

# Improved Efficiency of a Multi-Index FEM for Computational Uncertainty Quantification

J. Dick and M. Feischl and Ch. Schwab

Research Report No. 2018-22  
June 2018

Seminar für Angewandte Mathematik  
Eidgenössische Technische Hochschule  
CH-8092 Zürich  
Switzerland

# IMPROVED EFFICIENCY OF A MULTI-INDEX FEM FOR COMPUTATIONAL UNCERTAINTY QUANTIFICATION\*

JOSEF DICK<sup>†</sup>, MICHAEL FEISCHL<sup>‡</sup>, AND CHRISTOPH SCHWAB<sup>§</sup>

**Abstract.** We propose a multi-index algorithm for the Monte Carlo discretization of a linear, elliptic PDE with affine-parametric input. We prove an error vs. work analysis which allows a multi-level finite-element approximation in the physical domain, and apply the multi-index analysis with isotropic, unstructured mesh refinement in the physical domain for the solution of the forward problem, for the approximation of the random field, and for the Monte-Carlo quadrature error. Our approach allows general spatial domains and unstructured mesh hierarchies. The improvement in complexity is obtained from combining spacial discretization, dimension truncation and MC sampling in a multi-index fashion. Our analysis improves cost estimates compared to multi-level algorithms for similar problems and mathematically underpins the outstanding practical performance of multi-index algorithms for partial differential equations with random coefficients.

**Key words.** multi-index, Monte Carlo, finite element method, uncertainty quantification

**AMS subject classifications.**

**1. Introduction.** The term *multi-index Monte Carlo method* MIMC was first coined in the work [4] as an extension of the multi-level Monte Carlo method MLMC developed in [3]. Since then, the multi-level idea has been applied in many areas including high-dimensional integration, stochastic differential equations, and several types of PDEs with random coefficients. Most of these works used multi-level Monte Carlo algorithms, while few papers considered multi-level quasi-Monte Carlo (MLQMC) algorithms [1, 2]. In the framework of PDEs with random coefficients, the idea of the multi-level approach is to introduce a sequence of geometrically refined grids and to compute finite element approximations of a given partial differential equation (PDE) with random coefficients on each level. By varying the Monte Carlo (MC) sample size on each level of the finite element method (FEM) and by recombination of the individual approximation, it is possible to reduce the total cost (up to logarithmic factors) from  $\text{cost}(\text{sampling}) \times \text{cost}(\text{FEM})$  to  $\text{cost}(\text{sampling}) + \text{cost}(\text{FEM})$ , where the individual cost terms are measured on the finest level. This idea was further extended in [4] to include more than one parameter which is quantized into levels. One possible example for this approach, presented in [4], is to introduce anisotropic discretizations in the physical domain (as, e.g., sparse grid FE discretizations) for which two (three) parameters control the element size in the coordinate direction. This ‘sparse grid’ approach has been combined with a heuristic, adaptive algorithm and a Quasi-Monte Carlo algorithm in [6]. More examples of variations of this approach can be found in [5, 7]. In these approaches, the construction of sparse grid hierarchies in the physical domain to access the multi-index efficiency could impose obstructions on

---

\*Submitted to the editors DATE.

**Funding:** Supported by the Australian Research Council (ARC) under grant number DP150101770 (to JD), by the Swiss National Science Foundation (SNSF) under grant number 200021\_159940 (to CS), and by the Deutsche Forschungsgemeinschaft (DFG) through CRC 1173 (to MF)

<sup>†</sup>School of Mathematics and Statistics, The University of New South Wales, Sydney 2052, Australia ([dick@unsw.edu.au](mailto:dick@unsw.edu.au)).

<sup>‡</sup>Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany ([michael.feischl@kit.edu](mailto:michael.feischl@kit.edu)).

<sup>§</sup>SAM, ETH Zürich, ETH Zentrum HG G57.1, CH 8092 Zürich, Switzerland ([christoph.schwab@sam.math.ethz.ch](mailto:christoph.schwab@sam.math.ethz.ch)).

the shape of the physical domains which are amenable to this kind of discretization.

In the present work, we follow a different (but, as we will show, very natural) approach: we include the approximation of the random coefficients into the multi-index discretization and convergence analysis. As we show, this is effective due to the following consideration: apart from toy problems, it is almost impossible to obtain exact (MC or also collocation or QMC) samples of the random coefficients, and often, this particular part constitutes a major bottleneck in computations. It is therefore of practical importance to improve efficiency of algorithms.

Although the presently proposed approach is, in principle, more general, we develop it here for affine-parametric random coefficients in a standard, linear Poisson model problem. We parametrize the uncertain diffusion coefficient by a dimensionally truncated KL expansion, i.e.

$$A(x, \omega) = \phi_0(x) + \sum_{j=1}^{\infty} \phi_j(x) \psi_j(\omega_j) \approx A^\nu(x, \omega) := \phi_0(x) + \sum_{j=1}^{s_\nu} \phi_j(x) \psi_j(\omega_j),$$

where  $\{s_\nu\}_{\nu \in \mathbb{N}} \subset \mathbb{N}$  is an increasing sequence of “dimension truncation” parameters.

Given a quantity of interest  $G(\cdot)$ , the idea is to approximate the expectation of the exact solution  $\mathbb{E}(G(u))$  by several instances of the “double difference”  $D_\ell^\nu = (u_\ell^\nu - u_{\ell-1}^\nu) - (u_{\ell-1}^{\nu-1} - u_{\ell-2}^{\nu-1})$ , where  $u_\ell^\nu$  denotes the FEM approximation of  $u$  on a mesh of size  $h_\ell$  and with respect to the approximation  $A^\nu$  of the exact random coefficient. This leads to

$$\mathbb{E}(G(u)) \approx \sum_{0 \leq \ell + \nu \leq N} Q_{m_{N-\ell-\nu}}(G(D_\ell^\nu)),$$

where  $Q_{m_{N-\ell-\nu}}$  denotes a MC sampling rule with given sample size  $m_{N-\ell-\nu} \in \mathbb{N}$  such that  $m_0 < m_1 < \dots < m_N$ . The main result of this work is to prove that the above approximation is (up to logarithmic factors) optimal in the sense that it is as good as the approximation given by the naive approach  $Q_{m_N}(G(u_N^N))$ , where all components are computed on the finest level, while reducing the computational cost.

The error/cost estimates from Section 6 show that the distribution of work among the individual levels is optimal up to logarithmic factors. This can be seen from the fact that the multi-index algorithm achieves the same (up to logarithmic factors) cost versus error ratio than the worst ratio of each of the involved algorithms (FEM, Monte Carlo, approximation of the random coefficient). Since a combined algorithm of this form cannot be more efficient than each of its components, this shows optimality.

**2. Model problem.** We chose a simple Poisson model problem to give a concise presentation of the ideas and proof techniques. The authors are confident that very similar techniques can be used to include more general model problems. Moreover, we focus on the standard case of  $H^2$ -regularity of the Poisson problem. Intermediate cases with less regularity can be included with the same arguments, but are left out for the sake of clarity.

**2.1. Abstract setting.** Let  $D \subseteq \mathbb{R}^d$  be a bounded Lipschitz domain for  $d \in \{2, 3\}$  and let  $(\Omega, \Sigma, \mathbb{P})$  denote a probability space.

Define the parametrized bilinear form

$$a(A, \omega; u, v) := \int_D A(x, \omega) \nabla u(x) \cdot \nabla v(x) dx$$

for some diffusion coefficient  $A: D \times \Omega \rightarrow [0, \infty)$  which satisfies  $A(\cdot, \omega) \in L^\infty(D)$  for almost all  $\omega \in \Omega$ . Here,  $H_0^1(D)$  denotes the usual Sobolev space of weak derivatives given by

$$H_0^1(D) := \{v \in L^2(D) : \nabla v \in L^2(D)^d, v|_{\partial D} = 0\}.$$

We assume a sequence of approximations  $(A^\nu)_{\nu \in \mathbb{N}}$  of  $A = A^\infty$  which satisfy  $A^\nu(\cdot, \omega) \in W^{1, \infty}(D)$  for almost all  $\omega \in \Omega$  as well as

$$(2.1) \quad \lim_{\nu \rightarrow \infty} \|A - A^\nu\|_{L^\infty(\Omega; W^{1, \infty}(D))} = 0.$$

Furthermore, we assume for  $\nu \in \mathbb{N} \cup \{\infty\}$

$$(2.2) \quad 0 < a_{\min}(\omega) \leq \inf_{x \in D} A^\nu(x, \omega) \leq \sup_{x \in D} A^\nu(x, \omega) \leq a_{\max}(\omega) < \infty.$$

For brevity, we define  $a_\omega^\nu(\cdot, \cdot) := a(A^\nu, \omega, \cdot, \cdot)$ . Finally, suppose the right-hand side  $f \in H^{-1}(D)$ . We embed  $L^2(D)$  in  $H^{-1}(D)$  via the compact embedding  $v \mapsto \langle v, \cdot \rangle_D$  for all  $v \in L^2(D)$ .

The assumptions imply ellipticity and continuity of the bilinear form, i.e., for almost all  $\omega \in \Omega$

$$(2.3) \quad \inf_{\nu \in \mathbb{N} \cup \infty} \inf_{u \in H_0^1(D)} \frac{a_\omega^\nu(u, u)}{\|u\|_{H^1(D)}^2} \geq a_{\min}(\omega)$$

as well as

$$(2.4) \quad \sup_{\nu \in \mathbb{N} \cup \infty} \sup_{u, v \in H_0^1(D)} \frac{a_\omega^\nu(u, v)}{\|u\|_{H^1(D)} \|v\|_{H^1(D)}} \leq a_{\max}(\omega).$$

The Lax-Milgram lemma proves unique solvability of the problem: Find  $u: D \times \Omega \rightarrow \mathbb{R}$  with  $u(\omega) \in H_0^1(D)$  such that

$$a(A, \omega; u(\omega), v) = \langle f, v \rangle_D \quad \text{for all } v \in H_0^1(D) \text{ a.e. in } \Omega.$$

Finally, we are interested in the expectation of a certain quantity of interest  $G(\cdot)$  which is a deterministic, bounded linear functional  $G(\cdot): H_0^1(D) \rightarrow \mathbb{R}$ , i.e.

$$\mathbb{E}(G(u)) \in \mathbb{R}.$$

We assume that  $G$  has an  $L^2$  representer, i.e.,  $g \in L^2(D)$  such that

$$G(v) = \int_D gv \, dx \quad \text{for all } v \in H_0^1(D).$$

**2.2. Finite element discretization.** We assume a sequence of nested triangulations  $(\mathcal{T}_\ell)_{\ell \in \mathbb{N}}$  with corresponding spaces  $(\mathcal{X}_\ell)_{\ell \in \mathbb{N}}$  (such that  $\mathcal{X}_\ell \subseteq \mathcal{X}_k \subset H_0^1(D)$  for all  $\ell \leq k$ ). We assume the following approximation property of the spaces  $\mathcal{X}_\ell$ : There exists a constant  $C_{\text{approx}} > 0$  and a monotone sequence  $(h_\ell)_{\ell \in \mathbb{N}}$  with  $h_\ell > 0$  and  $\lim_{\ell \rightarrow \infty} h_\ell = 0$  such that all  $u \in H^2(D)$  satisfy

$$(2.5) \quad \inf_{v \in \mathcal{X}_\ell} \|u - v\|_{H^1(D)} \leq C_{\text{approx}} h_\ell \|u\|_{H^2(D)}.$$

For convenience, we assume  $h_{\ell+1} \geq C_{\text{unif}} h_\ell$  for all  $\ell \in \mathbb{N}$  and for some constant  $C_{\text{unif}} > 0$ . (A popular example would be to have uniform triangulations  $\mathcal{T}_\ell$  with mesh-width  $h_\ell$  and to define the spaces  $\mathcal{X}_\ell$  as piecewise-linear functions on  $\mathcal{T}_\ell$ .)

