

# Higher order Quasi-Monte Carlo integration for Bayesian Estimation

J. Dick and R. Gantner and Q. Le Gia and Ch. Schwab

Research Report No. 2016-13  
February 2016

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

---

# Higher order Quasi-Monte Carlo integration for Bayesian Estimation

Josef Dick, Robert N. Gantner, Quoc T. Le Gia, and Christoph Schwab

February 18, 2016

## Abstract

We analyze combined Quasi-Monte Carlo quadrature and Finite Element approximations in Bayesian estimation of solutions to countably-parametric operator equations with holomorphic dependence on the parameters as considered in [Cl. Schillings and Ch. Schwab: Sparsity in Bayesian Inversion of Parametric Operator Equations. *Inverse Problems*, **30**, (2014)]. Such problems arise in numerical uncertainty quantification and in Bayesian inversion of operator equations with distributed uncertain inputs, such as uncertain coefficients, uncertain domains or uncertain source terms and boundary data. We show that the parametric Bayesian posterior densities belong to a class of weighted Bochner spaces of functions of countably many variables, with a particular structure of the QMC quadrature weights: up to a (problem-dependent, and possibly large) finite dimension  $S$  product weights can be used, and beyond this dimension, weighted spaces with so-called SPOD weights, recently introduced in [J. Dick, F.Y. Kuo, Q.T. Le Gia, D. Nuyens and Ch. Schwab, Christoph Higher order QMC Petrov-Galerkin discretization for affine parametric operator equations with random field inputs. *SIAM J. Numer. Anal.* **52** (2014), 2676–2702.], are used to describe the solution regularity. We establish error bounds for higher order Quasi-Monte Carlo quadrature for the Bayesian estimation based on [J. Dick, Q.T. LeGia and Ch. Schwab, Higher order Quasi-Monte Carlo integration for holomorphic, parametric operator equations, Report 2014-23, SAM, ETH Zürich]. It implies, in particular, regularity of the parametric solution and of the countably-parametric Bayesian posterior density in SPOD weighted spaces. This, in turn, implies that the Quasi-Monte Carlo quadrature methods in [J. Dick, F.Y. Kuo, Q.T. Le Gia, D. Nuyens, Ch. Schwab, Higher order QMC Galerkin discretization for parametric operator equations, *SINUM* (2014)] are applicable to these problem classes, with dimension-independent convergence rates  $\mathcal{O}(N^{-1/p})$  of  $N$ -point HoQMC approximated Bayesian estimates, where  $0 < p < 1$  depends only on the sparsity class of the uncertain input in the Bayesian estimation. Fast component-by-component (CBC for short) construction [R. N. Gantner and Ch. Schwab Computational Higher Order Quasi-Monte Carlo Integration, Report 2014-25, SAM, ETH Zürich] allow efficient Bayesian estimation with up to  $10^3$  parameters.

Key words: Quasi-Monte Carlo, Lattice rules, digital nets, parametric operator equations, infinite-dimensional quadrature, Bayesian inverse problems, Uncertainty Quantification, CBC construction, SPOD weights.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Forward UQ for Parametric Operator Equations</b>	<b>2</b>
2.1	Operator equations with uncertain input . . . . .	2
2.2	Uncertainty parametrization . . . . .	4
2.3	Dimension truncation . . . . .	5
2.4	Petrov-Galerkin discretization . . . . .	7
<b>3</b>	<b>Bayesian Inverse UQ</b>	<b>8</b>
3.1	General setup . . . . .	8
3.2	Parametric Bayesian posterior . . . . .	9
3.3	Well-posedness and approximation . . . . .	11
<b>4</b>	<b>Holomorphic parameter dependence</b>	<b>12</b>
4.1	Holomorphic Families of Countably-Parametric Maps . . . . .	13
4.2	Examples of $(\mathbf{b}, p, \varepsilon)$ -holomorphic families . . . . .	14
<b>5</b>	<b>Higher order QMC-PG method for Bayesian inverse problems</b>	<b>15</b>
5.1	Regularity of parametric solutions . . . . .	15
5.2	Holomorphy of Bayesian posterior $\Theta(\mathbf{y})$ . . . . .	16
5.3	Higher order quasi-Monte Carlo integration . . . . .	17
5.4	Combined QMC-PG Error Bound . . . . .	20
<b>6</b>	<b>Model Problems</b>	<b>22</b>
6.1	Affine-parametric diffusion problem . . . . .	22
6.2	Parametric Regularity . . . . .	23
6.3	Example 1: Karhunen-Loève series . . . . .	24
6.4	Example 2: Indicator functions . . . . .	24
<b>7</b>	<b>Numerical Results</b>	<b>24</b>
7.1	Construction of Interlaced Polynomial Lattice Rules . . . . .	24
7.2	Approximation of Prior Expectation . . . . .	25
7.3	Approximation of Posterior Expectation . . . . .	25
7.4	Results for Example 1: Karhunen-Loève Series . . . . .	26
7.4.1	Finite Element Approximation . . . . .	26
7.4.2	Dimension Truncation . . . . .	27
7.4.3	QMC Convergence . . . . .	27
7.5	Results for Example 2: Indicator Functions . . . . .	28
7.5.1	Choice of QMC weights . . . . .	29
7.5.2	QMC Convergence . . . . .	29

# 1 Introduction

The statistical estimation of solutions of operator equations which depend on *uncertain inputs*, subject to given noisy data, is a key task in computational uncertainty quantification. In the present paper we consider the particular case when the uncertain input is *distributed*. Specifically, we allow the distributed, uncertain input  $u$  to take values in an infinite-dimensional, separable Banach space  $X$ . The *forward responses* resulting from (instances of) uncertain data  $u \in X$  then take values in a second Banach space, the *state space*  $\mathcal{X}$ . In Bayesian estimation, one is interested in computing the expected value of a Quantity of Interest (QoI for short) taking values in  $\mathbb{R}$ . The mathematical expectation (or *ensemble average*) is conditional on the given, noisy observation data  $\delta \in Y$ .

The efficient computation of such QoI's in either forward or inverse problems involves two basic steps: i) approximate (numerical) solution of the operator equation in the forward problem, and ii) approximate evaluation of the mathematical expectation w.r.t. the posterior over all possible realizations of the uncertain input, conditional on given data, by some form of numerical integration.

Due to the high (infinite) dimensionality of the integration domain, Monte-Carlo methods have been widely used. In the present paper, building on our previous work [13] on high-dimensional Quasi-Monte Carlo integration and on numerical Bayesian estimation [35, 36] we propose a novel deterministic computational approach towards these aims. It consists in i) *uncertainty parametrization*: through an unconditional basis  $\{\psi_j\}_{j \geq 1}$  of  $X$ , the forward problem is transformed formally to an infinite-dimensional, parametric deterministic problem. ii) *dimension-truncation*: the uncertain input  $u$  is restricted to a finite number  $s$  of parameters. iii) *(Petrov-)Galerkin discretization* of the parametric operator equation and, finally, iv) *Quasi-Monte Carlo* (QMC) integration in  $s$  parameters from step ii) to compute approximate Bayesian estimates for the quantity of interest (QoI).

The present paper is motivated in part by [28], where QMC integration using a family of *randomly shifted lattice rules* was combined with Petrov-Galerkin Finite Element discretization for a model parametric diffusion equation, and in part by [38], where the methodology of [28] was extended to forward problems described by an abstract family of *affine-parametric, linear* operator equations.

The treatment of inverse problems is based on the infinite-dimensional Bayesian framework as developed in [44, 7]. In this work, in contrast to [28, 38], we analyze *deterministic, interlaced polynomial lattice rules* for the numerical evaluation of Bayesian estimates. As we show here, these higher order QMC quadratures can provide dimension-independent convergence rates beyond order one for smooth integrands (cf. [10, 11]); convergence order 1 was the limitation in [28, 38] and order 1/2 is an intrinsic limitation of Monte-Carlo based methods (here, convergence order is meant in terms of the number  $N$  of “samples”, i.e. of forward solves). We prove that sparsity of uncertainty parametrization implies higher order, dimension-independent convergence rates for QMC evaluation of ratio estimators for expectations of QoI's under the Bayesian posterior, for a broad class of smooth, nonlinear, and possibly nonaffine-parametric operator equations with distributed uncertain input data. Our results imply that unlike MCMC and filtering methods, the presently proposed QMC evaluation of ratio estimators can provide convergence rates larger than 1/2 regardless of the dimension of the parameter space, while also allowing for “embarrassing parallelism”.

The structure of this paper is as follows: In Section 2, we introduce a class of smooth,

nonlinear, holomorphic-parametric operator equations admitted in our approach. We present sufficient conditions on the nonlinear operators and on the uncertainty for the forward problems to be well-posed, as in [19, 34]. We require these conditions uniformly in the set  $\tilde{X}$  of admissible uncertainties. We give a parametrization of the uncertain inputs which reduces the forward problem with distributed uncertain input to a countably-parametric, deterministic problem. These parameters, denoted by  $y_j$ , are assumed scaled so as to take values in  $[-1, 1]$ .

We review several approximations of these equations which are required in their computational Bayesian inversion, in particular, (Petrov-)Galerkin discretizations of the parametric forward equations with discretization error estimates from [19, 34].

In Section 3, we review Bayesian inversion for these operator equations, based on [44, 7] and on [41, 35, 36]. The countably-parametric representation of the uncertain inputs allows us to write the integrals arising in Bayesian estimation as countably iterated parametric integrals. The principal result of the present paper, a convergence rate bound of the Quasi Monte-Carlo integration for these integrals, requires precise derivative bounds for the integrand functions. These are proved based on analytic continuation of the integrand functions into the complex domain. To this end, in Section 4, we review the notion of *holomorphy* of countably parametric integrand functions in both forward and inverse problems. Section 5.1 gives the holomorphy and resulting derivative bounds on parametric forward solutions, whereas Section 5.2 contains the corresponding holomorphy results for the Bayesian posterior densities. Section 5.3 reviews recent results on convergence theory for higher-order QMC quadratures (based on [17]) and for the countably-parametric integrands which arise in Bayesian estimation (based on [13]). Section 5.4 presents the combined error bound for the QMC-PG approximation of the Bayesian estimate.

Based on the results in Sections 5.1 and 5.2, in [13, 20] variants of the fast CBC constructions of generating vectors are developed based on [13, 22, 21] which are tailored to the ‘hybrid’ nature of the QMC weights.

Elliptic model problems illustrating the general theory of holomorphic forward maps are presented in Section 6. Numerical results for Bayesian estimation for these problems are presented in Section 7.

## 2 Forward UQ for Parametric Operator Equations

We introduce a class of smooth, nonlinear operator equations with distributed uncertain input data  $u$  taking values in a separable Banach space  $X$ . Upon appropriate uncertainty parametrization, these equations become countably-parametric operator equations.

### 2.1 Operator equations with uncertain input

Let  $X, \mathcal{X}$  and  $\mathcal{Y}$  be real, separable Banach spaces. For a distributed, uncertain parameter  $u \in X$ , assume a *nominal parameter instance*  $\langle u \rangle \in X$  (such as, for example, the expectation of an  $X$ -valued random field  $u$ ), is known.

Let  $B_X(\langle u \rangle; R)$  be an open ball of radius  $R > 0$  in  $X$  centered at a nominal input  $\langle u \rangle \in X$ . We consider the following problem:

$$\text{given } u \in B_X(\langle u \rangle; R), \text{ find } q \in \mathcal{X} \text{ s.t. } \mathcal{Y}\langle \mathcal{R}(u; q), v \rangle_{\mathcal{Y}} = 0 \quad \forall v \in \mathcal{Y}, \quad (2.1)$$

where  $\mathcal{R} : X \times \mathcal{X} \rightarrow \mathcal{Y}'$  is the residual of a *forward operator*, depending on  $u$  and acting on  $q \in \mathcal{X}$ .

Given  $u \in B_X(\langle u \rangle; R)$ , a solution  $q_0$  of (2.1) is called *regular at  $u$*  if and only if the map  $\mathcal{R}(u; \cdot)$  is Fréchet differentiable with respect to  $q$  at  $q_0$  and if the differential is an isomorphism:  $(D_q \mathcal{R})(u; q_0) \in \mathcal{L}_{iso}(\mathcal{X}; \mathcal{Y}')$ . Here and in what follows,  $\mathcal{L}_{iso}$  shall denote the set of linear isomorphisms between the Banach space arguments.

We assume the map  $\mathcal{R}(u; \cdot) : \mathcal{X} \rightarrow \mathcal{Y}'$  admits a family of regular solutions *locally, in an open neighborhood of the nominal parameter instance*  $\langle u \rangle \in X$  so that the operator equations involving  $\mathcal{R}(u; q)$  are well-posed. For all  $u$  in a sufficiently small, closed neighborhood  $\tilde{X} \subseteq X$  of  $\langle u \rangle \in X$  we impose the following structural assumption on the parametric forward problem:

**Assumption 1.** *For every  $u \in \tilde{X} \subseteq X$ , we assume given maps  $A(u; q) : X \times \mathcal{X} \rightarrow \mathcal{Y}'$  and  $F(u) : X \rightarrow \mathcal{Y}'$  such that*

$$\mathcal{R}(u; q) = A(u; q) - F(u) \quad \text{in } \mathcal{Y}' . \quad (2.2)$$

We denote by  $\mathbf{a} : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$  the bilinear form associated with  $A$ . For every fixed  $u \in \tilde{X} \subset X$ , and for every  $F(u) \in \mathcal{Y}'$ , the problem to find  $q(u) \in \mathcal{X}$  such that the residual equation (2.1) is well-posed, i.e. there exists a unique solution  $q(u)$  of (2.1) which depends continuously on  $u$ .

The set  $\{(u, q(u)) : u \in \tilde{X}\} \subset X \times \mathcal{X}$  is called a *regular branch of solutions* of (2.1) if

$$\begin{aligned} \tilde{X} \ni u &\mapsto q(u) \text{ is continuous as mapping from } \tilde{X} \rightarrow \mathcal{X} , \\ \mathcal{R}(u; q(u)) &= 0 \quad \text{in } \mathcal{Y}' . \end{aligned} \quad (2.3)$$

The regular branch of solutions (2.3) is called *nonsingular* if, in addition, the differential

$$(D_q \mathcal{R})(u; q(u)) \in \mathcal{L}(\mathcal{X}, \mathcal{Y}') \text{ is an isomorphism from } \mathcal{X} \text{ onto } \mathcal{Y}' , \text{ for all } u \in \tilde{X} . \quad (2.4)$$

Well-known sufficient conditions for well-posedness of (2.1) are stated in the following proposition. More precisely, for regular branches of nonsingular solutions given by (2.1) - (2.4), the differential  $D_q \mathcal{R}$  satisfies the so-called *inf-sup conditions*.

**Proposition 2.1.** *Assume that  $\mathcal{Y}$  is reflexive and that, for some nominal value  $\langle u \rangle \in X$  of the uncertainty, the operator equation (2.1) with (2.2) admits a regular branch of solutions (2.3) with  $q_0 \in \mathcal{X}$  denoting the “nominal” solution corresponding to the data  $\langle u \rangle \in X$ . Then the differential  $D_q \mathcal{R}$  at  $(\langle u \rangle, q_0)$  given by the bilinear map*

$$\mathcal{X} \times \mathcal{Y} \ni (\varphi, \psi) \mapsto {}_{\mathcal{Y}'} \langle D_q \mathcal{R}(\langle u \rangle; q_0) \varphi, \psi \rangle_{\mathcal{Y}}$$

*is boundedly invertible, uniformly with respect to  $u \in \tilde{X}$  where  $\tilde{X} \subset X$  is an open neighborhood of the nominal instance  $\langle u \rangle \in X$  of the uncertain parameter. In particular, there exists a constant  $\kappa > 0$  such that there holds*

$$\begin{aligned} \forall u \in \tilde{X} : \quad & \inf_{0 \neq \varphi \in \mathcal{X}} \sup_{0 \neq \psi \in \mathcal{Y}} \frac{{}_{\mathcal{Y}'} \langle (D_q \mathcal{R})(u; q_0) \varphi, \psi \rangle_{\mathcal{Y}}}{\|\varphi\|_{\mathcal{X}} \|\psi\|_{\mathcal{Y}}} \geq \kappa > 0 , \\ & \inf_{0 \neq \psi \in \mathcal{Y}} \sup_{0 \neq \varphi \in \mathcal{X}} \frac{{}_{\mathcal{Y}'} \langle (D_q \mathcal{R})(u; q_0) \varphi, \psi \rangle_{\mathcal{Y}}}{\|\varphi\|_{\mathcal{X}} \|\psi\|_{\mathcal{Y}}} \geq \kappa > 0 , \end{aligned} \quad (2.5)$$

and

$$\forall u \in \tilde{X} : \quad \|(D_q \mathcal{R})(u; q_0)\|_{\mathcal{L}(\mathcal{X}, \mathcal{Y}')} = \sup_{0 \neq \varphi \in \mathcal{X}} \sup_{0 \neq \psi \in \mathcal{Y}} \frac{{}_{\mathcal{Y}'} \langle (D_q \mathcal{R})(u; q_0) \varphi, \psi \rangle_{\mathcal{Y}}}{\|\varphi\|_{\mathcal{X}} \|\psi\|_{\mathcal{Y}}} \leq \kappa^{-1} . \quad (2.6)$$

For every  $u \in \tilde{X} \subseteq X$ , under conditions (2.5) and (2.6), there exists a unique, regular solution  $q(u)$  of (2.2) which is uniformly bounded with respect to  $u \in \tilde{X}$  in the sense that there exists a constant  $C(F, \tilde{X}) > 0$  such that

$$\sup_{u \in \tilde{X}} \|q(u)\|_{\mathcal{X}} \leq C(F, \tilde{X}) . \quad (2.7)$$

The set  $\{(u, q(u)) : u \in \tilde{X}\} \subset \tilde{X} \times \mathcal{X}$  is a *regular branch of nonsingular solutions* when (2.5) - (2.7) hold.

