Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Seminar for Applied Mathematics SAM

Extrapolated Lattice Rule Integration in Computational Uncertainty Quantification

J. Dick and M. Longo and Ch. Schwab

Research Report No. 2020-29

May 2020
Latest revision: June 2021

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

Funding: JD is partly supported by the Australian Research Council Discovery Project DP190101197

Extrapolated Lattice Rule Integration in Computational Uncertainty Quantification

Josef Dick, Marcello Longo, Christoph Schwab

June 22, 2021

Abstract

We present an extension of the convergence analysis for Richardson-extrapolated polynomial lattice rules from [Josef Dick, Takashi Goda and Takehito Yoshiki: Richardson extrapolation of polynomial lattice rules, SIAM J. Numer. Anal. **57**(2019) 44-69] for high-dimensional, numerical integration to classes of integrand functions with quantified smoothness and Quasi-Monte Carlo ("QMC" for short) integration rules with so-called smoothness-driven, product and order dependent (SPOD for short) weights. We establish in particular sufficient conditions for the existence of an asymptotic expansion of the QMC integration error with respect to suitable powers of N, the number of QMC integration nodes. We derive a dimension-separated criterion for a fast component-by-component ("CBC" for short) construction algorithm for the computation of the QMC generating vector with quadratic scaling with respect to the integration dimension s.

We prove that the proposed QMC integration strategies a) are free from the curse of dimensionality, b) afford higher-order convergence rates subject to suitable summability conditions on the QMC weights, c) allow for certain classes of high-dimensional integrands functions a computable, asymptotically exact numerical estimate of the QMC quadrature error, with reliability and efficiency independent of the dimension of the integration domain and d) accommodate fast, FFT-based matrix-vector multiplication from [Dick, Josef; Kuo, Frances Y.; Le Gia, Quoc T.; Schwab, Christoph: Fast QMC matrix-vector multiplication. SIAM J. Sci. Comput. 37 (2015), no. 3, A1436-A1450] when applied to parametric operator equations.

The integration methods are applicable for large classes of many-parametric integrand functions with quantified parametric smoothness. We verify all hypotheses and present numerical examples arising from the Galerkin Finite-Element discretization of a model, linear parametric elliptic PDE illustrating a) - d). We verify computationally the scaling of the fast CBC construction algorithm with SPOD QMC weights, and examine the extrapolation-based a-posteriori numerical estimation of the QMC quadrature error. We find in examples with parameter spaces of dimension $s = 10, ..., 128$ that the extrapolation-based, computable QMC integration error indicator has an efficiency index between 0.9 and 1.1, for a moderate number N of QMC points.

Key Words: High-dimensional Quadrature, Quasi-Monte Carlo, Richardson Extrapolation, A-posterior Error Estimation

AMS Subject Classification: 65C05, 65N30, 35J25

1 Introduction

The efficient numerical analysis of partial differential equations (PDEs for short) with *distributed uncertain inputs*, i.e. uncertain input data from function spaces, has emerged as one key element in the field of computational uncertainty quantification.

We consider a physical process described by a governing equation (assumed to be known), the *forward model* $\mathcal{P}$. We assume that $\mathcal{P}$ depends on *empirical input data* to be determined by observations or experiments, and therefore prone to (observational) uncertainty. For a given instance of such uncertain input ψ into $\mathcal{P}$, we consider a parametric operator equation of generic form: given $\psi \in L$, find $u \in X$ such that

$$\mathcal{P}\left(u, \psi\right) = 0 \quad \text{in } Y'. \tag{1}$$

Here, L, X, Y are suitable Banach spaces. We assume that the forward model is *locally well-posed*, i.e. it is well-posed for a (assumed known) *nominal input* $\psi_0 \in L$ and the unique solution $u \in X$ is assumed to depend continuously on the data $\psi \in L$, i.e. the *data-to-solution map* $S : L \to X$, where $S : \psi \mapsto u$, is assumed to be locally Lipschitz continuous as a map from L to X, on a (sufficiently small) neighborhood of $\psi_0 \in L$. More precisely, we assume that (1) is well-posed for all $\psi \in B_R(\psi_0) \subset L$, with the usual notation of $B_R(\psi)$ denoting an open ball of radius $R > 0$ about ψ in a Banach space (here: L).

The numerical analysis of (1) will require further hypotheses. We assume that all admissible inputs $\psi \in B_R(\psi_0) \subset L$ for (1) are parametrized in terms of an *affine representation system* $\boldsymbol{\Psi} = \{\psi_j\}_{j\geq 1}$, where the index j ranges over a set $\{1 : s\} := \{1, 2, ..., s\} \subseteq \mathbb{N}$ (understood as all of $\mathbb{N}$ in the case that $s = \infty$). Then, we consider (1) for input data $\psi \in B_R(\psi_0) \subset L$ of affine-parametric form

$$\psi(\boldsymbol{y}) := \psi_0 + \sum_{j\geq 1} y_j \psi_j, \tag{2}$$

where the parameter sequence $\boldsymbol{y} := (y_j)_{j\geq 1} \subset U$ lies in the parameter domain $U = [-1/2, 1/2]^s$, and where the parameter dimension $s \in \mathbb{N}$ is either finite or, in case that sequences of parameters are considered, infinite, in which case $s = \infty$. Inserting the affine-parametric representation (2) into the forward operator equation (1), we obtain the *parametric forward operator equation*: given $\boldsymbol{y} \in U$, find $u(\boldsymbol{y}) \in X$ such that

$$\mathcal{P}\left(u(\boldsymbol{y}), \psi(\boldsymbol{y})\right) = 0 \quad \text{in } Y'. \tag{3}$$

Examples of affine-parametric representations (2) comprise in particular so-called *Karhunen-Loeve* (KL for short) expansions of random fields ψ, but also multiresolution representations of ψ.

The purpose of the present paper is to study the numerical approximation of integrals over (functionals of) parametric solution families of the parametric operator equations (3) on possibly high-dimensional parameter spaces U. Our goal is an accurate numerical approximation, with low computational cost, of the quantity

$$I_s(G(u)) = \int_U G(u(\cdot, \boldsymbol{y}))\mathrm{d}\boldsymbol{y} \approx \frac{1}{N} \sum_{\boldsymbol{y}_n \in P} G(u_h(\cdot, \boldsymbol{y}_n)) =: Q_{N,\bar{s}}(G(u_h)), \tag{4}$$

for some $\bar{s} \leq s$, $\bar{s} < \infty$, where u_h is the finite element solution of (3) for the case where ψ depends only on the first $\bar{s}$ elements of $\boldsymbol{y}$. Here, the linear functional $G \in X'$ shall be referred to as *Quantity of Interest* ("QoI" for short).

The sampling set $P \subset [-1/2, 1/2]^{\bar{s}}$ in (4) in the present shall be a *deterministic QMC point set* of cardinality N. Specifically, we choose P to be the *extrapolated polynomial lattice* as proposed recently in [9]. For the numerical approximation of (4), the parametric solution u of (3) must be approximated numerically by discretizing the operator equation (3) for each instance of the parameter sequence $\boldsymbol{y}$. We denote by h a generic discretization parameter that describes, for example, the meshwidth of a Galerkin discretization of the parametric problem.

In recent years, the mathematical analysis of QMC integration methods as applied to PDEs with distributed uncertain inputs (such as diffusion coefficient fields in heterogeneous media, spatiotemporally varying source term and boundary data, etc.) has seen significant development, starting with [30, 21]. However, the Richardson extrapolation method based on an asymptotic expansion of the QMC integration error, which was first proposed in [9], has not been studied so far in the context of PDEs with random coefficients. It allows one to obtain QMC integration rules which achieve convergence rates greater than 1 independently of the dimension s of the integration domain thereby overcoming the curse of dimensionality, for certain classes of smooth integrands. In the present paper we develop the Richardson extrapolation for QMC from [9] further and apply it to PDEs with random coefficients.

1.1 Contributions

The contributions of the present paper are as follows.

Firstly, we extend the QMC error analysis for extrapolated polynomial lattice rules given in [9] for the function space setting with so-called "product weights " to the more general, so-called "smoothness-driven, product and order dependent weights" (SPOD weights, for short). The main result, Theorem 2.4, constitutes an extension of [9] to the case of SPOD weights. We remark that both, product and SPOD weights do appear in partial differential equations with parametric random field input data. We refer to the discussion in [17, 16], depending on the support properties of the representation system for the parametric input data: localized supports allow for the use of product weights whereas globally supported representation systems (such as Karhunen-Loeve eigensystems [36], or reduced basis representations computed by greedy searches [35]) entail SPOD type QMC weights in order to ensure the maximal (dimension-independent) convergence rates for given sparsity of the coefficient representation.

Secondly, we show that the asymptotic expansion of the QMC quadrature error which is furnished by the QMC theory allows for computable, a posteriori estimation of the QMC integration error. Under suitable hypotheses on the parametric integrand functions that we verify for a model, parametric elliptic PDE with uncertain coefficient, we prove that the computable QMC quadrature error estimate is asymptotically exact. Furthermore, the efficiency of the extrapolation-based, computable QMC integration error estimator is independent of the dimension s of the QMC integration domain. In numerical experiments we show that very good efficiencies, i.e., ratios between numerically estimated QMC integration error and the exact value of the integral, between 0.9 and 1.1, are achieved already with a moderate number of QMC lattice points, with performance which is uniform in the quadrature dimension $s = 16, ..., 128$.

Thirdly, we argue that the particular structure of the lattice points employed in the base QMC integration rule, on which the extrapolation process is based, facilitates higher order QMC quadrature and a so-called *fast matrix-vector multiplication*, which is accelerated by FFT algorithms, as proposed in [11] for first order lattice QMC integration rules as used here. Extension of [11] to higher order QMC quadrature rules, such as the interlaced polynomial lattice rules (IPLs) in [10], was not feasible due to digit interlacing used in the construction of the generating vectors for these higher order QMC integration techniques. The extrapolation-based QMC algorithms proposed here and in [9] do allow us to achieve higher order, dimension-independent QMC convergence rates while at the same time facilitating use of FFT accelerated matrix-vector multiplication. In numerical experiments for a linear, affine-parametric elliptic model PDE, we find significant quantitative advantages of the FFT accelerated algorithms.

1.2 Outline

The outline of this paper is as follows. In Section 2, we recapitulate the function space setting and the basic results from [9] on extrapolated polynomial lattice rules. The main result is contained in Theorem 2.4 in Section 2.

In Section 3, we verify the assumptions in Theorem 2.4 for a particular, model class of operator equations (1), namely a linear, elliptic diffusion problem in a bounded, physical domain D. Section 4 will present a novel, computable a posteriori QMC integration error estimator and establishes its asymptotic exactness. Section 5 is devoted to several sets of numerical experiments, indicating the sharpness of the summability conditions of the extrapolated lattice rules based on SPOD QMC weights, establishing the viability and the asymptotic exactness of the computable QMC a-posterior error estimators and demonstrating an application to a model, linear elliptic parametric PDE problem in two space dimensions. Section 6 will present several conclusions and perspectives for further work.

2 Richardson extrapolation of polynomial lattice rules for SPOD weights

In this section, to prepare the subsequent developments of the present paper, we recall the setting of [9] and the references there in Sections 2.1–2.3. Then in Sections 2.4 and 2.5, we develop the extension of the Richardson extrapolation of the QMC error, that was developed in [9] for product weights, to SPOD weights.

Throughout this section, we assume that the parametric dimension $s \in \mathbb{N}$ is arbitrary.

2.1 Polynomial lattice rules

Polynomial lattice rules provide a special construction of QMC quadrature rules introduced by Niederreiter [33] and employed in many instances, for a comprehensive overview we refer to [13]. In the following let $b \geq 2$ be a prime number, $\mathbb{F}_b$ be the finite field with b elements, $\mathbb{F}_b[x]$ be the set of all polynomials with coefficients in $\mathbb{F}_b$ and $\mathbb{F}_b((x^{-1}))$ be the set of all formal Laurent series $\sum_{i=w}^{\infty} a_i x^{-i}$, $w \in \mathbb{Z}$, and with coefficients a_i in $\mathbb{F}_b$. We identify the integers $0, 1, \ldots, b-1$ with the elements in the finite field $0, 1, \ldots, b-1 \pmod b$. For an integer $0 \leq n < b^m$ given by the base b expansion $n = n_0 + n_1 b + \cdots + n_{m-1} b^{m-1}$, with $n_0, \ldots, n_{m-1} \in \{0, 1, \ldots, b-1\}$, we define $n(x) \in \mathbb{F}_b[x]$ given by $n(x) = n_0 + n_1 x + \cdots + n_{m-1} x^{m-1}$, where we now consider $n_0, \ldots, n_{m-1} \in \mathbb{F}_b$.

Definition 2.1. *Let $m \geq 2$ be an integer and $\mathfrak{p} \in \mathbb{F}_b[x]$ be a polynomial with $\deg(\mathfrak{p}) = m$. Let $\boldsymbol{q} = (q_1, \ldots, q_s)$ be a vector of polynomials over $\mathbb{F}_b$ with degree $\deg q_j < m$. We define the map $v_m : \mathbb{F}_b((x^{-1})) \to [0, 1)$ by*

$$v_m \left(\sum_{i=w}^{\infty} a_i x^{-i} \right) = \sum_{i=\max\{1,w\}}^{m} a_i b^{-i}.$$

For $0 \leq n < b^m$, we put

$$\boldsymbol{x}_n = \left(v_m \left(\frac{n(x) q_1(x)}{\mathfrak{p}(x)} \right), \ldots, v_m \left(\frac{n(x) q_s(x)}{\mathfrak{p}(x)} \right) \right) \in [0, 1)^s.$$

Then the point set $P(\mathfrak{p}, \boldsymbol{q}) = \{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{b^m-1}\}$ is called a polynomial lattice point set and a QMC rule using this point set is called a polynomial lattice rule.

Since our integrands are defined on $[-\frac{1}{2},\frac{1}{2}]^s$ rather than $[0,1]^s$, we sometimes use the point sets

$$\boldsymbol{x}_n = \left(v_m\left(\frac{n(x)q_1(x)}{\mathfrak{p}(x)}\right) - \frac{1}{2}, \dots, v_m\left(\frac{n(x)q_s(x)}{\mathfrak{p}(x)}\right) - \frac{1}{2}\right), \quad n = 0, 1, \dots, b^m - 1.$$

In the analysis of the following sections, we will denote by $P^{\perp}(\mathfrak{p}, \boldsymbol{q}) \subset \mathbb{N}_0^s$ the dual lattice of $P(\mathfrak{p}, \boldsymbol{q})$, defined as in [9, Definition 2.6]. Moreover, we denote $P_{\mathfrak{u}}^{\perp}(\mathfrak{p}, \boldsymbol{q}) := \{\boldsymbol{k}_{\mathfrak{u}} \in \mathbb{N}^{|\mathfrak{u}|} : (\boldsymbol{k}_{\mathfrak{u}}, \boldsymbol{0}) \in P^{\perp}(\mathfrak{p}, \boldsymbol{q})\}$.

2.2 Extrapolated polynomial lattice rules and CBC construction

Given $1 \le r, q \le \infty$, a set of positive weights $\boldsymbol{\gamma} = (\gamma_{\mathfrak{u}})_{\mathfrak{u} \subset \mathbb{N}, |\mathfrak{u}| < \infty}$ and $\alpha \in \mathbb{N}, \alpha \ge 2$, the QMC error analysis is based on the *weighted unanchored Sobolev space* $\mathcal{W}_{s,\alpha,\boldsymbol{\gamma},q,r}$ which is equipped with the norm

$$\|F\|_{s,\alpha,\boldsymbol{\gamma},q,r} := \left(\sum_{\mathfrak{u} \subseteq \{1:s\}} \left(\gamma_{\mathfrak{u}}^{-q} \sum_{\mathfrak{v} \subseteq \mathfrak{u}} \sum_{\boldsymbol{\nu}_{\mathfrak{u}\setminus\mathfrak{v}} \in \{1:\alpha\}^{|\mathfrak{u}\setminus\mathfrak{v}|}} \int_{[-\frac{1}{2},\frac{1}{2}]^{|\mathfrak{v}|}} \left| \int_{[-\frac{1}{2},\frac{1}{2}]^{s-|\mathfrak{v}|}} \partial_{\boldsymbol{y}}^{(\boldsymbol{\nu}_{\mathfrak{u}\setminus\mathfrak{v}}, \boldsymbol{\alpha}_{\mathfrak{v}})} F(\boldsymbol{y}) \right|^q \right)^{r/q} \right)^{1/r}. \tag{5}$$

These function spaces were also found to be crucial in the mathematical convergence rate analysis for so-called interlaced polynomial lattice rules (IPLs for short) in [10, 8, 12] and the references there. Assume that the integrand F has finite norm $\|F\|_{s,\alpha,\boldsymbol{\gamma},q,r} < \infty$. Then in [9, Equation (3.1)] it was shown that the following equality holds

$$Q_{b^m,s}(F) = \frac{1}{b^m} \sum_{n=0}^{b^m-1} F(\boldsymbol{x}_n) = I_s(F) + \sum_{\tau=1}^{\alpha-1} \frac{\sigma_\tau(F)}{b^{\tau m}} + S_{\mathfrak{p}_m}(\boldsymbol{q}_m)(F) + R_{s,\alpha,b^m}, \tag{6}$$

where $P(\mathfrak{p}_m, \boldsymbol{q}_m) = \{\boldsymbol{x}_0, \boldsymbol{x}_1, \dots, \boldsymbol{x}_{b^m-1}\}$ is the polynomial lattice with generating vector $\boldsymbol{q}_m$ and modulus $\mathfrak{p}_m$, and $Q_{b^m,s}$ is the corresponding rule on the nodes $P(\mathfrak{p}_m, \boldsymbol{q}_m)$. We have that $\sigma_\tau(F)$ depends on the function F and τ but not on the polynomial lattice point set, R_{s,α,b^m} decays with order $b^{-\alpha m}$, and $S_{\mathfrak{p}_m}(\boldsymbol{q}_m)(F)$ depends on the polynomial lattice rule and the integrand F (see (49) below for a precise definition). The work [9] uses a so-called component-by-component (CBC) algorithm to find a sequence of polynomial lattices $P(\mathfrak{p}_m, \boldsymbol{q}_m)$, $m \in \mathbb{N}$, such that $S_{\mathfrak{p}_m}(\boldsymbol{q}_m)(F)$ is of order $C_\delta b^{-\alpha m + \delta}$, for some constant $C_\delta > 0$ and any $\delta > 0$, where the constant C_δ goes to ∞ as $\delta \to 0^+$. $S_{\mathfrak{p}_m}(\boldsymbol{q}_m)(F)$ is also related to the search criterion $B_{\boldsymbol{\gamma}}(\mathfrak{p}_m, (q_{1,m}, \dots, q_{d-1,m}, q_{d,m}))$ defined in (12) below.

