ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Seminar for Applied Mathematics SAM

Adaptive Quasi-Monte Carlo Finite Element Methods for parametric elliptic PDEs

M. Longo

Research Report No. 2021-03

January 2021
Latest revision: August 2021

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

Adaptive Quasi-Monte Carlo Finite Element Methods for parametric elliptic PDEs

Marcello Longo*

August 28, 2021

Abstract

We introduce novel adaptive methods to approximate moments of solutions of Partial Differential Equations (PDEs) with uncertain parametric inputs. A typical problem in Uncertainty Quantification is the approximation of the expected values of Quantities of Interest of the solution, which requires the efficient numerical approximation of high-dimensional integrals. We perform this task by a class of deterministic Quasi-Monte Carlo integration rules derived from Polynomial lattices, that allows to control a-posteriori the integration error without querying the governing PDE and does not incur in the curse of dimensionality. Based on an abstract formulation of Adaptive Finite Element methods for deterministic problems, we infer convergence of the combined adaptive algorithms in the parameter and physical space. We propose a selection of examples of PDEs admissible for these algorithms. Finally, we present numerical evidence of convergence for a model diffusion PDE.

1 Introduction

The study of problems governed by parametric PDEs has seen a steady development in recent years with an eye on applications to computational sciences and engineering. The general common methodology is to treat the parametric equation as a family of equations with given data and to query a possibly large number of them, by well-known solvers.

For deterministic PDEs, quasi-optimality of Adaptive Finite Element Methods (AFEM) has been extensively studied, we refer to [6] for classical results on elliptic diffusion PDEs and to [5, 17, 21] and the references therein for more recent developments towards an abstract analysis. When including uncertainty in the underlying PDE, in order to maintain the computational cost to a minimum, it is crucial to additionally estimate the error for the parametric solution and in particular to determine adaptively a finite sampling set P contained in U and a suitable Finite Element space of the PDE, such that a given error tolerance is met a-posteriori. Existing approaches involve adaptive stochastic Galerkin, studied in [15, 16, 2] and more recently adaptive collocation methods on sparse grids [14, 18]. We extend these results to sampling based on Quasi-Monte Carlo (QMC) rules, while leveraging the aforementioned abstract AFEM framework of the *Axioms of adaptivity*.

In the spirit of Uncertainty Quantification (UQ), we aim at the approximation of low-order moments of a goal functional $G \in V^*$ (also called quantity of interest or observable) of the solution $u\colon U \to V$ of a parametric PDE given by $\mathcal{P} : V \times U \to W^*$

$$\mathcal{P}(u(\boldsymbol{y}), \boldsymbol{y}) = 0 \quad \text{in } W^*, \tag{1}$$

*Seminar for Applied Mathematics (SAM), D-MATH
ETH Zürich, Rämistrasse 101, Zürich-8092, Switzerland

for some reflexive Banach spaces V, W of functions defined on a *physical* domain $D \subset \mathbb{R}^d$ and a parameter space $U \subseteq \mathbb{R}^s$, $d, s \in \mathbb{N}$.

We are particularly interested in the case of a large number s of parameters, all independent uniformly distributed on the interval $[-\frac{1}{2}, \frac{1}{2}]$. Moments of u are then expressed as high-dimensional integrals. In such setting, deterministic QMC integration is proven to outperform standard Monte-Carlo: suitable assumptions on the regularity of the parameter to solution map $u : U \to V$ are known to grant dimension independent and higher order decay of the quadrature error, for deterministic QMC rules derived from Polynomial lattices, comprising Interlaced Polynomial lattices [12, 24, 23] and Extrapolated Polynomial lattices [11, 10].

Moreover, Polynomial lattices allow for an easily computable a-posteriori error estimator, that is known to be asymptotically exact and free of the curse of dimensionality [10]. We remark that other a-posteriori estimation techniques were developed for Sobol' points and Rank-1 Lattices in [27, 28].

The scope of this work is to introduce a family of adaptive algorithms, to approximate solutions of *many-query* problems, based on deterministic QMC sampling on the parameter box $U := [-\frac{1}{2}, \frac{1}{2}]^s$. Our contribution is to provide convergence results of these algorithms, without incurring in the curse of dimensionality, in a generic framework comprising several common PDE problems, where the parametric error estimator is independent of the underlying PDE. We employ parametric error estimators that only depend on the computed discrete solution $u_{\mathcal{T}} \colon U \to V_{\mathcal{T}}$ (where $V_{\mathcal{T}}$ is a finite dimensional space), while its computation is independent of a) the specific discretization space $V_{\mathcal{T}}$, b) the equation (1) satisfied by u and c) the PDE solver used. Moreover, we pay particular attention to modularity and non-intrusivity of the algorithm, i.e. we break the overall computation into smaller parts, each with its requirements, in order to be able to reuse existing implementations.

The structure of the paper is as follows: in Section 2 we introduce the problem and we summarize the notation and results from Quasi-Monte Carlo integration with Polynomial Lattice rules and convergence of Adaptive FEM. Section 3 is devoted to the description and proof of convergence of 3 different adaptive procedures; for each of them, we show that it is possible to leverage Goal oriented adaptivity, as in [17, 3]. Additionally, we indicate a few examples of problems that can be solved with our method. In Section 4 we present numerical experiments for a model PDE with random diffusion.

2 Preliminaries

In this section we formulate the problem and illustrate our working assumptions. For $s \in \mathbb{N}$, define the Lebesgue measure μ on $U := [-\frac{1}{2}, \frac{1}{2}]^s$, which is a probability measure. Let V, W be reflexive Banach spaces of functions defined on a Lipschitz domain $D \subset \mathbb{R}^d$, $d \in \{2, 3\}$. For $\boldsymbol{y} \in U$, let $\mathfrak{a}_{\boldsymbol{y}} \colon V \times W \to \mathbb{R}$ be a bilinear form and $\mathfrak{l}_{\boldsymbol{y}} \in W^*$, W^* denoting the topological dual of W. Moreover, let $G \in V^*$ be the sought Quantity of Interest.

We consider problems of the form: find

$$I(G(u)) = \int_U G(u(\boldsymbol{y}))\, \mathrm{d}\mu(\boldsymbol{y}), \tag{2}$$

where, for all $\boldsymbol{y} \in U$, $u(\boldsymbol{y}) \in V$ solves the variational problem

$$\mathfrak{a}_{\boldsymbol{y}}(u(\boldsymbol{y}), w) = \mathfrak{l}_{\boldsymbol{y}}(w) \quad \forall w \in W. \tag{3}$$

In order to ensure that the problem is well-posed we shall impose the following, [4].

Assumption 2.1. *The data $\mathfrak{a}_{\boldsymbol{y}}, \mathfrak{l}_{\boldsymbol{y}}$ satisfy uniform, with respect to $\boldsymbol{y} \in U$, inf-sup conditions*

$$\begin{aligned} \inf_{0\neq v\in V} \sup_{0\neq w\in W} \frac{\mathfrak{a}_{\boldsymbol{y}}(v,w)}{\|v\|_V \|w\|_W} \geq \lambda > 0 \\ \inf_{0\neq w\in W} \sup_{0\neq v\in V} \frac{\mathfrak{a}_{\boldsymbol{y}}(v,w)}{\|v\|_V \|w\|_W} \geq \lambda > 0 \end{aligned} \tag{4}$$

and continuity

$$\mathfrak{a}_{\boldsymbol{y}}(v,w) \leq \Lambda \|v\|_V \|w\|_W, \qquad \forall v \in V, w \in W, \tag{5}$$

for some $0 < \lambda < \Lambda < \infty$ independent of $\boldsymbol{y}$. Moreover, we assume, for some $0 < C_{\mathfrak{l}} < \infty$

$$\sup_{\boldsymbol{y}\in U} \|\mathfrak{l}_{\boldsymbol{y}}\|_{W^*} \leq C_{\mathfrak{l}}. \tag{6}$$

Under Assumption 2.1, we have the following a-priori estimate

$$\sup_{\boldsymbol{y}\in U} \|u(\boldsymbol{y})\|_V \leq \frac{C_{\mathfrak{l}}}{\lambda}. \tag{7}$$

Then, given a small tolerance $\varepsilon > 0$, we want to compute a $Q \in \mathbb{R}$ such that

$$|I(G(u)) - Q| \approx \varepsilon.$$

It is clear that we have multiple sources of error to take into consideration. First, we have the quadrature error in approximating the expectation by sampling with Quasi-Monte Carlo rules. Second, we include the discretization error as the solution $u(\boldsymbol{y})$ comes from a PDE problem and we cannot expect in general to recover it exactly.

Additionally, one could consider dimension truncation error, that arises in the treatment of countably parametric problems by means of a quadrature rule over a finite dimensional set U [22]. We exclude this error from the analysis and we assume that the dimension $s \in \mathbb{N}$ is known a-priori throughout the rest of the discussion.

2.1 Quasi-Monte Carlo a posteriori error estimation

In order to determine a stopping criterion for the QMC-AFEM algorithm below, we make use of the asymptotically exact a-posteriori estimator from [10, Section 4]. Here, we recall briefly the relevant notation and results.

Fix $b \geq 2$ prime and denote by $\mathbb{Z}_b$ the field with b elements; let $\mathbb{Z}_b[x], \mathbb{Z}_b((x))$ be respectively polynomials and formal Laurent series with coefficients in $\mathbb{Z}_b$. Given $m \in \mathbb{N}$, let

$$v_m \colon \mathbb{Z}_b((x)) \to [0,1) \quad \text{with} \quad v_m\left(\sum_{i=-\infty}^{\infty} \xi_i x^{-i}\right) := \sum_{i=1}^{m} \xi_i b^{-i}. \tag{8}$$

A *Polynomial Lattice* on U is defined as $P_m(\boldsymbol{q}, p) = \{\boldsymbol{y}_0, \dots, \boldsymbol{y}_{b^m-1}\} \subset U$ where $p \in \mathbb{Z}_b[x]$ is an irreducible polynomial of degree m, $\boldsymbol{q} = (q_1, \dots, q_s) \in (\mathbb{Z}_b[x])^s$ is componentwise of degree at most $m-1$ and

$$\boldsymbol{y}_n := \left(v_m\left(\frac{q_1(x)n(x)}{p(x)}\right) - \frac{1}{2}, \dots, v_m\left(\frac{q_s(x)n(x)}{p(x)}\right) - \frac{1}{2}\right). \tag{9}$$

The cardinality of $P_m(\boldsymbol{q},p)$ is b^m. Given a space $\mathcal{W}$ of functions $F\colon U \to \mathbb{R}$ for which point values are well-defined, a Quasi-Monte Carlo rule is a quadrature rule with equal weights and deterministic nodes. We denote by $Q_{b^m,\boldsymbol{q},p}\colon \mathcal{W} \to \mathbb{R}$ the QMC rule that uses $P_m(\boldsymbol{q},p)$ as nodes

$$Q_{b^m,\boldsymbol{q},p}(F) := \frac{1}{b^m} \sum_{\boldsymbol{y}\in P_m(\boldsymbol{q},p)} F(\boldsymbol{y}) \tag{10}$$

and we call it *Polynomial Lattice Rule*.

Next, we introduce the weighted unanchored Sobolev space $\mathcal{W}_{s,\alpha,\boldsymbol{\gamma}}(U,V)$. For $\boldsymbol{\nu} = (\nu_j)_{j\in\mathbb{N}} \in \mathbb{N}_0^{\mathbb{N}}$ with finite support $\operatorname{supp}(\boldsymbol{\nu}) := |\{j : \nu_j > 0\}| < \infty$, we will use the notation $\boldsymbol{\nu}! := \prod_{j\in\operatorname{supp}(\boldsymbol{\nu})} \nu_j!$, $|\boldsymbol{\nu}| := \sum_j \nu_j$, and the partial derivatives $\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} := \partial_{y_1}^{\nu_1}\partial_{y_2}^{\nu_2}\cdots$. We also write, for a real valued sequence $\boldsymbol{\beta}$, $\boldsymbol{\beta}^{\boldsymbol{\nu}} := \prod_{j\in\operatorname{supp}(\boldsymbol{\nu})} \beta_j^{\nu_j}$.