If the nonlinear functional  $\mathcal{R}$  is Fréchet differentiable with respect to  $u$  and Fréchet differentiable with respect to  $q$  at every point of the regular branch  $\{(u, q(u)) : u \in \tilde{X}\} \subset \tilde{X} \times \mathcal{X}$ , then the mapping relating  $u$  to  $q(u)$  with the branch of nonsingular solutions is locally Lipschitz on  $\tilde{X}$ : i.e. there exists a Lipschitz constant  $L(F, \tilde{X})$  such that

$$\forall u, v \in \tilde{X} : \quad \|q(u) - q(v)\|_{\mathcal{X}} \leq L(F, \tilde{X}) \|u - v\|_X . \quad (2.8)$$

This follows from the identity  $(D_u q)(u) = -(D_q \mathcal{R})^{-1}(D_u \mathcal{R})$ , and from the isomorphism property  $(D_u \mathcal{R}_q)(\langle u \rangle; q_0) \in \mathcal{L}_{iso}(\mathcal{X}, \mathcal{Y}')$  which is implied by (2.5) and (2.6), and from the continuity of the differential  $D_q \mathcal{R}$  on the regular branch.

In what follows, we will place ourselves in the abstract setting (2.2) with uniformly continuously differentiable mapping  $\mathcal{R}(u; q)$  in a product of neighborhoods  $B_X(\langle u \rangle; R) \times B_{\mathcal{X}}(q(\langle u \rangle); R)$  of sufficiently small radius  $R > 0$ . The quantity  $q(\langle u \rangle) \in \mathcal{X}$  is the corresponding regular solution of (2.2) at the nominal input  $\langle u \rangle \in X$ .

## 2.2 Uncertainty parametrization

We shall be concerned with the particular case where  $u \in X$  is a random variable taking values in a subset  $\tilde{X}$  of the Banach space  $X$ . We assume that  $X$  is separable, infinite-dimensional, and admits an unconditional Schauder basis  $\{\psi_j\}_{j \geq 1} : X = \text{span}\{\psi_j : j \geq 1\}$ . Then, every  $u \in \tilde{X} \subset X$  can be parametrized in this basis, i.e.

$$u = u(\mathbf{y}) := \langle u \rangle + \sum_{j \geq 1} y_j \psi_j \quad \text{for some } \mathbf{y} = (y_j)_{j \geq 1} \in U . \quad (2.9)$$

Examples of representations (2.9) are Karhunen-Loève expansions (see, e.g., [42, 39, 44, 7]) or unconditional Schauder bases (see, e.g., [4]). Note that the representation (2.9) is not unique: rescaling  $y_j$  and  $\psi_j$  will not change  $u$ . We will assume, therefore, throughout what follows that *the sequence  $\{\psi_j\}_{j \geq 1}$  is such that  $U = [-1, 1]^{\mathbb{N}}$* . For any  $\mathbf{y} \in U$ , norm-convergence in  $X$  of the series (2.9) in  $X$  is implied by the *summability condition*

$$\sum_{j \geq 1} \|\psi_j\|_X < \infty . \quad (2.10)$$

Condition (2.10) will be assumed throughout in what follows.

To obtain convergence rate estimates for the discretization of the forward problem, we shall restrict uncertain inputs  $u$  to sets  $X_t \subset X$  of inputs  $u$  with “higher regularity” (measured in a smoothness scale  $\{X_t\}_{t \geq 0}$  with  $X = X_0 \supset X_1 \supset X_2 \supset \dots$ ), so that  $u \in X_t$  will imply in Assumption 1 that  $F(\cdot) \in \mathcal{Y}'_t$  and  $q(u) \in \mathcal{X}_t$ , with corresponding subspaces  $\mathcal{X}_t \subset \mathcal{X}$  and  $\mathcal{Y}'_t \subset \mathcal{Y}'$  with extra regularity from suitable scales.

We remark that  $u \in X_t$  in general corresponds to stronger decay of the  $\psi_j$  in (2.9) which is relevant for optimal convergence estimates for multi-level QMC discretizations. In the present paper, we consider only single-level algorithms. For  $u \in X$ , in (2.9) the  $\{\psi_j\}_{j \geq 1}$  are thus assumed scaled such that (2.10) is strengthened to

$$\mathbf{b} := \{\|\psi_j\|_X\}_{j \geq 1} \in \ell^p(\mathbb{N}) \text{ for some } 0 < p < 1. \quad (2.11)$$

We also introduce the subset<sup>1</sup>

$$U = \{\mathbf{y} \in [-1, 1]^{\mathbb{N}} : u(\mathbf{y}) := \langle u \rangle + \sum_{j \geq 1} y_j \psi_j \in \tilde{X}\}. \quad (2.12)$$

Once an unconditional Schauder basis  $\{\psi_j\}_{j \geq 1}$  has been chosen, every realization  $u \in X$  can be identified in a one-to-one fashion with the pair  $(\langle u \rangle, \mathbf{y})$  where  $\langle u \rangle$  denotes the *nominal instance* of the uncertain datum  $u$  and  $\mathbf{y}$  is the coordinate vector of the unique representation (2.9).

**Remark 2.1.** In what follows, by a slight abuse of notation, *we identify the subset  $U$  in (2.12) with the countable set of parameters from the infinite-dimensional parameter domain  $U \subseteq \mathbb{R}^{\mathbb{N}}$  without explicitly writing so.* The operator  $A(u; q)$  in (2.2) then becomes, via the parametric dependence  $u = u(\mathbf{y})$ , a parametric operator family  $A(u(\mathbf{y}); q)$  which we denote (with slight abuse of notation) by  $\{A(\mathbf{y}; q) : \mathbf{y} \in U\}$ , with the parameter set  $U = [-1, 1]^{\mathbb{N}}$  (again, we use in what follows this definition in place of the set  $U$  as defined in (2.12)). In the particular case that the parametric operator family is linear, we have  $A(\mathbf{y}; q) = A(\mathbf{y})q$  with  $A(\mathbf{y}) \in \mathcal{L}_{iso}(\mathcal{X}, \mathcal{Y})$ . We do not assume, however, that the maps  $q \mapsto A(\mathbf{y}; q)$  are linear in what follows, unless explicitly stated.

With this understanding, and under the assumptions (2.7) and (2.8), the operator equation (2.2) will admit, for every  $\mathbf{y} \in U$ , a unique solution  $q(\mathbf{y}; F)$ , which is, due to (2.7) and (2.8), uniformly bounded and depends Lipschitz continuously on the parameter sequence  $\mathbf{y} \in U$ : there holds

$$\sup_{\mathbf{y} \in U} \|q(\mathbf{y}; F)\|_{\mathcal{X}} \leq C(F, U). \quad (2.13)$$

If the local Lipschitz condition (2.8) holds, there exists a Lipschitz constant  $L > 0$  such that

$$\|q(\mathbf{y}; F) - q(\mathbf{y}'; F)\|_{\mathcal{X}} \leq L(F, U) \|u(\mathbf{y}) - u(\mathbf{y}')\|_X. \quad (2.14)$$

The Lipschitz constant  $L > 0$  in (2.14) is not, in general, equal to  $L(F, \tilde{X})$  in (2.8): it depends on the nominal instance  $\langle u \rangle \in X$  and on the choice of basis  $\{\psi_j\}_{j \geq 1}$ .

*Unless explicitly stated otherwise, throughout what follows, we shall identify  $q_0 = q(\mathbf{0}; F) \in \mathcal{X}$  in Proposition 2.1 with the solution of (2.1) at the nominal input  $\langle u \rangle \in X$ .*

## 2.3 Dimension truncation

For a *truncation dimension*  $s \in \mathbb{N}$ , denote the  $s$ -term truncation of parametric representation (2.9) of the uncertain datum  $u$  by  $u^s \in X$ . Dimension truncation is equivalent to setting  $y_j = 0$  for  $j > s$  in (2.9) and we denote by  $q^s(\mathbf{y})$  the solution of

<sup>1</sup>For QMC quadrature, ahead, we rescale this set to  $[-1/2, 1/2]^{\mathbb{N}}$



the corresponding parametric weak problem (2.20). Unique solvability of (2.20) implies  $q^s(\mathbf{y}) = q(\{y_1, y_2, \dots, y_s, 0, \dots\})$ . For  $\mathbf{y} \in U$ , define  $\mathbf{y}_{\{1:s\}} := (y_1, y_2, \dots, y_s, 0, 0, \dots)$ . Proposition 2.1 holds when  $u(\mathbf{y})$  is replaced by  $u^s(\mathbf{y})$ , with  $\kappa > 0$  in (2.5) independent of  $s$  for sufficiently large  $s$ .

Our estimation of the *dimension truncation error*  $q(\mathbf{y}) - q^s(\mathbf{y})$  relies on two assumptions: (i) We assume the  $p$ -summability (2.11) of the sequence  $\mathbf{b}$  given by  $b_j := \|\psi_j\|_X$  in (2.9). From the definition of the sequence  $\mathbf{b} = (b_j)_{j \geq 1}$  in (2.11), the condition is equivalent to  $\sum_{j \geq 1} b_j^p < \infty$ ; (ii) the  $b_j$  in (2.11) are enumerated so that

$$b_1 \geq b_2 \geq \dots \geq b_j \geq \dots . \quad (2.15)$$

Consider the  $s$ -term truncated problem: given  $u^s \in \tilde{X}$ ,

$$\text{find } q^s \in \mathcal{X} : \quad \mathcal{Y}' \langle \mathcal{R}(u^s; q^s), w \rangle_{\mathcal{Y}} = 0 \quad \forall w \in \mathcal{Y} . \quad (2.16)$$

**Proposition 2.2.** *Under Assumptions (2.10), (2.11), for every  $f \in \mathcal{Y}'$ , for every  $\mathbf{y} \in U$  and for every  $s \in \mathbb{N}$ , the parametric solution  $q^s(\mathbf{y})$  of the dimensionally truncated, parametric weak problem (2.20) with  $s$ -term truncated parametric expansion (2.9) satisfies, with  $b_j$  as defined in (2.11),*

$$\sup_{\mathbf{y} \in U} \|q(\mathbf{y}) - q^s(\mathbf{y})\|_{\mathcal{X}} \leq C(F, X) \sum_{j \geq s+1} b_j \quad (2.17)$$

for some constant  $C > 0$  independent of  $f$ . Moreover, for every  $G(\cdot) \in \mathcal{X}'$ , we have

$$|I(G(q)) - I(G(q^s))| \leq \tilde{C} \sum_{j \geq s+1} b_j, \quad (2.18)$$

where

$$I(G(q)) = \int_U G(q(\mathbf{y})) \, d\mathbf{y} \quad \text{and} \quad I(G(q^s)) = \int_{[-1,1]^s} G(q(y_1, \dots, y_s, 0, \dots)) \, dy_1 \cdots dy_s,$$

for some constant  $\tilde{C} > 0$  independent of  $s$ . In addition, if conditions (2.10), (2.11) and (2.15) hold, then in (2.17) and (2.18) holds

$$\sum_{j \geq s+1} b_j \leq \min\left(\frac{1}{1/p - 1}, 1\right) \left(\sum_{j \geq 1} b_j^p\right)^{1/p} s^{-(1/p-1)}. \quad (2.19)$$

*Proof.* Assumption 1 on well-posedness of the forward problem (2.1) uniformly for all  $u \in B_X(\langle u \rangle; R)$  and the basis property (2.9) of the sequence  $\{\psi_j\}$  imply that for sufficiently large  $s$ ,  $u^s \in \tilde{X}$  and therefore (2.16) admits a unique solution,  $q^s \in \mathcal{X}$ , for these  $s$ . The unique local solvability of (2.1) and of (2.16) implies that  $q^s(\mathbf{y}) = q(\mathbf{y}_{\{1:s\}})$  where we recall for  $\mathbf{y} \in U$  the notation  $\mathbf{y}_{\{1:s\}} = (y_1, \dots, y_s, 0, \dots)$ . From (2.14) we obtain

$$\begin{aligned} \|q(\mathbf{y}; F) - q^s(\mathbf{y}; F)\|_{\mathcal{X}} &= \|q(\mathbf{y}; F) - q(\mathbf{y}_{\{1:s\}}; F)\|_{\mathcal{X}} \leq L \sup_{\mathbf{y} \in U} \|u(\mathbf{y}) - u(\mathbf{y}_{\{1:s\}})\|_X \\ &= L \sup_{\mathbf{y} \in U} \left\| \sum_{j=s+1}^{\infty} y_j \psi_j \right\|_X \leq L \sum_{j=s+1}^{\infty} b_j, \end{aligned}$$

which is (2.17). The bound (2.18) follows from  $G(\cdot) \in \mathcal{X}'$  and from (2.17).

With  $\mathbf{b} \in \ell^p(\mathbb{N})$  and the assumption (2.15), Stechkin's lemma (see also [28, p. 3363]) implies (2.19).  $\square$

## 2.4 Petrov-Galerkin discretization

Assuming that the infinite-dimensional space  $X$  of uncertain inputs to admit an unconditional Schauder basis  $\{\psi_j\}_{j \in \mathbb{N}}$ , the uncertain parameter  $u \in X$  in (2.2) can be equivalently expressed as (2.9) which turns (2.2) into an *equivalent, deterministic, countably parametric operator equation*: given  $\mathbf{y} \in U$ , find  $q(\mathbf{y}) \in \mathcal{X}$  such that

$$\mathcal{R}(\mathbf{y}; q(\mathbf{y})) = 0 \quad \text{in } \mathcal{Y}' . \quad (2.20)$$

Based on the theory in [19, Chap. IV.3] and in [34], an error analysis of Galerkin discretizations of (2.20) for the approximation of regular branches of solutions of smooth, nonlinear forward problems (2.2) will be presented in this section. Building on this, in the next section we generalize the results [28, 30] of Quasi-Monte Carlo, Petrov-Galerkin approximation to direct and inverse problems for operator equations (2.20) with countably-parametric uncertain inputs.

To this end, we assume, as in [38, 13], that we are given two sequences  $\{\mathcal{X}^h\}_{h>0} \subset \mathcal{X}$  and  $\{\mathcal{Y}^h\}_{h>0} \subset \mathcal{Y}$  of finite dimensional subspaces which are dense in  $\mathcal{X}$  and in  $\mathcal{Y}$ , respectively. For the computational complexity analysis, we also assume the following *approximation properties*: there is a scale  $\{\mathcal{X}_t\}_{t \geq 0}$  of subspaces such that  $\mathcal{X}_{t'} \subset \mathcal{X}_t \subset \mathcal{X}_0 = \mathcal{X}$  for any  $0 < t < t' < \infty$  and such that, for  $0 < t \leq \bar{t}$  and  $0 < t' \leq \bar{t}'$ , and for  $0 < h \leq h_0$ , there hold

$$\forall v \in \mathcal{X}_t : \quad \inf_{v^h \in \mathcal{X}^h} \|v - v^h\|_{\mathcal{X}} \leq C_t h^t \|v\|_{\mathcal{X}_t} . \quad (2.21)$$

Typical examples of smoothness scales  $\{\mathcal{X}_t\}_{t \geq 0}$  and  $\{\mathcal{Y}_t\}_{t \geq 0}$  are furnished by the Sobolev scale  $\mathcal{X}_t = H^{1+t}(D)$  in smooth domains or by its weighted counterparts in polyhedra [32].

**Proposition 2.3.** *Assume that the subspace sequences  $\{\mathcal{X}^h\}_{h>0} \subset \mathcal{X}$  and  $\{\mathcal{Y}^h\}_{h>0} \subset \mathcal{Y}$  are stable, i.e. there exist  $\bar{\mu} > 0$  and  $h_0 > 0$  such that for every  $0 < h \leq h_0$ , there hold the uniform (with respect to  $\mathbf{y} \in U$ ) discrete inf-sup conditions*

$$\forall \mathbf{y} \in U : \quad \inf_{0 \neq v^h \in \mathcal{X}^h} \sup_{0 \neq w^h \in \mathcal{Y}^h} \frac{\mathcal{Y}' \langle (D_q \mathcal{R})(u(\mathbf{y}); q_0) v^h, w^h \rangle_{\mathcal{Y}}}{\|v^h\|_{\mathcal{X}} \|w^h\|_{\mathcal{Y}}} \geq \bar{\mu} > 0 , \quad (2.22)$$

$$\forall \mathbf{y} \in U : \quad \inf_{0 \neq w^h \in \mathcal{Y}^h} \sup_{0 \neq v^h \in \mathcal{X}^h} \frac{\mathcal{Y}' \langle (D_q \mathcal{R})(u(\mathbf{y}); q_0) v^h, w^h \rangle_{\mathcal{Y}}}{\|v^h\|_{\mathcal{X}} \|w^h\|_{\mathcal{Y}}} \geq \bar{\mu} > 0 . \quad (2.23)$$

Then, for every  $0 < h \leq h_0$  the Galerkin approximations: given  $\mathbf{y} \in U$ ,

$$\text{find } q^h(\mathbf{y}) \in \mathcal{X}^h : \quad \mathcal{Y}' \langle \mathcal{R}(\mathbf{y}; q^h(\mathbf{y})), w^h \rangle_{\mathcal{Y}} = 0 \quad \forall w^h \in \mathcal{Y}^h , \quad (2.24)$$

are uniquely defined and converge quasioptimally: there exists a constant  $C > 0$  such that for all  $\mathbf{y} \in U$

$$\|q(\mathbf{y}) - q^h(\mathbf{y})\|_{\mathcal{X}} \leq \frac{C}{\bar{\mu}} \inf_{0 \neq v^h \in \mathcal{X}^h} \|q(\mathbf{y}) - v^h\|_{\mathcal{X}} . \quad (2.25)$$

If  $q(\mathbf{y}) \in \mathcal{X}_t$  uniformly w.r.t.  $\mathbf{y}$  and if (2.21) holds, then

$$\|q(\mathbf{y}) - q^h(\mathbf{y})\|_{\mathcal{X}} \leq \frac{C}{\bar{\mu}} h^t \sup_{\mathbf{y} \in U} \|q(\mathbf{y})\|_{\mathcal{X}_t} . \quad (2.26)$$

In the ensuing QMC convergence analysis we shall also require error bounds for the dimensionally truncated parameter sequences.

