Construction Algorithms for Higher Order Polynomial Lattice Rules

Jan Baldeaux*, Josef Dick†,
Julia Greslehner‡ and Friedrich Pillichshammer§

Dedicated to Gerhard Larcher on the occasion of his 50th birthday

Abstract

Higher order polynomial lattice point sets are special types of digital higher order nets which are known to achieve almost optimal convergence rates when used in a quasi-Monte Carlo algorithm to approximate high-dimensional integrals over the unit cube. The existence of higher order polynomial lattice point sets of "good" quality has recently been established, but their construction was not addressed.

We use a component-by-component approach to construct higher order polynomial lattice rules achieving optimal convergence rates for functions of arbitrarily high smoothness and at the same time – under certain conditions on the weights – (strong) polynomial tractability. Combining this approach with a sieve-type algorithm yields higher order polynomial lattice rules adjusting themselves to the smoothness of the integrand up to a certain given degree. Higher order Korobov polynomial lattice rules achieve analogous results.

*The support of the Australian Research Council under its Centre of Excellence Program is gratefully acknowledged.

†The support of the Australian Research Council under its Centre of Excellence Program is gratefully acknowledged. The author is supported by an Australian Research Council Queen Elizabeth II Research Fellowship.

‡J.G. is supported by the Austrian Science Foundation (FWF), Project P21943

§F.P. is supported by the Austrian Science Foundation (FWF), Project S9609, that is part of the Austrian National Research Network "Analytic Combinatorics and Probabilistic Number Theory".

1 Introduction

Quasi-Monte Carlo rules are equal weight integration formulas used to approximate integrals over the unit cube $[0,1]^s$, where the dimension s is typically large. In particular, one approximates an integral $I_s(f)=\int_{[0,1]^s} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ by

$$Q_{N,s}(f)=\frac{1}{N}\sum_{n=0}^{N-1} f(\boldsymbol{x}_n) \quad \text{where } \boldsymbol{x}_0,\ldots,\boldsymbol{x}_{N-1}\in[0,1)^s.$$

Popular choices for the underlying integration nodes $\boldsymbol{x}_0,\ldots,\boldsymbol{x}_{N-1}\in[0,1)^s$ are either lattice point sets (see [16, 17]) or digital (t,m,s)-nets (see [14, 16]); in this paper, we focus on digital nets.

Recently, digital higher order nets were introduced by Dick [5] which include digital (t,m,s)-nets as special cases and have the appealing property that they can exploit the smoothness of the integrand under consideration. This is not possible with ordinary digital (t,m,s)-nets. To be more precise, if the integrand under consideration has square integrable mixed partial derivatives of order α in each variable, then digital higher order nets consisting of N points can produce integration errors converging at a rate of $N^{-\alpha+\varepsilon}$ with arbitrarily small $\varepsilon>0$.

Having established the desirable properties of digital higher order nets, we now address their construction. One possible method based on classical digital nets was shown in [5, Section 4.4]. In this paper, we present constructions independent of classical digital nets, instead we employ polynomial lattice point sets, first introduced by Niederreiter [15, 16] as special cases of classical digital nets, and later generalized in [10] as special cases of digital higher order nets. Quasi-Monte Carlo rules using such point sets as integration nodes are nowadays known as (higher order) polynomial lattice rules, see [11, 16] for more information. In [10] the existence of higher order polynomial lattice rules achieving optimal convergence rates was established and furthermore these rules were shown to achieve (strong) polynomial tractability. However, being of probabilistic nature, the approach does not show how to construct such point sets; see also [7] for a further nonconstructive existence result based on the concept of a figure of merit.

In this paper we use a component-by-component (CBC) approach (an idea first used in [18]) to produce higher order polynomial lattice rules achieving the optimal rate of convergence for functions having higher order mixed partial derivatives, see Algorithm 1 and Theorem 3.1. Furthermore, by combining the CBC approach with a "sieve"-type algorithm (as used in [4, 12]) we can even construct higher order polynomial lattice rules which automatically adjust themselves to the smoothness of the integrand in terms of the convergence of the integration error within a certain (arbitrarily high) range; see Algorithm 2 and Theorem 4.2. We point out already here, that an analogous result for lattice rules is not known. Finally, analogous results are obtained using so-called higher order Korobov polynomial lattice rules.