For $s, \alpha \in \mathbb{N}$, $\alpha \geq 2$ and $\boldsymbol{\gamma} = (\gamma_{\mathfrak{u}})_{\mathfrak{u}\subset\mathbb{N},|\mathfrak{u}|<\infty}$ a set of positive weights, we define the space $\mathcal{W}_{s,\alpha,\boldsymbol{\gamma}}(U,V) \subseteq C^\infty(U,V)$ of functions F with finite norm

$$\|F\|_{s,\alpha,\boldsymbol{\gamma},V} := \max_{\mathfrak{u}\subseteq\{1:s\}} \frac{1}{\gamma_{\mathfrak{u}}} \sum_{\mathfrak{v}\subseteq\mathfrak{u}} \sum_{\boldsymbol{\nu}_{\mathfrak{u}\setminus\mathfrak{v}}\in\{1:\alpha\}^{|\mathfrak{u}\setminus\mathfrak{v}|}} \int_{[-\frac{1}{2},\frac{1}{2}]^{|\mathfrak{v}|}} \left\| \int_{[-\frac{1}{2},\frac{1}{2}]^{s-|\mathfrak{v}|}} \partial_{\boldsymbol{y}}^{(\alpha_{\mathfrak{v}},\boldsymbol{\nu}_{\mathfrak{u}\setminus\mathfrak{v}})} F(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}_{-\mathfrak{v}} \right\|_V \mathrm{d}\boldsymbol{y}_{\mathfrak{v}}, \tag{11}$$

where the j-th component of $(\alpha_{\mathfrak{v}}, \boldsymbol{\nu}_{\mathfrak{u}\setminus\mathfrak{v}})$ is α for $j \in \mathfrak{v}$, equals ν_j if $j \in \mathfrak{u}\setminus\mathfrak{v}$ and vanishes otherwise. Note that the inner integral is understood in the Bochner sense. For simplicity we write $\mathcal{W}_{s,\alpha,\boldsymbol{\gamma}} := \mathcal{W}_{s,\alpha,\boldsymbol{\gamma}}(U,\mathbb{R})$ and $\|F\|_{s,\alpha,\boldsymbol{\gamma}} := \|F\|_{s,\alpha,\boldsymbol{\gamma},\mathbb{R}}$. In the context of higher-order QMC error analysis, the weights are usually defined in terms of a positive sequence $\boldsymbol{\beta} = (\beta_j)_{j\in\mathbb{N}} \in \ell^p(\mathbb{N})$, with $p \in (0,1)$ and $\kappa \geq 0, c > 0$, by

$$\gamma_{\mathfrak{u}} = \prod_{j\in\mathfrak{u}} c\beta_j, \qquad \text{or} \tag{12}$$

$$\gamma_{\mathfrak{u}} = \sum_{\boldsymbol{\nu}\in\{1:\alpha\}^{|\mathfrak{u}|}} (|\boldsymbol{\nu}|!)^{1+\kappa} \prod_{j\in\mathfrak{u}} c\beta_j^{\nu_j}. \tag{13}$$

We refer to the former as *product* weights and to the latter as *smoothness driven, product and order dependent* (or SPOD) weights. Weights are given as input to a Component-By-Component (CBC) construction of the generating vector $\boldsymbol{q} = (q_1,\ldots,q_s)$ of a Polynomial Lattice rule. In particular, we operate inductively (over the parametric dimension $\bar{s} = 1,\ldots,s$) a minimization of (a computable bound for) the worst case quadrature error $B_m(\boldsymbol{q},p)$ as follows: we initialize $q_1 = 1$, $p \in \mathbb{Z}_b[x]$ irreducible of degree m and we compute for $\bar{s} = 2,\ldots,s$

$$q_{\bar{s}} = \operatorname*{argmin}_{q\in\mathbb{Z}_b[x]\setminus\{0\}:\deg(q)<m} B_m((q_1,\ldots,q_{\bar{s}-1},q),p). \tag{14}$$

Furthermore, the construction can be accelerated employing Fast Fourier transform. We refer to [11] for the description of the (Fast) CBC algorithm for product weights, and to [10] to the case of weights of SPOD type used here.

In what follows, we omit the explicit dependence of $P_m(\boldsymbol{q},p), Q_{b^m,\boldsymbol{q},p}$ on $\boldsymbol{q},p$ and write P_m, Q_{b^m}, respectively, to denote polynomial lattices constructed by a CBC algorithm for given weights. The fundamental result from QMC theory that we will employ is the following Proposition from [11, 10].

Proposition 2.1. *Let $s \in \mathbb{N}, \alpha \geq 2$. Let $\boldsymbol{\gamma}$ be a set of product weights such that $\boldsymbol{\beta} \in \ell^p(\mathbb{N})$ for all $p > \frac{1}{2}$, or a set of SPOD weights with $\boldsymbol{\beta} \in \ell^p(\mathbb{N})$ for some $0 < p < \frac{1}{2+\kappa}$. Then, for all $F \in \mathcal{W}_{s,\alpha,\boldsymbol{\gamma}}$ a sequence of Polynomial Lattice rules $(Q_{b^m})_{m\in\mathbb{N}}$ can be constructed with a CBC algorithm such that*

$$|I(F) - Q_{b^m}(F)| \leq C\,\|F\|_{s,\alpha,\boldsymbol{\gamma}}\, b^{-m} \tag{15}$$

for a constant C independent of m, s, F and $\sup_{s<\infty} \|F\|_{s,\alpha,\boldsymbol{\gamma}} < \infty$. Moreover,

$$I(F) - Q_{b^m}(F) = \frac{Q_{b^m}(F) - Q_{b^{m-1}}(F)}{b-1} + \mathcal{O}\big(b^{-2m+\delta}\big) \qquad \text{as } m \to \infty \tag{16}$$

for all $\delta > 0$, with constant in $\mathcal{O}(\cdot)$ independent of $s \in \mathbb{N}$ but dependent on δ.

Proof. (15) follows from [11, Equation 3.1] and (16) is [10, Theorem 4.1]. □

We denote the QMC a-posteriori estimator by

$$E_{b^m}(F) := \frac{Q_{b^m}(F) - Q_{b^{m-1}}(F)}{b-1}. \tag{17}$$

For $u \in \mathcal{W}_{s,\alpha,\boldsymbol{\gamma}}(U,V)$, $G \in V^*$ it holds $G(u) \in \mathcal{W}_{s,\alpha,\boldsymbol{\gamma}}$ with $\|G(u)\|_{s,\alpha,\boldsymbol{\gamma}} \le \|G\|_{V^*} \|u\|_{s,\alpha,\boldsymbol{\gamma},V}$. As a consequence, a typical strategy to verify the viability of QMC integration consists in determining derivative bounds of the form

$$\sup_{\boldsymbol{y}\in U} \big\|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\boldsymbol{y})\big\|_V \le C(|\boldsymbol{\nu}|!)^{1+\kappa} \boldsymbol{\beta}^{\boldsymbol{\nu}} \quad \forall \boldsymbol{\nu} \in \mathbb{N}_0^{\mathbb{N}}, |\boldsymbol{\nu}| < \infty \tag{18}$$

for some $\kappa \ge 0, C > 0$ independent of $\boldsymbol{\nu}, s$ and $\boldsymbol{\beta} \in \ell^p(\mathbb{N})$ for p as in Proposition (2.1). This gives rise to (13), see [12]. Analogously, (12) is originated by bounds of the form [23]

$$\sup_{\boldsymbol{y}\in U} \big\|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\boldsymbol{y})\big\|_V \le C\boldsymbol{\nu}!\boldsymbol{\beta}^{\boldsymbol{\nu}} \quad \forall \boldsymbol{\nu} \in \mathbb{N}_0^{\mathbb{N}}, |\boldsymbol{\nu}| < \infty. \tag{19}$$

For completeness, we mention a criterion to verify (18) in the next Theorem.

Theorem 2.2. *Let V, W be reflexive Banach spaces and $A_{\boldsymbol{y}} \in \mathcal{L}(V, W^*)$ be affine parametric, that is there exist a family $\{A_j\}_{j=0,\dots,s} \subset \mathcal{L}(V, W^*)$ such that A_0 is boundedly invertible and*

$$A_{\boldsymbol{y}} = A_0 + \sum_{j=1}^{s} y_j A_j \quad \text{in } W^*. \tag{20}$$

Assume that $\tilde{\boldsymbol{\beta}} = (\tilde{\beta}_j) \in \ell^p(\mathbb{N})$, $p \in (0,1]$ defined by $\tilde{\beta}_j := \big\|A_0^{-1} A_j\big\|_{\mathcal{L}(V,V)}$ satisfies

$$\sum_{j\ge 1} \tilde{\beta}_j < 2. \tag{21}$$

Moreover, for a sequence $\boldsymbol{\beta}_{\mathfrak{l}} \in \ell^p(\mathbb{N})$ assume that $\mathfrak{l}_{\boldsymbol{y}} \in C^\infty(U, W^*)$ and

$$\sup_{\boldsymbol{y}\in U} \big\|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} \mathfrak{l}_{\boldsymbol{y}}\big\|_{W^*} \le C_{\mathfrak{l}} |\boldsymbol{\nu}|! \boldsymbol{\beta}_{\mathfrak{l}}^{\boldsymbol{\nu}}, \quad \forall \boldsymbol{\nu} \in \mathbb{N}_0^{\mathbb{N}}, |\boldsymbol{\nu}| < \infty. \tag{22}$$

Then, (18) holds with $\kappa = 0$, $C = \frac{2C_{\mathfrak{l}} \big\|A_0^{-1}\big\|_{\mathcal{L}(W^*,V)}}{2 - \sum_{j\ge 1} \tilde{\beta}_j}$ and $\boldsymbol{\beta} := \boldsymbol{\beta}_{\mathfrak{l}} + \tilde{\boldsymbol{\beta}}$. If $\mathfrak{l} = \mathfrak{l}_{\boldsymbol{y}}$ is independent of $\boldsymbol{y}$, the same estimate holds with $\boldsymbol{\beta} := \tilde{\boldsymbol{\beta}}$.

Proof. The proof follows the same arguments of [8, Theorem 4.3], see also [32]. □

In our setting, if $\mathfrak{a}_{\boldsymbol{y}}$ is affine parametric, then we can define $A_{\boldsymbol{y}} \in \mathcal{L}(V, W^*)$ by $A_{\boldsymbol{y}} v := \mathfrak{a}_{\boldsymbol{y}}(v, \cdot) \in W^*$ for all $v \in V$. Hence, it is sufficient to verify (21) and (22) to obtain our working Assumption 2.1 and (18).

However, the latter assumptions can be verified with alternative methods, also for non-affine parametric operators, based on holomorphic extensions of $\mathfrak{a}_{\boldsymbol{y}}$ for complex parameters $\boldsymbol{y} \in \tilde{U} \subseteq \mathbb{C}^s$. For more details we refer to [13]. On the other hand, (19) can also be verified in some situations [23]. In what follows, we will assume that Assumption 2.1 and either (18) or (19) are available for the parametric solution map $u \in L^\infty(U, V)$.

2.2 Modules of AFEM

We mentioned that discretization error occurs in the solution of (3), for any instance of $\boldsymbol{y} \in U$. In this section we precise our discretization method of choice.

For simplicity we restrict to polyhedral Lipschitz domains $D \subset \mathbb{R}^d$, $d \in \{2, 3\}$. A mesh $\mathcal{T}$ on D is defined as a finite collection of compact sets $T \in \mathcal{T}$, $|T| > 0$ such that $\bigcup_{T\in\mathcal{T}} T = \overline{D}$ and $|T \cap T'| = 0$, for all $T, T' \in \mathcal{T}$ with $T \neq T'$. We assume availability of finite-dimensional spaces $V_\mathcal{T} \subset V, W_\mathcal{T} \subset W$ linked to a mesh $\mathcal{T}$ on D with $\dim(V_\mathcal{T}) = \dim(W_\mathcal{T})$ and such that the following stable discrete inf-sup condition hold: for $\tilde{\lambda} > 0$ independent of $\mathcal{T} \in \mathbb{T}$ and $\boldsymbol{y} \in U$,

$$\begin{aligned} \inf_{0\neq v\in V_\mathcal{T}} \sup_{0\neq w\in W_\mathcal{T}} \frac{\mathfrak{a}_{\boldsymbol{y}}(v,w)}{\|v\|_V \|w\|_W} \geq \tilde{\lambda} > 0, \\ \inf_{0\neq w\in W_\mathcal{T}} \sup_{0\neq v\in V_\mathcal{T}} \frac{\mathfrak{a}_{\boldsymbol{y}}(v,w)}{\|v\|_V \|w\|_W} \geq \tilde{\lambda} > 0. \end{aligned} \tag{23}$$

Then, $u_\mathcal{T}(\boldsymbol{y}) \in V_\mathcal{T}$ denotes the unique solution of the problem

$$\mathfrak{a}_{\boldsymbol{y}}(u_\mathcal{T}(\boldsymbol{y}), w) = \mathfrak{l}_{\boldsymbol{y}}(w) \quad \forall w \in V_\mathcal{T}, \tag{24}$$

corresponding to (3).

We will often use the shorthand notation $\mathcal{T} \leq \mathcal{T}'$, meaning that the mesh $\mathcal{T}'$ can be obtained from another mesh $\mathcal{T}$ by possibly multiple applications of the module REFINE, as described in Assumption 2.2 below. Further, we fix an initial mesh $\mathcal{T}_0$ of D and we denote by $\mathbb{T} := \{\mathcal{T} : \mathcal{T}_0 \leq \mathcal{T}\}$ the set of admissible refinements of the initial mesh $\mathcal{T}_0$.

The well-established Adaptive FEM algorithm, see Algorithm 1, is composed of the four modules SOLVE, ESTIMATE, MARK and REFINE, plus a stopping criterion determined by a given tolerance ε.

Algorithm 1 AFEM

Input: $\mathfrak{a}, \mathfrak{l}, \varepsilon, \mathcal{T}_0$
Output: $u_\mathcal{T}, \mathcal{T}$

```
T ← T_0
while True do
    u_T ← SOLVE(a, l, T)
    {η(T)} ← ESTIMATE(u_T)
    if Σ η²(T) ≤ ε² then
        return u_T, T
    end if
    M ← MARK({η(T)})
    T ← REFINE(T, M)
end while
```

Following the description in [21], we state the abstract assumptions for Algorithm 1 to ensure error convergence of AFEM, pointwise for all $\boldsymbol{y} \in U$.

