Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Seminar for Applied Mathematics

A-posteriori QMC-FEM error estimation for Bayesian inversion and optimal control with entropic risk measure

M. Longo and Ch. Schwab and A. Stein

Research Report No. 2023-04
January 2023

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

A-posteriori QMC-FEM error estimation for Bayesian inversion and optimal control with entropic risk measure

Marcello Longo, Christoph Schwab, Andreas Stein

January 9, 2023

Abstract

We propose a novel a-posteriori error estimation technique where the target quantities of interest are ratios of high-dimensional integrals, as occur e.g. in PDE constrained Bayesian inversion and PDE constrained optimal control subject to an entropic risk measure. We consider in particular parametric, elliptic PDEs with affine-parametric diffusion coefficient, on high-dimensional parameter spaces. We combine our recent a-posteriori Quasi-Monte Carlo (QMC) error analysis, with Finite Element a-posteriori error estimation. The proposed approach yields a computable a-posteriori estimator which is reliable, up to higher order terms. The estimator's reliability is uniform with respect to the PDE discretization, and robust with respect to the parametric dimension of the uncertain PDE input.

1 Introduction

The efficient numerical approximation of high-dimensional, parametric partial differential equations (PDEs for short) received increasing attention during the past years. In this work, we address two classes of high-dimensional numerical integration problems which arise in connection with data assimilation and PDE constrained optimization. We illustrate the abstract concepts for a parametric, linear elliptic PDE. The first class is the so-called *Bayesian inverse problem* (BIP). There, we are interested in the posterior expectation of a (linear) functional of the solution u of a parametric PDE, conditional on observation data subject to additive, centered Gaussian observation noise [18, 19]. See also [10]. A second problem class is PDE-constrained optimization. Specifically, the optimal control problem (OCP) of parametric PDEs under an *entropic risk* measure [7], where the state variable satisfies a parametric PDE constraint [9].

When numerically approximating solutions of a BIP or an OCP, it is essential to quantify the error due to the numerical discretization in order to meet a prescribed numerical tolerance without wasting computational resources. As solving PDEs exactly is in general not possible, discretizations such as Finite Element Methods (FEM) must be used instead. Additionally, the parametric uncertainty in the forward PDE model is passed on to the solution, and this must be taken into account both in the computation and in the error estimation. This justifies the need for an *a-posteriori* error analysis.

Assuming that the uncertain PDE coefficients can be described by means of a parameter vector $\boldsymbol{y} \in U := \left[-\frac{1}{2}, \frac{1}{2}\right]^s$ where $s \in \mathbb{N}$, $s \gg 1$, e.g. (6) below, we can employ suitable quasi-Monte Carlo (QMC) rules to approximate integrals over U. Here, we select extrapolated polynomial lattice (EPL) rules as first introduced in [3,5]. This choice is motivated by the deterministic nature or their quadrature nodes P_m, $|P_m| = b^m$ for some prime b [15], and good convergence properties under quantified parametric regularity of the integrand functions with respect to $\boldsymbol{y} \in U$, uniformly in the dimension s [3]. Moreover, it was shown in [5,14] that under assumptions, EPL quadratures allow for computable a-posteriori quadrature error estimators that are asymptotically exact as $m \to \infty$, with dimension robust ratio between estimated and actual quadrature error.

Both, BIP and OCP problems for PDEs with parametric input take the form

$$\frac{Z'}{Z} \in \mathcal{Y}, \quad \text{where } Z = \int_U \Theta(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}, \quad Z' = \int_U \Theta'(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}, \tag{1}$$

for some suitable integrable functions $\Theta\colon U \to \mathbb{R}$ and $\Theta'\colon U \to \mathcal{Y}$, where $\mathcal{Y}$ is a separable Hilbert space and Z, Z' are Bochner integrals with respect to a product measure $\mathrm{d}\boldsymbol{y}$ on the possibly high-dimensional parameter space U. In particular, we have $\mathcal{Y} = \mathbb{R}$ for the BIP case and $\mathcal{Y} \in L^2(D)$ for the OCP case, with D being the physical domain of the considered PDE. Approximating the high-dimensional integrals with averages over polynomial lattices $P_m \subset U$ yields a first approximation

$$\frac{Z'_m}{Z_m} \in \mathcal{Y}, \quad \text{where } Z_m = \frac{1}{b^m} \sum_{\boldsymbol{y} \in P_m} \Theta(\boldsymbol{y}), \quad Z'_m = \frac{1}{b^m} \sum_{\boldsymbol{y} \in P_m} \Theta'(\boldsymbol{y}). \tag{2}$$

Then, since the integrands Θ, Θ' depend on the solution of a $\boldsymbol{y}$-parametric PDE, we discretize the parametric PDEs for $\boldsymbol{y} \in P_m$, resulting in computable, parametric integrand functions $\Theta_h(\boldsymbol{y}), \Theta'_h(\boldsymbol{y})$ and in the computable estimates

$$\frac{Z'_{m,h}}{Z_{m,h}} \in \mathcal{Y}, \quad \text{where } Z_{m,h} = \frac{1}{b^m} \sum_{\boldsymbol{y} \in P_m} \Theta_h(\boldsymbol{y}), \quad Z'_{m,h} = \frac{1}{b^m} \sum_{\boldsymbol{y} \in P_m} \Theta'_h(\boldsymbol{y}). \tag{3}$$

Here, the parameter $h > 0$ denotes the meshwidth of conforming Lagrangian Finite Element discretizations. We present a *computable a-posteriori estimator for the combined Finite Element discretization and quadrature error*

$$\mathrm{err} = \left\| \frac{Z'}{Z} - \frac{Z'_{m,h}}{Z_{m,h}} \right\|_{\mathcal{Y}}. \tag{4}$$

In the rest of this section we introduce the setting and we describe the two problems of interest, namely the BIP and the OCP with entropic risk measure. Section 2 and Section 3 are devoted to the QMC and the FEM a-posteriori error analysis, respectively, and these results will be combined in Section 4. We present numerical experiments in Section 5 and summary and conclusions in Section 6.

1.1 Affine-Parametric Forward PDE

For brevity of presentation, we consider a model, linear elliptic PDE with homogeneous Dirichlet boundary conditions. Given a bounded polygon $D \subseteq \mathbb{R}^2$

and a parameter sequence $\boldsymbol{y} \in U$, $s \in \mathbb{N}$, consider the following parametric, linear, second order elliptic PDE in variational form: find $u(\cdot, \boldsymbol{y}) \in \mathcal{X} = H_0^1(D)$ such that

$$\int_D a(x, \boldsymbol{y}) \nabla_x u(x, \boldsymbol{y}) \cdot \nabla_x v(x) \, \mathrm{d}x = \int_D f(x) v(x) \, \mathrm{d}x \qquad \forall v \in \mathcal{X}. \tag{5}$$

We assume that a is affine-parametric, namely that we are given a family of functions $\{\psi_j\}_{j \in \mathbb{N}_0} \subseteq L^\infty(D)$ such that, with $\operatorname{essinf} \psi_0 > \kappa > 0$ and $b_j := \frac{1}{\kappa} \|\psi_j\|_{L^\infty(D)}$, we have

$$a(x, \boldsymbol{y}) = \psi_0(x) + \sum_{j=1}^{s} y_j \psi_j, \qquad \sum_{j \geq 1} b_j < 2. \tag{6}$$

Then $\operatorname{essinf} a(\cdot, \boldsymbol{y}) > \operatorname{essinf} \psi_0 - \kappa > 0$ for all $\boldsymbol{y} \in U$. By the Lax-Milgram lemma, the parametric weak solution $u(\cdot, \boldsymbol{y}) \in \mathcal{X}$ is well defined for any $f \in \mathcal{X}^* = H^{-1}(D)$, where * denotes the topological dual. To justify the a-posteriori QMC error analysis of Section 2, we will additionally require the summability

$$\boldsymbol{b} = (b_j)_{j \geq 1} \in \ell^p(\mathbb{N}) \quad \text{ for some } p \in (0, 1/2]. \tag{7}$$

For the FEM approximation, we consider conforming subspaces[1] $\mathcal{X}_h \subseteq \mathcal{X}$ $h \in H \subseteq (0, \infty)$, $\dim(\mathcal{X}_h) < \infty$, that are linked to shape-regular, simplicial partitions $\{\mathcal{T}_h\}_{h \in H}$ of D [6, Section 8]. Assume that the resulting spaces are nested and conforming, that is $\mathcal{X}_h \subseteq \mathcal{X}_{h'}$ for any $h, h' \in H, h > h'$ and that H accumulates at 0. We construct the Galerkin discretizations $u_h(\boldsymbol{y}) \in \mathcal{X}_h$ of (5), by solving

$$\int_D a(x, \boldsymbol{y}) \nabla_x u_h(x, \boldsymbol{y}) \cdot \nabla_x v(x) \, \mathrm{d}x = \int_D f(x) v(x) \, \mathrm{d}x \qquad \forall v \in \mathcal{X}_h. \tag{8}$$

To simplify notation, we write $a(\boldsymbol{y}) = a(\cdot, \boldsymbol{y})$ and $u(\boldsymbol{y}) = u(\cdot, \boldsymbol{y})$ and we omit the variable x for $\nabla_x = \nabla$ and $\operatorname{div}_x = \operatorname{div}$.

1.2 Bayesian inverse problem (BIP)

Let $X = \{a \in L^\infty(D) : \operatorname{essinf} a > 0\}$ and fix $f \in \mathcal{X}^*$. Then, we can define the data-to-solution map $\mathcal{S} \colon X \to \mathcal{X}$ for the forward problem (5). We also define the observation functional $\mathcal{O} \in (\mathcal{X}^*)^K$, with a finite number $K \in \mathbb{N}$ of observations (e.g. representing sensors), and a goal functional (also called quantity of interest) $G \in \mathcal{X}^*$. We define the prior measure π_0 to be the uniform distribution on U.

