FE},I}(U) &:= \left(\sum_{T \in \mathcal{T}} \eta_T^2(U) \right)^{1/2}, & \eta_T(U) &:= \|\eta_T(\cdot; U)\|_{L^\infty(\Gamma)}, \\ \eta_T^2(\mathbf{y}; U) &:= h_T^2 \|S_I [f + \nabla \cdot (a \nabla U)](\mathbf{y})\|_{L^2(T)}^2 + \sum_{e \subset \partial T} h_T \|S_I [[a \nabla U \cdot \mathbf{n}_e]_{\mathbf{n}_e}](\mathbf{y})\|_{L^2(e)}^2. \end{aligned}$$

Since we use a single mesh for all collocation points $\mathbf{y} \in \mathcal{H}_I$, we replace $\mathcal{T}_{\mathbf{y}}$ in Algorithm 2 by \mathcal{T} . We change the Dörfler marking in Line 5 (Alg. 2) to: Find minimal $\mathcal{K} \subseteq \mathcal{T}$ such that

$$\sum_{T \in \mathcal{K}} \eta_T(U)^2 \geq \theta \eta_{\text{FE},I}^2.$$

Moreover, we replace the refinement loop in Line 6 (Alg. 2) by a single refinement of the mesh \mathcal{T} with marked elements \mathcal{K} .

Due to the fact that $U: \Gamma \rightarrow \mathcal{S}_0^1(\mathcal{T})$ admits a holomorphic extension to $\Sigma(\Gamma, \boldsymbol{\tau})$ just as does u (the same arguments work also for the discrete approximation), the convergence analysis of the parametric enrichment algorithm remains unchanged (Section 2), we now have to show convergence of the adaptive finite element subroutine. With this, we may analogously employ the results of Section 3.2 to obtain convergence of the full algorithm. Note that we can not directly transfer the proof of Proposition 3.2 as the definition of $\eta_{\text{FE},I}$ in this section mixes L^2 -norms and L^∞ -norms.

In this setting, the multi-index set $I \subset \mathbb{N}^N$ is fixed. We denote by \mathcal{T}_ℓ the finite element mesh at step $\ell > 0$ (the same for every collocation point). U_ℓ represents the discrete solution at step ℓ and $U_{\ell, \mathbf{y}} \in \mathcal{S}_0^1(\mathcal{T}_\ell)$ its value on a collocation point $\mathbf{y} \in \mathcal{H}_I$. We simplify the notation for the estimator as $\eta_\ell := \eta_{\text{FE},I}(U_\ell)$, $\eta_\ell(U) := \left(\sum_{T \in \mathcal{T}_\ell} \eta_T(U)^2 \right)^{1/2}$.

We first give a perturbation estimate localized on one element T of a mesh \mathcal{T} , analogously to [8, Proposition 3.3].

Lemma 3.7. *Consider a shape-regular mesh \mathcal{T} obtained by NVB from a mesh $\mathcal{T}_{\text{init}}$. There holds for $U, W \in C^0(\Gamma, \mathcal{S}_0^1(\mathcal{T}))$ that*

$$(59) \quad \eta_T(U) \leq \eta_T(W) + C \|S_I\|_{\mathcal{L}(L^\infty(\Gamma, L^2(D)))} \max_{\mathbf{y} \in \mathcal{H}} \|\nabla(U(\mathbf{y}) - W(\mathbf{y}))\|_{L^2(\omega(T))} \quad \forall T \in \mathcal{T},$$

where $\omega(T)$ is the union of the elements sharing an edge with T , $C > 0$ depends only on a and $\mathcal{T}_{\text{init}}$.

Proof. For any fixed $\mathbf{y} \in \Gamma$, the linearity of S_I and the triangle inequality yield

$$\begin{aligned} \eta_T(\mathbf{y}; U) &\leq \eta_T(\mathbf{y}; W) + h_T \|S_I [\nabla \cdot (a \nabla(U - W))](\mathbf{y})\|_{L^2(T)} \\ &\quad + h_T^{1/2} \sum_{e \subset \partial T} \|S_I [[a \nabla(U - W) \cdot \mathbf{n}_e]_{\mathbf{n}_e}](\mathbf{y})\|_{L^2(e)}. \end{aligned}$$

With the operator norm of S_I , we obtain

$$\|S_I [\nabla \cdot (a \nabla(U - W))](\mathbf{y})\|_{L^2(T)} \leq \|S_I\|_{\mathcal{L}(L^\infty(\Gamma, L^2(D)))} \max_{\mathbf{y} \in \mathcal{H}} \|\nabla a(\mathbf{y})\|_{L^\infty(T)} \|\nabla(U - W)(\mathbf{y})\|_{L^2(T)}.$$

Analogously, for the jump terms $[a\nabla(U - W) \cdot \mathbf{n}_e]_{\mathbf{n}_e}$ with $e \subset \partial T$, we obtain, following the same steps as in [8, Proposition 3.3],

$$\sum_{e \subset \partial T} \|S_I [[a\nabla(U - W) \cdot \mathbf{n}_e]_{\mathbf{n}_e}](\mathbf{y})\|_{L^2(e)} \lesssim \|S_I\|_{\mathcal{L}(L^\infty(\Gamma, L^2(D)))} \max_{\mathbf{y} \in \mathcal{H}} \|a(\mathbf{y})\|_{L^\infty(\omega(T))} \|\nabla(U - W)(\mathbf{y})\|_{L^2(\omega(T))}.$$

This concludes the proof. \square

Proposition 3.8. *The sequence of finite element estimators η_ℓ obtained from the single mesh adaptive algorithm satisfies*

$$\lim_{\ell \rightarrow \infty} \eta_\ell = 0.$$

Proof. With the perturbation estimate from Lemma 3.7, we may follow [7, Section 4.3] to show estimator reduction

$$(60) \quad \eta_{\ell+1}^2 \leq q\eta_\ell^2 + C^2 \|S_I\|^2 \sum_{T \in \mathcal{T}_\ell} \max_{\mathbf{y} \in \mathcal{H}} \|\nabla(U_{\ell+1}(\mathbf{y}) - U_\ell(\mathbf{y}))\|_{L^2(\omega(T))}^2$$

for some universal $0 < q < 1$ and all $\ell \in \mathbb{N}$.