The structure of the paper is as follows: In Section 2 we recall higher order polynomial lattice rules, discuss the function space under consideration and present a result on numerical integration in this function space employing higher order polynomial lattice rules. In Section 3 we use a CBC approach to construct higher order polynomial lattice rules achieving optimal rates of convergence for functions of a given smoothness and in Section 4 we show how to construct higher order polynomial lattice rules achieving optimal convergence rates for a given range of smoothness parameters using a CBC sieve algorithm. Finally, in Section 5, analogous results for higher order Korobov polynomial lattice rules are established.

2 Preliminaries

In this section we introduce higher order polynomial lattice rules which can achieve arbitrarily high convergence rates, the function space under consideration, and a result on numerical integration in this function space when using higher order polynomial lattice rules.

2.1 Polynomial lattice rules for arbitrarily smooth functions

For a prime b let $\mathbb{Z}_b$ be the finite field with b elements and let $\mathbb{Z}_b((x^{-1}))$ be the field of formal Laurent series over $\mathbb{Z}_b$. Elements of $\mathbb{Z}_b((x^{-1}))$ are formal Laurent series,

$$L = \sum_{l=w}^{\infty} t_l x^{-l},$$

where w is an arbitrary integer and all $t_l \in \mathbb{Z}_b$. Note that $\mathbb{Z}_b((x^{-1}))$ contains the field of rational functions over $\mathbb{Z}_b$ as a subfield. Further let $\mathbb{Z}_b[x]$ be the set of all polynomials over $\mathbb{Z}_b$.

For an integer n let v_n be the map from $\mathbb{Z}_b((x^{-1}))$ to the interval $[0,1)$ defined by

$$v_n\left(\sum_{l=w}^{\infty} t_l x^{-l}\right) = \sum_{l=\max(1,w)}^{n} t_l b^{-l}.$$

Furthermore, we write $\vec{h}$ for vectors over $\mathbb{Z}_b$, $\boldsymbol{h}$ for vectors over $\mathbb{Z}$ or $\mathbb{R}$ and denote polynomials over $\mathbb{Z}_b$ by $h(x)$ and vectors of polynomials by $\boldsymbol{h}(x)$. Given an integer h with b-adic expansion $h = \sum_{r=0}^{\infty} h_r b^r$, we denote the associated polynomial by $\overline{h}(x)$, which is given by

$$\overline{h}(x) = \sum_{r=0}^{n-1} h_r x^r$$

and vectors of associated polynomials are denoted by $\overline{\boldsymbol{h}}(x)$. For arbitrary $\boldsymbol{k}(x) = (k_1(x), \ldots, k_s(x)) \in \mathbb{Z}_b[x]^s$ and $\boldsymbol{q}(x) = (q_1(x), \ldots, q_s(x)) \in \mathbb{Z}_b[x]^s$, we define the "inner product"

$$\boldsymbol{k}(x) \cdot \boldsymbol{q}(x) = \sum_{j=1}^{s} k_j(x) q_j(x) \in \mathbb{Z}_b[x],$$

and we write $q(x) \equiv 0 \pmod{p(x)}$ if $p(x)$ divides $q(x)$ in $\mathbb{Z}_b[x]$. The following definition of higher order polynomial lattice rules given in [10] is a slight generalization of the definition from [15], see also [16].

Definition 2.1 Let b be prime and let $1 \leq m \leq n$ be integers. For a given dimension $s \geq 1$, choose $p(x) \in \mathbb{Z}_b[x]$ with $\deg(p(x)) = n$ and let $q_1(x), \ldots, q_s(x) \in \mathbb{Z}_b[x]$. Then $\mathcal{S}_{p,m,n}(\boldsymbol{q})$, where $\boldsymbol{q} = (q_1(x), \ldots, q_s(x))$, is the point set consisting of the b^m points

$$\boldsymbol{x}_{h(x)} = \left(v_n\left(\frac{h(x)q_1(x)}{p(x)}\right), \ldots, v_n\left(\frac{h(x)q_s(x)}{p(x)}\right) \right) \in [0,1)^s,$$

for $h(x) \in \mathbb{Z}_b[x]$ with $\deg(h(x)) < m$. A quasi-Monte Carlo rule using the point set $\mathcal{S}_{p,m,n}(\boldsymbol{q})$ is called a *polynomial lattice rule*.

