

Higher order QMC Galerkin discretization for parametric operator equations

J. Dick and F.Y. Kuo and Q.T. Le Gia and D. Nuyens and Ch. Schwab

Research Report No. 2013-29
September 2013

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

HIGHER ORDER QMC GALERKIN DISCRETIZATION FOR PARAMETRIC OPERATOR EQUATIONS*

JOSEF DICK[†], FRANCES Y. KUO[†], QUOC T. LE GIA[†], DIRK NUYENS[‡], AND
CHRISTOPH SCHWAB[§]

Abstract. We construct quasi-Monte Carlo methods to approximate the expected values of linear functionals of Galerkin discretizations of parametric operator equations which depend on a possibly infinite sequence of parameters. Such problems arise in the numerical solution of differential and integral equations with random field inputs. We analyze the regularity of the solutions with respect to the parameters in terms of the rate of decay of the fluctuations of the input field. If $p \in (0, 1]$ denotes the “summability exponent” corresponding to the fluctuations in affine-parametric families of operators, then we prove that deterministic “interlaced polynomial lattice rules” of order $\alpha = \lfloor 1/p \rfloor + 1$ in s dimensions with N points can be constructed using a fast component-by-component algorithm, in $\mathcal{O}(\alpha s N \log N + \alpha^2 s^2 N)$ operations, to achieve a convergence rate of $\mathcal{O}(N^{-1/p})$, with the implied constant independent of s . This dimension-independent convergence rate is superior to the rate $\mathcal{O}(N^{-1/p+1/2})$, for $2/3 \leq p \leq 1$ recently established for randomly shifted lattice rules under comparable assumptions. In our analysis we use a non-standard Banach space setting and introduce “smoothness-driven product and order dependent (SPOD)” weights for which we show fast CBC construction.

Key words. Quasi Monte-Carlo methods, interlaced polynomial lattice rules, higher order digital nets, parametric operator equations, infinite dimensional quadrature, Galerkin discretization

AMS subject classifications. 65D30, 65D32, 65N30

1. Introduction. The efficient numerical computation of statistical quantities for solutions of partial differential and of integral equations with random inputs is a key task in uncertainty quantification in engineering and in the sciences. The quantity of interest is expressed as a mathematical expectation, and the efficient computation of these quantities involves two basic steps: i) the approximate (numerical) solution of the operator equation, and ii) the approximate evaluation of the mathematical expectation by numerical integration. In the present paper, we outline a strategy towards these two aims which is based on i) *Galerkin discretization* of the operator equation and on ii) *Quasi-Monte Carlo* (QMC) integration.

The present paper is motivated in part by [18], where QMC integration using a family of *randomly shifted lattice rules* was combined with Finite Element discretization for a model parametric diffusion equation, and in part by [27], where the methodology was extended to an abstract family of parametric operator equations. In this paper, in contrast to [18, 27], we use *deterministic*, “*interlaced polynomial lattice rules*”, which provide a convergence rate beyond order one for smooth integrands; whereas order one was the limitation in [18, 27].

Contrary to Monte Carlo methods which require uniformly distributed samples of random input functions, QMC (and other) quadrature methods require the introduction of coordinates of integration prior to numerical quadrature. In the context

*September 19, 2013

[†]School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052, Australia (josef.dick@unsw.edu.au, f.kuo@unsw.edu.au, qlegia@unsw.edu.au). The work of these authors was supported by Australian Research Council Discovery Projects. The first author’s work was additionally supported by an Australian Research Council QEII Fellowship.

[‡]Department of Computer Science, KU Leuven, Belgium (dirk.nuyens@cs.kuleuven.be). This author is a fellow of the Research Foundation Flanders (FWO).

[§]Seminar for Applied Mathematics, ETH Zürich, ETH Zentrum, HG G57.1, CH8092 Zürich, Switzerland (christoph.schwab@sam.math.ethz.ch). This author’s work was supported by the European Research Council under grant AdG247277.

of random field inputs with non-degenerate covariance operators, a *countable* number of coordinates is required to describe the random input data, e.g., by a Karhunen-Loève expansion. Therefore, in the present work, we consider in particular that the operator equation contains not only a finite number of random input parameters, but rather depends on *random field inputs*, i.e., it contains random functions of space and, in evolution problems, of time which describe uncertainty in the problem under consideration.

More precisely, let $\mathbf{y} := (y_j)_{j \geq 1}$ denote the possibly countable set of parameters from a domain $U \subseteq \mathbb{R}^{\mathbb{N}}$, and let $A(\mathbf{y})$ denote a \mathbf{y} -parametric bounded linear operator between suitably defined spaces \mathcal{X} and \mathcal{Y}' . Then we wish to solve the following parametric operator equation: given $f \in \mathcal{Y}'$, for every $\mathbf{y} \in U$ find $u(\mathbf{y}) \in \mathcal{X}$ such that

$$A(\mathbf{y})u(\mathbf{y}) = f. \quad (1.1)$$

Such parametric operator equations arise from partial differential equations with random field input, see, e.g., [28]. We assume in this paper the simplest case, namely, that $A(\mathbf{y})$ has “*affine*” *parameter dependence*, i.e., there exists a sequence $\{A_j\}_{j \geq 0} \subset \mathcal{L}(\mathcal{X}, \mathcal{Y}')$ such that for every $\mathbf{y} \in U$ we can write

$$A(\mathbf{y}) = A_0 + \sum_{j \geq 1} y_j A_j. \quad (1.2)$$

A concrete example is the diffusion problem considered, e.g., in [18], in which the diffusion in random media is modeled by equation (1.1) with $A(\mathbf{y}) = -\nabla \cdot (a(\mathbf{y})\nabla)$, and where the diffusion coefficients are expanded in terms of a Karhunen-Loève expansion $a(\mathbf{y}) = \bar{a} + \sum_{j \geq 1} y_j \psi_j$, leading to $A_0 = -\nabla \cdot (\bar{a}\nabla)$ and $A_j = -\nabla \cdot (\psi_j \nabla)$. Here, as in [18], we restrict ourselves to the (infinite-dimensional) parameter domain

$$U = [-\frac{1}{2}, \frac{1}{2}]^{\mathbb{N}}.$$

Some assumptions on the “*nominal*” (or “*mean field*”) operator A_0 and the “*fluctuation*” operators A_j are required to ensure that the sum in (1.2) converges, and to ensure existence and uniqueness of the solution $u(\mathbf{y})$ in (1.1) for all $\mathbf{y} \in U$; these will be given in §2. In addition, we shall consider the parametric Galerkin approximation $u^h(\mathbf{y}) \in \mathcal{X}_h \subset \mathcal{X}$, to be defined in (2.14) below, as well as $u_s^h(\mathbf{y})$, corresponding to the Galerkin approximation of the problem with the sum in (1.2) truncated to s terms (this is equivalent to setting $y_j = 0$ for $j > s$). Further assumptions on A_0 and A_j are required for our regularity and approximation results; these will all be given in §2. For now we mention only one key assumption, namely, that there exists $p \in (0, 1]$ for which

$$\sum_{j \geq 1} \|A_j\|_{\mathcal{L}(\mathcal{X}, \mathcal{Y}')}^p < \infty, \quad (1.3)$$

where $\|\cdot\|_{\mathcal{L}(\mathcal{X}, \mathcal{Y}')}$ denotes the operator norm for the set of all bounded linear mappings from \mathcal{X} to \mathcal{Y}' . This assumption implies a decay of the fluctuation coefficients A_j , with stronger decay as the value of $p \in (0, 1]$ decreases.

For a given bounded linear functional $G(\cdot) : \mathcal{X} \rightarrow \mathbb{R}$, we are interested in computing expected values of $G(u(\mathbf{y}))$ with respect to $\mathbf{y} \in U$, i.e., an integral of the functional $G(\cdot)$ of the parametric solution,

$$I(G(u)) := \int_U G(u(\mathbf{y})) \, d\mathbf{y}, \quad (1.4)$$

over the infinite dimensional domain of integration U . We truncate the infinite sum in (1.2) to s terms and solve the corresponding operator equation (1.1) using Galerkin discretization from a dense, one-parameter family $\{\mathcal{X}^h\}$ of subspaces of \mathcal{X} . Denoting this dimension-truncated Galerkin solution by u_s^h , we then approximate the corresponding s -dimensional integral using QMC quadrature,

$$\frac{1}{N} \sum_{n=0}^{N-1} G(u_s^h(\mathbf{y}_n - \frac{1}{2})), \quad (1.5)$$

where $\mathbf{y}_0, \dots, \mathbf{y}_{N-1} \in [0, 1]^s$ denote N points from a properly chosen QMC rule, and the shift of coordinates by $\frac{1}{2}$ takes care of the translation from $[0, 1]^s$ to $[-\frac{1}{2}, \frac{1}{2}]^s$. Note that each evaluation of the integrand at a single QMC point \mathbf{y}_n requires the approximate (Galerkin) solution of one operator equation for $u_s^h(\mathbf{y}_n)$.

There are three sources of error in approximating (1.4) by (1.5): a *Galerkin discretization error* depending on h , a *dimension truncation error* depending on s , and a *QMC quadrature error* depending on N . The main focus of this paper will be on the analysis of the QMC error: we prove that *interlaced polynomial lattice rules* [12, 13] can be constructed using a component-by-component (CBC) algorithm to achieve a rate of convergence of

$$\mathcal{O}(N^{-1/p}),$$

with $p \in (0, 1]$ as in (1.3), and with the implied constant independent of N , h , and s .

The function space setting for QMC integration considered in this paper uses a Banach space norm with two parameters $1 \leq q \leq \infty$ and $1 \leq r \leq \infty$, corresponding to the L_q norm of functions and an ℓ_r norm of vectors combining these L_q norms. Often q and r are taken to be the same value in the literature, with $q = r = 2$ giving the Hilbert space setting. However, as discussed in [19], decoupling q and r allows more flexibility in the analysis, since the two parameters play different roles. (The L_q norm increases with increasing q , while the ℓ_r norm increases with decreasing r .) The results in [18, 27] are based on the Hilbert space setting, with a convergence rate of $\mathcal{O}(N^{-1/p+1/2})$. The main result of this paper is based on $r = \infty$ and it holds for all values of q . The convergence rate of $\mathcal{O}(N^{-1/p})$ obtained in this paper is an improvement by a factor of $N^{-1/2}$.

In comparison, under the same assumption (1.3), the paper [5] establishes p -summability of generalized (Legendre) polynomial chaos expansions of the integrand $G(u(\cdot))$ in (1.4), and shows that N -term approximation of the integrand has a convergence rate in L_2 norm of $\mathcal{O}(N^{-1/p+1/2})$, with the implied constant independent of the dimension of the integration domain. This rate could be realized, for example, by adaptive Galerkin projections. It also suggests an N -term approximation rate in the (natural for integration) L_1 norm of $\mathcal{O}(N^{-1/p})$.

Our QMC quadrature approach requires the use of interlaced polynomial lattice rules of order $\alpha = \lceil 1/p \rceil + 1$, which is at least 2. Similar to the analysis in [18, 27], we need to choose “weights” for the function space setting to ensure that the implied constant for the convergence rate is bounded independently of the truncation dimension s . The regularity analysis reveals the need to use weights that are *not* of “POD” form (namely, “product and order dependent” form) as in [18], but of a more general form which we call “SPOD weights”, for “smoothness-driven product and order dependent” weights, see (3.17) ahead. For these SPOD weights, we develop a new *fast component-by-component construction* of interlaced polynomial lattice rules, with cost of $\mathcal{O}(\alpha s N \log N + \alpha^2 s^2 N)$ operations.

The outline of this paper is as follows. In §2 we present a class of parametric operator equations, review the parametric and spatial regularities of their solutions, give a synopsis of the Galerkin discretization of these equations, and outline some estimates relating to dimension truncation. In §3 we derive a worst case error bound for digital nets in a novel weighted Banach space setting, prove that interlaced polynomial lattice rules can be constructed by a CBC algorithm to achieve a dimension-independent error bound with a good convergence rate, and explain how to implement the algorithm in an efficient way. Finally in §4 we summarize the combined QMC Galerkin error bound.

2. Problem formulation. Generalizing results of [5], we study well-posedness, regularity and polynomial approximation of solutions for a family of abstract parametric saddle point problems, with operators depending on a sequence of parameters. The results cover a wide range of operator equations: among them are (stationary and time-dependent) diffusion in random media [5], wave propagation [15], parametric, nonlinear PDEs [4] and optimal control problems for uncertain systems [17].

2.1. Parametric operator equations. We denote by \mathcal{X} and \mathcal{Y} two separable and reflexive Banach spaces over \mathbb{R} (all results will hold with the obvious modifications also for spaces over \mathbb{C}) with (topological) duals \mathcal{X}' and \mathcal{Y}' , respectively. By $\mathcal{L}(\mathcal{X}, \mathcal{Y}')$, we denote the set of bounded linear operators $A : \mathcal{X} \rightarrow \mathcal{Y}'$.

As we explained in the introduction, let $\mathbf{y} := (y_j)_{j \geq 1} \in U = [-\frac{1}{2}, \frac{1}{2}]^{\mathbb{N}}$ be a countable set of parameters. For every $f \in \mathcal{Y}'$ and every $\mathbf{y} \in U$, we wish to solve the parametric operator equation (1.1), where the operator $A(\mathbf{y}) \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$ is of affine parameter dependence, see (1.2). In order for the sum in (1.2) to converge, we impose the following assumptions on the sequence $\{A_j\}_{j \geq 0} \subset \mathcal{L}(\mathcal{X}, \mathcal{Y}')$. In doing so, we associate with the operator A_j the bilinear forms $\mathbf{a}_j(\cdot, \cdot) : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ via

$$\forall v \in \mathcal{X}, w \in \mathcal{Y} : \quad \mathbf{a}_j(v, w) = {}_{\mathcal{Y}}\langle w, A_j v \rangle_{\mathcal{Y}'}, \quad j = 0, 1, 2, \dots .$$

ASSUMPTION 1. *The sequence $\{A_j\}_{j \geq 0}$ in (1.2) satisfies the following conditions:*

1. *The nominal operator $A_0 \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$ is boundedly invertible, i.e., there exists $\mu_0 > 0$ such that (cf. the inf-sup conditions in [3])*

$$\inf_{0 \neq v \in \mathcal{X}} \sup_{0 \neq w \in \mathcal{Y}} \frac{\mathbf{a}_0(v, w)}{\|v\|_{\mathcal{X}} \|w\|_{\mathcal{Y}}} \geq \mu_0, \quad \inf_{0 \neq w \in \mathcal{Y}} \sup_{0 \neq v \in \mathcal{X}} \frac{\mathbf{a}_0(v, w)}{\|v\|_{\mathcal{X}} \|w\|_{\mathcal{Y}}} \geq \mu_0. \quad (2.1)$$

2. *The fluctuation operators $\{A_j\}_{j \geq 1}$ are small with respect to A_0 in the following sense: there exists a constant $0 < \kappa < 2$ such that for μ_0 as in (2.1) and some $s \geq 1$ we have*

$$\sum_{j \geq s+1} \beta_j \leq \kappa < 2, \quad \text{where } \beta_j := \|A_0^{-1} A_j\|_{\mathcal{L}(\mathcal{X}, \mathcal{Y}')} , \quad j = 1, 2, \dots . \quad (2.2)$$

Assumption 1 is sufficient for the bounded invertibility of $A(\mathbf{y})$, uniformly with respect to the parameter sequence $\mathbf{y} \in U$.