To show that the second term in (60) vanishes, we first observe that $(U_\ell(\mathbf{y}))_{\ell \in \mathbb{N}}$ converges in V for all $\mathbf{y} \in \mathcal{H}$. Indeed, for any fixed $\mathbf{y} \in \mathcal{H}$, the nestedness of the finite element spaces V_ℓ guarantees the existence of $U_\infty(\mathbf{y}) \in \overline{\bigcup_\ell V_\ell} \subset V$ such that $\lim_{\ell \rightarrow \infty} \|U_\infty(\mathbf{y}) - U_\ell(\mathbf{y})\|_V = 0$ by Céa's lemma (see, e.g., [7, Section 3.6]). This implies that

$$(61) \quad \lim_{\ell \rightarrow \infty} \|\nabla(U_{\ell+1} - U_\ell)(\mathbf{y})\|_{L^2(D)} = 0$$

for all $\mathbf{y} \in \mathcal{H}$. Since $\#\mathcal{H}$ is fixed in the finite element refinement loop of the adaptive algorithm, we have

$$\begin{aligned} \sum_{T \in \mathcal{T}_\ell} \max_{\mathbf{y} \in \mathcal{H}} \|\nabla(U_{\ell+1}(\mathbf{y}) - U_\ell(\mathbf{y}))\|_{L^2(\omega(T))}^2 &\leq \sum_{T \in \mathcal{T}_\ell} \sum_{\mathbf{y} \in \mathcal{H}} \|\nabla(U_{\ell+1}(\mathbf{y}) - U_\ell(\mathbf{y}))\|_{L^2(\omega(T))}^2 \\ &\lesssim \sum_{\mathbf{y} \in \mathcal{H}} \|\nabla(U_{\ell+1}(\mathbf{y}) - U_\ell(\mathbf{y}))\|_{L^2(D)}^2 \rightarrow 0 \end{aligned}$$

as $\ell \rightarrow \infty$. Passing to the limit superior in (60) shows $0 \leq \limsup_{\ell \rightarrow \infty} \eta_{\ell+1} \leq q \limsup_{\ell \rightarrow \infty} \eta_\ell$ and thus concludes $\lim_{\ell \rightarrow \infty} \eta_\ell = 0$. \square

Altogether, we obtain the convergence result analogously to Theorem 3.6.

Theorem 3.9. *The single mesh SCFE algorithm discussed in this section satisfies the following: The sequence of parametric a-posteriori error estimators $(\zeta_{SC, I_\ell}(U_\ell))_\ell$ vanishes*

$$\lim_{\ell \rightarrow \infty} \zeta_{SC, I_\ell}(U_\ell) = 0.$$

Thus, also the finite element error estimator vanishes

$$\lim_{\ell \rightarrow \infty} \eta_{FE, I_\ell}(U_\ell) = 0,$$

and the reliability of the a-posteriori error estimator implies error convergence

$$\lim_{\ell \rightarrow \infty} \|u - S_{I_\ell}[U_\ell]\|_{L^\infty(\Gamma, V)} = 0.$$

3.5. Cost of the stochastic collocation algorithms. Under the assumption that the pointwise estimators $\zeta_{i,I}(U)$ and $\eta_{\mathbf{y},T}(U)$ can be computed from the discrete solution in $\mathcal{O}(1)$, each step of the adaptive loop (all algorithms) is linear with respect to the number of degrees of freedom of the current sparse grid and spatial meshes. Indeed, a properly preconditioned iterative solver computes U in linear cost (depending on a_{\min} and a_{\max}). The Dörfler marking in Algorithm 2 requires sorting when done in a naive way, but can be improved to linear cost by binning [43] or by a clever variation of the quick-select algorithm [41]. Finally, the refinement of the finite element meshes $\mathcal{T}_{\mathbf{y}}$ via newest-vertex-bisection can be done in linear cost [44].

As discussed in Section 4 below, the computation of the $L^\infty(\Gamma)$ and $L^2(D)$ -norms for $\zeta_{i,I}$ is done via a random sample/Monte-Carlo procedure. This results in constant cost $\mathcal{O}(1)$ and the numerical experiments below show that the approximation error is negligible. A precise convergence analysis of this procedure would be interesting but is beyond the scope of this work. Each random sample requires the evaluation of the sparse grid interpolant. Theoretically, the cost of the evaluation of the sparse grid interpolation operator is linear in terms of collocation points, after a quadratic set-up cost. Practically, however, the cost of computing the discrete solutions is expected to dominate significantly.

4. NUMERICAL EXPERIMENTS

The Matlab implementation of Algorithm 1 used to produce the numerical results presented in this section is based on two Matlab libraries. For sparse grid algorithms, the *Sparse Grids Kit* [3] was used. The implementation of the adaptive P1 finite element methods is from the *p1afem* Matlab package [26], which uses Matlab's direct solver for sparse matrices. For further details about parameters and algorithm used within these libraries, the reader is referred to the respective documentations. The parts of the algorithm that deal with parameter enrichment (e.g. Algorithm 3) were implemented following the guidelines from [32].

In order to compute the $L^\infty(\Gamma)$ norm approximately, we consider a set Θ of 500 uniformly distributed random points in Γ and approximate, for any $g \in C^0(\Gamma)$, $\|g\|_{L^\infty(\Gamma)} \approx \max_{\mathbf{y} \in \Theta} |g(\mathbf{y})|$. The computation of the $L^2(D)$ norm is carried out with Monte Carlo integration: Given $f \in L^2(D)$, we denote by Π a set of 500 uniformly distributed random points in D and approximate $\|f\|_{L^2(D)}^2 \approx \frac{1}{\#\Pi} \sum_{x \in \Pi} f(x)^2$. The reason Monte Carlo integration is used is that for a generic $\mathbf{y} \in \Gamma$ the discrete solution $S_I[U](\mathbf{y})$ belongs to the finite element space $S_0^1(\mathcal{T})$, where \mathcal{T} is the coarsest common refinement of the meshes $\mathcal{T}_{\mathbf{y}}$. Therefore, in order to compute the exact $L^2(D)$ norm of the function, it would be necessary to compute \mathcal{T} , which would lead to a significant computational overhead. In numerical experiments, we have observed that increasing $\#\Pi$ does not lead to a significant improvement in the approximation of the $L^2(D)$ norm, thus suggesting that the approximation error can be neglected. In the numerical examples presented in the next sections, we approximate the error between the exact solution u and a discrete solution $S_I[U]$ by $\|u - S_I[U]\|_{L^\infty(\Gamma,V)} \approx \|u_{\text{approx}} - S_I[U]\|_{L^\infty(\Gamma,V)}$, where u_{approx} is a discrete solution obtained as the last iteration of the single mesh version of SCFE. To approximate the $L^\infty(\Gamma,V)$ -norm appearing in the error, we use the same method detailed above, just with $\#\Theta = \#\Pi = 5000$.

To drive parametric refinement, we employ only profits with work as defined in (15). As observed in Section 2.1.4, workless profits lead to a tensor-product interpolant and thus less interesting results. For all examples, we consider the finite element estimator is $\eta_{FE,I} = \eta_{2,I} \Lambda_{2,I}$ as defined in Section 3.3. The Dörfler parameter for refinement is chosen as $\theta = 0.7$ and, as default mesh $\mathcal{T}_{\text{init}}$, a quasi-uniform mesh with 512 triangles and 289 vertices.