Remark 2.1 Using similar arguments as for the classical case $n = m$, see [15, 16], it can be shown that the point set $S_{p,m,n}(\boldsymbol{q})$ is a digital net in the sense of [5] which can be seen as a generalization of the classical definition of digital nets according to Niederreiter [11, 14, 15, 16]. The generating matrices $C_1, \ldots, C_s \in \mathbb{Z}_b^{n\times m}$ of this digital net can be obtained in the following way: For $1 \le j \le s$ consider the expansions

$$\frac{q_j(x)}{p(x)} = \sum_{l=w_j}^{\infty} u_l^{(j)} x^{-l} \in \mathbb{Z}_b((x^{-1})),$$

where $w_j \in \mathbb{Z}$. Then the elements $c_{l,r}^{(j)}$ of the $n \times m$ matrix C_j over $\mathbb{Z}_b$ are given by

$$c_{l,r}^{(j)} = u_{r+l}^{(j)} \in \mathbb{Z}_b,$$

for $1 \le j \le s$, $1 \le l \le n$, $0 \le r \le m-1$.

We remark here that for our results only the degree of the polynomial $p(x)$ is important and not the specific choice of $p(x)$ itself (we assume though that $p(x)$ is irreducible, but this assumption could be removed by a more complicated analysis).

2.2 Walsh functions and the function space $\mathscr{W}_{\alpha,s,\boldsymbol{\gamma}}$

We now define the space of functions we are going to study. This function space is based on Walsh functions whose definition is recalled in the following.

Let $\mathbb{N}_0$ denote the set of nonnegative and $\mathbb{N}$ the set of positive integers. Each $k \in \mathbb{N}$ has a unique b-adic representation $k = \sum_{i=0}^{a} \kappa_i b^i$ with digits $\kappa_i \in \{0, \ldots, b-1\}$ for $0 \le i \le a$, where $\kappa_a \neq 0$. For $k = 0$ we have $a = 0$ and $\kappa_0 = 0$. Similarly, each $x \in [0,1)$ has a b-adic representation $x = \sum_{i=1}^{\infty} \xi_i b^{-i}$ with digits $\xi_i \in \{0, \ldots, b-1\}$ for $i \ge 1$. This representation is unique in the sense that infinitely many of the ξ_i must differ from $b-1$. We define the kth Walsh function in base b, $\mathrm{wal}_k : [0,1) \to \mathbb{C}$ by

$$\mathrm{wal}_k(x) := \exp(2\pi \mathtt{i}(\xi_1\kappa_0 + \cdots + \xi_{a+1}\kappa_a)/b).$$

For dimension $s \ge 2$ and vectors $\boldsymbol{k} = (k_1, \ldots, k_s) \in \mathbb{N}_0^s$ and $\boldsymbol{x} = (x_1, \ldots, x_s) \in [0,1)^s$ we define $\mathrm{wal}_{\boldsymbol{k}} : [0,1)^s \to \mathbb{C}$ by

$$\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}) := \prod_{j=1}^{s} \mathrm{wal}_{k_j}(x_j).$$

It follows from the definition above that Walsh functions are piecewise constant functions. For more information on Walsh functions, see, e.g., [2, 20] or [11, Appendix A].

When studying integration errors resulting from the approximation of an integral based on a digital net or digital higher order net or a (higher order) polynomial lattice rule, it is convenient to consider the Walsh series of the integrand f. In particular, for $f \in L_2([0,1]^s)$, the Walsh series of f is given by

$$f(\boldsymbol{x}) \sim \sum_{\boldsymbol{k} \in \mathbb{N}_0^s} \widehat{f}(\boldsymbol{k}) \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}), \tag{1}$$

where the Walsh coefficients $\widehat{f}(\boldsymbol{k})$ are given by

$$\widehat{f}(\boldsymbol{k}) = \int_{[0,1]^s} f(\boldsymbol{x}) \overline{\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})} \,\mathrm{d}\boldsymbol{x}.$$

In general, the Walsh series given in Equation (1) need not converge to f, however, for the space of Walsh series $\mathscr{W}_{\alpha,s,\boldsymbol{\gamma}}$, which we define in the following, it does, see also [5]. For more details on the convergence of Walsh series, we refer to [5] or [11].