THEOREM 2.1. *Under Assumption 1, for every realization $\mathbf{y} \in U$ of the parameter vector, the affine parametric operator $A(\mathbf{y})$ given by (1.2) is boundedly invertible. Specifically, for the bilinear form $\mathbf{a}(\mathbf{y}; \cdot, \cdot) : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$ associated with $A(\mathbf{y}) \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$ via*

$$\mathbf{a}(\mathbf{y}; v, w) := {}_{\mathcal{Y}}\langle w, A(\mathbf{y})v \rangle_{\mathcal{Y}'}, \quad (2.3)$$

there hold the uniform (with respect to $\mathbf{y} \in U$) inf-sup conditions with $\mu = (1 - \kappa/2)\mu_0$,

$$\forall \mathbf{y} \in U : \quad \inf_{0 \neq v \in \mathcal{X}} \sup_{0 \neq w \in \mathcal{Y}} \frac{\mathbf{a}(\mathbf{y}; v, w)}{\|v\|_{\mathcal{X}} \|w\|_{\mathcal{Y}}} \geq \mu, \quad \inf_{0 \neq w \in \mathcal{Y}} \sup_{0 \neq v \in \mathcal{X}} \frac{\mathbf{a}(\mathbf{y}; v, w)}{\|v\|_{\mathcal{X}} \|w\|_{\mathcal{Y}}} \geq \mu. \quad (2.4)$$

In particular, for every $f \in \mathcal{Y}'$ and for every $\mathbf{y} \in U$, the parametric operator equation

$$\text{find } u(\mathbf{y}) \in \mathcal{X} : \quad \mathbf{a}(\mathbf{y}; u(\mathbf{y}), w) = {}_{\mathcal{Y}}\langle w, f \rangle_{\mathcal{Y}'} \quad \forall w \in \mathcal{Y} \quad (2.5)$$

admits a unique solution $u(\mathbf{y})$ which satisfies the a-priori estimate

$$\|u(\mathbf{y})\|_{\mathcal{X}} \leq \frac{1}{\mu} \|f\|_{\mathcal{Y}'}. \quad (2.6)$$

For a proof of the theorem, we refer to [27, Theorem 2].

2.2. Parametric regularity of solutions. In this subsection we study the dependence of the solution $u(\mathbf{y})$ of the parametric, variational problem (2.5) on the parameter vector \mathbf{y} . In the following, let $\mathbb{N}_0^{\mathbb{N}}$ denote the set of sequences $\boldsymbol{\nu} = (\nu_j)_{j \geq 1}$ of nonnegative integers ν_j , and let $|\boldsymbol{\nu}| := \sum_{j \geq 1} \nu_j$. For $|\boldsymbol{\nu}| < \infty$, we denote the partial derivative of order $\boldsymbol{\nu}$ of $u(\mathbf{y})$ with respect to \mathbf{y} by $\partial_{\mathbf{y}}^{\boldsymbol{\nu}} u := (\partial^{|\boldsymbol{\nu}|} u) / (\partial_{y_1}^{\nu_1} \partial_{y_2}^{\nu_2} \dots)$.

THEOREM 2.2. [5, 17] *Under Assumption 1, there exists a constant $C_0 > 0$ such that for every $f \in \mathcal{Y}'$ and for every $\mathbf{y} \in U$, the partial derivatives of the parametric solution $u(\mathbf{y})$ of the parametric operator equation (1.1) with affine operator (1.2) satisfy the bounds*

$$\|(\partial_{\mathbf{y}}^{\boldsymbol{\nu}} u)(\mathbf{y})\|_{\mathcal{X}} \leq C_0 |\boldsymbol{\nu}|! \boldsymbol{\beta}^{\boldsymbol{\nu}} \|f\|_{\mathcal{Y}'}, \quad \text{for all } \boldsymbol{\nu} \in \mathbb{N}_0^{\mathbb{N}} \text{ with } |\boldsymbol{\nu}| < \infty, \quad (2.7)$$

where $0! := 1$, $\boldsymbol{\beta}^{\boldsymbol{\nu}} := \prod_{j \geq 1} \beta_j^{\nu_j}$, with β_j as in (2.2), and $|\boldsymbol{\nu}| = \sum_{j \geq 1} \nu_j$.

2.3. Spatial regularity of solutions. We assume given scales of smoothness spaces $\{\mathcal{X}_t\}_{t \geq 0}$, $\{\mathcal{Y}_t\}_{t \geq 0}$, with

$$\begin{aligned} \mathcal{X} &= \mathcal{X}_0 \supset \mathcal{X}_1 \supset \mathcal{X}_2 \supset \dots, & \mathcal{Y} &= \mathcal{Y}_0 \subset \mathcal{Y}_1 \subset \mathcal{Y}_2 \subset \dots, & \text{and} \\ \mathcal{X}' &= \mathcal{X}'_0 \subset \mathcal{X}'_1 \subset \mathcal{X}'_2 \subset \dots, & \mathcal{Y}' &= \mathcal{Y}'_0 \supset \mathcal{Y}'_1 \supset \mathcal{Y}'_2 \supset \dots. \end{aligned} \quad (2.8)$$

The scales are assumed to be defined also for noninteger values of the smoothness parameter $t \geq 0$ by interpolation. For self-adjoint operators, usually $\mathcal{X}_t = \mathcal{Y}_t$. For example, in diffusion problems in *convex domains* D considered in [5, 18], the smoothness scales (2.8) are $\mathcal{X} = \mathcal{Y} = H_0^1(D)$, $\mathcal{X}_1 = \mathcal{Y}_1 = (H^2 \cap H_0^1)(D)$, $\mathcal{Y}' = H^{-1}(D)$, $\mathcal{Y}'_1 = L^2(D)$. In a nonconvex polygon (or polyhedron), analogous smoothness scales are available, but involve Sobolev spaces with weights¹.

In [22], this kind of abstract regularity result was established for a wide range of second order parametric, elliptic systems in 2D and 3D, also for higher order regularity. Importantly, the smoothness scales are then weighted Sobolev spaces $\mathcal{K}_{a+1}^{t+1}(D)$ of Kondratiev type in D , and hence $\mathcal{X}_t = \mathcal{K}_{a+1}^{t+1}(D)$, $\mathcal{Y}'_t = \mathcal{K}_{a-1}^{t-1}(D)$ in this case. The Finite Element spaces which realize the maximal convergence rates (beyond order one) are regular, simplicial families in the sense of Ciarlet, on suitably refined meshes which compensate for the corner and edge singularities.

In the ensuing convergence analysis of Galerkin discretizations of (1.1), we will assume that the data regularity $f \in \mathcal{Y}'_t$ for some $t > 0$ implies that

$$\forall \mathbf{y} \in U : \quad u(\mathbf{y}) = A(\mathbf{y})^{-1} f \in \mathcal{X}_t. \quad (2.9)$$

¹Not to be confused with the weighted Sobolev spaces in QMC error analysis, see, e.g., [29, 19].

Parametric regularity is available for numerous parametric differential equations (see [28, 14, 16, 4, 17] and the references there) as well as for posterior densities in Bayesian inverse problems [25, 26].

2.4. Galerkin discretization. Let $\{\mathcal{X}^h\}_{h>0} \subset \mathcal{X}$ and $\{\mathcal{Y}^h\}_{h>0} \subset \mathcal{Y}$ be two families of finite dimensional subspaces which are dense in \mathcal{X} and in \mathcal{Y} , respectively. We will also assume the *approximation properties*: for $0 < t \leq \bar{t}$ and $0 < t' \leq \bar{t}'$, and for $0 < h \leq h_0$, there hold

$$\begin{aligned} \forall v \in \mathcal{X}_t : \quad \inf_{v^h \in \mathcal{X}^h} \|v - v^h\|_{\mathcal{X}} &\leq C_t h^t \|v\|_{\mathcal{X}_t}, \\ \forall w \in \mathcal{Y}_{t'} : \quad \inf_{w^h \in \mathcal{Y}^h} \|w - w^h\|_{\mathcal{Y}} &\leq C_{t'} h^{t'} \|w\|_{\mathcal{Y}_{t'}}. \end{aligned} \quad (2.10)$$

The maximum amount of smoothness in the scale \mathcal{X}_t , denoted by \bar{t} , depends on the problem class under consideration and on the Sobolev scale: e.g., for elliptic problems in polygonal domains, it is well known that choosing for \mathcal{X}_t the usual Sobolev spaces will allow (2.9) with t only in a possibly small interval $0 < t \leq \bar{t}$, whereas choosing \mathcal{X}_t as Sobolev spaces with weights will allow rather large values of \bar{t} (see, e.g., [22]). Corresponding to (2.9), we shall assume that

$$\forall 0 \leq t \leq \bar{t} : \quad \sup_{\mathbf{y} \in U} \|A(\mathbf{y})^{-1}\|_{\mathcal{L}(\mathcal{Y}'_t, \mathcal{X}_t)} < \infty. \quad (2.11)$$

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

$$\forall \mathbf{y} \in U : \quad \inf_{0 \neq v^h \in \mathcal{X}^h} \sup_{0 \neq w^h \in \mathcal{Y}^h} \frac{\mathbf{a}(\mathbf{y}; v^h, w^h)}{\|v^h\|_{\mathcal{X}} \|w^h\|_{\mathcal{Y}}} \geq \bar{\mu} > 0, \quad (2.12)$$

$$\forall \mathbf{y} \in U : \quad \inf_{0 \neq w^h \in \mathcal{Y}^h} \sup_{0 \neq v^h \in \mathcal{X}^h} \frac{\mathbf{a}(\mathbf{y}; v^h, w^h)}{\|v^h\|_{\mathcal{X}} \|w^h\|_{\mathcal{Y}}} \geq \bar{\mu} > 0. \quad (2.13)$$

Then, for every $0 < h \leq h_0$ and for every $\mathbf{y} \in U$, the Galerkin approximations $u^h(\mathbf{y}) \in \mathcal{X}^h$, given by

$$\text{find } u^h(\mathbf{y}) \in \mathcal{X}^h : \quad \mathbf{a}(\mathbf{y}; u^h(\mathbf{y}), w^h) = {}_{\mathcal{Y}'} \langle w^h, f \rangle_{\mathcal{Y}'} \quad \forall w^h \in \mathcal{Y}^h, \quad (2.14)$$

admits a unique solution $u^h(\mathbf{y})$ which satisfies the a-priori estimate

$$\|u^h(\mathbf{y})\|_{\mathcal{X}} \leq \frac{1}{\bar{\mu}} \|f\|_{\mathcal{Y}'}. \quad (2.15)$$

Moreover, there exists a constant $C > 0$ such that for all $\mathbf{y} \in U$ quasioptimality holds,

$$\|u(\mathbf{y}) - u^h(\mathbf{y})\|_{\mathcal{X}} \leq \frac{C}{\bar{\mu}} \inf_{0 \neq v^h \in \mathcal{X}^h} \|u(\mathbf{y}) - v^h\|_{\mathcal{X}}. \quad (2.16)$$

We remark that under Assumption 1, the validity of the discrete inf-sup conditions for the nominal bilinear form $\mathbf{a}_0(\cdot, \cdot)$, see (2.1), with constant $\bar{\mu}_0 > 0$ independent of h , implies (2.12) and (2.13) for the bilinear form $\mathbf{a}(\mathbf{y}; \cdot, \cdot)$ with $\bar{\mu} = (1 - \kappa/2)\bar{\mu}_0 > 0$.

THEOREM 2.4. *Under Assumption 1 and condition (2.11), for every $f \in \mathcal{Y}'$ and for every $\mathbf{y} \in U$, the approximations $u^h(\mathbf{y})$ are stable, i.e., (2.15) holds. For every $f \in \mathcal{Y}'_t$ with $0 < t \leq \bar{t}$, there exists a constant $C > 0$ such that as $h \rightarrow 0$ there holds*

$$\|u(\mathbf{y}) - u^h(\mathbf{y})\|_{\mathcal{X}} \leq C h^t \|f\|_{\mathcal{Y}'_t}. \quad (2.17)$$

Since we are interested in the expectations of functionals of the parametric solution, see (1.4), we will also impose a regularity assumption on the functional $G(\cdot) \in \mathcal{X}'$:

$$\exists 0 < t' \leq \bar{t}: \quad G(\cdot) \in \mathcal{X}'_{t'}, \quad (2.18)$$

and the *adjoint regularity*: for t' as in (2.18) there exists $C_{t'} > 0$ such that for every $\mathbf{y} \in U$,

$$w(\mathbf{y}) = (A^*(\mathbf{y}))^{-1}G \in \mathcal{Y}_{t'}, \quad \|w(\mathbf{y})\|_{\mathcal{Y}_{t'}} \leq C_{t'} \|G\|_{\mathcal{X}'_{t'}}. \quad (2.19)$$

Moreover, we see from (1.5) that the discretization error of $G(u(\mathbf{y}))$ is of interest as well. It is known that $|G(u(\mathbf{y})) - G(u^h(\mathbf{y}))|$ may converge faster than $\|u(\mathbf{y}) - u^h(\mathbf{y})\|_{\mathcal{X}}$.

THEOREM 2.5. *Under Assumption 1 and the conditions (2.11) and (2.19), for every $f \in \mathcal{Y}'_t$ with $0 < t \leq \bar{t}$, for every $G(\cdot) \in \mathcal{X}'_{t'}$ with $0 < t' \leq \bar{t}$ and for every $\mathbf{y} \in U$, as $h \rightarrow 0$, there exists a constant $C > 0$ independent of $h > 0$ and of $\mathbf{y} \in U$ such that the Galerkin approximations $G(u^h(\mathbf{y}))$ satisfy*

$$|G(u(\mathbf{y})) - G(u^h(\mathbf{y}))| \leq C h^\tau \|f\|_{\mathcal{Y}'_t} \|G\|_{\mathcal{X}'_{t'}}. \quad (2.20)$$

where $0 < \tau := t + t'$.

The result follows from a (classical) Aubin-Nitsche duality argument [23].