In order to decrease the memory requirements of the program, the finite element refinement tolerance from Section 3.3 is modified as follows:

$$(62) \quad \text{Tol} := \alpha \Lambda_{2,I}^{-1} \zeta_{SC,I},$$

i.e. we neglect the term depending on the margin of I . In the experiments below, we observe that this choice does not compromise convergence. Further investigations will have to be carried out in order to understand whether or not the sufficient condition for convergence can be weakened. The constant α appearing in (62) is chosen as $\alpha = 0.9$. A value of α close to one shifts the balance between finite element refinement and parameter enrichment towards the latter one.

In order to improve the computational efficiency, we use the following shortcut in the implementation of Algorithm 2: Instead of re-computing the tolerance Tol at each iteration of the loop, we update it only at the end and, if needed, keep refining the finite element solutions. We alternate these two steps until the finite element estimator falls below the tolerance.

In the following two sections, we consider a physical domain $D = (0, 1)^2$ and denote $x = (x_1, x_2) \in D$. The parametric domain is $\Gamma = [-1, 1]^N$ for an integer N representing the number of parametric dimensions of the problem.

We recall that for a numerical solution $S_I[U]$ obtained with SCFE, its number of degrees of freedom is proportional to $M := \sum_{\mathbf{y} \in \mathcal{H}_I} \#\mathcal{T}_{\mathbf{y}}$, where $\#\mathcal{T}_{\mathbf{y}}$ is the number of vertices of the mesh corresponding to the collocation point \mathbf{y} (or equivalently the dimension of the finite element space $V_{\mathbf{y}}$ up to boundary conditions).

4.1. First example: Karhunen–Loève expansion with $N=5, 11$. We consider a constant forcing term $f(x) \equiv 1$ and the following diffusion coefficient with affine dependence on the parameter $\mathbf{y} \in \Gamma$:

$$(63) \quad a(x, \mathbf{y}) = a_0(x) + \frac{1}{3} \left(a_1(x)y_1 + \sum_{n=2}^N a_n(x)y_n \right),$$

where $a_0(x) \equiv 1$, $a_1(x) \equiv \left(\frac{\sqrt{\pi}L}{2}\right)^{1/2}$ and, for $n > 1$,

$$\lambda_n = (\sqrt{\pi}L)^{1/2} \exp\left(-\frac{\left(\lfloor \frac{n}{2} \rfloor \pi L\right)^2}{8}\right),$$

$$a_n(x) = \begin{cases} \sqrt{\lambda_n} \sin(n\pi x_1) & \text{if } n \text{ even} \\ \sqrt{\lambda_n} \cos(n\pi x_1) & \text{if } n \text{ odd,} \end{cases}$$

where $L \in (0, 1)$ is a constant. Such a diffusion coefficient is the result of the Karhunen–Loève expansion [42] of the random field $a(x, \omega)$ with mean a_0 and covariance $Cov(x, x') = \frac{1}{3^2} \exp\left(-\frac{(x_1 - x'_1)^2}{L^2}\right)$, for $x, x' \in D$. The constant L denotes the “correlation length” of the stochastic parameter. We choose $L = 0.5$, which implies

$$a_1 \approx 6.7 \cdot 10^{-1}, \quad \lambda_1 \approx 6.9 \cdot 10^{-1}, \quad \lambda_2 \approx 2.7 \cdot 10^{-1}, \quad \lambda_3 \approx 5.8 \cdot 10^{-2}, \quad \lambda_4 \approx 6.8 \cdot 10^{-3}, \quad \lambda_5 \approx 4.2 \cdot 10^{-4}.$$

In the rest of this section, we truncate the expansion to $N = 5$ and $N = 11$ terms. The aim is to study how the algorithm performs for different numbers of parametric dimensions N on an anisotropic problem, where the first parameters are more relevant than the last ones.

In Figure 3, we use the problem with $N = 5$ parameters to provide the reader with a concrete example of the steps of the algorithm. On the left, we plot the evolution of the estimators with respect to the number of degrees of freedom. We plot the values of the estimators any time they are computed (not only once per iteration). The algorithm alternates between steps of parameter enrichment and mesh refinement. The spikes in the value of the finite element estimator correspond

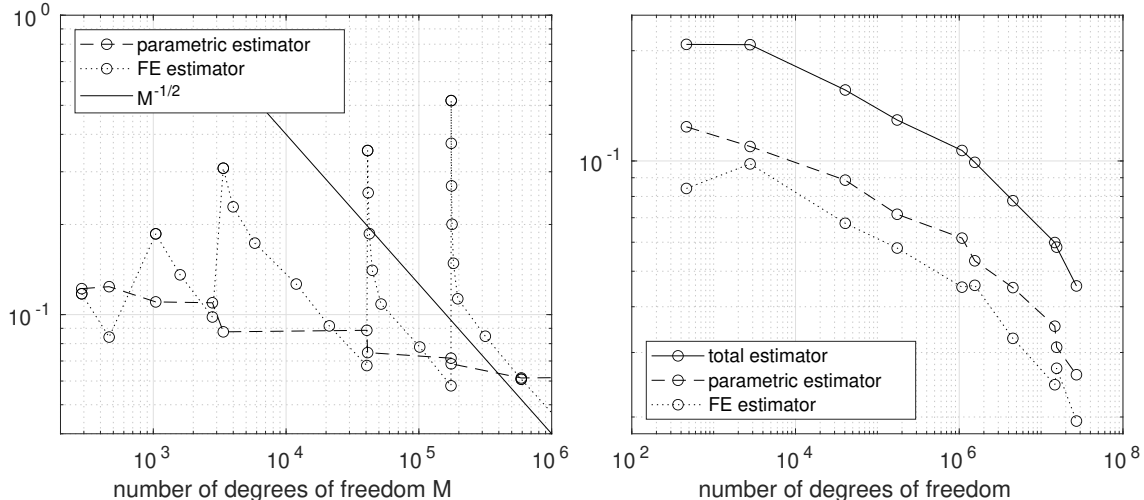


Figure 3. First results for SCFE applied to the problem with Karhunen–Loève expansion $N = 5$. Left: “detailed” evolution of the estimators, i.e. reporting their values any time they are computed during the execution. Right: Total, parametric and finite element estimators at every iteration.

to the parametric enrichment steps, when new collocation points are added to the sparse grid with the initial (coarse) mesh $\mathcal{T}_{\text{init}}$. When finite element refinement is carried out, the finite element estimator eventually decreases with order $M^{-1/2}$, as has to be expected for lowest order adaptive FEM (see also Theorem 3.4). On the right-hand side of Figure 3, we plot the estimator only once per iteration. As prescribed in (62), the finite element estimator is bounded from above by the parametric estimator after each finite element refinement loop.