Assumption 2.2. AFEM *modules for parametric problems:*

- *For given $\boldsymbol{y} \in U$ and $u_\mathcal{T}(\boldsymbol{y}) \in V_\mathcal{T}$,* ESTIMATE *computes positive real numbers $\{\eta_{\boldsymbol{y},\mathcal{T}}(T)\}_{T\in\mathcal{T}}$, called indicators. We assume that the indicators satisfy, for all $\mathcal{T}, \mathcal{T}'$ with $\mathcal{T}_0 \leq \mathcal{T} \leq \mathcal{T}'$ the stability over non-refined elements*

$$\left(\sum_{T\in\mathcal{T}\cap\mathcal{T}'} \eta^2_{\boldsymbol{y},\mathcal{T}'}(T)\right)^{\frac{1}{2}} \leq \left(\sum_{T\in\mathcal{T}\cap\mathcal{T}'} \eta^2_{\boldsymbol{y},\mathcal{T}}(T)\right)^{\frac{1}{2}} + S(\|u_{\mathcal{T}'}(\boldsymbol{y}) - u_\mathcal{T}(\boldsymbol{y})\|_V), \tag{25}$$

and reduction over refined elements

$$\sum_{T\in\mathcal{T}'\setminus\mathcal{T}} \eta^2_{\boldsymbol{y},\mathcal{T}'}(T) \le q_{red} \sum_{T\in\mathcal{T}\setminus\mathcal{T}'} \eta^2_{\boldsymbol{y},\mathcal{T}}(T) + R(\|u_{\mathcal{T}'}(\boldsymbol{y}) - u_{\mathcal{T}}(\boldsymbol{y})\|_V), \tag{26}$$

where $q_{red} \in (0,1)$ *and the functions* $S, R\colon [0,\infty) \to [0,\infty)$ *are continuous at* 0 *with* $S(0) = R(0) = 0$ *and monotocally increasing. We assume that* $S(\cdot), R(\cdot), q_{red}$ *are independent of* $\boldsymbol{y} \in U$*. Furthermore, we assume reliability: there exists a constant* $c^* > 0$ *such that* $\forall\, \mathcal{T} \in \mathbb{T}, \forall\, \boldsymbol{y} \in U$

$$\|u(\boldsymbol{y}) - u_{\mathcal{T}}(\boldsymbol{y})\|_V \le c^* \left(\sum_{T\in\mathcal{T}} \eta^2_{\boldsymbol{y},\mathcal{T}}(T) \right)^{\frac{1}{2}}. \tag{27}$$

- *The marking procedure* MARK *selects, based on a set of indicators* $\{\eta_{\mathcal{T}}(T)\}$ *computed in the previous step, a subset* $\mathcal{M} \subset \mathcal{T}$ *of simplices that will be refined. We assume that there exists a function* $M\colon [0,\infty) \to [0,\infty)$ *continuous at* 0 *with* $M(0) = 0$ *such that*

$$\max_{T\in\mathcal{T}\setminus\mathcal{M}} \eta_{\mathcal{T}}(T) \le M\left(\left(\sum_{T\in\mathcal{M}} \eta^2_{\mathcal{T}}(T) \right)^{\frac{1}{2}} \right). \tag{28}$$

- *The* REFINE *module, for a given mesh* $\mathcal{T}$ *and a set of marked elements* $\mathcal{M} \subseteq \mathcal{T}$*, produces a new mesh* $\mathcal{T}'$ *such that* $\mathcal{T}' \cap \mathcal{M} = \emptyset$*. We assume that parents are union of their children, that is* $T = \bigcup \{T' \in \mathcal{T}' : T' \subseteq T\}$ *for all* $T \in \mathcal{T}$*. We stress that* $\mathcal{M} \subseteq \mathcal{T} \setminus \mathcal{T}'$*, that is* REFINE *can in principle refine more than the marked set. To simplify the presentation, we further assume conformity* $V_{\mathcal{T}} \subseteq V_{\mathcal{T}'} \subset V$ *for all* $\mathcal{T} \le \mathcal{T}'$*,* $\mathcal{T}, \mathcal{T}' \in \mathbb{T}$*.*

- *For the module* SOLVE*, we assume that the Galerkin solution* $u_{\mathcal{T}}(\boldsymbol{y})$ *of* (24) *can be recovered exactly for every* $\boldsymbol{y} \in U$*, which entails exact integration and linear algebra.*

We stress that the availability of c^* (27) depends implicitly on the set $\mathbb{T}$, and hence on the REFINE module. In practice, usually c^* depends on λ, Λ from (4),(5) and on the shape regularity of a mesh $\mathcal{T}$, and hence it is often required that REFINE does not generate strongly anisotropic meshes, i.e. $\mathbb{T}$ is uniformly shape-regular. Typical MARK strategies, as the Dörfler criterion, are known to satisfy (28), see e.g. [21].

Let $(u_{\mathcal{T}_\ell(\boldsymbol{y})}(\boldsymbol{y}))_{\ell\in\mathbb{N}}$, $\mathcal{T}_\ell := \mathcal{T}_\ell(\boldsymbol{y})$ be the sequence of approximations produced by the AFEM loop with $\mathcal{T}_{\ell+1}(\boldsymbol{y}) = \text{REFINE}(\mathcal{T}_\ell(\boldsymbol{y}), \mathcal{M}_\ell(\boldsymbol{y}))$, $\mathcal{M}_\ell(\boldsymbol{y}) = \text{MARK}(\{\eta_{\boldsymbol{y},\mathcal{T}_\ell(\boldsymbol{y})}(T)\}) \subseteq \mathcal{T}_\ell(\boldsymbol{y})$ for all $\ell \in \mathbb{N}$; then, as a corollary of [21, Theorem 3.1] we get the following pointwise convergence result.

Lemma 2.3. *Consider a problem of the form* (3) *satisfying Assumption 2.1. Let* AFEM *satisfy* (23) *and Assumption 2.2. Then, for all* $\boldsymbol{y} \in U$ *and all initial meshes* $\mathcal{T} \in \mathbb{T}$*, it returns in finite time* $\mathcal{T}(\boldsymbol{y})$ *and* $u_{\mathcal{T}(\boldsymbol{y})}(\boldsymbol{y})$ *such that*

$$\left\|u(\boldsymbol{y}) - u_{\mathcal{T}(\boldsymbol{y})}(\boldsymbol{y})\right\|_V \le c^* tol, \tag{29}$$

for a constant $c^* > 0$ *independent of* $tol, \boldsymbol{y}$ *and* $\mathcal{T} \in \mathbb{T}$*.*

Proof. From [31], for all $\boldsymbol{y}$ there exists $u_\infty(\boldsymbol{y}) \in V$ such that

$$\left\|u_\infty(\boldsymbol{y}) - u_{\mathcal{T}_\ell(\boldsymbol{y})}(\boldsymbol{y})\right\|_V \to 0 \quad \text{as } \ell \to \infty. \tag{30}$$

Hence, the result follows from [21, Theorem 3.1] and reliability (27). □

3 QMC-AFEM algorithms

3.1 A first convergence result

In this section we present a first combined QMC-AFEM algorithm, that outputs an approximation of $I(G(u))$ for a given tolerance ε. Algorithm 2 is in fact not efficient for implementation, but it illustrates effectively the key ideas.

First of all, we observe that E_{b^m} can be fully evaluated by means of quantities $Q_{b^m}, Q_{b^{m-1}}$ that have already been computed, when we loop over m. In other words, when adding more QMC points we reuse part of the work done previously so that the cost to compute E_{b^m} is negligible. A second crucial observation is that each call of the Algorithm 1 results (in principle) in a different mesh $\mathcal{T}(\boldsymbol{y}) \geq \mathcal{T}_0$, starting from a common coarse mesh $\mathcal{T}_0$. In particular, $G(u_{\mathcal{T}(\boldsymbol{y})}(\boldsymbol{y}))$ may not be even continuous with respect to $\boldsymbol{y} \in U$, and hence in general $G(u_{\mathcal{T}(\cdot)}(\cdot)) \notin \mathcal{W}_{s,\alpha,\boldsymbol{\gamma}}$ regardless of the discretization scheme.

Algorithm 2 QMC-AFEM

Input: $[\boldsymbol{y} \mapsto \mathfrak{a}_{\boldsymbol{y}}], [\boldsymbol{y} \mapsto \mathfrak{l}_{\boldsymbol{y}}], G, \varepsilon, \mathcal{T}_0$
Output: Approximation of $I(G(u))$ within tolerance $\propto \varepsilon$

$m \leftarrow 1$
while True **do**
  Generate lattice P_m
  for $\boldsymbol{y} \in P_m$ **do**
    $u_{\mathcal{T}(\boldsymbol{y})}(\boldsymbol{y}) \leftarrow \mathrm{AFEM}(\mathfrak{a}_{\boldsymbol{y}}, \mathfrak{l}_{\boldsymbol{y}}, \varepsilon_F, \mathcal{T}_0)$ ▷ Algorithm 1
    Evaluate $G(u_{\mathcal{T}(\boldsymbol{y})}(\boldsymbol{y}))$
  end for
  if $m \geq 2$ **and** $|E_{b^m}(G(u_{\mathcal{T}}))| \leq \varepsilon_Q$ **then**
    return $Q_{b^m}(G(u_{\mathcal{T}}))$
  end if
  $m \leftarrow m + 1$
end while

Proposition 3.1. *Let $s, \alpha \in \mathbb{N}$, $\alpha \geq 2$ and given initial mesh $\mathcal{T}_0 \in \mathbb{T}$. Assume that $G \in V^*$, $u \in \mathcal{W}_{s,\alpha,\boldsymbol{\gamma}}(U, V)$ for some weights $\boldsymbol{\gamma}$ as in Proposition 2.1 and that $\forall\, \boldsymbol{y} \in U$ and any tolerance $tol > 0$,* AFEM *returns in finite time $u_{\mathcal{T}(\boldsymbol{y})}(\boldsymbol{y})$ such that* (29) *holds for a constant c^* independent of tol. Then, for any $\varepsilon > 0$, there exist choices ε_Q and $\varepsilon_F := tol$, with $\varepsilon^{-1}\varepsilon_F, \varepsilon^{-1}\varepsilon_Q$ independent of ε, such that*

1. *Algorithm 2 stops in finite time and*
2. *it produces an approximation of $I(G(u))$ within tolerance $c^*\varepsilon + \mathcal{O}\big(b^{-2m+\delta}\big)$.*

Proof. Let ε_F be the tolerance for AFEM. To prove the first item it is sufficient to show that, for any $\varepsilon_Q > \frac{2}{b-1} c^* \|G\|_{V^*} \varepsilon_F$ there exists m sufficiently large such that $|E_{b^m}(G(u_{\mathcal{T}}))| \leq \varepsilon_Q$. By linearity of G,

$$\begin{aligned}
|E_{b^m}(G(u_{\mathcal{T}}))| &\leq |E_{b^m}(G(u - u_{\mathcal{T}}))| + |E_{b^m}(G(u))| \\
&\leq \frac{2}{b-1} \max_{\boldsymbol{y} \in P_{m-1} \cup P_m} G(u - u_{\mathcal{T}(\boldsymbol{y})})(\boldsymbol{y}) + |E_{b^m}(G(u))| \\
&\leq \frac{2}{b-1} c^* \|G\|_{V^*} \varepsilon_F + |E_{b^m}(G(u))|.
\end{aligned}$$

Proposition 2.1 also implies that $|E_{b^m}(G(u))| \to 0$ as $m \to \infty$ and hence the claim.

Now we show the second item: we separate the error due to the Finite Element discretization from the QMC integration error as follows

$$|I(G(u)) - Q_{b^m}(G(u_{\mathcal{T}}))| \leq |Q_{b^m}(G(u - u_{\mathcal{T}}))| + |I(G(u)) - Q_{b^m}(G(u))|.$$

For the FEM error we have

$$|Q_{b^m}(G(u - u_{\mathcal{T}}))| \leq \max_{\boldsymbol{y} \in P_m} |G(u(\boldsymbol{y}) - u_{\mathcal{T}}(\boldsymbol{y}))| \leq c^* \|G\|_{V^*} \varepsilon_F.$$

For the QMC error we apply Proposition 2.1 to get for all $\delta > 0$

$$\begin{aligned} |I(G(u)) - Q_{b^m}(G(u))| &\leq \left(|E_{b^m}(G(u_{\mathcal{T}}))| + \tfrac{2}{b-1} c^* \|G\|_{V^*} \varepsilon_F\right) + \mathcal{O}\left(b^{-2m+\delta}\right) \\ &\leq \left(\varepsilon_Q + \tfrac{2}{b-1} c^* \|G\|_{V^*} \varepsilon_F\right) + \mathcal{O}\left(b^{-2m+\delta}\right). \end{aligned}$$

Hence, for given ε we can choose $\varepsilon_F := \frac{(b-1)\varepsilon}{2(b+1)\|G\|_{V^*}}$ and $\varepsilon_Q := \frac{2}{5} c^* \varepsilon > \frac{2}{b-1} c^* \|G\|_{V^*} \varepsilon_F$ and obtain

$$|I(G(u)) - Q_{b^m}(G(u_{\mathcal{T}}))| \leq \left(\varepsilon_Q + (\tfrac{2}{b-1} + 1) c^* \|G\|_{V^*} \varepsilon_F\right) + \mathcal{O}\left(b^{-2m+\delta}\right) \leq c^* \varepsilon + \mathcal{O}\left(b^{-2m+\delta}\right). \quad (31)$$

□

We remark that a sharp value for the reliability constant c^* is usually not known but (potentially pessimistic) upper bounds exist. The size of c^* can be controlled for structured meshes and refinement by bisection in spatial dimension $d = 2$.