2.5. Dimension truncation. In order to approximate the integral (1.4) by QMC methods, we truncate the infinite sum in (1.2) to s terms, as indicated in (1.5). We denote by $u_s(\mathbf{y})$ the solution of the corresponding parametric weak problem (2.5). Then Theorem 2.1 holds when $u(\mathbf{y})$ is replaced by $u_s(\mathbf{y})$. In addition to the assumption (1.3), which implies $\sum_{j \geq 1} \beta_j^p < \infty$ with β_j defined as in (2.2), we assume that the operators A_j are enumerated so that

$$\beta_1 \geq \beta_2 \geq \dots \geq \beta_j \geq \dots. \quad (2.21)$$

THEOREM 2.6. *Under Assumption 1, for every $f \in \mathcal{Y}'$, for every $\mathbf{y} \in U$ and for every $s \in \mathbb{N}$, the solution $u_s(\mathbf{y})$ of the s -term truncated parametric weak problem (2.5) satisfies, with β_j as defined in (2.2),*

$$\|u(\mathbf{y}) - u_s(\mathbf{y})\|_{\mathcal{X}} \leq \frac{C}{\mu} \|f\|_{\mathcal{Y}'} \sum_{j \geq s+1} \beta_j \quad (2.22)$$

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

$$|I(G(u)) - I(G(u^s))| \leq \frac{\tilde{C}}{\mu} \|f\|_{\mathcal{Y}'} \|G\|_{\mathcal{X}'} \left(\sum_{j \geq s+1} \beta_j \right)^2 \quad (2.23)$$

for some constant $\tilde{C} > 0$ independent of f and G . In addition, if conditions (1.3) and (2.21) hold, then

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

The proof is a generalization of [18, Theorem 5.1].

3. Higher order QMC error analysis. Throughout this section, we consider a *general* s -variate integrand F defined over the unit cube $[0, 1]^s$, and we approximate the s -dimensional integral

$$I_s(F) := \int_{[0,1]^s} F(\mathbf{y}) \, d\mathbf{y} \quad (3.1)$$

by an N -point QMC method, i.e., an equal-weight quadrature rule of the form

$$Q_{N,s}(F) := \frac{1}{N} \sum_{n=0}^{N-1} F(\mathbf{y}_n), \quad (3.2)$$

with judiciously chosen points $\mathbf{y}_0, \dots, \mathbf{y}_{N-1} \in [0, 1]^s$. We shall always bear in mind the special case where the integrand $F(\mathbf{y}) = G(u_s^h(\mathbf{y} - \frac{1}{2}))$ is a linear functional applied to the solution of a parametric operator equation, see (1.4) and (1.5).

THEOREM 3.1 (Main Result). *Let $s \geq 1$ and $N = b^m$ for $m \geq 1$ and prime b . Let $\boldsymbol{\beta} = (\beta_j)_{j \geq 1}$ be a sequence of positive numbers, let $\boldsymbol{\beta}_s = (\beta_j)_{1 \leq j \leq s}$, and assume that*

$$\exists 0 < p \leq 1 : \sum_{j=1}^{\infty} \beta_j^p < \infty. \quad (3.3)$$

Define

$$\alpha := \lfloor 1/p \rfloor + 1. \quad (3.4)$$

If (3.3) holds with $p = 1$, we assume additionally that $\sum_{j=1}^{\infty} \beta_j$ is small as in (3.41) below. Suppose we have an integrand F whose partial derivatives satisfy

$$\forall \boldsymbol{\nu} \in \{0, 1, \dots, \alpha\}^s : |(\partial_{\mathbf{y}}^{\boldsymbol{\nu}} F)(\mathbf{y})| \leq c |\boldsymbol{\nu}|! \boldsymbol{\beta}_s^{\boldsymbol{\nu}} \quad (3.5)$$

for some constant $c > 0$. Then, an interlaced polynomial lattice rule of order α with N points can be constructed using a fast component-by-component algorithm, with cost $\mathcal{O}(\alpha s N \log N + \alpha^2 s^2 N)$ operations, such that

$$|I_s(F) - Q_{N,s}(F)| \leq C_{\alpha, \boldsymbol{\beta}, b, p} N^{-1/p},$$

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

THEOREM 3.2. *If (3.5) in Theorem 3.1 is replaced by*

$$\forall \boldsymbol{\nu} \in \{0, 1, \dots, \alpha\}^s : \sup_{\mathbf{y} \in U} |(\partial_{\mathbf{y}}^{\boldsymbol{\nu}} F)(\mathbf{y})| \leq c \boldsymbol{\nu}! \boldsymbol{\beta}_s^{\boldsymbol{\nu}}, \quad (3.6)$$

then the result still holds, but the cost of the CBC algorithm is only $\mathcal{O}(\alpha s N \log N)$ operations, and the additional condition (3.41) is not required when $p = 1$.

We shall prove these theorems in stages in the following subsections, starting by introducing a new function space setting motivated by (3.5) and (3.6). Before we proceed, we note that the ‘‘interlaced polynomial lattice rules’’ used in the theorems above, to be formally introduced in §3.2, are *deterministic*. We always have $\alpha \geq 2$, where $\alpha = 2$ is obtained with $p = 1$. This indicates that we require interlaced polynomial lattice rules of order 2 to achieve a convergence rate of N^{-1} , with the implied constant independent of the dimension.

We also remark that (2.7) in Theorem 2.2 yields an integrand $F(\mathbf{y}) = G(u_s(\mathbf{y}))$ (after dimension truncation) which satisfies (3.5), with $c = C_0 \|f\|_{\mathcal{Y}'} \|G\|_{\mathcal{X}'}$.

3.1. A new function space setting for smooth integrands. In this subsection we consider numerical integration for (3.1) for smooth integrands F of s variables using a family of QMC rules called *digital nets*, see, e.g., [21, 11]. We define in the following a class of function spaces on the unit cube in finite dimension s which contain the integrands $F(\mathbf{y}) = G(u_s^h(\mathbf{y} - \frac{1}{2}))$, that is, integrands which arise from linear functionals of solutions of the parametric operator equation (1.1).

DEFINITION 3.3 (Norm and function space). *Let $\alpha, s \in \mathbb{N}$, $1 \leq q \leq \infty$ and $1 \leq r \leq \infty$, and let $\gamma = (\gamma_{\mathbf{u}})_{\mathbf{u} \subset \mathbb{N}}$ be a collection of nonnegative real numbers, known as weights. Assume further that $F : [0, 1]^s \rightarrow \mathbb{R}$ has partial derivatives of orders up to α with respect to each variable. Then we define the norm of F by a higher order unanchored Sobolev norm*

$$\begin{aligned} & \|F\|_{s, \alpha, \gamma, q, r}^r \\ & := \sum_{\mathbf{u} \subseteq \{1:s\}} \left\| \gamma_{\mathbf{u}}^{-1} \sum_{\mathbf{v} \subseteq \mathbf{u}} \sum_{\tau_{\mathbf{u} \setminus \mathbf{v}} \in \{1:\alpha\}^{|\mathbf{u} \setminus \mathbf{v}|}} \int_{[0,1]^{s-|\mathbf{v}|}} (\partial_{\mathbf{y}}^{(\alpha_{\mathbf{v}}, \tau_{\mathbf{u} \setminus \mathbf{v}}, \mathbf{0})} F)(\mathbf{y}) \, d\mathbf{y}_{\{1:s\} \setminus \mathbf{v}} \right\|_{L_q}^r, \end{aligned} \quad (3.7)$$

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

DEFINITION 3.4 (Digital net). *Let b be prime and $\alpha, s, m \in \mathbb{N}$. Let C_1, \dots, C_s be $\alpha m \times m$ matrices over \mathbb{Z}_b ; these are known as the generating matrices. For each integer $0 \leq n < b^m$, let $n = \eta_0 + \eta_1 b + \dots + \eta_{m-1} b^{m-1}$ be the b -adic expansion of n . For each $1 \leq j \leq s$ we compute $(\zeta_1, \zeta_2, \dots, \zeta_{\alpha m})^\top = C_j (\eta_0, \eta_1, \dots, \eta_{m-1})^\top$, set $y_j^{(n)} = \frac{\zeta_1}{b} + \frac{\zeta_2}{b^2} + \dots + \frac{\zeta_{\alpha m}}{b^{\alpha m}}$ and set $\mathbf{y}_n = (y_1^{(n)}, y_2^{(n)}, \dots, y_s^{(n)})$.*

Then, the resulting point set $\mathcal{S} = \{\mathbf{y}_n\}_{n=0}^{b^m-1} \subset [0, 1]^s$ is called a digital net.

We derive an upper bound on the worst case error of a digital net in $\mathcal{W}_{s, \alpha, \gamma, q, r}$.

THEOREM 3.5 (Worst case error bound). *Let $\alpha, s \in \mathbb{N}$ with $\alpha > 1$, $1 \leq q \leq \infty$ and $1 \leq r \leq \infty$, and let $\gamma = (\gamma_{\mathbf{u}})_{\mathbf{u} \subset \mathbb{N}}$ denote a collection of weights. Let $r' \geq 1$ satisfy $1/r + 1/r' = 1$. Let b be prime, $m \in \mathbb{N}$, and let $\mathcal{S} = \{\mathbf{y}_n\}_{n=0}^{b^m-1}$ denote a digital net with generating matrices $C_1, \dots, C_s \in \mathbb{Z}_b^{\alpha m \times m}$. Then we have*

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

with

$$e_{s, \alpha, \gamma, r'}(\mathcal{S}) := \left(\sum_{\emptyset \neq \mathbf{u} \subseteq \{1:s\}} \left(C_{\alpha, b}^{|\mathbf{u}|} \gamma_{\mathbf{u}} \sum_{\mathbf{k}_{\mathbf{u}} \in \mathfrak{D}_{\mathbf{u}}^*} b^{-\mu_{\alpha}(\mathbf{k}_{\mathbf{u}})} \right)^{r'} \right)^{1/r'}. \quad (3.8)$$

Here $\mathfrak{D}_{\mathbf{u}}^*$ is the “dual net without 0 components” projected to the components in \mathbf{u} , defined by

$$\mathfrak{D}_{\mathbf{u}}^* := \left\{ \mathbf{k}_{\mathbf{u}} \in \mathbb{N}^{|\mathbf{u}|} : \sum_{j \in \mathbf{u}} C_j^\top \text{tr}_{\alpha m}(k_j) = \mathbf{0} \in \mathbb{Z}_b^m \right\}, \quad (3.9)$$

where $\text{tr}_{\alpha m}(k) := (\varkappa_0, \varkappa_1, \dots, \varkappa_{\alpha m-1})^\top$ if $k = \varkappa_0 + \varkappa_1 b + \varkappa_2 b^2 + \dots$ with $\varkappa_i \in$

$\{0, \dots, b-1\}$. Moreover, we have $\mu_\alpha(\mathbf{k}_u) = \sum_{j \in u} \mu_\alpha(k_j)$ with

$$\mu_\alpha(k) := \begin{cases} 0 & \text{if } k = 0, \\ a_1 + \dots + a_{\min(\alpha, \rho)} & \text{if } k = \kappa_1 b^{a_1-1} + \dots + \kappa_\rho b^{a_\rho-1} \text{ with} \\ & \kappa_i \in \{1, \dots, b-1\} \text{ and } a_1 > \dots > a_\rho > 0, \end{cases} \quad (3.10)$$

and

$$C_{\alpha, b} := \max \left(\frac{2}{(2 \sin \frac{\pi}{b})^\alpha}, \max_{1 \leq z \leq \alpha-1} \frac{1}{(2 \sin \frac{\pi}{b})^z} \right) \times \left(1 + \frac{1}{b} + \frac{1}{b(b+1)} \right)^{\alpha-2} \left(3 + \frac{2}{b} + \frac{2b+1}{b-1} \right). \quad (3.11)$$

Proof. Assume that $\|F\|_{s, \alpha, \gamma, q, r} < \infty$. Let B_τ denote the Bernoulli polynomial of degree τ and let $b_\tau = (\tau!)^{-1} B_\tau$. Furthermore, let \tilde{b}_τ denote the one-periodic extension of the polynomial $b_\tau : [0, 1) \rightarrow \mathbb{R}$. We claim that F can be represented by

$$F(\mathbf{y}) = \sum_{\mathbf{u} \subseteq \{1:s\}} F_{\mathbf{u}}(\mathbf{y}_{\mathbf{u}}), \quad (3.12)$$

where

$$F_{\mathbf{u}}(\mathbf{y}_{\mathbf{u}}) = \sum_{\mathbf{v} \subseteq \mathbf{u}} \sum_{\tau_{\mathbf{u} \setminus \mathbf{v}} \in \{1:\alpha\}^{|\mathbf{u} \setminus \mathbf{v}|}} \left(\prod_{j \in \mathbf{u} \setminus \mathbf{v}} b_{\tau_j}(y_j) \right) (-1)^{(\alpha+1)|\mathbf{v}|} \times \int_{[0,1]^s} (\partial_{\mathbf{x}}^{(\tau_{\mathbf{u} \setminus \mathbf{v}}, \boldsymbol{\alpha}_{\mathbf{v}}, \mathbf{0})} F)(\mathbf{x}) \prod_{j \in \mathbf{v}} \tilde{b}_\alpha(x_j - y_j) d\mathbf{x}. \quad (3.13)$$

To see this, consider the set \mathcal{P} of all polynomials defined on $[0, 1]^s$. For these functions (3.12) and (3.13) hold. The set \mathcal{P} is dense in $\mathcal{W}_{s, \alpha, \gamma, q, r}$. Let $F \in \mathcal{W}_{s, \alpha, \gamma, q, r}$. Since \mathcal{P} is dense, there exists a sequence of functions $(F_n)_{n \geq 1}$ in \mathcal{P} such that $\|F_n - F\|_{s, \alpha, \gamma, q, r} \rightarrow 0$ as $n \rightarrow \infty$. Since the set $\mathcal{W}_{s, \alpha, \gamma, q, r}$ is complete with respect to the norm $\|\cdot\|_{s, \alpha, \gamma, q, r}$, (F_n) is a Cauchy sequence. It follows that $(\partial_{\mathbf{x}}^{(\tau_{\mathbf{u} \setminus \mathbf{v}}, \boldsymbol{\alpha}_{\mathbf{v}}, \mathbf{0})} F_n)(\mathbf{x})$ is a Cauchy sequence in L_q . Let the limit of these sequences be denoted by $(\partial_{\mathbf{x}}^{(\tau_{\mathbf{u} \setminus \mathbf{v}}, \boldsymbol{\alpha}_{\mathbf{v}}, \mathbf{0})} \tilde{F})(\mathbf{x})$ and let \tilde{F} be defined via (3.12) and (3.13). Then $\|F - \tilde{F}\|_{s, \alpha, \gamma, q, r} \leq \|F - F_n\|_{s, \alpha, \gamma, q, r} + \|F_n - \tilde{F}\|_{s, \alpha, \gamma, q, r}$. Then $\|F - \tilde{F}\|_{s, \alpha, \gamma, q, r} \rightarrow 0$ by the definition of F_n and $\|F_n - \tilde{F}\|_{s, \alpha, \gamma, q, r} \rightarrow 0$ by the definition of $(\partial_{\mathbf{x}}^{(\tau_{\mathbf{u} \setminus \mathbf{v}}, \boldsymbol{\alpha}_{\mathbf{v}}, \mathbf{0})} \tilde{F})(\mathbf{x})$. Therefore the claim is shown.