In Figure 4 we compare the results for $N = 5$ and $N = 11$. On the left, the value of total estimator and reference error are plotted as a function of the number of degrees of freedom. The problem with $N = 11$ gives larger estimator and reference error. However, the difference is marginal, suggesting that the algorithm successfully detects the anisotropy of the problem. On the right, we plot the effectivity index (ratio between estimator and error). As observed in [32], the number of problem dimensions affects the efficiency of the estimator. In view of these facts, the algorithm may benefit from an adaptive dimension selection step as the one proposed in [32, Section 7].

In Figure 5, we consider the problem with $N = 11$ and plot projections of the final multi-index set I . The projections are obtained selecting pairs of parametric dimensions $n_1, n_2 \in 1, \dots, N$ and plotting the 2D set $\{(i_{n_1}, i_{n_2}), i \in I\}$. Observe how larger values are achieved by the first parametric dimensions, confirming that the algorithm manages to detect the anisotropy of the problem.

4.2. Second example: Inclusion problem with $N=8$. We consider an inclusion problem with $N = 8$ parameters similar to that in [32]. Within D , we identify nine disjoint subdomains F and $\{C_n\}_{n=1}^8$ depicted in Figure 6. The diffusion coefficient reads

$$(64) \quad a(x, \mathbf{y}) = a_0(x) + \sum_{n=1}^8 \gamma_n \chi_n y_n \quad \text{with } a_0 \equiv 1.1,$$

where $(\gamma_n)_{n=1}^8 = (1, 0.8, 0.4, 0.2, 0.1, 0.05, 0.02, 0.01)$ are constants used to introduce anisotropy in the problem and χ_n is the characteristic function of C_n , for all $n \in 1, \dots, 8$. The forcing term reads $f(x) := 100\chi_F(x)$, where χ_F is the characteristic function of F .

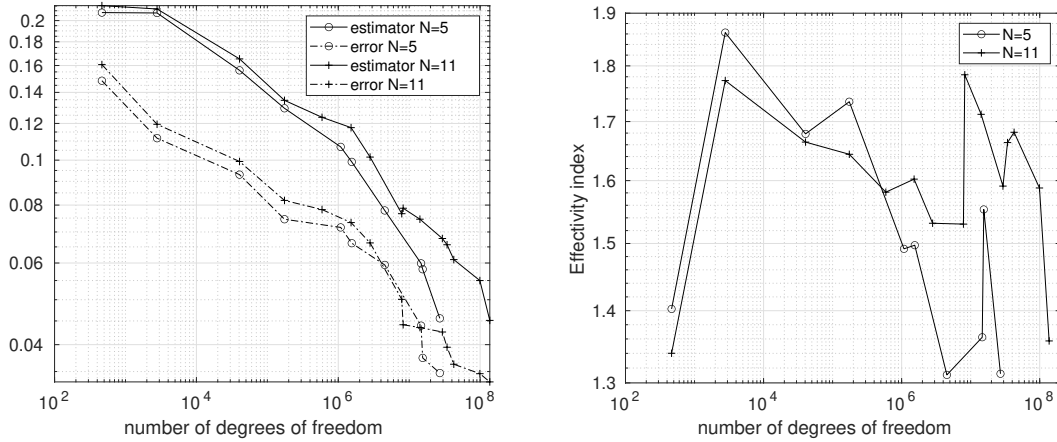


Figure 4. Comparing SCFE applied to the problem with Karhunen–Loève expansion for $N = 5$ and $N = 11$. Left: Total estimator and error. Right: Effectivity index.

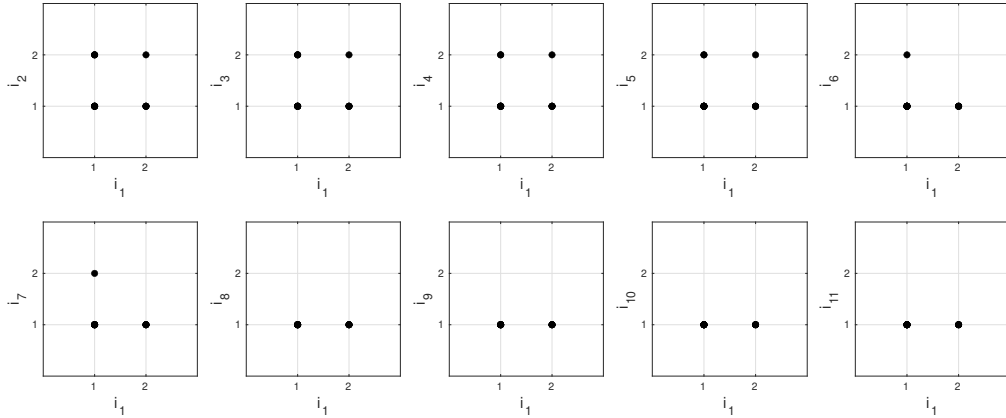


Figure 5. Projections of the final multi-index set from SCFE applied to the the problem with Karhunen–Loève expansion for $N = 11$.

In order to highlight the importance of adaptive finite element refinement in space, we present a comparison between the single mesh version of SCFE from Section 3.4, where the unique mesh is adaptively refined with Dörfler marking, and an analogous version where only uniform refinement on the whole mesh is allowed. In Figure 7 (top left) we report for both algorithms the value of the estimator and reference error. The adaptive version clearly outperforms the one with uniform refinement. In Figure 6 (right) we show a density plot of a mesh produced by the algorithm with $\approx 2 \cdot 10^7$ degrees of freedom. We see that mesh refinement occurs along the boundary of the inclusions and is more pronounced for the inclusions corresponding to larger anisotropy parameter γ_n , confirming that the algorithm detects the parametric structure of the problem.