Algorithm 2 entails a decoupling of the QMC sampling with a AFEM solver. In practice, this implies that an adaptive software can be integrated into such algorithm in a *non-intrusive* manner, provided that the reliability (29) is satisfied for some variational space V and $G \in V^*$.

This feature can be advantageous in many situations, especially when a solver is complex to implement. However, it presents two main computational difficulties:

- Algorithm 2 recomputes a mesh $\mathcal{T} \geq \mathcal{T}_0$ for the domain D, for each QMC sample $\boldsymbol{y} \in P_m$ as well as for all iterations over m, which for complex geometries is an expensive step.
- Imposing the same AFEM threshold (29) for all QMC points can be unnecessary since we are primarily interested in the average over the parameter space.

In the following we propose two alternative algorithms that improve upon Algorithm 2 under these aspects. The first is a modification of Algorithm 2, that recycles part of the computation from previous iterations over m.

Algorithm 3 is motivated by the following heuristics. If there exists a metric $d_{\boldsymbol{\gamma}} \colon U \times U \to [0, \infty)$ and a Lipschitz constant $L > 0$ satisfying

$$\max(\|u_{\mathcal{T}}(\boldsymbol{y}) - u_{\mathcal{T}}(\boldsymbol{y}')\|_V, \|u(\boldsymbol{y}) - u(\boldsymbol{y}')\|_V) \leq L d_{\boldsymbol{\gamma}}(\boldsymbol{y}, \boldsymbol{y}') \qquad \forall\, \boldsymbol{y}, \boldsymbol{y}' \in U, \forall\, \mathcal{T} \in \mathbb{T}, \quad (32)$$

then

$$\begin{aligned} \left\|u(\boldsymbol{y}) - u_{\mathcal{T}(\boldsymbol{y}')}(\boldsymbol{y})\right\|_V &\leq 2L d_{\boldsymbol{\gamma}}(\boldsymbol{y}, \boldsymbol{y}') + \left\|u(\boldsymbol{y}') - u_{\mathcal{T}(\boldsymbol{y}')}(\boldsymbol{y}')\right\|_V \\ &\leq 2L d_{\boldsymbol{\gamma}}(\boldsymbol{y}, \boldsymbol{y}') + c^* \varepsilon_F. \end{aligned}$$

In particular, for a small distance of the parameters we have a good chance to meet the AFEM tolerance by just one call of the SOLVE module, starting from the mesh $\mathcal{T}(\boldsymbol{y}')$.

Following verbatim the proof of Proposition 3.1, we get convergence Algorithm 3. The parameter $q \in \mathbb{N}$ in line 6 regulates how much information from previous iterations we use. A discussion of possible choices $q = q(\varepsilon)$ depending on the tolerance is given in Section 3.5 below.

Algorithm 3 QMC-AFEM (v2)

Input: $[\boldsymbol{y} \mapsto \mathfrak{a}_{\boldsymbol{y}}], [\boldsymbol{y} \mapsto \mathfrak{l}_{\boldsymbol{y}}], G, \varepsilon, \mathcal{T}_0, q$
Output: Approximation of $I(G(u))$ within tolerance $\propto \varepsilon$

$m \leftarrow 1$; generate P_1
$\mathcal{T}(\boldsymbol{y}') \leftarrow \mathcal{T}_0 \; \forall \boldsymbol{y}' \in P_1$
while True **do**
  for $\boldsymbol{y} \in P_m$ **do**
    if $m > 1$ **then**
      $\boldsymbol{y}' \leftarrow \operatorname{argmin} \left\{ d_{\boldsymbol{\gamma}}(\boldsymbol{y}, \boldsymbol{z}) : \boldsymbol{z} \in P_{\min(m-1,q)} \right\}$
    end if
    $u_{\mathcal{T}(\boldsymbol{y})}(\boldsymbol{y}), \mathcal{T}(\boldsymbol{y}) \leftarrow \mathrm{AFEM}(\mathfrak{a}_{\boldsymbol{y}}, \mathfrak{l}_{\boldsymbol{y}}, \varepsilon_F, \mathcal{T}(\boldsymbol{y}'))$ ▷ Algorithm 1
    Evaluate $G(u_{\mathcal{T}(\boldsymbol{y})}(\boldsymbol{y}))$
  end for
  if $m \geq 2$ **and** $|E_{b^m}(G(u_{\mathcal{T}}))| \leq \varepsilon_Q$ **then**
    return $Q_{b^m}(G(u_{\mathcal{T}}))$
  end if
  $m \leftarrow m + 1$
  Generate lattice P_m
end while

Proposition 3.2. *Let $s, \alpha \in \mathbb{N}$, $\alpha \geq 2$ and given $\mathcal{T}_0 \in \mathbb{T}$. Assume that $G \in V^*$, $u \in \mathcal{W}_{s,\alpha,\boldsymbol{\gamma}}(U, V)$ for some weights $\boldsymbol{\gamma}$ as in Proposition 2.1 and that $\forall\, \boldsymbol{y} \in U$, $\forall\, \mathcal{T} \in \mathbb{T}$ and any tolerance tol, AFEM, starting from the initial mesh $\mathcal{T}$, returns in finite time a mesh $\mathcal{T}(\boldsymbol{y}) \geq \mathcal{T}$ and $u_{\mathcal{T}(\boldsymbol{y})}(\boldsymbol{y})$ such that* (29) *holds for a constant c^* independent of tol and $\mathcal{T}$. Then, for any ε, there exist choices ε_Q and $\varepsilon_F := tol$, with $\varepsilon^{-1}\varepsilon_F, \varepsilon^{-1}\varepsilon_Q$ independent of ε such that*

1. *Algorithm 3 stops in finite time and*
2. *it produces an approximation of $I(G(u))$ within tolerance $c^*\varepsilon + \mathcal{O}\big(b^{-2m+\delta}\big)$.*

3.2 Goal oriented AFEM – part 1

For the convergence of Algorithm 2 and 3, we assumed (29). This assumption alone does not yield optimal convergence rate of the AFEM module; as a consequence, the Finite Element error is overestimated and the spatial domain D could be overrefined in the algorithms. Nevertheless, we only require a reliable upper bound for the difference $|G(u(\boldsymbol{y})) - G(u_{\mathcal{T}}(\boldsymbol{y}))|$, that in many situations converges to 0 faster than $\|u(\boldsymbol{y}) - u_{\mathcal{T}}(\boldsymbol{y})\|_V$ as we refine $\mathcal{T}$, by an Aubin-Nitsche duality argument.

Let $\mathcal{T} \in \mathbb{T}, \boldsymbol{y} \in U$, then we define $z(\boldsymbol{y}) \in W$ as the unique solution of the dual problem

$$\mathfrak{a}_{\boldsymbol{y}}(v, z(\boldsymbol{y})) = G(v) \quad \forall v \in V. \tag{33}$$

Then, for all $w_{\mathcal{T}} \in W_{\mathcal{T}}$,

$$\begin{aligned} |G(u(\boldsymbol{y})) - G(u_{\mathcal{T}}(\boldsymbol{y}))| &= |\mathfrak{a}_{\boldsymbol{y}}(u(\boldsymbol{y}), z(\boldsymbol{y})) - \mathfrak{a}_{\boldsymbol{y}}(u_{\mathcal{T}}(\boldsymbol{y}), z(\boldsymbol{y}))| \\ &= |\mathfrak{a}_{\boldsymbol{y}}(u(\boldsymbol{y}) - u_{\mathcal{T}}(\boldsymbol{y}), z(\boldsymbol{y}) - w_{\mathcal{T}})| \\ &\leq \Lambda \, \|u(\boldsymbol{y}) - u_{\mathcal{T}}(\boldsymbol{y})\|_V \, \|z(\boldsymbol{y}) - w_{\mathcal{T}}\|_W . \end{aligned} \tag{34}$$

When the goal functional $G \in V^*$ has additional regularity, i.e. it belongs to a suitable subspace $H \subseteq V^*$, then for $h_{\mathcal{T}} := \max_{T \in \mathcal{T}} \operatorname{diam}(T)$

$$\lim_{h_{\mathcal{T}} \to 0} \inf_{w_{\mathcal{T}} \in V_{\mathcal{T}}} \|z(\boldsymbol{y}) - w_{\mathcal{T}}\|_W = 0. \tag{35}$$

However, in general AFEM produces non quasi-uniform meshes. Hence we can exploit regularity of G as follows: we pick $w_{\mathcal{T}} := z_{\mathcal{T}}(\boldsymbol{y}) \in W_{\mathcal{T}}$ the FE solution of

$$\mathfrak{a}_{\boldsymbol{y}}(v, z_{\mathcal{T}}(\boldsymbol{y})) = G(v) \quad \forall v \in V_{\mathcal{T}} \tag{36}$$

and find reliable indicators $\{\zeta_{\boldsymbol{y},\mathcal{T}}(T)\}_{T\in\mathcal{T}}$ such that, for a constant $c^{**} > 0$ independent of $\boldsymbol{y}, \mathcal{T} \in \mathbb{T}$,

$$\|z(\boldsymbol{y}) - z_{\mathcal{T}}(\boldsymbol{y})\|_W \leq c^{**} \left(\sum_{T\in\mathcal{T}} \zeta^2_{\boldsymbol{y},\mathcal{T}}(T) \right)^{\frac{1}{2}}. \tag{37}$$

Combining (34) and (37) we can use the following a-posteriori estimator as termination criterion for AFEM, in Algorithms 2 and 3

$$|G(u(\boldsymbol{y})) - G(u_{\mathcal{T}}(\boldsymbol{y}))| \lesssim \left(\sum_{T\in\mathcal{T}} \eta^2_{\boldsymbol{y},\mathcal{T}}(T) \sum_{T\in\mathcal{T}} \zeta^2_{\boldsymbol{y},\mathcal{T}}(T) \right)^{\frac{1}{2}} \leq \varepsilon_F. \tag{38}$$

Remark 3.3. *Note that in this case we must solve numerically the dual problem for each sample $\boldsymbol{y} \in P_m, m = 1, 2, \ldots$, until the tolerance is met. However, the stiffness matrix of the dual problem coincides with the transpose of the stiffness matrix of the primal, thus the additional work for the solution of* (36) *includes only the construction of the load vector corresponding to G (independent of $\boldsymbol{y}$) and one linear solver per sample – in particular it is independent of the parametric dimension $s \in \mathbb{N}$.*

Furthermore, as shown in [17], suitable marking strategies driven by both indicators $\eta_{\boldsymbol{y},\mathcal{T}}, \zeta_{\boldsymbol{y},\mathcal{T}}$ yield optimal convergence of the resulting *goal oriented AFEM* (or goAFEM) algorithm, provided that the axioms of adaptivity (A1-A4 in [5]) hold for the indicators $\eta_{\boldsymbol{y},\mathcal{T}}(T), \zeta_{\boldsymbol{y},\mathcal{T}}(T)$.

Remark 3.4. *(A1), (A2) are analogous to* (25) *and* (26)*, while (A4) is a discrete version of* (27)*. Quasi-orthogonality (A3) holds trivially for symmetric bilinear forms $\mathfrak{a}_{\boldsymbol{y}}$, $\boldsymbol{y} \in U$, although here we do not assume symmetry and we must verify (A3) on a case by case basis, so to obtain optimal convergence of* goAFEM*.*

If in Algorithms 2 and 3, we replace AFEM by goAFEM, then the results of Propositions 3.1 and 3.2 remain valid. As a side advantage, we do not need to include $\|G\|_{V^*}$ to the FEM tolerance ε_F.

3.3 Indicator averaging

Next, we design an iterative algorithm that refines the mesh or increases the number of samples at each step. Conversely to the previous algorithms, at any given time we employ only one mesh of the domain D for all $\boldsymbol{y} \in U$. In this case, we will assume *a-priori uniform convergence*, slightly stronger than the a-priori convergence in (30).