Throughout the paper we assume that b is a fixed prime, all polynomials are over $\mathbb{Z}_b$ and all Walsh functions are also considered in the same base b.

The function space under consideration in this paper is the space $\mathscr{W}_{\alpha,s,\boldsymbol{\gamma}} \subseteq L_2([0,1]^s)$ as introduced in [5]. Here $\boldsymbol{\gamma} = (\gamma_j)_{j=1}^{\infty}$ is a sequence of positive non-increasing weights, which are introduced to model the importance of different variables for our approximation problem, see [19]. For $s \in \mathbb{N}$ let $[s] := \{1, \ldots, s\}$ and for $\mathfrak{u} \subseteq [s]$ let $\gamma_{\mathfrak{u}} := \prod_{j \in \mathfrak{u}} \gamma_j$ be the weight associated with the projection onto components whose index is contained in $\mathfrak{u}$. The parameter α, which assumes values in $\mathbb{N}$ and satisfies $\alpha \geq 2$, determines the smoothness of the function space via the function $\mu_\alpha(.)$, which we now define.

Given a positive integer k with base b expansion $k = \kappa_1 b^{a_1-1} + \kappa_2 b^{a_2-1} + \cdots + \kappa_v b^{a_v-1}$, $1 \leq a_v < \cdots < a_1$, $v \geq 1$ and $\kappa_1, \ldots, \kappa_v \in \{1, \ldots, b-1\}$, we define $\mu_\alpha(k) := a_1 + \cdots + a_{\min(v,\alpha)}$. Furthermore we put $\mu_\alpha(0) := 0$.

For $k \in \mathbb{N}_0$ and a weight $\gamma > 0$, we define a function

$$r_\alpha(\gamma, k) := \begin{cases} 1 & \text{if } k = 0, \\ \gamma b^{-\mu_\alpha(k)} & \text{otherwise.} \end{cases} \tag{2}$$

If we consider a vector $\boldsymbol{k} \in \mathbb{N}_0^s$ of the form $\boldsymbol{k} = (k_1, \ldots, k_s)$, we set

$$r_\alpha(\boldsymbol{\gamma}, \boldsymbol{k}) := \prod_{j=1}^{s} r_\alpha(\gamma_j, k_j).$$

Definition 2.2 The space $\mathscr{W}_{\alpha,s,\boldsymbol{\gamma}} \subseteq L_2([0,1]^s)$ consists of all Walsh series $f = \sum_{\boldsymbol{k}\in\mathbb{N}_0^s} \widehat{f}(\boldsymbol{k})\mathrm{wal}_{\boldsymbol{k}}$ for which the norm

$$\|f\|_{\mathscr{W}_{\alpha,s,\boldsymbol{\gamma}}} := \sup_{\boldsymbol{k}\in\mathbb{N}_0^s} \frac{|\widehat{f}(\boldsymbol{k})|}{r_\alpha(\boldsymbol{\gamma}, \boldsymbol{k})} \tag{3}$$

is finite.

For $\alpha \geq 2$, the following property was shown in [5]: Let $f : [0,1]^s \to \mathbb{R}$ be such that all mixed partial derivatives up to order α in each variable are square integrable, then $f \in \mathscr{W}_{\alpha,s,\boldsymbol{\gamma}}$. Furthermore, an inequality using a Sobolev type norm and the norm in Equation (3) was shown in [5] establishing that $\mathscr{W}_{\alpha,s,\boldsymbol{\gamma}}$ contains certain Sobolev spaces, see also [3, 6]. Consequently, the results we are going to establish in the following for functions in $\mathscr{W}_{\alpha,s,\boldsymbol{\gamma}}$ also apply automatically to smooth functions. The assumption $\alpha > 1$ is needed to ensure that the sum of the absolute values of the Walsh coefficients converges. For the case $\alpha = 1$, which requires a different analysis, we refer to [9] or to [11].