The observations $\mathcal{O}(\mathcal{S}(a))$ are assumed to be additionally subject to additive observation noise η, which we assume to be centered Gaussian, i.e., $\eta \sim \mathcal{N}(0, \Gamma)$ for some known, nondegenerate covariance matrix $\Gamma \in \mathbb{R}^{K \times K}$. In other words, we assume given noisy observation data $\delta \in \mathbb{R}^K$ modeled as

$$\delta = \mathcal{O}(\mathcal{S}(a)) + \eta \in L^2_\Gamma(\mathbb{R}^K). \tag{9}$$

We consider the Bayesian inverse problem of recovering the expected value of $G(u)$, given observation data δ, that is $\mathbb{E}^{\pi_0}[G(u)|\delta]$. By Bayes' theorem [19], the

[1] In practice, h either parametrizes the local mesh-size $\max_{T \in \mathcal{T}_h} |T|^{1/2}$, for quasi-uniform collections of partitions, or it relates to the refinement level in case of adaptive refinement [16].

posterior distribution π^δ of $\boldsymbol{y}|\delta$ is absolutely continuous with respect to π_0 and its Radon-Nikodym derivative with respect to the prior π_0 is

$$\frac{\mathrm{d}\pi^\delta}{\mathrm{d}\pi_0}(\boldsymbol{y}) = \frac{\Theta(\boldsymbol{y})}{Z}, \tag{10}$$

where $\Theta(\boldsymbol{y}) := \exp(-\frac{1}{2}|\delta - \mathcal{O}(\mathcal{S}(a(\boldsymbol{y})))|_\Gamma^2) = \exp(-\frac{1}{2}|\delta - \mathcal{O}(u(\boldsymbol{y}))|_\Gamma^2)$ denotes the likelihood with the observation noise covariance-weighted, data-to-observation misfit, where $|x|_\Gamma^2 := x^\top \Gamma^{-1} x$ and Z is defined in (1). As $\Theta(\boldsymbol{y}) > 0$ for all $\boldsymbol{y} \in U$, $Z > 0$. In the present setting, Bayesian inversion amounts to the numerical evaluation of the posterior mean

$$\mathbb{E}^{\pi^\delta}[G(u)] = \frac{1}{Z}\int_U G(u(\boldsymbol{y}))\Theta(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}.$$

This is (1) upon setting $\Theta'(\boldsymbol{y}) := G(u(\boldsymbol{y}))\Theta(\boldsymbol{y})$ and $\mathcal{Y} = \mathbb{R}$. Define the FE solution operator $\mathcal{S}_h\colon X \to \mathcal{X}_h$ as the mapping $\mathcal{S}_h a(\boldsymbol{y}) = u_h(\boldsymbol{y})$ via (8). The FE approximations of Θ, Θ' used in (3) are then $\Theta_h = \exp(-\frac{1}{2}|\delta - \mathcal{O}(u_h(\boldsymbol{y}))|_\Gamma^2)$ and $\Theta'_h = G(u_h(\boldsymbol{y}))\Theta_h(\boldsymbol{y})$, respectively.

1.3 Optimal control with entropic risk measure (OCP)

Let $\mathcal{Y} = L^2(D)$, assume a parameter independent target state $\hat{u} \in \mathcal{Y}$ and a nonempty, closed and convex set $X \subseteq \mathcal{Y}$ of admissible controls. Throughout the rest of the paper, we identify $\mathcal{Y}$ with its dual via Riesz representation and write $\langle\cdot,\cdot\rangle$ for the inner product on $\mathcal{Y}$. Once the affine parametric diffusion coefficient $a(\boldsymbol{y})$ is fixed, (5) defines a linear solution operator $\mathcal{L}^{\boldsymbol{y}}\colon \mathcal{Y} \to \mathcal{Y}$ by $\mathcal{L}^{\boldsymbol{y}} f = \iota \circ u(\boldsymbol{y})$ for all $f \in \mathcal{Y}$, where ι denotes the continuous embedding $\mathcal{X} \subset \mathcal{Y}$. In particular, we view $u(\boldsymbol{y})$ as a function of the right-hand side f of (5). For a function $\Phi\colon U \to \mathbb{R}$ and some $\theta \in (0,\infty)$, the entropic risk measure [12] is defined by

$$\mathcal{R}(\Phi) = \frac{1}{\theta}\log\left(\int_U \exp(\theta\Phi(\boldsymbol{y}))\,\mathrm{d}\boldsymbol{y}\right). \tag{11}$$

The entropic risk is especially relevant when favoring a risk averse behavior [7]. We consider the following minimization problem, for fixed constants $\alpha_1, \alpha_2 > 0$

$$f^* := \operatorname*{argmin}_{f\in X} J(f), \qquad J(f) := \mathcal{R}(\tfrac{\alpha_1}{2}\left\|\mathcal{L}^{\boldsymbol{y}} f - \hat{u}\right\|_{\mathcal{Y}}^2) + \tfrac{\alpha_2}{2}\left\|f\right\|_{\mathcal{Y}}^2. \tag{12}$$

Due to convexity of $\mathcal{R}$ and $\alpha_2 > 0$, the functional J is strongly convex so that (12) is a well-posed minimization problem [9, 12].

Define the shorthand notation $\Phi_f(\boldsymbol{y}) = \frac{\alpha_1}{2}\left\|\mathcal{L}^{\boldsymbol{y}} f - \hat{u}\right\|_{\mathcal{Y}}^2$ and the adjoint state given by $q_f(\boldsymbol{y}) = \alpha_1 \mathcal{L}^{\boldsymbol{y}}(\mathcal{L}^{\boldsymbol{y}} f - \hat{u}) \in \mathcal{Y}$. Under the above conditions on X, (12) is equivalent to the inequality $\langle J'(f^*), f - f^*\rangle \geq 0$ for all $f \in X$, where in analogy with [9, Lemma 3.6] the Fréchet derivative $J'(f) \in \mathcal{Y}$ of J at $f \in X$ is

$$J'(f) = \frac{1}{\int_U \exp(\theta\Phi_f(\boldsymbol{y}))\,\mathrm{d}\boldsymbol{y}}\int_U \exp(\theta\Phi_f(\boldsymbol{y}))q_f(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} + \alpha_2 f. \tag{13}$$

Next, we replace in (12) the integral over U by QMC rules and the exact solution operator $\mathcal{L}^{\boldsymbol{y}}$ by the Galerkin solution $\mathcal{L}_h^{\boldsymbol{y}}\colon \mathcal{Y} \to \mathcal{X}_h$ defined by $\mathcal{L}_h^{\boldsymbol{y}} f = u_h(\boldsymbol{y})$. Then, we obtain the discrete formulation

$$f_{m,h}^* := \operatorname*{argmin}_{f\in X} J_{m,h}(f), \qquad J_{m,h}(f) := \mathcal{R}_m(\tfrac{\alpha_1}{2}\left\|\mathcal{L}_h^{\boldsymbol{y}} f - \hat{u}\right\|_{\mathcal{Y}}^2) + \tfrac{\alpha_2}{2}\left\|f\right\|_{\mathcal{Y}}^2, \tag{14}$$

where $\mathcal{R}_m(\Phi) = \frac{1}{\theta}\log\left(\frac{1}{b^m}\sum_{\boldsymbol{y}\in P_m}\exp(\theta\Phi(\boldsymbol{y}))\right)$ is again convex, due to positivity of the QMC quadrature weights. The derivative $J'_{m,h}(f)\in\mathcal{Y}$ of $J_{m,h}$ is analogous to (13), again replacing the integrals by sample averages over P_m, $\Phi_{h,f}(\boldsymbol{y}) = \frac{\alpha_1}{2}\left\|\mathcal{L}_h^{\boldsymbol{y}} f - \hat{u}\right\|_{\mathcal{Y}}^2$ and $q_{h,f}(\boldsymbol{y}) = \alpha_1\mathcal{L}_h^{\boldsymbol{y}}(\mathcal{L}_h^{\boldsymbol{y}} f - \hat{u}) \in \mathcal{Y}$. The next proposition recasts the error in the approximation $f^*_{m,h}\approx f^*$ to the form (4). Whenever it does not cause confusion, we will write $q(\boldsymbol{y}) = q_{f^*_{m,h}}(\boldsymbol{y})$ and $q_h(\boldsymbol{y}) = q_{h,f^*_{m,h}}(\boldsymbol{y})$.