Assumption 3.1. *Denote by $(u_{\mathcal{T}_\ell}(\boldsymbol{y}))_{\ell\in\mathbb{N}_0}$ the sequence of approximations produced by Algorithm 4 with $\mathcal{T}_{\ell+1} = \mathrm{REFINE}(\mathcal{T}_\ell, \mathcal{M}_\ell)$, $\mathcal{M}_\ell = \mathrm{MARK}(\{\eta_{\mathcal{T}_\ell}(T)\}) \subseteq \mathcal{T}_\ell$ for all $\ell \in \mathbb{N}_0$. We assume that there exists $u_\infty \in C^0(U, V)$ such that*

$$\|u_\infty - u_{\mathcal{T}_\ell}\|_{L^\infty(U,V)} \to 0 \quad \text{as } \ell \to \infty. \tag{39}$$

Theorem 3.5. *Let $s, \alpha \in \mathbb{N}$, $\alpha \geq 2$ and given $\mathcal{T}_0 \in \mathbb{T}$. Assume that $G \in V^*$, that $u, u_{\mathcal{T}} \in \mathcal{W}_{s,\alpha,\boldsymbol{\gamma}}(U, V)$, $\forall \mathcal{T} \in \mathbb{T}$ for $\boldsymbol{\gamma}$ as in Proposition 2.1 and that the AFEM modules satisfy Assumption 2.2 and Assumption 3.1. Then, for all $\varepsilon > 0$ there exist $\varepsilon_F, \varepsilon_Q$, with $\varepsilon^{-1}\varepsilon_F, \varepsilon^{-1}\varepsilon_Q$ independent of ε, such that*

Algorithm 4 AQMC-FEM

Input: $[\boldsymbol{y} \mapsto \mathfrak{a}_{\boldsymbol{y}}], [\boldsymbol{y} \mapsto \mathfrak{l}_{\boldsymbol{y}}], G, \varepsilon, \mathcal{T}_0$
Output: Approximation of $I(G(u))$ within tolerance $\propto \varepsilon$

$m \leftarrow 2$; generate P_1, P_2
$\mathcal{T} \leftarrow \mathcal{T}_0$
while True **do**
  for $\boldsymbol{y} \in P_m$ **do**
    $u_{\mathcal{T}}(\boldsymbol{y}) \leftarrow \mathrm{SOLVE}(\mathfrak{a}_{\boldsymbol{y}}, \mathfrak{l}_{\boldsymbol{y}}, \mathcal{T})$
    $\{\eta_{\boldsymbol{y},\mathcal{T}}(T)\} \leftarrow \mathrm{ESTIMATE}(u_{\mathcal{T}}(\boldsymbol{y}))$
    Evaluate $G(u_{\mathcal{T}}(\boldsymbol{y}))$
  end for
  $\eta^2_{\mathcal{T}}(T) \leftarrow \frac{1}{b^m} \sum_{\boldsymbol{y} \in P_m} \eta^2_{\boldsymbol{y},\mathcal{T}}(T)$
  if $\sqrt{\sum_{T \in \mathcal{T}} \eta^2_{\mathcal{T}}(T)} > \varepsilon_F$ **then**
    $\mathcal{T} \leftarrow \mathrm{REFINE}(\mathcal{T}, \mathrm{MARK}(\{\eta_{\mathcal{T}}(T)\}))$
  else
    if $Q_{b^{m-1}}(G(u_{\mathcal{T}}))$ was not computed for the current $\mathcal{T}$ **then**
      for $\boldsymbol{y} \in P_{m-1}$ **do**
        $u_{\mathcal{T}}(\boldsymbol{y}) \leftarrow \mathrm{SOLVE}(\mathfrak{a}_{\boldsymbol{y}}, \mathfrak{l}_{\boldsymbol{y}}, \mathcal{T})$
        Evaluate $G(u_{\mathcal{T}}(\boldsymbol{y}))$
      end for
    end if
    if $|E_{b^m}(G(u_{\mathcal{T}}))| \leq \varepsilon_Q$ **then**
      return $Q_{b^m}(G(u_{\mathcal{T}}))$
    end if
    $m \leftarrow m + 1$
    Generate lattice P_m
  end if
end while

1. *Algorithm 4 stops in finite time and*

2. *it produces an approximation of $I(G(u))$ within tolerance $c^*\varepsilon + \mathcal{O}\big(b^{-2m+\delta}\big)$.*

Proof. Fix $m \in \mathbb{N}$. For a mesh $\mathcal{T}$ on D define $\eta_{\mathcal{T}}(T) := \Big(\frac{1}{b^m}\sum_{\boldsymbol{y}\in P_m}\eta^2_{\boldsymbol{y},\mathcal{T}}(T)\Big)^{\frac{1}{2}}$ the quadratic mean over $\boldsymbol{y} \in P_m$ of the local indicators. Since $\{\eta_{\boldsymbol{y},\mathcal{T}}(T)\}_T$ satisfy (26) for all $\boldsymbol{y} \in P_m$, and any $\mathcal{T}_0 \le \mathcal{T} \le \mathcal{T}'$, we get

$$\begin{aligned}\sum_{T\in\mathcal{T}'\setminus\mathcal{T}}\eta^2_{\mathcal{T}'}(T) &\le q_{red}\sum_{T\in\mathcal{T}\setminus\mathcal{T}'}\eta^2_{\mathcal{T}}(T) + \frac{1}{b^m}\sum_{\boldsymbol{y}\in P_m} R(\|u_{\mathcal{T}'}(\boldsymbol{y}) - u_{\mathcal{T}}(\boldsymbol{y})\|_V)\\ &\le q_{red}\sum_{T\in\mathcal{T}\setminus\mathcal{T}'}\eta^2_{\mathcal{T}}(T) + R(\|u_{\mathcal{T}'} - u_{\mathcal{T}}\|_{L^\infty(U,V)})\end{aligned}$$

as R is increasing. Hence, also $\eta_{\mathcal{T}}(T)$ has the reduction property (26), but with respect to the $L^\infty(U,V)$-norm. Similarly, from (25), monotonicity of S and Jensen inequality

$$\begin{aligned}\sum_{T\in\mathcal{T}\cap\mathcal{T}'}\eta^2_{\mathcal{T}'}(T) \le& \sum_{T\in\mathcal{T}\cap\mathcal{T}'}\eta^2_{\mathcal{T}}(T) + \frac{1}{b^m}\sum_{\boldsymbol{y}\in P_m} S(\|u_{\mathcal{T}'}(\boldsymbol{y}) - u_{\mathcal{T}}(\boldsymbol{y})\|_V)^2\\ &+ \frac{1}{b^m}\sum_{\boldsymbol{y}\in P_m} 2S(\|u_{\mathcal{T}'}(\boldsymbol{y}) - u_{\mathcal{T}}(\boldsymbol{y})\|_V)\left(\sum_{T\in\mathcal{T}\cap\mathcal{T}'}\eta^2_{\boldsymbol{y},\mathcal{T}}(T)\right)^{\frac{1}{2}}\\ \le& \sum_{T\in\mathcal{T}\cap\mathcal{T}'}\eta^2_{\mathcal{T}}(T) + S(\|u_{\mathcal{T}'} - u_{\mathcal{T}}\|_{L^\infty(U,V)})^2\\ &+ 2S(\|u_{\mathcal{T}'} - u_{\mathcal{T}}\|_{L^\infty(U,V)})\left(\sum_{T\in\mathcal{T}\cap\mathcal{T}'}\eta^2_{\mathcal{T}}(T)\right)^{\frac{1}{2}}.\end{aligned}$$

Taking square roots on both sides, we obtain that $\eta_{\mathcal{T}}(T)$ has the stability property (25) with respect to the $L^\infty(U,V)$-norm. Note that the modules MARK, REFINE are independent of $\boldsymbol{y} \in U$ and we assumed a-priori convergence in (39) in the same norm; therefore, from the proof [21, Theorem 3.1], for the sequence of meshes $(\mathcal{T}_\ell)_{\ell\in\mathbb{N}_0}$ constructed by

$$\mathcal{T}_{\ell+1} = \mathrm{REFINE}(\mathcal{T}_\ell, \mathrm{MARK}(\{\eta_{\mathcal{T}_\ell}(T)\})) \tag{40}$$

we obtain,

$$\sum_{T\in\mathcal{T}_\ell}\eta^2_{\mathcal{T}_\ell}(T) \to 0 \quad \text{as } \ell\to\infty. \tag{41}$$

Thus for all $m \in \mathbb{N}$, any FEM tolerance ε_F is met in finite time. Since $G(u_{\mathcal{T}}) \in \mathcal{W}_{s,\alpha,\boldsymbol{\gamma}}$ for all $\mathcal{T} \in \mathbb{T}$ (conversely to Algorithms 2,3, here there is only one mesh $\mathcal{T}$ at a time, used for all points $\boldsymbol{y} \in P_m \cup P_{m-1}$), Proposition 2.1 implies that $E_{b^m}(G(u_{\mathcal{T}})) \to 0$ as $m \to \infty$, showing that Algorithm 4 stops in finite time. The error bound follows as in Proposition 3.1: denote by $\mathcal{T}^{(m)} := \mathcal{T}_{\ell(m)}, m \in \mathbb{N}$ the mesh that meets the FEM error tolerance for the lattice P_m, i.e.

$$\frac{1}{b^m}\sum_{T\in\mathcal{T}^{(m)}}\sum_{\boldsymbol{y}\in P_m}\eta^2_{\boldsymbol{y},\mathcal{T}^{(m)}}(T) \le \varepsilon_F^2. \tag{42}$$

For all $\delta > 0$,

$$|I(G(u)) - Q_{b^m}(G(u_{\mathcal{T}^{(m)}}))| \le |Q_{b^m}(G(u - u_{\mathcal{T}^{(m)}}))| + |E_{b^m}(G(u - u_{\mathcal{T}^{(m)}}))|$$

$$+ |E_{b^m}(G(u_{\mathcal{T}^{(m)}}))| + \mathcal{O}\big(b^{-2m+\delta}\big)\,. \tag{43}$$

Jensen inequality and (27) give

$$|Q_{b^m}(G(u-u_{\mathcal{T}^{(m)}}))| \le \|G\|_{V^*} \frac{1}{b^m} \sum_{\boldsymbol{y}\in P_m} \|u(\boldsymbol{y}) - u_{\mathcal{T}^{(m)}}(\boldsymbol{y})\|_V \le c^* \|G\|_{V^*} \varepsilon_F. \tag{44}$$

Note that, since $\mathcal{T}^{(m-1)} \le \mathcal{T}^{(m)}$ as we never coarsen meshes, Galerkin orthogonality implies, for $C(\tilde{\lambda},\Lambda) = 1 + \frac{\Lambda}{\tilde{\lambda}}$

$$\frac{1}{b^{m-1}} \sum_{\boldsymbol{y}\in P_{m-1}} \|u(\boldsymbol{y}) - u_{\mathcal{T}^{(m)}}(\boldsymbol{y})\|_V \le \frac{C(\tilde{\lambda},\Lambda)}{b^{m-1}} \sum_{\boldsymbol{y}\in P_{m-1}} \|u(\boldsymbol{y}) - u_{\mathcal{T}^{(m-1)}}(\boldsymbol{y})\|_V \le C(\tilde{\lambda},\Lambda) c^* \varepsilon_F, \tag{45}$$

whence

$$|E_{b^m}(G(u-u_{\mathcal{T}^{(m)}}))| \le \frac{1+C(\tilde{\lambda},\Lambda)}{b-1} c^* \|G\|_{V^*} \varepsilon_F. \tag{46}$$

The stopping criterion gives $|E_{b^m}(G(u_{\mathcal{T}^{(m)}}))| \le \varepsilon_Q$ so that

$$|I(G(u)) - Q_{b^m}(G(u_{\mathcal{T}^{(m)}}))| \le \left(\frac{1+C(\tilde{\lambda},\Lambda)}{b-1} + 1\right) c^* \|G\|_{V^*} \varepsilon_F + \varepsilon_Q + \mathcal{O}\big(b^{-2m+\delta}\big) \tag{47}$$

and it is sufficient to pick $\varepsilon_F := \frac{(b-1)\varepsilon}{2(b+C(\tilde{\lambda},\Lambda))\|G\|_{V^*}}$, $\varepsilon_Q := \frac{c^*\varepsilon}{2}$ to get the claim. □

3.4 Goal oriented AFEM – part 2

We now include a goal oriented adaptivity approach in Algorithm 4. Given the estimator average $\varphi_{\mathcal{T}}(T) := \left(\frac{1}{b^m}\sum_{\boldsymbol{y}\in P_m} \varphi^2_{\boldsymbol{y},\mathcal{T}}(T)\right)^{\frac{1}{2}}$, $\varphi \in \{\eta,\zeta\}$, we define the following indicators from [1]

$$\rho^2_{\mathcal{T}}(T) := \eta^2_{\mathcal{T}}(T) \sum_{T'\in\mathcal{T}} \zeta^2_{\mathcal{T}}(T') + \zeta^2_{\mathcal{T}}(T) \sum_{T'\in\mathcal{T}} \eta^2_{\mathcal{T}}(T').$$

Proposition 3.6. *Let $\{\mathcal{T}_\ell\}_{\ell\in\mathbb{N}_0}$ be a sequence of meshes produced with the indicators $\rho_{\mathcal{T}}(T)$ by a marking and refinement strategy as in Assumption 2.2. Let $K_0 := \max_{\boldsymbol{y}\in U}(\eta^2_{\boldsymbol{y},\mathcal{T}_0} + \zeta^2_{\boldsymbol{y},\mathcal{T}_0}) < \infty$. Assume that both estimators $\eta_{\boldsymbol{y},\mathcal{T}}, \zeta_{\boldsymbol{y},\mathcal{T}}$ for the primal and dual problems satisfy reliability* (27)*, and* (37) *and the properties* (25), (26)*. Then*

$$\eta_{\mathcal{T}_\ell}\zeta_{\mathcal{T}_\ell} = \left(\sum_{T\in\mathcal{T}_\ell} \zeta^2_{\mathcal{T}_\ell}(T)\right)^{\frac{1}{2}} \left(\sum_{T\in\mathcal{T}_\ell} \eta^2_{\mathcal{T}_\ell}(T)\right)^{\frac{1}{2}} \to 0, \qquad \text{as } \ell\to\infty$$

Proof. Due to (23), we have quasi-optimality of the primal and dual problems

$$\|u(\boldsymbol{y}) - u_{\mathcal{T}}(\boldsymbol{y})\|_V \le C(\tilde{\lambda},\Lambda) \inf_{v_{\mathcal{T}}\in V_{\mathcal{T}}} \|u(\boldsymbol{y}) - v_{\mathcal{T}}\|_V$$
$$\|z(\boldsymbol{y}) - z_{\mathcal{T}}(\boldsymbol{y})\|_W \le C(\tilde{\lambda},\Lambda) \inf_{w_{\mathcal{T}}\in W_{\mathcal{T}}} \|z(\boldsymbol{y}) - w_{\mathcal{T}}\|_W\,,$$

for all $\mathcal{T}\in\mathbb{T}$. Hence, we get from [5, Lemma 3.6], quasi-monotonicity of the estimators: there exists $C>0$ independent of $\boldsymbol{y}\in U, \mathcal{T}\in\mathbb{T}$ such that

$$\sum_{T\in\mathcal{T}} \varphi^2_{\boldsymbol{y},\mathcal{T}}(T) \le C \sum_{T\in\mathcal{T}} \varphi^2_{\boldsymbol{y},\mathcal{T}_0}(T) < CK_0 \qquad \text{with } \varphi\in\{\eta,\zeta\}. \tag{48}$$

The axioms (25) and (26) for the indicators $\rho_{\mathcal{T}}(T)$ are verified as in Theorem 3.5, and using that $K_0 < \infty$. Therefore we conclude with [21, Theorem 3.1] the claim, $\sum_{T\in\mathcal{T}_\ell} \rho^2_{\mathcal{T}_\ell}(T) = 2\eta^2_{\mathcal{T}_\ell}\zeta^2_{\mathcal{T}_\ell} \to 0$ as $\ell \to \infty$. □

As termination criterion for the spatial refinement we impose

$$\frac{1}{b^m}\sum_{\boldsymbol{y}\in P_m} |G(u(\boldsymbol{y})) - G(u_{\mathcal{T}}(\boldsymbol{y}))| \lesssim \left(\sum_{T\in\mathcal{T}} \zeta^2_{\mathcal{T}}(T)\right)^{\frac{1}{2}} \left(\sum_{T\in\mathcal{T}} \eta^2_{\mathcal{T}}(T)\right)^{\frac{1}{2}} \leq \varepsilon_F. \tag{49}$$

Convergence of a goal oriented adaptive QMC-FEM Algorithm follows replacing (44) with the latter equation.