**Corollary 2.1.** *Under the assumptions of Proposition 2.3, for sufficiently large truncation dimension  $s \in \mathbb{N}$ , for given  $\mathbf{y}_{\{1:s\}} \in U$  the dimensionally truncated Galerkin approximations*

$$\text{find } q^h(\mathbf{y}_{\{1:s\}}) \in \mathcal{X}^h : \quad \mathcal{Y}' \langle \mathcal{R}(\mathbf{y}_{\{1:s\}}; q^h(\mathbf{y}_{\{1:s\}})), w^h \rangle_{\mathcal{Y}} = 0 \quad \forall w^h \in \mathcal{Y}^h, \quad (2.27)$$

*admit unique solutions  $q^h(\mathbf{y}_{\{1:s\}}) \in \mathcal{X}^h$  which converge, as  $h \downarrow 0$ , quasioptimally to  $q(\mathbf{y}_{\{1:s\}}) \in \mathcal{X}$ , i.e. (2.25) and (2.26) hold with  $\mathbf{y}_{\{1:s\}}$  in place of  $\mathbf{y}$ , with the same constants  $C > 0$  and  $\bar{\mu}$  independent of  $s$  and of  $h$ .*

### 3 Bayesian Inverse UQ

The nonlinear, parametric problems considered in Section 2 were *forward problems*: for a single instance of the uncertain datum  $u \in X$ , and for given input data  $F$ , the quantity of interest was the parametric solution  $q(u)$ , or  $q(\mathbf{y})$  in terms of the parametrization (2.9). Often, however, also the corresponding *inverse problem* is of interest: given observational data  $\delta$ , predict a ‘‘most likely’’ value of a Quantity of Interest (‘QoI’ for short)  $\phi$  which, typically, is a continuously (Fréchet-)differentiable functional of the input  $u \in X$ .

#### 3.1 General setup

Following [44, 7, 41, 35, 36], we equip the space of uncertain inputs  $X$  and the space of solutions  $\mathcal{X}$  of the forward maps with norms  $\|\cdot\|_X$  and with  $\|\cdot\|_{\mathcal{X}}$ , respectively. We consider the abstract (possibly nonlinear) operator equation (2.2) where the system’s forcing  $F \in \mathcal{Y}'$  is allowed to depend on the uncertain input  $u$ .

The uncertain operator  $A(u; \cdot) \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$  is assumed to be boundedly invertible, at least locally for the uncertain input  $u$  sufficiently close to a nominal input  $\langle u \rangle \in X$ , i.e. for  $\|u - \langle u \rangle\|_X$  sufficiently small so that, for such  $u$ , the response of the forward problem (2.2) is uniquely defined. We define the *forward response map*, which maps a given uncertain input  $u$  and a given forcing  $F$  to the response  $q$  in (2.2) by

$$X \ni u \mapsto q(u) := G(u; F) : X \times \mathcal{Y}' \rightarrow \mathcal{X}.$$

We omit the dependence of the response on  $F$  and simply denote the dependence of the forward solution on the uncertain input as  $q(u) = G(u)$ . We assume that we are given an observation functional  $\mathcal{O}(\cdot) : \mathcal{X} \rightarrow Y$ , which denotes a *bounded linear observation operator* on the space  $\mathcal{X}$  of observed system responses in  $Y$ . Throughout the remainder of this paper, we assume that there is a finite number  $K$  of sensors, so that  $Y = \mathbb{R}^K$  with  $K < \infty$ . We equip  $Y = \mathbb{R}^K$  with the Euclidean norm, denoted by  $|\cdot|$ . Then  $\mathcal{O} \in \mathcal{L}(\mathcal{X}; Y) \simeq (\mathcal{X}^*)^K$ , i.e.  $\mathcal{O}(\cdot)$  is a  $K$ -dimensional vector of observation functionals  $\mathcal{O}(\cdot) = (o_k(\cdot))_{k=1}^K$ .

In this setting, we wish to predict *computationally* an expected (under the Bayesian posterior, defined below) system response of the QoI, conditional on given, noisy measurement data  $\delta$ . Specifically, we assume the data  $\delta$  to consist of observations of system responses in the data space  $Y$ , corrupted by additive observation noise, e.g. by a realization of a random variable  $\eta$  taking values in  $Y$  with law  $\mathbb{Q}_0$ . We assume the following form of observed data, composed of the observed system response and the additive noise  $\eta$

$$\delta = \mathcal{O}(G(u)) + \eta \in Y. \quad (3.1)$$

We assume that the additive observation noise process  $\eta$  is Gaussian, i.e. a random vector  $\eta \sim \mathbb{Q}_0 \sim \mathcal{N}(0, \Gamma)$  with a positive definite covariance  $\Gamma$  on  $\mathbb{R}^K$  (i.e., a symmetric, positive definite covariance matrix  $\Gamma \in \mathbb{R}_{sym}^{K \times K}$  which we assume to be known). Henceforth, with a slight abuse of notation, we say  $\Gamma > 0$  (which means that  $\Gamma$  is positive definite).

The *uncertainty-to-observation map*  $\mathcal{G} : X \rightarrow \mathbb{R}^K$  of the system is  $\mathcal{G} = \mathcal{O} \circ G$ , so that

$$\delta = \mathcal{G}(u) + \eta = (\mathcal{O} \circ G)(u) + \eta \in L_{\Gamma}^2(\mathbb{R}^K),$$

where  $L_{\Gamma}^2(\mathbb{R}^K)$  denotes random vectors taking values in  $Y = \mathbb{R}^K$  which are square integrable with respect to the Gaussian measure with covariance matrix  $\Gamma > 0$  on the finite-dimensional observation space  $Y = \mathbb{R}^K$ . Bayes' formula [44, 7] yields a density of the Bayesian posterior with respect to the prior whose negative log-likelihood equals the observation noise covariance-weighted, least squares functional (also referred to as “potential” in what follows)  $\Phi_{\Gamma} : X \times Y \rightarrow \mathbb{R}$  given by  $\Phi_{\Gamma}(u; \delta) = \frac{1}{2}|\delta - \mathcal{G}(u)|_{\Gamma}^2$ , i.e.

$$\Phi_{\Gamma}(u; \delta) = \frac{1}{2}|\delta - \mathcal{G}(u)|_{\Gamma}^2 := \frac{1}{2}((\delta - \mathcal{G}(u))^{\top} \Gamma^{-1} (\delta - \mathcal{G}(u))) . \quad (3.2)$$

In [44, 7], an infinite-dimensional version of Bayes' rule was shown to hold in the present setting. In particular, the local Lipschitz assumption (2.8) on the solutions' dependence on the data implies a corresponding Lipschitz dependence of the Bayesian potential (3.2) on  $u \in X$ . Bayes' Theorem states that, under appropriate continuity conditions on the uncertainty-to-observation map  $\mathcal{G} = (\mathcal{O} \circ G)(\cdot)$  and on the prior measure  $\pi_0$  on  $u \in X$ , for positive observation noise covariance in (3.2), the posterior  $\pi^{\delta}$  of  $u \in X$  given data  $\delta \in Y$  is absolutely continuous with respect to the prior  $\pi_0$ . The following result is a version of Bayes' theorem, from [7, Thm. 3.4].

**Theorem 3.1.** *Assume that the potential  $\Phi_{\Gamma} : X \times Y \rightarrow \mathbb{R}$  is, for fixed data  $\delta \in Y$ ,  $\pi_0$  measurable and that, for  $\mathbb{Q}_0$ -a.e. data  $\delta \in Y$  there holds*

$$Z := \int_X \exp(-\Phi_{\Gamma}(u; \delta)) \pi_0(du) > 0 .$$

*Then the conditional distribution of  $u|\delta$  ( $u$  given  $\delta$ ) exists and is denoted by  $\pi^{\delta}$ . It is absolutely continuous with respect to  $\pi_0$  and there holds*

$$\frac{d\pi^{\delta}}{d\pi_0}(u) = \frac{1}{Z} \exp(-\Phi_{\Gamma}(u; \delta)) . \quad (3.3)$$

In particular, then, the Radon-Nikodym derivative of the Bayesian posterior w.r.t. the prior measure admits a bounded density w.r.t. the prior  $\pi_0$  which we denote by  $\Theta$ , and which is given by (3.3).

## 3.2 Parametric Bayesian posterior

We parametrize the uncertain datum  $u$  in the forward equation (2.2) as in (2.9). Motivated by [35, 36], the basis for the presently proposed deterministic quadrature approaches for Bayesian estimation via the computational realization of Bayes' formula is a *parametric, deterministic representation* of the derivative of the posterior measure  $\pi^{\delta}$  with respect to the *uniform prior measure*  $\pi_0$  on the set  $U$  of coordinates in the uncertainty

parametrization (2.12). The prior measure  $\pi_0$  being uniform, we admit in (2.9) sequences  $\mathbf{y}$  which take values in the parameter domain  $U = [-1, 1]^{\mathbb{N}}$ . As explained in sections 3.1 and 2.2, we consider the parametric, deterministic forward problem in the probability space

$$(U, \mathcal{B}, \pi_0) . \quad (3.4)$$

We assume throughout what follows that the prior measure  $\pi_0$  on the uncertain input  $u \in X$ , parametrized in the form (2.9), is the uniform measure. With the parameter domain  $U$  as in (3.4), the parametric uncertainty-to-observation map  $\Xi : U \rightarrow Y = \mathbb{R}^K$  is given by

$$\Xi(\mathbf{y}) = \mathcal{G}(u) \Big|_{u=\langle u \rangle + \sum_{j \in \mathbb{J}} y_j \psi_j} . \quad (3.5)$$

Our QMC quadrature approach will be based on a parametric version of Bayes' Theorem 3.1, in terms of the uncertainty parametrization (2.9). To do so, we view  $U$  as the unit ball of all sequences in  $U$  with respect to the  $\ell^\infty$ -norm, i.e. the Banach space of bounded sequences taking values in  $U$ .

**Theorem 3.2.** *Assume that  $\Xi : U \rightarrow Y = \mathbb{R}^K$  is bounded and continuous. Then  $\pi^\delta(d\mathbf{y})$ , the distribution of  $\mathbf{y} \in U$  given data  $\delta \in Y$ , is absolutely continuous with respect to  $\pi_0(d\mathbf{y})$ , i.e. there exists a parametric density  $\Theta(\mathbf{y})$  such that*

$$\frac{d\pi^\delta}{d\pi_0}(\mathbf{y}) = \frac{1}{Z} \Theta(\mathbf{y}) \quad (3.6)$$

with  $\Theta(\mathbf{y})$  given by

$$\Theta(\mathbf{y}) = \exp(-\Phi_\Gamma(u; \delta)) \Big|_{u=\langle u \rangle + \sum_{j \in \mathbb{J}} y_j \psi_j} , \quad (3.7)$$

with Bayesian potential  $\Phi_\Gamma$  as in (3.2) and with normalization constant  $Z$  given by

$$Z = \mathbb{E}^{\pi_0}[1] = \int_U \Theta(\mathbf{y}) \pi_0(d\mathbf{y}) > 0 . \quad (3.8)$$

Bayesian estimation is concerned with the approximation of a ‘‘most likely’’ Quantity of Interest (QoI)  $\phi : X \rightarrow \mathbb{R}$  conditional on given (noisy) observation data  $\delta \in Y$ . With the QoI  $\phi$  we associate the deterministic, infinite-dimensional, parametric map

$$\begin{aligned} \Psi(\mathbf{y}) &= \Theta(\mathbf{y}) \phi(q(u)) \Big|_{u=\langle u \rangle + \sum_{j \in \mathbb{J}} y_j \psi_j} \\ &= \exp(-\Phi_\Gamma(u; \delta)) \phi(q(u)) \Big|_{u=\langle u \rangle + \sum_{j \in \mathbb{J}} y_j \psi_j} : U \rightarrow \mathbb{R} . \end{aligned} \quad (3.9)$$

Then the Bayesian estimate of the QoI  $\phi$ , given noisy data  $\delta$ , takes the form

$$\mathbb{E}^{\pi^\delta}[\phi] = Z'/Z, \quad Z' := \int_U \Psi(\mathbf{y}) \pi_0(d\mathbf{y}) . \quad (3.10)$$

The task in computational Bayesian estimation is therefore to approximate the ratio  $Z'/Z \in \mathbb{R}$  in (3.10). In the parametrization with respect to  $\mathbf{y} \in U$ ,  $Z$  and  $Z'$  take the form of infinite-dimensional, iterated integrals with respect to the uniform prior  $\pi_0(d\mathbf{y})$ .

### 3.3 Well-posedness and approximation

For the computational viability of Bayesian inversion the quantity  $\mathbb{E}^{\pi^\delta}[\phi]$  should be stable under perturbations of the data  $\delta$  and under changes in the forward problem stemming, for example, from discretizations as considered in Section 2.4.

Unlike deterministic inverse problems where the data-to-solution maps can be severely ill-posed, for  $\Gamma > 0$  the expectations (3.10) are Lipschitz continuous with respect to the data  $\delta$ , *provided that the potential  $\Phi_\Gamma$  in (3.2) is locally Lipschitz with respect to the data  $\delta$  in the following sense.*

**Assumption 2.** [7, Assumption 4.2] *Let  $\tilde{X} \subseteq X$  and assume  $\Phi_\Gamma \in C(\tilde{X} \times Y; \mathbb{R})$  is Lipschitz on bounded sets. Assume also that there exist functions  $\mathcal{M}_i : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ ,  $i = 1, 2$ , (depending on  $\Gamma > 0$ ) which are monotone, non-decreasing separately in each argument, and with  $\mathcal{M}_2$  strictly positive, such that for all  $u \in \tilde{X}$ , and for all  $\delta, \delta_1, \delta_2 \in B_Y(0, r)$*

$$\Phi_\Gamma(u; \delta) \geq -\mathcal{M}_1(r, \|u\|_X), \quad (3.11)$$

and

$$|\Phi_\Gamma(u; \delta_1) - \Phi_\Gamma(u; \delta_2)| \leq \mathcal{M}_2(r, \|u\|_X) \|\delta_1 - \delta_2\|_Y. \quad (3.12)$$

We remark that in the present context, (3.12) follows from (3.2); for convenient reference, we include (3.12) in Assumption 2. Under Assumption 2, the expectation (3.10) depends Lipschitz continuously on  $\delta$  (see, e.g., [7, Sec. 4.1] for a proof):

$$\forall \phi \in L^2(\pi^{\delta_1}, X; \mathbb{R}) \cap L^2(\pi^{\delta_2}, X; \mathbb{R}) : \quad |\mathbb{E}^{\pi^{\delta_1}}[\phi] - \mathbb{E}^{\pi^{\delta_2}}[\phi]| \leq C(\Gamma, r) \|\delta_1 - \delta_2\|_Y. \quad (3.13)$$

Ahead, we shall be interested in the impact of approximation errors in the forward response of the system (e.g. due to discretization and approximate numerical solution of system responses) on the Bayesian predictions (3.10). To ensure continuity of the expectations (3.10) w.r.t. changes in the potential, we impose the following assumption.

**Assumption 3.** [7, Assumption 4.6] *Let  $\tilde{X} \subseteq X$  and assume  $\Phi_\Gamma \in C(\tilde{X}; \mathbb{R})$  is Lipschitz on bounded sets. Assume also that there exist functions  $\mathcal{M}_i : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ ,  $i = 1, 2$ , independent of the number  $M$  of degrees of freedom in the discretization of the forward problem, where the functions  $\mathcal{M}_i$  are monotonically non-decreasing separately in each argument, and with  $\mathcal{M}_2$  strictly positive, such that for all  $u \in \tilde{X}$ , and all  $\delta \in B_Y(0, r)$ ,*

$$\Phi_\Gamma(u; \delta) \geq -\mathcal{M}_1(\|u\|_X), \quad (3.14)$$

and there is a positive, monotonically decreasing  $\varphi(\cdot)$  such that  $\varphi(M) \rightarrow 0$  as  $M \rightarrow \infty$ , monotonically and uniformly w.r.t.  $u \in \tilde{X}$  (resp. w.r.t.  $\mathbf{y} \in U$ ) and such that

$$|\Phi_\Gamma(u; \delta) - \Phi_\Gamma^M(u; \delta)| \leq \mathcal{M}_2(\|u\|_X) \varphi(M). \quad (3.15)$$

Denote by  $\pi_M^\delta$  the Bayesian posterior, for given data  $\delta \in Y$ , with respect to the approximate potential  $\Phi_\Gamma^M$  obtained from a Petrov-Galerkin discretization (2.27).

**Proposition 3.1.** *Suppose that Assumption 3 holds, and assume that for  $\tilde{X} \subseteq X$  and for some bounded set  $B \subset X$  we have  $\pi_0(\tilde{X} \cap B) > 0$  and*

$$X \ni u \mapsto \exp(\mathcal{M}_1(\|u\|_X))(1 + \mathcal{M}_2^2(\|u\|_X)) \in L_{\pi_0}^1(X; \mathbb{R}).$$

Then there holds, for every QoI  $\phi : X \rightarrow \mathbb{R}$  such that  $\phi \in L^2_{\pi^\delta}(X; \mathbb{R}) \cap L^2_{\pi^{\delta_M}}(X; \mathbb{R})$  uniformly w.r.t.  $M$  and such that  $Z > 0$  in (3.8), the consistency error bound

$$\|\mathbb{E}^{\pi^\delta}[\phi] - \mathbb{E}^{\pi^{\delta_M}}[\phi]\|_{\mathbb{R}} \leq C(\Gamma, r)\varphi(M). \quad (3.16)$$

For a proof of Proposition 3.1, we refer to [7, Thm. 4.8, Rem. 4.9].

Below, we shall present concrete choices for the convergence rate function  $\varphi(M)$  in estimate (3.15), in terms of i) “*dimension truncation*” of the uncertainty parametrization (2.9), i.e. to a finite number of  $s \geq 1$  terms in (2.9), and ii) *discretization* of the dimensionally truncated problem for particular classes of forward problems. The verification of the consistency condition (3.15) in either of these cases will be based on the following observation.

**Proposition 3.2.** *Assume we are given a sequence  $\{q_M\}_{M \geq 1}$  of approximations to the forward response  $X \ni u \mapsto q(u) \in \mathcal{X}$  such that, with the parametrization (2.9),*

$$\sup_{u \in \tilde{X}} \|(q - q_M)(\mathbf{y})\|_{\mathcal{X}} \leq \varphi(M) \quad (3.17)$$

with a consistency error bound  $\varphi(M) \downarrow 0$  as in Assumption 3. Denote by  $G^M$  the corresponding (Galerkin) approximations of the parametric forward maps. Then the approximate Bayesian potential is

$$\Phi_\Gamma^M(u; \delta) = \frac{1}{2}(\delta - \mathcal{G}^M(u))^\top \Gamma^{-1}(\delta - \mathcal{G}^M(u)) : X \times Y \rightarrow \mathbb{R}, \quad (3.18)$$

where  $\mathcal{G}^M := \mathcal{O} \circ G^M$ , satisfies (3.15).