2.3 Numerical Integration in $\mathscr{W}_{\alpha,s,\boldsymbol{\gamma}}$

We are interested in the worst-case error of multivariate integration in $\mathscr{W}_{\alpha,s,\boldsymbol{\gamma}}$ using a quasi-Monte Carlo rule $Q_{b^m,s}$, which is given by

$$e(Q_{b^m,s}, \mathscr{W}_{\alpha,s,\boldsymbol{\gamma}}) = \sup_{\substack{f\in\mathscr{W}_{\alpha,s,\boldsymbol{\gamma}} \\ \|f\|_{\mathscr{W}_{\alpha,s,\boldsymbol{\gamma}}}\leq 1}} |I_s(f) - Q_{b^m,s}(f)|. \tag{4}$$

The initial error is given by

$$e(Q_{0,s}, \mathscr{W}_{\alpha,s,\boldsymbol{\gamma}}) = \sup_{\substack{f\in\mathscr{W}_{\alpha,s,\boldsymbol{\gamma}} \\ \|f\|_{\mathscr{W}_{\alpha,s,\boldsymbol{\gamma}}}\leq 1}} |I_s(f)| = \|I_s\|.$$

We denote the quasi-Monte Carlo rule based on a polynomial lattice point set $S_{p,m,n}(\boldsymbol{q})$ by $Q_{b^m,s}(\boldsymbol{q})$ and the associated worst-case integration error by $e_{b^m,\alpha}(\boldsymbol{q}, p)$. The next proposition gives information on this quantity.

Proposition 2.1 *Let b be a prime and $\alpha \geq 2$ an integer. Then the worst-case integration error for multivariate integration in $\mathscr{W}_{\alpha,s,\boldsymbol{\gamma}}$ using the polynomial lattice point set $S_{p,m,n}(\boldsymbol{q})$ is given by*

$$e_{b^m,\alpha}(\boldsymbol{q},p) = \sum_{\boldsymbol{k} \in \mathscr{D}_p(\boldsymbol{q})} r_\alpha(\boldsymbol{\gamma},\boldsymbol{k}),$$

where

$$\begin{aligned}\mathscr{D}_p(\boldsymbol{q}) \quad := \quad & \big\{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\} : \overline{\boldsymbol{k}}(x) \cdot \boldsymbol{q}(x) \equiv u(x) \pmod{p(x)} \\ & \qquad\qquad \text{with } \deg(u(x)) < n-m\}\,. \end{aligned} \tag{5}$$

Proof. Combine [5, Equation (5.2)] with the determination of the dual net $\mathscr{D}$ of a polynomial lattice from [10, Section 4]. □

Finally, the next proposition presents an expression for $e_{b^m,\alpha}(\boldsymbol{q},p)$ which is computable; of course, such an expression is needed to implement the algorithms presented in this paper.

Proposition 2.2 *The worst-case integration error in $\mathscr{W}_{\alpha,s,\boldsymbol{\gamma}}$ associated with the polynomial lattice point set $S_{p,m,n}(\boldsymbol{q})$ satisfies*

$$e_{b^m,\alpha}(\boldsymbol{q}_s,p) = -1 + \frac{1}{b^m}\sum_{h=0}^{b^m-1}\prod_{j=1}^{s}(1+\gamma_j\omega(x_{h,j},\alpha)), \tag{6}$$

where, for $x \in [0,1)$, $\omega(x,\alpha) = \sum_{k=1}^\infty r_\alpha(1,k)\mathrm{wal}_k(x)$.

Proof. Using Proposition 2.1 and [5, Lemma 4.2], we get

$$\begin{aligned} e_{b^m,\alpha}(\boldsymbol{q}_s,p) \quad &= \quad \sum_{\boldsymbol{k} \in \mathscr{D}_p(\boldsymbol{q})} r_\alpha(\boldsymbol{\gamma},\boldsymbol{k}) \\ &= \quad \sum_{\boldsymbol{k} \in \mathscr{D}_p(\boldsymbol{q})} r_\alpha(\boldsymbol{\gamma},\boldsymbol{k}) \sum_{h=0}^{b^m-1} \frac{\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}_h)}{b^m} \\ &= \quad \sum_{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}} r_\alpha(\boldsymbol{\gamma},\boldsymbol{k}) \sum_{h=0}^{b^m-1} \frac{\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}_h)}{b^m} \\ &= \quad -1 + \frac{1}{b^m}\sum_{h=0}^{b^m-1}\sum_{\boldsymbol{k} \in \mathbb{N}_0^s} r_\alpha(\boldsymbol{\gamma},\boldsymbol{k})\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}_h) \end{aligned}$$

$$= -1 + \frac{1}{b^m} \sum_{h=0}^{b^m-1} \prod_{j=1}^{s} \left[\sum_{k=0}^{\infty} r_\alpha(\gamma_j, k) \mathrm{wal}_k(x_{h,j}) \right]$$
$$= -1 + \frac{1}{b^m} \sum_{h=0}^{b^m-1} \prod_{j=1}^{s} (1 + \gamma_j \omega(x_{h,j}, \alpha)).$$

□

We conclude this subsection by noting that an efficient implementation of the function $\omega(\cdot,\cdot)$ is presented in [1].