Define the Galerkin approximation  $u_\ell^\nu(\omega) \in \mathcal{X}_\ell$  as the unique solution of

$$a_\omega^\nu(u_\ell^\nu(\omega), v) = \langle f, v \rangle_D \quad \text{for all } v \in \mathcal{X}_\ell \text{ and almost all } \omega \in \Omega.$$

(Note that unique solvability follows from the Lax-Milgram lemma and (2.3)–(2.4).) Consider the solution operators  $\mathbb{S}_\ell^\nu(\omega): H^{-1}(D) \rightarrow \mathcal{X}_\ell$  defined by  $\mathbb{S}_\ell^\nu(\omega)f := u_\ell^\nu(\omega)$ . Moreover, let  $(\mathbb{S}_\ell^\nu(\omega))^{-1}: \mathcal{X}_\ell \rightarrow H^{-1}(D)$  be defined by

$$((\mathbb{S}_\ell^\nu(\omega))^{-1}u)(v) := a_\omega^\nu(u, v) \quad \text{for all } u \in \mathcal{X}_\ell, v \in H_0^1(D).$$

For brevity, we will omit the random parameter and just write  $\mathbb{S}_\ell^\nu := \mathbb{S}_\ell^\nu(\omega)$ . Moreover, we write  $\mathbb{S}_\infty^\nu f := u^\nu$ , where  $u^\nu(\omega) \in H_0^1(D)$  is the unique solution of

$$a_\omega^\nu(u^\nu(\omega), v) = \langle f, v \rangle_D \quad \text{for all } v \in H_0^1(D).$$

Thus,  $u^\nu$  denotes the exact solution corresponding to  $A^\nu$  and  $((\mathbb{S}_\infty^\nu(\omega))^{-1}\cdot)(v) := a_\omega^\nu(\cdot, v) \in H^{-1}(D)$ .

We restrict to domains  $D \subseteq \mathbb{R}^d$ , which imply  $H^2$ -regularity of the exact solution as long as  $f \in L^2(D)$ , i.e., there exists a constant  $C_{\text{reg}} > 0$  such that all  $\omega \in \Omega$  satisfy

$$(2.6) \quad \sup_{\nu \in \mathbb{N}} \|\mathbb{S}_\infty^\nu f\|_{H^2(D)} \leq \frac{C_{\text{reg}}}{a_{\min}(\omega)^2} (1 + \|A^\nu(\omega)\|_{W^{1,\infty}(D)}) \|f\|_{L^2(D)}$$

for all  $f \in L^2(D)$  (Note that if the standard Poisson equation satisfies  $H^2$  regularity, (2.6) follows as an immediate consequence). Possible examples of domains  $D$  which satisfy this property include domains with  $C^2$ -boundary  $\partial D$  or convex Lipschitz domains.

LEMMA 2.1. *We have*

$$\|\mathbb{S}_\ell^\nu\|_{H^{-1}(D) \rightarrow H^1(D)} \leq a_{\min}(\omega)^{-1}$$

as well as

$$\|(\mathbb{S}_\ell^\nu)^{-1}\|_{\mathcal{X}_\ell \rightarrow H^{-1}(D)} \leq a_{\max}(\omega).$$

*Proof.* The result follows immediately from (2.3)–(2.4).  $\square$

**3. Product structure of the approximation error.** The main purpose of this section is to prove the product error estimate of Theorem 3.9 below at the end of this section. This error estimate factors the total error into error contributions of the approximation of the random coefficient  $A \approx A^\nu$  and finite element approximation error  $h_\ell \rightarrow 0$ . We will restate several well-known results from finite-element analysis, as we will make use of the exact dependence on the constants.

In view of the multi-index decomposition in Section 6, we consider the difference of differences

$$D_\ell^\nu := (u_\ell^\nu - u_{\ell-1}^\nu) - (u_\ell^{\nu-1} - u_{\ell-1}^{\nu-1}): \Omega \rightarrow \mathcal{X}_\ell.$$

The goal is to get an error estimate of product form, as this allows us to obtain nearly optimal complexity estimates. The key observation is that there holds

$$D_\ell^\nu \approx (\mathbb{S}_\ell^\nu - \mathbb{S}_{\ell-1}^\nu)(\mathbb{S}_\ell^\nu)^{-1}(\mathbb{S}_\ell^\nu - \mathbb{S}_{\ell-1}^{\nu-1})f,$$

where the error term can be controlled.

In the following, we use the operator norm for bilinear forms  $b(\cdot, \cdot): \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$  for a Hilbert space  $\mathcal{X}$ , i.e.,

$$\|b\| := \sup_{x, y \in \mathcal{X} \setminus \{0\}} \frac{|b(x, y)|}{\|x\|_{\mathcal{X}} \|y\|_{\mathcal{X}}}.$$

LEMMA 3.1. *Given  $A, B: \Omega \rightarrow L^\infty(D)$ , there holds*

$$\|a(A, \omega, \cdot, \cdot) - a(B, \omega, \cdot, \cdot)\| \leq \|A(\omega) - B(\omega)\|_{L^\infty(D)} \quad \text{for all } \omega \in \Omega.$$

Moreover, we have for all  $\ell, \nu, \mu \in \mathbb{N}$

$$\|\mathbb{S}_\ell^\nu f - \mathbb{S}_\ell^\mu f\|_{H^1(D)} \leq a_{\min}(\omega)^{-2} \|A^\nu(\omega) - A^\mu(\omega)\|_{L^\infty(D)} \|f\|_{L^2(D)}.$$

*Proof.* We have for almost all  $\omega \in \Omega$  that

$$\begin{aligned} |a(A, \omega, u, v) - a(B, \omega, u, v)| &\leq \int_D |A(x, \omega) - B(x, \omega)| |\nabla u| |\nabla v| dx \\ &\leq \|A(\omega) - B(\omega)\|_{L^\infty(D)} \|u\|_{H^1(D)} \|v\|_{H^1(D)}. \end{aligned}$$

With this, (2.3), and Lemma 2.1, there holds for  $\omega \in \Omega$

$$\begin{aligned} a_{\min}(\omega) \|\mathbb{S}_\ell^\nu f - \mathbb{S}_\ell^\mu f\|_{H^1(D)}^2 &\leq a_\omega^\nu (\mathbb{S}_\ell^\nu f - \mathbb{S}_\ell^\mu f, \mathbb{S}_\ell^\nu f - \mathbb{S}_\ell^\mu f) \\ &= \langle f, \mathbb{S}_\ell^\nu f - \mathbb{S}_\ell^\mu f \rangle_D - a_\omega^\nu (\mathbb{S}_\ell^\mu f, \mathbb{S}_\ell^\nu f - \mathbb{S}_\ell^\mu f) \\ &= (a_\omega^\mu - a_\omega^\nu) (\mathbb{S}_\ell^\mu f, \mathbb{S}_\ell^\nu f - \mathbb{S}_\ell^\mu f) \\ &\leq a_{\min}(\omega)^{-1} \|A^\nu - A^\mu\|_{L^\infty(D)} \|f\|_{L^2(D)} \|\mathbb{S}_\ell^\nu f - \mathbb{S}_\ell^\mu f\|_{H^1(D)}. \end{aligned}$$

This concludes the proof.  $\square$

LEMMA 3.2 (Galerkin orthogonality). *There holds Galerkin orthogonality for all  $k, \ell \in \mathbb{N} \cup \{\infty\}$ ,  $\nu \in \mathbb{N}$  and all  $f \in H^{-1}(D)$  in the form*

$$a_\omega^\nu (\mathbb{S}_k^\nu f, v) = a_\omega^\nu (\mathbb{S}_\ell^\nu f, v) \quad \text{for all } v \in \mathcal{X}_{\min\{\ell, k\}} \text{ and all } \omega \in \Omega.$$

Particularly, this implies  $\mathbb{S}_\ell^\nu (\mathbb{S}_k^\nu)^{-1} = \text{id}_{\mathcal{X}_k}$  for all  $\ell \geq k$  and  $k < \infty$ .

*Proof.* By definition, we have

$$a_\omega^\nu (\mathbb{S}_k^\nu f, v) = \langle f, v \rangle_D = a_\omega^\nu (\mathbb{S}_\ell^\nu f, v).$$

To see the second statement, note that for  $v \in \mathcal{X}_k$  and  $w \in \mathcal{X}_\ell$ , there holds by definition of the inverse

$$a_\omega^\nu (\mathbb{S}_\ell^\nu (\mathbb{S}_k^\nu)^{-1} v, w) = ((\mathbb{S}_k^\nu)^{-1} v)(w) = a_\omega^\nu (v, w).$$

This and the positive definiteness of  $a_\omega^\nu$  conclude the proof.  $\square$

For the next lemma, we define the energy norm

$$\|u\|_{\omega, \nu} := (a_\omega^\nu(u, u))^{1/2}.$$

Note that (2.3)–(2.4) ensure  $a_{\min}(\omega) \|\cdot\|_{H^1(D)} \leq \|\cdot\|_{\omega, \nu} \leq a_{\max}(\omega)^{1/2} \|\cdot\|_{H^1(D)}$  for almost all  $\omega \in \Omega$  and for all  $\nu \in \mathbb{N}$ .

There holds the following variant of Céa's lemma:

LEMMA 3.3 (Céa's lemma). For  $v: \Omega \rightarrow \mathcal{X}_\ell$ ,  $\omega \in \Omega$ , and  $k \leq \ell$ , we have

$$\begin{aligned} \|(\mathbb{S}_\ell^\mu(\mathbb{S}_\ell^\mu)^{-1} - \mathbb{S}_k^\mu(\mathbb{S}_\ell^\mu)^{-1})v(\omega)\|_{H^1(D)} &\leq a_{\min}(\omega)^{-1} \inf_{w \in \mathcal{X}_k} \|v(\omega) - w\|_{\omega, \mu} \\ &\leq a_{\min}(\omega)^{-1} a_{\max}(\omega)^{1/2} \inf_{w \in \mathcal{X}_k} \|v(\omega) - w\|_{H^1(D)}. \end{aligned}$$

*Proof.* For almost all  $\omega \in \Omega$ , Galerkin orthogonality guarantees

$$\begin{aligned} a_\omega^\mu((\mathbb{S}_\ell^\mu(\mathbb{S}_\ell^\mu)^{-1} - \mathbb{S}_k^\mu(\mathbb{S}_\ell^\mu)^{-1})v, (\mathbb{S}_\ell^\mu(\mathbb{S}_\ell^\mu)^{-1} - \mathbb{S}_k^\mu(\mathbb{S}_\ell^\mu)^{-1})v) \\ = a_\omega^\mu((\mathbb{S}_\ell^\mu(\mathbb{S}_\ell^\mu)^{-1} - \mathbb{S}_k^\mu(\mathbb{S}_\ell^\mu)^{-1})v, \mathbb{S}_\ell^\mu(\mathbb{S}_\ell^\mu)^{-1}v - w) \end{aligned}$$

for all  $w \in \mathcal{X}_k$ . Since  $a_\omega^\nu$  is a scalar product with respective norm  $\|\cdot\|_{\omega, \nu}$ , we have

$$\begin{aligned} a_\omega^\mu((\mathbb{S}_\ell^\mu(\mathbb{S}_\ell^\mu)^{-1} - \mathbb{S}_k^\mu(\mathbb{S}_\ell^\mu)^{-1})v, \mathbb{S}_\ell^\mu(\mathbb{S}_\ell^\mu)^{-1}v - w) \\ \leq \|(\mathbb{S}_\ell^\mu(\mathbb{S}_\ell^\mu)^{-1} - \mathbb{S}_k^\mu(\mathbb{S}_\ell^\mu)^{-1})v\|_{\omega, \mu} \|\mathbb{S}_\ell^\mu(\mathbb{S}_\ell^\mu)^{-1}v - w\|_{\omega, \mu}. \end{aligned}$$

Ellipticity (2.3), norm equivalence  $a_{\min}(\omega)\|\cdot\|_{H^1(D)} \leq \|\cdot\|_{\omega, \nu} \leq a_{\max}(\omega)^{1/2}\|\cdot\|_{H^1(D)}$ , and the fact that  $\omega$  was arbitrary conclude the proof.  $\square$

The following lemma bounds the difference of the Galerkin projections  $\mathbb{S}_k^\nu(\mathbb{S}_\ell^\nu)^{-1}$  for different parameters  $\nu$ .