Proposition 1. *For $\mathcal{Y} = L^2(D)$, assume Z, Z' in* (1) *are defined by $\Theta(\boldsymbol{y}) = \exp(\theta\Phi_{f^*_{m,h}}(\boldsymbol{y}))$ and $\Theta'(\boldsymbol{y}) = q(\boldsymbol{y})\Theta(\boldsymbol{y})$. Similarly, let $Z_{m,h}, Z'_{m,h}$ in* (3) *be defined by $\Theta_h(\boldsymbol{y}) = \exp(\theta\Phi_{h,f^*_{m,h}}(\boldsymbol{y}))$ and $\Theta'_h(\boldsymbol{y}) = q_h(\boldsymbol{y})\Theta_h(\boldsymbol{y})$.*

Then

$$\left\|f^* - f^*_{m,h}\right\|_{\mathcal{Y}} \le \frac{1}{\alpha_2}\left\|J'(f^*_{m,h}) - J'_{m,h}(f^*_{m,h})\right\|_{\mathcal{Y}} = \frac{1}{\alpha_2}\left\|\frac{Z'}{Z} - \frac{Z'_{m,h}}{Z_{m,h}}\right\|_{\mathcal{Y}}. \tag{15}$$

Proof. The inequalities $\langle J'(f^*), f - f^*\rangle \ge 0$ and $\langle J'_{m,h}(f^*_{m,h}), f - f^*_{m,h}\rangle \ge 0$ for all $f \in X$ imply that $\langle J'_{m,h}(f^*_{m,h}) - J'(f^*), f^* - f^*_{m,h}\rangle \ge 0$. Moreover, strong convexity of J yields the relation $\langle J'(f^*) - J'(f^*_{m,h}) - \alpha_2(f^* - f^*_{m,h}), f^* - f^*_{m,h}\rangle \ge 0$. Thus, we get

$$\begin{aligned}\alpha_2\left\|f^* - f^*_{m,h}\right\|_{\mathcal{Y}}^2 &\le \langle J'_{m,h}(f^*_{m,h}) - J'(f^*) + \alpha_2(f^* - f^*_{m,h}), f^* - f^*_{m,h}\rangle\\ &\le \langle J'_{m,h}(f^*_{m,h}) - J'(f^*_{m,h}), f^* - f^*_{m,h}\rangle\\ &\le \left\|J'_{m,h}(f^*_{m,h}) - J'(f^*_{m,h})\right\|_{\mathcal{Y}}\left\|f^* - f^*_{m,h}\right\|_{\mathcal{Y}},\end{aligned}$$

which implies the inequality in (15). The equality follows by substituting $J'_{m,h}(f^*_{m,h}) = \frac{Z'_{m,h}}{Z_{m,h}} + \alpha_2 f^*_{m,h}$ and (13). □

Remark 1. Note that Θ, Θ' as defined in Proposition 1, and hence Z, Z', also implicitly depend on h via the discrete minimizer $f^*_{m,h}$. In particular, the exact minimizer f^* does not appear in the right hand side of (15). This fact will be crucial for the ensuing a-posteriori error estimation methodology.

2 A-posteriori QMC error estimation

We develop computable a-posteriori error estimators for the PDE discretization error and for the QMC-quadrature error, the latter being reliable independent of the integration dimension s.

2.1 Parametric regularity

To leverage the results from [5, 14] and overcome the curse of dimensionality, we need to quantify the regularity with respect to $\boldsymbol{y}\in U$. To this end, we write $|\boldsymbol{\nu}| = \sum_j \nu_j$, $\operatorname{supp}(\boldsymbol{\nu}) = \{j : \nu_j \neq 0\}$ and, given smooth $F\colon U\to\mathcal{Y}$ and $\boldsymbol{\nu}\in\mathcal{F} := \{\boldsymbol{\nu}\in\mathbb{N}_0^{\mathbb{N}} : |\operatorname{supp}(\boldsymbol{\nu})| < \infty\}$ we introduce the multi-index notation $\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} F(\boldsymbol{y}) = \prod_{j\in\operatorname{supp}(\boldsymbol{\nu})}\partial_{y_j}^{\nu_j} F(\boldsymbol{y})$ for the derivatives with respect to $\boldsymbol{y}$. Moreover, given $\boldsymbol{\beta} = (\beta_j)_{j\in\mathbb{N}}\in\ell^p(\mathbb{N})$ for some $p\in(0,1)$, $n\in\mathbb{N}$ and $c>0$, we define the SPOD weights $\boldsymbol{\gamma} = (\gamma_{\mathfrak{u}})_{\mathfrak{u}\subseteq\{1:s\}}$ via

$$\gamma_{\mathfrak{u}} = \sum_{\boldsymbol{\nu}\in\{1:\alpha\}^{|\mathfrak{u}|}} (|\boldsymbol{\nu}| + n)!\prod_{j\in\mathfrak{u}} c\beta_j^{\nu_j}. \tag{16}$$

Definition 1. Let $\mathcal{Y}$ be a separable Hilbert space, $\alpha \in \mathbb{N}$, $\alpha \geq 2$. We define the weighted unanchored Sobolev space $\mathcal{W}_{s,\alpha,\boldsymbol{\gamma}}$ with dominating mixed smoothness as the completion of $C^\infty(U, \mathcal{Y})$ with respect to the norm

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

where $\boldsymbol{\mu} = (\boldsymbol{\nu}_{\mathfrak{u}\setminus\mathfrak{v}}, \boldsymbol{\alpha}_{\mathfrak{v}}) \in \mathcal{F}$ is such that $\mu_j = \nu_j$ if $j \in \mathfrak{u}\setminus\mathfrak{v}$, $\mu_j = \alpha$ if $j \in \mathfrak{v}$ and $\mu_j = 0$ otherwise. The inner integral is interpreted as a Bochner integral.

The relevance of the space $\mathcal{W}_{s,\alpha,\boldsymbol{\gamma}}$ in our context is justified by the following result, which will be the starting point of our analysis. This is a consequence of the so-called component-by-component (CBC) construction as described in [5,14], which takes as input $s, \alpha, \boldsymbol{\gamma}$ and m and returns a polynomial lattice P_m.

Theorem 1. *Let $\alpha \in \mathbb{N}, \alpha \geq 2$ and $F\colon U \to \mathcal{Y}$ for some separable Hilbert space $\mathcal{Y}$ be such that $F \in \mathcal{W}_{s,\alpha,\boldsymbol{\gamma}}$ for some weights $\boldsymbol{\gamma}$ of the form (16) for $\boldsymbol{\beta} \in \ell^p(\mathbb{N})$, $p \in (0, 1/2]$. Then, there exists a sequence $(P_m)_{m\in\mathbb{N}}$ of polynomial lattice rules such that as $m \to \infty$*

$$\mathfrak{E}_m(F) := \int_U F - \frac{1}{b^m} \sum_{\boldsymbol{y}\in P_m} F(\boldsymbol{y}) = \|F\|_{s,\alpha,\boldsymbol{\gamma}}\, O(b^{-m}),$$

as well as for all $\varepsilon > 0$

$$\mathfrak{E}_m(F) = \frac{1}{b^m} \sum_{\boldsymbol{y}\in P_m} F(\boldsymbol{y}) - \frac{1}{b^{m-1}} \sum_{\boldsymbol{y}\in P_{m-1}} F(\boldsymbol{y}) + \|F\|_{s,\alpha,\boldsymbol{\gamma}}\, O(b^{-2m+\varepsilon}).$$

Here the constants hidden in $O(\cdot)$ are independent of $s, m \in \mathbb{N}$ and F, but depend on ε and $\boldsymbol{\gamma}$. In addition, the point sets $(P_m)_{m\in\mathbb{N}}$ can be constructed explicitly by a CBC construction in $O(smb^m + s^2 b^m)$ operations.

Proof. For the case $\mathcal{Y} = \mathbb{R}$, this is proved in [5, Theorem 4.1]. Otherwise, let $v \in \mathcal{Y}$ arbitrary such that $\|v\|_{\mathcal{Y}} = 1$. Then, $\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} \langle F(\boldsymbol{y}), v\rangle = \langle \partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} F(\boldsymbol{y}), v\rangle$ for all $\boldsymbol{\nu} \in \mathcal{F}$ implies $\|\langle F, v\rangle\|_{s,\alpha,\boldsymbol{\gamma}} \leq \|F\|_{s,\alpha,\boldsymbol{\gamma}}$. Hence, we reduced to the case $\mathcal{Y} = \mathbb{R}$ and by linearity of the integral and the quadrature we conclude. □

Corollary 1. *Assume that the PDE (5) satisfies (6) and (7). Then we can construct polynomial lattice rules $(P_m)_{m\in\mathbb{N}}$ (depending on $\boldsymbol{b}$) in $O(smb^m + s^2 b^m)$ operations such that for all $\varepsilon > 0$ it holds for the BIP*

$$Z - Z_m = Z_m - Z_{m-1} + O(b^{-2m+\varepsilon}), \qquad Z' - Z'_m = Z'_m - Z'_{m-1} + O(b^{-2m+\varepsilon}),$$

as $m \to \infty$. The hidden constants in $O(\cdot)$ are independent of $s, m \in \mathbb{N}$.

Proof. From [18, Section 4.1], [4, Theorem 3.1] and $\boldsymbol{\nu}! := \prod_{j\in\mathrm{supp}(\boldsymbol{\nu})} \nu_j! \leq |\boldsymbol{\nu}|!$, we have that $\sup_s \|\Theta\|_{s,\alpha,\boldsymbol{\gamma}} + \sup_s \|\Theta'\|_{s,\alpha,\boldsymbol{\gamma}} < \infty$ for $\alpha = 1 + \lfloor 1/p \rfloor$ and some SPOD weights (16) defined by a sequence $\boldsymbol{\beta} \sim \boldsymbol{b}$ and $n = 0$. Hence we can apply Theorem 1 and conclude. □

A similar result can be given for the OCP, based on the results in [8, 9].