3 Component-by-component construction of polynomial lattice rules

We propose the following algorithm to construct a polynomial lattice rule that achieves higher order convergence. We remark that unlike the results presented in Section 4, we only deal with a fixed α in this section. For ease of notation, we proceed as follows: We use $q = q(x) \in \mathbb{Z}_b[x]$, $p = p(x) \in \mathbb{Z}_b[x]$ and $u = u(x) \in \mathbb{Z}_b[x]$; also, if we consider the polynomial associated with an integer k, we use $\overline{k} = \overline{k}(x) \in \mathbb{Z}_b[x]$. We put

$$G_{b,n} := \{q \in \mathbb{Z}_b[x] : \deg(q) < n\}.$$

We also make use of the following lemma, which appeared in a weaker and non-explicit form as [10, Lemma 4.2]. The constant $C_{b,\alpha,\lambda}$ introduced in the following lemma will be used repeatedly throughout the paper.

Lemma 3.1 *Let $\alpha \geq 2$ be an integer. Then for every $1/\alpha < \lambda \leq 1$ we have*

$$\sum_{l=1}^{\infty} r_\alpha^\lambda(\gamma, l) \leq \gamma^\lambda C_{b,\alpha,\lambda},$$

where

$$C_{b,\alpha,\lambda} := \widetilde{C}_{b,\alpha,\lambda} + \frac{(b-1)^\alpha}{b^{\lambda\alpha} - b} \prod_{i=1}^{\alpha-1} \frac{1}{b^{\lambda i} - 1},$$
$$\widetilde{C}_{b,\alpha,\lambda} := \begin{cases} \alpha - 1 & \text{if } \lambda = 1, \\ \frac{(b-1)((b-1)^{\alpha-1} - (b^\lambda - 1)^{\alpha-1})}{(b-b^\lambda)(b^\lambda-1)^{\alpha-1}} & \text{if } \lambda < 1. \end{cases}$$

Furthermore, the series $\sum_{l=1}^{\infty} r_\alpha^\lambda(\gamma, l)$ *diverges to* ∞ *as* λ *goes to* $1/\alpha$ *from the right.*

Proof. Let $l = \lambda_1 b^{a_1-1} + \cdots + \lambda_v b^{a_v-1}$ where $v \geq 1$, $0 < a_v < \cdots < a_1$ and $\lambda_i \in \{1, \ldots, b-1\}$. We divide the sum over all $l \in \mathbb{N}$ into two parts, namely firstly where $1 \leq v \leq \alpha - 1$ and secondly where $v > \alpha - 1$. For the first part, it follows from Equation (2) that

$$\begin{aligned}
\sum_{\substack{l=1 \\ 1 \leq v \leq \alpha-1}}^{\infty} r_\alpha^\lambda(\gamma, l) &= \gamma^\lambda \sum_{v=1}^{\alpha-1} (b-1)^v \sum_{0 < a_v < \cdots < a_1} \frac{1}{b^{\lambda(a_1 + \cdots + a_v)}} \\
&= \gamma^\lambda \sum_{v=1}^{\alpha-1} (b-1)^v \sum_{a_1=v}^{\infty} \frac{1}{b^{\lambda a_1}} \sum_{a_2=v-1}^{a_1-1} \frac{1}{b^{\lambda a_2}} \cdots \sum_{a_v=1}^{a_{v-1}-1} \frac{1}{b^{\lambda a_v}} \\
&\leq \gamma^\lambda \sum_{v=1}^{\alpha-1} \left(\frac{b-1}{b^\lambda - 1} \right)^v = \begin{cases} \gamma^\lambda (\alpha - 1) & \text{if } \lambda = 1, \\ \gamma^\lambda \frac{(b-1)((b-1)^{\alpha-1} - (b^\lambda - 1)^{\alpha-1})}{(b - b^\lambda)(b^\lambda - 1)^{\alpha-1}} & \text{if } \lambda < 1, \end{cases} \\
&= \gamma^\lambda \widetilde{C}_{b,\alpha,\lambda}.
\end{aligned}$$