3.5 Computational cost

We analyze the computation cost under the working assumption that we have available SOLVE and ESTIMATE modules for the Galerkin formulation (24) that run in $\mathcal{O}(s|\mathcal{T}|)$ operations. On the other hand, MARK and REFINE have typically cost $\mathcal{O}(|\mathcal{T}|)$.

Fix a tolerance $\varepsilon > 0$ and consider Alg. 2. Let $N_\ell := \max_{\boldsymbol{y}\in P_m} |\mathcal{T}_\ell(\boldsymbol{y})|$, where $\mathcal{T}_\ell(\boldsymbol{y})$ is the mesh obtained after ℓ iterations of AFEM for a sample $\boldsymbol{y}$. From Proposition 3.1, we have $\varepsilon \sim \varepsilon_F$ independent of m, that justifies imposing N_ℓ independent of m: then the number of operation required is

$$\begin{aligned}\text{Work(QMC-AFEM)} &= \mathcal{O}\left(\sum_{m=1}^{M(\varepsilon)} b^m \left(\underbrace{s^2 + sm}_{\text{CBC construction (SPOD weights)}} + \underbrace{s\sum_{\ell=0}^{\ell(\varepsilon)} N_\ell}_{\text{AFEM}}\right)\right)\\ &= \mathcal{O}\left(s^2 b^{M(\varepsilon)+1} + sM(\varepsilon)b^{M(\varepsilon)+1} + sb^{M(\varepsilon)+1}\sum_{\ell=0}^{\ell(\varepsilon)} N_\ell\right).\end{aligned}$$

Here, $M(\varepsilon), \ell(\varepsilon) \in \mathbb{N}$ are the maximum number of iterations in the (outer) QMC and (inner) AFEM loop, respectively. In the last step we used $\sum_{m=1}^{M(\varepsilon)} mb^m \sim M(\varepsilon)b^{M(\varepsilon)+1}$. For spaces $V_{\mathcal{T}}$ of piecewise polynomials, we have $N_{\ell(\varepsilon)} \sim \varepsilon^{-d/k}$ for some $k > 0$, which is typically determined by the polynomial degree and the spatial regularity of the data (see e.g. [6]). Moreover, due to Proposition 2.1 we have

$$M(\varepsilon) \sim \log(\varepsilon^{-1}). \tag{50}$$

Similarly, in the case of Alg. 3 we get, with $\ell(0,\varepsilon) := 0$ and $\ell(m,\varepsilon) = \ell(\varepsilon)$ for $m \in \mathbb{N}$,

$$\begin{aligned}\text{Work(QMC-AFEMv2)} &= \mathcal{O}\left(\sum_{m=1}^{M(\varepsilon)} b^m \left(\underbrace{s^2 + sm}_{\text{CBC construction (SPOD weights)}} + \underbrace{s\sum_{\ell=\ell(m-1,\varepsilon)}^{\ell(m,\varepsilon)} N_\ell}_{\text{AFEM solver}} + \underbrace{sb^{\min(m-1,q)}}_{\arg\min d_{\boldsymbol{\gamma}}}\right)\right)\\ &= \mathcal{O}\left(s^2 b^{M(\varepsilon)+1} + sM(\varepsilon)b^{M(\varepsilon)+1} + sb\sum_{\ell=0}^{\ell(\varepsilon)} N_\ell + sb^{M(\varepsilon)+1}N_{\ell(\varepsilon)} + sb^{M(\varepsilon)+1+\min(M(\varepsilon),q)}\right).\end{aligned}$$

If we set $q > M(\varepsilon)$, the asymptotic cost increases quadratically with respect to $b^{M(\varepsilon)}$, and the argmin computation dominates the cost. Therefore, such choice is only possible when the metric

$d_{\boldsymbol{\gamma}}$ is cheap to compute. Another possibility is imposing q such that $b^q \sim N_{\ell(\varepsilon)}$, so that the dominating contribution to the computational cost is due to the AFEM solver.

We finally turn to Algorithm 4. The mesh that meets the FEM tolerance for P_m is denoted by $\mathcal{T}^{(m)} := \mathcal{T}_{\ell(m,\varepsilon)}$ (cf. (42)). Hence we get

$$\begin{aligned}\text{Work(AQMC-FEM)} = \mathcal{O}\left(\sum_{m=1}^{M(\varepsilon)} b^m \left(\underbrace{s^2+sm}_{\text{CBC construction (SPOD weights)}} + \underbrace{s\sum_{\ell=\ell(m-1,\varepsilon)}^{\ell(m,\varepsilon)} N_\ell}_{\text{SOLVE,ESTIMATE}}\right)\right) \\ + \mathcal{O}\left(\sum_{m=1}^{M(\varepsilon)} \underbrace{b^{m-1}\left(1+sN_{\ell(m)}\chi_{\{\mathcal{T}^{(m)}\neq\mathcal{T}^{(m-1)}\}}\right)}_{\text{QMC estimator}} + \sum_{\ell=0}^{\ell(M(\varepsilon),\varepsilon)} \underbrace{N_\ell}_{\text{MARK,REFINE}}\right),\end{aligned}$$

where χ_A denotes the indicator function of a set A. We observe that the extra cost, due to the condition of Algorithm 4 line 13, is asymptotically lower than the cost for SOLVE and ESTIMATE. Thus we conclude

$$\text{Work(AQMC-FEM)} = \mathcal{O}\left(s^2 b^{M(\varepsilon)+1} + sM(\varepsilon)b^{M(\varepsilon)+1} + (1+sb)\sum_{\ell=0}^{\ell(\varepsilon)} N_\ell + sb^{M(\varepsilon)+1}N_{\ell(\varepsilon)}\right).$$

Remark 3.7. *When $\sum_{\ell=0}^{\ell(\varepsilon)} N_\ell \sim N_{\ell(\varepsilon)}$, i.e. the cost of the adaptive loop is dominated by the last iteration (cf. [20]), all three algorithms require (asymptotically) the same computational effort $\mathcal{O}\big(s^2\varepsilon^{-1} + s\varepsilon^{-1-d/k}\big)$.*

Remark 3.8. *Assuming that $\boldsymbol{\beta} \in \ell^p(\mathbb{N})$ for $0 < p < \frac{1}{\alpha}$, for some $\alpha \in \mathbb{N}, \alpha \geq 2$, it is possible to employ higher-order QMC integration of order α [12, 10], and a-posteriori QMC error estimation up to order $\alpha - 1$ [10]. Thus, after slight modifications to the algorithms,* (50) *improves to $M(\varepsilon) \sim \log(\varepsilon^{-\alpha+1})$.*

3.6 Examples

In the present section we illustrate the framework in a selection of model problems.

Parametric diffusion. We consider a parametric stationary diffusion equation: given $\boldsymbol{y} \in U$, find $u(\cdot, \boldsymbol{y})$ such that

$$-\operatorname{div}(a(x,\boldsymbol{y})\nabla u(x,\boldsymbol{y})) = f(x) \quad x \in D, \quad u(\cdot,\boldsymbol{y})\big|_{\partial D} = 0, \tag{51}$$

where $a(\cdot,\boldsymbol{y}) \in W^{1,\infty}(D)$ and $f \in L^2(D)$. We select an affine-parametric diffusion: for $\{\psi_j\}_{j\in\mathbb{N}_0} \in W^{1,\infty}(D)$,

$$a(x,\boldsymbol{y}) = \psi_0(x) + \sum_{j=1}^{s} y_j\psi_j(x). \tag{52}$$

Assume that the $\psi_0 > \psi_{0,\min}$ a.e. in D for a constant $\psi_{0,\min} > 0$ and the sequence $\boldsymbol{\beta}$ given by $\beta_j = \frac{\|\psi_j\|_{L^\infty(D)}}{\psi_{0,\min}}$, $j \geq 1$ satisfies $\|\boldsymbol{\beta}\|_{\ell^1(\mathbb{N})} := \sum_{j\geq 1}\beta_j < 2$ and $\boldsymbol{\beta} \in \ell^p(\mathbb{N})$, for some $p \in (0, \frac{1}{2})$. The weak formulation of equation (51) reads, for all $\boldsymbol{y} \in U$ find $u(\cdot,\boldsymbol{y}) \in V := H_0^1(D)$ such that

$$\mathfrak{a}_{\boldsymbol{y}}(u(\cdot,\boldsymbol{y}),v) := \int_D a(\cdot,\boldsymbol{y})\nabla u(\cdot,\boldsymbol{y})\cdot\nabla v = \int_D fv =: \mathfrak{l}_{\boldsymbol{y}}(v) \quad \forall v \in V. \tag{53}$$

This model problem satisfies (2.1) and the derivative bound (18) with $\kappa = 0$, follows from [8] or Theorem 2.2. AFEM can be performed (with quasi-optimal convergence) for example by 1st order Lagrangian elements, standard residual indicators, Dörfler marking and refinement by newest vertex bisection, as e.g. derived in [6].

For completeness we verify Lipschitz continuity (32) for the model problem: denote $u(\boldsymbol{y}) = u(\cdot, \boldsymbol{y}) \in V$, then affine parametric structure of $a(\cdot, \boldsymbol{y})$ gives

$$\begin{aligned}\psi_{0,\min}\left(1 - \frac{\|\boldsymbol{\beta}\|_{\ell^1(\mathbb{N})}}{2}\right)\|u(\boldsymbol{y}) - u(\boldsymbol{y}')\|_V^2 \leq& \mathfrak{a}_{\boldsymbol{y}}(u(\boldsymbol{y}) - u(\boldsymbol{y}'), u(\boldsymbol{y}) - u(\boldsymbol{y}')) \\ =& \langle f, u(\boldsymbol{y}) - u(\boldsymbol{y}')\rangle - \mathfrak{a}_{\boldsymbol{y}}(u(\boldsymbol{y}'), u(\boldsymbol{y}) - u(\boldsymbol{y}')) \\ =& \langle f, u(\boldsymbol{y}) - u(\boldsymbol{y}')\rangle - \mathfrak{a}_{\boldsymbol{y}'}(u(\boldsymbol{y}'), u(\boldsymbol{y}) - u(\boldsymbol{y}')) \\ &+ \sum_{j\geq 1}(y_j' - y_j)\int_D \psi_j \nabla u(\boldsymbol{y}')\nabla(u(\boldsymbol{y}) - u(\boldsymbol{y}')).\end{aligned}$$

The first two terms cancel since f is independent of $\boldsymbol{y} \in U$. Furthermore,

$$\left|\sum_{j\geq 1}(y_j' - y_j)\int_D \psi_j \nabla u(\boldsymbol{y}')\nabla(u(\boldsymbol{y}) - u(\boldsymbol{y}'))\right| \leq \psi_{0,\min}\|u(\boldsymbol{y}) - u(\boldsymbol{y}')\|_V \|u(\boldsymbol{y})\|_V \sum_{j\geq 1}|y_j' - y_j|\beta_j.$$

Therefore defining $d_{\boldsymbol{\gamma}}(\boldsymbol{y}, \boldsymbol{y}') := \sum_{j\geq 1}|y_j' - y_j|\beta_j$ and $L := \frac{4}{\psi_{0,\min}\left(2 - \|\boldsymbol{\beta}\|_{\ell^1(\mathbb{N})}\right)^2}\|f\|_{V^*}$ we have the claim. The same steps hold for a FE solution $u_{\mathcal{T}}(\boldsymbol{y})$, for any $\mathcal{T} \in \mathbb{T}$. This also implies Assumption 3.1 due to compactness of U.