Given $m \in \mathbb{N}$, the CBC algorithm consists of the steps:

1. Initialize $\mathfrak{p} \in \mathbb{F}_b[x]$ irreducible of degree m and $q_1^* = 1$,
2. For $d = 2, \dots, s$, define $q_d^* = \operatorname{argmin}_{q_d \in \mathbb{F}_b[x]\setminus\{0\}, \deg(q_d) < m} B_{\boldsymbol{\gamma}}(\mathfrak{p}, (q_1^*, \dots, q_{d-1}^*, q_d))$.

Moreover, a suitable reformulation of the criterion $B_{\boldsymbol{\gamma}}$, allows to accelerate the argmin search using FFT, in the case of product or SPOD weights [34, 29, 10]. Its analysis in the present setting is subject of Section 2.5.

The only terms in (6), which are not of order $b^{-\alpha m + \delta}$, $\delta > 0$, are $\sum_{\tau=1}^{\alpha-1} \frac{\sigma_\tau(F)}{b^{\tau m}}$. The basic idea of the Richardson extrapolation rests on the following formula

$$\begin{aligned} Q_{b^m,s}^{(2)}(F) =& \frac{b\, Q_{b^m,s}(F) - Q_{b^{m-1},s}(F)}{b-1} \\ =& I_s(F) + \sum_{\tau=1}^{\alpha-1} \frac{\sigma_\tau(F)}{b^{\tau m}} \frac{b - b^\tau}{b-1} + \frac{b S_{p_m}(\boldsymbol{q}_m)(F) - S_{p_{m-1}}(\boldsymbol{q}_{m-1})(F)}{b-1} + \frac{b R_{s,\alpha,m} - R_{s,\alpha,m-1}}{b-1}. \end{aligned} \tag{7}$$

Since the term in the sum for $\tau = 1$ now cancels out, we get that $Q^{(2)}_{b^m,s}(F) - I_s(F)$ converges with order $b^{-2m+\delta}$ for any $\delta > 0$. Hence we have improved the convergence rate of our approximation algorithm. Repeated application of this idea, namely,

$$Q^{(\tau+1)}_{b^n,s}(F) = \frac{b^\tau Q^{(\tau)}_{b^n,s}(F) - Q^{(\tau)}_{b^{n-1},s}(F)}{b^\tau - 1}, \quad m - \alpha + \tau < n \le m,$$

then yields an integration rule $Q^{(\alpha)}_{b^m,s}$ which achieves a convergence rate of the integration error of order $C_\delta b^{-\alpha m+\delta}$ for any $\delta > 0$. Here, we set $Q^{(1)}_{b^m,s} = Q_{b^m,s}$. Therefore, we can rewrite the extrapolated sequence as linear combinations of the original sequence

$$Q^{(\alpha)}_{b^m,s}(F) = \sum_{\tau=1}^{\alpha} a^{(\alpha)}_\tau Q_{b^{m-\tau+1},s}(F), \tag{8}$$

for some constants $a^{(\alpha)}_\tau$ which are independent of b, m, s (these constants arise from the Richardson extrapolation, see [9, Section 2.4]). In Section 4 we show that this method also yields a computable a-posteriori estimation of the integration error.

2.3 Previous results

In [9, Section 3.4], it is shown that for every $\alpha \in \mathbb{N}$, $\alpha \ge 2$, and for every prime basis $b \in \mathbb{N}$, there exists an extrapolated polynomial lattice rule $Q^{(\alpha)}_{b^m,s}$ such that, for all $1/\alpha < \lambda \le 1$ and for every integrand function $F \in \mathcal{W}_{s,\alpha,\boldsymbol{\gamma},q,\infty}$, there exists a constant $C > 0$ independent of m, F and of the integration dimension s such that

$$|I_s(F) - Q^{(\alpha)}_{b^m,s}(F)| \le C \frac{\|F\|_{s,\alpha,\boldsymbol{\gamma},q,\infty}}{(b^m - 1)^{1/\lambda}} \left(J_{s,\lambda,\boldsymbol{\gamma}} + H_{s,\boldsymbol{\gamma},q,\infty}\right), \tag{9}$$

where the constant C depends only on b and α and

$$J_{s,\lambda,\boldsymbol{\gamma}} := \left[\sum_{u \subseteq \{1:s\}} \gamma_u^\lambda C_\alpha^{\lambda|u|} E^{|u|}_{\alpha,\lambda}\right]^{1/\lambda}, \quad H_{s,\boldsymbol{\gamma},q,\infty} := \sum_{\mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}} (\alpha+1)^{|\mathfrak{u}|/q'} D_\alpha^{|\mathfrak{u}|}. \tag{10}$$

In [9, Theorem 4.1], it was shown for *product weights* $\gamma_{\mathfrak{u}} = \prod_{j\in\mathfrak{u}} \gamma_j$ that it is possible to construct a generating vector with a so-called fast CBC algorithm satisfying (9). Moreover, it is sufficient to have $(\gamma_j)_j \in \ell^\lambda(\mathbb{N})$ for some $\lambda > 1/\alpha$ to obtain the convergence rate $\mathcal{O}(b^{-m/\lambda})$, which is free from the curse of dimensionality, i.e. it holds with rate and constant independent of the parametric dimension s.

Here, we extend this result to SPOD weights. We recall that the error bound in [9] was restricted to product weights due to a technical obstruction (see [9, Remark 4.2]).

2.4 Extrapolated polynomial lattice rule error analysis with SPOD weights

We use Richardson extrapolation in the context of PDEs with random coefficients, which are represented by *dictionaries with globally supported elements*. Such representations arise, for example, in parametric input functions which are obtained from reduced basis (RB) or from model order reduction (MOR) approaches which typically result in parsimonius representation of input manifolds in terms of globally supported basis functions. We refer to [26, 35] and the references there for such representations of distributed, parametric inputs.

We need a corresponding error bound also for SPOD weights $\gamma_{\mathfrak{u}}$ in the weighted norm (5), where

$$\gamma_{\mathfrak{u}} = \sum_{\boldsymbol{\nu} \in \{1:\alpha\}^{|\mathfrak{u}|}} ((|\boldsymbol{\nu}| + c_1)!)^{c_2} \prod_{j \in \mathfrak{u}} c_3 \beta_j^{\nu_j}, \tag{11}$$

where $\boldsymbol{\beta} = (\beta_j)_{j \in \mathbb{N}}$ is a sequence of non-increasing, non-negative real numbers, c_1 is a non-negative integer, and $c_2, c_3 > 0$ are real numbers.

In [9, Theorem 3.1] it was shown that the following quantity, for $d \in \mathbb{N}, d \leq s$

$$B_{\boldsymbol{\gamma}}(\mathfrak{p}_m, (q_{1,m}, \ldots, q_{d-1,m}, q_{d,m})) = \sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:d\}} \gamma_{\mathfrak{u}} C_\alpha^{|\mathfrak{u}|} \sum_{\substack{\boldsymbol{k}_{\mathfrak{u}} \in P_{\mathfrak{u}}^{\perp}(\mathfrak{p}_m, (q_{1,m}, \ldots, q_{d-1,m}, q_{d,m})) \\ \exists j \in \mathfrak{u}: b^m \nmid k_j}} b^{-\mu_\alpha(\boldsymbol{k}_{\mathfrak{u}})}, \tag{12}$$

where $\mathfrak{p}_m \in \mathbb{F}_b[x]$ is the modulus of degree m and the generating vector $(q_{1,m}, \ldots, q_{s-1,m}, q_{s,m}) \in (\mathbb{F}_b[x])^s$ depends on m, is the main term in the bound on the QMC integration error for Richardson-extrapolated lattice QMC integration rules, i.e.

$$|I_s(F) - Q_{b^m,s}^{(\alpha)}(F)| \leq \sum_{\tau=1}^{\alpha} |a_\tau^{(\alpha)}| \left(B_{\boldsymbol{\gamma}}(\mathfrak{p}_{m-\tau+1}, (q_{1,m-\tau+1}, \ldots, q_{s-1,m-\tau+1}, q_{s,m-\tau+1})) + R_{s,\alpha,b^{m-\tau+1}} \right). \tag{13}$$

The second term in the above bound, arising from R_{s,α,b^m} in (6) is bounded up to a constant independent of s, F and the number of QMC points, by

$$b^{-\alpha m} \|F\|_{s,\alpha,\boldsymbol{\gamma},q,\infty} H_{s,\boldsymbol{\gamma},q,\infty},$$

where $H_{s,\boldsymbol{\gamma},q,\infty}$ given in (10), and hence already converges with the optimal rate. Since the second term R_{s,α,b^m} is independent of the choice of $(q_1, \ldots, q_{s-1}, q_s)$, we focus on $B_{\boldsymbol{\gamma}}$. In the following we show that there is a component-by-component algorithm for SPOD weights such that $B_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q})$ is bounded by $C(b^m - 1)^{1/\lambda} \widetilde{J}_{s,\lambda,\boldsymbol{\gamma}}$, where $\widetilde{J}_{s,\lambda,\boldsymbol{\gamma}}$ is similar to $J_{s,\lambda,\boldsymbol{\gamma}}$ given in (10).

We need the following lemma, which is [20, Lemma 7].

Lemma 2.2. *For $\alpha \geq 2$ and $1/\alpha < \lambda \leq 1$, we have*

$$\sum_{k=1}^{\infty} b^{-\lambda \mu_\alpha(k)} = \sum_{w=1}^{\alpha-1} \prod_{i=1}^{w} \left(\frac{b-1}{b^{\lambda i} - 1} \right) + \left(\frac{b^{\lambda \alpha} - 1}{b^{\lambda \alpha} - b} \right) \prod_{i=1}^{\alpha} \left(\frac{b-1}{b^{\lambda i} - 1} \right) =: E_{\alpha,\lambda}.$$

We obtain the following extension of [9, Theorem 4.1] to SPOD weights.

Lemma 2.3. *Let $\boldsymbol{\beta}$ be a sequence of non-increasing, non-negative real numbers. For $\mathfrak{u} \subset \mathbb{N}$ with $|\mathfrak{u}| < \infty$ let $\gamma_{\mathfrak{u}}$ be given by (11).*

Let $\alpha, s \in \mathbb{N}$, b be a prime number and let $\mathfrak{p} \in \mathbb{F}_b[x]$ be an irreducible polynomial of degree $m \in \mathbb{N}$. Assume that $q_1^*, q_2^*, \ldots, q_s^* \in \mathbb{F}_b[x]$ were constructed using a component-by-component algorithm based on the criterion (12).

Then, for any $1/\alpha < \lambda \leq 1$ with $E = C_\alpha c_3 E_{\alpha,1} \alpha^{\alpha c_2}$ we have

$$B_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}^*) \leq \frac{1}{(b^m - 1)^{1/\lambda}} \left(\sum_{\mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}}^\lambda C_\alpha^{\lambda |\mathfrak{u}|} E_{\alpha,\lambda}^{|\mathfrak{u}|} \prod_{j \notin \mathfrak{u}} \left(1 + E \sum_{\nu=1}^{\alpha} ((j + c_1/\alpha)^{c_2} \beta_j)^\nu \right)^\lambda \right)^{1/\lambda}.$$

The proof follows along the lines of the proof of [9, Theorem 4], with some modifications to avoid the obstruction outlined in [9, Remark 4.2].

Proof. Without loss of generality we may assume that the modulus $\mathfrak{p} \in \mathbb{F}_b[x]$ is monic. We prove the result by induction on s. The dual polynomial lattice for $q^* = 1$ is given by

$$P^{\perp}(\mathfrak{p}, 1) = \{k \in \mathbb{N}_0 : \mathrm{tr}_m(k) = 0 \pmod{\mathfrak{p}}\} = \{k \in \mathbb{N}_0 : b^m | k\}.$$

Hence we have

$$B_{\boldsymbol{\gamma}}(\mathfrak{p}, 1) = C_\alpha \gamma_1 \sum_{\substack{k \in P^{\perp}(\mathfrak{p},1)\setminus\{0\} \\ b^m \nmid k}} b^{-\mu_\alpha(k)} = 0.$$

Now assume that we have already fixed the first $d-1$ components $\boldsymbol{q}_{d-1}^* = (q_1^*, \ldots, q_{d-1}^*) \in (G_{b,m}^*)^{d-1}$, $2 \le d \le s$ of the generating vector, such that

$$(B_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}_{d-1}^*))^\lambda \le \frac{1}{b^m - 1} \sum_{\mathfrak{u} \subseteq \{1:d-1\}} \gamma_{\mathfrak{u}}^\lambda C_\alpha^{\lambda|\mathfrak{u}|} E_{\alpha,\lambda}^{|\mathfrak{u}|} \prod_{j \notin \mathfrak{u}} \left(1 + E \sum_{\nu=1}^{\alpha} ((j + c_1/\alpha)^{c_2} \beta_j)^\nu\right)^\lambda$$

holds for any $1/\alpha < \lambda \le 1$. Put $\boldsymbol{q}_d = (\boldsymbol{q}_{d-1}^*, q_d)$ with $q_d \in G_{b,m}^* := \{q \in \mathbb{F}_b[x] \colon \deg(q) < m\} \setminus \{0\}$. Then we have

$$\begin{aligned}
B_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}_d) = & \sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:d-1\}} \gamma_{\mathfrak{u}} C_\alpha^{|\mathfrak{u}|} \sum_{\substack{\boldsymbol{k}_{\mathfrak{u}} \in P_{\mathfrak{u}}^{\perp}(\mathfrak{p}, \boldsymbol{q}_d) \\ \exists j \in \mathfrak{u} \colon b^m \nmid k_j}} b^{-\mu_\alpha(\boldsymbol{k}_{\mathfrak{u}})} \\
& + \sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:d-1\}} \gamma_{\mathfrak{u} \cup \{d\}} C_\alpha^{|\mathfrak{u}|+1} \sum_{\substack{\boldsymbol{k}_{\mathfrak{u} \cup \{d\}} \in P_{\mathfrak{u} \cup \{d\}}^{\perp}(\mathfrak{p}, \boldsymbol{q}_d) \\ \exists j \in \mathfrak{u} \colon b^m \nmid k_j \\ b^m | k_d}} b^{-\mu_\alpha(\boldsymbol{k}_{\mathfrak{u} \cup \{d\}})} \\
& + \sum_{\mathfrak{u} \subseteq \{1:d-1\}} \gamma_{\mathfrak{u} \cup \{d\}} C_\alpha^{|\mathfrak{u}|+1} \sum_{\substack{\boldsymbol{k}_{\mathfrak{u} \cup \{d\}} \in P_{\mathfrak{u} \cup \{d\}}^{\perp}(\mathfrak{p}, \boldsymbol{q}_d) \\ b^m \nmid k_d}} b^{-\mu_\alpha(\boldsymbol{k}_{\mathfrak{u} \cup \{d\}})} \\
= & B_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}_{d-1}^*) + \sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:d-1\}} \gamma_{\mathfrak{u} \cup \{d\}} C_\alpha^{|\mathfrak{u}|+1} \sum_{\substack{\boldsymbol{k}_{\mathfrak{u}} \in P_{\mathfrak{u}}^{\perp}(\mathfrak{p}, \boldsymbol{q}_{d-1}^*) \\ \exists j \in \mathfrak{u} \colon b^m \nmid k_j}} \sum_{\substack{k_d \in \mathbb{N} \\ b^m | k_d}} b^{-\mu_\alpha(\boldsymbol{k}_{\mathfrak{u}}, k_d)} \\
& + \sum_{\mathfrak{u} \subseteq \{1:d-1\}} \gamma_{\mathfrak{u} \cup \{d\}} C_\alpha^{|\mathfrak{u}|+1} \sum_{\substack{\boldsymbol{k}_{\mathfrak{u} \cup \{d\}} \in P_{\mathfrak{u} \cup \{d\}}^{\perp}(\mathfrak{p}, \boldsymbol{q}_d) \\ b^m \nmid k_d}} b^{-\mu_\alpha(\boldsymbol{k}_{\mathfrak{u} \cup \{d\}})} \\
\le & B_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}_{d-1}^*) \left(1 + \underbrace{\sum_{\substack{k_d \in \mathbb{N} \\ b^m | k_d}} b^{-\mu_\alpha(k_d)} \sum_{\nu_d=1}^{\alpha} \left(\frac{(\alpha(d-1) + c_1 + \nu_d)!}{(\alpha(d-1) + c_1)!}\right)^{c_2} C_\alpha c_3 \beta_d^{\nu_d}}_{=: A(\beta_d)}\right) \\
& + \sum_{\mathfrak{u} \subseteq \{1:d-1\}} \gamma_{\mathfrak{u} \cup \{d\}} C_\alpha^{|\mathfrak{u}|+1} \sum_{\substack{\boldsymbol{k}_{\mathfrak{u} \cup \{d\}} \in P_{\mathfrak{u} \cup \{d\}}^{\perp}(\mathfrak{p}, \boldsymbol{q}_d) \\ b^m \nmid k_d}} b^{-\mu_\alpha(\boldsymbol{k}_{\mathfrak{u} \cup \{d\}})}, \qquad (14)
\end{aligned}$$

where the second equality stems from the fact that since $b^m \mid k_d$, we have $\mathrm{tr}_m(k_d) = 0$ and thus $\mathrm{tr}_m(\boldsymbol{k}_{\mathfrak{u} \cup \{d\}}) \cdot (\boldsymbol{q}_{\mathfrak{u}}^*, q_d) = \mathrm{tr}_m(\boldsymbol{k}_{\mathfrak{u}}) \cdot \boldsymbol{q}_{\mathfrak{u}}^*$, which yields

$$\{\boldsymbol{k}_{\mathfrak{u} \cup \{d\}} \in P_{\mathfrak{u} \cup \{d\}}^{\perp}(\mathfrak{p}, \boldsymbol{q}_d) \colon b^m \mid k_d\} = \{(\boldsymbol{k}_{\mathfrak{u}}, k_d) \in \mathbb{N}^{|\mathfrak{u}|+1} \colon \boldsymbol{k}_{\mathfrak{u}} \in P_{\mathfrak{u}}^{\perp}(\mathfrak{p}, \boldsymbol{q}_{d-1}^*), b^m \mid k_d\}.$$