LEMMA 3.4. There holds for  $\ell, k, \nu, \mu \in \mathbb{N}$ , all  $v: \Omega \rightarrow \mathcal{X}_\ell$ , and all  $\omega \in \Omega$

$$\begin{aligned} \|(\mathbb{S}_k^\nu(\mathbb{S}_\ell^\nu)^{-1} - \mathbb{S}_k^\mu(\mathbb{S}_\ell^\mu)^{-1})v(\omega)\|_{H^1(D)} \\ \leq C_{\text{proj}}(\omega) \|(A^\nu - A^\mu)(\omega)\|_{L^\infty(D)} \inf_{w \in \mathcal{X}_k} \|v(\omega) - w\|_{H^1(D)}, \end{aligned}$$

where  $C_{\text{proj}}(\omega) := a_{\min}(\omega)^{-2} a_{\max}(\omega)$ .

*Proof.* For  $k \geq \ell$ , we have  $\mathbb{S}_k^\nu(\mathbb{S}_\ell^\nu)^{-1} = \text{id}_{\mathcal{X}_\ell} = \mathbb{S}_k^\mu(\mathbb{S}_\ell^\mu)^{-1}$  and thus the assertion holds trivially. Assume  $k < \ell$ . Define  $v_k := (\mathbb{S}_k^\nu(\mathbb{S}_\ell^\nu)^{-1} - \mathbb{S}_k^\mu(\mathbb{S}_\ell^\mu)^{-1})v: \Omega \rightarrow \mathcal{X}_\ell$ . Ellipticity (2.3) of  $a_\omega^\nu(\cdot, \cdot)$  together with Galerkin orthogonality shows for  $\omega \in \Omega$

$$a_{\min}(\omega) \|v_k(\omega)\|_{H^1(D)}^2 \leq a_\omega^\nu(v_k(\omega), v_k(\omega)) = a_\omega^\nu((\mathbb{S}_\ell^\nu(\mathbb{S}_\ell^\nu)^{-1} - \mathbb{S}_k^\mu(\mathbb{S}_\ell^\mu)^{-1})v(\omega), v_k(\omega)).$$

Since  $\mathbb{S}_\ell^\nu(\mathbb{S}_\ell^\nu)^{-1} = \text{id}_{\mathcal{X}_\ell} = \mathbb{S}_\ell^\mu(\mathbb{S}_\ell^\mu)^{-1}$ , we have

$$\begin{aligned} a_{\min}(\omega) \|v_k(\omega)\|_{H^1(D)}^2 &\leq a_\omega^\nu((\mathbb{S}_\ell^\mu(\mathbb{S}_\ell^\mu)^{-1} - \mathbb{S}_k^\mu(\mathbb{S}_\ell^\mu)^{-1})v(\omega), v_k(\omega)) \\ &= a_\omega^\mu((\mathbb{S}_\ell^\mu(\mathbb{S}_\ell^\mu)^{-1} - \mathbb{S}_k^\mu(\mathbb{S}_\ell^\mu)^{-1})v(\omega), v_k(\omega)) \\ &\quad + (a_\omega^\nu - a_\omega^\mu)((\mathbb{S}_\ell^\mu(\mathbb{S}_\ell^\mu)^{-1} - \mathbb{S}_k^\mu(\mathbb{S}_\ell^\mu)^{-1})v(\omega), v_k(\omega)). \end{aligned}$$

The first term on the right-hand side above is zero due to Galerkin orthogonality. Therefore, we obtain

$$(3.1) \quad \|v_k(\omega)\|_{H^1(D)}^2 \lesssim \|a_\omega^\nu - a_\omega^\mu\| \|(\mathbb{S}_\ell^\mu(\mathbb{S}_\ell^\mu)^{-1} - \mathbb{S}_k^\mu(\mathbb{S}_\ell^\mu)^{-1})v(\omega)\|_{H^1(D)} \|v_k(\omega)\|_{H^1(D)}.$$

As shown in Lemma 3.1, there holds  $\|a_\omega^\nu - a_\omega^\mu\| \leq \|(A^\nu - A^\mu)(\omega)\|_{L^\infty(D)}$ . Moreover, we have by Céa's lemma (Lemma 3.3)

$$\|(\mathbb{S}_\ell^\mu(\mathbb{S}_\ell^\mu)^{-1} - \mathbb{S}_k^\mu(\mathbb{S}_\ell^\mu)^{-1})v(\omega)\|_{H^1(D)} \leq a_{\min}(\omega)^{-1} a_{\max}(\omega)^{1/2} \inf_{w \in \mathcal{X}_k} \|v(\omega) - w\|_{H^1(D)}.$$

This together with (3.1) concludes the proof.  $\square$

For the statement of the next result, we recall the definition of the double difference

$$D_\ell^\nu := (u_\ell^\nu - u_{\ell-1}^\nu) - (u_\ell^{\nu-1} - u_{\ell-1}^{\nu-1}): \Omega \rightarrow \mathcal{X}_\ell.$$

LEMMA 3.5. *There holds for all  $\omega \in \Omega$*

$$(3.2) \quad \begin{aligned} \|D_\ell^\nu(\omega)\|_{H^1(D)} &\leq \|(\mathbb{S}_\ell^\nu - \mathbb{S}_{\ell-1}^\nu)(\mathbb{S}_\ell^\nu)^{-1}(\mathbb{S}_\ell^\nu - \mathbb{S}_\ell^{\nu-1})f\|_{H^1(D)} \\ &\quad + C_{\text{proj}}(\omega)\|(A^\nu - A^{\nu-1})(\omega)\|_{L^\infty(D)} \inf_{w \in \mathcal{X}_k} \|u_\ell^{\nu-1}(\omega) - w\|_{H^1(D)}. \end{aligned}$$

*Proof.* Straightforward expansion of the equation and  $\mathbb{S}_\ell^\nu(\mathbb{S}_k^\nu)^{-1} = \text{id}_{\mathcal{X}_k}$ ,  $k \leq \ell$  from Lemma 3.2 show

$$\begin{aligned} D_\ell^\nu &= ((\mathbb{S}_\ell^\nu - \mathbb{S}_{\ell-1}^\nu) - (\mathbb{S}_\ell^{\nu-1} - \mathbb{S}_{\ell-1}^{\nu-1}))f \\ &= (\mathbb{S}_\ell^\nu - \mathbb{S}_{\ell-1}^\nu)(\mathbb{S}_\ell^\nu)^{-1}(\mathbb{S}_\ell^\nu - \mathbb{S}_\ell^{\nu-1})f - (\mathbb{S}_{\ell-1}^{\nu-1}(\mathbb{S}_\ell^{\nu-1})^{-1}\mathbb{S}_\ell^{\nu-1} - \mathbb{S}_{\ell-1}^{\nu-1})f. \end{aligned}$$

The last term on the right-hand side satisfies

$$(3.3) \quad \begin{aligned} &\|(\mathbb{S}_{\ell-1}^{\nu-1}(\mathbb{S}_\ell^{\nu-1})^{-1}\mathbb{S}_\ell^{\nu-1} - \mathbb{S}_{\ell-1}^{\nu-1})f\|_{H^1(D)} \\ &\leq \|(\mathbb{S}_{\ell-1}^{\nu-1}(\mathbb{S}_\ell^{\nu-1})^{-1}\mathbb{S}_\ell^{\nu-1} - \mathbb{S}_{\ell-1}^{\nu-1})f\|_{H^1(D)} \\ &\quad + \|(\mathbb{S}_{\ell-1}^{\nu-1}(\mathbb{S}_\ell^{\nu-1})^{-1} - \mathbb{S}_{\ell-1}^{\nu-1}(\mathbb{S}_\ell^{\nu-1})^{-1})\mathbb{S}_\ell^{\nu-1}f\|_{H^1(D)}. \end{aligned}$$

The first term on the right-hand side satisfies for all  $v \in \mathcal{X}_{\ell-1}$

$$a_\omega^\nu((\mathbb{S}_{\ell-1}^{\nu-1}(\mathbb{S}_\ell^{\nu-1})^{-1}\mathbb{S}_\ell^{\nu-1} - \mathbb{S}_{\ell-1}^{\nu-1})f, v) = a_\omega^\nu((\mathbb{S}_{\ell-1}^{\nu-1}(\mathbb{S}_\ell^{\nu-1})^{-1}\mathbb{S}_\ell^{\nu-1} - \mathbb{S}_{\ell-1}^{\nu-1})f, v) = 0$$

and thus  $\|(\mathbb{S}_{\ell-1}^{\nu-1}(\mathbb{S}_\ell^{\nu-1})^{-1}\mathbb{S}_\ell^{\nu-1} - \mathbb{S}_{\ell-1}^{\nu-1})f\|_{H^1(D)} = 0$ . For the second term on the right-hand side of (3.3), Lemma 3.4 with  $\mu = \nu - 1$  and  $k = \ell - 1$  proves

$$\begin{aligned} &\|(\mathbb{S}_{\ell-1}^{\nu-1}(\mathbb{S}_\ell^{\nu-1})^{-1} - \mathbb{S}_{\ell-1}^{\nu-1}(\mathbb{S}_\ell^{\nu-1})^{-1})\mathbb{S}_\ell^{\nu-1}f\|_{H_0^1(D)} \\ &\lesssim \|A^\nu(\omega) - A^{\nu-1}(\omega)\|_{L^\infty(D)} \inf_{v \in \mathcal{X}_{\ell-1}} \|u_\ell^{\nu-1}(\omega) - v\|_{H^1(D)}. \end{aligned}$$

Altogether, this concludes the proof.  $\square$

The following result is well-known and we reprove it in our setting for the convenience of the reader.