Note that (3.12) and (3.13) together is the ANOVA decomposition of F , since for any nonempty \mathbf{u} we have $\int_0^1 F_{\mathbf{u}}(\mathbf{y}_{\mathbf{u}}) dy_j = 0$ whenever $j \in \mathbf{u}$. This follows from the property that $\int_0^1 b_\tau(y) dy = 0$ for all $\tau \geq 1$. Moreover, we have

$$\begin{aligned} & \|F_{\mathbf{u}}\|_{s, \alpha, \gamma, q, r} \\ &= \gamma_{\mathbf{u}}^{-1} \left(\sum_{\mathbf{v} \subseteq \mathbf{u}} \sum_{\tau_{\mathbf{u} \setminus \mathbf{v}} \in \{1:\alpha\}^{|\mathbf{u} \setminus \mathbf{v}|}} \int_{[0,1]^{|\mathbf{v}|}} \left| \int_{[0,1]^{s-|\mathbf{v}|}} (\partial_{\mathbf{y}}^{(\tau_{\mathbf{u} \setminus \mathbf{v}}, \boldsymbol{\alpha}_{\mathbf{v}}, \mathbf{0})} F_{\mathbf{u}})(\mathbf{y}) d\mathbf{y}_{\{1:s\} \setminus \mathbf{v}} \right|^q d\mathbf{y}_{\mathbf{v}} \right)^{1/q}. \end{aligned}$$

Thus we have the norm decomposition $\|F\|_{s, \alpha, \gamma, q, r} = \left(\sum_{\mathbf{u} \subseteq \{1:s\}} \|F_{\mathbf{u}}\|_{s, \alpha, \gamma, q, r}^r \right)^{1/r}$.

Let $\hat{F}(\mathbf{k})$ denote the \mathbf{k} th Walsh coefficient of F and $\hat{F}_{\mathbf{u}}(\mathbf{k}_{\mathbf{u}})$ denote the $\mathbf{k}_{\mathbf{u}}$ th Walsh coefficient of $F_{\mathbf{u}}$. (We refer to [8, 11] and the references there for more information

on Walsh function expansions.) Then

$$F(\mathbf{y}) = \sum_{\mathbf{k} \in \mathbb{N}_0^s} \widehat{F}(\mathbf{k}) \text{wal}_{\mathbf{k}}(\mathbf{y}) = \sum_{\mathbf{u} \subseteq \{1:s\}} \sum_{\mathbf{k}_{\mathbf{u}} \in \mathbb{N}^{|\mathbf{u}|}} \widehat{F}_{\mathbf{u}}(\mathbf{k}_{\mathbf{u}}) \text{wal}_{\mathbf{k}_{\mathbf{u}}}(\mathbf{y}_{\mathbf{u}}) = \sum_{\mathbf{u} \subseteq \{1:s\}} F_{\mathbf{u}}(\mathbf{y}_{\mathbf{u}}).$$

Note that $\int_0^1 b_{\tau}(y) dy = 0$ for all $\tau \geq 1$ and $\int_0^1 \text{wal}_k(y) dy = 0$ for all $k \geq 1$. Thus

$$\widehat{F}(\mathbf{k}_{\mathbf{u}}, \mathbf{0}) = \int_{[0,1]^s} F(\mathbf{y}) \overline{\text{wal}_{\mathbf{k}_{\mathbf{u}}}(\mathbf{y}_{\mathbf{u}})} d\mathbf{y}_{\mathbf{u}} = \int_{[0,1]^{|\mathbf{u}|}} F_{\mathbf{u}}(\mathbf{y}_{\mathbf{u}}) \overline{\text{wal}_{\mathbf{k}_{\mathbf{u}}}(\mathbf{y}_{\mathbf{u}})} d\mathbf{y}_{\mathbf{u}} = \widehat{F}_{\mathbf{u}}(\mathbf{k}_{\mathbf{u}}).$$

From the character property of digital nets, see, e.g., [11, Lemma 4.75], we obtain

$$\left| \frac{1}{b^m} \sum_{n=0}^{b^m-1} F(\mathbf{y}_n) - \int_{[0,1]^s} F(\mathbf{y}) d\mathbf{y} \right| \leq \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:s\}} \sum_{\mathbf{k}_{\mathbf{u}} \in \mathcal{D}_{\mathbf{u}}^*} |\widehat{F}(\mathbf{k}_{\mathbf{u}}, \mathbf{0})|. \quad (3.14)$$

We now explain how to obtain a bound on the Walsh coefficients $\widehat{F}(\mathbf{k}_{\mathbf{u}}, \mathbf{0})$. For $s = 1$, [8, Theorem 14] states that for $F \in \mathcal{W}_{1,\alpha,\gamma,q,r}$ and $k \in \mathbb{N}$ we have

$$\begin{aligned} |\widehat{F}(k)| &\leq \sum_{z=\rho}^{\alpha} \left| \int_0^1 F^{(z)}(x) dx \right| \frac{b^{-\mu_{z,\text{per}}(k)}}{(2 \sin \frac{\pi}{b})^z} \left(1 + \frac{1}{b} + \frac{1}{b(b+1)} \right)^{\max(0,z-2)} \\ &\quad + \int_0^1 |F^{(\alpha)}(x)| dx \frac{2b^{-\mu_{\alpha,\text{per}}(k)}}{(2 \sin \frac{\pi}{b})^{\alpha}} \left(1 + \frac{1}{b} + \frac{1}{b(b+1)} \right)^{\alpha-2} \left(3 + \frac{2}{b} + \frac{2b+1}{b-1} \right), \end{aligned} \quad (3.15)$$

where $\rho = 0$ if $k = 0$, otherwise ρ is given by the expansion of k in (3.10), and where

$$\mu_{z,\text{per}}(k) := \begin{cases} 0 & \text{for } z = 0 \text{ and } k \geq 0, \\ 0 & \text{for } k = 0 \text{ and } z \geq 0, \\ a_1 + \cdots + a_{\rho} + (z - \rho)a_{\rho} & \text{for } 1 \leq \rho < z, \\ a_1 + \cdots + a_{\rho} & \text{for } \rho \geq z. \end{cases}$$

Moreover, for $\rho > \alpha$ the empty sum (3.15) is defined as 0. Note that for $k \in \mathbb{N}$ with ρ nonzero digits in its base b expansion, we have for $z \geq \rho$ that $\mu_{z,\text{per}}(k) \geq \mu_{\alpha}(k)$ and therefore $b^{-\mu_{z,\text{per}}(k)} \leq b^{-\mu_{\alpha}(k)}$. Some further estimates yield

$$|\widehat{F}(k)| \leq C_{\alpha,b} b^{-\mu_{\alpha}(k)} \gamma_{\{1\}} \|F_{\{1\}}\|_{1,\alpha,\gamma,q,r}.$$

As mentioned in [8, Remark 15], this bound can be extended to $s > 1$. In this case one uses the representation (3.13). Since the proof of [8, Theorem 14] uses bounds on the Walsh coefficients of the Bernoulli polynomials, which appear in (3.13) in product form, we obtain a bound of the following product form:

$$|\widehat{F}(\mathbf{k}_{\mathbf{u}}, \mathbf{0})| \leq C_{\alpha,b}^{|\mathbf{u}|} b^{-\mu_{\alpha}(\mathbf{k}_{\mathbf{u}})} \gamma_{\mathbf{u}} \|F_{\mathbf{u}}\|_{s,\alpha,\gamma,q,r}.$$

Substituting this into (3.14) gives

$$\begin{aligned} \left| \frac{1}{b^m} \sum_{n=0}^{b^m-1} F(\mathbf{y}_n) - \int_{[0,1]^s} F(\mathbf{y}) d\mathbf{y} \right| &\leq \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:s\}} \|F_{\mathbf{u}}\|_{s,\alpha,\gamma,q,r} C_{\alpha,b}^{|\mathbf{u}|} \gamma_{\mathbf{u}} \sum_{\mathbf{k}_{\mathbf{u}} \in \mathcal{D}_{\mathbf{u}}^*} b^{-\mu_{\alpha}(\mathbf{k}_{\mathbf{u}})} \\ &\leq \left(\sum_{\mathbf{u} \subseteq \{1:s\}} \|F_{\mathbf{u}}\|_{s,\alpha,\gamma,q,r}^r \right)^{1/r} \left(\sum_{\emptyset \neq \mathbf{u} \subseteq \{1:s\}} \left(C_{\alpha,b}^{|\mathbf{u}|} \gamma_{\mathbf{u}} \sum_{\mathbf{k}_{\mathbf{u}} \in \mathcal{D}_{\mathbf{u}}^*} b^{-\mu_{\alpha}(\mathbf{k}_{\mathbf{u}})} \right)^{r'} \right)^{1/r'}, \end{aligned}$$

which yields the worst case error bound in the theorem. \square

REMARK 3.1. *Theorem 3.5 also holds for any digitally shifted digital net with any digital shift. To define the digital shift, consider $y, \sigma \in [0, 1)$ with b -adic expansion $y = y_1 b^{-1} + y_2 b^{-2} + \dots$ and $\sigma = \sigma_1 b^{-1} + \sigma_2 b^{-2} + \dots$ with the assumption that infinitely many digits are different from $b - 1$. Then we set $y \oplus \sigma = z_1 b^{-1} + z_2 b^{-2} + \dots$ where $z_i = y_i + \sigma_i \pmod{b}$. For vectors $\mathbf{y}, \boldsymbol{\sigma} \in [0, 1)^s$ we define $\mathbf{y} \oplus \boldsymbol{\sigma}$ component-wise. If $\mathbf{y}_0, \mathbf{y}_1, \dots, \mathbf{y}_{b^m-1}$ is a digital net, then we call $\mathbf{y}_0 \oplus \boldsymbol{\sigma}, \mathbf{y}_1 \oplus \boldsymbol{\sigma}, \dots, \mathbf{y}_{b^m-1} \oplus \boldsymbol{\sigma}$ a digitally shifted digital net with digital shift $\boldsymbol{\sigma} \in [0, 1)^s$.*

If we replace the points \mathbf{y}_n for $0 \leq n < b^m$ in Theorem 3.5 by $\mathbf{y}_n \oplus \boldsymbol{\sigma}$ for $0 \leq n < b^m$ for some arbitrary vector $\boldsymbol{\sigma} \in [0, 1)^s$, then the statement of the theorem still holds. The only change in the proof is in (3.14). There we have

$$\begin{aligned} & \left| \frac{1}{b^m} \sum_{n=0}^{b^m-1} F(\mathbf{y}_n \oplus \boldsymbol{\sigma}) - \int_{[0,1]^s} F(\mathbf{y}) \, d\mathbf{y} \right| \leq \left| \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:s\}} \sum_{\mathbf{k}_u \in \mathcal{D}_u^*} \widehat{F}(\mathbf{k}_u, \mathbf{0}) \text{wal}_{(\mathbf{k}_u, \mathbf{0})}(\boldsymbol{\sigma}) \right| \\ & \leq \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:s\}} \sum_{\mathbf{k}_u \in \mathcal{D}_u^*} |\widehat{F}(\mathbf{k}_u, \mathbf{0})| |\text{wal}_{(\mathbf{k}_u, \mathbf{0})}(\boldsymbol{\sigma})|. \end{aligned} \quad (3.16)$$

Since $|\text{wal}_{(\mathbf{k}_u, \mathbf{0})}(\boldsymbol{\sigma})| = 1$, (3.16) coincides with the right-hand side of (3.14). The remaining part of the proof stays unchanged.

Note that we require $\alpha > 1$ in Theorem 3.5 to ensure the convergence of the expression $\sum_{\mathbf{k}_u \in \mathcal{D}_u^*} b^{-\mu_\alpha(\mathbf{k}_u)}$. This expression does not converge for $\alpha = 1$.

The quantity $e_{s,\alpha,\gamma,r'}(\mathcal{S})$ can be used as an error criterion to obtain good digital nets for our function space setting. It is a generalization of the criterion used in [1, 2]: there the weights are of a product form $\gamma_u = \prod_{j \in u} \gamma_j$ for some nonnegative sequence $(\gamma_j)_{j \geq 1}$, whereas here the weights take a general form.

We stress that this quantity $e_{s,\alpha,\gamma,r'}(\mathcal{S})$ does not depend on the parameter q . Furthermore, it is more convenient to work with an upper bound which can be obtained by taking $r = \infty$ and $r' = 1$, i.e., $e_{s,\alpha,\gamma,r'}(\mathcal{S}) \leq e_{s,\alpha,\gamma,1}(\mathcal{S})$. On the other hand, for any q and r we have $\|F\|_{s,\alpha,\gamma,q,r} \geq \|F\|_{s,\alpha,\gamma,q,\infty}$. Hence by restricting ourselves to the case $r = \infty$, we are working with the larger quantity $e_{s,\alpha,\gamma,1}(\mathcal{S})$, but we benefit from having a smaller norm $\|F\|_{s,\alpha,\gamma,q,\infty}$. This is the main reason why we are able to obtain an improved convergence rate compared with other papers.

SPOD weights. For a function F satisfying (3.5), its norm, with $r = \infty$ and any q , can be bounded by

$$\begin{aligned} \|F\|_{s,\alpha,\gamma,q,\infty} & \leq \max_{\mathbf{u} \subseteq \{1:s\}} \gamma_u^{-1} \sum_{\mathbf{v} \subseteq \mathbf{u}} \sum_{\boldsymbol{\tau}_{\mathbf{u} \setminus \mathbf{v}} \in \{1:\alpha\}^{|\mathbf{u} \setminus \mathbf{v}|}} |(\boldsymbol{\alpha}_{\mathbf{v}}, \boldsymbol{\tau}_{\mathbf{u} \setminus \mathbf{v}}, \mathbf{0})!| \beta_s^{(\boldsymbol{\alpha}_{\mathbf{v}}, \boldsymbol{\tau}_{\mathbf{u} \setminus \mathbf{v}}, \mathbf{0})} \\ & = \max_{\mathbf{u} \subseteq \{1:s\}} \gamma_u^{-1} \sum_{\boldsymbol{\nu}_{\mathbf{u}} \in \{1:\alpha\}^{|\mathbf{u}|}} |\boldsymbol{\nu}_{\mathbf{u}}!| \prod_{j \in \mathbf{u}} \left(2^{\delta(\nu_j, \alpha)} \beta_j^{\nu_j} \right), \end{aligned}$$

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

$$\gamma_u := \sum_{\boldsymbol{\nu}_{\mathbf{u}} \in \{1:\alpha\}^{|\mathbf{u}|}} |\boldsymbol{\nu}_{\mathbf{u}}!| \prod_{j \in \mathbf{u}} \left(2^{\delta(\nu_j, \alpha)} \beta_j^{\nu_j} \right). \quad (3.17)$$

We shall refer to our new form of weights (3.17) as “smoothness-driven product and order dependent weights”, or “SPOD weights” for short. This new form of weights has similar characteristics to POD weights—product and order dependent weights, which were first introduced in [18] in the analysis of QMC methods for PDEs with random coefficients.

Product weights. In a similar way, we deduce that for a function F satisfying (3.6) we should choose

$$\gamma_{\mathbf{u}} := \sum_{\nu_{\mathbf{u}} \in \{1:\alpha\}^{|\mathbf{u}|}} \nu_{\mathbf{u}}! \prod_{j \in \mathbf{u}} \left(2^{\delta(\nu_j, \alpha)} \beta_j^{\nu_j} \right) = \prod_{j \in \mathbf{u}} \sum_{\nu=1}^{\alpha} \left(\nu! 2^{\delta(\nu, \alpha)} \beta_j^{\nu} \right), \quad (3.18)$$

which is of product form.