In the last step we used the estimate

$$\gamma_{\mathfrak{u}\cup\{d\}} \leq \sum_{\boldsymbol{\nu}\in\{1:\alpha\}^{|\mathfrak{u}|}} ((|\boldsymbol{\nu}|+c_1)!)^{c_2} \left[\prod_{j\in\mathfrak{u}} c_3\beta_j^{\nu_j}\right] \sum_{\nu_d=1}^{\alpha} \left(\frac{(\alpha(d-1)+c_1+\nu_d)!}{(\alpha(d-1)+c_1)!}\right)^{c_2} c_3\beta_d^{\nu_d}.$$

It is clear that the first term of (14) does not depend on the choice of q_d. Thus, denoting the second term of (14) by

$$\psi_{\mathfrak{p},\boldsymbol{q}_{d-1}^*}(q_d) := \sum_{\mathfrak{u}\subseteq\{1:d-1\}} \gamma_{\mathfrak{u}\cup\{d\}} C_\alpha^{|\mathfrak{u}|+1} \sum_{\substack{\boldsymbol{k}_{\mathfrak{u}\cup\{d\}}\in P^\perp_{\mathfrak{u}\cup\{d\}}(\mathfrak{p},\boldsymbol{q}_d)\\ b^m\nmid k_d}} b^{-\mu_\alpha(\boldsymbol{k}_{\mathfrak{u}\cup\{d\}})},$$

we have

$$q_d^* = \arg\min_{q_d\in G^*_{b,m}} B_{\boldsymbol{\gamma}}(\mathfrak{p},\boldsymbol{q}_d) = \arg\min_{q_d\in G^*_{b,m}} \psi_{\mathfrak{p},\boldsymbol{q}^*_{d-1}}(q_d).$$

Using Jensen's inequality, as long as $1/\alpha<\lambda\leq 1$, we have

$$\begin{aligned}
&(\psi_{\mathfrak{p},\boldsymbol{q}^*_{d-1}}(q_d^*))^\lambda\\
&\leq \frac{1}{b^m-1}\sum_{q_d\in G^*_{b,m}} (\psi_{\mathfrak{p},\boldsymbol{q}^*_{d-1}}(q_d))^\lambda\\
&\leq \frac{1}{b^m-1}\sum_{q_d\in G^*_{b,m}}\sum_{\mathfrak{u}\subseteq\{1:d-1\}} \gamma^\lambda_{\mathfrak{u}\cup\{d\}} C_\alpha^{\lambda(|\mathfrak{u}|+1)} \sum_{\substack{\boldsymbol{k}_{\mathfrak{u}\cup\{d\}}\in P^\perp_{\mathfrak{u}\cup\{d\}}(\mathfrak{p},\boldsymbol{q}_d)\\ b^m\nmid k_d}} b^{-\lambda\mu_\alpha(\boldsymbol{k}_{\mathfrak{u}\cup\{d\}})}\\
&= \frac{1}{b^m-1}\sum_{\mathfrak{u}\subseteq\{1:d-1\}} \gamma^\lambda_{\mathfrak{u}\cup\{d\}} C_\alpha^{\lambda(|\mathfrak{u}|+1)} \sum_{\substack{\boldsymbol{k}_{\mathfrak{u}\cup\{d\}}\in \mathbb{N}^{|\mathfrak{u}|+1}\\ b^m\nmid k_d}} b^{-\lambda\mu_\alpha(\boldsymbol{k}_{\mathfrak{u}\cup\{d\}})}\\
&\qquad\times \sum_{\substack{q_d\in G^*_{b,m}\\ \mathrm{tr}_m(\boldsymbol{k}_\mathfrak{u})\cdot\boldsymbol{q}^*_\mathfrak{u}+\mathrm{tr}_m(k_d)q_d=0 \pmod{\mathfrak{p}}}} 1.
\end{aligned}$$

Since $b^m\nmid k_d$, we have $\mathrm{tr}_m(k_d)\neq 0$. For $\boldsymbol{k}_\mathfrak{u}\in P^\perp_\mathfrak{u}(\mathfrak{p},\boldsymbol{q}^*_{d-1})$, it follows from the definition of the dual polynomial lattice that $\mathrm{tr}_m(\boldsymbol{k}_\mathfrak{u})\cdot\boldsymbol{q}^*_\mathfrak{u}=0 \pmod{\mathfrak{p}}$, and thus there is no polynomial $q_d\in G^*_{b,m}$ such that the condition $\mathrm{tr}_m(k_d)q_d=0 \pmod{\mathfrak{p}}$ is satisfied. For $\boldsymbol{k}_\mathfrak{u}\notin P^\perp_\mathfrak{u}(\mathfrak{p},\boldsymbol{q}^*_{d-1})$, there exists exactly one $q_d\in G^*_{b,m}$ such that $\mathrm{tr}_m(k_d)q_d=-\mathrm{tr}_m(\boldsymbol{k}_\mathfrak{u})\cdot\boldsymbol{q}^*_\mathfrak{u} \pmod{\mathfrak{p}}$. From these facts and Lemma 2.2, we obtain

$$\begin{aligned}
(\psi_{\mathfrak{p},\boldsymbol{q}^*_{d-1}}(q_d^*))^\lambda &\leq \frac{1}{b^m-1}\sum_{\mathfrak{u}\subseteq\{1:d-1\}} \gamma^\lambda_{\mathfrak{u}\cup\{d\}} C_\alpha^{\lambda(|\mathfrak{u}|+1)} \sum_{\substack{\boldsymbol{k}_\mathfrak{u}\in\mathbb{N}^{|\mathfrak{u}|}\\ \boldsymbol{k}_\mathfrak{u}\notin P^\perp_\mathfrak{u}(\mathfrak{p},\boldsymbol{q}^*_{d-1})}} \sum_{\substack{k_d\in\mathbb{N}\\ b^m\nmid k_d}} b^{-\lambda\mu_\alpha(\boldsymbol{k}_\mathfrak{u},k_d)}\\
&\leq \frac{1}{b^m-1}\sum_{\mathfrak{u}\subseteq\{1:d-1\}} \gamma^\lambda_{\mathfrak{u}\cup\{d\}} C_\alpha^{\lambda(|\mathfrak{u}|+1)} \sum_{\boldsymbol{k}_\mathfrak{u}\in\mathbb{N}^{|\mathfrak{u}|}} b^{-\lambda\mu_\alpha(\boldsymbol{k}_\mathfrak{u})} \sum_{\substack{k_d\in\mathbb{N}\\ b^m\nmid k_d}} b^{-\lambda\mu_\alpha(k_d)}\\
&= \frac{1}{b^m-1}\sum_{\mathfrak{u}\subseteq\{1:d-1\}} \gamma^\lambda_{\mathfrak{u}\cup\{d\}} C_\alpha^{\lambda(|\mathfrak{u}|+1)} E_{\alpha,\lambda}^{|\mathfrak{u}|+1}.
\end{aligned}$$

We now study the expression $A(\beta_d)$ from (14) in more detail. The sum over k_d is bounded by $E_{\alpha,1}$ from Lemma 2.2. Then we have

$$A(\beta_d) \leq C_\alpha c_3 E_{\alpha,1}\sum_{\nu_d=1}^{\alpha}\prod_{\ell=1}^{\nu_d}\beta_d(\alpha(d-1)+c_1+\ell)^{c_2}.$$

Hence

$$A(\beta_d) \le C_\alpha c_3 E_{\alpha,1} \sum_{\nu_d=1}^{\alpha} \prod_{\ell=1}^{\nu_d} \beta_d(\alpha d + c_1 + \ell - \alpha)^{c_2} \le C_\alpha c_3 E_{\alpha,1} \alpha^{\alpha c_2} \sum_{\nu_d=1}^{\alpha} ((d + c_1/\alpha)^{c_2} \beta_d)^{\nu_d}.$$

To simplify the notation we collect all the constants in a new constant $E = C_\alpha c_3 E_{\alpha,1} \alpha^{\alpha c_2}$.

Finally by applying Jensen's inequality to (14) and using Lemma 2.2, we have

$$\begin{aligned}(B_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}_d^*))^\lambda &\le (B_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}_{d-1}^*))^\lambda \left(1 + E \sum_{\nu_d=1}^{\alpha} ((d + c_1/\alpha)^{c_2} \beta_d)^{\nu_d}\right)^\lambda \\ &\quad + \frac{1}{b^m - 1} \sum_{\mathfrak{u} \subseteq \{1:d-1\}} \gamma_{\mathfrak{u} \cup \{d\}}^\lambda C_\alpha^{\lambda(|\mathfrak{u}|+1)} E_{\alpha,\lambda}^{|\mathfrak{u}|+1} \\ &\le \frac{1}{b^m - 1} \sum_{\mathfrak{u} \subseteq \{1:d\}} \gamma_{\mathfrak{u}}^\lambda C_\alpha^{\lambda|\mathfrak{u}|} E_{\alpha,\lambda}^{|\mathfrak{u}|} \prod_{j \in \{1:d\} \setminus \mathfrak{u}} \left(1 + E \sum_{\nu=1}^{\alpha} ((j + c_1/\alpha)^{c_2} \beta_j)^\nu\right)^\lambda.\end{aligned}$$

This completes the proof. □

Theorem 2.4. *Let $c_2 \ge 0$ and $\boldsymbol{\beta}$ be a sequence of non-increasing, non-negative real numbers such that*

$$\sum_{j=1}^{\infty} j^{c_2} \beta_j < \infty.$$

For $\mathfrak{u} \subset \mathbb{N}$ with $|\mathfrak{u}| < \infty$ let $\gamma_{\mathfrak{u}}$ be given by (11). Let $\alpha \in \mathbb{N}$, b be a prime number and $\mathfrak{p} \in \mathbb{F}_b[x]$ be an irreducible polynomial of degree $m \in \mathbb{N}$. Assume that $q_1^*, q_2^*, \ldots, q_s^* \in \mathbb{F}_b[x]$ were constructed using a component-by-component algorithm based on the criterion (12). Then, for any $1/\alpha < \lambda \le 1$ we have

$$B_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}^*) \le \frac{K}{(b^m - 1)^{1/\lambda}} \left(\sum_{\mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}}^\lambda C_\alpha^{\lambda|\mathfrak{u}|} E_{\alpha,\lambda}^{|\mathfrak{u}|}\right)^{1/\lambda},$$

where the constant $K := \prod_{j=1}^{\infty} \left(1 + E \sum_{\nu=1}^{\alpha} ((j + c_1/\alpha)^{c_2} \beta_j)^\nu\right)$ is independent of b, m, s, λ.

Proof. We bound the term

$$\begin{aligned}\prod_{j \notin \mathfrak{u}} \left(1 + E \sum_{\nu=1}^{\alpha} ((j + c_1/\alpha)^{c_2} \beta_j)^\nu\right) &\le \prod_{j=1}^{\infty} \left(1 + E \sum_{\nu=1}^{\alpha} ((j + c_1/\alpha)^{c_2} \beta_j)^\nu\right) \\ &\le \exp\left(E \sum_{\nu=1}^{\alpha} \sum_{j=1}^{\infty} ((j + c_1/\alpha)^{c_2} \beta_j)^\nu\right) < \infty,\end{aligned}$$

where we used the inequality $\log(1 + x) \le x$ for $x > 0$. The assumption $\sum_{j=1}^{\infty} (j + c_1/\alpha)^{c_2} \beta_j < \infty$ implies that $\sum_{j=1}^{\infty} ((j + c_1/\alpha)^{c_2} \beta_j)^\nu < \infty$ for any $\nu \ge 1$. Further we have for any $j \in \mathbb{N}$ that

$$(j + c_1/\alpha)^{c_2} \le (1 + c_1/\alpha)^{c_2} j^{c_2}.$$

Hence $\sum_{j=1}^{\infty} j^{c_2} \beta_j < \infty$ implies that $\sum_{j=1}^{\infty} (j + c_1/\alpha)^{c_2} \beta_j < \infty$. □

Theorem 2.5. *Let $\boldsymbol{\beta}$ be a sequence of non-increasing, non-negative real numbers. Let $c_2 > 0$ and $0 < p < 1/(1+c_2)$ such that $\sum_{j=1}^{\infty} \beta_j^p < \infty$. For $\mathfrak{u} \subset \mathbb{N}$ with $|\mathfrak{u}| < \infty$ let $\gamma_{\mathfrak{u}}$ be given by (11). Let $\alpha = 1 + \lfloor 1/p \rfloor$, b be a prime number and $\mathfrak{p} \in \mathbb{F}_b[x]$ be an irreducible polynomial of degree $m \in \mathbb{N}$. Assume that $q_1^*, q_2^*, \ldots, q_s^* \in \mathbb{F}_b[x]$ were constructed using a component-by-component algorithm based on the criterion (12). Then for any $p \leq \lambda < 1/c_2$ there is a constant $C(\lambda) > 0$, which does not depend on s, m, such that*

$$B_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}^*) \leq \frac{C(\lambda)}{b^{m/\lambda}}.$$

Proof. The bound on the sum $\sum_{\mathfrak{u}} \gamma_{\mathfrak{u}}^{\lambda} C_{\alpha}^{\lambda|\mathfrak{u}|} E_{\alpha,\lambda}^{|\mathfrak{u}|}$ follows as in [10, Section 3]. In order to obtain a bound which is independent of the dimension, we need to bound $\sum_{\mathfrak{u} \subset \mathbb{N}, |\mathfrak{u}| < \infty} \gamma_{\mathfrak{u}}^{\lambda} C_{\alpha}^{\lambda|\mathfrak{u}|} E_{\alpha,\lambda}^{|\mathfrak{u}|}$. Let $B := \max\{1, c_3 C_\alpha E_{\alpha,\lambda}^{1/\lambda}\}$ and define $\tilde{\gamma}_1, \tilde{\gamma}_2, \ldots$ to be the sequence

$$\underbrace{B\beta_1, B\beta_1, \ldots, B\beta_1}_{\alpha \text{ times}}, \underbrace{B\beta_2, B\beta_2, \ldots, B\beta_2}_{\alpha \text{ times}}, \ldots$$

i.e., $\tilde{\gamma}_1 = \cdots = \tilde{\gamma}_\alpha = B\beta_1$, $\tilde{\gamma}_{\alpha+1} = \cdots = \tilde{\gamma}_{2\alpha} = B\beta_2$, $\ldots$. Then $\sum_{j=1}^{\infty} \beta_j^p < \infty$ if and only if $\sum_{j=1}^{\infty} \tilde{\gamma}_j^p < \infty$. We have

$$\begin{aligned}
\sum_{\substack{\mathfrak{u} \subset \mathbb{N} \\ |\mathfrak{u}| < \infty}} \gamma_{\mathfrak{u}}^{\lambda} C_{\alpha}^{\lambda|\mathfrak{u}|} E_{\alpha,\lambda}^{|\mathfrak{u}|} &\leq \sum_{\substack{\mathfrak{u} \subset \mathbb{N} \\ |\mathfrak{u}| < \infty}} \sum_{\boldsymbol{\nu} \in \{1:\alpha\}^{|\mathfrak{u}|}} ((|\boldsymbol{\nu}| + c_1)!)^{c_2\lambda} \prod_{j \in \mathfrak{u}} (B\beta_j^{\nu_j})^{\lambda} \\
&\leq \sum_{\substack{\mathfrak{v} \subset \mathbb{N} \\ |\mathfrak{v}| < \infty}} ((|\mathfrak{v}| + c_1)!)^{c_2\lambda} \prod_{j \in \mathfrak{v}} \tilde{\gamma}_j^{\lambda} \leq \sum_{\ell=0}^{\infty} ((\ell + c_1)!)^{c_2\lambda} \frac{1}{\ell!} \left(\sum_{j=1}^{\infty} \tilde{\gamma}_j^{\lambda} \right)^{\ell}.
\end{aligned}$$

As long as $\lambda \geq p$, the sum $S = \sum_{j=1}^{\infty} \tilde{\gamma}_j^{\lambda} < \infty$. From Stirling's formula we have

$$\begin{aligned}
\frac{((\ell + c_1)!)^{c_2\lambda}}{\ell!} &\asymp \frac{(\ell + c_1)^{(\ell + c_1 + 1/2)c_2\lambda} \mathrm{e}^{-\ell c_2 \lambda}}{\ell^{\ell + 1/2} \mathrm{e}^{-\ell}} \\
&\asymp \frac{(\ell + c_1)^{\ell c_2 \lambda}}{\ell^{\ell}} \mathrm{e}^{\ell(1 - c_2\lambda)} \frac{(\ell + c_1)^{(c_1 + 1/2)c_2\lambda}}{\ell^{1/2}}, \quad \text{as } \ell \to \infty.
\end{aligned}$$

This expression converges to 0 superexponentially fast as long as $c_2\lambda < 1$. Hence

$$\sum_{\substack{\mathfrak{u} \subset \mathbb{N} \\ |\mathfrak{u}| < \infty}} \gamma_{\mathfrak{u}}^{\lambda} C_{\alpha}^{\lambda|\mathfrak{u}|} E_{\alpha,\lambda}^{|\mathfrak{u}|} < \infty \tag{15}$$

for any $p \leq \lambda < 1/c_2$. We now show that $\sum_{j=1}^{\infty} \beta_j^p < \infty$ for some $0 < p < 1/(1+c_2)$ implies that $\sum_{j=1}^{\infty} j^{c_2} \beta_j < \infty$. We have $j\beta_j^p \leq \sum_{i=1}^{j} \beta_i^p$ and therefore $\beta_j \leq C j^{-1/p}$, where $C = (\sum_{j=1}^{\infty} \beta_j^p)^{1/p}$. Hence

$$\sum_{j=1}^{\infty} j^{c_2} \beta_j \leq C \sum_{j=1}^{\infty} j^{c_2 - 1/p}.$$

Now $p < 1/(1+c_2)$ implies that $c_2 - 1/p < -1$ and the result follows. □

Remark 2.6. *To have a guaranteed convergence rate of the QMC approximation of $1/\lambda$, we have the constraints*

- $1/\alpha < \lambda \leq 1$, *coming from the CBC construction*

- $p < \frac{1}{c_2+1}$, *to verify the summability hypothesis of Theorem 2.4*

- $p \le \lambda < \frac{1}{c_2}$, *for the summability required in Theorem 2.5.*

Therefore, in the case $c_2 = 1$ *and* $\alpha = 2$, *we also obtain convergence order arbitrarily close to* $\mathcal{O}(N^{-2})$ *provided that* $p < \frac{1}{2}$.