LEMMA 3.6 (Aubin-Nitsche duality). *There holds for all  $v \in H_0^1(D)$  that*

$$\|v - \mathbb{S}_\ell^\nu(\mathbb{S}_\infty^\nu)^{-1}v\|_{L^2(D)} \leq C_{\text{approx}} \frac{C_{\text{reg}}}{a_{\min}(\omega)^2} (1 + \|A^\nu(\omega)\|_{W^{1,\infty}(D)}) h_\ell \|v\|_{H^1(D)}.$$

*Proof.* Let  $\iota: L^2(D) \rightarrow H^{-1}(D)$  be the usual embedding via the  $L^2(D)$ -scalar product. Define  $V := v - \mathbb{S}_\ell^\nu(\mathbb{S}_\infty^\nu)^{-1}v$ . We have with Galerkin orthogonality and by symmetry of  $a_\omega^\nu$  for all  $w \in \mathcal{X}_\ell$

$$\begin{aligned} \|v - \mathbb{S}_\ell^\nu(\mathbb{S}_\infty^\nu)^{-1}v\|_{L^2(D)}^2 &= a_\omega^\nu(\mathbb{S}_\infty^\nu \circ \iota(V), V) = a_\omega^\nu(\mathbb{S}_\infty^\nu \circ \iota(V) - w, V) \\ &\leq \|\mathbb{S}_\infty^\nu \circ \iota(V) - w\|_{H^1(D)} \|V\|_{H^1(D)}. \end{aligned}$$



Since  $w \in \mathcal{X}_\ell$  was arbitrary, we get with (2.5) and (2.6)

$$\begin{aligned} & \|v - \mathbb{S}_\ell^\nu (\mathbb{S}_\infty^\nu)^{-1} v\|_{L^2(D)}^2 \\ & \leq C_{\text{approx}} h_\ell \|\mathbb{S}_\infty^\nu \circ \iota(V)\|_{H^2(D)} \|V\|_{H^1(D)} \\ & \leq C_{\text{approx}} \frac{C_{\text{reg}}}{a_{\min}(\omega)^2} (1 + \|A^\nu(\omega)\|_{W^{1,\infty}(D)}) h_\ell \|v - \mathbb{S}_\ell^\nu (\mathbb{S}_\infty^\nu)^{-1} v\|_{L^2(D)} \|V\|_{H^1(D)}. \end{aligned}$$

With Lemma 2.1, we show  $\|V\|_{H^1(D)} \leq (1 + a_{\min}(\omega)^{-1} a_{\min}(\omega)) \|v\|_{H^1(D)}$  and thus conclude the proof.  $\square$

The following result bounds the first term on the right-hand side of the estimate in Lemma 3.5 by an error estimate in product form.

LEMMA 3.7. *There holds for all  $\omega \in \Omega$*

$$\begin{aligned} & \|(\mathbb{S}_\ell^\nu - \mathbb{S}_{\ell-1}^\nu)(\mathbb{S}_\ell^\nu)^{-1}(\mathbb{S}_\ell^\nu - \mathbb{S}_\ell^{\nu-1})f\|_{H^1(D)} \\ & \leq \tilde{C}_{\text{prod}}(\omega) h_\ell \|(A^\nu - A^{\nu-1})(\omega)\|_{W^{1,\infty}(D)} \|f\|_{L^2(D)}, \end{aligned}$$

where  $\tilde{C}_{\text{prod}}(\omega) \simeq C_{\text{unif}} a_{\min}(\omega)^{-5} a_{\max}(\omega)^{1/2} (1 + \max_{i \in \{0,1\}} \|A^{\nu-i}(\omega)\|_{W^{1,\infty}(D)})^2 > 0$ .

*Proof.* First, Céa's lemma (Lemma 3.3) shows for  $v: \Omega \rightarrow \mathcal{X}_\ell$

$$\|(\mathbb{S}_\ell^\nu - \mathbb{S}_{\ell-1}^\nu(\omega))(\mathbb{S}_\ell^\nu)^{-1} v\|_{H^1(D)} \leq a_{\min}(\omega)^{-1} \inf_{w \in \mathcal{X}_{\ell-1}} \|v(\omega) - w\|_{\omega, \nu}.$$

Let  $v := (\mathbb{S}_\ell^\nu - \mathbb{S}_\ell^{\nu-1})f$  and choose  $w := \mathbb{S}_{\ell-1}^\nu (\mathbb{S}_\infty^\nu)^{-1} v$ . Then, there holds with Galerkin orthogonality  $a_\omega^\nu(w, v - w) = a_\omega^\nu(v - \mathbb{S}_{\ell-1}^\nu (\mathbb{S}_\infty^\nu)^{-1} v, w) = 0$  and hence

$$\begin{aligned} \|v - w\|_{\omega, \nu}^2 & = a_\omega^\nu(v, v - w) = a_\omega^\nu(u^\nu - \mathbb{S}_\ell^{\nu-1} f, v - w) \\ & = a_\omega^{\nu-1}(u^\nu - \mathbb{S}_\ell^{\nu-1} f, v - w) + (a_\omega^\nu - a_\omega^{\nu-1})(u^\nu - \mathbb{S}_\ell^{\nu-1} f, v - w) \\ & = a_\omega^{\nu-1}(u^\nu, v - w) - \langle f, v - w \rangle_D + (a_\omega^\nu - a_\omega^{\nu-1})(u^\nu - \mathbb{S}_\ell^{\nu-1} f, v - w), \end{aligned}$$

where we just inserted and subtracted  $a_\omega^{\nu-1}(\cdot, \cdot)$ . This, then leads to

$$\begin{aligned} \|v - w\|_{\omega, \nu}^2 & = a_\omega^{\nu-1}(u^\nu, v - w) - a_\omega^\nu(u^\nu, v - w) + (a_\omega^\nu - a_\omega^{\nu-1})(u^\nu - \mathbb{S}_\ell^{\nu-1} f, v - w) \\ & = -(a_\omega^\nu - a_\omega^{\nu-1})(\mathbb{S}_\ell^{\nu-1} f, v - w) \\ & = -(a_\omega^\nu - a_\omega^{\nu-1})(u^{\nu-1}, v - w) - (a_\omega^\nu - a_\omega^{\nu-1})(u_\ell^{\nu-1} - u^{\nu-1}, v - w), \end{aligned}$$

where we used  $\mathbb{S}_\ell^{\nu-1} f = u_\ell^{\nu-1}$  and we added and subtracted the corresponding exact solution  $u^{\nu-1}$ . Using the definition of the bilinear forms as well as integration by parts, the above reads

$$\begin{aligned} \|v - w\|_{\omega, \nu}^2 & = \int_D (\nabla(A^\nu - A^{\nu-1}) \cdot \nabla u^{\nu-1} + (A^\nu - A^{\nu-1}) \Delta u^{\nu-1})(v - w) dx \\ & \quad - (a_\omega^\nu - a_\omega^{\nu-1})(u_\ell^{\nu-1} - u^{\nu-1}, v - w) \\ & \leq \|A^\nu - A^{\nu-1}\|_{W^{1,\infty}(D)} \|u^{\nu-1}\|_{H^2(D)} \|v - w\|_{L^2(D)} \\ & \quad + \|a_\omega^\nu - a_\omega^{\nu-1}\| \|u_\ell^{\nu-1} - u^{\nu-1}\|_{H^1(D)} \|v - w\|_{H^1(D)}. \end{aligned}$$

Finally, Lemma 3.6 shows

$$\begin{aligned} \|v - w\|_{L^2(D)} & \lesssim h_{\ell-1} a_{\min}(\omega)^{-2} (1 + \|A^{\nu-1}(\omega)\|_{W^{1,\infty}(D)}) \|v\|_{H^1(D)} \\ & \lesssim h_{\ell-1} a_{\min}(\omega)^{-3} (1 + \|A^{\nu-1}(\omega)\|_{W^{1,\infty}(D)}) \|f\|_{L^2(D)}, \end{aligned}$$

where the last estimate uses Lemma 2.1. Assumption (2.5), together with the Céa lemma (Lemma 3.3), implies

$$\|u_\ell^{\nu-1} - u^{\nu-1}\|_{H^1(D)} \lesssim a_{\min}(\omega)^{-1} a_{\max}(\omega)^{1/2} h_\ell \|u^{\nu-1}\|_{H^2(D)}.$$

Together with (2.6), we obtain

$$\|u^{\nu-1}\|_{H^2(D)} \lesssim a_{\min}(\omega)^{-2} (1 + \|A^{\nu-1}(\omega)\|_{W^{1,\infty}(D)}) \|f\|_{L^2(D)} \quad \square$$

and thus conclude the proof.

Finally, we have collected all the ingredients to obtain the full error estimate in product form.

PROPOSITION 3.8. *There holds for all  $\omega \in \Omega$*

$$\|D_\ell^\nu(\omega)\|_{H^1(D)} \leq C_{\text{prod}}(\omega) h_\ell \|(A^\nu - A^{\nu-1})(\omega)\|_{W^{1,\infty}(D)} \|f\|_{L^2(D)},$$

where  $C_{\text{prod}}(\omega) \simeq \widetilde{C}_{\text{prod}}(\omega)(1 + a_{\max}(\omega)) > 0$ .

*Proof.* The first term on the right-hand side of (3.2) is bounded by Lemma 3.7. For the second term, we use (2.5) together with (2.6) to obtain a similar bound. Finally, we exploit that  $h_\ell \geq C_{\text{unif}} h_{\ell-1}$  and conclude the proof.  $\square$

Since we are interested in the error of the goal functional  $G(\cdot)$ , we may exploit a standard Aubin-Nitsche duality argument to double the rate of convergence.

THEOREM 3.9. *There holds for all  $\omega \in \Omega$*

$$|G(D_\ell^\nu(\omega))| \leq \overline{C}_{\text{prod}}(\omega) h_\ell^2 \min\{1, \|(A^\nu - A^{\nu-1})(\omega)\|_{W^{1,\infty}(D)}\} \|f\|_{L^2(D)} \|g\|_{L^2(D)}$$

with  $\overline{C}_{\text{prod}}(\omega) > 0$  depending on  $C_{\text{prod}}(\omega)$  from Proposition 3.8 via  $\overline{C}_{\text{prod}}(\omega) \simeq a_{\min}(\omega)^{-5} a_{\max}(\omega) \|A^\nu(\omega)\|_{W^{1,\infty}(D)} \|A^{\nu-1}(\omega)\|_{W^{1,\infty}(D)} C_{\text{prod}}(\omega)$ .