*Proof.* By definition (3.2) of the Bayesian potential, for every  $u \in \tilde{X}$  (i.e. every  $\mathbf{y} \in U$  defined in (2.12))

$$\begin{aligned} & |\Phi_\Gamma(u; \delta) - \Phi_\Gamma^M(u; \delta)| \\ &= \frac{1}{2} \left| (\delta - (\mathcal{O} \circ G)(u))^\top \Gamma^{-1}(\delta - (\mathcal{O} \circ G)(u)) - (\delta - (\mathcal{O} \circ G^M)(u))^\top \Gamma^{-1}(\delta - (\mathcal{O} \circ G^M)(u)) \right| \\ &\leq \frac{1}{2} \|\Gamma^{-1/2} \mathcal{O}\|_{\mathcal{X}^*} \|(G - G^M)(u)\|_{\mathcal{X}} |2\delta - \mathcal{O} \circ (G + G^M)(u)|_\Gamma \\ &= \frac{1}{2} \|\Gamma^{-1/2} \mathcal{O}\|_{\mathcal{X}^*} \|(q - q_M)(\mathbf{y})\|_{\mathcal{X}} |2\delta - \mathcal{O} \circ (q + q_M)(\mathbf{y})|_\Gamma. \end{aligned}$$

□

Throughout the following, we will denote by  $q_M$  the Petrov-Galerkin discretization (2.27) in Section 2.4, with  $M = M_h = \dim(\mathcal{X}^h) = \dim(\mathcal{Y}^h)$  degrees of freedom.

## 4 Holomorphic parameter dependence

As indicated, a key role in the present paper is played by holomorphy of countably parametric families of operator equations and their solution families. By this we mean that the parametric family of solutions permits, with respect to each parameter  $y_j$ , a holomorphic extension into the complex domain  $\mathbb{C}$ ; for purposes of QMC integration, in

addition some uniform bounds on these holomorphic extensions must be satisfied in order to prove approximation rates and QMC quadrature error bounds which are independent of the number of parameters which are “activated” in the approximation, i.e. in the QMC quadrature process. In [23, 3], the notion of  $(\mathbf{b}, p, \varepsilon)$ -holomorphy of parametric solutions has been introduced to this end. *In the remainder of Section 4 and throughout the next Section 5.1, all spaces  $X, Y, \mathcal{X}$  and  $\mathcal{Y}$  will be understood as Banach spaces over  $\mathbb{C}$ , without notationally indicating so.*

## 4.1 Holomorphic Families of Countably-Parametric Maps

**Definition 4.1.** ( $(\mathbf{b}, p, \varepsilon)$ -holomorphy) Let  $\varepsilon > 0$  and  $0 < p < 1$  be given. For a positive sequence  $\mathbf{b} = (b_j)_{j \geq 1} \in \ell^p(\mathbb{N})$ , we say that a parametric solution family  $q(\mathbf{y}) : U \rightarrow \mathcal{X}$  of (2.2) satisfies the  $(\mathbf{b}, p, \varepsilon)$ -holomorphy assumption if and only if all of the following conditions hold:

1. The map  $\mathbf{y} \mapsto q(\mathbf{y})$  from  $U$  to  $\mathcal{X}$ , for each  $\mathbf{y} \in U$ , is uniformly bounded with respect to the parameter sequence  $\mathbf{y}$ , i.e. there is a bound  $C_0 > 0$  such that

$$\sup_{\mathbf{y} \in U} \|q(\mathbf{y})\|_X \leq C_0. \quad (4.1)$$

2. There exists a positive sequence  $\mathbf{b} = (b_j)_{j \geq 1} \in \ell^p(\mathbb{N})$  such that, for any sequence  $\boldsymbol{\rho} := (\rho_j)_{j \geq 1}$  of numbers  $\rho_j > 1$  that satisfies

$$\sum_{j \geq 1} (\rho_j - 1) b_j \leq \varepsilon, \quad (4.2)$$

for sufficiently small  $\varepsilon > 0$ , the parametric solution map  $U \ni \mathbf{y} \mapsto q(\mathbf{y})$  admits an extension  $\mathbf{z} \mapsto q(\mathbf{z})$  to the complex domain that is holomorphic with respect to each variable  $z_j$  in a cylindrical set of the form  $O_{\boldsymbol{\rho}} := \bigotimes_{j \geq 1} O_{\rho_j}$  where, for every  $j \geq 1$ ,  $[-1, 1] \subset O_{\rho_j} \subset \mathbb{C}$  with open sets  $O_{\rho_j} \subset \mathbb{C}$ .

3. For any poly-radius  $\boldsymbol{\rho}$  satisfying (4.2), there is a second family  $\tilde{O}_{\boldsymbol{\rho}} := \bigotimes_{j \geq 1} \tilde{O}_{\rho_j}$  of open, cylindrical sets

$$[-1, 1] \subset O_{\rho_j} \subset \tilde{O}_{\rho_j} \subset \mathbb{C}$$

(strict inclusions), such that the extension is bounded on  $\overline{\tilde{O}_{\boldsymbol{\rho}}}$  according to

$$\sup_{\mathbf{z} \in \tilde{O}_{\boldsymbol{\rho}}} \|q(\mathbf{z})\|_X \leq C_{\varepsilon}, \quad (4.3)$$

where the bounds  $C_{\varepsilon} > 0$  depend on  $\varepsilon$ , but are independent of  $\boldsymbol{\rho}$ .

The notion of  $(\mathbf{b}, p, \varepsilon)$ -holomorphy depends implicitly on the choice of sets  $\tilde{O}_{\rho_j}$ . Depending on the approximation process in the parameter domain  $U$  under consideration, a particular choice of the sets  $\tilde{O}_{\rho_j}$  has to be made in order to obtain sharp convergence bounds under minimal holomorphy requirements.

Some particular choices of sets  $\tilde{O}_{\rho_j}$  are as follows. For a real number  $\kappa > 1$ , we denote by  $\mathcal{D}_{\kappa} = \{z \in \mathbb{C} \mid |z| \leq \kappa\}$  the closed disc of radius  $\kappa > 1$  in  $\mathbb{C}$ . Choosing  $\tilde{O}_{\rho_j} = \mathcal{D}_{\rho_j}$  implies, for example,  $p$ -summability of partial sums of finitely truncated Taylor



expansions of  $q(\mathbf{y})$  about  $\mathbf{y} = 0$ , cf. [3] and the references there. This choice arises also in connection with Chebyshev approximation, cf. [23]. The existence of analytic continuations to polydiscs arises naturally in the context of *affine parametric operator equations* which were considered in [13].

A second choice is the *Bernstein ellipse* in the complex plane, which is defined for some  $\kappa > 1$  by (cp. [8, Sec. 1.13])

$$\tilde{O}_\kappa = \mathcal{E}_\kappa := \left\{ \frac{w + w^{-1}}{2} : w \in \mathbb{C}, 1 \leq |w| \leq \kappa \right\}, \quad (4.4)$$

which has foci at  $\pm 1$  and semi axes of length  $\frac{1}{2}(\kappa + 1/\kappa)$  and  $\frac{1}{2}(\kappa - 1/\kappa)$ . This class of holomorphy domains is well known to afford sharp results on the convergence of approximations by truncated Legendre series, cf. [8, 3, 6]. Importantly, this class of domains can be obtained by *local* analytic continuation into discs of radii larger than  $\rho_j - 1$  of the parametric mapping  $q(\mathbf{y})$  about a finite number of points  $y_m \in [-1, 1]$  (cp. [3]).

For the derivative bounds which arise in connection with higher order Quasi-Monte Carlo error analysis (see, e.g., [13, 28]) we use the following family of smaller continuation domains, as in [15]: for  $\kappa > 1$ , consider by  $\mathcal{T}_\kappa$  the set of points

$$\mathcal{T}_\kappa = \{z \in \mathbb{C} | \text{dist}(z, [-1, 1]) \leq \kappa - 1\} = \bigcup_{-1 \leq y \leq 1} \{z \in \mathbb{C} | |z - y| \leq \kappa - 1\} \subset \mathbb{C}. \quad (4.5)$$

Then, once more, for a poly-radius  $\boldsymbol{\rho}$  satisfying (4.2), we denote by  $\mathcal{T}_\boldsymbol{\rho}$  the corresponding cylindrical set  $\mathcal{T}_\boldsymbol{\rho} := \bigotimes_{j \geq 1} \mathcal{T}_{\rho_j} \subset \mathbb{C}^{\mathbb{N}}$ .

## 4.2 Examples of $(\mathbf{b}, p, \varepsilon)$ -holomorphic families

We present several classes of examples of parametric equations  $\mathcal{R}(u; q)$  in (2.2) for which  $(\mathbf{b}, p, \varepsilon)$ -holomorphy can be verified. To this end, we denote by  $\mathcal{X}$  and  $\mathcal{Y}$  separable and reflexive Banach spaces over  $\mathbb{R}$  (all results will hold with the obvious modifications also for spaces over  $\mathbb{C}$ ) with (topological) duals  $\mathcal{X}'$  and  $\mathcal{Y}'$ , respectively. By  $\mathcal{L}(\mathcal{X}, \mathcal{Y}')$ , we denote the set of bounded linear operators  $A : \mathcal{X} \rightarrow \mathcal{Y}'$ . We next consider parametric forward models and the regularity of their (countably-) parametric solution families. The following result, proved in [3], shows that holomorphy of the solution map  $\mathbf{y} \mapsto q(\mathbf{y})$  follows from the holomorphy of the maps  $A$  and  $F$  in (2.2). Ahead, in Section 5.2, it will be seen to imply holomorphy of the parametric Bayesian posterior which, in turn, will be seen in Section 5.3 to yield dimension independent convergence rates for higher order QMC quadrature approximations of  $Z$  and  $Z'$  in the Bayesian estimate (3.10).

**Theorem 4.1.** *For  $\varepsilon > 0$  and  $0 < p < 1$ , assume that there exist a positive sequence  $\mathbf{b} = (b_j)_{j \geq 1} \in \ell^p(\mathbb{N})$ , and two constants  $0 < r \leq R < \infty$  independent of  $u \in \tilde{X}$  such that the following holds:*

1. *For any sequence  $\boldsymbol{\rho} := (\rho_j)_{j \geq 1}$  of numbers strictly greater than 1 that satisfies (4.2), the parametric maps  $\mathbf{a}(\mathbf{y}; \cdot, \cdot)$  (corresponding to the linear operator  $A(\mathbf{y}) \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$ ) and  $F$  admit extensions to complex parameters that are holomorphic with respect to every variable  $\mathbf{z}$  on a set of the form  $O_\boldsymbol{\rho} = \bigotimes_{j \geq 1} O_{\rho_j}$ , where  $O_{\rho_j} \subset \mathbb{C}$  is an open set containing  $\tilde{O}_{\rho_j}$ .*

2. These extensions satisfy for all  $\mathbf{z} \in O_\rho$  the uniform continuity conditions

$$\sup_{w \in \mathcal{Y} \setminus \{0\}} \frac{|F(\mathbf{z}; w)|}{\|w\|_{\mathcal{Y}}} \leq R, \quad \sup_{v \in \mathcal{X} \setminus \{0\}, w \in \mathcal{Y} \setminus \{0\}} \frac{|\mathbf{a}(\mathbf{z}; v, w)|}{\|v\|_{\mathcal{X}} \|w\|_{\mathcal{Y}}} \leq R, \quad (4.6)$$

and the uniform inf-sup conditions: there exists  $r > 0$  such that for every  $\mathbf{z} \in O_\rho$  there hold the uniform inf-sup conditions

$$\inf_{v \in \mathcal{X} \setminus \{0\}} \sup_{w \in \mathcal{Y} \setminus \{0\}} \frac{|\mathbf{a}(\mathbf{z}; v, w)|}{\|v\|_{\mathcal{X}} \|w\|_{\mathcal{Y}}} \geq r \quad \text{and} \quad \inf_{w \in \mathcal{Y} \setminus \{0\}} \sup_{v \in \mathcal{X} \setminus \{0\}} \frac{|\mathbf{a}(\mathbf{z}; v, w)|}{\|v\|_{\mathcal{X}} \|w\|_{\mathcal{Y}}} \geq r. \quad (4.7)$$

Then, the nonlinear, parametric residual operator  $\mathcal{R}(u(\mathbf{z}); q) = A(u(\mathbf{z}); q) - F(u(\mathbf{z}); q)$  in (2.2) (where  $\mathbf{z} \in O_\rho$ ) satisfies the  $(\mathbf{b}, p, \varepsilon)$ -holomorphy assumptions with the same  $p$  and  $\varepsilon$  and with the same sequence  $\mathbf{b}$ .

We refer to Section 6 ahead for concrete examples.

## 5 Higher order QMC-PG method for Bayesian inverse problems

### 5.1 Regularity of parametric solutions

The dependence of the solution  $q(\mathbf{y})$  of the parametric, variational problem (2.2) on the parameter vector  $\mathbf{y}$  is studied in this subsection. Precise bounds on the growth of the partial derivatives of  $q(\mathbf{y})$  with respect to  $\mathbf{y}$  will be given. These bounds will, as in [28], imply dimension independent convergence rates for QMC quadratures.

In the following, let  $\mathbb{N}_0^{\mathbb{N}}$  denote the set of sequences  $\boldsymbol{\nu} = (\nu_j)_{j \geq 1}$  of nonnegative integers  $\nu_j$ , and let  $|\boldsymbol{\nu}| := \sum_{j \geq 1} \nu_j$ . For  $|\boldsymbol{\nu}| < \infty$ , we denote the partial derivative of order  $\boldsymbol{\nu}$  of  $q(\mathbf{y})$  with respect to  $\mathbf{y}$  by

$$\partial_{\mathbf{y}}^{\boldsymbol{\nu}} q(\mathbf{y}) := \frac{\partial^{|\boldsymbol{\nu}|}}{\partial y_1^{\nu_1} \partial y_2^{\nu_2} \dots} q(\mathbf{y}). \quad (5.1)$$

In [5, 28, 27], bounds on the derivatives (5.1) were obtained by an induction argument which strongly relied on affine-parametric dependence of the parametric operator.

Alternative bounds on  $\|(\partial_{\mathbf{y}}^{\boldsymbol{\nu}} q)(\mathbf{y})\|_{\mathcal{X}}$  based on complex variable methods from [6, 41, 35, 36, 3], which give rise to product weights at least for a finite (possibly large, but in general operator-dependent) “leading” dimension of the parameter space are derived in [15]. These bounds are based on a *holomorphic extension of the parametric integrand functions to the complex domain* (we add that not all PDE problems afford such extensions and refer to [25] for an example).

For the particular case of linear, countably affine-parametric operator families  $(\mathbf{b}, p, \varepsilon)$ -holomorphy as in Definition 4.1 holds on polydiscs  $\mathcal{D}_\rho$ .

We remark that the smaller tubes  $\mathcal{T}_\rho$  of holomorphy in (4.5) are, nevertheless, important: on the one hand, the ensuing result about  $(\mathbf{b}, p, \varepsilon)$ -holomorphic, countably parametric functions belonging to SPOD-weighted spaces of integrands admissible for QMC quadrature by higher order digital nets is stronger (being valid under weaker hypotheses). On the other hand, in certain cases the possibility of covering the parameter intervals  $[-1, 1]$  by a finite number of small balls (whose union is contained in a tube  $\mathcal{T}_{\rho_j}$  for radius

sufficiently close to  $\rho_j = 1$ ) is crucial to verify holomorphy of the parametric solution families for certain nonlinear operator equations, see for example [3, Sec. 5.2]. The following result is the main result from [15]; it shows that  $(\mathbf{b}, p, \varepsilon)$ -holomorphy with respect to the domains  $\mathcal{T}_{\rho_j} \subset \mathbb{C}$  for  $j \geq 1$  implies derivative bounds for higher order QMC integration with dimension-independent rates of convergence.

**Theorem 5.1.** *For every mapping  $q(\mathbf{y}) : U \rightarrow \mathcal{X}$  which is  $(\mathbf{b}, p, \varepsilon)$ -holomorphic on a polytube  $\mathcal{T}_{\boldsymbol{\rho}}$  of poly-radius  $\boldsymbol{\rho} = (\rho_j)_{j \geq 1}$  with  $\rho_j > 1$  satisfying (4.2), there exists a sequence  $\boldsymbol{\beta} \in \ell^p(\mathbb{N})$  (depending on the sequence  $\mathbf{b}$  in (4.2)) and a partition  $\mathbb{N} = E \cup E^c$  such that the parametric solution  $q(\mathbf{y})$  satisfies, for every  $\boldsymbol{\nu} \in \mathbb{N}_0^{\mathbb{N}}$  with  $|\boldsymbol{\nu}| < \infty$ , the bound*

$$\sup_{\mathbf{y} \in U} \|(\partial_{\mathbf{y}}^{\boldsymbol{\nu}} q)(\mathbf{y})\|_{\mathcal{X}} \leq C_{\varepsilon} \boldsymbol{\nu}_E! \prod_{j \in E} \beta_j^{\nu_j} \times |\boldsymbol{\nu}_{E^c}|! \prod_{j \in E^c} \beta_j^{\nu_j}. \quad (5.2)$$

Here,  $E = \{1, 2, \dots, J\}$  for some  $J = J(\mathbf{b}) < \infty$  depending on the sequence  $\mathbf{b}$  in (4.2), and for  $\boldsymbol{\nu} \in \mathbb{N}_0^{\mathbb{N}}$ , we set  $\boldsymbol{\nu}_E := \{\nu_j : j \in E\}$ . The sequence  $\boldsymbol{\beta} = (\beta_j)_{j \geq 1}$  satisfies  $\beta_j = 4\|\mathbf{b}\|_{\ell^1(\mathbb{N})}/\varepsilon$ , i.e. it is in particular independent of  $j$  for  $1 \leq j \leq J$ . Moreover,  $\beta_j \lesssim b_j$  for  $j > J$  with the implied constant depending only on  $J(\mathbf{b})$  and on  $\|\mathbf{b}\|_{\ell^1(\mathbb{N})}$ .