Since the quantity $e_{s, \alpha, \gamma, r'}(\mathcal{S})$ generalizes the criterion used in [1, 2], the results in [1, 2] can potentially be adapted to show that a “higher order polynomial lattice rule” (a kind of digital net), can be constructed using a component-by-component (CBC) algorithm to achieve the convergence rate, with $N = b^m$,

$$e_{s, \alpha, \gamma, r'}(\mathcal{S}) = \mathcal{O}(N^{-\tau}) \quad \text{for any } \tau \in [1, \alpha), \quad (3.19)$$

where the implied constant depends on τ , γ and s . This approach works for product weights, but for general weights this should be understood only as an *existence* result, since the CBC construction for general weights is prohibitively expensive. Furthermore, the construction for such a rule of order α has a cost which scales with N^α , see [2], making it harder to obtain higher order rules. We do not take this approach, but use an analogous approach to [12, 13].

3.2. Interlaced polynomial lattice rules. In this subsection we formally introduce (interlaced) polynomial lattice rules. Polynomial lattice rules were first introduced by Niederreiter, see [21].

DEFINITION 3.6 (Polynomial lattice rules). *For a prime b and any $m \in \mathbb{N}$, let $P \in \mathbb{Z}_b[x]$ be an irreducible polynomial with $\deg(P) = m$; this is known as the modulus. For a given dimension $s \geq 1$, select s polynomials $q_1(x), \dots, q_s(x)$ from the set*

$$G_{b,m} := \{q(x) \in \mathbb{Z}_b[x] \setminus \{0\} : \deg(q) < m\}, \quad (3.20)$$

and write collectively

$$\mathbf{q} = \mathbf{q}(x) = (q_1(x), \dots, q_s(x)) \in G_{b,m}^s; \quad (3.21)$$

this is known as the generating vector. For each integer $0 \leq n < b^m$, let $n = \eta_0 + \eta_1 b + \dots + \eta_{m-1} b^{m-1}$ be the b -adic expansion of n , and we associate with n the polynomial

$$n(x) = \sum_{r=0}^{m-1} \eta_r x^r \in \mathbb{Z}_b[x].$$

Furthermore, we denote by v_m the map from $\mathbb{Z}_b((x^{-1}))$ to the interval $[0, 1)$ defined for any integer w by

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

Then, the QMC point set $\mathcal{S}_{P,b,m,s}(\mathbf{q})$ of a (classical) polynomial lattice rule comprises the points

$$\mathbf{y}_n = \left(v_m \left(\frac{n(x)q_1(x)}{P(x)} \right), \dots, v_m \left(\frac{n(x)q_s(x)}{P(x)} \right) \right) \in [0, 1)^s, \quad n = 0, \dots, b^m - 1.$$

In the following we define interlaced polynomial lattice rules [12, 13], belonging to the family of higher order digital nets, which were first introduced in [6, 7].

DEFINITION 3.7 (Interlaced polynomial lattice rules). *Define the digit interlacing function with interlacing factor $\alpha \in \mathbb{N}$ by*

$$\begin{aligned} \mathcal{D}_\alpha : [0, 1)^\alpha &\rightarrow [0, 1) \\ (x_1, \dots, x_\alpha) &\mapsto \sum_{a=1}^{\infty} \sum_{j=1}^{\alpha} \xi_{j,a} b^{-j-(a-1)\alpha}, \end{aligned} \quad (3.22)$$

where $x_j = \xi_{j,1}b^{-1} + \xi_{j,2}b^{-2} + \dots$ for $1 \leq j \leq \alpha$. We also define such a function for vectors by setting

$$\begin{aligned} \mathcal{D}_\alpha : [0, 1)^{\alpha s} &\rightarrow [0, 1)^s \\ (x_1, \dots, x_{\alpha s}) &\mapsto (\mathcal{D}_\alpha(x_1, \dots, x_\alpha), \dots, \mathcal{D}_\alpha(x_{(s-1)\alpha+1}, \dots, x_{s\alpha})). \end{aligned} \quad (3.23)$$

Then, an interlaced polynomial lattice rule of order α with b^m points in s dimensions is a QMC rule using $\mathcal{D}_\alpha(\mathcal{S}_{P,b,m,\alpha s}(\mathbf{q}))$ as quadrature points, for some given modulus P and generating vector $\mathbf{q} \in G_{b,m}^{\alpha s}$.

Note that the generating vector \mathbf{q} is of length αs , and it can be interpreted as the concatenation of α different generating vectors of (classical) polynomial lattice rules in s dimensions. Interlaced scrambled polynomial lattice rules were first used in [13] and interlaced polynomial lattice rules were first used in [12].

An illustration of how interlacing works is as follows: if we have α numbers in base b representation $(0.\xi_{1,1}\xi_{1,2}\xi_{1,3}\dots)_b, (0.\xi_{2,1}\xi_{2,2}\xi_{2,3}\dots)_b, \dots, (0.\xi_{\alpha,1}\xi_{\alpha,2}\xi_{\alpha,3}\dots)_b$, then the result of interlacing is $(0.\xi_{1,1}\xi_{2,1}\dots\xi_{\alpha,1}\xi_{1,2}\xi_{2,2}\dots\xi_{\alpha,2}\xi_{1,3}\xi_{2,3}\dots\xi_{\alpha,3}\dots)_b$, that is, we take the first digit of all the α numbers, followed by the second digit of all the α numbers, and then the third digit, and so on.

REMARK 3.2 (Higher order polynomial lattice rules). *Higher order polynomial lattice rules were first introduced in [10]. It is important to note that they are not obtained by interlacing (classical) polynomial lattice rules. A higher-order polynomial lattice rule of order α differs from Definition 3.6 in two ways: the modulus P has a higher degree $\deg(P) = \alpha m$, and the set of polynomials (3.20) now include all polynomials with $\deg(q) < \alpha m$. An interlaced polynomial lattice rule as defined in Definition 3.7 is not a higher order polynomial lattice rule.*

To proceed with our analysis, we note that the worst case error bound in Theorem 3.5 holds for interlaced polynomial lattice rules, but the dual net (3.9) is expressed in terms of the generating matrices, which is inconvenient for our analysis or computation. We now derive an alternative expression for $e_{s,\alpha,\gamma,r'}(\mathcal{S})$ in (3.8) which is expressed in terms of the generating polynomials.

We start by extending the definition of the interlacing function \mathcal{D}_α to nonnegative integers by setting

$$\begin{aligned} \mathcal{E}_\alpha : \mathbb{N}_0^\alpha &\rightarrow \mathbb{N}_0 \\ (\ell_1, \dots, \ell_\alpha) &\mapsto \sum_{a=0}^{\infty} \sum_{j=1}^{\alpha} \ell_{j,a} b^{j-1+a\alpha}, \end{aligned} \quad (3.24)$$

where $\ell_j = \ell_{j,0} + \ell_{j,1}b + \ell_{j,2}b^2 + \dots$ for $1 \leq j \leq \alpha$. We also extend this function to vectors via

$$\begin{aligned} \mathcal{E}_\alpha : \mathbb{N}_0^{\alpha s} &\rightarrow \mathbb{N}_0^s \\ (\ell_1, \dots, \ell_{\alpha s}) &\mapsto (\mathcal{E}_\alpha(\ell_1, \dots, \ell_\alpha), \dots, \mathcal{E}_\alpha(\ell_{(s-1)\alpha+1}, \dots, \ell_{\alpha s})). \end{aligned} \quad (3.25)$$

For a given set $\emptyset \neq \mathbf{v} \subseteq \{1 : \alpha s\}$, we define

$$\mathbf{u}(\mathbf{v}) := \{\lfloor j/\alpha \rfloor : j \in \mathbf{v}\} \subseteq \{1 : s\}, \quad (3.26)$$

where each element appears only once as is typical for sets. The set $\mathbf{u}(\mathbf{v})$ can be viewed as an indicator on whether the set \mathbf{v} includes any element from each block of α components from $\{1 : \alpha s\}$.

The dual net can be obtained by interlacing the dual nets corresponding to the α different (classical) polynomial lattice rules. By [11, Lemma 10.6] and the definition (3.25) of the interlacing function, we can rewrite (3.8) for an interlaced polynomial lattice rule with $r' = 1$ as

$$e_{s,\alpha,\gamma,1}(\mathcal{S}) = \sum_{\emptyset \neq \mathbf{v} \subseteq \{1:\alpha s\}} C_{\alpha,b}^{|\mathbf{u}(\mathbf{v})|} \gamma_{\mathbf{u}(\mathbf{v})} \sum_{\ell_{\mathbf{v}} \in \mathcal{D}_{\mathbf{v}}^*} b^{-\mu_{\alpha}(\mathcal{E}_{\alpha}(\ell_{\mathbf{v}}, \mathbf{0}))}, \quad (3.27)$$

where $(\ell_{\mathbf{v}}, \mathbf{0})$ denotes a vector of length αs whose j th component is ℓ_j if $j \in \mathbf{v}$ and 0 if $j \notin \mathbf{v}$, and the ‘‘dual net without 0 components’’ is now defined in terms of the generating polynomials as

$$\mathcal{D}_{\mathbf{v}}^* := \{\ell_{\mathbf{v}} \in \mathbb{N}^{|\mathbf{v}|} : \text{tr}_m(\ell_{\mathbf{v}}) \cdot \mathbf{q}_{\mathbf{v}} \equiv 0 \pmod{P}\}. \quad (3.28)$$

The ‘‘inner-product’’ in (3.28) denotes

$$\text{tr}_m(\mathbf{k}_{\mathbf{v}}) \cdot \mathbf{q}_{\mathbf{v}} = \sum_{j \in \mathbf{v}} \text{tr}_m(k_j)(x) q_j(x) \in \mathbb{Z}_b[x],$$

with the polynomial $\text{tr}_m(k)(x)$ defined as follows: to any nonnegative integer k with b -adic expansion $k = \kappa_0 + \kappa_1 b + \dots + \kappa_{\rho-1} b^{\rho-1}$, we associate a unique polynomial $k(x) = \kappa_0 + \kappa_1 x + \dots + \kappa_{\rho-1} x^{\rho-1}$ and its truncated version

$$\text{tr}_m(k)(x) = \kappa_0 + \kappa_1 x + \dots + \kappa_{m-1} x^{m-1},$$

which is obtained by setting $\kappa_{\rho} = \dots = \kappa_{m-1} = 0$ if $\rho < m$.

3.3. Component-by-component construction. In this subsection we introduce and analyze a component-by-component (CBC) construction of interlaced polynomial lattice rules. The expression (3.27) could be used as our search criterion. However, to reduce the computational cost from a scaling of N^{α} to αN , we shall instead work with an upper bound to (3.27), which is based on the following lemma.

LEMMA 3.8. *For any $z_1, \dots, z_{\alpha} \in \mathbb{N}_0$ we have*

$$\mu_{\alpha}(\mathcal{E}_{\alpha}(z_1, \dots, z_{\alpha})) \geq \alpha \sum_{j=1}^{\alpha} \mu_1(z_j) - \frac{\alpha(\alpha-1)}{2}. \quad (3.29)$$

Proof. For $\alpha = 1$ we obviously have equality. For $\alpha \geq 2$, we define z'_j by retaining only the most significant base b digit in z_j , and we set $z'_j = 0$ if $z_j = 0$. Then we have $\mu_1(z'_j) = \mu_1(z_j)$ and $z'_j \neq 0 \Leftrightarrow z_j \neq 0$, and thus

$$\begin{aligned} \mu_{\alpha}(\mathcal{E}_{\alpha}(z_1, \dots, z_{\alpha})) &\geq \mu_{\alpha}(\mathcal{E}_{\alpha}(z'_1, \dots, z'_{\alpha})) = \sum_{\substack{j=1 \\ z'_j \neq 0}}^{\alpha} (\alpha(\mu_1(z_j) - 1) + j) \\ &= \alpha \sum_{j=1}^{\alpha} \mu_1(z_j) - \sum_{\substack{j=1 \\ z_j \neq 0}}^{\alpha} (\alpha - j) \geq \alpha \sum_{j=1}^{\alpha} \mu_1(z_j) - \sum_{j=1}^{\alpha} (\alpha - j), \end{aligned}$$

which proves the result. \square

We now apply (3.29) to the vector $(\ell_{\mathbf{v}}, \mathbf{0})$ in blocks of α components, noting that $\mu_{\alpha}(\mathcal{E}_{\alpha}(\mathbf{0})) = 0$, to obtain

$$\mu_{\alpha}(\mathcal{E}_{\alpha}(\ell_{\mathbf{v}}, \mathbf{0})) \geq \alpha\mu_1(\ell_{\mathbf{v}}) - \frac{\alpha(\alpha-1)}{2}|\mathbf{u}(\mathbf{v})|.$$

Substituting this into (3.27) then yields the upper bound

$$e_{s,\alpha,\gamma,1}(\mathcal{S}) \leq \sum_{\emptyset \neq \mathbf{v} \subseteq \{1:\alpha s\}} C_{\alpha,b}^{|\mathbf{u}(\mathbf{v})|} \gamma_{\mathbf{u}(\mathbf{v})} b^{\alpha(\alpha-1)|\mathbf{u}(\mathbf{v})/2} \sum_{\ell_{\mathbf{v}} \in \mathcal{D}_{\mathbf{v}}^*} b^{-\alpha\mu_1(\ell_{\mathbf{v}})}. \quad (3.30)$$

We shall use the right-hand side of (3.30) as our search criterion in the CBC construction. To simplify our notation, we define

$$E_d(\mathbf{q}) := \sum_{\emptyset \neq \mathbf{v} \subseteq \{1:d\}} \tilde{\gamma}_{\mathbf{v}} \sum_{\ell_{\mathbf{v}} \in \mathcal{D}_{\mathbf{v}}^*} b^{-\alpha\mu_1(\ell_{\mathbf{v}})}. \quad (3.31)$$

In particular, we are interested in the case $d = \alpha s$ and weights

$$\tilde{\gamma}_{\mathbf{v}} := C_{\alpha,b}^{|\mathbf{u}(\mathbf{v})|} \gamma_{\mathbf{u}(\mathbf{v})} b^{\alpha(\alpha-1)|\mathbf{u}(\mathbf{v})/2}. \quad (3.32)$$

However, the theorem below holds for any d and general weights $\tilde{\gamma}_{\mathbf{v}}$.