*Proof.* Let  $g^\nu \in H_0^1(\Omega)$  such that  $G(\cdot) = a_\omega^\nu(\cdot, g^\nu)$  (note that such a function always exists due to the ellipticity (2.3) of  $a_\omega^{\nu-1}$ ). There holds for  $v, w \in \mathcal{X}_{\ell-1}$

$$\begin{aligned} G(D_\ell^\nu) &= a_\omega^\nu(u_\ell^\nu - u_{\ell-1}^\nu, g^\nu) - a_\omega^{\nu-1}(u_\ell^{\nu-1} - u_{\ell-1}^{\nu-1}, g^{\nu-1}) \\ &= a_\omega^\nu(u_\ell^\nu - u_{\ell-1}^\nu, g^\nu - v) - a_\omega^{\nu-1}(u_\ell^{\nu-1} - u_{\ell-1}^{\nu-1}, g^{\nu-1} - v), \end{aligned}$$

where we used Galerkin orthogonality (Lemma 3.2) to insert  $v \in \mathcal{X}_{\ell-1}$ . Adding and subtracting of  $a_\omega^\nu(\cdot, \cdot)$  leads to

$$\begin{aligned} G(D_\ell^\nu) &= a_\omega^\nu(u_\ell^\nu - u_{\ell-1}^\nu, g^\nu - v) - a_\omega^\nu(u_\ell^{\nu-1} - u_{\ell-1}^{\nu-1}, g^{\nu-1} - v) \\ &\quad + (a_\omega^\nu - a_\omega^{\nu-1})(u_\ell^{\nu-1} - u_{\ell-1}^{\nu-1}, g^{\nu-1} - v) \\ &= a_\omega^\nu(u_\ell^\nu - u_{\ell-1}^\nu, g^{\nu-1} - v) - a_\omega^\nu(u_\ell^{\nu-1} - u_{\ell-1}^{\nu-1}, g^{\nu-1} - v) \\ &\quad + (a_\omega^\nu - a_\omega^{\nu-1})(u_\ell^{\nu-1} - u_{\ell-1}^{\nu-1}, g^{\nu-1} - v) + a_\omega^\nu(u_\ell^\nu - u_{\ell-1}^\nu, g^\nu - g^{\nu-1} - w), \end{aligned}$$

where we added and subtracted  $a_\omega^\nu(u_\ell^\nu - u_{\ell-1}^\nu, g^{\nu-1})$  and inserted  $w \in \mathcal{X}_{\ell-1}$  using Galerkin orthogonality (Lemma 3.2). Recalling the definition of  $D_\ell^\nu$ , we end up with

$$\begin{aligned} G(D_\ell^\nu) &= a_\omega^\nu(D_\ell^\nu, g^{\nu-1} - v) + (a_\omega^\nu - a_\omega^{\nu-1})(u_\ell^{\nu-1} - u_{\ell-1}^{\nu-1}, g^{\nu-1} - v) \\ &\quad + a_\omega^\nu(u_\ell^\nu - u_{\ell-1}^\nu, g^\nu - g^{\nu-1} - w). \end{aligned}$$

Lemma 3.1 and the C ea lemma (Lemma 3.3) together with (2.5) and (2.6) allows us to estimate

$$\begin{aligned}
|G(D_\ell^\nu)| &\lesssim a_{\max}(\omega) \|D_\ell^\nu\|_{H^1(D)} \|g^{\nu-1} - v\|_{H^1(D)} \\
&\quad + \|A^\nu - A^{\nu-1}\|_{L^\infty(D)} \|u_\ell^{\nu-1} - u_{\ell-1}^{\nu-1}\|_{H^1(D)} \|g^{\nu-1} - v\|_{H^1(D)} \\
&\quad + \|u_\ell^\nu - u_{\ell-1}^\nu\|_{H^1(D)} \|g^\nu - g^{\nu-1} - w\|_{H^1(D)} \\
(3.4) \quad &\lesssim a_{\max}(\omega) \|D_\ell^\nu\|_{H^1(D)} \|g^{\nu-1} - v\|_{H^1(D)} \\
&\quad + a_{\min}(\omega)^{-3} a_{\max}(\omega)^{1/2} (1 + \|A^{\nu-1}(\omega)\|_{W^{1,\infty}(D)}) \|f\|_{L^2(D)} h_\ell \\
&\quad \left( \|A^\nu - A^{\nu-1}\|_{L^\infty(D)} \|g^{\nu-1} - v\|_{H^1(D)} + \|g^\nu - g^{\nu-1} - w\|_{H^1(D)} \right).
\end{aligned}$$

Since  $G(\cdot) = \int_D g(x)(\cdot) dx$  for some  $g \in L^2(D)$ , we obtain from (2.6) that  $g^\nu, g^{\nu-1} \in H^2(D)$ . Therefore, and since  $v \in \mathcal{X}_{\ell-1}$  was arbitrary, (2.5) and (2.6) show

$$\inf_{v \in \mathcal{X}_{\ell-1}} \|g^{\nu-1} - v\|_{H^1(D)} \lesssim a_{\min}(\omega)^{-2} (1 + \|A^{\nu-1}(\omega)\|_{W^{1,\infty}(D)}) h_\ell \|g\|_{L^2(D)}.$$

Moreover, there holds for all  $v \in H_0^1(D)$

$$\begin{aligned}
a_\omega^\nu(g^\nu - g^{\nu-1}, v) &= \langle g, v \rangle_D - a_\omega^\nu(g^{\nu-1}, v) = (a^{\nu-1} - a^\nu)(g^{\nu-1}, v) \\
&= \int_D (\nabla(A^\nu - A^{\nu-1}) \cdot \nabla g^{\nu-1} + (A^\nu - A^{\nu-1}) \Delta g^{\nu-1}) v dx.
\end{aligned}$$

It is easy to see that the right-hand side is of the form  $\langle r, v \rangle_D$  for some  $r \in L^2(D)$  with

$$\|r\|_{L^2(D)} \leq 2 \|A^\nu - A^{\nu-1}\|_{W^{1,\infty}(D)} \|g^{\nu-1}\|_{H^2(D)} \lesssim \|A^\nu - A^{\nu-1}\|_{W^{1,\infty}(D)} \|g\|_{L^2(D)}.$$

Therefore, (2.6) shows

$$\|g^\nu - g^{\nu-1}\|_{H^2(D)} \lesssim a_{\min}(\omega)^{-2} (1 + \|A^\nu(\omega)\|_{W^{1,\infty}(D)}) \|A^\nu - A^{\nu-1}\|_{W^{1,\infty}(D)} \|g\|_{L^2(D)}.$$

Since  $w \in \mathcal{X}_{\ell-1}$  in (3.4) was arbitrary, the same argument and (2.5) show

$$\begin{aligned}
\inf_{w \in \mathcal{X}_{\ell-1}} \|g^\nu - g^{\nu-1} - w\|_{H^1(D)} \\
\lesssim h_\ell a_{\min}(\omega)^{-2} (1 + \|A^\nu(\omega)\|_{W^{1,\infty}(D)}) \|A^\nu - A^{\nu-1}\|_{W^{1,\infty}(D)} \|g\|_{L^2(D)}.
\end{aligned}$$

Altogether, we conclude the proof by use of Proposition 3.8, the above estimates, and insertion in (3.4). The minimum in the statement follows from standard arguments which we will sketch briefly. There holds

$$G(u_\ell^\nu - u_{\ell-1}^\nu) = a_\omega^\nu(u_\ell^\nu - u_{\ell-1}^\nu, g^\nu) = a_\omega^\nu(u_\ell^\nu - u_{\ell-1}^\nu, g^\nu - v)$$

for all  $v \in \mathcal{X}_{\ell-1}$ . As above, choosing  $v = \mathbb{S}_\ell^\nu(\mathbb{S}_\infty^\nu)^{-1} g^\nu$  and Lemma 3.3 together with (2.5) leads to

$$\begin{aligned}
|G(u_\ell^\nu - u_{\ell-1}^\nu)| &\lesssim \|u_\ell^\nu - u_{\ell-1}^\nu\|_{H^1(D)} h_{\ell-1} \|g\|_{L^2(D)} \\
&\lesssim h_{\ell-1}^2 \|f\|_{L^2(D)} \|g\|_{L^2(D)}.
\end{aligned}$$

This concludes the proof.  $\square$

**4. Approximation of the random coefficient.** This section gives two examples of how to choose the random coefficient  $A(x, \omega)$  as well as the approximations  $A^\nu(x, \omega)$ . The first one will be the standard KL-expansion, whereas the second one will come from a novel technique to generate random fields by use of  $H$ -matrices.

**4.1. KL expansion.** In this section, we assume  $\Omega = [0, 1]^\mathbb{N}$  and define  $\omega = (\omega_i)_{i \in \mathbb{N}}$ . We assume that  $A^\nu$  is of the form

$$(4.1) \quad A^\nu(x, \omega) := \phi_0(x) + \sum_{j=1}^{s_\nu} \psi_j(\omega_j) \phi_j(x)$$

for functions  $\phi_j \in W^{1, \infty}(D)$  and  $\psi_j \in L^\infty([0, 1], [-C_\psi, C_\psi])$  for some fixed  $C_\psi > 0$ . We assume that the series converges absolutely in  $W^{1, \infty}(D)$  for all  $\omega \in \Omega$  and hence define

$$A(x, \omega) := A^\infty(x, \omega) := \phi_0(x) + \sum_{j=1}^{\infty} \psi_j(\omega_j) \phi_j(x).$$

Moreover, we assume that (2.2) holds.

**THEOREM 4.1.** *Under the assumptions of the current section, there holds*

$$(4.2) \quad \|G(D_\ell^\nu)\|_{L^\infty(\Omega)} \leq C_{\text{KL}} h_\ell^2 \sum_{i=s_\nu-1+1}^{s_\nu} \|\phi_i\|_{W^{1, \infty}(D)} \|f\|_{L^2(D)} \|g\|_{L^2(D)}.$$

The constant  $C_{\text{KL}} > 0$  depends on  $C_\psi$  but is independent of  $\ell$ ,  $\nu$ , and  $\omega$ .

*Proof.* The estimate follows immediately by definition of  $A^\nu$  and Theorem 3.9.  $\square$

**4.2. KL expansion with uniform random variables.** In many cases, it is possible to reduce (4.1) to the simplified form

$$(4.3) \quad A^\nu(x, \omega) := \phi_0(x) + \sum_{j=1}^{s_\nu} \omega_j \phi_j(x),$$

where now  $\Omega = [-1/2, 1/2]^\mathbb{N}$  and  $\text{ess inf}_{x \in D} \phi_0(x) > 0$ .

*Remark 4.2.* Note that theoretically, the case from Section 4.1 can always be reduced to the present case. However, in many cases, this requires the user to compute all function  $\phi_j$  in advance, which of course is computationally impractical.

It turns out that in this case, an improved version of Theorem 3.9 (see Theorem 4.7 at the end of this section) can be derived by arguments already used for quasi-Monte Carlo estimates (see, e.g., the works [1, 2] and the references therein). Given a subset  $\Omega' \subseteq \prod_{j \in \mathbb{N}} \mathbb{C}$ , we define for all  $j \in \mathbb{N}$

$$\Omega'_j := \{\omega_j \in \mathbb{C} : \exists \omega_i \in \mathbb{C}, i \in \mathbb{N} \setminus \{j\} \text{ such that } \omega = (\omega_1, \omega_2, \dots) \in \Omega'\}.$$

**LEMMA 4.3.** *Assume that  $\Omega' \supseteq \Omega$  is such that all results of Section 3 hold true with  $\Omega'$  instead of  $\Omega$ . Then the function  $F: \Omega'_j \rightarrow \mathbb{C}, \omega_j \mapsto G(S_\ell^\nu(\omega)f)$  is holomorphic for all  $j \in \mathbb{N}$ .*

*Proof.* Fix  $j \in \mathbb{N}$ . Given  $z \in \mathbb{C}$ , define  $\omega + z \in \mathbb{C}^{\mathbb{N}}$  by  $(\omega + z)_i = \omega_i$  for all  $i \neq j$  and  $(\omega + z)_j = \omega_j + z$ . Let  $z$  be sufficiently small such that there exists  $\varepsilon \geq 2|z|$  with  $B_\varepsilon(\omega) \subseteq \Omega'$ . By definition, we have for  $v \in \mathcal{X}_\ell$

$$\begin{aligned} 0 &= a_{\omega+z}^\nu(\mathbb{S}_\ell^\nu(\omega+z)f, v) - a_\omega^\nu(\mathbb{S}_\ell^\nu(\omega)f, v) \\ &= \int_D (A^\nu(x, \omega+z) - A^\nu(x, \omega)) \nabla \mathbb{S}_\ell^\nu(\omega+z)f \cdot \nabla v \, dx + a_\omega^\nu(\mathbb{S}_\ell^\nu(\omega+z)f - \mathbb{S}_\ell^\nu(\omega)f, v). \end{aligned}$$