Shape Uncertainty Quantification for the Poisson equation. Consider the following domain uncertainty problem from [26]: define a family of domains $\{D(\boldsymbol{y}) : \boldsymbol{y} \in U\}$ contained in a hold-all domain $\mathcal{D} := \bigcup D(\boldsymbol{y})$. Given a reference Lipschitz polyhedron $\hat{D} \subset \mathbb{R}^d$ $d \in \{2, 3\}$, we assume that the family is parametrized by a $C^2(\hat{D})$ diffeomorphism $\Psi\colon \hat{D} \times U \to \mathcal{D}$ by the relations $D(\boldsymbol{y}) := \Psi(\hat{D}, \boldsymbol{y})$ and

$$\Psi(x, \boldsymbol{y}) = x + \sum_{j=1}^{s} y_j \psi_j(x), \quad x \in \hat{D}, \boldsymbol{y} \in U \tag{54}$$

for functions $\{\psi_j\}_{j\in\mathbb{N}} \subset W^{1,\infty}(D)$ satisfying $\boldsymbol{\beta} \in \ell^p(\mathbb{N})$, $p \in (0, \frac{1}{2})$ with $\beta_j := \|\psi_j\|_{W^{1,\infty}(\hat{D})}$. For all $\boldsymbol{y} \in U$, let $u(\cdot, \boldsymbol{y}) \in H_0^1(D(\boldsymbol{y}))$ solve the Poisson equation, given a source $f \in C^\infty(\mathcal{D})$ analytic (as in [26, Lemma 5]),

$$-\Delta u(x, \boldsymbol{y}) = f(x) \quad x \in D(\boldsymbol{y}), \quad u\big|_{\partial D(\boldsymbol{y})} = 0. \tag{55}$$

This problem can be recast by a change of variables to the reference domain: for $V = W = H_0^1(\hat{D})$ and for any $\boldsymbol{y} \in U$, we seek $\hat{u}(\cdot, \boldsymbol{y}) := u(\cdot, \boldsymbol{y}) \circ \Psi \in V$ such that (3) holds with

$$\begin{aligned}\mathfrak{l}_{\boldsymbol{y}}(\hat{w}) &:= \int_{\hat{D}} f \circ \Psi(x, \boldsymbol{y})\hat{w}(x)\det(J(x, \boldsymbol{y}))\,\mathrm{d}x, \\ \mathfrak{a}_{\boldsymbol{y}}(\hat{v}, \hat{w}) &:= \int_{\hat{D}} A(x, \boldsymbol{y})\nabla\hat{v}(x)\cdot\nabla\hat{w}(x)\,\mathrm{d}x,\end{aligned} \tag{56}$$

where $A(x, \boldsymbol{y}) := (J^\top(x, \boldsymbol{y})J(x, \boldsymbol{y}))^{-1}\det(J(x, \boldsymbol{y}))$ and $J(x, \boldsymbol{y}) := \nabla_x\Psi(x, \boldsymbol{y})$ is the Jacobian matrix of Ψ. In [26, Theorem 5], the authors provided a derivative bound in the form (18), $\kappa = 0$, for $\hat{u}$. The AFEM modules are analogous to the previous example (but here with parametric matrix valued diffusion coefficient); the applicability of Algorithms 2 and 4 is straightforward.

Linear elasticity of nearly incompressible materials. Robust approximation of linear elasticity in the incompressible limit, (that is Poisson ratio $\nu \in (0, \frac{1}{2})$ approaching $\frac{1}{2}$), was studied in [30, 29] by the following three-field-formulation. Let $E(x, \boldsymbol{y}) = e_0(x) + \sum_{j=1}^{s} y_j e_j(x) \in L^\infty(D)$ be the (affine-parametric) Young modulus, with $0 < e_{0,\min} < e_0(x) < e_{0,\max}$ a.e., for constants $e_{0,\min}, e_{0,\max}$. Define $\varepsilon(v) := \frac{1}{2}[\nabla v + (\nabla v)^\top]$ the strain tensor (for a vector field $v\colon D \to \mathbb{R}^d$) and $f \in L^2(D)^d$. Assume that the boundary conditions are of mixed type Dirichlet-Neumann given respectively on Γ_D, Γ_N, both of positive length, with $\Gamma_D \cup \Gamma_N = \partial D$, $\Gamma_D \cap \Gamma_N = \emptyset$. Introducing the extra variable $\tilde{p}(x, \boldsymbol{y}) = p(x, \boldsymbol{y})/E(x, \boldsymbol{y})$, where $p(x, \boldsymbol{y})$ is the (parameter-dependent) Herrmann pressure, we can write the linear elasticity equations as

$$\begin{cases} -\mathrm{div}\left(\frac{E(x,\boldsymbol{y})}{(1+\nu)}\varepsilon(u(x,\boldsymbol{y}))\right) + \nabla p(x,\boldsymbol{y}) = f(x) & x \in D, \boldsymbol{y} \in U \\ \mathrm{div}(u(x,\boldsymbol{y})) + c^{-1}\tilde{p}(x,\boldsymbol{y}) = 0 & x \in D, \boldsymbol{y} \in U \\ c^{-1}p(x,\boldsymbol{y}) - c^{-1}E(x,\boldsymbol{y})\tilde{p}(x,\boldsymbol{y}) = 0 & x \in D, \boldsymbol{y} \in U \\ u(x,\boldsymbol{y}) = 0 & x \in \Gamma_D, \boldsymbol{y} \in U \\ \left(\frac{E(x,\boldsymbol{y})}{1+\nu}\varepsilon(u(x,\boldsymbol{y})) - p(x,\boldsymbol{y})I\right)\boldsymbol{n}(x) = 0 & x \in \Gamma_N, \boldsymbol{y} \in U \end{cases} \tag{57}$$

for an absolute constant $c > 0$ only dependent on $\nu \in (0, \frac{1}{2})$. The weak formulation is: for all $\boldsymbol{y} \in U$, find $(u(\cdot,\boldsymbol{y}), p(\cdot,\boldsymbol{y}), \tilde{p}(\cdot,\boldsymbol{y})) \in V := W := H^1_{\Gamma_D}(D)^d \times L^2(D) \times L^2(D)$ such that (3) holds for the bilinear form

$$\begin{aligned} \mathfrak{a}_{\boldsymbol{y}}((v,g,\tilde{g}),(w,q,\tilde{q})) = \int_D E(\cdot,\boldsymbol{y})\varepsilon(v) : \varepsilon(w) - \int_D g\,\mathrm{div} w - \int_D q\,\mathrm{div} v - c^{-1}\int_D \tilde{g}\,q \\ - c^{-1}\int_D g\,\tilde{q} + c^{-1}\int_D E(\cdot,\boldsymbol{y})\tilde{g}\,\tilde{q} \qquad \forall (v,g,\tilde{g}),(w,q,\tilde{q}) \in V \end{aligned} \tag{58}$$

and $\mathfrak{l}_{\boldsymbol{y}}((w,q,\tilde{q})) = \int_D fw$. We equip V with the norm (related to [30, Equation (2.21)], but without integrating out the parameter space)

$$\|(w,q,\tilde{q})\|_V^2 := \frac{1}{1+\nu}\|\nabla w\|_{L^2(D)}^2 + (1+\nu+c^{-1})\|q\|_{L^2(D)}^2 + c^{-1}\|\tilde{q}\|_{L^2(D)}^2 . \tag{59}$$

The main motivation to introduce the three-field formulation is that E only appears in the numerator, and it is in particular affine-parametric. We verify the criteria of Theorem 2.2. The nominal operator $A_0\colon V \to V^*$, induced by $\mathfrak{a}_{\boldsymbol{0}}$ is linear and boundedly invertible by [30, Theorem 2.4], with norm $\|A_0^{-1}\| \le \frac{K_0(1+\nu)^{1/2}}{e_{0,\min}}$, for a constant $K_0 > 0$ dependent on D and $\|e_0\|_{L^\infty(D)}$. Moreover, the fluctuations $\{A_j\}_j$ in the notation of Theorem 2.2 satisfy, for all triples $(v,g,\tilde{g}),(w,q,\tilde{q}) \in V$

$$\langle A_j(v,g,\tilde{g}),(w,q,\tilde{q})\rangle = \frac{1}{1+\nu}\int_D e_j\varepsilon(v) : \varepsilon(w) + c^{-1}\int_D e_j\tilde{g}\tilde{q} \le \|e_j\|_{L^\infty(D)}\|(v,g,\tilde{g})\|_V\|(w,q,\tilde{q})\|_V , \tag{60}$$

that is $\|A_j\| \le \|e_j\|_{L^\infty(D)}$. Therefore, to obtain (18) we assume $\|\boldsymbol{\beta}\|_{\ell^1(\mathbb{N})} < 2, \boldsymbol{\beta} \in \ell^p(\mathbb{N})$ for some $p \in (0, \frac{1}{2})$, where

$$\beta_j := \frac{K_0(1+\nu)^{1/2}}{e_{0,\min}}\|e_j\|_{L^\infty(D)} . \tag{61}$$

With these choices, this formulation fits (2.1) by [30, Lemma 2.3] and Theorem 2.2; hence the problem is well-posed and (18) holds with $p < \frac{1}{2}$. Lipschitz continuity (32) follows as in the first example. Any converging AFEM solver (not necessarily conforming) for (57) ensures that Algorithms 2, 3 are applicable. In particular, the reliability and efficiency of [29, Theorem 5.1]

(applied with $\Gamma = \{0\}$, in the notation there, i.e. for a deterministic equation) and suitable inf-sup stable discretization spaces as $V_{\mathcal{T}} := (\mathbb{Q}_2(\mathcal{T}))^d \times \mathbb{Q}_1(\mathcal{T}) \times \mathbb{Q}_1(\mathcal{T})$ satisfying in (23) give, for all $\boldsymbol{y} \in U$, an AFEM algorithm based on hierarchical spatial refinement. Here, $\mathbb{Q}_2(\mathcal{T})$ denotes the space of continuous piecewise biquadratic functions on $\mathcal{T}$ and $\mathbb{Q}_1(\mathcal{T})$ the continuous piecewise bilinear functions, on quadrilatelar meshes.

4 Numerical experiments

We consider the model problem (51) on a polygon $D \subseteq \mathbb{R}^2$, for the solution of (24) we employ Lagrangian $\mathbb{P}_1$-FEM on regular triangulations $\mathcal{T} \in \mathbb{T}$ of D. AFEM is driven by the residual indicators from, e.g. [33, Section 1.4] and the Dörfler MARK strategy with marking parameter $\theta \in (0,1)$, where larger θ corresponds to more aggressive refinement. In all the computations, we select $\theta = 0.25$. The REFINE module is the Newest Vertex Bisection as from the MATLAB implementation in [19], that gives uniform shape regularity of $\mathbb{T}$. We run on a machine equipped with Intel(R) Core(TM) i7-10510U CPU @ 1.80GHz (OctaCore) and MATLAB R2019a.

4.1 Convex domain

As a first example we select an affine-parametric diffusion, (51), (52) with $\psi_0 \equiv 1$ and for $j \geq 1$,

$$\psi_j(x) := \frac{1}{(k_{j,1}^2 + k_{j,2}^2)^\eta} \sin(k_{j,1}\pi x_1) \sin(k_{j,2}\pi x_2), \tag{62}$$

where the pairs $(k_{j,1}, k_{j,2}) \in \mathbb{N}^2$ are defined by the ordering of $\mathbb{N}^2$ such that for $j \in \mathbb{N}$, $k_{j,1}^2 + k_{j,2}^2 \leq k_{j+1,1}^2 + k_{j+1,2}^2$, and the ordering is arbitrary when equality holds. With this choice $\beta_j \sim j^{-\eta}$, that is $\boldsymbol{\beta} \in \ell^p(\mathbb{N}), p > \frac{1}{\eta}$. Let $D = (0,1)^2$, $f(x) = e^{-|x|^2}$, $s = 32$, $\eta = 2.1$ and goal functional $G(v) = 4\int_{(0,\frac{1}{2})^2} v$. In this case we expect the mesh to be approximately uniformly refined by AFEM, starting from a structured mesh $\mathcal{T}_0$ with 128 elements, since $u(\cdot, \boldsymbol{y}) \in H^2(D)$ for all $\boldsymbol{y} \in U$, due to convexity of the domain and smoothness of the data. We also use the stopping criteria (38), (49) for the FEM error exploiting symmetry of the stiffness matrix – see Remark 3.3 – thus avoiding excessive spatial refinement. We compare the algorithms in Figure 1, for various tolerances $\varepsilon_F = \varepsilon_Q$. For convenience of the reader, we compute a reference value with $|\mathcal{T}| \approx 10^5$ many $\mathbb{P}_2$ (i.e. quadratic) elements obtained by uniform refinement of $\mathcal{T}_0$ and $|P_m| = 2^m$, $m = 8$ samples, obtaining $I(G(u)) \approx 0.024411631814585$. Since we observe that the cost of computing $d_{\boldsymbol{\gamma}}(\boldsymbol{y}, \boldsymbol{y}') := \sum_{j\geq 1} |y_j' - y_j| \beta_j$ is negligible, we formally set $q = \infty$ in Algorithm 3. As predicted, they all produce outputs well within the tolerance $\varepsilon = 2\varepsilon_F$. We also observe that, for the finest tolerance ($\varepsilon_F = 10^{-5}$), all 3 algorithms produce meshes with $\approx 2 \cdot 10^5$ degrees of freedom and they stop at $m = 7$, that is $N = 128$ samples are sufficient to meet the tolerance. In terms of computing time, Algorithm 2 lags behind the other 2 algorithms, which in turn offer comparable performance. The rates in Figure 1 (right) are estimated excluding the coarsest tolerance ($\varepsilon_F = 10^{-3}$) and favor slightly Algorithm 4.