The derivative bounds (5.2) were deduced from  $(\mathbf{b}, p, \varepsilon)$ -holomorphy of the forward map  $U \ni \mathbf{y} \mapsto q(\mathbf{y})$ . The same argument immediately implies corresponding derivative estimates for bounded, linear observation functionals  $\mathcal{O}(\cdot)$  of the parametric solution. With the  $(\mathbf{b}, p, \varepsilon)$ -holomorphy of the posterior densities  $\Theta(\mathbf{y})$  and  $\Psi(\mathbf{y})$  in Proposition 5.1, corresponding bounds for the parametric integrand functions in the Bayesian estimate (3.10) follow analogously. We sum up these observations in the following.

**Corollary 5.1.** *Assume that the parametric solution map  $U \ni \mathbf{y} \mapsto q(\mathbf{y})$  of the forward problem is  $(\mathbf{b}, p, \varepsilon)$ -holomorphic. Then, for every bounded, linear observation functional  $\mathcal{O}(\cdot) : \mathcal{X} \rightarrow Y$ , the countably-parametric uncertainty-to-observation map  $G(\mathbf{y}) := \mathcal{O}(q(\mathbf{y})) : U \rightarrow Y$  satisfies the estimates (5.2): for given  $0 < \varepsilon < 1$  and  $\mathbf{b} \in \ell^p(\mathbb{N})$  there exist a constant  $C_{\varepsilon} > 0$ , a sequence  $\boldsymbol{\beta}(\varepsilon) \in \ell^p(\mathbb{N})$  and a partition  $\mathbb{N} = E \cup E^c$  depending only on  $\varepsilon$  and on  $\mathbf{b}$  such that for every  $\boldsymbol{\nu} \in \mathbb{N}_0^{\mathbb{N}}$  with  $|\boldsymbol{\nu}| < \infty$ , there holds*

$$\sup_{\mathbf{y} \in U} \|\partial_{\mathbf{y}}^{\boldsymbol{\nu}} G(\mathbf{y})\|_Y \leq C_{\varepsilon} \boldsymbol{\nu}_E! \prod_{j \in E} \beta_j^{\nu_j} \times |\boldsymbol{\nu}_{E^c}|! \prod_{j \in E^c} \beta_j^{\nu_j} \leq C_{\varepsilon} |\boldsymbol{\nu}|! \boldsymbol{\beta}^{\boldsymbol{\nu}}. \quad (5.3)$$

## 5.2 Holomorphy of Bayesian posterior $\Theta(\mathbf{y})$

Our verification of existence and  $(\mathbf{b}, p, \varepsilon)$ -holomorphy of analytic continuations  $\Theta(\mathbf{z})$  and  $\Psi(\mathbf{z})$  defined in (3.7) and (3.9), respectively, which appear in the parametric version (3.10) of Bayes' formula will be based on  $(\mathbf{b}, p, \varepsilon)$ -holomorphy of the forward map  $\mathbf{y} \mapsto q(\mathbf{y})$  for the parametric posterior density  $\Theta(\mathbf{y})$  defined in (3.6) and (3.7). Sufficient conditions for this were given in Theorem 4.1 above.

**Proposition 5.1.** [36, Thm. 4.1] *Consider the Bayesian inversion of the parametric operator equation (2.2) with uncertain input  $u \in X$ , parametrized by the sequence  $\mathbf{y} = (y_j)_{j \in \mathbb{J}} \in U$ . Assume further that the corresponding forward solution map  $U \ni \mathbf{y} \mapsto q(\mathbf{y})$  is  $(\mathbf{b}, p, \varepsilon)$ -holomorphic for some sequence  $\mathbf{b} \in \ell^p(\mathbb{N})$  for some  $0 < p < 1$  and some  $\varepsilon > 0$ .*

Then the Bayesian posterior densities  $\Theta(\mathbf{y})$  and  $\Psi(\mathbf{y})$  defined in (3.7) and (3.9), respectively, are, as a function of the parameter  $\mathbf{y}$ , likewise  $(\mathbf{b}, p, \varepsilon)$ -holomorphic, with the same  $p$  and the same  $\varepsilon$ .

The modulus of the holomorphic extension of the Bayesian posterior  $\Theta(\mathbf{y})$  over the polyellipse  $\mathcal{E}_\rho$  for a  $(\mathbf{b}, \varepsilon)$ -admissible poly-radius  $\rho$  as in (4.2) is bounded as

$$B_\varepsilon = \sup_{\mathbf{z} \in \partial \mathcal{E}_\rho} |\Theta(\mathbf{z})| \leq C \exp(\theta^2 \|\Gamma^{-1}\|), \quad (5.4)$$

with  $\Gamma > 0$  denoting the positive definite covariance matrix in the additive, Gaussian observation noise model (3.1). The constants  $\theta, C > 0$  in (5.4) depend on the condition number of the uncertainty-to-observation map  $\mathcal{G}(\cdot) = (\mathcal{O} \circ G)(\cdot)$  but are independent of  $\Gamma$  in (3.1). The densities  $\Theta(\mathbf{y}) : U \rightarrow \mathbb{R}$  and  $\Psi(\mathbf{y}) : U \rightarrow \mathbb{R}$  in the Bayesian estimate (3.10) also satisfy estimates (5.2), with norms taken in the respective spaces  $\mathbb{R}$  and  $\mathbb{R}$ .

### 5.3 Higher order quasi-Monte Carlo integration

In order to prove error bounds using QMC quadrature, we require the integrand to be smooth. A result on the smoothness of the  $(\mathbf{b}, p, \varepsilon)$ -holomorphic solution of families of nonlinear parametric operator equations with  $(\mathbf{b}, p, \varepsilon)$ -holomorphic operators has been shown in [15] and is restated above (Theorem 5.1). For our purposes it is important to obtain error bounds which are independent of the dimension  $s$ , where  $s$  is the *truncation dimension*, i.e. we consider (2.16) and its (Petrov-)Galerkin discretization (2.27). This means that we truncate the infinite sum in (2.9) to a finite number  $s \geq 1$  of terms and then estimate the resulting  $s$  dimensional integral to approximate the mathematical expectation of the random solutions in (3.8) and (3.10).

For a real-valued integrand function  $g \in C^0([0, 1]^s; \mathbb{R})$ , we consider the  $s$ -variate integration problem

$$I_s(g) := \int_{[0,1]^s} g(\mathbf{y}) \, d\mathbf{y}. \quad (5.5)$$

We approximate this integral by an equal weight QMC quadrature rule

$$Q_{N,s}(g) := \frac{1}{N} \sum_{n=0}^{N-1} g(\mathbf{y}_n), \quad (5.6)$$

where the quadrature points  $\mathbf{y}_0, \dots, \mathbf{y}_{N-1} \in [0, 1]^s$  are judiciously chosen. In the following we restate the necessary definitions and results from [13].

**Definition 5.1.** Let nonnegative integers  $\alpha, s \in \mathbb{N}$ , and real numbers  $1 \leq q \leq \infty$  and  $1 \leq r \leq \infty$  be given. Let  $\boldsymbol{\gamma} = (\gamma_{\mathbf{u}})_{\mathbf{u} \subset \mathbb{N}}$  be a collection of nonnegative real numbers called *weights*. For every  $s \in \mathbb{N}$ , assume that the integrand function  $g : [0, 1]^s \rightarrow \mathbb{R}$  has partial derivatives of orders up to  $\alpha$  with respect to each variable. Define  $0/0 := 0$  and  $a/0 := \infty$  for  $a > 0$ . The smoothness of the integrand function  $g$  in (5.5) is quantified by the unanchored Sobolev norm

$$\|g\|_{s,\alpha,\boldsymbol{\gamma},q,r} := \left( \sum_{\mathbf{u} \subseteq \{1:s\}} \left( \gamma_{\mathbf{u}}^{-q} \sum_{\mathbf{v} \subseteq \mathbf{u}} \sum_{\boldsymbol{\tau}_{\mathbf{u} \setminus \mathbf{v}} \in \{1:\alpha\}^{|\mathbf{u} \setminus \mathbf{v}|}} \int_{[0,1]^{|\mathbf{v}|}} \left| \int_{[0,1]^{s-|\mathbf{v}|}} (\partial_{\mathbf{y}}^{(\alpha_{\mathbf{v}}, \boldsymbol{\tau}_{\mathbf{u} \setminus \mathbf{v}}, \mathbf{0})} g)(\mathbf{y}) \, d\mathbf{y}_{\{1:s\} \setminus \mathbf{v}} \right|^q d\mathbf{y}_{\mathbf{v}} \right)^{r/q} \right)^{1/r}, \quad (5.7)$$

with the obvious modifications if  $q$  or  $r$  is infinite. Here  $\{1 : s\}$  is a shorthand notation for the set  $\{1, 2, \dots, s\}$ , and  $(\boldsymbol{\alpha}_{\mathbf{v}}, \boldsymbol{\tau}_{\mathbf{u} \setminus \mathbf{v}}, \mathbf{0})$  denotes a sequence  $\boldsymbol{\nu}$  with  $\nu_j = \alpha$  for  $j \in \mathbf{v}$ ,  $\nu_j = \tau_j$  for  $j \in \mathbf{u} \setminus \mathbf{v}$ , and  $\nu_j = 0$  for  $j \notin \mathbf{u}$ . Let  $\mathcal{W}_{s, \alpha, \gamma, q, r}$  denote the Banach space of all such functions  $g$  with finite norm.

Note that if  $\gamma_{\mathbf{u}} = 0$  for some set  $\mathbf{u}$  then the corresponding projection term

$$\sum_{\mathbf{v} \subseteq \mathbf{u}} \sum_{\boldsymbol{\tau}_{\mathbf{u} \setminus \mathbf{v}} \in \{1 : \alpha\}^{|\mathbf{u} \setminus \mathbf{v}|}} \int_{[0, 1]^{s - |\mathbf{v}|}} (\partial_{\mathbf{y}}^{(\boldsymbol{\alpha}_{\mathbf{v}}, \boldsymbol{\tau}_{\mathbf{u} \setminus \mathbf{v}}, \mathbf{0})} g)(\mathbf{y}) \, d\mathbf{y}_{\{1 : s\} \setminus \mathbf{v}}$$

has to be 0 for all  $g \in \mathcal{W}_{s, \alpha, \gamma, q, r}$ . The next theorem, from [13, Thm. 3.5], states an upper bound on the worst-case integration error in  $\mathcal{W}_{s, \alpha, \gamma, q, r}$  using a QMC rule based on a digital net.

**Proposition 5.2.** *Let nonnegative integers  $\alpha, s \in \mathbb{N}$  with  $\alpha > 1$ , and real numbers  $1 \leq q \leq \infty$  and  $1 \leq r \leq \infty$  be given. Let  $\boldsymbol{\gamma} = (\gamma_{\mathbf{u}})_{\mathbf{u} \subseteq \mathbb{N}}$  be a collection of weights. Let  $r' \geq 1$  be the Hölder conjugate of  $r$ , i.e.  $1/r + 1/r' = 1$ . Let  $b$  be prime,  $m \in \mathbb{N}$ , and let  $\mathcal{S} = \{\mathbf{y}_n\}_{n=0}^{b^m - 1}$  denote a digital net with generating matrices  $C_1, \dots, C_s \in \mathbb{Z}_b^{\alpha m \times m}$ . Then*

$$\sup_{\|g\|_{s, \alpha, \gamma, q, r} \leq 1} \left| \frac{1}{b^m} \sum_{n=0}^{b^m - 1} g(\mathbf{y}_n) - \int_{[0, 1]^s} g(\mathbf{y}) \, d\mathbf{y} \right| \leq e_{s, \alpha, \gamma, r'}(\mathcal{S}),$$

with

$$e_{s, \alpha, \gamma, r'}(\mathcal{S}) := \left( \sum_{\emptyset \neq \mathbf{u} \subseteq \{1 : s\}} \left( C_{\alpha, b}^{|\mathbf{u}|} \gamma_{\mathbf{u}} \sum_{\mathbf{k}_{\mathbf{u}} \in \mathfrak{D}_{\mathbf{u}}^*} b^{-\mu_{\alpha}(\mathbf{k}_{\mathbf{u}})} \right)^{r'} \right)^{1/r'}. \quad (5.8)$$

Here  $\mathfrak{D}_{\mathbf{u}}^*$  is the “dual net without 0 components” projected to the components in  $\mathbf{u}$ . Moreover, we have  $\mu_{\alpha}(\mathbf{k}_{\mathbf{u}}) = \sum_{j \in \mathbf{u}} \mu_{\alpha}(k_j)$  with

$$\mu_{\alpha}(k) := \begin{cases} 0 & \text{if } k = 0, \\ a_1 + \dots + a_{\min(\alpha, \rho)} & \text{if } k = \kappa_1 b^{a_1 - 1} + \dots + \kappa_{\rho} b^{a_{\rho} - 1} \text{ with} \\ & \kappa_i \in \{1, \dots, b - 1\} \text{ and } a_1 > \dots > a_{\rho} > 0, \end{cases} \quad (5.9)$$

and

$$C_{\alpha, b} := \max \left( \frac{2}{(2 \sin \frac{\pi}{b})^{\alpha}}, \max_{1 \leq z \leq \alpha - 1} \frac{1}{(2 \sin \frac{\pi}{b})^z} \right) \times \left( 1 + \frac{1}{b} + \frac{1}{b(b+1)} \right)^{\alpha - 2} \left( 3 + \frac{2}{b} + \frac{2b+1}{b-1} \right). \quad (5.10)$$

We are interested in the case where the integrand  $g(\mathbf{y})$  is a composition of a continuous, linear functional  $\mathcal{O}(\cdot) \in \mathcal{X}'$  with the (Petrov-)Galerkin approximation  $q_s^h(2\mathbf{y} - \mathbf{1})$  of the dimension-truncated, parametric and  $(\mathbf{b}, p, \varepsilon)$ -holomorphic, operator equation (2.1). For every  $s \in \mathbb{N}$ , the dimension-truncated integrand functions  $g(\mathbf{y}) := (\mathcal{O} \circ q_s)(\mathbf{y}_{\{1 : s\}})$  are  $(\mathbf{b}, p, \varepsilon)$ -holomorphic *uniformly w.r.t.*  $s \in \mathbb{N}$  (see Section 2.3). It follows from Theorem 5.1 that they satisfy the derivative estimates (5.2) uniformly w.r.t.  $s \in \mathbb{N}$ . In [13, Sec. 3] and [15, Prop. 4.1] we showed the following result on convergence rates of QMC quadratures based on higher order digital nets for functions  $g(\mathbf{y})$  which satisfy (5.2).

**Proposition 5.3.** *Let  $s \geq 1$  and  $N = b^m$  for  $m \geq 1$  be integers and  $b$  be prime. Let  $\boldsymbol{\beta} = (\beta_j)_{j \geq 1}$  be a sequence of positive numbers, and denote by  $\boldsymbol{\beta}_s = (\beta_j)_{1 \leq j \leq s}$  its truncation after  $s$  terms. Assume that*

$$\exists 0 < p \leq 1 : \quad \sum_{j=1}^{\infty} \beta_j^p < \infty . \quad (5.11)$$

Define, for  $0 < p \leq 1$  as in (5.11),

$$\alpha := \lfloor 1/p \rfloor + 1 . \quad (5.12)$$

Consider integrand functions  $g(\mathbf{y})$  whose mixed partial derivatives of order  $\alpha$  satisfy

$$\forall \mathbf{y} \in U \forall s \in \mathbb{N} \forall \boldsymbol{\nu} \in \{0, 1, \dots, \alpha\}^s : \quad |(\partial_{\mathbf{y}}^{\boldsymbol{\nu}} g)(\mathbf{y})| \leq c(g) \boldsymbol{\nu}_E! \prod_{j \in E} \beta_j^{\nu_j} \times |\boldsymbol{\nu}_{E^c}|! \prod_{j \in E^c} \beta_j^{\nu_j} \quad (5.13)$$

for some fixed integer  $J \in \mathbb{N}$  where  $E = \{1, 2, \dots, J\}$  and  $E^c = \mathbb{N} \setminus E$ , and where  $c(g) > 0$  is independent of  $\mathbf{y}$ ,  $s$  and of  $\boldsymbol{\nu}$ . Then, for every  $N \in \mathbb{N}$ , one can construct an interlaced polynomial lattice rule of order  $\alpha$  with  $N$  points using a fast component-by-component algorithm, using  $\mathcal{O}(\alpha(\min\{s, J\} + \alpha(s - J)_+) N \log N)$  operations, plus  $\mathcal{O}(\alpha^2(s - J)_+^2 N)$  update cost, plus  $\mathcal{O}(N + \alpha(s - J)_+ N)$  memory cost, where  $(w)_+ = \max\{0, w\}$ , such that there holds the error bound

$$\forall s, N \in \mathbb{N} : \quad |I_s(g) - Q_{N,s}(g)| \leq C_{\alpha, \beta, b, p} N^{-1/p} , \quad (5.14)$$

where  $C_{\alpha, \beta, b, p} < \infty$  is a constant independent of  $s$  and  $N$ .

**Remark 5.1.** It has been observed, see for instance [20], that in implementations of the CBC construction of good generating vectors components tend to repeat, resulting in rather nonuniform lower dimensional projections. To avoid this problem, [20] used a ‘‘pruned’’ CBC algorithm where components of the generating vector are forced to differ from each other. A theoretical justification of this algorithm was provided in [12]. If we use this modified algorithm in Proposition 5.3 then, provided that  $N > 2s$ , the upper bound (5.14) still remains valid albeit with a slightly larger constant: the constant  $C_{\alpha, b}$  is replaced by  $2C_{\alpha, b}$ , which is still independent of the dimension  $s$ .

**Remark 5.2.** The bound (5.13) in Theorem 5.1 was shown for functions defined on the domain  $[-1, 1]^{\mathbb{N}}$ . However, QMC theory uses the domain  $[0, 1]^s$ . The change from  $[-1, 1]$  to  $[0, 1]$  can be achieved by the simple linear transformation  $y \mapsto (y + 1)/2$ . Using (5.2) together with this change of variable in Proposition 5.3 increases the constant in (5.10) by a factor of at most  $2^\alpha$ . Thus, in order for the theory to apply to the integrands from Sections 2 and 5.1, we need to scale the Walsh constant  $C_{\alpha, b}$  in (5.10) by a factor  $2^\alpha$ .