Let  $g^\nu \in \mathcal{X}_\ell$  denote the representer of  $G(\cdot)|_{\mathcal{X}_\ell}$  with respect to  $a_\omega^\nu$ . This and the above allows us to compute

$$(4.4) \quad \begin{aligned} \frac{G(\mathbb{S}_\ell^\nu(\omega+z)f) - G(\mathbb{S}_\ell^\nu(\omega)f)}{z} &= \frac{a_\omega^\nu(\mathbb{S}_\ell^\nu(\omega+z)f - \mathbb{S}_\ell^\nu(\omega)f, g^\nu)}{z} \\ &= - \int_D \frac{A^\nu(x, \omega+z) - A^\nu(x, \omega)}{z} \nabla \mathbb{S}_\ell^\nu(\omega+z)f \cdot \nabla g^\nu \, dx. \end{aligned}$$

Since  $A^\nu$  is holomorphic, Cauchy's integral formula shows for  $B_\varepsilon(\omega_j) \subset \Omega'_j$  that

$$\begin{aligned} & \left| \frac{A^\nu(x, \omega+z) - A^\nu(x, \omega)}{z} - \partial_{\omega_j} A^\nu(x, \omega) \right| \\ &= \frac{1}{2\pi} \left| \int_{\partial B_\varepsilon(\omega_j)} \frac{1}{z} \left( \frac{A^\nu(x, y)}{(y - (\omega_j + z))} - \frac{A^\nu(x, y)}{(y - \omega_j)} \right) - \frac{A^\nu(x, y)}{(y - \omega_j)^2} \, dy \right| \\ &= \frac{1}{2\pi} \left| \int_{\partial B_\varepsilon(\omega_j)} \frac{A^\nu(x, y)}{(y - \omega_j - z)(y - \omega_j)} - \frac{A^\nu(x, y)}{(y - \omega_j)^2} \, dy \right| \\ &= \frac{1}{2\pi} \left| \int_{\partial B_\varepsilon(\omega_j)} \frac{A^\nu(x, y)z}{(y - \omega_j - z)(y - \omega_j)^2} \, dy \right| \\ &\lesssim \varepsilon^{-2} \|A^\nu\|_{L^\infty(\Omega \times D)} |z|. \end{aligned}$$

This uniform convergence in  $|z|$  together with Lemma 3.1 shows that passing to the limit  $z \rightarrow 0$  in (4.4) leads to

$$\partial_{\omega_j} G(\mathbb{S}_\ell^\nu f) = - \int_D \partial_{\omega_j} A^\nu(x, \omega) \nabla \mathbb{S}_\ell^\nu(\omega)f \cdot \nabla g^\nu \, dx \in \mathbb{C}.$$

This shows that  $F$  is complex differentiable and thus holomorphic.  $\square$

LEMMA 4.4. *Let  $(\varrho_j)_{j \in \mathbb{N}}$  be a positive sequence such that*

$$\Omega \subset \Omega' := \prod_{j \in \mathbb{N}} B_{1+\varrho_j}(0)$$

*and that all the results of Section 3 hold true with  $\Omega'$  instead of  $\Omega$ . Given  $\ell, \nu \in \mathbb{N}$ , the function  $F_\ell^\nu: \Omega \rightarrow \mathbb{R}$ ,  $\omega \mapsto G(D_\ell^\nu(\omega))$  satisfies*

$$\begin{aligned} & \frac{\|\partial_\omega^\alpha F_\ell^\nu\|_{L^\infty(\Omega)}}{\|f\|_{L^2(D)} \|g\|_{L^2(D)}} \\ & \leq \begin{cases} 0 & \sum_{i=s_\nu+1}^\infty \alpha_i > 0, \\ C_{\text{der}} \frac{\alpha h_\ell^2}{\prod_{i=1}^\infty \varrho_i^{\alpha_i}} \min\{1, \sup_{\omega \in \Omega'} \|A^\nu - A^{\nu-1}\|_{W^{1,\infty}(D)}\} & \text{else,} \end{cases} \end{aligned}$$

*for all multi-indices  $\alpha \in \mathbb{N}^{\mathbb{N}}$  with  $|\alpha| < \infty$ . The constant  $C_{\text{der}} > 0$  depends only on  $C_{\text{prod}}$  from Theorem 3.9.*

*Proof.* For brevity of presentation, we fix  $\ell$  and  $\nu$  and write  $F := F_\ell^\nu$ . Lemma 4.3 shows that  $F$  can be extended to a function  $F: \Omega' \rightarrow \mathbb{C}$ , which is holomorphic in each coordinate  $\omega_j$ . Moreover, Lemma 3.1 proves that  $F$  is uniformly continuous in  $\Omega$ . Therefore, we obtain immediately by induction that  $F$  satisfies the multidimensional analog of Cauchy's integral formula for all  $\omega \in \Omega'$

$$F(\omega) = (2\pi i)^{-n} \int_{\partial B_{\varepsilon_1}(\omega_{d_1})} \cdots \int_{\partial B_{\varepsilon_n}(\omega_{d_n})} \frac{F(z)}{(z_1 - \omega_{d_1}) \cdots (z_n - \omega_{d_n})} dz_1 \cdots dz_n,$$

where  $(d_1, \dots, d_n) \in \mathbb{N}^n$  contains exactly  $n$ -different dimensions and  $\varepsilon_i > 0$ ,  $i = 1, \dots, n$  are chosen sufficiently small such that the integration domains of the contour integrals above are contained in  $\Omega'$ . This shows immediately that for any multi-index  $\alpha \in \mathbb{N}^{\mathbb{N}}$  with  $|\alpha| < \infty$ ,  $\partial_\omega^\alpha F$  is holomorphic in each variable. Thus, iterated application of Cauchy's integral formula shows for all  $\omega \in \Omega$  that

$$\partial_\omega^\alpha F(\omega) = \left( \prod_{\substack{i=1 \\ \alpha_i \neq 0}}^{\infty} \frac{\alpha_i!}{2\pi i} \right) \int_{\prod_{\substack{i=1 \\ \alpha_i \neq 0}}^{\infty} \partial B_{\varepsilon_i}(\omega_i)} \frac{F(z)}{\prod_{\substack{i=1 \\ \alpha_i \neq 0}}^{\infty} (z_i - \omega_i)^{\alpha_i+1}} dz.$$

This shows immediately

$$|\partial_\omega^\alpha F(\omega)| \leq \left( \prod_{\substack{i=1 \\ \alpha_i \neq 0}}^{\infty} \frac{\alpha_i!}{2\pi} 2\pi \varrho_i^{-\alpha_i} \right) \|F\|_{L^\infty(\Omega')} \leq \alpha! \left( \prod_{i=1}^{\infty} \varrho_i^{-\alpha_i} \right) \|F\|_{L^\infty(\Omega')}.$$

This and Theorem 3.9 with  $A^\nu(\omega) = \phi_0 + \sum_{i=1}^\nu \omega_i \phi_i$  conclude the proof.  $\square$

LEMMA 4.5. *Define for sufficiently small  $\delta > 0$*

$$\beta_i := \frac{\|\phi_i\|_{W^{1,\infty}(D)}}{(\text{ess inf}_{x \in D} \phi_0(x) - 2\delta)}.$$

*Given  $\ell, \nu \in \mathbb{N}$ , the function  $F: \Omega \rightarrow \mathbb{R}$ ,  $\omega \mapsto G(D_\ell^\nu(\omega))$  satisfies*

$$\|\partial_\omega^\alpha F\|_{L^\infty(\Omega)} \leq \tilde{C}_{\text{der}} \begin{cases} 0 & \sum_{i=s_\nu+1}^{\infty} \alpha_i > 0, \\ \left( \prod_{i=1}^{s_\nu} \beta_i^{\alpha_i} \right) h_\ell^2 \|f\|_{L^2(D)} \|g\|_{L^2(D)} & \text{else,} \end{cases}$$

*for all multi-indices  $\alpha \in \mathbb{N}_0^{\mathbb{N}}$  with  $|\alpha| \leq 2$ . The constant  $\tilde{C}_{\text{der}} > 0$  depends only on  $C_{\text{der}}$ ,  $\delta$ , and  $(\phi_j)_{j \in \mathbb{N}}$ .*

*Proof.* Given  $\alpha \in \mathbb{N}_0^{\mathbb{N}}$  with  $|\alpha| \leq 2$  an admissible sequence  $(\varrho_j)_{j \in \mathbb{N}}$  in Lemma 4.4 is, given  $\varepsilon > 0$ ,

$$\varrho_j := \begin{cases} (\text{inf}_{x \in D} \phi_0(x) - 2\delta) \alpha_j / 2 \|\phi_j\|_{W^{1,\infty}(D)}^{-1} & \text{for all } j \in \mathbb{N} \text{ with } \alpha_j > 0, \\ \varepsilon & \text{for all } j \in \mathbb{N} \text{ with } \alpha_j = 0. \end{cases}$$

This sequence satisfies

$$\inf_{\omega_i \in B_{1+\varrho_i}(0): i \in \mathbb{N}} \Re(\phi_0 + \sum_{i=1}^\nu \omega_i \phi_i) \geq \phi_0 - (\text{ess inf}_{x \in D} \phi_0(x) - 2\delta) - \varepsilon \sum_{i=1}^{\infty} \|\phi_j\|_{L^\infty(D)} \geq \delta$$

for sufficiently small  $\varepsilon > 0$  (here  $\Re$  denotes the real part). Moreover, the term  $\|\phi_0 + \sum_{i=1}^\nu \omega_i \phi_i\|_{W^{1,\infty}(D)}$  remains uniformly bounded in  $\Omega' := \prod_{i=1}^{\infty} B_{1+\varrho_i}(0)$ . This

ensures that  $\Omega'$  satisfies all the assumptions required for  $\Omega$  and thus all results of Section 3 remain valid for  $\Omega'$  instead of  $\Omega$ . In particular, the constant  $C_{\text{prod}}(\omega)$  from Theorem 3.9 is uniformly bounded in  $\omega \in \Omega'$ . It is also obvious that  $\omega \mapsto A^\nu(x, \omega)$  is holomorphic in each coordinate in  $\Omega'$ , with derivative

$$\partial_{\omega_j} A^\nu(x, \omega) = \begin{cases} \phi_j(x) & \text{for } j \leq s_\nu, \\ 0 & \text{else.} \end{cases}$$

Moreover, since  $|\alpha| \leq 2$  there holds

$$\prod_{i=1}^{\infty} \varrho_i^{-\alpha_i} \leq \prod_{i=1}^{\infty} \beta_i^{\alpha_i}.$$

This, together with Lemma 4.4 concludes the proof.  $\square$

LEMMA 4.6. *Let  $g \in L^\infty(\Omega)$  be sufficiently smooth and let  $g$  depend only on the first  $s \in \mathbb{N}$  dimensions, i.e.,  $\partial_{\omega_i} g = 0$  for all  $i > s$ . For  $0 \leq r \leq s$  and  $x = (x_1, x_2, \dots, x_s) \in \Omega^s$ , define the function space*

$$\mathcal{P}_r^s(\Omega) := \text{span}\{f \in L^\infty(\Omega) : f(x) = \sum_{i=r+1}^s \alpha(x_1, \dots, x_r) x_i, \alpha(x_1, \dots, x_r) \in \mathbb{R}\}.$$