Corollary 2. *Assume that the PDE* (5) *satisfies* (6) *and* (7)*. Then we can construct polynomial lattice rules* $(P_m)_{m\in\mathbb{N}}$ *(depending on* $\boldsymbol{b}$*) in* $O(smb^m + s^2b^m)$ *operations such that for all* $\varepsilon > 0$ *it holds for the OCP*

$$Z - Z_m = Z_m - Z_{m-1} + O(b^{-2m+\varepsilon}), \qquad Z' - Z'_m = Z'_m - Z'_{m-1} + O(b^{-2m+\varepsilon}),$$

as $m \to \infty$*. The hidden constants in* $O(\cdot)$ *are independent of* $s, m \in \mathbb{N}$*.*

Proof. Applying the result of [8, Lemma 4.6] in [9, Theorems 5.4 and 5.6], we conclude that $\sup_s \|\Theta\|_{s,\alpha,\gamma} + \sup_s \|\Theta'\|_{s,\alpha,\gamma} < \infty$ for $\alpha = 1 + \lfloor 1/p \rfloor$ and for some SPOD weights (16) defined by a sequence $\boldsymbol{\beta} \sim \boldsymbol{b}$ and $n = 2$. □

2.2 Ratio error estimator

Proposition 2. *Assume that the PDE* (5) *satisfies* (6) *and* (7)*. Then, for the BIP and the OCP, we can construct polynomial lattice rules* $(P_m)_{m\in\mathbb{N}}$ *such that*

$$\frac{Z'}{Z} - \frac{Z'_m}{Z_m} = \frac{1}{bZ_m - Z_{m-1}} \left(\frac{Z_{m-1}Z'_m - Z_m Z'_{m-1}}{Z_m} \right) + O(b^{-2m+\varepsilon}), \qquad (17)$$

holds for all $\varepsilon > 0$ *as* $m \to \infty$*. The hidden constant in* $O(\cdot)$ *is independent of* s, m*.*

Proof. First, $Z - Z_m \to 0$ as $m \to \infty$ implies that $Z_m > Z/2 > 0$ for m sufficiently large. Next, due to either Corollary 1 or 2, for all $\varepsilon > 0$ we have

$$\begin{aligned}
&\frac{Z'}{Z} - \frac{Z'_m}{Z_m}\\
&= \frac{(Z' - Z'_m)Z_m - (Z - Z_m)Z'_m}{(Z - Z_m)Z_m + Z_m^2}\\
&= \frac{\frac{1}{b-1}(Z'_m - Z'_{m-1} + O(b^{-2m+\varepsilon}))Z_m - \frac{1}{b-1}(Z_m - Z_{m-1} + O(b^{-2m+\varepsilon}))Z'_m}{\frac{1}{b-1}(Z_m - Z_{m-1} + O(b^{-2m+\varepsilon}))Z_m + Z_m^2}\\
&= \frac{\frac{1}{b-1}(Z'_m - Z'_{m-1})}{\frac{1}{b-1}(bZ_m - Z_{m-1}) + O(b^{-2m+\varepsilon})}\\
&\quad - \frac{\frac{1}{b-1}(Z_m - Z_{m-1})Z'_m}{(\frac{1}{b-1}(bZ_m - Z_{m-1}) + O(b^{-2m+\varepsilon}))Z_m} + O(b^{-2m+\varepsilon}).
\end{aligned}$$

Clearly $A_m := \frac{1}{b-1}(bZ_m - Z_{m-1}) \to Z$ as $m \to \infty$, so that $A_m > Z/2 > 0$ for m sufficiently large. Hence, by a geometric sum argument, collecting in B_m all terms contained in $O(b^{-2m+\varepsilon})$ in the denominator, we obtain for $m \to \infty$ that

$$\frac{1}{A_m + B_m} = \frac{1}{A_m} \sum_{k=0}^{\infty} (-1)^k \left(\frac{B_m}{A_m} \right)^k = \frac{1}{A_m} + O(b^{-2m+\varepsilon}).$$

Therefore, as $m \to \infty$

$$\frac{Z'}{Z} - \frac{Z'_m}{Z_m} = \frac{\frac{1}{b-1}(Z'_m - Z'_{m-1})}{A_m} - \frac{\frac{1}{b-1}(Z_m - Z_{m-1})Z'_m}{A_m Z_m} + O(b^{-2m+\varepsilon}), \qquad (18)$$

which, upon rearranging the terms, is the claim. □

Since $\frac{Z'}{Z} - \frac{Z'_m}{Z_m} = O(b^{-m}) \gg O(b^{-2m+\varepsilon})$, Proposition 2 states that

$$E_{b^m}(\Theta', \Theta) := \frac{1}{bZ_m - Z_{m-1}} \left(\frac{Z_{m-1}Z'_m - Z_m Z'_{m-1}}{Z_m} \right) \tag{19}$$

is a computable, asymptotically exact error estimator.

3 A-posteriori FEM error estimation

In practice the parametric solution $u(\boldsymbol{y}) \in \mathcal{X}$ is not exactly available, and hence Z_m, Z'_m are not computable. For any $\boldsymbol{y} \in U$, $u(\boldsymbol{y})$ will be approximated by the corresponding Galerkin discretizations $u_h(\boldsymbol{y}) \in \mathcal{X}_h$.

3.1 Ratio error estimator

Let $\Theta_h, \Theta'_h, Z_h, Z'_h$ be defined replacing u and q by u_h, q_h in the definitions of Θ, Θ', Z, Z', respectively. Similarly, let $Z_{m,h}, Z'_{m,h}$ be defined replacing u and q by u_h, q_h in the definitions of Z_m, Z'_m, respectively.

Proposition 3. *Assume there exist $\zeta_{m,h}, \zeta'_{m,h}$ such that for some $c > 0$ independent of $h \in H$, $|Z_m - Z_{m,h}| \leq c\zeta_{m,h}$ and $\left\| Z'_m - Z'_{m,h} \right\|_{\mathcal{Y}} \leq c\zeta'_{m,h}$. Assume that $\zeta_{m,h} \to 0$ as $h \to 0$. Then, there exists $h_0 \in H$ such that for all $h \leq h_0$*

$$\left\| \frac{Z'_m}{Z_m} - \frac{Z'_{m,h}}{Z_{m,h}} \right\|_{\mathcal{Y}} \leq c \frac{Z_{m,h}\zeta'_{m,h} + \left\| Z'_{m,h} \right\|_{\mathcal{Y}} \zeta_{m,h}}{Z^2_{m,h} - c\zeta_{m,h} Z_{m,h}}.$$

Proof. Due to the limit $\zeta_{m,h} \to 0$ there exists $h_1 \in H$ such that $Z_{m,h} \geq Z_m/2 > 0$ for all $h \leq h_1, h \in H$. Therefore, we can pick $h_0 \in H$ such that $Z_{m,h} > c\zeta_{m,h}$ for all $h \leq h_0, h \in H$. Thus

$$\begin{aligned} \left\| \frac{Z'_m}{Z_m} - \frac{Z'_{m,h}}{Z_{m,h}} \right\|_{\mathcal{Y}} &= \left\| \frac{(Z'_m - Z'_{m,h})Z_{m,h} - (Z_m - Z_{m,h})Z'_{m,h}}{(Z_m - Z_{m,h})Z_{m,h} + Z^2_{m,h}} \right\|_{\mathcal{Y}} \\ &\leq \frac{\left\| Z'_m - Z'_{m,h} \right\|_{\mathcal{Y}} Z_{m,h} + |Z_m - Z_{m,h}| \left\| Z'_{m,h} \right\|_{\mathcal{Y}}}{Z^2_{m,h} - |Z_m - Z_{m,h}| Z_{m,h}} \\ &\leq c \frac{Z_{m,h}\zeta'_{m,h} + \left\| Z'_{m,h} \right\|_{\mathcal{Y}} \zeta_{m,h}}{Z^2_{m,h} - c\zeta_{m,h} Z_{m,h}}. \end{aligned}$$

□

Due to this result, we are left with the task of finding computable $\zeta_{m,h}, \zeta'_{m,h}$ satisfying the conditions $|Z_m - Z_{m,h}| \leq c\zeta_{m,h}$ and $\left\| Z'_m - Z'_{m,h} \right\|_{\mathcal{Y}} \leq c\zeta'_{m,h}$ for some $c > 0$ independent of m, h. Thus, in the following sections we will provide such error estimators for BIP and OCP.

3.2 FEM error estimators for BIP

For a finite collection of observation functionals $\mathcal{O} = (\mathcal{O}_1, \ldots, \mathcal{O}_K) \in (\mathcal{X}^*)^K$, define $\|\mathcal{O}\|_{\mathcal{X}^*} = \sqrt{\sum_{k=1}^K \|\mathcal{O}_k\|_{\mathcal{X}^*}^2}$. The starting point to estimate the FEM error will be the following well-known result, e.g. [16]. Let $\{\mathcal{T}_h\}_{h\in H}$ be a family of shape-regular, simplicial meshes of D and let $\mathbb{P}_k(\mathcal{T}_h)$ be the set of piecewise polynomial functions on $\mathcal{T}_h$ of degree at most $k \in \mathbb{N}_0$ in each $T \in \mathcal{T}_h$. Let $\mathcal{E}_h$ be the set of interior edges of all elements $T \in \mathcal{T}_h$. We assume that $\mathcal{X}_h := \mathbb{P}_1(\mathcal{T}_h) \cap \mathcal{X}$, that $f \in L^2(D)$ and that $a(\boldsymbol{y}) \in W^{1,\infty}(D)$. Let $h_T = |T|^{1/2}$ for $T \in \mathcal{T}_h$ and h_e the length of an edge $e \in \mathcal{E}_h$. Then we define the a-posteriori error estimator

$$\begin{aligned}\eta_{\boldsymbol{y},h}^2 := \sum_{T\in\mathcal{T}_h} \Bigg[& h_T^2 \, \|f + \operatorname{div}(a(\boldsymbol{y})\nabla u_h(\boldsymbol{y}))\|_{L^2(T)}^2 \\ & + \frac{1}{2} \sum_{e\subseteq\partial T, e\in\mathcal{E}_h} h_e \, \|[\![a(\boldsymbol{y})\nabla u_h(\boldsymbol{y})]\!]\|_{L^2(e)}^2 \Bigg].\end{aligned} \tag{20}$$