THEOREM 3.9 (CBC error bound). *Let $b \geq 2$ be prime, and $\alpha \geq 2$ and $m, d \geq 1$ be integers, and let $P \in \mathbb{Z}_b[x]$ be an irreducible polynomial with $\deg(P) = m$. Let $(\tilde{\gamma}_{\mathbf{v}})_{\mathbf{v} \subseteq \{1:d\}}$ be positive real numbers. Then a generating vector $\mathbf{q}^* = (1, q_2^*, \dots, q_d^*) \in G_{b,m}^d$ can be constructed using a component-by-component approach, minimizing $E_d(\mathbf{q})$ in each step, such that*

$$E_d(\mathbf{q}^*) \leq \left(\frac{2}{b^m - 1} \sum_{\emptyset \neq \mathbf{v} \subseteq \{1:d\}} \tilde{\gamma}_{\mathbf{v}}^{\lambda} \left(\frac{b-1}{b^{\alpha\lambda} - b} \right)^{|\mathbf{v}|} \right)^{1/\lambda} \quad \text{for all } \lambda \in (1/\alpha, 1]. \quad (3.33)$$

Proof. We prove the result by induction. For $d = 1$, we calculate

$$E_1(1) = \tilde{\gamma}_{\{1\}} \sum_{\ell=1}^{\infty} b^{-\alpha\mu_1(\ell b^m)} = \tilde{\gamma}_{\{1\}} \sum_{a=0}^{\infty} b^{-\alpha(m+a+1)} (b-1)b^a = \tilde{\gamma}_{\{1\}} b^{-\alpha m} \frac{b-1}{b^{\alpha} - b}.$$

Thus the result holds for $d = 1$.

Suppose now that (3.33) holds for some vector $\mathbf{q}^* \in G_{b,m}^d$ for some $d \geq 1$. For $\mathbf{k} \in \mathbb{N}_0^d$ we denote the support of the multi-index \mathbf{k} by $\mathbf{v}(\mathbf{k}) := \{1 \leq j \leq d : k_j > 0\}$. Then we write (3.31), with d replaced by $d+1$, in an equivalent formulation

$$\begin{aligned} E_{d+1}(\mathbf{q}^*, q_{d+1}) &= \sum_{\substack{(\mathbf{k}, k_{d+1}) \in \mathbb{N}_0^{d+1} \setminus \{\mathbf{0}\} \\ \text{tr}_m(\mathbf{k}, k_{d+1}) \cdot (\mathbf{q}^*, q_{d+1}) \equiv 0 \pmod{P}}} \tilde{\gamma}_{\mathbf{v}(\mathbf{k}, k_{d+1})} b^{-\alpha\mu_1(\mathbf{k}, k_{d+1})} \\ &= E_d(\mathbf{q}^*) + \theta(q_{d+1}), \end{aligned} \quad (3.34)$$

where we have separated out the $k_{d+1} = 0$ terms so that

$$\theta(q_{d+1}) = \sum_{k_{d+1}=1}^{\infty} \left(b^{-\alpha\mu_1(k_{d+1})} \sum_{\substack{\mathbf{k} \in \mathbb{N}_0^d \\ \text{tr}_m(\mathbf{k}) \cdot \mathbf{q} \equiv -\text{tr}_m(k_{d+1}) \cdot q_{d+1} \pmod{P}}} \tilde{\gamma}_{\mathbf{v}(\mathbf{k}, k_{d+1})} b^{-\alpha\mu_1(\mathbf{k})} \right).$$

By the induction assumption, we may assume that a minimizer $\mathbf{q}^* = (q_1^*, \dots, q_d^*)$ of $E_d(\mathbf{q}^*)$ in (3.34) has already been determined. Then, the CBC algorithm chooses q_{d+1}^* such that $E_{d+1}(\mathbf{q}^*, q_{d+1})$ is minimized. By (3.34), the only dependency on q_{d+1} of $E_{d+1}(\mathbf{q}^*, q_{d+1})$ enters via $\theta(q_{d+1})$. Therefore, we conclude $\theta(q_{d+1}^*) \leq \theta(q_{d+1})$ for all $q_{d+1} \in G_{b,m}$, which implies that for any $\lambda \in (0, 1]$ we have $[\theta(q_{d+1}^*)]^\lambda \leq [\theta(q_{d+1})]^\lambda$ for all $q_{d+1} \in G_{b,m}$. This leads to

$$\theta(q_{d+1}^*) \leq \left(\frac{1}{b^m - 1} \sum_{q_{d+1} \in G_{b,m}} [\theta(q_{d+1})]^\lambda \right)^{1/\lambda}. \quad (3.35)$$

We will obtain a bound on $\theta(q_{d+1}^*)$ through this last inequality.

Let $\lambda \in (1/\alpha, 1]$. It follows from Jensen's inequality² that

$$[\theta(q_{d+1})]^\lambda \leq \sum_{k_{d+1}=1}^{\infty} b^{-\alpha\lambda\mu_1(k_{d+1})} \sum_{\substack{\mathbf{k} \in \mathbb{N}_0^d \\ \text{tr}_m(\mathbf{k}) \cdot \mathbf{q} \equiv -\text{tr}_m(k_{d+1}) \cdot q_{d+1} \pmod{P}}} \tilde{\gamma}_{\mathbf{v}(\mathbf{k}, k_{d+1})}^\lambda b^{-\alpha\lambda\mu_1(\mathbf{k})}.$$

If k_{d+1} is a multiple of b^m , then $\text{tr}_m(k_{d+1}) = 0$ and the corresponding term in the sum is independent of q_{d+1} . If k_{d+1} is not a multiple of b^m , then $\text{tr}_m(k_{d+1})$ can have any value between 1 and $b^m - 1$. Moreover, since $q_{d+1} \neq 0$ and P is irreducible, $\text{tr}_m(k_{d+1}) \cdot q_{d+1}$ is never a multiple of P . Thus

$$\begin{aligned} \frac{1}{b^m - 1} \sum_{q_{d+1} \in G_{b,m}} [\theta(q_{d+1})]^\lambda &\leq \sum_{\substack{k_{d+1}=1 \\ b^m | k_{d+1}}}^{\infty} b^{-\alpha\lambda\mu_1(k_{d+1})} \sum_{\substack{\mathbf{k} \in \mathbb{N}_0^d \\ \text{tr}_m(\mathbf{k}) \cdot \mathbf{q} \equiv 0 \pmod{P}}} \tilde{\gamma}_{\mathbf{v}(\mathbf{k}, k_{d+1})}^\lambda b^{-\alpha\lambda\mu_1(\mathbf{k})} \\ &\quad + \frac{1}{b^m - 1} \sum_{\substack{k_{d+1}=1 \\ b^m \nmid k_{d+1}}}^{\infty} b^{-\alpha\lambda\mu_1(k_{d+1})} \sum_{\substack{\mathbf{k} \in \mathbb{N}_0^d \\ \text{tr}_m(\mathbf{k}) \cdot \mathbf{q} \not\equiv 0 \pmod{P}}} \tilde{\gamma}_{\mathbf{v}(\mathbf{k}, k_{d+1})}^\lambda b^{-\alpha\lambda\mu_1(\mathbf{k})} \\ &\leq \left(b^{-\alpha\lambda m} + \frac{1}{b^m - 1} \right) \sum_{d+1 \in \mathbf{v} \subseteq \{1:d+1\}} \tilde{\gamma}_{\mathbf{v}}^\lambda \left(\frac{b-1}{b^{\alpha\lambda} - b} \right)^{|\mathbf{v}|}, \end{aligned} \quad (3.36)$$

where we used the following estimates

$$\begin{aligned} \sum_{\substack{k_{d+1}=1 \\ b^m | k_{d+1}}}^{\infty} b^{-\alpha\lambda\mu_1(k_{d+1})} &= b^{-\alpha\lambda(m+1)} \frac{b^{\alpha\lambda}(b-1)}{b^{\alpha\lambda} - b}, \quad \sum_{\substack{k_{d+1}=1 \\ b^m \nmid k_{d+1}}}^{\infty} b^{-\alpha\lambda\mu_1(k_{d+1})} \leq \frac{b-1}{b^{\alpha\lambda} - b}, \\ \sum_{\substack{\mathbf{k} \in \mathbb{N}_0^d \\ \text{tr}_m(\mathbf{k}) \cdot \mathbf{q} \not\equiv 0 \pmod{P}}} \tilde{\gamma}_{\mathbf{v}(\mathbf{k})}^\lambda b^{-\alpha\lambda\mu_1(\mathbf{k})} &\leq \sum_{\mathbf{v} \subseteq \{1:s\}} \tilde{\gamma}_{\mathbf{v}}^\lambda \left(\frac{b-1}{b^{\alpha\lambda} - b} \right)^{|\mathbf{v}|}. \end{aligned}$$

Hence we have from (3.35) and (3.36) that

$$\theta(q_{d+1}^*) \leq \left(\frac{2}{b^m - 1} \sum_{d+1 \in \mathbf{v} \subseteq \{1:d+1\}} \tilde{\gamma}_{\mathbf{v}}^\lambda \left(\frac{b-1}{b^{\alpha\lambda} - b} \right)^{|\mathbf{v}|} \right)^{1/\lambda},$$

² $(\sum_k a_k)^\lambda \leq \sum_k a_k^\lambda$ for $0 < \lambda \leq 1$ and $a_k \geq 0$.

which, together with (3.33) and (3.34), yields the required estimate for $E_{d+1}(\mathbf{q}^*, \mathbf{q}_{d+1}^*)$, that is, (3.33) with d replaced by $d+1$. This completes the proof. \square

Theorem 3.9 states that an interlaced polynomial lattice rule with interlacing factor α in s dimensions can be constructed by the CBC algorithm, with weights (3.32), such that

$$\begin{aligned} e_{\alpha, \gamma, s, 1}(\mathcal{S}) &\leq E_{\alpha s}(\mathbf{q}^*) \\ &\leq \left(\frac{2}{b^m - 1} \sum_{\emptyset \neq \mathbf{v} \subseteq \{1:s\}} \left(C_{\alpha, b}^{|\mathbf{u}(\mathbf{v})|} \gamma_{\mathbf{u}(\mathbf{v})} b^{\alpha(\alpha-1)|\mathbf{u}(\mathbf{v})/2} \right)^\lambda \left(\frac{b-1}{b^{\alpha\lambda} - b} \right)^{|\mathbf{v}|} \right)^{1/\lambda} \\ &= \left(\frac{2}{b^m - 1} \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:s\}} \left(C_{\alpha, b}^{|\mathbf{u}|} \gamma_{\mathbf{u}} b^{\alpha(\alpha-1)|\mathbf{u}|/2} \right)^\lambda \left(\left(1 + \frac{b-1}{b^{\alpha\lambda} - b} \right)^\alpha - 1 \right)^{|\mathbf{u}|} \right)^{1/\lambda}. \end{aligned}$$

In the following we discuss the two specific choices of weights (3.17) and (3.18).

SPOD weights. Substituting in our choice of $\gamma_{\mathbf{u}}$ from (3.17) and applying Jensen's inequality, we obtain

$$\begin{aligned} e_{\alpha, \gamma, s, 1}(\mathcal{S}) &\leq \left(\frac{2}{b^m - 1} \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:s\}} \sum_{\nu_{\mathbf{u}} \in \{1:\alpha\}^{|\mathbf{u}|}} (|\nu_{\mathbf{u}}|!)^\lambda \prod_{j \in \mathbf{u}} \left(B 2^{\delta(\nu_j, \alpha)} \beta_j^{\nu_j} \right)^\lambda \right)^{1/\lambda} \\ &= \left(\frac{2}{b^m - 1} \sum_{\mathbf{0} \neq \nu \in \{0:\alpha\}^s} (|\nu|!)^\lambda \prod_{\substack{j=1 \\ \nu_j > 0}}^s \left(B 2^{\delta(\nu_j, \alpha)} \beta_j^{\nu_j} \right)^\lambda \right)^{1/\lambda}, \end{aligned} \quad (3.37)$$

where

$$B := C_{\alpha, b} b^{\alpha(\alpha-1)/2} \left(\left(1 + \frac{b-1}{b^{\alpha\lambda} - b} \right)^\alpha - 1 \right)^{1/\lambda}. \quad (3.38)$$

We now choose λ to ensure that the sum in (3.37) is bounded independently of s . Let $\tilde{\beta}_j := 2 \max(B, 1) \beta_j$. Then the sum in (3.37) is bounded by

$$\sum_{\mathbf{0} \neq \nu \in \{0:\alpha\}^s} \left(|\nu|! \prod_{j=1}^s \tilde{\beta}_j^{\nu_j} \right)^\lambda,$$

where each term in the sum to be raised to the power of λ is of the form

$$(\nu_1 + \nu_2 + \cdots + \nu_s)! \underbrace{\tilde{\beta}_1 \cdots \tilde{\beta}_1}_{\nu_1} \underbrace{\tilde{\beta}_2 \cdots \tilde{\beta}_2}_{\nu_2} \cdots \underbrace{\tilde{\beta}_s \cdots \tilde{\beta}_s}_{\nu_s}. \quad (3.39)$$

We now define a sequence $d_j := \tilde{\beta}_{\lceil j/\alpha \rceil}$ so that $d_1 = \cdots = d_\alpha = \tilde{\beta}_1$ and $d_{\alpha+1} = \cdots = d_{2\alpha} = \tilde{\beta}_2$, and so on. Then any term of the form (3.39) can be written as $|\mathbf{v}|! \prod_{j \in \mathbf{v}} d_j$ for some subset of indices $\mathbf{v} \subset \mathbb{N}$. Thus we conclude that

$$\begin{aligned} \sum_{\mathbf{0} \neq \nu \in \{0:\alpha\}^s} \left(|\nu|! \prod_{\substack{j=1 \\ \nu_j > 0}}^s \tilde{\beta}_j^{\nu_j} \right)^\lambda &< \sum_{\substack{\mathbf{v} \subset \mathbb{N} \\ |\mathbf{v}| < \infty}} \left(|\mathbf{v}|! \prod_{j \in \mathbf{v}} d_j \right)^\lambda \\ &= \sum_{\ell=0}^{\infty} (\ell!)^\lambda \sum_{\substack{\mathbf{v} \subset \mathbb{N} \\ |\mathbf{v}| = \ell}} \prod_{j \in \mathbf{v}} d_j^\lambda \leq \sum_{\ell=0}^{\infty} (\ell!)^{\lambda-1} \left(\sum_{j=1}^{\infty} d_j^\lambda \right)^\ell. \end{aligned} \quad (3.40)$$

Note that $\sum_{j=1}^{\infty} \beta_j^p < \infty$ holds if and only if $\sum_{j=1}^{\infty} d_j^p < \infty$. By the ratio test, the last expression in (3.40) is finite if $\lambda \leq p$ and $\lambda < 1$. Alternatively, using the geometric series formula, the last expression in (3.40) is finite if $\lambda = 1$ and $\sum_{j=1}^{\infty} d_j < 1$. Recall that λ also needs to satisfy $1/\alpha < \lambda \leq 1$. Hence we take

$$\lambda = p \quad \text{and} \quad \alpha = \lfloor 1/p \rfloor + 1,$$

and for $p = 1$ we assume additionally that $\sum_{j=1}^{\infty} d_j < 1$, which is equivalent to

$$\sum_{j=1}^{\infty} \beta_j < \frac{1}{2\alpha \max(B, 1)} = \left(4 \max \left(\frac{2}{(2 \sin \frac{\pi}{b})^2}, \frac{1}{2 \sin \frac{\pi}{b}} \right) \left(3 + \frac{2}{b} + \frac{2b+1}{b-1} \right) \left(2 + \frac{1}{b} \right) \right)^{-1}. \quad (3.41)$$

Thus we obtain $\mathcal{O}(N^{-1/p})$ convergence, with the implied constant independent of s .