In Figure 7 (top right) we study the fully adaptive SCFE algorithm and observe a clear performance benefit for the single mesh algorithm from Section 3.4. Additional insight is given in the plot on the bottom left of Figure 7. Here we show the value of the parametric estimator with respect to the number of collocation point for both the fully adaptive and single mesh versions of SCFE. This shows that the fully adaptive SCFE algorithm seems to overrefine the finite-element meshes. We

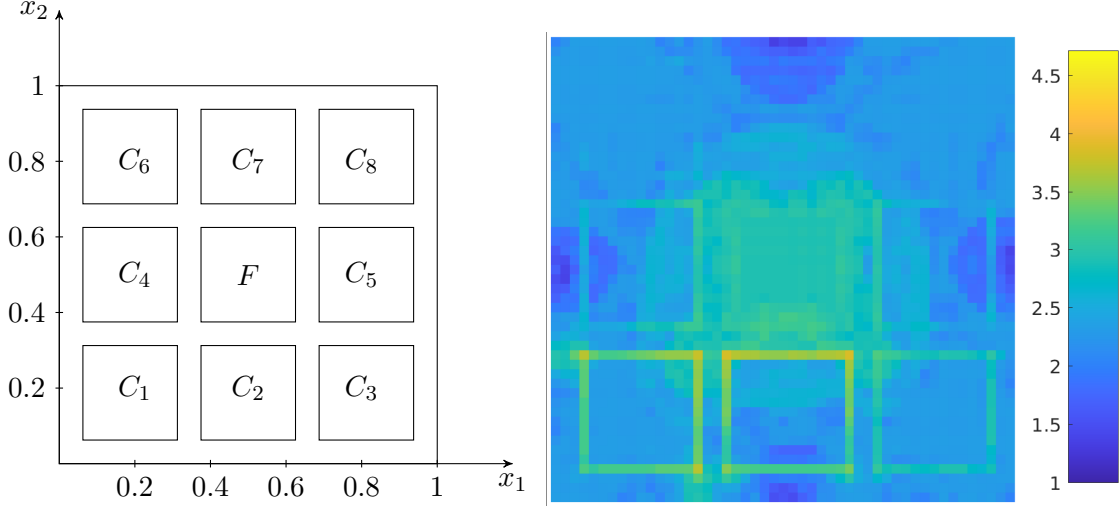


Figure 6. Left: Domain for the inclusion problem. Right: Logarithmic density plot of a mesh of SCFE with single adaptive mesh. The colors refer to the number of mesh elements within one pixel of the plot.

suspect that this is due to the fact that in the derivation of $\eta_{\text{FE},I}$ in Section 1.3, one is required to use the triangle inequality and thus sacrifices local information of the sparse grid interpolant. This is not necessary in the single-mesh estimator from Section 3.4.

Finally, we ran Algorithm 1 with fixed tolerance Tol (not depending on the parametric estimator) and plot the results in Figure 7 (bottom right). We observe that the results are very much comparable to the standard fully adaptive SCFE algorithm, except for significant over refinement in the early stages of the computation (see the flat line of the finite element error estimator satisfying the tolerance). In terms of computational effort, the algorithms are nearly identical, as the same spatial refinements are performed, only at different stages of the algorithm.

We also tested the algorithms with respect to the $L^2(\Gamma)$ -norm instead of the $L^\infty(\Gamma)$ -norm. The necessary changes in the estimators are straightforward, essentially we replace the search for the maximum by a Monte Carlo quadrature. The theoretical results of this manuscript all hold verbatim for the $L^2(\Gamma)$ -norm. Figure 8 shows the results. Again the single mesh algorithm outperforms the fully adaptive algorithm.

4.2.1. Distribution of computational cost. In Figure 8 (right-hand side) we compare the total number of degrees of freedom and collocation points achieved by the three methods, i.e. the single adaptive mesh algorithm from Section 3.4, adaptivity in the parameter space but uniform refinement in the spatial domain, and the fully adaptive SCFE algorithm (Algorithm 1). The adaptive strategy with a single adaptive mesh performs parametric refinement more often than the other two, leading to a higher number of collocation points and lower average number of degrees of freedom per collocation point. In Figure 9 (compare also Figure 6) we provide logarithmic density plots of the meshes produced by the multiple adaptive mesh algorithm (with $\approx 2 \cdot 10^7$ degrees of freedom). We observe that the mesh corresponding to a collocation point is locally refined along the edges of the corresponding inclusion. Furthermore, the intensity of the refinement around a certain inclusion is related to the constant γ_n of the diffusion coefficient, confirming that the numerical methods detects the anisotropy of the problem.