Remark 2.7 (Low summability – Part I)**.** *Under the same assumptions of Lemma 2.3, we also have, the estimate*

$$B_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}^*) \le \frac{1}{b^m - 1} \sum_{\mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}} C_\alpha^{|\mathfrak{u}|} E_{\alpha,1}^{|\mathfrak{u}|}. \tag{16}$$

Therefore, following the arguments from Theorem 2.5, given SPOD weights (11) *such that* $c_2 = 1$ *and* $\boldsymbol{\beta} \in \ell^1(\mathbb{N})$ *with the smallness condition (analogous to [10, Equation (3.42)])*

$$\|\boldsymbol{\beta}\|_{\ell^1(\mathbb{N})} < \frac{1}{\alpha \max\{1, c_3 C_\alpha E_{\alpha,1}\}}, \tag{17}$$

this yields

$$\sum_{\mathfrak{u} \subseteq \{1:s\}} \gamma_{\mathfrak{u}} C_\alpha^{|\mathfrak{u}|} E_{\alpha,1}^{|\mathfrak{u}|} \le \sum_{\ell=0}^{\infty} \frac{(\ell + c_1)!}{\ell!} \left(\alpha \max\{1, c_3 C_\alpha E_{\alpha,1}\} \sum_{j=1}^{\infty} \beta_j \right)^{\ell} < \infty. \tag{18}$$

The same argument applies to $H_{s,\boldsymbol{\gamma},q,\infty}$ *from* (10)*. Hence, we obtain at least first order convergence of the QMC quadrature with extrapolated polynomial lattices. In particular, the constraint* $\boldsymbol{\beta} \in \ell^p(\mathbb{N})$, $p < \frac{1}{1+c_2} = \frac{1}{2}$ *can be omitted in this case.*

Proof of (16). In the proof of Lemma 2.3 we observe

$$\begin{aligned}
B_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}_d) &= B_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}_{d-1}^*) + \sum_{\emptyset \ne \mathfrak{u} \subseteq \{1:d-1\}} \gamma_{\mathfrak{u} \cup \{d\}} C_\alpha^{|\mathfrak{u}|+1} \sum_{\substack{\boldsymbol{k}_{\mathfrak{u}} \in P_{\mathfrak{u}}^{\perp}(\mathfrak{p}, \boldsymbol{q}_{d-1}^*) \\ \exists j \in \mathfrak{u}:\, b^m \nmid k_j}} \sum_{\substack{k_d \in \mathbb{N} \\ b^m | k_d}} b^{-\mu_\alpha(\boldsymbol{k}_{\mathfrak{u}}, k_d)} \\
&\quad + \sum_{\mathfrak{u} \subseteq \{1:d-1\}} \gamma_{\mathfrak{u} \cup \{d\}} C_\alpha^{|\mathfrak{u}|+1} \sum_{\substack{\boldsymbol{k}_{\mathfrak{u} \cup \{d\}} \in P_{\mathfrak{u} \cup \{d\}}^{\perp}(\mathfrak{p}, \boldsymbol{q}_d) \\ b^m \nmid k_d}} b^{-\mu_\alpha(\boldsymbol{k}_{\mathfrak{u} \cup \{d\}})} \\
&=: B_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}_{d-1}^*) + \phi_{\mathfrak{p}, \boldsymbol{q}_{d-1}^*} + \psi_{\mathfrak{p}, \boldsymbol{q}_{d-1}^*}(q_d),
\end{aligned} \tag{19}$$

where the first two terms of (19) do not depend on the choice of q_d. In the same lemma we also show that for the choice $q_d = q_d^*$,

$$\psi_{\mathfrak{p}, \boldsymbol{q}_{d-1}^*}(q_d^*) \le \frac{1}{b^m - 1} \sum_{\mathfrak{u} \subseteq \{1:d-1\}} \gamma_{\mathfrak{u} \cup \{d\}} C_\alpha^{|\mathfrak{u}|+1} E_{\alpha,1} \sum_{\substack{\boldsymbol{k}_{\mathfrak{u}} \in \mathbb{N}^{|\mathfrak{u}|} \\ \boldsymbol{k}_{\mathfrak{u}} \notin P_{\mathfrak{u}}^{\perp}(\mathfrak{p}, \boldsymbol{q}_{d-1}^*)}} b^{-\mu_\alpha(\boldsymbol{k}_{\mathfrak{u}})}.$$

Next, using the estimate $\mu_\alpha(kb^m) \ge m + \mu_\alpha(k)$ for all $k \in \mathbb{N}$ and Lemma 2.2, we bound

$$\begin{aligned}
\phi_{\mathfrak{p}, \boldsymbol{q}_{d-1}^*} &= \sum_{\emptyset \ne \mathfrak{u} \subseteq \{1:d-1\}} \gamma_{\mathfrak{u} \cup \{d\}} C_\alpha^{|\mathfrak{u}|+1} \sum_{\substack{\boldsymbol{k}_{\mathfrak{u}} \in P_{\mathfrak{u}}^{\perp}(\mathfrak{p}, \boldsymbol{q}_{d-1}^*) \\ \exists j \in \mathfrak{u}:\, b^m \nmid k_j}} b^{-\mu_\alpha(\boldsymbol{k}_{\mathfrak{u}})} \sum_{k \in \mathbb{N}} b^{-\mu_\alpha(kb^m)} \\
&\le b^{-m} \sum_{\emptyset \ne \mathfrak{u} \subseteq \{1:d-1\}} \gamma_{\mathfrak{u} \cup \{d\}} C_\alpha^{|\mathfrak{u}|+1} \sum_{\substack{\boldsymbol{k}_{\mathfrak{u}} \in P_{\mathfrak{u}}^{\perp}(\mathfrak{p}, \boldsymbol{q}_{d-1}^*) \\ \exists j \in \mathfrak{u}:\, b^m \nmid k_j}} b^{-\mu_\alpha(\boldsymbol{k}_{\mathfrak{u}})} \sum_{k \in \mathbb{N}} b^{-\mu_\alpha(k)}
\end{aligned}$$

$$= b^{-m} \sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:d-1\}} \gamma_{\mathfrak{u}\cup\{d\}} C_\alpha^{|\mathfrak{u}|+1} E_{\alpha,1} \sum_{\substack{\boldsymbol{k}_\mathfrak{u} \in P_\mathfrak{u}^\perp(\mathfrak{p}, \boldsymbol{q}_{d-1}^*) \\ \exists j \in \mathfrak{u}:\, b^m \nmid k_j}} b^{-\mu_\alpha(\boldsymbol{k}_\mathfrak{u})}. \tag{20}$$

Therefore, we obtain

$$\psi_{\mathfrak{p},\boldsymbol{q}_{d-1}^*}(q_d^*) + \phi_{\mathfrak{p},\boldsymbol{q}_{d-1}^*} \le \frac{1}{b^m - 1} \sum_{\mathfrak{u} \subseteq \{1:d-1\}} \gamma_{\mathfrak{u}\cup\{d\}} C_\alpha^{|\mathfrak{u}|+1} E_{\alpha,1} \sum_{\boldsymbol{k}_\mathfrak{u} \in \mathbb{N}^{|\mathfrak{u}|}} b^{-\mu_\alpha(\boldsymbol{k}_\mathfrak{u})}$$
$$= \frac{1}{b^m - 1} \sum_{\mathfrak{u} \subseteq \{1:d-1\}} \gamma_{\mathfrak{u}\cup\{d\}} C_\alpha^{|\mathfrak{u}|+1} E_{\alpha,1}^{|\mathfrak{u}|+1},$$

and with the inductive hypothesis this completes the proof. □

Remark 2.8 (Low summability – Part II)**.** *Applying the same estimate $\mu_\alpha(kb^m) \ge m + \mu_\alpha(k)$ to $A(\beta_d)$ defined in (14), we observe that the result of Lemma 2.3 is valid with the constant E replaced by Eb^{-m}. Hence, under the additional assumption that $s \le b^m < \infty$, the (dimension independent) convergence $B_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}^*) \le C(p) b^{-m/p}$ from Theorem 2.5 is also attained for $\boldsymbol{\beta} \in \ell^p(\mathbb{N})$, $p \in [\frac{1}{c_2+1}, \frac{1}{c_2})$, due to (15) and*

$$\prod_{j \in \{1:s\}\setminus\mathfrak{u}} \left(1 + Eb^{-m} \sum_{\nu=1}^{\alpha} ((j + c_1/\alpha)^{c_2} \beta_j)^\nu \right) \le \exp\left(E \sum_{\nu=1}^{\alpha} \max_{j \in \mathbb{N}} ((j + c_1/\alpha)^{c_2} \beta_j)^\nu \right) < \infty.$$

Since in practical applications the parameter space U is truncated to finite dimension $s < \infty$, the growth $m \ge \log_b s$ is a mild requirement that can be enforced easily in many situations, in particular in the parametric PDE setting of Theorem 3.8 below.

2.5 Fast component-by-component construction

We want to apply the fast CBC construction for SPOD weights for the construction of extrapolated polynomial lattice rules from [9]. The criterion in (12) is of the same form as the criterion $E_s^2(z_s)$ in [29, Section 5], so the fast CBC construction with POD weights can be performed in the same way as described there.

The general form of the SPOD weights (11) can be written as $\gamma_\emptyset = 1$ and, for any $\emptyset \neq \mathfrak{u} \subseteq \{1 : s\}$,

$$\gamma_\mathfrak{u} = \sum_{\boldsymbol{\nu} \in \{1:\bar{\alpha}\}^{|\mathfrak{u}|}} \Gamma_{|\boldsymbol{\nu}|} \prod_{j \in \mathrm{supp}(\boldsymbol{\nu})} \gamma_j(\nu_j),$$

where $\mathrm{supp}(\boldsymbol{\nu}) = \{j \in \{1 : s\} : \gamma_j(\nu_j) \neq 0\}$, and where $\gamma_j(\nu_j)$ is a non-negative real number which may depend on ν_j (cf. (36)).

The POD weights $\gamma_\mathfrak{u} := \Gamma_{|\mathfrak{u}|} \prod_{j\in\mathfrak{u}} \gamma_j$ correspond to the case of $\bar{\alpha} = 1$. For applications to PDEs with globally supported uncertain coefficients, we have $\bar{\alpha} = \alpha$ as in (11). However, in order to have greater flexibility of the results in this section, we distinguish α corresponding to the maximum derivative order in (5) and appearing in the Walsh bound, from $\bar{\alpha}$ for the parameter in the SPOD weights. By choosing the parameter $\bar{\alpha} = 1$ we obtain results for POD weights and by setting $\bar{\alpha} = \alpha$ we obtain results for SPOD weights.

As in [9, p.64], we perform the CBC construction for $d = 1, \ldots, s$ by adding the terms that do not depend on the new component q_d. We thus employ the following search criterion in the CBC algorithm,

$$\tilde{B}_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}_d) := B_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}_d) + \sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:d\}} \gamma_\mathfrak{u} C_\alpha^{|\mathfrak{u}|} \sum_{\substack{\boldsymbol{k}_\mathfrak{u} \in P_\mathfrak{u}^\perp(\mathfrak{p}, \boldsymbol{q}_d) \\ \forall j \in \mathfrak{u}:\, b^m \mid k_j}} b^{-\mu_\alpha(\boldsymbol{k}_\mathfrak{u})}$$

$$= \sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:d\}} \gamma_{\mathfrak{u}} C_\alpha^{|\mathfrak{u}|} \sum_{\boldsymbol{k}_{\mathfrak{u}} \in P_{\mathfrak{u}}^{\perp}(\mathfrak{p}, \boldsymbol{q}_d)} b^{-\mu_\alpha(\boldsymbol{k}_{\mathfrak{u}})}.$$

Therefore, using the dual lattice property we get

$$\begin{aligned}\tilde{B}_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}_d) &= \frac{1}{b^m} \sum_{n=0}^{b^m-1} \sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:d\}} \gamma_{\mathfrak{u}} C_\alpha^{|\mathfrak{u}|} \sum_{\boldsymbol{k}_{\mathfrak{u}} \in \mathbb{N}^{|\mathfrak{u}|}} b^{-\mu_\alpha(\boldsymbol{k}_{\mathfrak{u}})} \operatorname{wal}_{(\boldsymbol{k}_{\mathfrak{u}}, \boldsymbol{0})}(\boldsymbol{y}_n) \\ &= \frac{1}{b^m} \sum_{n=0}^{b^m-1} \sum_{\emptyset \neq \mathfrak{u} \subseteq \{1:d\}} \gamma_{\mathfrak{u}} C_\alpha^{|\mathfrak{u}|} \prod_{j \in \mathfrak{u}} \sum_{k \in \mathbb{N}} b^{-\mu_\alpha(k)} \operatorname{wal}_k(y_{n,j}).\end{aligned}$$

Define $w_\alpha(y) := \sum_{k \in \mathbb{N}} b^{-\mu_\alpha(k)} \operatorname{wal}_k(y)$ and $\Gamma_0 := 1$. Following the fast CBC construction in [10] we obtain

$$\begin{aligned}\tilde{B}_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}_d) &= -1 + \frac{1}{b^m} \sum_{n=0}^{b^m-1} \sum_{\mathfrak{u} \subseteq \{1:d\}} \sum_{\boldsymbol{\nu} \in \{1:\bar{\alpha}\}^{|\mathfrak{u}|}} \Gamma_{|\boldsymbol{\nu}|} \prod_{j \in \mathfrak{u}} \gamma_j(\nu_j) C_\alpha w_\alpha(y_{n,j}) \\ &= -1 + \frac{1}{b^m} \sum_{n=0}^{b^m-1} \sum_{l=0}^{\bar{\alpha}d} \Gamma_l \sum_{\substack{\boldsymbol{\nu} \in \{0:\bar{\alpha}\}^d \\ |\boldsymbol{\nu}| = l}} \prod_{j \in \operatorname{supp}(\boldsymbol{\nu})} \gamma_j(\nu_j) C_\alpha w_\alpha(y_{n,j}).\end{aligned}$$

Employing the convention that $U_{d,0}(n) := 1$ for all $d \in \mathbb{N}_0$, and $U_{d,l}(n) := 0$ for all $l > \bar{\alpha}d$, the definition

$$U_{d,l}(n) := \Gamma_l \sum_{\substack{\boldsymbol{\nu} \in \{0:\bar{\alpha}\}^d \\ |\boldsymbol{\nu}| = l}} \prod_{j \in \operatorname{supp}(\boldsymbol{\nu})} \gamma_j(\nu_j) C_\alpha w_\alpha(y_{n,j}) \tag{21}$$

implies

$$\tilde{B}_{\boldsymbol{\gamma}}(\mathfrak{p}, \boldsymbol{q}_d) = -1 + \frac{1}{b^m} \sum_{n=0}^{b^m-1} \sum_{l=0}^{\bar{\alpha}d} U_{d,l}(n). \tag{22}$$

We now isolate the summands that do not depend on the last component of the generating vector, that is the summands corresponding to $\boldsymbol{\nu}$ with $\nu_d = 0$. With the conventions above we obtain a recursive formula

$$\begin{aligned}U_{d,l}(n) =& U_{d-1,l}(n) + \Gamma_l \sum_{\nu_d=1}^{\min(\bar{\alpha},l)} \gamma_d(\nu_d) C_\alpha w_\alpha(y_{n,d}) \\ & \times \sum_{\substack{\boldsymbol{\nu} \in \{0:\bar{\alpha}\}^{d-1} \\ |\boldsymbol{\nu}| = l-\nu_d}} \prod_{j \in \operatorname{supp}(\boldsymbol{\nu})} \gamma_j(\nu_j) C_\alpha w_\alpha(y_{n,j}) \\ =& U_{d-1,l}(n) + w_\alpha(y_{n,d}) \sum_{\nu_d=1}^{\min(\bar{\alpha},l)} \gamma_d(\nu_d) C_\alpha \frac{\Gamma_l}{\Gamma_{l-\nu_d}} U_{d-1,l-\nu_d}(n) \\ =& U_{d-1,l}(n) + w_\alpha(y_{n,d}) V_{d,l}(n),\end{aligned} \tag{23}$$

where we defined

$$V_{d,l}(n) := \sum_{\nu_d=1}^{\min(\bar{\alpha},l)} \gamma_d(\nu_d) C_\alpha \frac{\Gamma_l}{\Gamma_{l-\nu_d}} U_{d-1,l-\nu_d}(n). \tag{24}$$

Therefore, the only term dependent on q_d in (22) is

$$\sum_{n=1}^{b^m-1} w_\alpha\left(v_m\left(\frac{q_d n}{\mathfrak{p}}\right)\right)\sum_{l=1}^{\bar{\alpha}d} V_{d,l}(n),$$

where $\frac{q_d n}{\mathfrak{p}}$ is computed in $\mathbb{F}_b((x^{-1}))$, i.e., with slight abuse of notation we identify $n \in \{1 : b^m - 1\}$ with $n(x) \in \mathbb{F}_b[x]$ defined in Section 2.1. Note that $n = 0$ is not included. Therefore, there exists a permutation Π of $n \in \{1 : b^m - 1\}$ that allows us to rewrite $q_d\Pi(n) = g^{z_d - n} \pmod{\mathfrak{p}}$ for some primitive element $g \in (\mathbb{F}_b[x]/\mathfrak{p}) \setminus \{0\}$, obtaining

$$\sum_{n=1}^{b^m-1} w_\alpha\left(v_m\left(\frac{g^{z_d-n}}{\mathfrak{p}}\right)\right)\sum_{l=1}^{\bar{\alpha}d} V_{d,l}(\Pi(n)). \tag{25}$$

Here, the values $w_\alpha(v_m(g^n/\mathfrak{p}))$ can be efficiently precomputed for $n = 1, \ldots, b^m - 1$ in $\mathcal{O}(\alpha m b^m)$ operations, as shown in [6, Theorem 2]. Next, the convolution above can be evaluated for all $z_d = 0, \ldots, b^m - 1$ with FFT in $\mathcal{O}(mb^m)$ operations. We then choose q_d^*, i.e. z_d^* that realizes the minimum. Next we compute $U_{d,l}(n), V_{d,l}(n)$ $\forall l = 1, \ldots, \bar{\alpha}d, \forall n = 0, \ldots, b^m - 1$ in $\mathcal{O}(\bar{\alpha}^2 d b^m)$ operations. Iterating over $d = 1, \ldots, s$, the computational cost for the CBC algorithm is then $\mathcal{O}(\bar{\alpha}^2 s^2 b^m + (s + \alpha) m b^m)$. Moreover, we can overwrite the quantities $U_{d,l}(n), V_{d,l}(n)$ as d increases; therefore, we require $\mathcal{O}(\bar{\alpha} s b^m)$ memory. The vector $w_\alpha(v_m(g^n/\mathfrak{p}))$ can be stored with $\mathcal{O}(b^m)$ memory. The cases of POD and SPOD weights are both covered, with $\bar{\alpha} = 1$ and $\bar{\alpha} = \alpha$, respectively.