Product weights. With the product weights given by (3.18), we obtain

$$\begin{aligned} e_{\alpha, \gamma, s, 1}(\mathcal{S}) &\leq \left(\frac{2}{b^m - 1} \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:s\}} \prod_{j \in \mathbf{u}} \left(B \sum_{\nu=1}^{\alpha} \nu! 2^{\delta(\nu, \alpha)} \beta_j^{\nu} \right)^{\lambda} \right)^{1/\lambda} \\ &\leq \left(\frac{2}{b^m - 1} \exp \left(\sum_{j=1}^s \left(B \sum_{\nu=1}^{\alpha} \nu! 2^{\delta(\nu, \alpha)} \beta_j^{\nu} \right)^{\lambda} \right) \right)^{1/\lambda}, \end{aligned}$$

which is bounded independently of s if $\sum_{j=1}^{\infty} \beta_j^{\lambda} < \infty$, where B is as defined in (3.38). Hence we take again

$$\lambda = p \quad \text{and} \quad \alpha = \lfloor 1/p \rfloor + 1,$$

to obtain the convergence rate of $\mathcal{O}(N^{-1/p})$, with the implied constant independent of s . Note that the condition (3.41) is not needed for product weights when $p = 1$.

3.4. Component-by-component algorithm. We first express $E_d(\mathbf{q})$ in (3.31) in a more convenient form for computation. Recall from Definition 3.6 that the j -th coordinate of the n -th point of the interlaced polynomial lattice point set is

$$y_j^{(n)} = v_m \left(\frac{n(x) q_j(x)}{P(x)} \right).$$

In the following we write $y_j^{(n)}$ for brevity; note however that $y_j^{(n)}$ depends on the j -th component q_j of the generating vector. We have

$$\sum_{\ell_{\mathbf{v}} \in \mathcal{D}_{\mathbf{v}}^*} b^{-\alpha \mu_1(\ell_{\mathbf{v}})} = \frac{1}{b^m} \sum_{n=0}^{b^m-1} \sum_{\ell \in \mathbb{N}^{|\mathbf{v}|}} b^{-\alpha \mu_1(\ell)} \text{wal}_{\ell}(\mathbf{y}_{\mathbf{v}}^{(n)}) = \frac{1}{b^m} \sum_{n=0}^{b^m-1} \prod_{j \in \mathbf{v}} \omega(y_j^{(n)}),$$

where $\mathbf{y}_{\mathbf{v}}^{(n)} = (y_j^{(n)})_{j \in \mathbf{v}}$ is the projection of n -th point $\mathbf{y}^{(n)}$ onto the coordinates in \mathbf{v} ,

$$\omega(y) = \sum_{\ell=1}^{\infty} b^{-\alpha \mu_1(\ell)} \text{wal}_{\ell}(y) = \frac{b-1}{b^{\alpha}-b} - b^{\lfloor \log_b y \rfloor (\alpha-1)} \frac{b^{\alpha}-1}{b^{\alpha}-b},$$

and where for $y = 0$ we set $b^{\lfloor \log_b 0 \rfloor (\alpha-1)} := 0$. The last equality can be obtained by multiplying [9, Eq. (2)] by $b^{-\alpha}$. Hence we can write

$$E_d(\mathbf{q}) = \frac{1}{b^m} \sum_{n=0}^{b^m-1} \sum_{\emptyset \neq \mathbf{v} \subseteq \{1:d\}} \tilde{\gamma}_{\mathbf{v}} \prod_{j \in \mathbf{v}} \omega(y_j^{(n)}). \quad (3.42)$$

To facilitate the CBC construction, it is important that $E_d(\mathbf{q})$ allows recursively separating out a term depending only on the highest dimension. The strategy depends on the form of weights.

SPOD weights. Combining (3.32) with (3.17), we obtain SPOD weights

$$\tilde{\gamma}_{\mathbf{v}} = \sum_{\nu_{\mathbf{u}(\mathbf{v})} \in \{1:\alpha\}^{|\mathbf{u}(\mathbf{v})|}} |\nu_{\mathbf{u}(\mathbf{v})}|! \prod_{j \in \mathbf{u}(\mathbf{v})} \gamma_j(\nu_j), \quad \text{with } \gamma_j(\nu_j) := C_{\alpha,b} b^{\alpha(\alpha-1)/2} 2^{\delta(\nu_j, \alpha)} \beta_j^{\nu_j}.$$

Substituting this into (3.42) yields

$$E_d(\mathbf{q}) = \frac{1}{b^m} \sum_{n=0}^{b^m-1} \sum_{\emptyset \neq \mathbf{v} \subseteq \{1:d\}} \sum_{\nu_{\mathbf{u}(\mathbf{v})} \in \{1:\alpha\}^{|\mathbf{u}(\mathbf{v})|}} |\nu_{\mathbf{u}(\mathbf{v})}|! \left(\prod_{j \in \mathbf{u}(\mathbf{v})} \gamma_j(\nu_j) \right) \left(\prod_{j \in \mathbf{v}} \omega(y_j^{(n)}) \right).$$

Recall that every block of α components in the generating vector \mathbf{q} yields one component for the interlaced polynomial lattice rule. For convenience, we replace the index d by a double index (s, t) such that s is the index for the block and t is the index within the block, that is, we have

$$s = \lceil d/\alpha \rceil \quad \text{and} \quad t = (d-1) \bmod \alpha + 1 \quad \text{such that} \quad d = \alpha(s-1) + t.$$

We then reorder the sums in $E_d(\mathbf{q})$ according to $\nu = (\nu_1, \dots, \nu_s) \in \{0:\alpha\}^s$ and $\mathbf{v} \subseteq \{1:d\}$ so that the set $\mathbf{u}(\mathbf{v})$ consists of the indices j for which $\nu_j > 0$. This yields

$$\begin{aligned} E_{s,t}(\mathbf{q}) &= \frac{1}{b^m} \sum_{n=0}^{b^m-1} \sum_{\substack{\nu \in \{0:\alpha\}^s \\ |\nu| \neq 0}} \sum_{\substack{\mathbf{v} \subseteq \{1:d\} \text{ s.t.} \\ \mathbf{u}(\mathbf{v}) = \{1 \leq j \leq s : \nu_j > 0\}}} |\nu|! \left(\prod_{j \in \mathbf{u}(\mathbf{v})} \gamma_j(\nu_j) \right) \left(\prod_{j \in \mathbf{v}} \omega(y_j^{(n)}) \right) \\ &= \frac{1}{b^m} \sum_{n=0}^{b^m-1} \sum_{\ell=1}^{\alpha s} \ell! \sum_{\substack{\nu \in \{0:\alpha\}^s \\ |\nu| = \ell}} \left(\prod_{\substack{j=1 \\ \nu_j > 0}}^s \gamma_j(\nu_j) \right) \sum_{\substack{\mathbf{v} \subseteq \{1:d\} \text{ s.t.} \\ \mathbf{u}(\mathbf{v}) = \{1 \leq j \leq s : \nu_j > 0\}}} \prod_{j \in \mathbf{v}} \omega(y_j^{(n)}). \end{aligned} \quad (3.43)$$

When $t = \alpha$, that is, when the final block is complete, we have

$$E_{s,\alpha}(\mathbf{q}) = \frac{1}{b^m} \sum_{n=0}^{b^m-1} \sum_{\ell=1}^{\alpha s} \ell! \underbrace{\sum_{\substack{\nu \in \{0:\alpha\}^s \\ |\nu| = \ell}} \prod_{\substack{j=1 \\ \nu_j > 0}}^s \left[\gamma_j(\nu_j) \left(\prod_{i=1}^{\alpha} (1 + \omega(y_{j,i}^{(n)})) - 1 \right) \right]}_{=: U_{s,\ell}(n)}, \quad (3.44)$$

where we defined the quantity $U_{s,\ell}(n)$, with $U_{0,\ell}(n) := 1$, $U_{s,0}(n) := 0$, and $U_{s,\ell}(n) := 0$ for $\ell > \alpha s$. When $t < \alpha$, that is, when the final block is incomplete, by separating out the case $\nu_s = 0$ in (3.43), we obtain

$$\begin{aligned} E_{s,t}(\mathbf{q}) &= \frac{1}{b^m} \sum_{n=0}^{b^m-1} \sum_{\ell=1}^{\alpha(s-1)} \ell! \sum_{\substack{\nu \in \{0:\alpha\}^{s-1} \\ |\nu| = \ell}} \prod_{\substack{j=1 \\ \nu_j > 0}}^{s-1} \left[\gamma_j(\nu_j) \left(\prod_{i=1}^{\alpha} (1 + \omega(y_{j,i}^{(n)})) - 1 \right) \right] \\ &\quad + \frac{1}{b^m} \sum_{n=0}^{b^m-1} \sum_{\ell=1}^{\alpha s} \sum_{\nu_s=1}^{\min(\alpha, \ell)} \ell! \sum_{\substack{\nu \in \{0:\alpha\}^{s-1} \\ |\nu| = \ell - \nu_s}} \left(\prod_{\substack{j=1 \\ \nu_j > 0}}^{s-1} \left[\gamma_j(\nu_j) \left(\prod_{i=1}^{\alpha} (1 + \omega(y_{j,i}^{(n)})) - 1 \right) \right] \right. \\ &\quad \left. \times \gamma_s(\nu_s) \left(\prod_{i=1}^t (1 + \omega(y_{s,i}^{(n)})) - 1 \right) \right), \end{aligned}$$

and thus

$$\begin{aligned}
E_{s,t}(\mathbf{q}) &= E_{s-1,\alpha}(\mathbf{q}) \\
&+ \frac{1}{b^m} \sum_{n=0}^{b^m-1} \underbrace{\left(\prod_{i=1}^t (1 + \omega(y_{s,i}^{(n)})) - 1 \right)}_{=: V_{s,t}(n)} \underbrace{\left(\sum_{\ell=1}^{\alpha s} \sum_{\nu_s=1}^{\min(\alpha,\ell)} \gamma_s(\nu_s) \frac{\ell!}{(\ell - \nu_s)!} U_{s-1,\ell-\nu_s}(n) \right)}_{=: X_{s,\ell}(n)}, \\
&\hspace{15em} \underbrace{\hspace{10em}}_{=: W_s(n)},
\end{aligned} \tag{3.45}$$

where we defined $V_{s,t}(n)$, $W_s(n)$, and $X_{s,\ell}(n)$ as indicated, with $V_{s,0}(n) := 1$.

Note that the polynomial $q_{s,t}$ only appears in the final factor of the products $V_{s,t}(n)$. In particular, the part of (3.45) that is affected by $q_{s,t}$ is

$$\sum_{n=1}^{b^m-1} \omega(y_{s,t}^{(n)}) V_{s,t-1}(n) W_s(n).$$

Computing this quantity for every $q_{s,t} \in G_{b,m}$ requires the matrix-vector multiplication with the matrix

$$\mathbf{\Omega} := \left[\omega \left(v_m \left(\frac{n(x)q(x)}{P(x)} \right) \right) \right]_{\substack{1 \leq n \leq b^m-1 \\ q \in G_{b,m}}}$$

and the vector $[V_{s,t-1}(n) W_s(n)]_{1 \leq n \leq b^m-1}$. The rows and columns of this matrix can be permuted to allow the matrix-vector multiplication to be carried out using the fast Fourier transform, see [24], with a cost of $\mathcal{O}(M \log M) = \mathcal{O}(N \log N)$ operations, where $M = b^m - 1$ and $N = b^m$. The strategy is based on the Rader transform, see also [11, Chapter 10.3].

Once $q_{s,t}$ is chosen for dimension $\alpha(s-1) + t$, we update the products $V_{s,t}(n)$ by

$$V_{s,t}(n) = (1 + \omega(y_{s,t}^{(n)})) V_{s,t-1}(n).$$

This requires $\mathcal{O}(N)$ operations. Once we have completed an entire block of α dimensions, we need to update the values $U_{s,\ell}(n)$ using

$$\begin{aligned}
U_{s,\ell}(n) &= \ell! \sum_{\substack{\nu \in \{0:\alpha\}^{s-1} \\ |\nu|=\ell}} \prod_{\substack{j=1 \\ \nu_j > 0}}^{s-1} \left[\gamma_j(\nu_j) \left(\prod_{i=1}^{\alpha} (1 + \omega(y_{j,i}^{(n)})) - 1 \right) \right] \\
&+ \ell! \sum_{\nu_s=1}^{\min(\alpha,\ell)} \sum_{\substack{\nu \in \{0:\alpha\}^{s-1} \\ |\nu|=\ell-\nu_s}} \left(\prod_{\substack{j=1 \\ \nu_j > 0}}^{s-1} \left[\gamma_j(\nu_j) \left(\prod_{i=1}^{\alpha} (1 + \omega(y_{j,i}^{(n)})) - 1 \right) \right] \right) \\
&\hspace{15em} \times \gamma_s(\nu_s) \left(\prod_{i=1}^{\alpha} (1 + \omega(y_{s,i}^{(n)})) - 1 \right) \\
&= U_{s-1,\ell}(n) + (V_{s,\alpha}(n) - 1) X_{s,\ell}(n).
\end{aligned}$$

Given that the quantities $V_{s,\alpha}(n)$ and $X_{s,\ell}(n)$ have been pre-computed and stored, this update requires $\mathcal{O}(\alpha s N)$ operations. We then need to initialize the products $V_{s+1,0}(n)$ by 1 with $\mathcal{O}(N)$ operations, and compute the quantities $W_{s+1}(n)$ and $X_{s+1,\ell}(n)$ with $\mathcal{O}(\alpha^2 s N)$ operations, before continuing the search in the new block.

The total computational cost for the CBC construction up to αs dimensions is

$$\mathcal{O}(\alpha s N \log N) \text{ search cost, plus } \mathcal{O}(\alpha^2 s^2 N) \text{ update cost.}$$

We need to store the quantities $U_{s,\ell}(n)$, $V_{s,t}(n)$, $W_s(n)$, and $X_{s,\ell}(n)$, which can be overwritten as we increase s and t . Hence, the total memory requirement is $\mathcal{O}(\alpha s N)$.

We summarize the algorithm in Pseudocode 1 below where $*$ means element wise multiplication. Note that $\mathbf{U}(\ell)$ for $\ell = 0, \dots, \alpha s_{\max}$, \mathbf{V} , \mathbf{W} , $\mathbf{X}(\ell)$ for $\ell = 1, \dots, \alpha s_{\max}$, and \mathbf{E} are all vectors of length $N - 1$, while $\mathbf{\Omega}^{\text{perm}}$ denotes the permuted version of the matrix $\mathbf{\Omega}$.