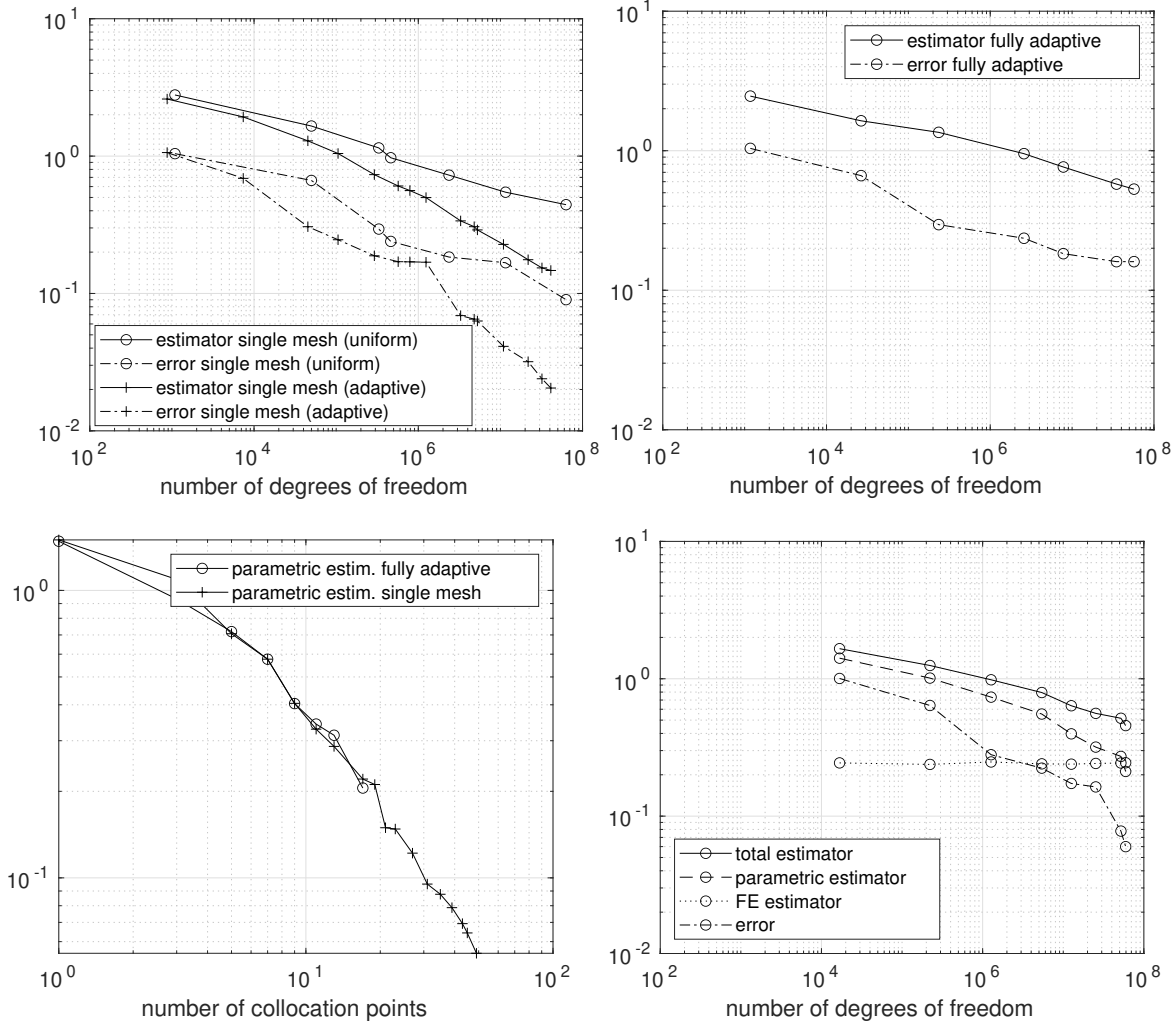


Figure 7. Results for SCFE on the 8D inclusion problem. Top left: Comparison between adaptive and uniform space refinement with the single mesh algorithm from Section 3.4. Top right: Total estimator and error for the fully adaptive algorithm. Bottom left: Parametric estimator as a function of the number of collocation points, for both the fully adaptive and the single mesh SCFE. Bottom right: the fully adaptive SCFE with fixed value for finite element tolerance Tol.

5. CONCLUSION

We analyze the adaptive stochastic collocation algorithm from [32] and prove convergence of several different versions of the algorithm:

- Convergence of the parametric enrichment algorithm without finite element refinement (Section 2)
- Convergence of the fully adaptive algorithm (Algorithm 1) even with optimal convergence of the finite element loop (Theorem 3.4)
- Convergence of a single-mesh variant of Algorithm 1 (Section 3.4) proposed in [32].

The numerical examples clearly show the superiority of spatial adaptive refinement combined with parametric enrichment over pure parametric enrichment algorithms. While the theoretical results are strongest for the fully adaptive algorithm (linear convergence in Proposition 3.2 for Algorithm 1)

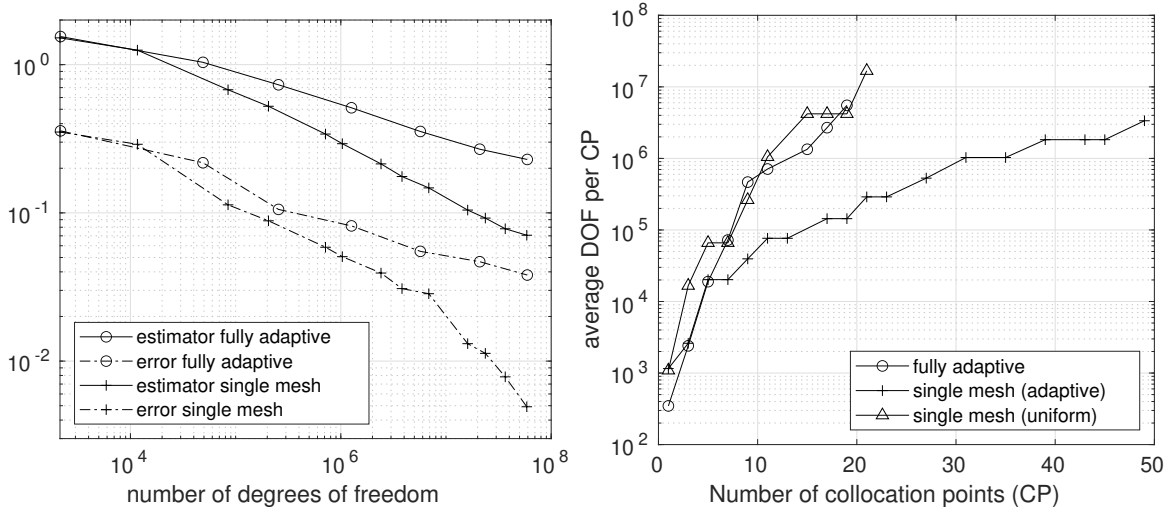


Figure 8. Left: Numerical results on the 8D inclusion problem in the $L^2(\Gamma)$ -norm. Right: Total estimator and error for SCFE (fully adaptive). Right: Average number of degrees of freedom (DOF) per collocation point (CP) plotted versus the number of collocation points. Each line corresponds to one of the three proposed algorithms. Each marker corresponds to one step of the adaptively refined discrete solution.

the single mesh algorithm from Section 3.4 seems to be more efficient. This is underlined by the numerical experiments in the previous section, which clearly show an advantage of the single mesh version over the fully adaptive version. Based on the theoretical results from Theorem 3.4 and the experiments, we come to the conclusion that the finite element error estimator of Algorithm 1 severely over-estimates the total error and hence leads to over-refinement of the finite element meshes. This does not seem to happen for the single-mesh error estimator. We suspect that the application of the triangle inequality in the derivation in Section 3.3 is mainly responsible for this over-estimation and further research is required to see whether this can be avoided.