To apply Richardson extrapolation, we need to construct polynomial lattice rules with α consecutive sizes of nodes $b^{m-\alpha+1}, \ldots, b^m$, so that we construct in total $N = b^{m-\alpha+1} + \ldots + b^m$ QMC points. Since

$$\sum_{\tau=1}^{\alpha}(s + \alpha)(m - \tau + 1)b^{m-\tau+1} \le (s + \alpha)mN \le (s + \alpha)N \log_b N$$

we have proven that the total cost is

$$\mathcal{O}\big(\bar{\alpha}^2 s^2 N + (s + \alpha)N \log N\big) \text{ operations and } \mathcal{O}(\bar{\alpha} s N) \text{ memory.} \tag{26}$$

Remark 2.9. *The error bound does not apply for $\alpha = 1$, since we require $1/\alpha < \lambda \le 1$. Moreover, for applications to parametric PDEs with global support of the fluctuations we usually have $\alpha = \bar{\alpha}$, to bound the derivatives of the solution up to order α. Therefore, Richardson extrapolation is not relevant for such PDE applications in the case of POD weights. In the following sections we will always work with $\alpha = \bar{\alpha}$.*

Remark 2.10. *The result* (26) *compares favorably to interlaced polynomial lattice (IPL) rules: IPL rules require $\mathcal{O}\big(\alpha^2 s^2 N + \alpha s N \log N\big)$ operations for SPOD weights (see, e.g., [10, 18, 19]).*

3 Linear affine-parametric PDEs

The error analysis of the extrapolated lattice rules for QMC integration of the previous section is now applied to forward UQ for a model linear, elliptic countably parametric PDE. Specifically, we consider the following model parametric elliptic PDE on a bounded physical domain $\mathrm{D} \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$

$$\begin{cases} -\mathrm{div}\,(a(x, \boldsymbol{y})\nabla u(x, \boldsymbol{y})) = f(x) & x \in \mathrm{D} \\ u(x, \boldsymbol{y}) = 0 & x \in \partial\mathrm{D} \end{cases} \tag{27}$$

where $\boldsymbol{y} \in U := \left[-\frac{1}{2}, \frac{1}{2}\right]^{\mathbb{N}}$ denotes the sequence of parameters of the uncertain diffusion coefficient.

We describe the uncertainty through an affine-parametric structure of the coefficients

$$a(x, \boldsymbol{y}) = \bar{a}(x) + \sum_{j \geq 1} y_j \psi_j(x) \qquad \text{for } \boldsymbol{y} \in U \tag{28}$$

for a sequence $(\psi_j)_{j\geq 1} \subset L^\infty(\mathrm{D})$. Examples of such sequences include Karhunen-Loeve expansions [36], which are generally described by globally supported functions, as well as locally supported bases as, for example, splines or wavelets. The former case will lead to the choice of SPOD weights, which is subject of Section 3.1. The latter will be analyzed in Section 3.2. Following the arguments in [17], in this case the QMC theory based on (5) for product weights will be sufficient.

For $f \in L^2(\mathrm{D})$ and for $a(\cdot, \boldsymbol{y}) \in L^\infty(\mathrm{D})$ for all $\boldsymbol{y} \in U$, we consider its variational formulation

$$\int_{\mathrm{D}} a(x, \boldsymbol{y}) \nabla u(x, \boldsymbol{y}) \cdot \nabla v(x) \mathrm{d}x = \int_{\mathrm{D}} f(x) v(x) \mathrm{d}x \quad \forall v \in H_0^1(\mathrm{D}). \tag{29}$$

To state its variational form, we introduce the space $V := H_0^1(\mathrm{D})$, with dual $V' := H^{-1}(\mathrm{D})$ with respect to the pivot space $L^2(\mathrm{D})$.

For any $f \in V'$, we can write the above equation in the generic form

$$\mathfrak{a}_{\boldsymbol{y}}(u(\cdot, \boldsymbol{y}), v) = \langle f, w \rangle_V \quad \forall v \in V, \tag{30}$$

where brackets denote the duality pairing in V and

$$\mathfrak{a}_{\boldsymbol{y}}(v, w) := \int_{\mathrm{D}} a(x, \boldsymbol{y}) \nabla v(x) \cdot \nabla w(x) \mathrm{d}x \tag{31}$$

is a bilinear form in V.

3.1 Globally supported fluctuations

In order to verify well-posedness of (30), we impose a set of additional assumptions. First, we assume in (32) *nominal invertibility*, i.e. there are constants $\bar{a}_{\min} \leq \bar{a}_{\max}$ such that

$$0 < \bar{a}_{\min} \leq \bar{a}(x) \leq \bar{a}_{\max} \qquad \text{for a.a. } x \in \mathrm{D}\,. \tag{32}$$

The smallness of the fluctuation in (32) with respect to the nominal operator is given by

$$\|\boldsymbol{\beta}\|_{\ell^1(\mathbb{N})} < 2 \quad \text{ for } \quad \beta_j := \frac{\|\psi_j\|_{L^\infty(\mathrm{D})}}{\bar{a}_{\min}}, \forall j \in \mathbb{N}\,. \tag{33}$$

With these assumptions we have that $a(x, \boldsymbol{y}) \geq a_{\min} > 0$ a.e. $x \in \mathrm{D}$ and for all $\boldsymbol{y} \in U$ where $a_{\min} := \bar{a}_{\min}(1 - \|\boldsymbol{\beta}\|_{\ell^1(\mathbb{N})}/2)$. A direct application of the Lax-Milgram lemma verifies that these conditions are sufficient for existence and uniqueness of solutions $u(\cdot, \boldsymbol{y}) \in V$ for all $\boldsymbol{y} \in U$. Furthermore, we have the uniform a-priori estimate

$$\sup_{\boldsymbol{y} \in U} \|u(\cdot, \boldsymbol{y})\|_V \leq \frac{\|f\|_{V'}}{a_{\min}}.$$

Moreover, we choose an ordering of the functions $\psi_j, j \in \mathbb{N}$, such that the sequence $\boldsymbol{\beta}$ is monotonically non-increasing and we assume that

$$\boldsymbol{\beta} \in \ell^p(\mathbb{N}), \quad p \in (0, 1). \tag{34}$$

The following theorem was obtained in [7, Theorem 4.3]. Such bounds on the derivatives with respect to the parameters allow to control the norm (5) of $F(\boldsymbol{y}) = G(u(\boldsymbol{y}))$.

Theorem 3.1. *Under the assumptions* (32), (33)*, for all* $f \in V'$ *the partial derivatives of the parametric solution* u *of* (27), (28) *satisfy*

$$\sup_{\boldsymbol{y}\in U} \left\|(\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u)(\boldsymbol{y})\right\|_V \le |\boldsymbol{\nu}|!\boldsymbol{\beta}^{\boldsymbol{\nu}} \frac{\|f\|_{V'}}{a_{\min}}.$$

Corollary 3.2. *Let* $\alpha, s \in \mathbb{N}$ *and* $f, G \in V'$*. Assume that* $\boldsymbol{\beta}$ *is a non-increasing sequence satisfying* (32), (33) *and define the positive SPOD weights* $\boldsymbol{\gamma} = (\gamma_{\mathfrak{u}})_{\mathfrak{u}\subset\mathbb{N},|\mathfrak{u}|<\infty}$ *by (cf. [10])*

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

where $\delta(\nu_j, \alpha) = 1$ *if* $\nu_j = \alpha$ *and* 0 *otherwise. Then there exist a positive constant* C *only dependent on the data* f, G *and* a *such that the solution* $u \in V$ *of* (27) *satisfies*

$$\|G(u)\|_{s,\alpha,\boldsymbol{\gamma},1,\infty} \le C. \tag{35}$$

Proof. Theorem 3.1 implies the bound

$$\begin{aligned}\|G(u)\|_{s,\alpha,\boldsymbol{\gamma},1,\infty} &\le \|G\|_{V'} \sup_{\mathfrak{u}\subseteq\{1:s\}} \gamma_{\mathfrak{u}}^{-1} \sum_{\boldsymbol{\nu}\in\{1:\alpha\}^{|\mathfrak{u}|}} 2^{|j:\ \nu_j=\alpha|} \sup_{\boldsymbol{y}\in U} \left\|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} u(\cdot,\boldsymbol{y})\right\|_V \\ &\le \frac{\|G\|_{V'}\|f\|_{V'}}{a_{\min}} \sup_{\mathfrak{u}\subseteq\{1:s\}} \gamma_{\mathfrak{u}}^{-1} \sum_{\boldsymbol{\nu}\in\{1:\alpha\}^{|\mathfrak{u}|}} |\boldsymbol{\nu}|! \prod_{j\in\mathfrak{u}} 2^{\delta(\nu_j,\alpha)} \beta_j^{\nu_j},\end{aligned}$$

which leads to the choice of SPOD weights for $\gamma_{\mathfrak{u}}$. Thus, $\|G(u)\|_{s,\alpha,\boldsymbol{\gamma},1,\infty}$ is bounded independently of s by $C := \frac{\|G\|_{V'}\|f\|_{V'}}{a_{\min}}$. □

Proposition 3.3. *Let* $f, G \in V'$ *and* $s \in \mathbb{N}$ *be given. Assume that* $\boldsymbol{\beta}$ *is a non-increasing sequence satisfying* (32), (33), (34) *with* $p \in (0, 1/2)$*. Then, there exist extrapolated polynomial lattice rules constructed with a CBC algorithm and with* $\alpha = \left\lfloor \frac{1}{p} \right\rfloor + 1$ *such that*

$$\left|(I_s - Q_{N,s}^{(\alpha)})(G(u))\right| \le CN^{-\frac{1}{p}},$$

where the constant C *is independent of* s*. If instead* $p \in [1/2, 1)$*, the same holds if additionally* $N \ge s$*.*

Proof. By Corollary 3.2, $\|G(u)\|_{s,\alpha,\boldsymbol{\gamma},1,\infty}$ is bounded independently of s for the SPOD weights

$$\gamma_{\mathfrak{u}} := \sum_{\boldsymbol{\nu}\in\{1:\alpha\}^{|\mathfrak{u}|}} |\boldsymbol{\nu}|! \prod_{j\in\mathfrak{u}} 2^{\delta(\nu_j,\alpha)} \beta_j^{\nu_j}. \tag{36}$$

We can then apply Theorem 2.5 or Remark 2.8, so that we can construct a QMC rule such that $B_{\boldsymbol{\gamma}}(p, \boldsymbol{q}) \le C(p)N^{-1/p}$ with $C(p)$ independent of s. Moreover, the residual term $H_{s,\boldsymbol{\gamma},1,\infty}$ in (9) is also bounded independently of s, since $p < 1$. Therefore, the claim follows by inserting these estimates into equation (13). □

3.2 Locally supported fluctuations

In this section, motivated by the results in [17, 16], we replace the assumptions (33) and (34) on the diffusion coefficient in (27), (28) by the following bound that takes into account possible local support of the $(\psi_j)_{j\geq 1}$

$$\left\| \frac{\sum_{j\geq 1} |\psi_j|/\bar{\beta}_j}{2\bar{a}} \right\|_{L^\infty(\mathrm{D})} \leq \kappa < 1. \tag{37}$$

Here, we assume that $(\bar{\beta}_j)_j$ is a non-increasing sequence in $\ell^p(\mathbb{N})$ for some $p \in (0,1)$, with $\bar{\beta}_j \leq 1$. Again we assume the invertibility of the nominal operator in (32). Under these assumptions it was proved in [17] that the problem is well-posed for every $\boldsymbol{y} \in U$ and that, for any $\eta \in (\kappa, 1)$ there holds

$$|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} G(u(\cdot,\boldsymbol{y}))| \leq C\|f\|_{V'}\|G\|_{V'} \left[\prod_{j\in\mathfrak{u}} \left(\frac{2\bar{\beta}_j}{1-\eta} \right)^{\nu_j} \nu_j! \right].$$

This bound on the derivatives is of product form. Defining $F(\boldsymbol{y}) := G(u(\cdot,\boldsymbol{y}))$, for $r=\infty$ and any $q \in [1,\infty]$, there holds

$$\begin{aligned}
\|F\|_{s,\alpha,\boldsymbol{\gamma},q,\infty} &\leq \sup_{\mathfrak{u}\subseteq\{1:s\}} \gamma_{\mathfrak{u}}^{-1} \left(\sum_{\boldsymbol{\nu}\in\{1:\alpha\}^{|\mathfrak{u}|}} \int_{[-\frac{1}{2},\frac{1}{2}]^s} 2^{|\{j\in\mathfrak{u}\,:\nu_j=\alpha\}|} \left|\partial_{\boldsymbol{y}}^{\boldsymbol{\nu}} F(\boldsymbol{y})\right|^q \,\mathrm{d}\boldsymbol{y} \right)^{1/q} \\
&\leq C\|f\|_{V'}\|G\|_{V'} \sup_{\mathfrak{u}\subseteq\{1:s\}} \gamma_{\mathfrak{u}}^{-1} \left(\sum_{\boldsymbol{\nu}\in\{1:\alpha\}^{|\mathfrak{u}|}} \prod_{j\in\mathfrak{u}} 2^{\delta(\nu_j,\alpha)} \left[\left(\frac{2\bar{\beta}_j}{1-\eta} \right)^{\nu_j} \nu_j! \right]^q \right)^{1/q} \\
&= C\|f\|_{V'}\|G\|_{V'} \sup_{\mathfrak{u}\subseteq\{1:s\}} \gamma_{\mathfrak{u}}^{-1} \prod_{j\in\mathfrak{u}} \left(\sum_{\nu=1}^{\alpha} 2^{\delta(\nu,\alpha)} \left[\left(\frac{2\bar{\beta}_j}{1-\eta} \right)^{\nu} \nu! \right]^q \right)^{1/q}.
\end{aligned}$$

We now consider two methods for obtaining upper bounds on these expressions which are adapted to particular integrand classes.

Method 1: (conservative upper bound) Set $q=1$, i.e. $\|F\|_{s,\alpha,\boldsymbol{\gamma},q,\infty} \leq \|F\|_{s,\alpha,\boldsymbol{\gamma},1,\infty}$ and choose product weights

$$\gamma_{\mathfrak{u}} := \prod_{j\in\mathfrak{u}} \sum_{\nu=1}^{\alpha} 2^{\delta(\nu,\alpha)} \left(\frac{2\bar{\beta}_j}{1-\eta} \right)^{\nu} \nu!.$$

Method 2: (sharper bound) the inequality above is valid for all $q \in [1,\infty]$; therefore, we let $q=\infty$ to minimize the weights. Then

$$\prod_{j\in\mathfrak{u}} \left(\sum_{\nu=1}^{\alpha} 2^{\delta(\nu,\alpha)} \left[\left(\frac{2\bar{\beta}_j}{1-\eta} \right)^{\nu} \nu! \right]^q \right)^{1/q} = \prod_{j\in\mathfrak{u}} \left(\left[\left(\frac{2\bar{\beta}_j}{1-\eta} \right)^{\alpha} \alpha! \right]^q + \sum_{\nu=1}^{\alpha} \left[\left(\frac{2\bar{\beta}_j}{1-\eta} \right)^{\nu} \nu! \right]^q \right)^{1/q}$$

that leads to the definition

$$\gamma_{\mathfrak{u}} := \prod_{j\in\mathfrak{u}} \max_{\nu=1,\ldots,\alpha} \left[\left(\frac{2\bar{\beta}_j}{1-\eta} \right)^{\nu} \nu! \right].$$

Note that method 2 results in a better constant but the convergence rate of the QMC approximation does not improve. Moreover, both methods above result in product weights, so that we can apply the results from [9]: $J_{s,\lambda,\boldsymbol{\gamma}}$ and $H_{s,\boldsymbol{\gamma},\infty,\infty}$ are bounded independently of s if and only if

$$\sum_{|\mathfrak{u}|<\infty} \gamma_{\mathfrak{u}}^{\lambda} C_{\alpha}^{\lambda|\mathfrak{u}|} E_{\alpha,\lambda}^{|\mathfrak{u}|} < \infty \qquad \text{and} \qquad \sum_{|\mathfrak{u}|<\infty} \gamma_{\mathfrak{u}} (\alpha+1)^{|\mathfrak{u}|} D_{\alpha}^{|\mathfrak{u}|} < \infty.$$

Since $\lambda \leq 1$, the first condition is stronger than the second. Let K be a generic constant, then we verify both as follows:

$$\begin{aligned}\sum_{|\mathfrak{u}|<\infty} \gamma_{\mathfrak{u}}^{\lambda} K^{|\mathfrak{u}|} &= \sum_{|\mathfrak{u}|<\infty} \prod_{j\in\mathfrak{u}} K \max_{\nu=1,\dots,\alpha} \left[\left(\frac{2\bar{\beta}_j}{1-\eta}\right)^{\nu} \nu!\right]^{\lambda} \\ &\leq \exp\left(K \sum_{j\geq 1} \max_{\nu=1,\dots,\alpha} \left[\left(\frac{2\bar{\beta}_j}{1-\eta}\right)^{\lambda\nu} (\nu!)^{\lambda}\right]\right) \\ &\leq \exp\left(K \sum_{\nu=1}^{\alpha} \left(\frac{2}{1-\eta}\right)^{\lambda\nu} (\nu!)^{\lambda} \sum_{j\geq 1} \bar{\beta}_j^{\lambda\nu}\right).\end{aligned}$$

The value $\nu = 1$ gives the asymptotically largest summand; hence the decay rate of the QMC error of $\mathcal{O}(N^{-1/\lambda})$ follows provided that $(\bar{\beta}_j)_j \in \ell^{\lambda}(\mathbb{N})$, that imposes $\lambda \geq p$. Since we also have the constraint $\lambda > 1/\alpha$ we get the rate $\mathcal{O}(N^{-1/p})$, with constant independent of s, using extrapolation of order $\alpha = 1 + \left\lfloor \frac{1}{p} \right\rfloor$. Observe that, conversely to Proposition 3.3, we do not require $p < \frac{1}{2}$ in this case.