Assume that  $\omega \in \Omega$  with  $\omega_i = 0$  for all  $i > r$  implies  $g(\omega) = 0$ . Then, there holds

$$\|g(\omega)\|_{L^\infty(\Omega)} \leq \sum_{i=r+1}^s \|\partial_{\omega_i} g\|_{L^\infty(\Omega)}.$$

Moreover, there exists  $g_0 \in \mathcal{P}_r^s(\Omega)$  such that

$$\|g(\omega) - g_0(\omega)\|_{L^\infty(\Omega)} \leq \frac{1}{2} \sum_{i=r+1}^s \sum_{j=r+1}^i \|\partial_{\omega_i} \partial_{\omega_j} g\|_{L^\infty(\Omega)}.$$

*Proof.* Let  $\omega \in \mathbb{R}^s$ . There holds

$$\begin{aligned} g(\omega) &= \underbrace{g(\omega_1, \dots, \omega_r, 0, \dots)}_{=0} + \sum_{i=r+1}^s \int_0^{\omega_i} \partial_{\omega_i} g(\omega_1, \dots, \omega_{i-1}, t_i, 0, \dots) dt_i \\ &= \sum_{i=r+1}^s \int_0^{\omega_i} \left( \partial_{\omega_i} g(\omega_1, \dots, \omega_r, 0, \dots) \right. \\ &\quad \left. + \int_0^{t_i} \partial_{\omega_i}^2 g(\omega_1, \dots, \omega_{i-1}, s_i, 0, \dots) ds_i \right. \\ &\quad \left. + \sum_{j=r+1}^{i-1} \int_0^{\omega_j} \partial_{\omega_j} \partial_{\omega_i} g(\omega_1, \dots, \omega_{j-1}, s_j, 0, \dots) ds_j \right) dt_i. \end{aligned}$$

Since the first integrand on the right-hand side does not depend on  $\omega_i$ , the above

implies

$$g(\omega) = \sum_{i=r+1}^s \left( \omega_i \partial_{\omega_i} g(\omega_1, \dots, \omega_r, 0, \dots) \right. \\ \left. + \int_0^{\omega_i} \left( \int_0^{t_i} \partial_{\omega_i}^2 g(\omega_1, \dots, \omega_{i-1}, s_i, 0, \dots) ds_i \right. \right. \\ \left. \left. + \sum_{j=r+1}^{i-1} \int_0^{\omega_j} \partial_{\omega_j} \partial_{\omega_i} g(\omega_1, \dots, \omega_{j-1}, s_j, 0, \dots) ds_j \right) dt_i \right).$$

Since there holds  $(\omega \mapsto \omega_i \partial_{\omega_i} g(\omega_1, \dots, \omega_r, 0, \dots)) \in \mathcal{P}_r^s(\Omega)$  for all  $i \geq r+1$ , we conclude the proof.  $\square$

**THEOREM 4.7.** *Under the assumptions of the current section, there holds*

$$(4.5) \quad \|G(D_\ell^\nu)\|_{L^\infty(\Omega)} \leq C_{\text{KL}} h_\ell^2 \sum_{i=s_{\nu-1}+1}^{s_\nu} \|\phi_i\|_{W^{1,\infty}(D)} \|f\|_{L^2(D)} \|g\|_{L^2(D)}.$$

Moreover, there exists  $g_0 \in \mathcal{P}_{s_{\nu-1}}^{s_\nu}(\Omega)$  such that

$$(4.6) \quad \|G(D_\ell^\nu) - g_0\|_{L^\infty(\Omega)} \\ \leq C_{\text{KL}} h_\ell^2 \sum_{i=s_{\nu-1}+1}^{s_\nu} \sum_{j=s_{\nu-1}+1}^{s_\nu} \|\phi_i\|_{W^{1,\infty}(D)} \|\phi_j\|_{W^{1,\infty}(D)} \|f\|_{L^2(D)} \|g\|_{L^2(D)}.$$

The constant  $C_{\text{KL}} > 0$  is independent of  $\ell$ ,  $\nu$ , and  $\omega$ .

*Proof.* The first estimate (4.5) follows from the definition of  $A^\nu$  and Theorem 3.9. For (4.6), the function  $g(\omega) := D_\ell^\nu(\omega)$  satisfies the requirements of Lemma 4.6 with  $r = s_{\nu-1}$ . Hence, the result follows immediately from Lemma 4.6 and Lemma 4.5.  $\square$

**5. Monte Carlo integration.** This section discusses the Monte Carlo quadrature rules. While for the  $H$ -matrix case, the standard rule will be used, the  $KL$ -expansion case allows us to increase the order of convergence by symmetrization of the Monte Carlo rule. This section defines the Monte Carlo integration for the case that the random coefficient is given by a  $KL$ -expansion as discussed in Sections 4.1–4.2.

We make the standard assumption that the functions  $\phi_i$  from (4.3) satisfy

$$(5.1) \quad \|\phi_j\|_{W^{1,\infty}(D)} \leq C_{\text{KL}} j^{-r} \quad \text{for all } j \in \mathbb{N}$$

for some  $r > 1$ .

**LEMMA 5.1.** *Define the Monte Carlo rule*

$$Q_M(g) := \frac{1}{M} \sum_{i=1}^M g(X^i)$$

for uniformly distributed i.i.d  $X^i \in [-1/2, 1/2]^{s_\nu}$ . Then, under the assumptions of Section 4.1 given  $\ell, \nu \in \mathbb{N}$ , the function  $F: \Omega \rightarrow \mathbb{R}$ ,  $\omega \mapsto G(D_\ell^\nu(\omega))$  satisfies

$$\sqrt{\mathbb{E}_{\text{MC}} |\mathbb{E}(F) - Q_M(F)|^2} \leq C_{\text{MC}} s_{\nu-1}^{1-r} \frac{h_\ell^2}{\sqrt{M}} \|f\|_{L^2(D)} \|g\|_{L^2(D)}.$$

Here,  $\mathbb{E}_{\text{MC}}(\cdot)$  denotes integration over the combined probability spaces of the  $X^i$ ,  $i = 1, \dots, M$ , whereas  $\mathbb{E}(\cdot)$  denotes integration over  $\Omega_\nu$ .



*Proof.* The statement follows immediately from the standard Monte Carlo error estimate, Theorem 4.1, and the fact that  $\sum_{j=s_{\nu-1}+1}^{s_\nu} j^{-r} \lesssim s_{\nu-1}^{1-r}$ .  $\square$

By symmetrization of the Monte Carlo sequence, we are able to increase the order of convergence in the truncation parameter  $\nu$ .

LEMMA 5.2. *Define the symmetric Monte Carlo rule*

$$Q_M(g) := \frac{1}{2M} \sum_{i=1}^M (g(X_1^i, \dots, X_{s_\nu}^i) + g(X_1^i, \dots, X_{s_{\nu-1}}^i, -X_{s_{\nu-1}+1}^i, \dots, -X_{s_\nu}^i)),$$

where the  $X^i \in [-1/2, 1/2]^{s_\nu}$  are i.i.d. and uniformly distributed. Under the assumptions of Section 4.2, there holds  $Q_M(g_0) = 0$  for all  $g_0 \in \mathcal{P}_{s_{\nu-1}}^{s_\nu}(\Omega)$ . Moreover, given  $\ell, \nu \in \mathbb{N}$ , the function  $F: \Omega \rightarrow \mathbb{R}$ ,  $\omega \mapsto G(D_\ell^\nu(\omega))$  satisfies

$$\sqrt{\mathbb{E}_{\text{MC}} |\mathbb{E}(F) - Q_M(F)|^2} \leq C_{\text{MC}} s_{\nu-1}^{2(1-r)} \frac{h_\ell^2}{\sqrt{M}} \|f\|_{L^2(D)} \|g\|_{L^2(D)}.$$

Here,  $\mathbb{E}_{\text{MC}}(\cdot)$  denotes integration over the combined probability spaces of the  $X_i$ ,  $i = 1, \dots, 2^m$ , whereas  $\mathbb{E}(\cdot)$  denotes integration over  $\Omega_\nu$ .

*Proof.* First, we notice that for  $g_0 \in \mathcal{P}_{s_{\nu-1}}^1(\Omega)$ , there holds

$$g_0(X_1^i, \dots, X_{s_\nu}^i) = -g_0(X_1^i, \dots, X_{s_{\nu-1}}^i, -X_{s_{\nu-1}+1}^i, \dots, -X_{s_\nu}^i).$$

Therefore, we have  $Q_M(g_0) = 0$  for all  $g_0 \in \mathcal{P}_{s_{\nu-1}}^1(\Omega)$ . Thus, the statement follows from the standard Monte Carlo error estimate and Theorem 4.7, where we note with (5.1)

$$\begin{aligned} & \sum_{i=s_{\nu-1}+1}^{s_\nu} \sum_{j=s_{\nu-1}+1}^{s_\nu} \|\phi_i\|_{W^{1,\infty}(D)} \|\phi_j\|_{W^{1,\infty}(D)} \\ & \lesssim \sum_{i=s_{\nu-1}+1}^{\infty} \sum_{j=s_{\nu-1}+1}^{\infty} i^{-r} j^{-r} \lesssim (s_{\nu-1})^{2(-r+1)}. \end{aligned} \quad \square$$

**6. Multi-Index error control.** The multi-index decomposition allows us to exploit the product error estimates and hence to improve the complexity of the finite-element/Monte Carlo algorithm. To that end, we rewrite the exact solution as ( $Q_m$  denotes one of the quadrature rules  $Q_M$  from Section 5 with  $M = 2^m$ )

$$\begin{aligned} \mathbb{E}(G(u)) &= \sum_{j=0}^{\infty} (Q_{m_j} - Q_{m_{j-1}})(G(u)) \\ &= \sum_{j=0}^{\infty} \sum_{\ell=0}^{\infty} (Q_{m_j} - Q_{m_{j-1}})(G(u_\ell - u_{\ell-1})) \\ &= \sum_{j=0}^{\infty} \sum_{\ell=0}^{\infty} \sum_{\nu=0}^{\infty} (Q_{m_j} - Q_{m_{j-1}})(G(D_\ell^\nu)), \end{aligned}$$

where  $m_j \in \mathbb{N}$  and  $Q_{m_{-1}} := 0$ . By truncation of the series, we achieve a sparse approximation, i.e., given  $N \in \mathbb{N}$

$$\mathbb{E}(G(u)) \approx G_N := \sum_{0 \leq j+\ell+\nu \leq N} (Q_{m_j} - Q_{m_{j-1}})G(D_\ell^\nu) = \sum_{0 \leq \ell+\nu \leq N} Q_{m_{N-\ell-\nu}}(G(D_\ell^\nu)).$$

We define two quantities to calculate the efficiency of the method: The error is defined by

$$E_N := \sqrt{\mathbb{E}_{\text{mc}}|\mathbb{E}(G(u)) - G_N|^2}$$

whereas the cost model is defined by

$$C_N := (\text{The number of computational operations necessary to compute } G_N)$$

and obviously depends on the chosen method discussed below.

First, we establish the cost model. A standard FEM will ensure  $h_\ell \simeq 2^{-\ell}$  which implies  $\#\mathcal{T}_\ell \simeq 2^{d\ell}$ . We assume a linear iterative solver such that solving the sparse FEM system costs  $\mathcal{O}(2^{d\ell})$ .