For the second part we have

$$\begin{aligned}
\sum_{\substack{l=1 \\ v > \alpha-1}}^{\infty} r_\alpha^\lambda(\gamma, l) &= \gamma^\lambda (b-1)^\alpha \sum_{0 < a_\alpha < \cdots < a_1} \frac{b^{a_\alpha - 1}}{b^{\lambda(a_1 + \cdots + a_\alpha)}} \\
&= \gamma^\lambda \frac{(b-1)^\alpha}{b} \sum_{a_1=\alpha}^{\infty} \frac{1}{b^{\lambda a_1}} \sum_{a_2=\alpha-1}^{a_1-1} \frac{1}{b^{\lambda a_2}} \cdots \sum_{a_\alpha=1}^{a_{\alpha-1}-1} \frac{b^{a_\alpha}}{b^{\lambda a_\alpha}} \\
&= \gamma^\lambda \frac{(b-1)^\alpha}{b} \sum_{a_\alpha=1}^{\infty} \frac{b^{a_\alpha}}{b^{\lambda a_\alpha}} \sum_{a_{\alpha-1}=a_\alpha+1}^{\infty} \frac{1}{b^{\lambda a_{\alpha-1}}} \cdots \sum_{a_2=a_3+1}^{\infty} \frac{1}{b^{\lambda a_2}} \sum_{a_1=a_2+1}^{\infty} \frac{1}{b^{\lambda a_1}} \\
&= \gamma^\lambda \frac{(b-1)^\alpha}{b} \prod_{i=1}^{\alpha-1} \frac{1}{b^{\lambda i} - 1} \sum_{a_\alpha=1}^{\infty} \frac{b^{a_\alpha}}{b^{\lambda a_\alpha}} \frac{1}{b^{\lambda(\alpha-1)a_\alpha}} \\
&= \gamma^\lambda \frac{(b-1)^\alpha}{b^{\lambda\alpha} - b} \prod_{i=1}^{\alpha-1} \frac{1}{b^{\lambda i} - 1}.
\end{aligned}$$

Hence, we have shown that

$$\gamma^\lambda \frac{(b-1)^\alpha}{b^{\lambda\alpha} - b} \prod_{i=1}^{\alpha-1} \frac{1}{b^{\lambda i} - 1} \leq \sum_{l=1}^{\infty} r_\alpha^\lambda(\gamma, l)$$

$$\leq \quad \gamma^\lambda \left(\widetilde{C}_{b,\alpha,\lambda} + \frac{(b-1)^\alpha}{b^{\lambda\alpha} - b} \prod_{i=1}^{\alpha-1} \frac{1}{b^{\lambda i} - 1} \right) =: \gamma^\lambda C_{b,\alpha,\lambda}.$$

As $\frac{(b-1)^\alpha}{b^{\lambda\alpha}-b} \prod_{i=1}^{\alpha-1} \frac{1}{b^{\lambda i}-1} \to \infty$ whenever $\lambda \to 1/\alpha$ from the right we also obtain the second assertion. $\square$

Now we show that a component-by-component approach can be used to construct a polynomial lattice rule that achieves higher order convergence, where for $1 \leq d \leq s$, we set $\boldsymbol{q}_d = (q_1, \ldots, q_d)$. Note that we consider this vector instead of $(1, q_2, \ldots, q_s)$, c.f. [8, Algorithm 4.3], as otherwise the projection onto the first coordinate does not achieve a convergence rate of $b^{-\alpha m}$, see also [10, Remark 2.3]. The component-by-component algorithm for a fixed α is summarized in Algorithm 1.

Algorithm 1 CBC algorithm for fixed α

Require: b a prime, $s, m \in \mathbb{N}, n \geq m$ and weights $\boldsymbol{\gamma} = (\gamma_j)_{j \geq 1}$.

1: Choose an irreducible polynomial $p \in \mathbb{Z}_b[x]$, with $\deg(p) = n$.
2: **for** $d = 1$ to s **do**
3: find $q_d \in G_{b,n}$ by minimizing $e_{b^m,\alpha}((q_1, \ldots, q_d), p)$ as a function of q_d.
4: **end for**
5: **return** $\boldsymbol{q} = (q_1, \ldots, q_s)$.