By [16, Theorem 6.3] there exists some $c^* > 0$, only depending on D and the shape regularity constant of $\{\mathcal{T}_h\}_{h\in H}$, and in particular independent of $\boldsymbol{y} \in U$ and $h \in H$, such that

$$\|u(\boldsymbol{y}) - u_h(\boldsymbol{y})\|_{\mathcal{X}} \le c^* \eta_{\boldsymbol{y},h}. \tag{21}$$

For the important special case $\mathcal{O} \in (L^2(D))^K$ and $G \in L^2(D)$ we may derive sharper estimates. To simplify the presentation, we assume here that the physical domain $D \subseteq \mathbb{R}^2$ is a convex polygon (see also Remark 3 below), and introduce the $L^2(D)-$residual estimator

$$\begin{aligned}\tilde{\eta}_{\boldsymbol{y},h}^2 := \sum_{T\in\mathcal{T}_h} \Bigg[& h_T^4 \, \left\|f_{m,h}^* + \operatorname{div}(a(\boldsymbol{y})\nabla u_h(\boldsymbol{y}))\right\|_{L^2(T)}^2 \\ & + \frac{1}{2} \sum_{e\subseteq\partial T, e\in\mathcal{E}_h} h_e^3 \, \|[\![a(\boldsymbol{y})\nabla u_h(\boldsymbol{y})]\!]\|_{L^2(e)}^2 \Bigg].\end{aligned} \tag{22}$$

The additional factors h_T, h_e are derived from a standard duality argument, see, e.g. [20, Section 1.11]. Then, there exists some $c^* > 0$ depending only on D and the shape regularity constant of $\{\mathcal{T}_h\}_{h\in H}$, and in particular independent of $\boldsymbol{y} \in U$ and $h \in H$, such that

$$\|u(\boldsymbol{y}) - u_h(\boldsymbol{y})\|_{L^2(D)} \le c^* \tilde{\eta}_{\boldsymbol{y},h}. \tag{23}$$

Lemma 1. *Fix a regular mesh $\mathcal{T}_h$ of simplices in D and a parameter vector $\boldsymbol{y} \in U$ and assume* (21). *Then*

$$|\Theta(\boldsymbol{y}) - \Theta_h(\boldsymbol{y})| \le \Theta_h(\boldsymbol{y})(e^{\chi_{\boldsymbol{y},h}} - 1) =: \zeta_{\boldsymbol{y},h}$$

holds for

$$\chi_{\boldsymbol{y},h} := \left\|\Gamma^{-1/2}\mathcal{O}\right\|_{\mathcal{X}^*} \left[|\delta - \mathcal{O}(u_h(\boldsymbol{y}))|_\Gamma + \frac{1}{2} \left\|\Gamma^{-1/2}\mathcal{O}\right\|_{\mathcal{X}^*} c^* \eta_{\boldsymbol{y},h} \right] c^* \eta_{\boldsymbol{y},h}.$$

Furthermore, if $\mathcal{O} \in (L^2(D))^K$ and $G \in L^2(D)$, then

$$|\Theta(\boldsymbol{y}) - \Theta_h(\boldsymbol{y})| \le \Theta_h(\boldsymbol{y})(e^{\tilde{\chi}_{\boldsymbol{y},h}} - 1) := \tilde{\zeta}_{\boldsymbol{y},h}$$

holds for

$$\tilde{\chi}_{\boldsymbol{y},h} := \left\|\Gamma^{-1/2}\mathcal{O}\right\|_{L^2(D)} \left[|\delta - \mathcal{O}(u_h(\boldsymbol{y}))|_\Gamma + \frac{1}{2}\left\|\Gamma^{-1/2}\mathcal{O}\right\|_{L^2(D)} c^*\tilde{\eta}_{\boldsymbol{y},h}\right] c^*\tilde{\eta}_{\boldsymbol{y},h}.$$

Proof. We define $\Delta_h(\boldsymbol{y}) := -\frac{1}{2}|\delta - \mathcal{O}(u(\boldsymbol{y}))|_\Gamma^2 + \frac{1}{2}|\delta - \mathcal{O}(u_h(\boldsymbol{y}))|_\Gamma^2$ to obtain

$$|\Theta(\boldsymbol{y}) - \Theta_h(\boldsymbol{y})| = |\Theta_h(\boldsymbol{y})(e^{\Delta_h(\boldsymbol{y})} - 1)| \leq \Theta_h(\boldsymbol{y})(e^{|\Delta_h(\boldsymbol{y})|} - 1).$$

The first part of the claim now follows with (21) and

$$\begin{aligned} |\Delta_h(\boldsymbol{y})| &\leq \frac{1}{2}\left\|\Gamma^{-1/2}\mathcal{O}\right\|_{\mathcal{X}^*} |2\delta - \mathcal{O}(u(\boldsymbol{y}) + u_h(\boldsymbol{y}))|_\Gamma \|u(\boldsymbol{y}) - u_h(\boldsymbol{y})\|_{\mathcal{X}} \\ &\leq \left\|\Gamma^{-1/2}\mathcal{O}\right\|_{\mathcal{X}^*} |\delta - \mathcal{O}(u_h(\boldsymbol{y}))|_\Gamma \|u(\boldsymbol{y}) - u_h(\boldsymbol{y})\|_{\mathcal{X}} \\ &\quad + \frac{1}{2}\left\|\Gamma^{-1/2}\mathcal{O}\right\|_{\mathcal{X}^*}^2 \|u(\boldsymbol{y}) - u_h(\boldsymbol{y})\|_{\mathcal{X}}^2. \end{aligned}$$

The second part of the proof follows analogously by replacing $\mathcal{X}$ by $L^2(D)$ and using (23) instead of (21). □

Lemma 2. *Fix a regular mesh of simplices $\mathcal{T}_h$ and $\boldsymbol{y} \in U$ and assume (21). Then there exists a constant $c^* > 0$ such that for all $\boldsymbol{y} \in U$ and $h \in H$*

$$|\Theta'(\boldsymbol{y}) - \Theta'_h(\boldsymbol{y})| \leq \|G\|_{\mathcal{X}^*} (c^*\eta_{\boldsymbol{y},h}\Theta_h(\boldsymbol{y})e^{\chi_{\boldsymbol{y},h}} + \zeta_{\boldsymbol{y},h}\|u_h(\boldsymbol{y})\|_{\mathcal{X}}) =: \zeta'_{\boldsymbol{y},h}.$$

Furthermore, if $\mathcal{O} \in (L^2(D))^K$ and $G \in L^2(D)$, then

$$|\Theta'(\boldsymbol{y}) - \Theta'_h(\boldsymbol{y})| \leq \|G\|_{L^2(D)} \left(c^*\tilde{\eta}_{\boldsymbol{y},h}\Theta_h(\boldsymbol{y})e^{\tilde{\chi}_{\boldsymbol{y},h}} + \tilde{\zeta}_{\boldsymbol{y},h}\|u_h(\boldsymbol{y})\|_{L^2(D)}\right) =: \tilde{\zeta}'_{\boldsymbol{y},h}.$$

Proof. Since $\Theta'(\boldsymbol{y}) = G(u(\boldsymbol{y}))\Theta(\boldsymbol{y}) = G(u(\boldsymbol{y})\Theta(\boldsymbol{y}))$, we get

$$\begin{aligned} |\Theta'(\boldsymbol{y}) - \Theta'_h(\boldsymbol{y})| &\leq \|G\|_{\mathcal{X}^*} \|u(\boldsymbol{y})\Theta(\boldsymbol{y}) - u_h(\boldsymbol{y})\Theta_h(\boldsymbol{y})\|_{\mathcal{X}} \\ &\leq \|G\|_{\mathcal{X}^*} (\|u(\boldsymbol{y}) - u_h(\boldsymbol{y})\|_{\mathcal{X}}\Theta(\boldsymbol{y}) + \|u_h(\boldsymbol{y})\|_{\mathcal{X}} |\Theta(\boldsymbol{y}) - \Theta_h(\boldsymbol{y})|), \end{aligned}$$

and hence the claim follows with Lemma 1 since $\Theta(\boldsymbol{y}) \leq \Theta_h(\boldsymbol{y})e^{\chi_{\boldsymbol{y},h}}$. The second part of the claim follows analogously by the second part of Lemma 1. □

Remark 2. We remark that the estimates in the proof of Lemma 2 are conservative: we used that $K > 1$ in Lemma 1. For $K = 1$, i.e. for a single observation functional, goal-oriented AFEM results from [1] and the references there, can be used to obtain sharper a-posteriori error bounds.

We now can define $\zeta_{m,h}$ by averaging $\zeta_{\boldsymbol{y},h}$ for $\boldsymbol{y} \in P_m$, that is $\zeta_{m,h} := \frac{1}{b^m}\sum_{\boldsymbol{y}\in P_m}\zeta_{\boldsymbol{y},h}$. Then we get from Lemma 1

$$|Z_m - Z_{m,h}| \leq \frac{1}{b^m}\sum_{\boldsymbol{y}\in P_m} |\Theta(\boldsymbol{y}) - \Theta_h(\boldsymbol{y})| \leq \zeta_{m,h}. \tag{24}$$

Analogously, with $\zeta'_{m,h} = \frac{1}{b^m}\sum_{\boldsymbol{y}\in P_m}\zeta'_{\boldsymbol{y},h}$, Lemma 2 implies

$$\left|Z'_m - Z'_{m,h}\right| \leq \frac{1}{b^m}\sum_{\boldsymbol{y}\in P_m} |\Theta'(\boldsymbol{y}) - \Theta'_h(\boldsymbol{y})| \leq \zeta'_{m,h}. \tag{25}$$

In particular, if we construct $\mathcal{T}_h$ such that it also holds $\eta_{\boldsymbol{y},h} \to 0$ as $h \to 0$ for all $\boldsymbol{y} \in U$, (24) and (25) verify the hypotheses of Proposition 3 with $c = 1$.