3.3 Galerkin discretization

We consider a bounded polygon $\mathrm{D} \subset \mathbb{R}^d$, $d = 2$ with corners $\xi_1, \dots, \xi_J$ and we fix $\omega \in \mathbb{R}$ satisfying $\omega < \frac{\pi}{\max_i \theta_i}$ where θ_i is the interior angle of D corresponding to ξ_i. In addition, given the weight function

$$r_{\mathrm{D}}(x) := \prod_{j=1}^{J} |x - \xi_j|$$

and $k \in \mathbb{N}_0$, we can define the Kondrat'ev spaces $\mathcal{K}_{\omega}^{k}(\mathrm{D}) \subset H_{loc}^{k}(\mathrm{D})$ via the norm

$$\|v\|_{\mathcal{K}_{\omega}^{k}(\mathrm{D})} := \sum_{|\alpha|=0}^{k} \left\| |\partial^{\alpha} v| r_{\mathrm{D}}^{|\alpha|-\omega} \right\|_{L^2(\mathrm{D})} \tag{38}$$

and the space $\mathcal{W}^{k,\infty}(\mathrm{D})$ with the norm

$$\|v\|_{\mathcal{W}^{k,\infty}(\mathrm{D})} := \sum_{|\alpha|=0}^{k} \left\| |\partial^{\alpha} v| r_{\mathrm{D}}^{|\alpha|} \right\|_{L^{\infty}(\mathrm{D})},$$

where we used the multiindex notation for derivatives with respect to x. We assume that there are $t, t' \in \mathbb{N}$ such that

$$f \in \mathcal{K}_{\omega-1}^{t-1}(\mathrm{D}),\ G \in \mathcal{K}_{\omega-1}^{t'-1}(\mathrm{D}),\ \sup_{\boldsymbol{y}\in U} \|a(\cdot,\boldsymbol{y})\|_{\mathcal{W}^{t,\infty}(\mathrm{D})} < \infty.$$

The regularity theory in [2, Theorem 4.4] implies that the solution of (27) satisfies

$$\sup_{\boldsymbol{y}\in U} \|u(\cdot,\boldsymbol{y})\|_{\mathcal{K}_{\omega+1}^{t+1}(\mathrm{D})} \leq C \|f\|_{\mathcal{K}_{\omega-1}^{t-1}(\mathrm{D})}, \tag{39}$$

as the full regularity shift of the elliptic operator holds in the weighted spaces (38) uniformly in the parameter $\boldsymbol{y} \in U$. In what follows, we will write $V^t_\pm := \mathcal{K}^{t\pm 1}_{\omega\pm 1}(\mathrm{D})$ and $V^t := \mathcal{K}^t_\omega(\mathrm{D})$. We define a sequence of nested, conforming, finite-dimensional FEM spaces $\{V_M\}_M$, $\dim(V_M) = M$, $V_M \subset V$. Then we consider the *Galerkin discretization of the parametric, elliptic PDE* (27): find

$$u_M(\cdot, \boldsymbol{y}) \in V_M \text{ such that } \mathfrak{a}_{\boldsymbol{y}}(u_M(\cdot, \boldsymbol{y}), v) = \langle f, v \rangle_V\,, \quad \forall v \in V_M. \tag{40}$$

This problem is also well posed due to the Lax-Milgram lemma and conformity of the FEM spaces, and we have the uniform stability estimate

$$\sup_{M\in\mathbb{N}} \sup_{\boldsymbol{y}\in U} \|u_M(\cdot, \boldsymbol{y})\|_V \leq \frac{\|f\|_{V'}}{a_{\min}}.$$

Furthermore, there exists a constant $C > 0$, independent of M and $\boldsymbol{y}$, such that there holds quasi-optimality

$$\|u(\cdot, \boldsymbol{y}) - u_M(\cdot, \boldsymbol{y})\|_V \leq C \inf_{v_M\in V_M} \|u(\cdot, \boldsymbol{y}) - v_M\|_V\,. \tag{41}$$

It was shown in [3] that suitably graded meshes can give an explicit construction of the spaces V_M, satisfying the approximation property,

$$\inf_{v_M\in V_M} \|v - v_M\|_V \leq CM^{-t/d} \|v\|_{V^t_+}\,, \tag{42}$$

for $d = 2$ and a constant C independent of v. This is done with piecewise polynomials of degree t in each element.

For the case of a polyhedron $\mathrm{D} \subset \mathbb{R}^d$, in space dimension $d = 3$ with plane faces, the definition of the solution space V^t_+ is more involved. It considers anisotropic regularity [4, 5, 1, 23]. Therefore, the approximation property (42) for $t \geq 2$ holds, provided that the data f belongs to the space $V^t_- = H^{t-1}(\mathrm{D})$ [5, Theorem 8.1], while the case $t = 1$ and less regular f was covered in [1, Theorem 4.6]. In both cases, the regular, simplicial triangulations of D which enter the construction of the Lagrangian FE spaces V_M must be graded towards the corners and, in space dimension $d = 3$, also *anisotropically* towards the edges of the domain (see [1] and the references there).

Combining the estimates (39), (41) and (42) we obtain the following bound for the Galerkin error.

Proposition 3.4. *Let $f \in V^t_-$ and u be the exact solution of* (27) *for $d = 2, 3$. Then, there exists a suitably graded mesh such that the corresponding Galerkin solution u_M on the space V_M satisfies*

$$\sup_{\boldsymbol{y}\in U} \|u(\cdot, \boldsymbol{y}) - u_M(\cdot, \boldsymbol{y})\|_V \leq CM^{-t/d} \|f\|_{V^t_-} \tag{43}$$

for all M. Moreover, for a $G \in V^{t'}_-$, an Aubin-Nitsche duality argument implies that there exists a constant $C > 0$ independent of M such that

$$\sup_{\boldsymbol{y}\in U} |G(u(\cdot, \boldsymbol{y})) - G(u_M(\cdot, \boldsymbol{y}))| \leq CM^{-(t+t')/d} \|f\|_{V^t_-} \|G\|_{V^{t'}_-} \tag{44}$$

in the family of FE spaces V_M of piecewise polynomials of degree $\max(t, t')$.

Remark 3.5. *Corollary 3.2 also holds for Galerkin solutions u_M with the same choice of weights. This follows from the fact that the proof of 3.1 only uses the variational formulation of the PDE. Therefore, restricting the test space to the finite dimensional space V_M leads to the same upper bound on the derivatives $\partial^{\boldsymbol{\nu}}_{\boldsymbol{y}} u_M(\cdot, \boldsymbol{y})$.*

3.4 Dimension truncation

Since the parameter space is infinite dimensional, the first step in approximating integrals of the goal functional is a truncation of the expansion for the fluctuation. In our setting, a complete theory is already available from [30, 31, 15]. Given $\boldsymbol{y} = (y_1, y_2, \ldots) \in U$, define $\boldsymbol{y}_s := (y_1, \ldots, y_s)$ and $u_s(\cdot, \boldsymbol{y}) := u(\cdot, (\boldsymbol{y}_s, 0, 0, \ldots))$ be the solution of (27) with truncated expansion of the uncertain coefficient. We recall here the main results of [15, Proposition 3, Theorem 1].

Theorem 3.6. *Under assumptions* (32) *and* (33)*, there exists a constant C such that for every $f \in V'$, every $\boldsymbol{y} \in U$ and $s \in \mathbb{N}$, we have*

$$\|u(\cdot, \boldsymbol{y}) - u_s(\cdot, \boldsymbol{y})\|_V \le C \frac{\|f\|_{V'}}{a_{\min}} s^{-(1/p-1)}.$$

Moreover, there exists another constant $\tilde{C}$ such that, if also $G \in V'$,

$$|I_\infty(G(u)) - I_s(G(u))| \le \tilde{C} \frac{\|f\|_{V'} \|G\|_{V'}}{a_{\min}} s^{-(2/p-1)}.$$

Remark 3.7. *For any fixed $s \in \mathbb{N}$, since $\left[-\frac{1}{2}, \frac{1}{2}\right]^s \times \{0\}^{\mathbb{N}\setminus\{1:s\}} \subset U$, then equation* (44) *is also valid for the solution $u_s(\cdot, \boldsymbol{y})$ of the truncated problem.*

3.5 Combined QMCFEM error bound

In the following theorem we summarize the preceding bounds on the QMC quadrature error, Galerkin error and dimension truncation error.

Theorem 3.8. *Let $s \in \mathbb{N}$. For $0 < t, t' \le \bar{t}$, let $a(\cdot, \boldsymbol{y}) \in \mathcal{W}^{\bar{t},\infty}(\mathrm{D})$, $f \in V^t_-$, $G \in V^{t'}_-$ and assume that* (32) *holds. Let $\boldsymbol{\beta}$ be a non-increasing sequence satisfying* (33) *and* (34) *for some $p \in (0, 1/2)$. Then, there exists an extrapolated polynomial lattice rule of order $\alpha = 1 + \left\lfloor \frac{1}{p} \right\rfloor$ such that*

$$\left| I_\infty(G(u)) - Q^{(\alpha)}_{N,s}(G(u_M)) \right| \le C \|f\|_{V^t_-} \|G\|_{V^{t'}_-} (M^{-(t+t')/d} + N^{-1/p} + s^{-(2/p-1)})$$

for a constant $C > 0$ independent of s, N, M and of the data f, G. If instead $p \in [1/2, 1)$, the same holds if additionally $N \ge s$.

Proof. We separate the sources of error so that

$$\begin{aligned} \left| I_\infty(G(u)) - Q^{(\alpha)}_{N,s}(G(u_M)) \right| \le & |I_\infty(G(u)) - I_s(G(u))| + \left| I_s(G(u)) - Q^{(\alpha)}_{N,s}(G(u)) \right| \\ & + \left| Q^{(\alpha)}_{N,s}(G(u)) - Q^{(\alpha)}_{N,s}(G(u_M)) \right|. \end{aligned}$$

Since (44) holds, we can bound the Galerkin error as follows

$$\left| Q^{(\alpha)}_{N,s}(G(u - u_M)) \right| \le C M^{-(t+t')/d} \|f\|_{V^t_-} \|G\|_{V^{t'}_-}.$$

On the other hand, since $V^t \subset V'$ with continuous embedding, we bound the truncation error and the QMC error using Theorem 3.6 and Proposition 3.3 and the claim follows. □

Coupling the number of degrees of freedom in the FEM space and the number of QMC samples should be done according to

$$N^{-1/p} \sim M^{-(t+t')/d} \sim s^{-(2/p-1)} = \mathcal{O}(\varepsilon)$$

where ε is a prescribed error tolerance. Note that, for given $p \in [1/2, 1)$, the requirement $N \geq s$ is thus readily satisfied. Therefore, assuming that the QMC points have been precomputed and that the solution of the linear FE system can be done in $\mathcal{O}(M)$ operations using sparse matrices, the computational work of the single level QMCFEM algorithm is

$$\text{work } = \mathcal{O}\Big(\text{work}_a + \varepsilon^{-p}\varepsilon^{-d/(t+t')}\Big)\Big) .$$

Here work_a is the cost for the assembly of all the linear FEM systems. In particular, the affine-parametric structure (28) implies that

$$A(\boldsymbol{y}_n) = \bar{A} + \sum_{j=1}^{s} y_{n,j}\Psi_j \qquad \forall n \in 0, \ldots, b^m - 1,$$

where $\bar{A}$ and Ψ_j are the stiffness matrices corresponding to $\bar{a}$ and ψ_j respectively. The Ψ_j are usually sparse and have $\mathcal{O}(M)$ non-zero entries. Moreover, since they have the same sparsity pattern as $\bar{A}$, dependent on the FEM basis, but not on n, we get

$$\text{work } = \mathcal{O}\Big(s\varepsilon^{-p-d/(t+t')}\Big) = \mathcal{O}\Big(\varepsilon^{-p-p/(2-p)-d/(t+t')}\Big) . \tag{45}$$

On the other hand, the main motivation to introduce Richardson extrapolation in [9] was the possibility to extend the fast matrix-vector multiplication in [11] to higher-order QMC quadrature. This is due to the fact that extrapolated lattice rules are linear combinations of first order polynomial lattice rules, see Section 2.2. As a consequence, the fast QMC matrix vector product can be used to reduce the complexity of the computation of the parametric stiffness matrices $A(\boldsymbol{y})$ corresponding to the PDE coefficient in (28). Using the standard approach, the overall computational cost is $\mathcal{O}(Mb^m s)$; however, the computation can be carried out in $\mathcal{O}(Mmb^m)$ operations plus at most $\mathcal{O}(M(s-1))$ additions with FFT (see [11, Section 3.2] for more details). On the other hand, this requires to store all the stiffness matrices in $\mathcal{O}(Mb^m)$ memory.

If we repeat the same steps for every $m' = m - \alpha + 1, \ldots, m$ and then we combine the partial results $Q^{(1)}_{b^{m'},s}(F)$, it is immediate to verify that the overall computational cost of the fast matrix-vector multiplication for extrapolated lattice rules is $\mathcal{O}(MN \log N)$ plus at most $\mathcal{O}(M(s-1))$ additions – that can be avoided when the generating vector has no repeated components – and $\mathcal{O}(MN)$ memory, with $N = b^m + \ldots + b^{m-\alpha+1}$. This is advantageous for $N \ll 2^s$, which holds in our setting since $N \sim s^{2-p}$. As a result, we obtain the following work vs error rate that improves (45)

$$\text{work } = \mathcal{O}\Big(\log(\varepsilon^{-1})\varepsilon^{-p-d/(t+t')}\Big) . \tag{46}$$

4 A-posteriori QMC error estimator

It is often required to control the (relative) error of a numerical approximation, aiming at an accuracy up to a predefined tolerance $tol > 0$. In the context of QMC integration using extrapolated polynomial lattice rules, we want to verify that

$$\frac{\left|I_s(F) - Q^{(\alpha)}_{N,s}(F)\right|}{|I_s(F)|} \leq tol$$

with reasonable computational effort. We show that it is possible to compute an estimate of the error that is asymptotically exact and we can use this quantity as a valid stopping criterion for the

QMC approximation. The key to numerical extrapolation of the QMC approximating sequence $Q^{(\alpha)}_{b^m,s}(F)$, for $m \in \mathbb{N}$, is the availability of the asymptotic Euler-MacLaurin expansion of the QMC rule. We therefore think of Richardson extrapolation of order α as an application of correction terms to the sequence $(Q^{(1)}_{b^m,s})_{m\in\mathbb{N}}$, based on previously computed quantities. We collect these corrections in the value $\Delta Q^{(\alpha)}_{b^m,s}$, defined by the relation

$$Q^{(\alpha)}_{b^m,s} = Q^{(1)}_{b^m,s} + \Delta Q^{(\alpha)}_{b^m,s}. \tag{47}$$

In particular, we interpret $\Delta Q^{(\alpha)}_{b^m,s}$ as an indicator of how far the originally computed QMC quadrature $Q^{(1)}_{b^m,s}$ lies from the exact integral, provided that we have $F \in \mathcal{W}_{s,\alpha,\boldsymbol{\gamma},q,\infty}$ so that the Euler-MacLaurin formula holds. This indicator is evaluated in the extrapolation algorithm with negligible overhead. Let us now proceed to the detailed derivation. We adopt the notation of Section 2.2.

Fix a natural prime number $b \geq 2$. Let furthermore $P_m := P(\mathfrak{p}, \boldsymbol{q})$ be a polynomial lattice point set with $\deg(\mathfrak{p}) = m$ and $\boldsymbol{q}$ constructed with the CBC algorithm of Section 2.4, and denote by $Q^{(1)}_{b^m,s}$ the corresponding QMC rule, obtained by shifting the points $\boldsymbol{y}_n = \boldsymbol{x}_n - \boldsymbol{\frac{1}{2}}$ to the hypercube $[-1/2, 1/2]^s$. In the following theorem, the term on the right hand side of (48) is (up to the remainder term $\mathcal{O}(\cdot)$) a computable expression for $\Delta Q^{(2)}_{b^m,s}$.

Theorem 4.1. *Let an integration dimension $s \in \mathbb{N}$ be given, and also $\alpha \geq 2$, $1 \leq q \leq \infty$, $\boldsymbol{\gamma} = (\gamma_{\mathfrak{u}})_{\mathfrak{u}}$ be a set of positive product weights $\gamma_{\mathfrak{u}} = \prod_{j\in\mathfrak{u}} \gamma_j$, with $(\gamma_j)_{j\in\mathbb{N}} \in \ell^p(\mathbb{N})$ for all $p > 1/2$, and let $F \in \mathcal{W}_{s,\alpha,\boldsymbol{\gamma},q,\infty}$. Then, for all fixed $n \in \mathbb{N}$ and for all $\varepsilon > 0$*

$$I_s(F) - Q^{(1)}_{b^m,s}(F) = \frac{1}{b^n - 1}\big(Q^{(1)}_{b^m,s}(F) - Q^{(1)}_{b^{m-n},s}(F)\big) + \mathcal{O}\big(b^{n-2m+\varepsilon}\big) \ \text{ as } m \to \infty, \tag{48}$$

with constant in the $\mathcal{O}(\cdot)$ notation independent of s. Furthermore, for weights in SPOD form (11) with $c_1 \in \mathbb{N}_0, c_2 \in \mathbb{N}$, the same estimate holds if we assume $(\beta_j)_{j\in\mathbb{N}} \in \ell^p(\mathbb{N})$ for some $0 < p < 1/(1+c_2)$.