Under the assumptions of Section 4.1 and 4.2, we assume that we can compute the bilinear forms

$$a_j(v, w) := \int_D \phi_j(x) \nabla v(x) \nabla w(x) dx \quad \text{for all } v, w \in \mathcal{X}_\ell$$

exactly in  $\mathcal{O}(\#\mathcal{T}_\ell)$ . Depending on the truncation parameters  $s_\nu$ , we have to compute  $s_\nu$  bilinear forms  $a_j(\cdot, \cdot)$  to obtain in the affine case

$$a_\omega^\nu(v, w) = \sum_{j=1}^{s_\nu} \omega_j a_j(v, w),$$

resulting in a cost of  $\mathcal{O}(2^{d\ell} s_\nu)$ . Altogether, this yields

$$C_N \simeq \sum_{0 \leq j+\ell+\nu \leq N} 2^{m_j} 2^{d\ell} s_\nu$$

Using Lemma 5.1, we see that the multi-index error satisfies

$$\begin{aligned} E_N &= \mathbb{E}_{\text{MC}} \left( \left| \sum_{N < j+\ell+\nu} (Q_{m_j} - Q_{m_{j-1}}) G(D_\ell^\nu) \right|^2 \right)^{1/2} \\ &\leq \sum_{0 \leq \ell+\nu} \mathbb{E}_{\text{MC}} \left( |(\mathbb{E} - Q_{m_{\max\{0, N-\ell-\nu+1\}}}) G(D_\ell^\nu)|^2 \right)^{1/2} \\ &\lesssim \|f\|_{L^2(D)} \|g\|_{L^2(D)} \sum_{0 \leq \ell+\nu} 2^{-m_{\max\{0, N-\ell-\nu+1\}}/2} 2^{-2\ell} s_{\nu-1}^{1-r}. \end{aligned}$$

An obvious choice of the parameters  $s_\nu$  and  $m_j$  is to balance the work spent on each of the two tasks such that the three error contributions (FEM-discretization error, truncation error, quadrature error) are of equal asymptotic order. Following this idea, we define

$$m_j := \lceil 4j \rceil \quad \text{and} \quad s_\nu := \lceil 2^{\frac{2\nu}{r-1}} \rceil.$$

With this, we have

$$\begin{aligned} (6.1) \quad E_N &\lesssim \|f\|_{L^2(D)} \|g\|_{L^2(D)} \sum_{0 \leq \ell+\nu} 2^{-2 \max\{0, N-\ell-\nu+1\}} 2^{-2\ell} 2^{-2\nu} \\ &\lesssim \|f\|_{L^2(D)} \|g\|_{L^2(D)} (N+1)^2 2^{-2N} \end{aligned}$$

as well as

$$(6.2) \quad C_N \simeq \sum_{0 \leq j+\ell+\nu \leq N} 2^{4j} 2^{d\ell} 2^{\frac{2\nu}{r-1}} \lesssim 2^{\max\{4, d, \frac{2}{r-1}\}N}.$$

Using the symmetrized Monte Carlo rule from Lemma 5.2, we see that the multi-index error improves to

$$E_N \lesssim \|f\|_{L^2(D)} \|g\|_{L^2(D)} \sum_{0 \leq \ell+\nu} 2^{-m_{\max\{0, N-\ell-\nu+1\}/2}} 2^{-2\ell} s_{\nu-1}^{2(1-r)}.$$

As above, we balance the contributions by

$$m_j := \lceil 4j \rceil \quad \text{and} \quad s_\nu := \lceil 2^{\frac{\nu}{r-1}} \rceil.$$

With this, we obtain the same error estimate as for the plain Monte Carlo rule (6.1), but with an improved cost estimate of

$$(6.3) \quad C_N^{\text{symm}} \lesssim 2^{\max\{4, d, \frac{1}{r-1}\}N}.$$

**7. Numerical experiment.** We provide numerical tests in 2D using a similar example as in [1, Section 5.2]. We choose  $D = [0, 1]^2$  and define the coefficient  $A$  by

$$\begin{aligned} A(x, \omega) &:= 1/2 + \sum_{k_1, k_2=1}^{\infty} \frac{\omega_{k_1, k_2}}{(k_1^2 + k_2^2)^2} \sin(k_1 \pi x_1) \sin(k_2 \pi x_2) \\ &:= 1/2 + \sum_{j=1}^{\infty} \frac{\omega_j}{\mu_j} \sin(k_{1,j} \pi x_1) \sin(k_{2,j} \pi x_2), \end{aligned}$$

where  $\mu_j := (k_{1,j}^2 + k_{2,j}^2)^2$  such that  $\mu_i \leq \mu_j$  for all  $i \leq j$  and ties are broken in an arbitrary fashion. This ensures that the  $\phi_j$  satisfy (5.1) with  $r = 2$ . The problem then reads

$$a(A, \omega; u, v) = f(v),$$

where  $f \in H^{-1/2-\varepsilon}(D)$  for all  $\varepsilon > 0$  is defined by

$$f(v) := \int_{\Gamma} v(x_1, x_2) x_1 \, d\Gamma(x_1, x_2) = \sqrt{2} \int_0^1 t v(t, 1-t) \, dt$$

for  $\Gamma = \{(0, 1) + r(1, -1) : 0 \leq r \leq 1\}$  being a diagonal of  $D$ . Note that we choose the weight  $x_1$  in the integral in the definition of the right-hand side to introduce some non-symmetric quantities and thus avoid super-convergence effects. We consider the quantity of interest  $G(u) := \int_{D'} u \, dx$ , where  $D' = (1/2, 1)^2 \subset D$ . Whereas the analysis of the present paper is focused on the full regularity case with right-hand side  $f \in L^2(D)$ , we would like to remark that all the arguments remain valid in case of reduced regularity (with corresponding reduced rates of convergence).

The discretization is done via linear finite elements  $\mathcal{X}_\ell$  on a uniform triangulation of  $[0, 1]^2$  into  $2^{2\ell}$ -elements (one example is shown in Figure 1). Note that the cost model applies as we can compute the stiffness matrix exactly since the gradients of the shape functions are constants and the anti-derivatives of products of sine functions are known over triangles. Hence we have  $h_\ell = 2^{-\ell}$ . The expected error for the FEM on

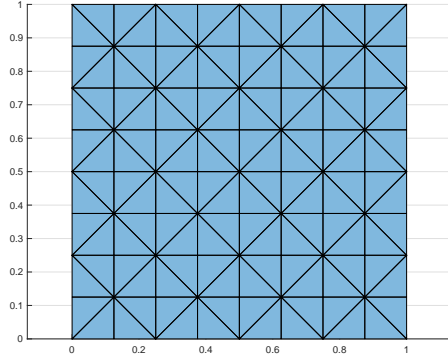


FIG. 1. One level of mesh-refinement for the 2D example.

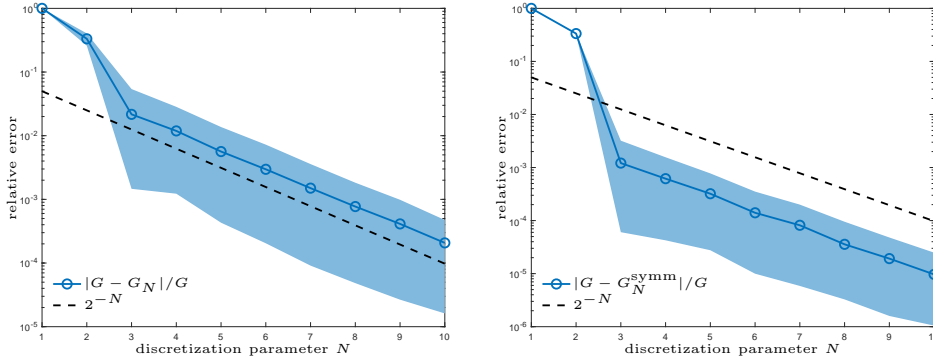


FIG. 2. Relative errors of the multi-index algorithms with respect to the reference solution  $G$  compared with the theoretical error bound  $2^{-N}$  (original algorithm (left) and symmetrized version (right)). Both plots show the average error curve of 500 runs of the algorithms. The shaded regions contain 95% of the computed error curves.

level  $\ell$  is  $\mathcal{O}(h_\ell) = \mathcal{O}(2^{-\ell})$ . Thus we choose the  $m_j := 2j$  as well as  $s_\nu = \lceil 2^{\nu/(r-1)} \rceil$  for the original algorithm and  $s_\nu = \lceil 2^{\nu/(2(r-1))} \rceil$  for the symmetrized version. Therefore we expect that the errors for both algorithms satisfy  $E_N = \mathcal{O}(2^{-N}) = \mathcal{O}(h_N)$ . This is confirmed in Figure 2. For the numerical experiments, we compare with a reference solution computed with a higher-order quasi-Monte Carlo method as proposed in [1]. The authors thank Dr. Hernandez for computing the reference value  $G = 0.011079\dots$ . To smooth out the effects of MC sampling, the plotted relative errors are averaged over 500 runs of the respective multi-index algorithm (we also plot a shaded region which contains 95% of all error curves).

**8. Conclusion.** The present work shows that the multi-index Monte Carlo algorithm with the indices being the discretization parameters of the finite element method, of the Monte Carlo method, and of the approximation of the random field is superior to its multi-level counterpart. The error estimates are rigorous and the product error estimate from Theorem 3.9 might be of independent interest. The method can be combined with existing multi-index techniques which focus on sparse grids in the domain to further reduce the computational effort.

## REFERENCES

- [1] J. DICK, R. N. GANTNER, Q. T. LE GIA, AND C. SCHWAB, *Multilevel higher-order quasi-Monte Carlo Bayesian estimation*, Math. Models Methods Appl. Sci., 27 (2017), pp. 953–995, <https://doi.org/10.1142/S021820251750021X>.
- [2] J. DICK, F. Y. KUO, Q. T. LE GIA, AND C. SCHWAB, *Multilevel higher order QMC Petrov-Galerkin discretization for affine parametric operator equations*, SIAM J. Numer. Anal., 54 (2016), pp. 2541–2568, <https://doi.org/10.1137/16M1078690>.
- [3] M. B. GILES, *Multilevel Monte Carlo path simulation*, Oper. Res., 56 (2008), pp. 607–617, <https://doi.org/10.1287/opre.1070.0496>.
- [4] A.-L. HAJI-ALI, F. NOBILE, AND R. TEMPONE, *Multi-index Monte Carlo: when sparsity meets sampling*, Numer. Math., 132 (2016), pp. 767–806, <https://doi.org/10.1007/s00211-015-0734-5>.
- [5] A.-L. HAJI-ALI AND R. TEMPONE, *Multilevel and Multi-index Monte Carlo methods for the McKean-Vlasov equation*, Stat. Comput., 28 (2018), pp. 923–935, <https://doi.org/10.1007/s11222-017-9771-5>, <https://doi.org/10.1007/s11222-017-9771-5>.
- [6] S. V. PIETERJAN ROBBE, DIRK NUYENS, *A multi-index quasi-monte carlo algorithm for lognormal diffusion problems*, arXiv:1608.03157, (2017).
- [7] C. REISINGER AND Z. WANG, *Analysis of multi-index Monte Carlo estimators for a Zakai SPDE*, J. Comput. Math., 36 (2018), pp. 202–236.