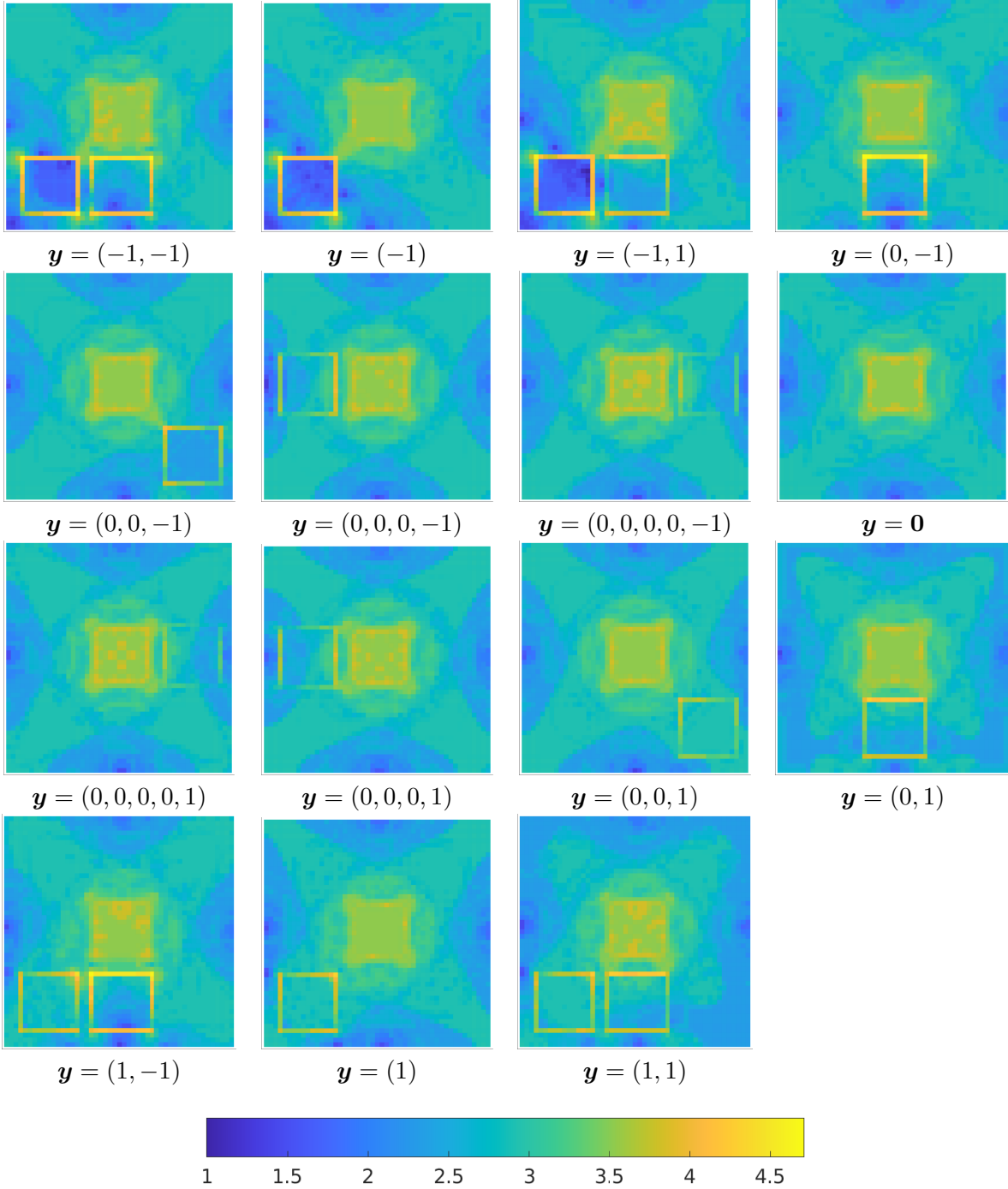


Figure 9. Density plot of the meshes produced by the fully adaptive SCFE algorithm. The corresponding collocation point is indicated below, ignoring the trailing components equal to zero. The color-bar at the bottom indicates the base-10 logarithm of the density of elements.

REFERENCES

- [1] I. Babuška, B. Andersson, P. J. Smith, and K. Levin. Damage analysis of fiber composites. I. Statistical analysis on fiber scale. *Comput. Methods Appl. Mech. Engrg.*, 172(1-4):27–77, 1999.
- [2] Ivo Babuška, Fabio Nobile, and Raul Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 45(3):1005–1034, 2007.
- [3] J. Bäck, F. Nobile, L. Tamellini, and R. Tempone. Stochastic spectral Galerkin and collocation methods for PDEs with random coefficients: a numerical comparison. In J.S. Hesthaven and E.M. Ronquist, editors, *Spectral and High Order Methods for Partial Differential Equations*, volume 76 of *Lecture Notes in Computational Science and Engineering*, pages 43–62. Springer, 2011. Selected papers from the ICOSAHOM '09 conference, June 22-26, Trondheim, Norway.
- [4] Volker Barthelmann, Erich Novak, and Klaus Ritter. High dimensional polynomial interpolation on sparse grids. *Advances in Computational Mathematics*, 12(4):273–288, 2000.
- [5] Joakim Beck, Raul Tempone, Fabio Nobile, and Lorenzo Tamellini. On the optimal polynomial approximation of stochastic PDEs by Galerkin and collocation methods. *Math. Models Methods Appl. Sci.*, 22(9):1250023, 33, 2012.
- [6] Alex Bespalov, Dirk Praetorius, Leonardo Rocchi, and Michele Ruggeri. Convergence of adaptive stochastic galerkin fem. *SIAM Journal on Numerical Analysis*, 57(5):2359–2382, 2019.
- [7] C. Carstensen, M. Feischl, M. Page, and D. Praetorius. Axioms of adaptivity. *Comput. Math. Appl.*, 67(6):1195–1253, 2014.
- [8] J Manuel Cascon, Christian Kreuzer, Ricardo H Nochetto, and Kunibert G Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM Journal on Numerical Analysis*, 46(5):2524–2550, 2008.
- [9] Abdellah Chkifa, Albert Cohen, and Christoph Schwab. High-dimensional adaptive sparse polynomial interpolation and applications to parametric pdes. *Foundations of Computational Mathematics*, 14(4):601–633, 2014.
- [10] Albert Cohen and Ronald DeVore. Approximation of high-dimensional parametric PDEs. *Acta Numer.*, 24:1–159, 2015.
- [11] Albert Cohen, Ronald DeVore, and Christoph Schwab. Convergence rates of best N -term Galerkin approximations for a class of elliptic sPDEs. *Found. Comput. Math.*, 10(6):615–646, 2010.
- [12] J. Dick, F. Y. Kuo, Q. T. Le Gia, D. Nuyens, and C. Schwab. Higher order QMC Petrov-Galerkin discretization for affine parametric operator equations with random field inputs. *SIAM J. Numer. Anal.*, 52(6):2676–2702, 2014.
- [13] Josef Dick, Michael Feischl, and Christoph Schwab. Improved efficiency of a multi-index FEM for computational uncertainty quantification. *SIAM J. Numer. Anal.*, 57(4):1744–1769, 2019.
- [14] Josef Dick, Robert N. Gantner, Quoc T. Le Gia, and Christoph Schwab. Multilevel higher-order quasi-Monte Carlo Bayesian estimation. *Math. Models Methods Appl. Sci.*, 27(5):953–995, 2017.
- [15] Josef Dick, Frances Y. Kuo, Quoc T. Le Gia, Dirk Nuyens, and Christoph Schwab. Higher order QMC Petrov-Galerkin discretization for affine parametric operator equations with random field inputs. *SIAM J. Numer. Anal.*, 52(6):2676–2702, 2014.
- [16] Josef Dick, Frances Y. Kuo, Quoc T. Le Gia, and Christoph Schwab. Multilevel higher order QMC Petrov-Galerkin discretization for affine parametric operator equations. *SIAM J. Numer. Anal.*, 54(4):2541–2568, 2016.
- [17] Sergey Dolgov, Boris N. Khoromskij, Alexander Litvinenko, and Hermann G. Matthies. Polynomial chaos expansion of random coefficients and the solution of stochastic partial differential equations in the tensor train format. *SIAM/ASA Journal on Uncertainty Quantification*, 3(1):1109–1135, 2015.
- [18] VK Dzijadyk and VV Ivanov. On asymptotics and estimates for the uniform norms of the lagrange interpolation polynomials corresponding to the chebyshev nodal points. *Analysis Mathematica*, 9(2):85–97, 1983.
- [19] Martin Eigel, Oliver Ernst, Björn Sprungk, and Lorenzo Tamellini. On the convergence of adaptive stochastic collocation for elliptic partial differential equations with affine diffusion. *arXiv:2008.07186*, 2020.
- [20] Martin Eigel, Claude Jeffrey Gittelsohn, Christoph Schwab, and Elmar Zander. A convergent adaptive stochastic Galerkin finite element method with quasi-optimal spatial meshes. *ESAIM Math. Model. Numer. Anal.*, 49(5):1367–1398, 2015.
- [21] I. Elishakoff, editor. *Whys and hows in uncertainty modelling*, volume 388 of *CISM Courses and Lectures*. Springer-Verlag, Vienna, 1999. Probability, fuzziness and anti-optimization.
- [22] Mike Espig, Wolfgang Hackbusch, Alexander Litvinenko, Hermann G. Matthies, and Philipp Wähnert. Efficient low-rank approximation of the stochastic Galerkin matrix in tensor formats. *Computers & Mathematics with Applications*, 67(4):818 – 829, 2014. High-order Finite Element Approximation for Partial Differential Equations.
- [23] Michael Feischl, Frances Y. Kuo, and Ian H. Sloan. Fast random field generation with H -matrices. *Numer. Math.*, 140(3):639–676, 2018.