If the function  $g$  satisfies (5.13), then its norm (5.7) with  $r = \infty$  and for any  $q$ , can be bounded by

$$\begin{aligned} \|g\|_{s, \alpha, \gamma, q, \infty} &\leq c \max_{\mathbf{u} \subseteq \{1:s\}} \gamma_{\mathbf{u}}^{-1} \sum_{\boldsymbol{\nu}_{\mathbf{u}} \in \{1:\alpha\}^{|\mathbf{u}|}} \boldsymbol{\nu}_{\mathbf{u} \cap E}! \prod_{j \in \mathbf{u} \cap E} (2^{\delta(\nu_j, \alpha)} \beta_j^{\nu_j}) |\boldsymbol{\nu}_{\mathbf{u} \cap E^c}|! \prod_{j \in \mathbf{u} \cap E^c} (2^{\delta(\nu_j, \alpha)} \beta_j^{\nu_j}) \\ &= c(g) \max_{\mathbf{u} \subseteq \{1:s\}} \gamma_{\mathbf{u}}^{-1} \sum_{\boldsymbol{\nu}_{\mathbf{u}} \in \{1:\alpha\}^{|\mathbf{u}|}} \boldsymbol{\nu}_{\mathbf{u} \cap E}! |\boldsymbol{\nu}_{\mathbf{u} \cap E^c}|! \prod_{j \in \mathbf{u}} (2^{\delta(\nu_j, \alpha)} \beta_j^{\nu_j}) , \end{aligned}$$

where  $\delta(\nu_j, \alpha)$  is 1 if  $\nu_j = \alpha$  and is 0 otherwise. To make  $\|g\|_{s,\alpha,\gamma,q,\infty} \leq c$ , we choose

$$\gamma_{\mathbf{u}} := \sum_{\nu_{\mathbf{u}} \in \{1:\alpha\}^{|\mathbf{u}|}} \nu_{\mathbf{u} \cap E}! |\nu_{\mathbf{u} \cap E^c}|! \prod_{j \in \mathbf{u}} (2^{\delta(\nu_j, \alpha)} \beta_j^{\nu_j}). \quad (5.15)$$

## 5.4 Combined QMC-PG Error Bound

We approximate the Bayesian estimate  $\mathbb{E}^{\pi^\delta}[\phi] = Z'/Z$  with the ratio estimator  $Z'_{N,s,h}/Z_{N,s,h}$  where

$$Z'_{N,s,h} = Q_{N,s}(\Theta^{s,h}(\cdot))\phi(q^{s,h}(\cdot)) \in \mathbb{R}, \quad (5.16)$$

$$Z_{N,s,h} = Q_{N,s}(\Theta^{s,h}(\cdot)) \in \mathbb{R}, \quad (5.17)$$

where  $\Theta^{s,h}$  is given by (cf.(3.7) and (3.18))

$$\Theta^{s,h}(\mathbf{y}) = \exp(-\Phi_{\Gamma}^M(u; \delta)) \Big|_{u=\langle u \rangle + \sum_{j=1}^s y_j \psi_j}, \quad (5.18)$$

and  $q^{s,h}(\mathbf{y}) := q^h(\mathbf{y}_{\{1:s\}})$  with  $q^h(\mathbf{y}_{\{1:s\}})$  being the Galerkin approximation of the forward problem with dimension-truncation to dimension  $s$  as defined in (2.27). We recall that  $M = M_h = \dim(\mathcal{X}^h) = \dim(\mathcal{Y}^h)$  (see Section 2.4). We are now in position to prove our main result: an error bound for the approximate Bayesian estimates which accounts for the errors of dimension truncation to finite dimension  $s$  given by (2.16), Petrov-Galerkin discretization (2.27) of (2.16) and, finally, the QMC integration (5.6) of the goal functional  $G(\cdot)$  evaluated at the dimensionally truncated solution, i.e. of the approximation  $Q_{N,s}(G(q^{s,h}))$  of  $I(G(q))$ . All implied constants in the error bounds of the ensuing theorem are independent of the truncation dimension  $s$ .

**Theorem 5.2.** *Suppose that Assumptions 1, 2 and 3 hold. Suppose further that the parametric forward problem (2.1) admits, for every  $\mathbf{y} \in U$ , a unique solution  $q(\mathbf{y})$  which belongs to the smoothness space  $\mathcal{X}_t$  for some  $t > 0$  affording the approximation property (2.21). Assume further the sparsity condition (2.11) for the uncertainty parametrization, and that conditions (2.15) hold. Then, for given positive definite observation noise covariance  $\Gamma$ , the normalization constant  $Z$  in (3.8) is positive.*

*Assume further that the QoI  $\phi : \mathcal{X} \rightarrow \mathbb{R}$  is Lipschitz continuous in an open ball in  $\mathcal{X}$  about the forward solution  $q(\langle u \rangle)$  at the nominal input  $\langle u \rangle \in X$  (resp. about the origin  $\mathbf{0} \in U$ ).*

*Then there exist  $N_0, s_0, h_0$  (depending on the bound  $r > 0$  on the size  $|\delta|$  of the data, and on the observation noise covariance  $\Gamma$  in Assumption 2), such that the QMC-PG approximation of the Bayesian estimate  $\mathbb{E}^{\pi^\delta}[\phi]$  in (3.10) obtained as approximate ratio estimate  $Z'_{N,s,h}/Z_{N,s,h}$  with QMC-PG approximations  $Z'_{N,s,h}$  and  $Z_{N,s,h}$  of the integrals  $Z'$  and  $Z$  in (3.10), satisfies, for  $N \geq N_0$ ,  $s \geq s_0$  and for  $h \leq h_0$ , the error bound*

$$\left| \mathbb{E}^{\pi^\delta}[\phi] - Z'_{N,s,h}/Z_{N,s,h} \right| \leq C(\Gamma, p) (h^t + s^{-(1/p-1)} + N^{-1/p}). \quad (5.19)$$

Before giving the proof, we remark that the constant  $C(\Gamma, p)$  depends on the observation noise covariance  $\Gamma > 0$  and on the summability exponent  $p$ , but is independent of the truncation dimension  $s$  and of the number  $N$  of QMC points. We refer to [37] for details.

We also remark that the ensuing proof does not require bounds (5.13) for (functionals of) the Petrov-Galerkin discretized solution and, therefore, the verification of  $h$ -independence of the domain of holomorphy of the PG approximation  $q^{s,h}(\mathbf{y})$  of the parametric forward solution is not necessary.

*Proof.* (of Theorem 5.2): Both integrands  $\Psi(\mathbf{y})$  and  $\Theta(\mathbf{y})$  in the definition (3.10) of  $Z$  and of  $Z'$  are  $(\mathbf{b}, p, \varepsilon)$ -holomorphic by Proposition 5.1. By Theorem 5.1, their approximations  $\Psi^{s,h}(\mathbf{y})$  and  $\Theta^{s,h}(\mathbf{y})$  satisfy the derivative estimates (5.2) and (5.3), uniformly with respect to the truncation dimension  $s$ , and the Petrov-Galerkin discretization parameter  $h$ : as dimension truncation amounts to restricting  $\mathbf{y}$  to  $\mathbf{y}_{\{1:s\}}$  (see Section 2.3), the QMC error bounds in Proposition 5.3 are applicable.

By the triangle inequality, we have

$$\begin{aligned} |Z'Z_{N,s,h} - Z'_{N,s,h}Z| &= |Z'Z_{N,s,h} - ZZ' + ZZ' - Z'_{N,s,h}Z| \\ &\leq |Z'||Z_{N,s,h} - Z| + |Z||Z' - Z'_{N,s,h}|. \end{aligned} \quad (5.20)$$

By condition (3.8), there is  $N_0, s_0 \in \mathbb{N}$  and  $h_0 > 0$  so that  $Z_{N,s,h} \geq Z/2$  for all  $N \geq N_0$ ,  $s \geq s_0$  and  $0 < h \leq h_0$ . Using this and (5.20) we obtain

$$\begin{aligned} \left| \frac{Z'}{Z} - \frac{Z'_{N,s,h}}{Z_{N,s,h}} \right| &= \frac{|Z'Z_{N,s,h} - Z'_{N,s,h}Z|}{|ZZ_{N,s,h}|} \leq \frac{|Z'||Z_{N,s,h} - Z|}{|ZZ_{N,s,h}|} + \frac{|Z' - Z'_{N,s,h}|}{|Z_{N,s,h}|} \\ &\leq \frac{2|Z'||Z_{N,s,h} - Z|}{|Z|^2} + \frac{2|Z' - Z'_{N,s,h}|}{|Z|}. \end{aligned}$$

With  $Z = I(\Theta)$  and  $Z_{N,s,h} = Q_{N,s}(\Theta^{s,h})$ , we bound the error  $|Z - Z_{N,s,h}|$  using Proposition 5.3. There exists a constant  $C > 0$  (independent of the parameter dimension  $s$ ) such that for all  $N$

$$|I(\Theta) - Q_{N,s}(\Theta)| \leq CN^{-1/p}.$$

With  $Z' = I(\Theta\phi(q))$  and  $Z'_{N,s,h}$  as in (5.16), we bound the combined error  $|Z' - Z'_{N,s,h}|$  incurred by dimension-truncation, QMC integration and Petrov-Galerkin discretization as follows:

$$\begin{aligned} |I(\Theta\phi(q)) - Q_{s,N}(\Theta\phi(q^{s,h}))| &\leq \|\Theta\|_\infty (|I(\phi(q)) - I(\phi(q^s))| + |I(\phi(q^s)) - Q_{N,s}(\phi(q^s))| \\ &\quad + |Q_{N,s}(\phi(q^s)) - Q_{N,s}(\phi(q^{s,h}))|). \end{aligned}$$

Here, the first term can be bounded using Proposition 2.2, and (2.18) and (2.19) yield

$$|I(\phi(q)) - I(\phi(q^s))| = |I(\phi(q)) - I_s(\phi(q))| \leq C(p)\|\phi\|_{\mathcal{L}(\mathcal{X},\mathbb{R})}s^{-(1/p-1)}.$$

The second term is bounded using Proposition 5.3 as

$$|I(\phi(q^s)) - Q_{N,s}(\phi(q^s))| = |I_s(\phi(q)) - Q_{N,s}(\phi(q^s))| \leq CN^{-1/p}.$$

The third term is estimated by the PG-discretization error

$$|Q_{N,s}(\phi(q^s)) - Q_{N,s}(\phi(q^{s,h}))| = |Q_{N,s}(\phi(q^s - q^{s,h}))| \leq \|\phi\|_{\mathcal{L}(\mathcal{X},\mathbb{R})} \sup_{\mathbf{y} \in U} \|q^s(\mathbf{y}) - q^{s,h}(\mathbf{y})\|_{\mathcal{X}},$$

which, in turn, is estimated by Proposition 2.3 as

$$|Q_{N,s}(\phi(q^s)) - Q_{N,s}(\phi(q^{s,h}))| \leq \frac{C}{\bar{\mu}} \|\phi\|_{\mathcal{L}(\mathcal{X},\mathbb{R})} h^t \sup_{\mathbf{y} \in U} \|q(\mathbf{y})\|_{\mathcal{X}_t}.$$

□



**Remark 5.3.** The preceding analysis used that  $\phi \in \mathcal{L}(\mathcal{X}, \mathbb{R}) = \mathcal{X}'$ . If the QoI  $\phi$  has more regularity, such as  $\phi \in \mathcal{L}(\mathcal{X}_{t'}, \mathbb{R})$  for some  $t' > 0$ , higher PG convergence rates  $O(h^{t+t'})$  in (5.19) are possible, provided that the adjoint of the differential  $(D_q \mathcal{R})(u; q_0)$  is boundedly invertible in suitable scales of spaces. We refer to [28, 30, 14] for a statement of results and proofs in the case of linear, affine-parametric forward problems.

## 6 Model Problems

We illustrate the hypotheses in Section 2, with a view towards the numerical experiments in Section 7 ahead, by a model, affine-parametric diffusion problem which was already considered in [6]. We emphasize that these model problems are selected to illustrate the preceding error analysis; the preceding theory is applicable to considerably more general problems.

### 6.1 Affine-parametric diffusion problem

To illustrate the preceding convergence estimates, we consider a linear elliptic diffusion equation on the physical domain  $D = (0, 1)^d$  in space dimension  $d = 1$  or  $d = 2$ , with uncertain diffusion coefficient  $u(x)$ , known source term  $f(x) \in L^2(D)$  and with homogeneous Dirichlet boundary conditions, i.e.,

$$-\nabla \cdot (u(x)\nabla q(x)) = f(x) \text{ in } D = (0, 1)^d, \quad q(x) = 0 \text{ for } x \in \partial D. \quad (6.1)$$

Problem (6.1) is a particular instance of the abstract setting introduced in Section 2.1, with  $\mathcal{X} = \mathcal{Y} = H_0^1(D)$  and with deterministic right-hand side  $F \in \mathcal{Y}'$  obtained by associating to  $f$  a continuous, linear functional  $F$  on  $\mathcal{X} = H_0^1(D)$ , i.e.

$$A(u; q) = -\nabla \cdot (u\nabla q) \in \mathcal{L}(H_0^1(D), H^{-1}(D)), \quad F(\cdot) = \langle f, \cdot \rangle.$$

We now parametrize the diffusion coefficient as in (2.9) with the  $s$ -dimensional parameter vector  $\mathbf{y} \in U_s = [-1, 1]^s$ , which we indicate by writing  $u(x, \mathbf{y})$ . Note that the QMC integration domain  $U_s = [-\frac{1}{2}, \frac{1}{2}]^s$  can be equivalently obtained by rescaling. For a nominal input  $\langle u \rangle \in X := L^\infty(D)$ , a finite truncation dimension  $s$  and a sequence  $(\psi_j)_{j \geq 1} \subset X$ , we consider affine-parametric uncertainties

$$u(\cdot, \mathbf{y}) = \langle u \rangle + \sum_{j=1}^s y_j \psi_j(\cdot). \quad (6.2)$$

We assume that the nominal coefficient  $\langle u \rangle$  is positive and bounded,

$$0 < \bar{u}_{\min} \leq \langle u \rangle \leq \bar{u}_{\max}. \quad (6.3)$$

We suppress in the following the explicit dependence on  $x$ . We also assume the *sparsity condition*  $(\|\psi_j\|_{L^\infty(D)})_j \in \ell^r(\mathbb{N})$  for some  $0 < r < 1$ . The parametric weak formulation reads: for  $\mathbf{y} \in U_s$  find  $q(\cdot, \mathbf{y}) \in H_0^1(D)$  such that for all  $v \in H_0^1(D)$  holds

$$\int_D u(x, \mathbf{y}) \nabla_x q(x, \mathbf{y}) \cdot \nabla_x v(x) \, dx = \int_D f(x) v(x) \, dx. \quad (6.4)$$

In [3], Problem (6.4), with  $U$  given by (6.2), was considered for complex parameter sequences  $\mathbf{z} = (z_j)_{j \geq 1}$  under the following assumption.

**Assumption 4** (Uniform Ellipticity). *There exist constants  $0 < u_- < u_+ < \infty$  such that for a.e.  $x \in D$  and for all  $\mathbf{z} \in \mathcal{U}_s = \{z_j \in \mathbb{C} : |z_j| \leq 1, 1 \leq j \leq s\}$ ,  $s \in \mathbb{N}$ , there holds*

$$0 < u_- \leq \Re u(\mathbf{z}) \leq u_+ < \infty. \quad (6.5)$$

We remind the reader that  $\langle u \rangle$  and  $\psi_j$  are assumed to be real-valued functions. The choice  $\mathbf{z} = \mathbf{0}$  in Assumption 4 and (6.3) imply  $u_- \leq \langle u \rangle \leq u_+$ . A sufficient condition for Assumption 4 to hold is (6.3) and the condition

$$\gamma := \left\| \frac{1}{\langle u \rangle} \sum_{j \geq 1} |\psi_j| \right\|_{L^\infty(D)} < 1. \quad (6.6)$$

The Lax–Milgram lemma implies, with (6.5), for every fixed  $\mathbf{z} \in \mathcal{U}_s$ , and for every  $s \in \mathbb{N}$  the existence of a unique solution to the variational problem (6.4). Symmetric Galerkin discretization as explained in Section 2.4, yields for each parameter instance  $\mathbf{y} \in U_s$  a unique, parametric Galerkin solution  $q^h(\mathbf{y}) \in \mathcal{X}^h \subset \mathcal{X} = H_0^1(D)$ . The family  $\{q^h(\mathbf{y}) : \mathbf{y} \in U_s\} \subset \mathcal{X}^h$  is uniformly bounded in  $\mathcal{X}$  with respect to  $s, h$  and  $\mathbf{y} \in U_s$ .

## 6.2 Parametric Regularity

The parametric solution family  $\mathbf{z} \rightarrow q(\mathbf{z})$  of the linear forward problem (6.1) with complex-parametric input (6.2) is, for any value of  $s$ , holomorphic in the sense of Definition 4.1. To verify this, we recall from [6, Eqn. (2.8)] the notion of  $\delta$ -admissibility of a poly-radius  $\boldsymbol{\rho} = (\rho_j)_{j \geq 1}$  with  $\rho_j > 1$ :  $\boldsymbol{\rho}$  is called  $\delta$ -admissible if there exists a  $\delta > 0$  such that

$$\sum_{j \geq 1} \rho_j |\psi_j(x)| \leq \langle u \rangle - \delta \quad \text{for almost all } x \in D. \quad (6.7)$$

As above,  $\delta$ -admissibility (6.7) of the poly-radius  $\boldsymbol{\rho}$  implies  $\boldsymbol{\rho}$ -weighted ellipticity

$$\gamma(\boldsymbol{\rho}) := \left\| \frac{1}{\langle u \rangle} \sum_{j \geq 1} \rho_j |\psi_j| \right\|_{L^\infty(D)} < 1. \quad (6.8)$$

Under (6.6), the sequence  $\rho_j = 1$  is  $\delta$ -admissible for some  $\delta > 0$  and (6.8) and (6.6) coincide.

It was shown in [6, Lemma 2.4] that under condition (6.7) or (6.8), the family of parametric solutions  $\{q(\mathbf{y}) \in \mathcal{X} : \mathbf{y} \in U\}$  admits a holomorphic extension to the polydisc  $\mathcal{U}_\boldsymbol{\rho} := \prod_{j \geq 1} \{|z_j| < \rho_j\} \subset \mathbb{C}^\mathbb{N}$ . In particular, for every fixed  $\mathbf{y} \in U$ ,  $q(\mathbf{y})$  admits an extension with respect to  $\mathbf{y}_{\{1:s\}}$  to the finite dimensional cylinder  $\mathcal{U}_{\boldsymbol{\rho},s} := \prod_{1 \leq j \leq s} \{|z_j| < \rho_j\} \subset \mathbb{C}^s$  for any finite dimension  $s \in \mathbb{N}$  with modulus  $\|q(\mathbf{z})\|_{\mathcal{X}}$  uniformly bounded with respect to  $\mathbf{z} \in \mathcal{U}_{\boldsymbol{\rho},s}$ , and uniformly with respect to  $s \in \mathbb{N}$ .