Product weights. Combining (3.32) with (3.18), we obtain product weights

$$\tilde{\gamma}_{\mathbf{v}} = \prod_{j \in \mathbf{u}(\mathbf{v})} \gamma_j, \quad \text{with } \gamma_j := C_{\alpha,b} b^{\alpha(\alpha-1)/2} \sum_{\nu=1}^{\alpha} \nu! 2^{\delta(\nu,\alpha)} \beta_j^{\nu}.$$

Substituting this into (3.42) yields

$$\begin{aligned} E_d(\mathbf{q}) &= \frac{1}{b^m} \sum_{n=0}^{b^m-1} \sum_{\emptyset \neq \mathbf{v} \subseteq \{1:d\}} \left(\prod_{j \in \mathbf{u}(\mathbf{v})} \gamma_j \right) \left(\prod_{j \in \mathbf{v}} \omega(y_j^{(n)}) \right) \\ &= \frac{1}{b^m} \sum_{n=0}^{b^m-1} \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:s\}} \left(\prod_{j \in \mathbf{u}} \gamma_j \right) \sum_{\substack{\mathbf{v} \subseteq \{1:d\} \\ \mathbf{u}(\mathbf{v}) = \mathbf{u}}} \left(\prod_{j \in \mathbf{v}} \omega(y_j^{(n)}) \right). \end{aligned}$$

Replacing d by the double index (s, t) as before, we obtain for $t = \alpha$ that

$$E_{s,\alpha}(\mathbf{q}) = \frac{1}{b^m} \sum_{n=0}^{b^m-1} \underbrace{\prod_{j=1}^s \left[1 + \gamma_j \left(\prod_{i=1}^{\alpha} (1 + \omega(y_{j,i}^{(n)})) - 1 \right) \right]}_{=: Y_s(n)} - 1,$$

where we defined the quantity $Y_s(n)$, with $Y_0(n) := 1$. For $t < \alpha$ we have

$$E_{s,t}(\mathbf{q}) = \frac{1}{b^m} \sum_{n=0}^{b^m-1} \left[1 + \gamma_s \underbrace{\left(\prod_{i=1}^t (1 + \omega(y_{s,i}^{(n)})) - 1 \right)}_{=: V_{s,t}(n)} \right] Y_{s-1}(n) - 1,$$

where $V_{s,t}(n)$ is as defined before. The part of $E_{s,t}(\mathbf{q})$ that is affected by $q_{s,t}$ is

$$\sum_{n=1}^{b^m-1} \omega(y_{s,t}^{(n)}) V_{s,t-1}(n) Y_{s-1}(n).$$

Computing this quantity for every $q_{s,t} \in G_{b,m}$ can be done using the fast Fourier transform as before, with a cost of $\mathcal{O}(N \log N)$ operations. Once $q_{s,t}$ is chosen for each dimension, we need to update the products $V_{s,t}(n)$ with $\mathcal{O}(N)$ operations. Once we have completed an entire block of α dimensions, we need to update the products $Y_s(n)$, again with $\mathcal{O}(N)$ operations. Hence the total computational cost is only $\mathcal{O}(\alpha s N \log N)$ operations, with the memory requirement of $\mathcal{O}(N)$. The algorithm for product weights is summarized in Pseudocode 2 below.

Pseudocode 1 (Fast CBC implementation for SPOD weights)

```

 $U(0) := \mathbf{1}$ 
 $U(1 : \alpha s_{\max}) := \mathbf{0}$ 
for  $s$  from 1 to  $s_{\max}$  do
   $V := \mathbf{1}$  ▷ initialize products and sums
   $W := \mathbf{0}$ 
  for  $\ell$  from 1 to  $\alpha s$  do
     $X(\ell) := \mathbf{0}$ 
    for  $\nu$  from 1 to  $\min(\alpha, \ell)$  do
       $X(\ell) := X(\ell) + \gamma_s(\nu) \frac{\ell!}{(\ell - \nu)!} U(\ell - \nu)$ 
    end for
     $W := W + X(\ell)$ 
  end for
  for  $t$  from 1 to  $\alpha$  do
     $E := \Omega^{\text{perm}}(V .* W)$  ▷ compute – use FFT
     $q_{s,t} := \operatorname{argmin}_{q \in G_{b,m}} E(q)$  ▷ select – pick the correct index
     $V := (\mathbf{1} + \Omega^{\text{perm}}(q_{s,t}, :)) .* V$  ▷ update products
  end for
  for  $\ell$  from 1 to  $\alpha s$  do ▷ update sums
     $U(\ell) := U(\ell) + (V - \mathbf{1}) .* X(\ell)$ 
  end for
end for

```

Pseudocode 2 (Fast CBC implementation for product weights)

```

 $Y := \mathbf{1}$ 
for  $s$  from 1 to  $s_{\max}$  do
   $V := \mathbf{1}$ 
  for  $t$  from 1 to  $\alpha$  do
     $E := \Omega^{\text{perm}}(V .* Y)$  ▷ compute – use FFT
     $q_{s,t} := \operatorname{argmin}_{q \in G_{b,m}} E(q)$  ▷ select – pick the correct index
     $V := (\mathbf{1} + \Omega^{\text{perm}}(q_{s,t}, :)) .* V$  ▷ update products
  end for
   $Y := (\mathbf{1} + \gamma_s(V - \mathbf{1})) .* Y$  ▷ update products
end for

```

4. Combined QMC Galerkin Error Bound. At last we return to the PDE problem where the goal is to approximate the integral (1.4) by a QMC Galerkin method (1.5). With a slight abuse of notation, here we denote (1.5) by $Q_{N,s}(G(u_s^h))$, suppressing from our notation the translation from $[0, 1]^s$ to $[-\frac{1}{2}, \frac{1}{2}]^s$. We write the overall error as

$$\begin{aligned}
& I(G(u)) - Q_{N,s}(G(u_s^h)) \\
&= [I(G(u)) - I(G(u_s))] + [I(G(u_s)) - I(G(u_s^h))] + [I(G(u_s^h)) - Q_{N,s}(G(u_s^h))].
\end{aligned} \tag{4.1}$$

The first term in (4.1) is the *dimension truncation error* which was analyzed in Theorem 2.6. The second term in (4.1) is the *Galerkin discretization error* which can be deduced from Theorem 2.5 by taking the vectors \mathbf{y} with $y_j = 0$ for $j > s$. The third term in (4.1) is the *QMC quadrature error* which can be estimated from Theorem 3.1 by noting that for the integrand $F(\mathbf{y}) = G(u_s^h(\mathbf{y}))$ we have $|(\partial_{\mathbf{y}}^\nu F)(\mathbf{y})| \leq \|G(\cdot)\|_{\mathcal{X}'} \|(\partial_{\mathbf{y}}^\nu u_s^h)(\mathbf{y})\|_{\mathcal{X}}$, while recognizing that Theorem 2.2 applies also to the truncated Galerkin solution u_s^h . We summarize the combined error estimate in the follow-

ing theorem.

THEOREM 4.1. *Under Assumption 1 and conditions (1.3), (2.11), (2.18), (2.21), if we approximate the integral (1.4) by (1.5) using an interlaced polynomial lattice rule of order $\alpha = \lfloor 1/p \rfloor + 1$ with $N = b^m$ points (with b prime) in s dimensions, combined with a Galerkin method in the domain D with one common subspace \mathcal{X}^h with $M_h = \dim(\mathcal{X}^h)$ degrees of freedom and with the approximation property (2.10) with linear cost $\mathcal{O}(M_h)$, then there holds the error bound*

$$|I(G(u)) - Q_{N,s}(G(u^h))| \leq C \left(\kappa(s, N) \|f\|_{\mathcal{Y}'} \|G(\cdot)\|_{\mathcal{X}'} + h^\tau \|f\|_{\mathcal{Y}_i'} \|G(\cdot)\|_{\mathcal{X}_i'} \right),$$

where $\tau = t + t'$, $C > 0$ is independent of s , h and N , and

$$\kappa(s, N) = \begin{cases} s^{-2(1/p-1)} + N^{-1/p} & \text{if } p \in (0, 1), \\ (\sum_{j=s+1}^{\infty} \beta_j)^2 + N^{-1} & \text{if } p = 1. \end{cases}$$

The cost for the evaluation of $Q_{N,s}(G(u^h))$ is $\mathcal{O}(sNM_h)$ operations. The cost for the CBC construction of the interlaced polynomial lattice rule is $\mathcal{O}(\alpha s N \log N + \alpha^2 s^2 N)$ operations with SPOD weights, plus $\mathcal{O}(\alpha s N)$ memory requirement.

REFERENCES

- [1] J. Baldeaux, J. Dick, J. Greslehner, and F. Pillichshammer, Construction algorithms for higher order polynomial lattice rules, *J. Complexity*, 27 (2011), pp. 281–299.
- [2] J. Baldeaux, J. Dick, G. Leobacher, D. Nuyens, and F. Pillichshammer, Efficient calculation of the worst-case error and (fast) component-by-component construction of higher order polynomial lattice rules, *Numer. Algorithms*, 59 (2012), pp. 403–431.
- [3] F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer Verlag, Berlin, 1991.
- [4] A. Cohen, A. Chkifa, and Ch. Schwab, Breaking the curse of dimensionality in sparse polynomial approximation of parametric PDEs, Report, Seminar for Applied Mathematics, ETH Zürich 2013 (in review).
- [5] A. Cohen, R. DeVore, and Ch. Schwab, Convergence rates of best N -term Galerkin approximation for a class of elliptic sPDEs, *Found. Comput. Math.*, 10 (2010), pp. 615–646.
- [6] J. Dick, Explicit constructions of Quasi-Monte Carlo rules for the numerical integration of high-dimensional periodic functions, *SIAM J. Numer. Anal.*, 45 (2007), pp. 2141–2176.
- [7] J. Dick, Walsh spaces containing smooth functions and Quasi-Monte Carlo rules of arbitrary high order, *SIAM J. Numer. Anal.*, 46 (2008), pp. 1519–1553.
- [8] J. Dick, The decay of the Walsh coefficients of smooth functions, *Bull. Aust. Math. Soc.*, 80 (2009), pp. 430–453.
- [9] J. Dick and F. Pillichshammer, Multivariate integration in weighted Hilbert spaces based on Walsh functions and weighted Sobolev spaces, *J. Complexity*, 21 (2005), pp. 149–195.
- [10] J. Dick and F. Pillichshammer, Strong tractability of multivariate integration of arbitrary high order using digitally shifted polynomial lattice rules, *J. Complexity*, 23 (2007), pp. 436–453.
- [11] J. Dick and F. Pillichshammer, *Digital Nets and Sequences. Discrepancy Theory and Quasi-Monte Carlo Integration*, Cambridge, Cambridge University Press, 2010.
- [12] T. Goda, Good interlaced polynomial lattice rules for numerical integration in weighted Walsh spaces, ArXiv Preprint arXiv:1306.4573v1 (in review).
- [13] T. Goda and J. Dick, Construction of interlaced scrambled polynomial lattice rules of arbitrary high order, ArXiv Preprint arXiv:1301.6441v2 (in review).
- [14] M. Hansen and Ch. Schwab, *Analytic regularity and best N -term approximation of high dimensional parametric initial value problems*, *Vietnam Journal of Mathematics*, 41 (2013), pp. 181–215.
- [15] V.H. Hoang and Ch. Schwab, Analytic regularity and polynomial approximation of stochastic, parametric elliptic multiscale PDEs, *Analysis and Applications (Singapore)*, 11 (2011), 1350001, pp. 50.
- [16] V. H. Hoang and Ch. Schwab, Regularity and Generalized Polynomial Chaos Approximation of Parametric and Random Second-Order Hyperbolic Partial Differential Equations, *Analysis and Applications (Singapore)*, 10 (2012), pp. 295–326.

- [17] A. Kunoth and Ch. Schwab, Analytic Regularity and GPC Approximation for Stochastic Control Problems Constrained by Linear Parametric Elliptic and Parabolic PDEs, *SIAM J. Control and Optimization*, 51 (2013), pp. 2442 – 2471.
- [18] F. Y. Kuo, Ch. Schwab, and I. H. Sloan, Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficient, *SIAM J. Numer. Anal.*, 50 (2012), pp. 3351-3374.
- [19] F. Y. Kuo, Ch. Schwab, and I. H. Sloan, Quasi-Monte Carlo methods for very high dimensional integration: the standard weighted-space setting and beyond, *ANZIAM Journal*, 53 (2011), pp. 1–37.
- [20] F. Y. Kuo, Ch. Schwab, and I. H. Sloan, Multi-Level Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficient, (in review).
- [21] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia, 1992.
- [22] V. Nistor and C. Schwab, High order Galerkin approximations for parametric second order elliptic partial differential equations, to appear in *Math. Mod. Meth. Appl. Sci.*, 2013.
- [23] J.A. Nitsche and A.H. Schatz, Interior Estimates for Ritz-Galerkin Methods, *Math. Comp.*, 28 (1974), pp. 937–958.
- [24] D. Nuyens and R. Cools, Fast algorithms for component-by-component construction of rank-1 lattice rules in shift-invariant reproducing kernel Hilbert spaces, *Math. Comp.*, 75 (2006), pp. 903–920.
- [25] Cl. Schillings and Ch. Schwab, Sparse, Adaptive Smolyak Algorithms for Bayesian Inverse Problems, *Inverse Problems*, (2013), 065011.
- [26] Cl. Schillings and Ch. Schwab, Sparsity in Bayesian Inversion of Parametric Operator Equations, Report 2013-17, Seminar for Applied Mathematics, ETH Zürich (in review).
- [27] Ch. Schwab, QMC Galerkin discretizations of parametric operator equations, *Proc. MCQMC 2012*, Sydney (to appear).
- [28] Ch. Schwab and C.J. Gittelsohn, Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs, *Acta Numerica*, 20 (2011), pp. 291–467.
- [29] I. H. Sloan and H. Woźniakowski, When are Quasi-Monte Carlo algorithms efficient for high-dimensional integrals?, *J. Complexity*, 14 (1998), pp. 1–33.

Recent Research Reports

Nr.	Authors/Title
2013-19	K. Schmidt and R. Hiptmair Asymptotic Boundary Element Methods for Thin Conducting Sheets
2013-20	R. Kruse Consistency and Stability of a Milstein-Galerkin Finite Element Scheme for Semilinear SPDE
2013-21	J. Sukys Adaptive load balancing for massively parallel multi-level Monte Carlo solvers
2013-22	R. Andreev and A. Lang Kolmogorov-Chentsov theorem and differentiability of random fields on manifolds
2013-23	P. Grohs and M. Sprecher Projection-based Quasiinterpolation in Manifolds
2013-24	P. Grohs and S. Keiper and G. Kutyniok and M. Schaefer α -Molecules: Curvelets, Shearlets, Ridgelets, and Beyond
2013-25	A. Cohen and A. Chkifa and Ch. Schwab Breaking the curse of dimensionality in sparse polynomial approximation of parametric PDEs
2013-26	A. Lang Isotropic Gaussian random fields on the sphere
2013-27	Ch. Schwab and C. Schillings Sparse Quadrature Approach to Bayesian Inverse Problems
2013-28	V. Kazeev and I. Oseledets The tensor structure of a class of adaptive algebraic wavelet transforms