Proof. For any $F \in \mathcal{W}_{s,\alpha,\boldsymbol{\gamma},q,\infty}$ the Euler-MacLaurin formula holds for regular s-dimensional grids [9, Equation 3.1] which gives the following asymptotic expansion for the QMC integral

$$Q^{(1)}_{b^m,s}(F) = I_s(F) + \sum_{\substack{\boldsymbol{k}\in P_m^\perp\setminus\{\boldsymbol{0}\}\\ \exists j:\ b^m \nmid k_j}} \hat{F}(\boldsymbol{k}) + \sum_{\tau=1}^{\alpha-1} \frac{\sigma_\tau(F)}{b^{\tau m}} + R_{s,\alpha,b^m}, \tag{49}$$

where the coefficients $\sigma_\tau(F)$ are defined in [9, Theorem 3.4]. Here, $P_m^\perp$ is the *dual lattice* of P_m and $R_{s,\alpha,b^m} = \mathcal{O}(b^{-m\alpha})$. Moreover, in [9, Theorem 3.6] it was shown that, for product weights, a suitable CBC constructed generating vector $\boldsymbol{q}$ satisfies for all $\lambda > 1/\alpha$

$$\sum_{\substack{\boldsymbol{k}\in P_m^\perp\setminus\{\boldsymbol{0}\}\\ \exists j:\ b^m \nmid k_j}} |\hat{F}(\boldsymbol{k})| \leq \frac{1}{(b^m-1)^{1/\lambda}} \|F\|_{s,\alpha,\boldsymbol{\gamma},q,\infty} \left[\prod_{j=1}^{s} \left(1 + \gamma_j^\lambda C_\alpha^\lambda E_{\alpha,\lambda}\right)\right]^{1/\lambda}.$$

If $(\gamma_j)_{j\in\mathbb{N}} \in \ell^p(\mathbb{N})\ \forall p > 1/2$, then the right-hand side decays at least with rate $\mathcal{O}\big(b^{-m(2-\varepsilon)}\big)$ for all $\varepsilon > 0$, with constant independent of s. The same decay property is satisfied for SPOD weights, with the constraint $(\beta_j)_{j\in\mathbb{N}} \in \ell^p(\mathbb{N})$ for some $p < \frac{1}{1+c_2} \leq \frac{1}{2}$ (see Theorem 2.5). Therefore, if we collect the higher order terms

$$\delta_m := R_{s,\alpha,b^m} + \sum_{\substack{\boldsymbol{k}\in P_m^\perp\setminus\{\boldsymbol{0}\}\\ \exists j:\ b^m \nmid k_j}} \hat{F}(\boldsymbol{k}) + \sum_{\tau=2}^{\alpha-1} \frac{\sigma_\tau(F)}{b^{\tau m}}$$

we get that $\delta_m = \mathcal{O}\big(b^{-2m+\varepsilon}\big)$ for any $\varepsilon > 0$. Applying the Euler-MacLaurin formula (49) for distinct values $m, m' \in \mathbb{N}$ with $m = m' + n$, we have

$$Q^{(1)}_{b^m,s}(F) - Q^{(1)}_{b^{m'},s}(F) = \sigma_1(F)(b^{-m} - b^{-m'}) - \delta_{m'} + \delta_m,$$

which yields

$$\sigma_1(F) = \frac{b^m}{1-b^n}\left(Q^{(1)}_{b^m,s}(F) - Q^{(1)}_{b^{m-n},s}(F)\right) + \mathcal{O}\big(b^{n-m+\varepsilon}\big)\,.$$

Thus, defining $\tilde{\sigma}_1(F) := \frac{b^m}{1-b^n}\left(Q^{(1)}_{b^m,s}(F) - Q^{(1)}_{b^{m-n},s}(F)\right)$, we get from (49) that

$$\begin{aligned} I_s(F) - Q^{(1)}_{b^m,s}(F) &= -\frac{\tilde{\sigma}_1(F) + \mathcal{O}(b^{n-m+\varepsilon})}{b^m} - \delta_m \\ &= \frac{1}{b^n-1}(Q^{(1)}_{b^m,s}(F) - Q^{(1)}_{b^{m-n},s}(F)) + \mathcal{O}\big(b^{n-2m+\varepsilon}\big) \end{aligned}$$

and the proof is complete. □

In a similar fashion, we can approximate the relative error: if the exact integral is unknown, we can compare the absolute error with the approximate integral $Q^{(1)}_{b^m,s}(F)$; then we obtain, for the choice $n = 1$

$$\begin{aligned} \frac{\left|I_s(F) - Q^{(1)}_{b^m,s}(F)\right|}{|I_s(F)|} &= \frac{1}{b-1}\frac{\left|Q^{(1)}_{b^m,s}(F) - Q^{(1)}_{b^{m-1},s}(F)\right|}{\left|Q^{(1)}_{b^m,s}(F)\right| + \mathcal{O}(b^{-m})} + \mathcal{O}\big(b^{1-2m+\varepsilon}\big) \\ &\approx \frac{1}{b-1}\frac{\left|Q^{(1)}_{b^m,s}(F) - Q^{(1)}_{b^{m-1},s}(F)\right|}{\left|Q^{(1)}_{b^m,s}(F)\right|}, \end{aligned} \tag{50}$$

which is an a-posteriori QMC error estimator, *asymptotically exact* for $m \to \infty$. Furthermore, if $\left|Q^{(1)}_{b^m,s}(F)\right| \neq 0$ then the approximation above is accurate up to $\mathcal{O}\big(b^{-2m+\varepsilon}\big)$.

A straightforward application of Theorem 4.1 implies the following result.

Corollary 4.2. *Under the assumptions of Theorem 4.1, for the computable QMC quadrature error estimator* $\Delta Q^{(1)}_{b^m,s} = Q^{(1)}_{b^m,s}(F) - Q^{(1)}_{b^{m-1},s}(F)$, *we have* asymptotic exactness, *i.e.*

$$\frac{\left|\Delta Q^{(1)}_{b^m,s}(F)\right|}{\left|I_s(F) - Q^{(1)}_{b^m,s}(F)\right|} \to 1 \quad \text{as } m \to \infty. \tag{51}$$

Remark 4.3. *If we assume that* $\alpha \geq 3$ *and we employ* α *different values* $m, \dots, m-\alpha+1$*, we can analogously approximate the quantities* $\sigma_1, \dots, \sigma_{\alpha-1}$*, by solving a linear system with* α *variables up to higher order terms. For the numerical extrapolation process, however, the knowledge of the numerical values of* $\sigma_1, \dots, \sigma_{\alpha-1}$ *is not required.*

The above approach can be extended to $Q^{(\tau)}_{b^m,s}$ for $\tau = 1, 2, \dots, \alpha-1$, since $Q^{(\tau)}_{b^m,s}$ also satisfies an expansion of the form (49), i.e.

$$Q^{(\tau)}_{b^m,s}(F) = I_s(F) + \sum_{\kappa=\tau}^{\alpha-1} \frac{\sigma_{\tau,\kappa}(F)}{b^{\kappa m}} + \delta_{\tau,m},$$

where $\delta_{\tau,m}$ decays with order $b^{-m/p+\varepsilon}$ independently of the dimension. See (7) for the case $\tau = 2$. Hence we have

$$I_s(F) - Q^{(\tau)}_{b^m,s}(F) = \frac{1}{b^{\tau n} - 1}(Q^{(\tau)}_{b^m,s}(F) - Q^{(\tau)}_{b^{m-n},s}(F)) + \mathcal{O}\Big(b^{\tau n - m/p+\varepsilon}\Big) \text{ as } m \to \infty. \qquad (52)$$

In the same way, we can also extend Corollary 4.2 to obtain for any $\tau = 1, 2, \dots, \alpha - 1$ that

$$\frac{(b^\tau - 1)^{-1}\Big|Q^{(\tau)}_{b^m,s}(F) - Q^{(\tau)}_{b^{m-1},s}(F)\Big|}{\Big|I_s(F) - Q^{(\tau)}_{b^m,s}(F)\Big|} \to 1, \quad \text{as } m \to \infty.$$

Notice that in a general setting for integrands with smoothness α, this only works for $\tau = 1, 2, \dots, \alpha - 1$, since the sum $\sum_{\tau=1}^{\alpha-1} \sigma_\tau(F) b^{-\tau m}$ is restricted by the smoothness of F, i.e. $\sigma_\alpha(F)$ is in general not defined anymore. However, in the context of PDEs with random coefficients, it is known that the integrands are actually infinitely many times differentiable, the limiting factor in this context is the dependence on the dimension. Hence, the formula (49) also holds with the sum extended to $\sum_{\tau=1}^{\alpha} \sigma_\tau(F) b^{-\tau m}$. Hence in this special situation, (52) also holds for $\tau = \alpha$.

5 Numerical experiments

In this section we present some numerical examples to illustrate applications of extrapolated polynomial lattice rules. In all experiments which we report here, we employ polynomial lattice rules constructed with base $b = 2$.

5.1 Fast CBC construction

As a first example, we measure the computational cost of the fast CBC algorithm to compute the generating vector of the polynomial lattice rule, for the choice of SPOD weights considered here. We are in particular interested in the verification of the cost of the CBC construction in (26) using FFT with respect to the integration dimension s. The computations were performed with MATLAB 2018a on the ETH Euler cluster[1], enforcing single thread computations by activating the option `-singleCompThread`.

We observe the asymptotic rate of $\mathcal{O}(s^2)$ for every fixed m, which confirms our analysis.

5.2 Explicit parametric integrand

We perform numerical integration of the following explicit parametric integrand function over $U = [-1/2, 1/2]^s$ for a range of integration dimensions s

$$F(\boldsymbol{y}) := \left(1 + \sigma \sum_{j=1}^{s} j^{-\eta} y_j\right)^{-1} \qquad (53)$$

for a parameter $\eta > 1$ and a constant $\sigma > 0$, that can be chosen so that the function is bounded uniformly in $\boldsymbol{y}$. In particular, we have the constraint $\sigma < 2/\zeta(\eta)$ with ζ denoting the Riemann zeta function. From [19, Section 4.1.1] we know

$$\left|\partial^{\boldsymbol{\nu}}_{\boldsymbol{y}} F(\boldsymbol{y})\right| \le \sup_{\boldsymbol{y}\in U} |F(\boldsymbol{y})||\boldsymbol{\nu}|!\boldsymbol{\beta}^{\boldsymbol{\nu}} \quad \text{with} \quad \beta_j = \sup_{\boldsymbol{y}\in U} |F(\boldsymbol{y})|\sigma j^{-\eta}. \qquad (54)$$

[1] https://scicomp.ethz.ch/wiki/Euler

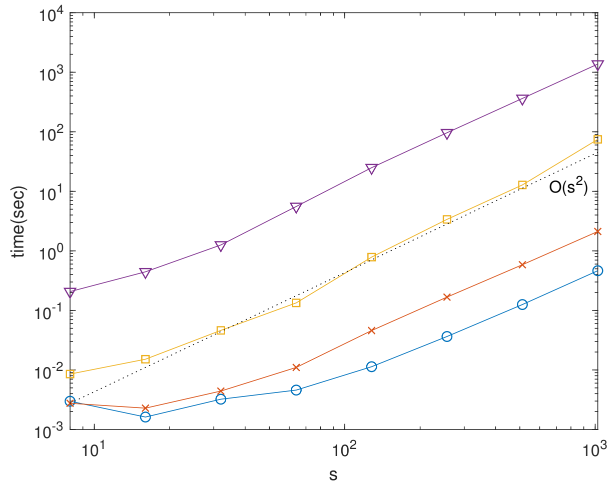

Figure 1:
Fast CBC construction: runtimes in seconds, versus number of dimensions, for $\alpha = 2$, SPOD weights with $c_1 = 0, c_2 = 1, c_3 = 1$ and $\beta_j = 0.2j^{-2}$ for the choices $m = 4, 8, 12, 16$ marked by circle, cross, square and down-triangle, respectively.

Therefore, the parametric integrand function F defined in (53) belongs to the weighted, unanchored Sobolev space $\mathcal{W}_{s,\alpha,\boldsymbol{\gamma},1,\infty}$ with SPOD weights

$$\gamma_{\mathfrak{u}} := \sum_{\boldsymbol{\nu} \in \{1:\alpha\}^{|\mathfrak{u}|}} |\boldsymbol{\nu}|! \prod_{j \in \mathfrak{u}} 2^{\delta(\nu_j, \alpha)} \beta_j^{\nu_j}. \tag{55}$$

We have $\boldsymbol{\beta} \in \ell^p(\mathbb{N})$ for $p > 1/\eta$. In the first experiment we work with $s = 16$. The reference value for the exact integral $I_s(F)$ was computed by adaptive Smolyak with tolerance $tol = 10^{-14}$ [19, Table 9.2], for the case $\eta = 2$, and by 2^{20} points of an interlaced polynomial lattice of order 2, for $\eta = 3$. In the same work it was also shown that it is advantageous to set the multiplicative constant in β_j below the value suggested by the theory. Therefore, we perform the fast CBC construction with the weights $\tilde{\gamma}_{\mathfrak{u}}$ obtained replacing β_j by the choice $\tilde{\beta}_j := 0.2j^{-\eta}$.

In Figure 2, we observe that the error decay reaches $\mathcal{O}(N^{-2.07})$ for $\sigma = 0.1$ and reduces slightly as σ gets larger, for the choice $\eta = 3$. In Figure 3 we set $\eta = 2$, so that we do not have a theoretical convergence of $\mathcal{O}(N^{-2+\varepsilon})$ for any $\varepsilon > 0$, but only for N sufficiently large (cp. Remark 2.8); the optimal rate is obtained for sufficiently small values of σ, showing robustness of extrapolation for $\alpha = 2$. Finally, convergence for varying dimension s is shown in Figure 4.

5.3 A-posteriori QMC Quadrature Error Estimation

We illustrate the efficiency of the a-posteriori computable QMC integration error estimator of Section 4 with an example for the same integrand (53), considering various choices of QMC weights, always with quadrature dimension $s = 16$.

We observe in Figure 5 that the ratio (51) converges to 1 for $\eta > 2$, which is the sufficient condition for the existence of the first term of the Euler-MacLaurin expansion. Furthermore, for $\eta = 1.9$, the estimator is still a good upper bound for the error, while as the summability decreases the estimator becomes less reliable. An analogous experiment employing the estimator $\Delta Q^{(\alpha)}_{b^m,s}$ for $\alpha = 3$ and the same summability is displayed in Figure 6. Here, we cannot expect to catch the coefficient $\sigma_2(F)$ in the estimator because $\eta \leq 3$; however, the ratio converges faster to 1 in the case $\eta = 2.5$.

Finally, the dimension independent convergence of the a-posteriori estimator to the QMC error is shown in Figure 7 up to dimension $s = 128$.

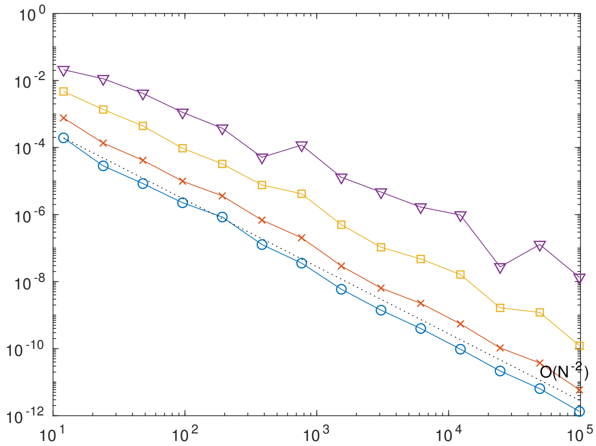

Figure 2:
Relative integration error versus number of QMC points. Extrapolated lattice rules with SPOD weights determined by the sequence $0.2j^{-\eta}$, $\eta = 3$, $\alpha = 2$, $s = 16$. The choices $\sigma = 0.1, 0.2, 0.5, 1$ are marked by circle, cross, square and down-triangle, respectively.

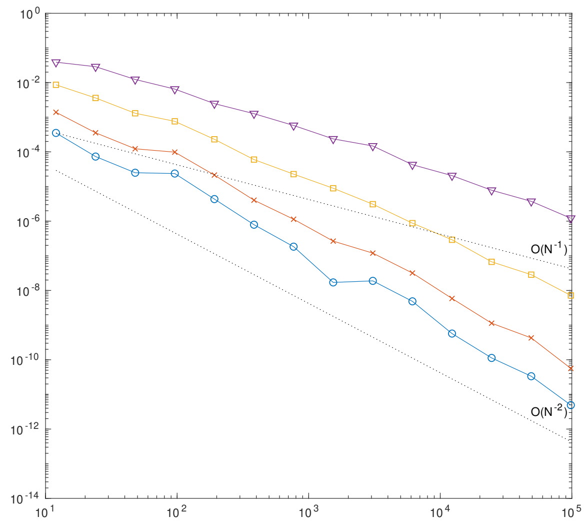

Figure 3:
Relative integration error versus number of QMC points. Extrapolated lattice rules with SPOD weights determined by the sequence $0.2j^{-\eta}$, $\eta = \alpha = 2$, $s = 16$. The choices $\sigma = 0.1, 0.2, 0.5, 1$ are marked by circle, cross, square and down-triangle, respectively.

5.4 Fast Matrix-vector multiplication

We compare the run times of the standard matrix-vector multiplication with the fast algorithm proposed in [11]. This algorithm is based on FFT, as explained in Section 3.5. All timings are performed in MATLAB R2019a, on an Intel(R) Core(TM) i7-7700T CPU @2.90GHz using the `timeit` tool. Since we need to compute α terms of a sequence to perform extrapolation, in each measurement we sum the runtimes corresponding to all α terms involved; here, we set $\alpha = 2$. On the interval $\mathrm{D} = (0, 1)$ we consider the model problem from Section 3 where we define the functions

$$\psi_j(x) = \frac{\sin(j\pi x)}{j^\eta}\ , \quad j = 1, 2, ... \tag{56}$$

with $\eta = 2.1$. Thus, the summability exponent of the sequence $(\beta_j)_{j\geq 1}$ satisfies $p < 1/2$. Based on Theorem 3.8, with first order, conforming FEM ($d = t = t' = 1$), we expect a dimension-

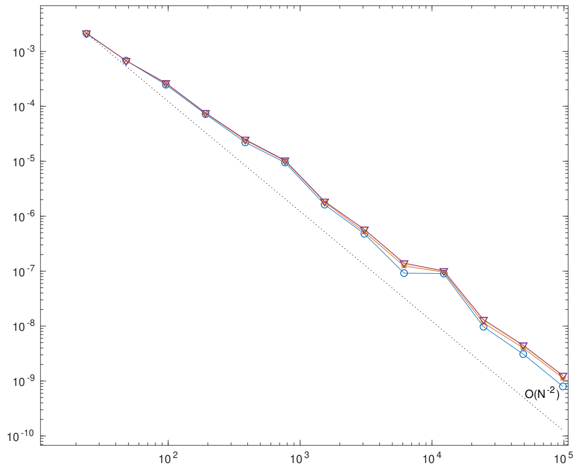

Figure 4:
Relative integration error versus number of QMC points. Extrapolated lattice rules with SPOD weights determined by the sequence $0.2j^{-\eta}$, $\eta = 2.5$, $\alpha = 2$, $\sigma = 0.5$. The choices $s = 16, 32, 64, 128$ are marked by circle, cross, square and down-triangle, respectively.

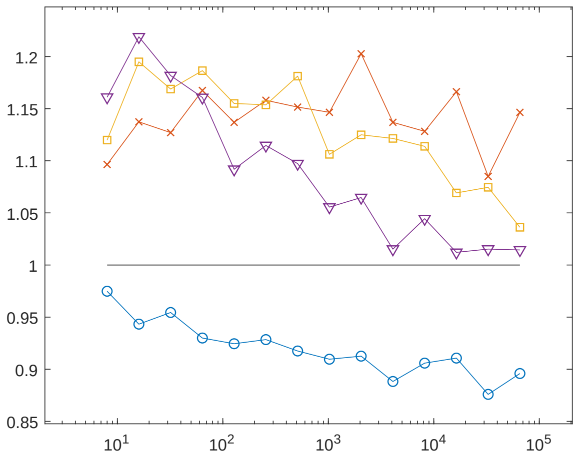

Figure 5:
QMC efficiency index (i.e., ratio between the QMC a-posteriori integration error estimator and actual integration error), versus number of QMC points, for the choices $\eta = 1.5, 1.9, 2.1, 2.5$ marked by circle, cross, square and down-triangle respectively. Here, $\alpha = 2$, $\sigma = 1$, $s = 16$, SPOD weights generated by the sequence $0.2j^{-\eta}$.

independent convergence rate arbitrarily close to $\mathcal{O}\big(N^{-2} + M^{-2} + s^{-3}\big)$. Equilibrating the (upper bounds on the) error contributions, we arrive at the choice $N \sim M \sim s^{3/2}$. On the other hand, if $d = 2$ or if we are interested in the FEM error measured in the $H^1(\mathrm{D})$ norm instead of the QoI, we have $\mathcal{O}\big(N^{-2} + M^{-1} + s^{-3}\big)$, which in turn implies $N \sim M^2 \sim s^{3/2}$.