In particular, for uncertainty parametrizations (6.2) with sequences  $(\psi_j)_{j \geq 1}$  satisfying (6.8) with  $\boldsymbol{\rho} \in \ell^p(\mathbb{N})$ , Theorem 5.1 then implies that the parametric solution family  $q(\mathbf{z})$  is  $(\mathbf{b}, p, \varepsilon)$ -holomorphic with  $\mathbf{b} = \left( \frac{1}{\langle u \rangle} \|\psi_j\|_{L^\infty(D)} \right)_{j \geq 1}$ . This in turn implies by Proposition 5.1 that the parametric Bayesian posterior densities  $\Theta(\mathbf{y})$  and  $\Psi(\mathbf{y})$  are  $(\mathbf{b}, p, \varepsilon)$  holomorphic, so that by Theorem 5.1 the parameter derivatives of  $q(\mathbf{y})$ ,  $\Theta(\mathbf{y})$  and  $\Psi(\mathbf{y})$  satisfy, for  $\mathbf{y} \in U_s$ , with any finite parameter dimension  $s \in \mathbb{N}$  the bounds (5.2) and (5.3).

### 6.3 Example 1: Karhunen-Loève series

Here, we choose  $d = 1$ ,  $D = (0, 1)$ ,  $\langle u \rangle = 1$  and, for some parameter  $\zeta > 1$  to be specified,

$$\psi_{2j}(x) = (2j)^{-\zeta} \sin(j\pi x), \quad \psi_{2j-1}(x) = (2j-1)^{-\zeta} \cos(j\pi x) \quad (6.9)$$

Then for every finite  $s$ , the  $s$ -term expansion (6.2) is smooth with respect to  $x$ ; this regularity is, however, not uniform with respect to the truncation dimension  $s$ : the choice of  $\zeta$  in (6.9) limits the spatial Sobolev regularity of  $u(\mathbf{y})$  in Sobolev and Hölder scales in the physical domain  $D$ . Here we have (6.6) and  $\mathbf{b} = (\|\psi_j\|_{L^\infty(D)})_{j \geq 1} \in \ell^p(\mathbb{N})$  holds for any  $1 \geq p > 1/\zeta$ .

### 6.4 Example 2: Indicator functions

When  $\psi_j$  are indicator functions of sets  $D_j \subset D$  which form a partition of  $D$  could be viewed as a model for material mixtures. We choose the basis in parametric space  $\{\psi_j(x)\}_{j=1}^s$  as “step functions” with supports  $D_j$ , i.e.

$$\psi_j(x) = b_j \chi_{D_j}(x), \quad b_j = \theta j^{-\zeta}, \quad 0 \leq \theta \leq u_-/4, \quad (6.10)$$

where  $\chi_I$  denotes the characteristic function of the interval  $I \subset D$  and where  $\{D_j\}_{j=1}^s$  denote a partition of the physical domain  $D$ . For the choice (6.10) of  $\psi_j$ , (6.8) holds with the particular sequence

$$\rho_j = 1 + \frac{u_-}{4\|\psi_j\|_{L^\infty(D)}} = 1 + \frac{u_-}{4b_j}. \quad (6.11)$$

This follows from

$$\left\| \sum_{j=1}^s \rho_j \psi_j \right\|_{L^\infty(D)} = \max_{1 \leq j \leq s} \left( 1 + \frac{u_-}{4b_j} \right) b_j = \frac{u_-}{4} + \|\mathbf{b}\|_{\ell^\infty} \leq \frac{u_-}{2}.$$

In this case  $E = \mathbb{N}$  and  $E^c = \emptyset$  in Proposition 5.3 so that HoQMC weights in (5.15) become product weights. This, in turn, implies in Proposition 5.3 that the complexity of the CBC construction for the HoQMC generating vectors scales linearly with the parameter dimension  $s$ .

## 7 Numerical Results

### 7.1 Construction of Interlaced Polynomial Lattice Rules

We consider the construction of interlaced polynomial lattice rules introduced in [13] by the component-by-component (CBC) algorithm. Proposition 5.3 shows that the CBC construction is feasible in  $\mathcal{O}(\alpha^2 s^2 N \log N)$  operations for SPOD weights and in  $\mathcal{O}(\alpha s N \log N)$  operations for product weights, which were both verified computationally in [20]. We consider the examples mentioned in Sections 6.3 and 6.4 above, which are of the SPOD and product type, respectively. The SPOD weights case corresponds to an empty set  $E$  in Proposition 5.3; for the product weights case we have  $J = \infty$ , i.e.  $E^c$  is empty.

In a computer implementation of the CBC algorithm for polynomial lattice rules, it is most natural to use the base  $b = 2$ , since then polynomials over  $\mathbb{Z}_2$  of degree less

than  $m$  can be represented as bit sequences of length  $m$ , allowing for very efficient bit-wise manipulations. In [45] it was shown that for  $b = 2$ , we have  $C_{\alpha,b} = 1$ , which yields significantly improved generating vectors (in terms of the observed convergence of the integration error for suitable test integrands) compared to the previous bounds with  $C_{\alpha,b} > 1$ . Further improvement is observed in practice by reducing the numerical value of  $C$  used in the CBC construction; in all experiments below, we use generating vectors based on the choice  $C = 0.1$ . A computational study of the impact on the choice of the Walsh constant was performed in [20].

## 7.2 Approximation of Prior Expectation

We consider here forward uncertainty quantification under the assumption of uniformly distributed parameters  $\mathbf{y} \in U_s$ , i.e. we choose the prior distribution  $\boldsymbol{\pi}_0 = \lambda^s$ , where  $\lambda^s$  denotes the  $s$ -dimensional Lebesgue measure. We denote in the following by  $q^h(x, \mathbf{y})$  the finite element approximation of the solution  $q$  of (6.1) with discretization parameter  $h$ . The goal of the computation is then the approximation of the expectation of a quantity of interest (QoI) function  $\phi(\mathbf{y})$ . We choose here and in the following as the quantity of interest the point evaluation of  $q^h$  at the point  $\bar{x} = 0.25$ ,  $\phi(\mathbf{y}) = q^h(\bar{x}, \mathbf{y})$ , and thus seek to compute the approximation

$$E^{\boldsymbol{\pi}_0}[\phi] := \frac{1}{N} \sum_{n=0}^{N-1} q^h(\bar{x}; \mathbf{y}^{(n)}) \approx \mathbb{E}^{\boldsymbol{\pi}_0}[\phi] = \int_{U_s} q^h(\bar{x}; \mathbf{y}) \, d\mathbf{y}, \quad (7.1)$$

where  $\{\mathbf{y}^{(n)}\}_{n=0}^{N-1}$  is the interlaced polynomial lattice point set. We are interested in the rate of convergence of the quadrature approximation  $E^{\boldsymbol{\pi}_0}$  to the true value  $\mathbb{E}^{\boldsymbol{\pi}_0}$ . Additionally, we consider the convergence of the dimension truncation and finite element errors, where applicable.

## 7.3 Approximation of Posterior Expectation

The approximation of the Bayesian inverse, as introduced in Section 3, is computationally very similar to the approximation of the prior expectation, since we must compute the ratio estimate  $\mathbb{E}^{\boldsymbol{\pi}^\delta}[\phi] = Z'/Z$ , where  $Z$  and  $Z'$  are high-dimensional integrals given by (3.8) and (3.10). We apply the same higher-order QMC rule mentioned above to both of these integrals.

Note that when using the same QMC rule for both integrals, the forward model is evaluated at exactly the same parameters  $\mathbf{y}^{(0)}, \dots, \mathbf{y}^{(N-1)}$  in both integrands; this allows a simple optimization leading to a reduction of the computational work by a factor two. Additionally storing the result of the prior expectation in the iteration over the quadrature points allows both forward and inverse UQ problems to be solved in one run, still requiring only  $N$  total evaluations of the forward model.

We briefly describe our choice of observation operator and specific noise model. As observation operator, we consider the evaluation of the solution  $q(\cdot, \mathbf{y})$  at the  $K = 3$  points  $\mathbf{x}_{obs} = (0.2, 0.5, 0.7)$ . As in the exposition in Section 3 above, we consider measurements perturbed by additive Gaussian noise  $\eta \sim \mathcal{N}(0, \Gamma)$  with known covariance matrix  $\Gamma \in \mathbb{R}_{sym}^{K \times K}$ . In the experiments, we consider one fixed instance of the measurement, i.e. we generate one realization of  $\eta \in \mathbb{R}^3$  which determines the measurement used throughout

the computation. The measurement  $\delta \in \mathbb{R}^3$  is then given for the unknown, exact value of the parameter  $\mathbf{y}^*$  as

$$\delta = \begin{pmatrix} q(0.2, \mathbf{y}^*) + \eta_1 \\ q(0.5, \mathbf{y}^*) + \eta_2 \\ q(0.7, \mathbf{y}^*) + \eta_3 \end{pmatrix}.$$

We assume independent noise for each component and unit variance, giving as covariance matrix simply the identity  $\Gamma = \mathbb{I}_{3 \times 3}$ . The parameters of the diffusion equation were chosen such that this makes sense, i.e. the size of the range of values of the observations are roughly  $\mathcal{O}(1)$ . As output quantity of interest, the solution evaluated at the point  $\bar{x} = 0.25$  is used, as in the prior expectation above.

## 7.4 Results for Example 1: Karhunen-Loève Series

We consider the trigonometric basis (6.9) where  $\bar{u}$  in (6.2) is chosen to be a constant of the size necessary to fulfill the uniform ellipticity conditions (6.5). For the Karhunen-Loève basis functions (6.9), we have  $\|\psi_j\|_{L^\infty(D)} = j^{-\zeta}$ ,  $\rho_j = 1$ , and thus  $(\|\psi_j\|_{L^\infty(D)})_j \in \ell^p(\mathbb{N})$  with  $1 \geq p > \frac{1}{\zeta}$ . The CBC construction of the generating vectors was based on SPOD weights (i.e., (5.15) with  $E = \emptyset$ ), given by

$$\gamma_{\mathbf{u}} := \sum_{\nu_{\mathbf{u}} \in \{1:\alpha\}^{|\mathbf{u}|}} |\nu_{\mathbf{u}}|! \prod_{j \in \mathbf{u}} (2^{\delta(\nu_j, \alpha)} \beta_j^{\nu_j}),$$

where  $\delta(\nu_j, \alpha) = 1$  if  $\nu_j = \alpha$  and  $\delta(\nu_j, \alpha) = 0$  otherwise and with the sequence  $\beta_j = \theta j^{-\zeta}$  with  $\theta = 0.2$  and  $\zeta = 2$ . We thus expect the QMC quadrature approximation to both prior and posterior expectation to converge with  $s$ -independent rate  $\mathcal{O}(N^{-\zeta})$  for digit interlacing parameter  $\alpha \geq 2$  in (5.12). To compare these QMC based results to the performance of existing methods, we also apply standard Monte Carlo sampling consisting of realizations of uniformly distributed points in  $U_s$ ; as a (rough) work measure we use in either case the number  $N$  of samples, which coincides with the number of forward PDE solves. As the algorithms considered in the present paper are single level algorithms, all PDE solves were performed with equal accuracy; multi-level extensions of either method are available and are known to deliver improved work versus accuracy [30, 14].

We briefly list the parameters used in the simulations. As right-hand side forcing term, we consider the function  $f(x) = 100x$ . Unless otherwise specified, we use a solution with  $N = 2^{20}$  quadrature points and meshwidth  $h = 2^{-20}$  as a reference. In the Monte Carlo results, we approximate the  $L^2$  error using 10 repetitions of the estimator. In the Bayesian inverse problem, we assume the measurement errors to result from a normal distribution with unit variance.

### 7.4.1 Finite Element Approximation

We solve the PDE (6.1) by the finite element method with piecewise linear basis functions and meshes obtained by regularly refining an initial mesh consisting of the points  $\{0, 1\}$ . The PG discretization error of the posterior approximation,  $|\mathbb{E}^{\mu^\delta}[q(\bar{x}; \mathbf{y})] - \mathbb{E}^{\mu^\delta}[q^h(\bar{x}; \mathbf{y})]|$ , is measured by replacing  $q(\bar{x}; \mathbf{y})$  with a reference solution obtained on a mesh with meshwidth  $h = 2^{-20}$ . Since as QoI we consider only the evaluation of the solution at the point  $\bar{x}$ , we only consider the absolute value of the (scalar) results. For QMC quadrature, we

use  $N = 2^{10}$  points. The convergence results of the finite element error of the posterior approximation, as well as of the individual integrands  $Z$  and  $Z'$  are shown in Figure 1a.

### 7.4.2 Dimension Truncation

To numerically verify the convergence rate of the error committed by dimension truncation to a finite dimension  $s < \infty$ , we consider the QMC-PG approximation of  $\mathbb{E}_N^{\pi^\delta}[\phi(q_s^h)]$  for varying  $s$ . In order to be able to neglect the other two error contributions, the finite element meshwidth is chosen as  $h = 2^{-20}$  and  $N = 2^{20}$  QMC points are used. By (2.19), we expect a convergence rate of  $s^{-(1/p-1)}$ ; for the case  $\zeta = 2$  in (6.10), we have  $p > 1/2$ , giving an expected rate of  $s^{-1+\varepsilon}$  for an  $\varepsilon > 0$ . In Figure 1b, this expected convergence rate can be clearly seen.

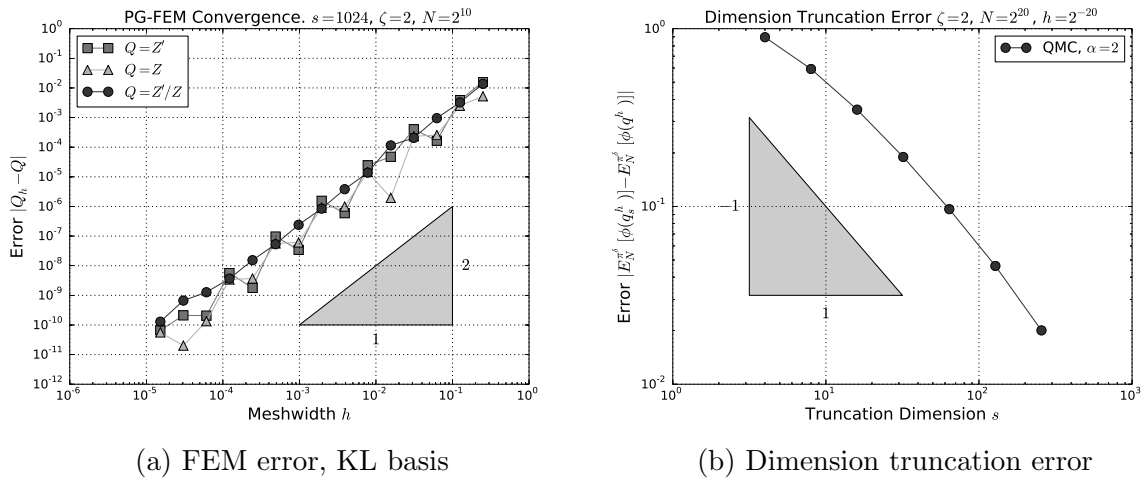
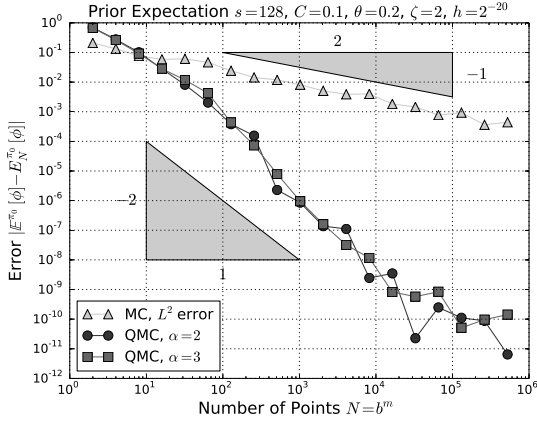


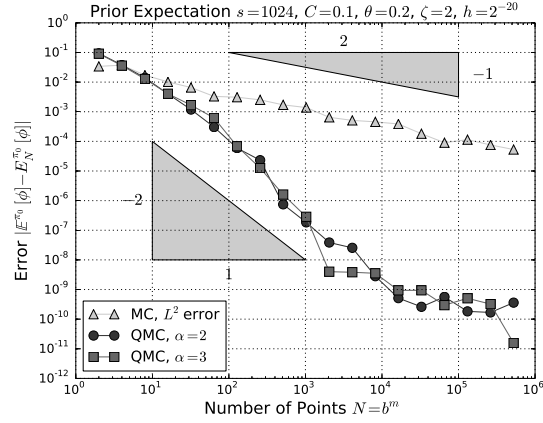
Figure 1: FEM and dimension truncation errors. Reference solutions were obtained using  $h = 2^{-20}$  for (a) and  $s = 1024$  dimensions for (b).

### 7.4.3 QMC Convergence

Figures 2 and 3 show the convergence of the QMC approximation to the prior and posterior expectations, respectively. In both cases, the convergence rate  $N^{-2}$  is clearly visible for the considered interlaced polynomial lattice rule with interlacing factors  $\alpha = 2, 3$ . This rate of convergence is in particular independent of the dimension  $s$  of the parametric space  $U_s$ , for the values considered here.