4.2 L-shape domain

We again pick the affine parametric diffusion in sin expansion of (62). Let $D = (-1,1)^2 \setminus [0,1) \times (-1,0]$, $f(x) = e^{-2|x+(1,0)|^2}$ be a localized source at $(-1,0)$. Assume homogeneous Neumann boundary conditions at $\Gamma_N = \{1\} \times (0,1) \cup (0,1] \times \{0\}$ and homogeneous Dirichlet at $\Gamma_D = \partial D \setminus \Gamma_N$. As Goal functional we pick $G(v) = \int_{D \cap B_{1/2}} v$, where B_r denotes the ball centered at the origin with radius $r > 0$. Again we choose $s = 32, \eta = 2.1$. We start from a uniform mesh $\mathcal{T}_0$ with $|\mathcal{T}_0| = 192$.

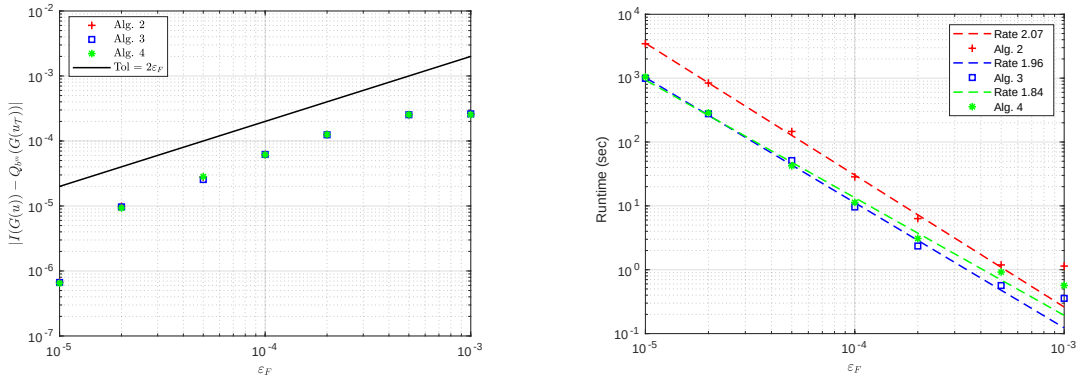

Figure 1: Errors committed by the 3 different algorithms for varying tolerances (left). Runtimes (in seconds) , averaged over 2 runs and estimated rate (right).

The evolution of the QMC and FEM error estimators run by AQMC-FEM for $\varepsilon_F = \varepsilon_Q = 5 \cdot 10^{-6}$ are displayed in Figure 2. Note that no QMC estimator is computed until the FEM tolerance is reached for $m = 2$. We measure the computational effort of each iteration of the algorithm AQMC-FEM (indexed by a pair $(m, \ell) \in \mathbb{N}^2$, corresponding to QMC and FEM refinement level, respectively) by

$$W(m, \ell) = |P_m||\mathcal{T}_\ell|. \tag{63}$$

This is proportional to the cost of the iteration (m, ℓ) of the SOLVE module, assuming that s is fixed and that a FEM solver that performs linearly with respect to $|\mathcal{T}_\ell|$ is available. We also show the mesh generated by AQMC-FEM for $\varepsilon_F = \varepsilon_Q = 10^{-3}$; as expected, it is strongly graded near the source and towards the corner of the domain, where a singularity of the solution occurs.

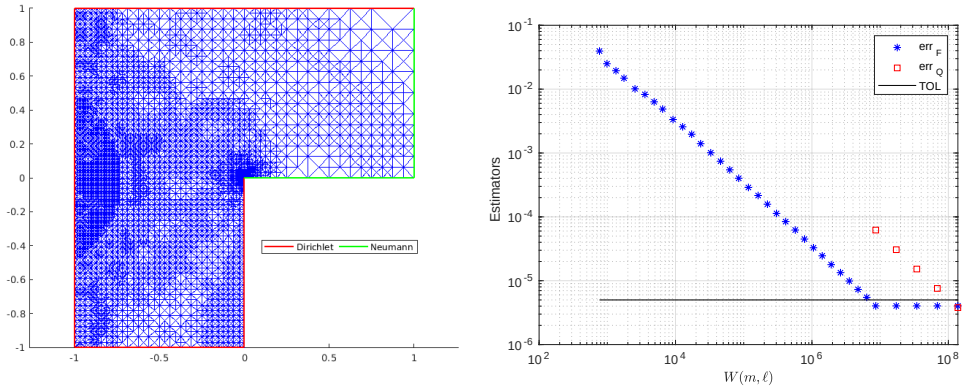

Figure 2: Mesh produced by AQMC-FEM for $\varepsilon_F = \varepsilon_Q = 10^{-3}$ (left) and decay of FEM and QMC estimators (asterisk and square, respectively) against $W(m, \ell)$, for $\varepsilon_F = \varepsilon_Q = \text{TOL} = 5 \cdot 10^{-6}$ (right).

5 Conclusions

We have presented a family of adaptive discretization methods that combine FEM error estimation in the spatial domain and deterministic Polynomial lattice rules in the parameter box $U = [-\frac{1}{2}, \frac{1}{2}]^s$. We recalled possible criteria to verify convergence of the AFEM iteration and to enable Quasi-Monte Carlo a-posteriori estimation. The convergence of the parametric estimator is free of the curse of dimensionality, allowing for arbitrary $s \in \mathbb{N}$, also in practical examples, under the assumption of quantified decay of the derivatives (19) or (18). Moreover, we stress that the parametric error is estimated without resorting to the specific problem formulation. This is the main feature that improves upon existing methods based on stochastic Galerkin or sparse grids [2, 29, 14, 18], in terms of versatility.

Thus, we expect our algorithms to be applicable in a wide range of problems, including, but not restricted to, those in the framework of Section 2, provided that a converging AFEM algorithm is available for the corresponding non-parametric equation. In particular, we mention parabolic equations (cp. a posteriori indicators in [33, Chapter 6]) and certain non-linear PDEs meeting the criteria exposed in [7], stationary Stokes (cp. [5, Section 6.2-6.3]) and Navier-Stokes (cp. [33, Chapter 5]) equations on uncertain domains [9] and elliptic eigenvalue problems [25], [5, Section 10.3].

Acknowledgement. The author would like to express his gratitude to Prof. Christoph Schwab for the useful input and advice throughout the redaction of this work.

References

[1] Roland Becker, Elodie Estecahandy, and David Trujillo. Weighted marking for goal-oriented adaptive finite element methods. *SIAM J. Numer. Anal.*, 49(6):2451–2469, 2011.

[2] Alex Bespalov, Dirk Praetorius, Leonardo Rocchi, and Michele Ruggeri. Convergence of adaptive stochastic Galerkin FEM. *SIAM J. Numer. Anal.*, 57(5):2359–2382, 2019.

[3] Alex Bespalov, Dirk Praetorius, Leonardo Rocchi, and Michele Ruggeri. Goal-oriented error estimation and adaptivity for elliptic PDEs with parametric or uncertain inputs. *Comput. Methods Appl. Mech. Engrg.*, 345:951–982, 2019.

[4] Franco Brezzi and Michel Fortin. *Mixed and hybrid finite element methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York, 1991.

[5] C. Carstensen, M. Feischl, M. Page, and D. Praetorius. Axioms of adaptivity. *Comput. Math. Appl.*, 67(6):1195–1253, 2014.

[6] J. Manuel Cascon, Christian Kreuzer, Ricardo H. Nochetto, and Kunibert G. Siebert. Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, 46(5):2524–2550, 2008.

[7] Abdellah Chkifa, Albert Cohen, and Christoph Schwab. Breaking the curse of dimensionality in sparse polynomial approximation of parametric PDEs. *J. Math. Pures Appl. (9)*, 103(2):400–428, 2015.

[8] 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.

[9] Albert Cohen, Christoph Schwab, and Jakob Zech. Shape holomorphy of the stationary Navier-Stokes equations. *SIAM J. Math. Anal.*, 50(2):1720–1752, 2018.

[10] J. Dick, M. Longo, and Ch. Schwab. Extrapolated lattice rule integration in computational uncertainty quantification. Technical Report 2020-29, Seminar for Applied Mathematics, ETH Zürich, Switzerland, 2020.

[11] Josef Dick, Takashi Goda, and Takehito Yoshiki. Richardson extrapolation of polynomial lattice rules. *SIAM J. Numer. Anal.*, 57(1):44–69, 2019.

[12] 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.

[13] Josef Dick, Quoc T. Le Gia, and Christoph Schwab. Higher order quasi-Monte Carlo integration for holomorphic, parametric operator equations. *SIAM/ASA J. Uncertain. Quantif.*, 4(1):48–79, 2016.

[14] Martin Eigel, Oliver Ernst, Björn Sprungk, and Lorenzo Tamellini. On the convergence of adaptive stochastic collocation for elliptic partial differential equations with affine diffusion, 2020, arXiv:2008.07186.

[15] Martin Eigel, Claude Jeffrey Gittelson, Christoph Schwab, and Elmar Zander. Adaptive stochastic Galerkin FEM. *Comput. Methods Appl. Mech. Engrg.*, 270:247–269, 2014.

[16] Martin Eigel, Claude Jeffrey Gittelson, 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.

[17] Michael Feischl, Dirk Praetorius, and Kristoffer G. van der Zee. An abstract analysis of optimal goal-oriented adaptivity. *SIAM J. Numer. Anal.*, 54(3):1423–1448, 2016.

[18] Michael Feischl and Andrea Scaglioni. Convergence of adaptive stochastic collocation with finite elements, 2020, arXiv:2008.12591.

[19] Stefan Funken, Dirk Praetorius, and Philipp Wissgott. Efficient implementation of adaptive P1-FEM in Matlab. *Comput. Methods Appl. Math.*, 11(4):460–490, 2011.

[20] Gregor Gantner, Alexander Haberl, Dirk Praetorius, and Stefan Schimanko. Rate optimality of adaptive finite element methods with respect to overall computational costs. *Math. Comp.*, 90(331):2011–2040, 2021.

[21] Gregor Gantner and Dirk Praetorius. Plain convergence of adaptive algorithms without exploiting reliability and efficiency. *IMA Journal of Numerical Analysis*, 03 2021, https://academic.oup.com/imajna/advance-article-pdf/doi/10.1093/imanum/drab010/36458721/drab010.pdf. drab010.

[22] Robert N. Gantner. Dimension truncation in QMC for affine-parametric operator equations. In *Monte Carlo and quasi–Monte Carlo methods*, volume 241 of *Springer Proc. Math. Stat.*, pages 249–264. Springer, Cham, 2018.

[23] Robert N. Gantner, Lukas Herrmann, and Christoph Schwab. Quasi-Monte Carlo integration for affine-parametric, elliptic PDEs: local supports and product weights. *SIAM J. Numer. Anal.*, 56(1):111–135, 2018.

[24] Robert N. Gantner and Christoph Schwab. Computational higher order quasi-Monte Carlo integration. In *Monte Carlo and quasi-Monte Carlo methods*, volume 163 of *Springer Proc. Math. Stat.*, pages 271–288. Springer, Cham, 2016.

[25] A. D. Gilbert, I. G. Graham, F. Y. Kuo, R. Scheichl, and I. H. Sloan. Analysis of quasi-Monte Carlo methods for elliptic eigenvalue problems with stochastic coefficients. *Numer. Math.*, 142(4):863–915, 2019.

[26] H. Harbrecht, M. Peters, and M. Siebenmorgen. Analysis of the domain mapping method for elliptic diffusion problems on random domains. *Numer. Math.*, 134(4):823–856, 2016.

[27] Fred J. Hickernell and Lluís Antoni Jiménez Rugama. Reliable adaptive cubature using digital sequences. In *Monte Carlo and quasi-Monte Carlo methods*, volume 163 of *Springer Proc. Math. Stat.*, pages 367–383. Springer, Cham, 2016.

[28] Lluís Antoni Jiménez Rugama and Fred J. Hickernell. Adaptive multidimensional integration based on rank-1 lattices. In *Monte Carlo and quasi-Monte Carlo methods*, volume 163 of *Springer Proc. Math. Stat.*, pages 407–422. Springer, Cham, 2016.

[29] Arbaz Khan, Catherine E. Powell, and David J. Silvester. Robust a posteriori error estimators for mixed approximation of nearly incompressible elasticity. *Internat. J. Numer. Methods Engrg.*, 119(1):18–37, 2019.

[30] Arbaz Khan, Catherine E. Powell, and David J. Silvester. Robust preconditioning for stochastic Galerkin formulations of parameter-dependent nearly incompressible elasticity equations. *SIAM J. Sci. Comput.*, 41(1):A402–A421, 2019.

[31] Pedro Morin, Kunibert G. Siebert, and Andreas Veeser. A basic convergence result for conforming adaptive finite elements. *Math. Models Methods Appl. Sci.*, 18(5):707–737, 2008.

[32] Christoph Schwab. QMC Galerkin discretization of parametric operator equations. In *Monte Carlo and quasi-Monte Carlo methods 2012*, volume 65 of *Springer Proc. Math. Stat.*, pages 613–629. Springer, Heidelberg, 2013.

[33] Rüdiger Verfürth. *A posteriori error estimation techniques for finite element methods*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2013.