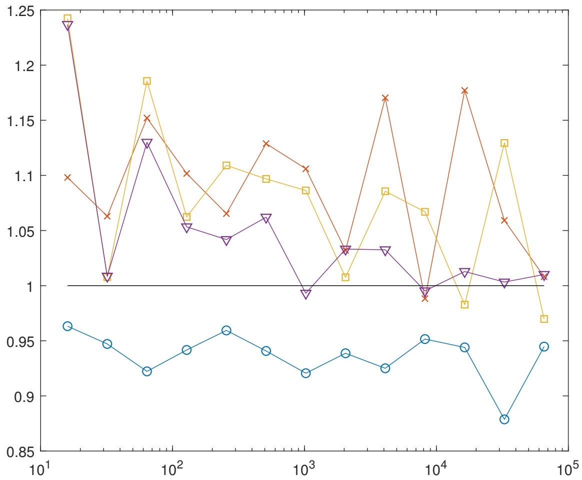

Figure 6:
QMC efficiency index (i.e., ratio between the QMC a-posteriori integration error estimator and actual integration error), versus number of QMC points, for the choices $\eta = 1.5, 1.9, 2.1, 2.5$ marked by circle, cross, square and down-triangle respectively. Here, $\alpha = 3$, $\sigma = 1$, $s = 16$, SPOD weights generated by the sequence $0.2j^{-\eta}$.

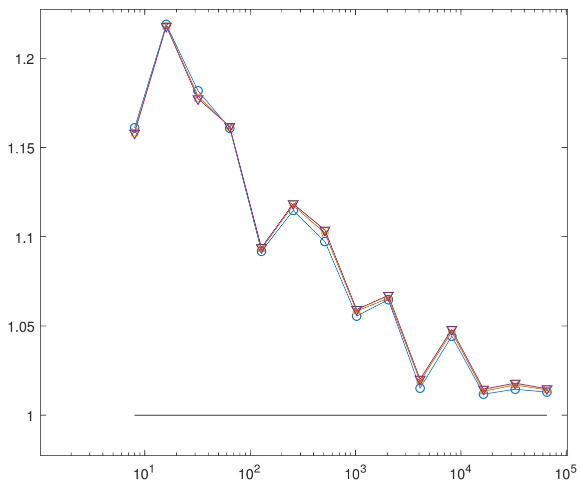

Figure 7:
QMC efficiency index (i.e., ratio between the QMC a-posteriori integration error estimator and actual integration error), versus number of QMC points, for $s = 16, 32, 64, 128$ dimensions marked by circle, cross, square and down-triangle respectively. Here, $\alpha = 2$, $\sigma = 1$, $\eta = 2.5$, SPOD weights generated by the sequence $0.2j^{-\eta}$. Reference values in high dimension are computed with 2^{20} IPL points with interlacing factor 2.

	Times (sec)					
	$M = N, s = \lceil N^{2/3} \rceil$		$M = \lceil N^{3/2} \rceil, s = \lceil N^{2/3} \rceil$		$M = N^2, s = \lceil N^{2/3} \rceil$	
N	Slow	Fast	Slow	Fast	Slow	Fast
48	0.0009	0.0014	0.0012	0.0021	0.0015	0.0057
96	0.0013	0.0016	0.0020	0.0052	0.0053	0.0371
192	0.0022	0.0046	0.0047	0.0412	0.3618	0.6595
384	0.0060	0.0111	0.1009	0.2000	5.3307	4.1966
768	0.0134	0.0349	1.8550	1.2186		
1536	0.0573	0.1484	18.6773	6.5630		
3072	1.8559	0.8218				
6144	17.9490	2.6575				
12288	118.2711	14.5797				

Table 1: Runtimes, in seconds, of the slow and fast matrix-vector multiplication for three sets of choices for N, M, s.

The results in Table 1 show a benefit of the fast MV algorithm for large values of the parameter dimension s and of N when $M = N$ or $M \sim N^{3/2}$. On the other hand, the memory demand increases with $\mathcal{O}(N^3)$ for the fast algorithm when $M = N^2$. This limits the range of N in the numerical experiments, for this choice of M. Moreover, compared to the numerical experiments in [11], we require stronger summability to achieve higher order convergence rates, which results in lower dimensionality of the problem and consequently smaller benefits of the fast algorithm.

5.5 Elliptic parametric PDE

We consider QMC-FE forward UQ for the model, linear, affine-parametric elliptic PDE (27) on the convex physical domain $\mathrm{D} = (0,1)^2$ with deterministic source $f \equiv 1$, QoI $G(u) := \int_{\mathrm{D}} u$ and with affine-parametric diffusion coefficient

$$a(x, \boldsymbol{y}) = 1 + \sum_{j=1}^{s} y_j \psi_j(x) \ , \quad x \in \mathrm{D}.$$

Here $\psi_j(x) := \psi_{\boldsymbol{k}}(x) = \frac{1}{(k_1^2+k_2^2)^\eta} \sin(k_1 \pi x_1) \sin(k_2 \pi x_2)$, $s = 16$ and the ordering is defined by $\boldsymbol{k} < \bar{\boldsymbol{k}}$ when $k_1^2 + k_2^2 < \bar{k}_1^2 + \bar{k}_2^2$ and is arbitrary when equality holds. We prescribe the asymptotic decay

$$\|\psi_j\|_{L^\infty(\mathrm{D})} \sim j^{-\eta}.$$

Due to the smoothness of f and of the parametric coefficient, i.e. $a(\cdot, \boldsymbol{y}) \in W^{1,\infty}(\mathrm{D})$ for all $\boldsymbol{y}$ which is implied by the preceding assumptions, the convexity of the physical domain D implies $u(\cdot, \boldsymbol{y}) \in H^2(\mathrm{D})$. This, in turn, ensures first order convergence in $H^1(\mathrm{D})$ of the $P1$-FEM on shape-regular, quasiuniform partitions of D into triangles.

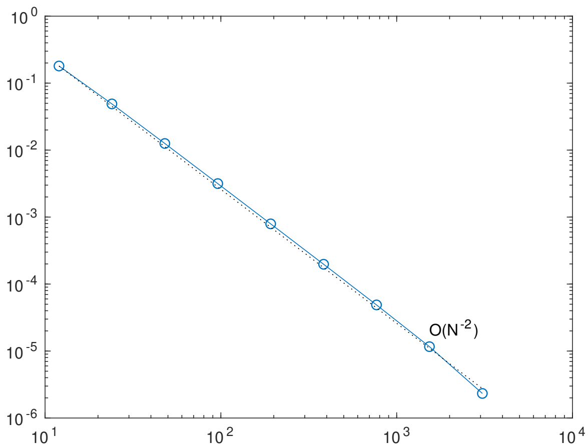

Figure 8:
Relative error versus number of QMC points. Extrapolated lattice rules with SPOD weights determined by the sequence $0.2j^{-\eta}$, $\eta = \alpha = 2$, $s = 16$ and $M \sim N^2$ for all N. The reference value was computed using 2^{12} IPL points with interlacing factor 2.

The resulting convergence of the QMCFEM algorithm is displayed in Figure 8, which confirms the accurate order of $\mathcal{O}(N^{-2})$.

6 Conclusion

We extended the error analysis for extrapolated polynomial lattice rules from [9] to classes of integrand functions with so-called SPOD QMC weights. Such classes typically arise in the computational uncertainty quantification for partial differential equations with distributed uncertain input data which is parametrized in terms of a representation system with globally supported (in the physical domain D) elements (we remark that the setting for integrand functions with product weights which was considered in [9] does accomodate inputs given in terms of locally supported representation systems. See, e.g., [17, 24]) .

We considered only the mathematical analysis of so-called single-level QMC FEM. It is, however, possible to obtain significant gains in error vs. work by combining the presently considered extrapolation methods with a multi-level discretization in physical space. We refer to [31, 12, 8] and to the references there.

The analysis of QMC integration with higher-order, extrapolated polynomial lattice rules in the present paper extends the work [9] to SPOD weights. Under the provision of sufficient summability of higher derivatives of the parametric integrand functions $F(\boldsymbol{y})$, we proved that there exist Richardson-extrapolated QMC integration schemes which afford, with N QMC integration points, convergence rate of $\mathcal{O}(N^{-\alpha})$ for any $\alpha \in \mathbb{N}$. In numerical experiments, however, we find the extrapolation formulas resulting from our analysis to be feasible only for moderate values of $\alpha = 2, 3, 4$. Considerably higher orders of integration are, in our view, theoretically justified, but are practically not feasible due to several reasons: first, large values of α require rather strong summability of the partial derivatives of the integrand function F as expressed in terms of the norm (5). This, in turn, implies that integrands in the class have low effective integration dimension (although formally depending on infinitely many co-ordinates $y_j \in \boldsymbol{y}$).

The presently developed Richardson extrapolated lattice rules afford convergence rates greater than 1 (under the provision of sufficient integrand sparsity, as quantified by the weighted function spaces (5)) without the curse of dimensionality *and* accomodate, due to the structure of their generating vectors, so-called *fast matrix-vector multiplication* developed for first order QMC methods in [11] for the efficient numerical evaluation of parametric solutions of the discretized PDEs at lattice point parameter inputs. QMC quadratures based on so-called interlaced polynomial lattice rules (IPLs) also afford higher order convergence rates without incurring the curse of dimensionality [10, 8]. However, the digit interlacing at the root of their construction precludes the Fast MV multiplication. The presently considered extrapolated polynomial lattice rules are, therefore, the first approach which allows to combine higher order convergence of the QMC integration with the computational advantages of the Fast MV multiplication.

In addition, we showed that the Richardson expansion of the QMC quadrature error can be leveraged to afford an asymptotically exact, computable estimate of the QMC quadrature error.

The analysis of extrapolated polynomial lattice rules in the present paper was developed only for forward uncertainty quantification for model, affine-parametric, linear elliptic boundary value problems, and for single-level Galerkin FEM discretizations of these. Natural extensions of the presently proposed analysis include multi-level QMC-FEM for such problems (e.g. [12, 16, 22]), Bayesian inverse problems (e.g. [8]), and non-affine parametric dependence of the forward PDEs on the parameters (e.g. [14, 25, 27, 28]). Furthermore, the presently proposed, extrapolation-based computable a-posteriori quadrature error estimator may be combined with a-posteriori discretization error estimators for the parametric PDE [32].

Acknowledgement

Josef Dick is partly supported by the Australian Research Council Discovery Project DP190101197 and thanks the FIM institute for the support during the stay at ETH Zürich in January 2020.

References

[1] Thomas Apel, Johannes Pfefferer, and Max Winkler. Local mesh refinement for the discretization of Neumann boundary control problems on polyhedra. *Math. Methods Appl. Sci.*, 39(5):1206–1232, 2016.

[2] Constantin Băcuţă, Hengguang Li, and Victor Nistor. Differential operators on domains with conical points: precise uniform regularity estimates. *Rev. Roumaine Math. Pures Appl.*, 62(3):383–411, 2017.

[3] Constantin Băcuţă, Victor Nistor, and Ludmil T. Zikatanov. Improving the rate of convergence of 'high order finite elements' on polygons and domains with cusps. *Numer. Math.*, 100(2):165–184, 2005.

[4] Constantin Băcuţă, Victor Nistor, and Ludmil T. Zikatanov. Improving the rate of convergence of high-order finite elements on polyhedra. I. A priori estimates. *Numer. Funct. Anal. Optim.*, 26(6):613–639, 2005.

[5] Constantin Băcuţă, Victor Nistor, and Ludmil T. Zikatanov. Improving the rate of convergence of high-order finite elements on polyhedra. II. Mesh refinements and interpolation. *Numer. Funct. Anal. Optim.*, 28(7-8):775–824, 2007.

[6] Jan Baldeaux, Josef Dick, Gunther Leobacher, Dirk Nuyens, and Friedrich Pillichshammer. Efficient calculation of the worst-case error and (fast) component-by-component construction of higher order polynomial lattice rules. *Numer. Algorithms*, 59(3):403–431, 2012.

[7] Albert Cohen, Ronald DeVore, and Christoph Schwab. Convergence rates of best N-term Galerkin approximations for a class of elliptic sPDEs. *Found. Comput. Math.*, 10(6):615–646, 2010.

[8] Josef Dick, Robert Nicholas Gantner, Quoc T. Le Gia, and Christoph Schwab. Multilevel higher-order Quasi-Monte Carlo Bayesian estimation. *Math. Mod. Meth. Appl. Sci.*, 27(5):953–995, 2017.

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

[10] Josef Dick, Frances Y. Kuo, Quoc T. Le Gia, Dirk Nuyens, and Christoph Schwab. Higher order QMC Petrov-Galerkin discretization for affine parametric operator equations with random field inputs. *SIAM J. Numer. Anal.*, 52(6):2676–2702, 2014.

[11] Josef Dick, Frances Y. Kuo, Quoc T. Le Gia, and Christoph Schwab. Fast QMC matrix-vector multiplication. *SIAM J. Sci. Comput.*, 37(3):A1436–A1450, 2015.

[12] Josef Dick, Frances Y. Kuo, Quoc T. Le Gia, and Christoph Schwab. Multilevel higher order QMC Petrov-Galerkin discretization for affine parametric operator equations. *SIAM J. Numer. Anal.*, 54(4):2541–2568, 2016.

[13] Josef Dick, Frances Y. Kuo, and Ian H. Sloan. High-dimensional integration: the quasi-Monte Carlo way. *Acta Numer.*, 22:133–288, 2013.

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

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

[16] Robert N. Gantner, Lukas Herrmann, and Christoph Schwab. Multilevel QMC with product weights for affine-parametric, elliptic PDEs. In *Contemporary computational mathematics—a celebration of the 80th birthday of Ian Sloan. Vol. 1, 2*, pages 373–405. Springer, Cham, 2018.

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

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

[19] Robert Nicholas Gantner. *Computational Bayesian Estimation for PDEs with Random Input Data*. PhD thesis, ETH Zürich, 2017. Dissertation 24529, Examiner Prof. Dr. Christoph Schwab.

[20] Takashi Goda. Quasi–Monte Carlo integration using digital nets with antithetics. *J. Comput. Appl. Math.*, 304:26–42, 2016.

[21] I. G. Graham, Frances Y. Kuo, J. A. Nichols, R. Scheichl, Christoph Schwab, and Ian H. Sloan. Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficients. *Numerische Mathematik*, 131(2):329–368, 2015.

[22] Lukas Herrmann and Christoph Schwab. Multilevel QMC Uncertainty Quantification for Advection-Reaction-Diffusion. Technical Report 2019-06 (revised), Seminar for Applied Mathematics, ETH Zürich, 2019. (to appear in Proc. MCQMC 2018, Springer Publ. 2020).

[23] Lukas Herrmann and Christoph Schwab. Multilevel quasi-Monte Carlo integration with product weights for elliptic PDEs with lognormal coefficients. *ESAIM Math. Model. Numer. Anal.*, 53(5):1507–1552, 2019.

[24] Lukas Herrmann and Christoph Schwab. QMC integration for lognormal-parametric, elliptic PDEs: local supports and product weights. *Numer. Math.*, 141(1):63–102, 2019.

[25] Lukas Herrmann, Christoph Schwab, and Jakob Zech. Uncertainty quantification for spectral fractional diffusion: sparsity analysis of parametric solutions. *SIAM/ASA J. Uncertain. Quantif.*, 7(3):913–947, 2019.

[26] Jan S. Hesthaven, Gianluigi Rozza, and Benjamin Stamm. *Certified reduced basis methods for parametrized partial differential equations*. SpringerBriefs in Mathematics. Springer, Cham; BCAM Basque Center for Applied Mathematics, Bilbao, 2016. BCAM SpringerBriefs.

[27] R. Hiptmair, L. Scarabosio, C. Schillings, and Ch. Schwab. Large deformation shape uncertainty quantification in acoustic scattering. *Adv. Comput. Math.*, 44(5):1475–1518, 2018.

[28] Carlos Jerez-Hanckes, Christoph Schwab, and Jakob Zech. Electromagnetic wave scattering by random surfaces: shape holomorphy. *Math. Models Methods Appl. Sci.*, 27(12):2229–2259, 2017.

[29] Frances Y. Kuo, Christoph Schwab, and Ian H. Sloan. Quasi-Monte Carlo methods for high-dimensional integration: the standard (weighted Hilbert space) setting and beyond. *ANZIAM J.*, 53(1):1–37, 2011.

[30] Frances Y. Kuo, Christoph Schwab, and Ian H. Sloan. Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients. *SIAM J. Numer. Anal.*, 50(6):3351–3374, 2012.

[31] Frances Y. Kuo, Christoph Schwab, and Ian H. Sloan. Multi-level Quasi-Monte Carlo Finite Element Methods for a Class of Elliptic PDEs with Random Coefficients. *Found. Comput. Math.*, 15(2):411–449, April 2015.

[32] M. Longo. Adaptive algorithms with a-posteriori Quasi-Monte Carlo estimation for parametric elliptic PDEs. Technical Report 2021-03, Seminar for Applied Mathematics, ETH Zürich, Switzerland, 2021.

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

[34] Dirk Nuyens and Ronald Cools. Fast algorithms for component-by-component construction of rank-1 lattice rules in shift-invariant reproducing kernel Hilbert spaces. *Math. Comp.*, 75(254):903–920, 2006.

[35] Alfio Quarteroni, Andrea Manzoni, and Federico Negri. *Reduced basis methods for partial differential equations*, volume 92 of *Unitext*. Springer, Cham, 2016. An introduction, La Matematica per il 3+2.

[36] Christoph Schwab and Radu Alexandru Todor. Karhunen-Loève Approximation of Random Fields by Generalized Fast Multipole Methods. *Journal of Computational Physics*, 217:100–122, 2006.