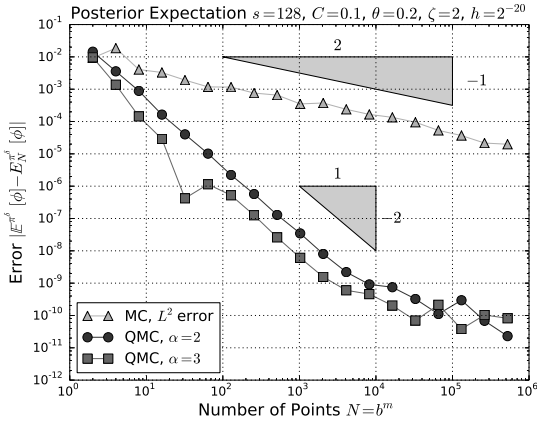


(a)  $s = 128$

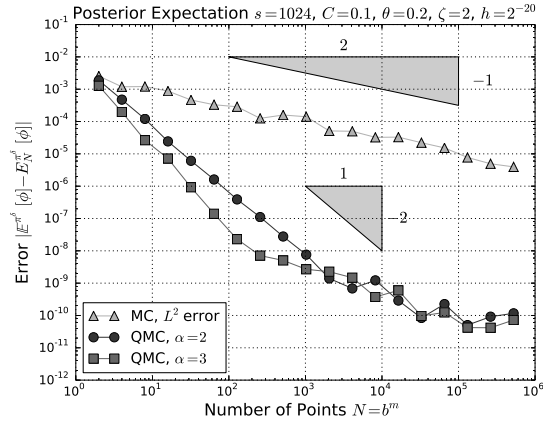


(b)  $s = 1024$

Figure 2: Convergence of the prior approximation vs. the number of samples  $N = b^m$  for  $C = 0.1$ ,  $\theta = 0.2$ ,  $s = 128, 1024$ ,  $\zeta = 2$ ,  $\alpha = 2, 3$ ,  $h = 2^{-20}$ . For MC, the  $L^2$  error was approximated using 10 repetitions.



(a)  $s = 128$



(b)  $s = 1024$

Figure 3: Convergence of the posterior approximation vs. the number of samples  $N = b^m$  for  $C = 0.1$ ,  $\theta = 0.2$ ,  $s = 128, 1024$ ,  $\zeta = 2$ ,  $\alpha = 2, 3$ ,  $h = 2^{-20}$ . For MC, the  $L^2$  error was approximated using 10 repetitions.

## 7.5 Results for Example 2: Indicator Functions

We consider the indicator function basis as in (6.10), with the intervals  $D_j = (x_{j-1}, x_j)$  chosen based on the points  $x_0, \dots, x_s$  of a graded mesh  $\mathcal{T} \subset (0, 1)$ . The points in  $\mathcal{T}$  are obtained by transforming an equidistant mesh with the function  $g(x) = x^a$  for an  $a \in \mathbb{R}$ . In the following, we use  $a = 0.2$ , yielding the points  $x_j = (j/s)^{1/5}$  for  $j = 0, \dots, s$ . This choice implies that the support of the first few parametric basis functions is relatively large, ensuring that the range of the observations of the solution is of the same order of magnitude for the values of  $s$  used in the experiments. This, together with the choice of  $f$  mentioned below, justifies the use of  $\Gamma = \mathbb{I}_{3 \times 3}$  in the measurement model of the Bayesian inverse problem. Note that for different “truncation” dimensions  $s$  the choice

(6.10) with intervals  $D_j = (x_{j-1}, x_j)$  and  $x_j = (j/s)^{1/5}$  implies that we solve different problems; convergence for  $s \rightarrow \infty$  is moot for this data.

In the following, we consider the right-hand side function  $f(x)$  to be constant. Together with the piecewise constant diffusion coefficient model, this implies that the solution is a piecewise quadratic function on the given mesh  $\mathcal{T}$ . Thus, if we use quadratic element basis functions in the finite element computations, we will obtain the exact solution, allowing us to ignore effects of the discretization error.

### 7.5.1 Choice of QMC weights

As mentioned in Section 6.4, this choice of diffusion coefficient model allows the use of product weights in the norm in Definition 5.1 of the form (cp. (5.15) with  $E^c = \emptyset$ )

$$\gamma_{\mathbf{u}} := \sum_{\nu_{\mathbf{u}} \in \{1:\alpha\}^{|\mathbf{u}|}} \nu_{\mathbf{u}}! \prod_{j \in \mathbf{u}} (2^{\delta(\nu_j, \alpha)} \beta_j^{\nu_j}),$$

which we construct with the sequence  $\beta_j = \theta j^{-\zeta}$  with  $\theta = 0.25$  and  $\zeta = 2$ . This is an advantage because the number of operations required for the construction of the generating vector is linear in the dimension  $s$ , whereas for the SPOD weights used in Section 7.4 it scaled quadratically with respect to  $s$  (see Proposition 5.3). This renders problems with large parametric dimension  $s$  computationally accessible.

### 7.5.2 QMC Convergence

Figures 4 and 5 show the convergence of the QMC approximation to the prior and posterior expectations, respectively. In both cases, we observe the expected rate  $N^{-2}$ , which seems to be independent of the parameter space dimension  $s$  for the values of  $s$  considered here.

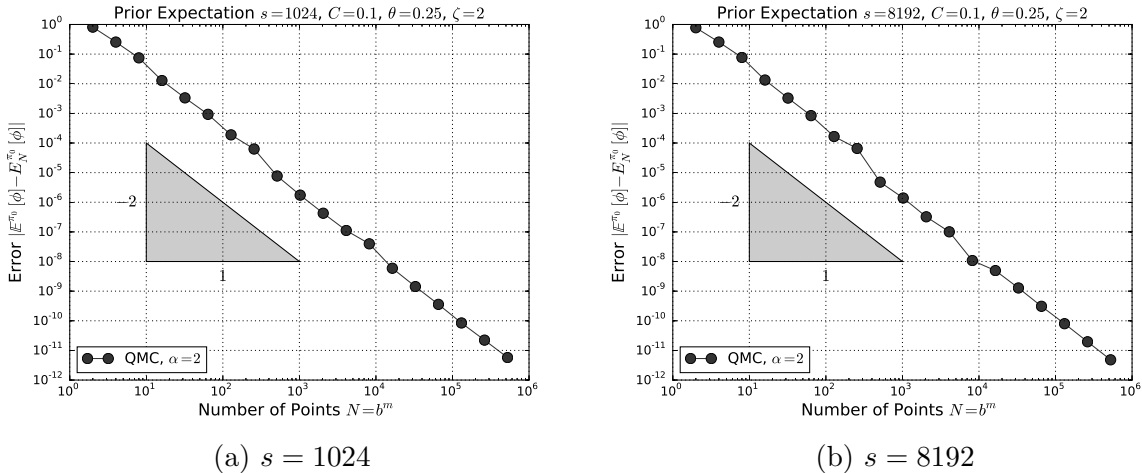


Figure 4: Convergence of the prior approximation vs. the number of samples  $N = b^m$  for  $C = 0.1$ ,  $\theta = 0.25$ ,  $s = 1024, 8192$ ,  $\zeta = 2$ ,  $\alpha = 2, 3$ .



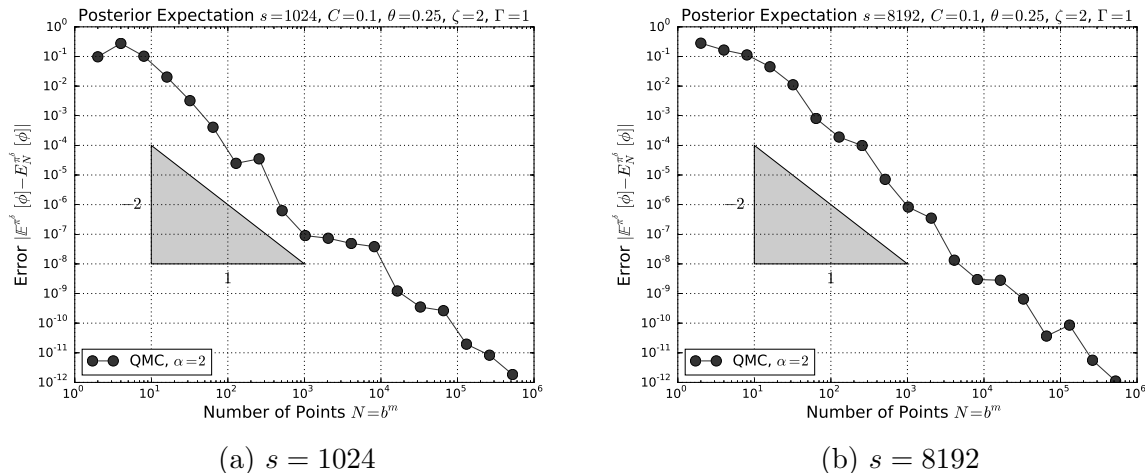


Figure 5: Convergence of the posterior approximation vs. the number of samples  $N = b^m$  for  $C = 0.1$ ,  $\theta = 0.25$ ,  $s = 1024, 8192$ ,  $\zeta = 2$ ,  $\alpha = 2, 3$ .

Comparing these convergence results with the corresponding results for the Karhunen-Loève basis, the ‘levelling’ of the total errors at around  $10^{-10}$  in Fig. 2 is due to the spatial discretization error, which is absent in the presently considered indicator function basis with  $P2$ -Finite Elements, suggesting that the additive structure of the combined error bound (5.19) (which resulted from the triangle inequality) is sharp in these cases.

**Acknowledgments** This work was supported by CPU time from the Swiss National Supercomputing Centre (CSCS) under project IDs s522 and d41, by the Swiss National Science Foundation (SNF) under Grant No. SNF149819, and by Australian Research Council’s Discovery Projects under project number DP150101770. This work has benefited from discussions between the authors at the conference “Approximation of High-Dimensional Numerical Problems: Algorithms, Analysis and Applications” at BIRS, Banff, Canada, September 27 - October 2, 2015.

## References

- [1] J. Baldeaux, J. Dick, J. Greslehner and F. Pillichshammer, Construction algorithms for higher order polynomial lattice rules. *J. Complexity*, 27, 281–299, 2011.
- [2] J. Baldeaux, J. Dick, G. Leobacher, D. Nuyens and F. Pillichshammer, Efficient calculation of the worst-case error and (fast) component-by-component construction of higher order polynomial lattice rules. *Numer. Algorithms*, 59, 403–431, 2012.
- [3] A. Chkifa, A. Cohen and Ch. Schwab, Breaking the curse of dimensionality in sparse polynomial approximation of parametric PDEs. *Journ. Math. Pures et Appliquees*, 103, 400–428, 2015.
- [4] Z. Ciesielski and J. Domsta, Construction of an orthonormal basis in  $C^m(I^d)$  and  $W_p^m(I^d)$ . *Studia Math.*, 41, 211–224, 1972.

- [5] A. Cohen, R. DeVore and Ch. Schwab, Convergence rates of best  $N$ -term Galerkin approximation for a class of elliptic sPDEs. *Found. Comput. Math.*, 10, 615–646, 2010.
- [6] A. Cohen, R. DeVore and Ch. Schwab, Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs. *Analysis and Applications*, 9, 1–37, 2011.
- [7] M. Dashti and A. M. Stuart, The Bayesian Approach to Inverse Problems (to appear in *Handbook of Uncertainty Quantification*, Springer Publ. 2016) Available at arXiv:1302.6989v2
- [8] P.J. Davis, *Interpolation and Approximation*. Dover Publications, Inc., New York, 1975.
- [9] J. Dick, Explicit constructions of Quasi-Monte Carlo rules for the numerical integration of high-dimensional periodic functions. *SIAM J. Numer. Anal.*, 45, 2141–2176, 2007.
- [10] J. Dick, Walsh spaces containing smooth functions and Quasi-Monte Carlo rules of arbitrary high order. *SIAM J. Numer. Anal.*, 46, 1519–1553, 2008.
- [11] J. Dick, The decay of the Walsh coefficients of smooth functions. *Bull. Aust. Math. Soc.*, 80, 430–453, 2009.
- [12] J. Dick and P. Kritzer, On a projection-corrected component-by-component construction. *J. Complexity*, 32, 74–80, 2016.
- [13] J. Dick, F.Y. Kuo, Q. T. Le Gia, D. Nuyens and Ch. Schwab, Higher order QMC Galerkin discretization for parametric operator equations. *SIAM J. Numer. Anal.* **52** 2676–2702, 2014.
- [14] J. Dick, F.Y. Kuo, Q. T. Le Gia and Ch. Schwab, Multi-level higher order QMC Galerkin discretization for affine parametric operator equations. Report 2014-14 Seminar for Applied Mathematics, ETH Zürich, Switzerland (in review). Available at arXiv:1406.4432 [math.NA]
- [15] J. Dick, Q. T. Le Gia and Ch. Schwab, Higher order Quasi Monte Carlo integration for holomorphic parametric operator equations, Report 2014-23, Seminar for Applied Mathematics, ETH Zürich, Switzerland (to appear in *SIAM Journ. Unc. Quantification*). Available at arXiv:1409.2180 [math.NA]
- [16] J. Dick, R. N. Gantner, Q. T. Le Gia and Ch. Schwab, Higher order Quasi-Monte Carlo integration for Bayesian Estimation. Report 2016 Seminar for Applied Mathematics, ETH Zürich, Switzerland (in preparation).
- [17] J. Dick and F. Pillichshammer, *Digital Nets and Sequences. Discrepancy Theory and Quasi-Monte Carlo Integration*, Cambridge University Press, 2010.

- [18] T.J. Dodwell, C. Ketelsen, R. Scheichl and A.L. Teckentrup, *A Hierarchical Multi-level Markov Chain Monte Carlo Algorithm with Applications to Uncertainty Quantification in Subsurface Flow*, SIAM/ASA Journal on Uncertainty Quantification, **3** (2015) 1075–1108.
- [19] V. Girault and P.A. Raviart, *Finite Element Methods for Navier-Stokes Equations*. Springer Verlag, Berlin, 1986.
- [20] R. N. Gantner and Ch. Schwab, Computational Higher Order Quasi-Monte Carlo Integration, Tech. Report 2014-25, Seminar for Applied Mathematics, ETH Zürich. To appear in Monte Carlo and Quasi-Monte Carlo Methods 2014, R. Cools and D. Nuyens (eds.), 2016.
- [21] T. Goda, Good interlaced polynomial lattice rules for numerical integration in weighted Walsh spaces. J. Comput. Appl. Math., 285, 279–294, 2015.
- [22] T. Goda and J. Dick, Construction of interlaced scrambled polynomial lattice rules of arbitrary high order. Found. Comput. Math., 15, 1245–1278, 2015.
- [23] M. Hansen and Ch. Schwab, Analytic regularity and best  $N$ -term approximation of high dimensional, parametric initial value problems. Vietnam Journal of Mathematics, 41, 181–215, 2013.
- [24] V.H. Hoang and Ch. Schwab, Analytic regularity and polynomial approximation of stochastic, parametric elliptic multiscale PDEs. Analysis and Applications (Singapore), 11, (01), 2011.
- [25] V. H. Hoang and Ch. Schwab, Regularity and Generalized Polynomial Chaos Approximation of Parametric and Random Second-Order Hyperbolic Partial Differential Equations. Analysis and Applications (Singapore), 10, (3), 2012.
- [26] V. H. Hoang and Ch. Schwab and A.M. Stuart, *Complexity analysis of accelerated MCMC methods for Bayesian inversion*, Inverse Problems **29**(8) 2013.
- [27] A. Kunoth and Ch. Schwab, Analytic Regularity and GPC Approximation for Stochastic Control Problems Constrained by Linear Parametric Elliptic and Parabolic PDEs. SIAM J. Control Optim., 51, 2442 – 2471, 2013.
- [28] F. Y. Kuo, Ch. Schwab and I. H. Sloan, Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficient. SIAM J. Numerical Analysis, 50, 3351–3374, 2012.
- [29] F. Y. Kuo, Ch. Schwab and I. H. Sloan, Quasi-Monte Carlo methods for very high dimensional integration: the standard weighted-space setting and beyond. ANZIAM Journal, 53, 1–37, 2011.
- [30] F. Y. Kuo, Ch. Schwab and I. H. Sloan, Multi-Level Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficient, Found. Comp. Math., 15, 411–449, 2015.
- [31] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*. SIAM, Philadelphia, 1992.

- [32] V. Nistor and Ch. Schwab, High order Galerkin approximations for parametric second order elliptic partial differential equations. *Math. Models Methods Appl. Sci.*, 23, 1729 – 1760, 2013.
- [33] D. Nuyens and R. Cools, Fast algorithms for component-by-component construction of rank-1 lattice rules in shift-invariant reproducing kernel Hilbert spaces. *Math. Comp.*, 75, 903 – 920, 2006.
- [34] J. Pousin and J. Rappaz, Consistency, Stability, apriori and aposteriori errors for Petrov-Galerkin methods applied to nonlinear problems. *Numer. Math.*, 69, 213–231, 1994.
- [35] Cl. Schillings and Ch. Schwab, Sparse, adaptive Smolyak quadratures for Bayesian inverse problems. *Inverse Problems*, 29, 065011, 28 pp, 2013.
- [36] Cl. Schillings and Ch. Schwab, Sparsity in Bayesian Inversion of Parametric Operator Equations, *Inverse Problems*, 30, 065007, 30 pp., 2014.
- [37] Cl. Schillings and Ch. Schwab, Scaling Limits in Computational Bayesian Inversion. Report 2014-26, Seminar for Applied Mathematics, ETH Zürich, (to appear in *M2AN*, 2016).
- [38] Ch. Schwab, QMC Galerkin discretizations of parametric operator equations. In J. Dick, F. Y. Kuo, G. W. Peters and I. H. Sloan (eds.), *Monte Carlo and Quasi-Monte Carlo methods 2012*, Springer Verlag, Berlin, 2013, pp. 613–630.
- [39] Ch. Schwab and C.J. Gittelsohn, Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs. *Acta Numerica*, 20, 291–467, 2011.
- [40] Ch. Schwab and R.S. Stevenson, Space-Time adaptive wavelet methods for parabolic evolution equations. *Math. Comp.*, 78, 1293–1318, 2009.
- [41] Ch. Schwab and A.M. Stuart, Sparse deterministic approximation of Bayesian inverse problems. *Inverse Problems*, 28, 045003, 2012.
- [42] Ch. Schwab and R. A. Todor, Karhunen-Loève approximation of random fields by generalized fast multipole methods. *J. Comput. Phy.*, 217, 100–122, 2006.
- [43] I. H. Sloan and H. Woźniakowski, When are Quasi-Monte Carlo algorithms efficient for high-dimensional integrals? *J. Complexity*, 14, 1–33.
- [44] A. M. Stuart, Inverse problems: a Bayesian perspective. *Acta Numerica*, 19, 451–559, 2010.
- [45] T. Yoshiki, Bounds on the Walsh coefficients by dyadic difference and a new Koksma-Hlawka type inequality for Quasi-Monte Carlo integration. Preprint arXiv:1504.03175.