3.3 FEM error estimators for OCP with entropic risk

In the case of OCP we require error estimates for the parametric state at the discrete optimal control $f^*_{m,h}$, i.e. $u(\boldsymbol{y}) = \mathcal{L}^{\boldsymbol{y}} f^*_{m,h}$ and for the corresponding adjoint state $q(\boldsymbol{y}) = \alpha_1 \mathcal{L}^{\boldsymbol{y}}(\mathcal{L}^{\boldsymbol{y}} f^*_{m,h} - \hat{u})$. The error will be measured in the $L^2(D)$-norm. Again we assume that $\mathcal{X}_h := \mathbb{P}_1(\mathcal{T}_h) \cap \mathcal{X}$, that $f \in L^2(D)$ and $a(\boldsymbol{y}) \in W^{1,\infty}(D)$, and that $D \subseteq \mathbb{R}^2$ is a convex polygon.

Lemma 3. *Fix a mesh $\mathcal{T}_h$ and $\boldsymbol{y} \in U$ and impose* (23). *With the notation of Proposition 1 and $\tilde{\eta}_{\boldsymbol{y},h}$ defined as in* (22), *we have*

$$|\Theta(\boldsymbol{y}) - \Theta_h(\boldsymbol{y})| \leq \Theta_h(\boldsymbol{y})(e^{\chi_{\boldsymbol{y},h}} - 1) =: \zeta_{\boldsymbol{y},h},$$

where

$$\chi_{\boldsymbol{y},h} := \theta c^* \left(\frac{\alpha_1}{2} \tilde{\eta}^2_{\boldsymbol{y},h} + \alpha_1 \left\| \mathcal{L}^{\boldsymbol{y}}_h f^*_{m,h} - \hat{u} \right\|_{L^2(D)} \tilde{\eta}_{\boldsymbol{y},h} \right).$$

Proof. By twofold application of the triangle inequality we have

$$\begin{aligned} \left| \Phi_{f^*_{m,h}}(\boldsymbol{y}) - \Phi_{h,f^*_{m,h}}(\boldsymbol{y}) \right| &\leq \frac{\alpha_1}{2} \left\| (\mathcal{L}^{\boldsymbol{y}} - \mathcal{L}^{\boldsymbol{y}}_h) f^*_{m,h} \right\|^2_{L^2(D)} \\ &\quad + \alpha_1 \left\| \mathcal{L}^{\boldsymbol{y}}_h f^*_{m,h} - \hat{u} \right\|_{L^2(D)} \left\| (\mathcal{L}^{\boldsymbol{y}} - \mathcal{L}^{\boldsymbol{y}}_h) f^*_{m,h} \right\|_{L^2(D)} \\ &\leq c^* \left(\frac{\alpha_1}{2} \tilde{\eta}^2_{\boldsymbol{y},h} + \alpha_1 \left\| \mathcal{L}^{\boldsymbol{y}}_h f^*_{m,h} - \hat{u} \right\|_{L^2(D)} \tilde{\eta}_{\boldsymbol{y},h} \right) =: c^* \xi_{\boldsymbol{y},h}. \end{aligned}$$

Note that $\xi_{\boldsymbol{y},h}$ is computable due to Remark 1. Then, it follows with $\Delta_h(\boldsymbol{y}) := \theta(\Phi_{f^*_{m,h}}(\boldsymbol{y}) - \Phi_{h,f^*_{m,h}}(\boldsymbol{y}))$ that

$$|\Theta(\boldsymbol{y}) - \Theta_h(\boldsymbol{y})| = \left| \Theta_h(\boldsymbol{y})(e^{\Delta_h(\boldsymbol{y})} - 1) \right| \leq \Theta_h(\boldsymbol{y})(e^{|\Delta_h(\boldsymbol{y})|} - 1).$$

□

Using a residual estimator in the form (22) for the adjoint problem yields

$$\|q(\boldsymbol{y}) - q_h(\boldsymbol{y})\|_{L^2(D)} \leq 2 \max(c^*, 1) c^* \tilde{\tilde{\eta}}_{\boldsymbol{y},h}, \tag{26}$$

where

$$\begin{aligned} \tilde{\tilde{\eta}}^2_{\boldsymbol{y},h} := \alpha_1^2 \sum_{T \in \mathcal{T}_h} \Bigg[& h_T^4 \left\| u_h(\boldsymbol{y}) - \hat{u} + \operatorname{div}(a(\boldsymbol{y}) \nabla q_h(\boldsymbol{y})) \right\|^2_{L^2(T)} \\ & + \frac{1}{2} \sum_{e \subseteq \partial T, e \in \mathcal{E}_h} h_e^3 \left\| [\![a(\boldsymbol{y}) \nabla q_h(\boldsymbol{y})]\!] \right\|^2_{L^2(e)} \Bigg] + (\max_{T \in \mathcal{T}_h} h_T^4) \tilde{\eta}^2_{\boldsymbol{y},h}. \end{aligned} \tag{27}$$

Lemma 4. *Let $\mathcal{Y} = L^2(D)$. Fix a mesh $\mathcal{T}_h$ and $\boldsymbol{y} \in U$ and impose* (23) *and* (26). *With the notation of Proposition 1, we have*

$$\|\Theta'(\boldsymbol{y}) - \Theta'_h(\boldsymbol{y})\|_{\mathcal{Y}} \leq \zeta_{\boldsymbol{y},h} \|q_h(\boldsymbol{y})\|_{\mathcal{Y}} + 2c^* \Theta_h(\boldsymbol{y}) e^{\chi_{\boldsymbol{y},h}} \tilde{\tilde{\eta}}_{\boldsymbol{y},h} =: \zeta'_{\boldsymbol{y},h}$$

Proof. As in the proof of Lemma 2, we get by $\Theta(\boldsymbol{y}) \leq \Theta_h(\boldsymbol{y}) e^{\chi_{\boldsymbol{y},h}}$ that

$$\begin{aligned} \|q(\boldsymbol{y})\Theta(\boldsymbol{y}) - q_h(\boldsymbol{y})\Theta_h(\boldsymbol{y})\|_{\mathcal{Y}} &\leq |\Theta(\boldsymbol{y}) - \Theta_h(\boldsymbol{y})| \, \|q_h(\boldsymbol{y})\|_{\mathcal{Y}} \\ &\quad + \Theta(\boldsymbol{y}) \|q(\boldsymbol{y}) - q_h(\boldsymbol{y})\|_{\mathcal{Y}} \\ &\leq \zeta_{\boldsymbol{y},h} \|q_h(\boldsymbol{y})\|_{\mathcal{Y}} + 2c^* \Theta_h(\boldsymbol{y}) e^{\chi_{\boldsymbol{y},h}} \tilde{\tilde{\eta}}_{\boldsymbol{y},h}. \end{aligned}$$

□

Remark 3. If $D \subseteq \mathbb{R}^2$ is a non-convex polygon, the reliability assumption (23) and the corresponding definitions (22), (27) of $\tilde{\eta}_{\boldsymbol{y},h}, \tilde{\tilde{\eta}}_{\boldsymbol{y},h}$ must be adapted by using weighted L^2 norms, with weights near the re-entrant corners. We refer to [21, Theorem 3.1] for a precise result in the case of the Poisson equation.

4 Combined QMC-FEM estimator

In view of Propositions 2 and 3 we employ the *computable a-posteriori estimator*

$$EST_{b^m,h} := \|E_{b^m}(\Theta'_h, \Theta_h)\|_{\mathcal{Y}} + \frac{Z_{m,h}\zeta'_{m,h} + \left\|Z'_{m,h}\right\|_{\mathcal{Y}} \zeta_{m,h}}{Z_{m,h}^2 - \zeta_{m,h} Z_{m,h}}. \tag{28}$$

Note that the QMC error estimator $\|E_{b^m}(\Theta', \Theta)\|_{\mathcal{Y}}$ derived from Proposition 2 is itself approximated by the computable expression $\|E_{b^m}(\Theta'_h, \Theta_h)\|_{\mathcal{Y}}$. In the next proposition we precise that the additional error committed due to this extra approximation is of higher asymptotic order, as $m \to \infty$. We equip the set $C^0(U, \mathcal{Y})$ with the norm $\|F\|_\infty = \sup_{\boldsymbol{y}\in U} \|F(\boldsymbol{y})\|_{\mathcal{Y}}$.