- [24] Philipp Frauenfelder, Christoph Schwab, and Radu Alexandru Todor. Finite elements for elliptic problems with stochastic coefficients. *Comput. Methods Appl. Mech. Engrg.*, 194(2-5):205–228, 2005.
- [25] Jerome H. Friedman. Multivariate adaptive regression splines. *Ann. Statist.*, 19(1):1–141, 1991. With discussion and a rejoinder by the author.
- [26] Stefan Funken, Dirk Praetorius, and Philipp Wissgott. Efficient implementation of adaptive p1-fem in matlab. *Computational Methods in Applied Mathematics*, 11(4):460–490, 2011.
- [27] Thomas Gerstner and Michael Griebel. Dimension-adaptive tensor-product quadrature. *Computing*, 71(1):65–87, September 2003.
- [28] Roger Ghanem. Ingredients for a general purpose stochastic finite elements implementation. *Comput. Methods Appl. Mech. Engrg.*, 168(1-4):19–34, 1999.
- [29] Roger G. Ghanem and Pol D. Spanos. *Stochastic finite elements: a spectral approach*. Springer-Verlag, New York, 1991.
- [30] Loïc Giraldi, Alexander Litvinenko, Dishu Liu, Hermann G. Matthies, and Anthony Nouy. To be or not to be intrusive? the solution of parametric and stochastic equations—the ”plain vanilla” galerkin case. *SIAM Journal on Scientific Computing*, 36(6):A2720–A2744, 2014.
- [31] I.G. Graham, F.Y. Kuo, D. Nuyens, R. Scheichl, and I.H. Sloan. Quasi-Monte Carlo methods for elliptic PDEs with random coefficients and applications. *Journal of Computational Physics*, 230(10):3668 – 3694, 2011.
- [32] Diane Guignard and Fabio Nobile. A posteriori error estimation for the stochastic collocation finite element method. *SIAM Journal on Numerical Analysis*, 56(5):3121–3143, 2018.
- [33] T. J. Hastie and R. J. Tibshirani. *Generalized additive models*, volume 43 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, Ltd., London, 1990.
- [34] Lukas Herrmann, Kristin Kirchner, and Christoph Schwab. Multilevel approximation of Gaussian random fields: fast simulation. *Math. Models Methods Appl. Sci.*, 30(1):181–223, 2020.
- [35] A. N. Kolmogorov. On the representation of continuous functions of several variables by superpositions of continuous functions of a smaller number of variables. *Amer. Math. Soc. Transl. (2)*, 17:369–373, 1961.
- [36] J. Lang, R. Scheichl, and D. Silvester. A fully adaptive multilevel stochastic collocation strategy for solving elliptic PDEs with random data. *J. Comput. Phys.*, 419:109692, 17, 2020.
- [37] Lionel Mathelin, M. Yousuff Hussaini, and Thomas A. Zang. Stochastic approaches to uncertainty quantification in CFD simulations. *Numer. Algorithms*, 38(1-3):209–236, 2005.
- [38] Hermann G. Matthies and Andreas Keese. Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations. *Comput. Methods Appl. Mech. Engrg.*, 194(12-16):1295–1331, 2005.
- [39] F. Nobile, R. Tempone, and C. G. Webster. An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. Numer. Anal.*, 46(5):2411–2442, 2008.
- [40] Fabio Nobile, Raúl Tempone, and Clayton G Webster. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 46(5):2309–2345, 2008.
- [41] Carl-Martin Pfeiler and Dirk Praetorius. Dörfler marking with minimal cardinality is a linear complexity problem. *Math. Comp.*, 89(326):2735–2752, 2020.
- [42] Christoph Schwab and Radu Alexandru Todor. Karhunen–Loève approximation of random fields by generalized fast multipole methods. *Journal of Computational Physics*, 217(1):100–122, 2006.
- [43] Rob Stevenson. Optimality of a standard adaptive finite element method. *Found. Comput. Math.*, 7(2):245–269, 2007.
- [44] Rob Stevenson. The completion of locally refined simplicial partitions created by bisection. *Math. Comp.*, 77(261):227–241, 2008.
- [45] A. L. Teckentrup, P. Jantsch, C. G. Webster, and M. Gunzburger. A multilevel stochastic collocation method for partial differential equations with random input data. *SIAM/ASA J. Uncertain. Quantif.*, 3(1):1046–1074, 2015.
- [46] Grace Wahba. *Spline models for observational data*, volume 59 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990.
- [47] Grzegorz W Wasilkowski and Henryk Wozniakowski. Explicit cost bounds of algorithms for multivariate tensor product problems. *Journal of Complexity*, 11(1):1–56, 1995.
- [48] Dongbin Xiu and Jan S. Hesthaven. High-order collocation methods for differential equations with random inputs. *SIAM J. Sci. Comput.*, 27(3):1118–1139, 2005.
- [49] Dongbin Xiu and George Em Karniadakis. Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos. *Comput. Methods Appl. Mech. Engrg.*, 191(43):4927–4948, 2002.
- [50] Rong-Xian Yue and Fred J. Hickernell. Designs for smoothing spline ANOVA models. *Metrika*, 55(3):161–176, 2002.