Proposition 4. *Fix a family of regular meshes $\{\mathcal{T}_h\}_{h\in H}$ such that for some $\tilde{C} > 0$ independent of $s \in \mathbb{N}, h \in H$ and some SPOD weights* (16)*, it holds*

$$\max(\|\Theta\|_{s,\alpha,\boldsymbol{\gamma}}, \|\Theta'\|_{s,\alpha,\boldsymbol{\gamma}}, \|\Theta_h\|_{s,\alpha,\boldsymbol{\gamma}}, \|\Theta'_h\|_{s,\alpha,\boldsymbol{\gamma}}) \leq \tilde{C}. \tag{29}$$

Assume that the spaces $\{\mathcal{X}_h\}_h$ are contained in $\mathcal{X}$ and that they are selected so that $\|\Theta - \Theta_h\|_\infty \to 0$ as $h \to 0$. Then we can construct a sequence of polynomial lattices $(P_m)_{m\in\mathbb{N}}$ in $O(smb^m + s^2b^m)$ operations such that, for some $h_0 \in H$ and some constant $C > 0$ (independent of m, h, s) we have for any $h < h_0$ that

$$\begin{aligned} &\|E_{b^m}(\Theta', \Theta) - E_{b^m}(\Theta'_h, \Theta_h)\|_{\mathcal{Y}} \\ &\quad \leq Cb^{-m}(\|\Theta - \Theta_h\|_\infty + \|\Theta' - \Theta'_h\|_\infty + \|\Theta - \Theta_h\|_{s,\alpha,\boldsymbol{\gamma}} + \|\Theta' - \Theta'_h\|_{s,\alpha,\boldsymbol{\gamma}}). \end{aligned}$$

Proof. Throughout the proof, $C > 0$ is a generic constant independent of m, h, s. We compute the difference of the numerators

$$\begin{aligned} \Delta_1 &:= Z_{m-1}Z'_m - Z_m Z'_{m-1} - Z_{m-1,h}Z'_{m,h} + Z_{m,h}Z'_{m-1,h} \\ &= -(Z_m - Z_{m-1})(Z'_m - Z'_{m,h}) - (Z_m - Z_{m,h} - Z_{m-1} + Z_{m-1,h})Z'_{m,h} \\ &\quad + (Z_m - Z_{m,h})(Z'_m - Z'_{m-1}) + Z_{m,h}(Z'_m - Z'_{m,h} - Z'_{m-1} + Z'_{m-1,h}). \end{aligned}$$

We have $|Z_m - Z_{m,h}| \leq \|\Theta - \Theta_h\|_\infty$ and $\left\|Z'_m - Z'_{m,h}\right\|_{\mathcal{Y}} \leq \|\Theta' - \Theta'_h\|_\infty$. From Theorem 1, we know that $Z_m = Z_{m-1} + \|\Theta\|_{s,\alpha,\boldsymbol{\gamma}} O(b^{-m})$ and $Z'_m = Z'_{m-1} + \|\Theta'\|_{s,\alpha,\boldsymbol{\gamma}} O(b^{-m})$ as $m \to \infty$, with hidden constants in $O(\cdot)$ independent of s, m, h. Furthermore $Z_m - Z_{m,h} - Z_{m-1} + Z_{m-1,h} = \|\Theta - \Theta_h\|_{s,\alpha,\boldsymbol{\gamma}} O(b^{-m})$, and $Z'_m - Z'_{m,h} - Z'_{m-1} + Z'_{m-1,h} = \|\Theta' - \Theta'_h\|_{s,\alpha,\boldsymbol{\gamma}} O(b^{-m})$ also follow by Theorem 1. Therefore, we have

$$\|\Delta_1\|_{\mathcal{Y}} \leq Cb^{-m}(\|\Theta - \Theta_h\|_\infty + \|\Theta' - \Theta'_h\|_\infty + \|\Theta - \Theta_h\|_{s,\alpha,\boldsymbol{\gamma}} + \|\Theta' - \Theta'_h\|_{s,\alpha,\boldsymbol{\gamma}}).$$

Next, we define $T_1 := Z_{m-1,h}Z'_{m,h} - Z_{m,h}Z'_{m-1,h}$ and obtain the estimate $\|T_1\|_{\mathcal{Y}} \leq C(\|\Theta_h\|_{s,\alpha,\boldsymbol{\gamma}} + \|\Theta'_h\|_{s,\alpha,\boldsymbol{\gamma}})b^{-m}$. Moreover,

$$\begin{aligned} \Delta_2 &:= (bZ_m - Z_{m-1})Z_m - (bZ_{m,h} - Z_{m-1,h})Z_{m,h} \\ &= (b(Z_m - Z_{m,h}) + (Z_{m-1,h} - Z_{m-1}))Z_m + (bZ_{m,h} - Z_{m-1,h})(Z_m - Z_{m,h}) \end{aligned}$$

gives $|\Delta_2| \leq C \|\Theta - \Theta_h\|_\infty$. Next, we observe that $T_2 = (bZ_{m,h} - Z_{m-1,h})Z_{m,h}$ is bounded from below away from 0, for h sufficiently small. Therefore, for h sufficiently small we apply the elementary inequality

$$\left\| \frac{T_1 + \Delta_1}{T_2 + \Delta_2} - \frac{T_1}{T_2} \right\|_{\mathcal{Y}} \leq \max(\|\Delta_1\|_{\mathcal{Y}}, \|T_1 \Delta_2\|_{\mathcal{Y}}) \frac{1 + |T_2|}{|T_2| \left(|T_2| - |\Delta_2|\right)},$$

valid for $T_1, T_2, \Delta_1, \Delta_2 \in \mathbb{R}$ with $|\Delta_2| < |T_2|$, which is satisfied since $|\Delta_2| \to 0$ as $h \to 0$. Combining all these observations we obtain the claim. □

Theorem 2. *For either the BIP or the OCP, assume that $D \subseteq \mathbb{R}^2$ is a convex polygon and that the PDE* (5) *satisfies* (6), (7), *$f \in L^2(D)$ and $\boldsymbol{b}' \in \ell^p(\mathbb{N}), p \in (0, 1/2]$, with ${b'}_j = \|\psi_j\|_{W^{1,\infty}(D)}$. Let $\mathcal{X}_h = \mathbb{P}_1(\mathcal{T}_h) \cap \mathcal{X}$ for a family of shape-regular meshes $\mathcal{T}_h$ such that $h = \max_{T \in \mathcal{T}_h} h_T$. Then, we can construct polynomial lattices $(P_m)_{m \in \mathbb{N}}$ such that the estimator $EST_{b^m,h}$ in* (28) *satisfies*

$$\left\| \frac{Z'}{Z} - \frac{Z'_{m,h}}{Z_{m,h}} \right\|_{\mathcal{Y}} \leq EST_{b^m,h} + O(b^{-2m+\varepsilon} + b^{-m}h),$$

for any $\varepsilon > 0$ as $m \to \infty$, $h \to 0$. The constant in $O(\cdot)$ is independent of s, m and h, but depends on ε.

Proof. Since $b_j \leq b'_j$, from either Corollary 1 or 2, there exist SPOD weights $\boldsymbol{\gamma}'$ as in (16) with $\boldsymbol{\beta}' \sim \boldsymbol{b}'$, such that $\sup_s \|\Theta\|_{s,\alpha,\boldsymbol{\gamma}'} + \|\Theta'\|_{s,\alpha,\boldsymbol{\gamma}'} < \infty$. Combining (20) with Lemma 1 and $h \to 0$ we have $\zeta_{m,h} \to 0$ for the BIP case for any $m \in \mathbb{N}$. Similarly, combining (22) with Lemma 3 yields the same observation for the OCP case. Therefore we can apply Propositions 2 and 3 to get that we can construct polynomial lattices so that as $m \to \infty$

$$\left\| \frac{Z'}{Z} - \frac{Z'_{m,h}}{Z_{m,h}} \right\|_{\mathcal{Y}} \leq \|E_{b^m}(\Theta', \Theta)\|_{\mathcal{Y}} + \frac{Z_{m,h}\zeta'_{m,h} + \left\|Z'_{m,h}\right\|_{\mathcal{Y}} \zeta_{m,h}}{Z_{m,h}^2 - \zeta_{m,h} Z_{m,h}} + O(b^{-2m+\varepsilon}).$$

Next, we say that $\boldsymbol{\rho} \in (1, \infty)^{\mathbb{N}}$ is $(\boldsymbol{b}', \varepsilon)$-admissible if $\sum_{j \geq 1} (\rho_j - 1) b'_j \leq \varepsilon$, see [4]. Then we define $\mathfrak{T}_{\boldsymbol{b}',\varepsilon} = \bigcup_{\boldsymbol{\rho}, (\boldsymbol{b}',\varepsilon)\text{-adm.}} \{\boldsymbol{y} \in \mathbb{C}^s : \operatorname{dist}(y_j, [-\frac{1}{2}, \frac{1}{2}]) \leq \rho_j - 1\}$. Following the computations in [2, Theorem 4.1], yield $h_0 \in H$ and $\varepsilon > 0$ sufficiently small such that for all $h \leq h_0, h \in H$, we have $\sup_{\boldsymbol{y} \in \mathfrak{T}_{\boldsymbol{b}',\varepsilon}} |\Theta(\boldsymbol{y}) - \Theta_h(\boldsymbol{y})| + \|\Theta'(\boldsymbol{y}) - \Theta'_h(\boldsymbol{y})\|_{\mathcal{Y}} \leq Ch$. By [4, Theorem 3.1], this implies that for a constant C independent of h, s,

$$\|\Theta - \Theta_h\| \infty + \|\Theta' - \Theta'_h\|_\infty + \|\Theta - \Theta_h\|_{s,\alpha,\boldsymbol{\gamma}'} + \|\Theta' - \Theta'_h\|_{s,\alpha,\boldsymbol{\gamma}'} \leq Ch.$$

Thus, (29) and $\|\Theta - \Theta_h\|_\infty \to 0$ hold, and we apply Proposition 4 to conclude. □

5 Numerical experiment

We consider the PDE (5) on the physical domain $D := (0, 1)^2$, with $f \equiv 10$, and parametric diffusion coefficient given by

$$a(x, \boldsymbol{y}) = \frac{1}{2} + \sum_{j=1}^{16} \frac{y_j}{(k_{j,1}^2 + k_{j,2}^2)^2} \sin(k_{j,1} x_1) \sin(k_{j,2} x_2).$$

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. We investigate a BIP with observation functional $\mathcal{O} = (\mathcal{O}_1, \dots, \mathcal{O}_4) \in (L^2(D))^4$, given by $\mathcal{O}_k(v) := \frac{1}{0.01} \int_{I_k} v dx$ for $v \in L^2(D)$ and $k = 1, \dots, 4$, where $I_1 := [0.1, 0.2]^2, I_2 := [0.1, 0.2] \times [0.8, 0.9], I_3 := [0.8, 0.9] \times [0.1, 0.2], I_4 := [0.8, 0.9]^2$. We draw a (random) sample of a to compute the "ground truth" of observations $\mathcal{O}(\mathcal{S}(a))$ on a sequence of regular FE meshes of triangles obtained by uniform refinement, with 525.313 degrees of freedom (dofs). We add random noise $\eta \sim \mathcal{N}(0, \sigma^2 \mathcal{I}_4)$ to the observations, where σ is set as 10% of the average of $\mathcal{O}(\mathcal{S}(a)) \in \mathbb{R}^4$. The realized synthetic data is then given by $\delta = (0.5205, 0.5037, 0.5443, 0.4609)^\top$.

Our aim is to approximate $\mathbb{E}^{\pi^\delta}[G(u)]$ by the ratio estimator $\frac{Z'_{m,h}}{Z_{m,h}}$, where $G \in L^2(D)$ is given by $G(v) := \frac{1}{0.5} \int_{[0.25,0.75]^2} v dx$ for $v \in L^2(D)$. The FE mesh and the polynomial lattice rule, that eventually determine h and m, are refined successively based on the combined estimator in (28). For tolerances $\tau_{FEM}, \tau_{QMC} > 0$, we start from an initial FE mesh of D, that is uniformly refined until the stopping criterion $\frac{Z_{m,h}\zeta'_{m,h} + \|Z'_{m,h}\|_{\mathcal{Y}} \zeta_{m,h}}{Z_{m,h}^2 - \zeta_{m,h} Z_{m,h}} \leq \tau_{FEM}$ is met. Thereafter, we increase the number b^m of lattice points until there holds $|E_{b^m}(\Theta'_h, \Theta_h)| \leq \tau_{QMC}$. We initialize by a FE mesh with 41 dofs and b^{m_0} QMC points with base $b = 2$ and $m_0 = 2$, and set the tolerances to $\tau_{FEM} = \tau_{QMC} = 2^{-6}$. To assess the total realized error, we compute a reference solution $\frac{Z'_{\text{ref}}}{Z_{\text{ref}}}$ by a multilevel Monte Carlo ratio estimator, see [17], and report absolute error $\left| \frac{Z'_{m,h}}{Z_{m,h}} - \frac{Z'_{\text{ref}}}{Z_{\text{ref}}} \right|$. The reference estimator uses 6 refinement levels with 545/525.313 dofs on the coarsest/finest level, respectively, and uniform (pseudo-) random numbers y. The number of samples is adjusted to balance statistical error and discretization bias on each level. The experiment has been implemented in MATLAB using the MOOAFEM library [11] for the FE discretization. All arising linear systems are solved directly by the \-operator in MATLAB.

The estimated and realized errors vs. the number of iterations (in the sense of refinement steps) are depicted Figure 1. Here, the FE a-posteriori estimator gives negative values on rather coarse meshes, where $c^* \tilde{\eta}_{\boldsymbol{y},h} > 1$. Therefore, we discarded these "pre-asymptotic" values in the plot. We see that the FE a-posteriori estimator from Proposition 3 is rather conservative at first, but eventually approaches the actual error for finer meshes. The QMC estimator $|E_{b^m}(\Theta'_h, \Theta'_h)|$ is of the same magnitude as σ at first, and only two more refinement steps are needed once the FE mesh is sufficiently fine. The combined error estimate $EST_{b^m,h}$ aligns well with the realized error, as expected from our theoretical analysis.

6 Conclusion

In this paper we outlined the construction of an a-posteriori QMC-FEM estimator, that allows to quantify the approximation error to a) posterior expectation in Bayesian inverse problems and b) the optimal control under the entropic risk measure. The estimator is computable and viable for large number of parameters s and it is asymptotically an upper bound for the errors in a) and b).

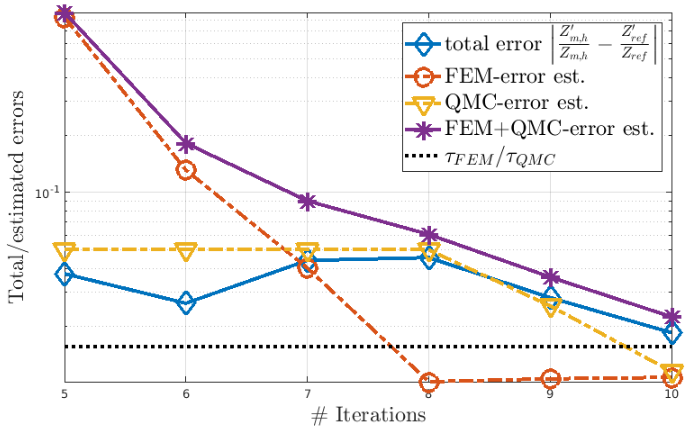

Figure 1: Results for the QMC-FEM ratio estimator with a-posterior ratio refinement. First the FE mesh is refined until the tolerance τ_{FEM} is achieved (dashed w. circles), then the QMC a-posterior refinement takes place (dashed w. triangles). The estimated error (solid w. stars) is conservative for coarse meshes, but eventually approaches the realized error (solid w. diamonds).

Furthermore, the particular ratio structure $\frac{Z'}{Z}$ of the sought quantities allows to tackle both the BIP and OCP, in a unified manner. In either case, we work under the assumption that the underlying model is a parametric elliptic PDE with affine-parametric diffusion. Nevertheless, the present QMC methodology to high-dimensional integration is applicable to non-affine parametric PDEs with *quantified, holomorphic-parametric dependence*, see [4] and the references there. Since the error estimators we consider $\eta_{\boldsymbol{y},h}, \tilde{\eta}_{\boldsymbol{y},h}, \tilde{\tilde{\eta}}_{\boldsymbol{y},h}$ are expressed as sums of local error contributions for $T \in \mathcal{T}_h$, a possible direction of research is to employ the presently proposed estimators $\zeta_{\boldsymbol{y},h}, \zeta'_{\boldsymbol{y},h}$ to steer an adaptive QMC-FEM algorithm [13].

References

[1] R. Becker, M. Brunner, M. Innerberger, J. M. Melenk, and D. Praetorius. Rate-optimal goal-oriented adaptive FEM for semilinear elliptic PDEs. *Comput. Math. Appl.*, 118:18–35, 2022.

[2] J. Dick, R. N. Gantner, Q. T. L. Gia, and Ch. Schwab. Multilevel higher-order quasi-Monte Carlo Bayesian estimation. *Math. Mod. Meth. Appl. Sci.*, 27(5):953–995, 2017.

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

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

[5] J. Dick, M. Longo, and Ch. Schwab. Extrapolated polynomial lattice rule integration in computational uncertainty quantification. *SIAM/ASA J. Uncertain. Quantif.*, 10(2):651–686, 2022.

[6] A. Ern and J.-L. Guermond. *Finite elements I—Approximation and interpolation*, volume 72 of *Texts in Applied Mathematics*. Springer, Cham, 2021.

[7] H. Föllmer and T. Knispel. Entropic risk measures: coherence vs. convexity, model ambiguity, and robust large deviations. *Stoch. Dyn.*, 11(2-3):333–351, 2011.

[8] P. A. Guth, V. Kaarnioja, F. Y. Kuo, C. Schillings, and I. H. Sloan. A quasi-Monte Carlo method for optimal control under uncertainty. *SIAM/ASA J. Uncertain. Quantif.*, 9(2):354–383, 2021.

[9] P. A. Guth, V. Kaarnioja, F. Y. Kuo, C. Schillings, and I. H. Sloan. Parabolic PDE-constrained optimal control under uncertainty with entropic risk measure using quasi-Monte Carlo integration. arXiv:2208.02767, 2022.

[10] L. Herrmann, M. Keller Ch. Schwab. Quasi-Monte Carlo Bayesian estimation under Besov priors in elliptic inverse problems. Math. Comp., **90**(2021) 1831–1860.

[11] M. Innerberger and D. Praetorius. "MooAFEM: An object oriented Matlab code for higher-order (nonlinear) adaptive FEM." , arXiv:2203.01845, 2022.

[12] D. P. Kouri and T. M. Surowiec. Existence and optimality conditions for risk-averse PDE-constrained optimization. *SIAM/ASA J. Uncertain. Quantif.*, 6(2):787–815, 2018.

[13] M. Longo. Adaptive Quasi-Monte Carlo Finite Element Methods for Parametric Elliptic PDEs. *J. Sci. Comput.*, 92(1), 2022.

[14] M. Longo. *Extrapolated polynomial lattices and adaptivity in computational Uncertainty Quantification*. PhD thesis, ETH Zürich, 2022.

[15] H. Niederreiter. Low-discrepancy point sets obtained by digital constructions over finite fields. *Czechoslovak Math. J.*, 42(117)(1):143–166, 1992.

[16] R. H. Nochetto, K. G. Siebert, and A. Veeser. Theory of adaptive finite element methods: an introduction. In *Multiscale, nonlinear and adaptive approximation*, pages 409–542. Springer, Berlin, 2009.

[17] R. Scheichl, A.M. Stuart, and A.L. Teckentrup. Quasi-Monte Carlo and multilevel Monte Carlo methods for computing posterior expectations in elliptic inverse problems. *SIAM/ASA J. Uncertain. Quantif.*, 5(1): 493-518, 2017.

[18] C. Schillings and Ch. Schwab. Sparsity in Bayesian inversion of parametric operator equations. *Inverse Problems*, 30(6):065007, 30, 2014.

[19] A. M. Stuart. Inverse problems: a Bayesian perspective. *Acta Numer.*, 19:451–559, 2010.

[20] R. Verfürth. *a-posteriori error estimation techniques for finite element methods*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford, 2013.

[21] T. P. Wihler. Weighted L^2-norm a-posteriori error estimation of FEM in polygons. *Int. J. Numer. Anal. Model.*, 4(1):100–115, 2007.