Theorem 3.1 *Let b be prime, let $s, n, m, \alpha \in \mathbb{N}$, $m \leq n$ and let $\alpha \geq 2$. Suppose $(q_1^*, \ldots, q_s^*) \in G_{b,n}^s$ is constructed using Algorithm 1 and p is chosen by Algorithm 1. Then for all $d = 1, \ldots, s$ we have:*

$$e_{b^m,\alpha}((q_1^*, \ldots, q_d^*), p) \leq \frac{1}{b^{\min(\tau m, n)}} \prod_{j=1}^{d} (1 + 3\gamma_j^{1/\tau} C_{b,\alpha,1/\tau})^\tau \quad \forall 1 \leq \tau < \alpha.$$

Proof. The proof is completed by induction and we first show the result for $d = 1$. By Proposition 2.1,

$$e_{b^m,\alpha}(q_1, p) = \sum_{k \in \mathscr{D}_p(q_1)} r_\alpha(\gamma, k).$$

The algorithm chooses q_1^* as to minimize the worst-case error, so we have

$$e_{b^m,\alpha}(q_1^*, p) \leq e_{b^m,\alpha}(q_1, p), \quad \forall q_1 \in G_{b,n}.$$

Hence for all $1/\alpha < \lambda \le 1$ we have

$$e_{b^m,\alpha}(q_1^*,p)^\lambda \le \frac{1}{b^n} \sum_{q_1 \in G_{b,n}} e_{b^m,\alpha}(q_1,p)^\lambda.$$

Using an argument very similar to the one used in the proof of [10, Proposition 4.3], it can be shown that for all $1/\alpha < \lambda \le 1$

$$e_{b^m,\alpha}(q_1^*,p)^\lambda \le \frac{1}{b^n} \sum_{q_1 \in G_{b,n}} e_{b^m,\alpha}(q_1,p)^\lambda \le \gamma_1^\lambda C_{b,\alpha,\lambda}(b^{-m} + b^{-\lambda n}).$$

Consequently, setting $\tau = 1/\lambda$ we obtain

$$\begin{aligned} e_{b^m,\alpha}(q_1^*,p) &\le (1+2\gamma_1^\lambda C_{b,\alpha,\lambda})^{1/\lambda} b^{-\min(m/\lambda,n)} \\ &\le (1+3\gamma_1^{1/\tau} C_{b,\alpha,1/\tau})^\tau b^{-\min(m\tau,n)}. \end{aligned}$$

We now assume that for some $1 \le d < s$ we have $\boldsymbol{q}_d^* \in G_{b,n}^d$ such that

$$e_{b^m,\alpha}(\boldsymbol{q}_d^*,p) \le b^{-\min(\tau m,n)} \prod_{j=1}^d (1+3\gamma_j^{1/\tau} C_{b,\alpha,1/\tau})^\tau.$$

We consider

$$\begin{aligned} & e_{b^m,\alpha}((\boldsymbol{q}_d^*,q_{d+1}),p) \\ &= \sum_{(\boldsymbol{k},k_{d+1}) \in \mathscr{D}_p(\boldsymbol{q}_d^*,q_{d+1})} r_\alpha(\boldsymbol{\gamma},\boldsymbol{k}) r_\alpha(\gamma_{d+1},k_{d+1}) \\ &= \sum_{\boldsymbol{k} \in \mathscr{D}_p(\boldsymbol{q}_d^*)} r_\alpha(\boldsymbol{\gamma},\boldsymbol{k}) + \sum_{k_{d+1}=1}^\infty r_\alpha(\gamma_{d+1},k_{d+1}) \sum_{\substack{\boldsymbol{k} \in \mathbb{N}_0^d \\ (\boldsymbol{k},k_{d+1}) \in \mathscr{D}_p(\boldsymbol{q}_d^*,q_{d+1})}} r_\alpha(\boldsymbol{\gamma},\boldsymbol{k}) \\ &= e_{b^m,\alpha}(\boldsymbol{q}_d^*,p) + \theta(\boldsymbol{q}_d^*,q_{d+1}), \end{aligned}$$

where we set

$$\theta(\boldsymbol{q}_d^*,q_{d+1}) := \sum_{k_{d+1}=1}^\infty r_\alpha(\gamma_{d+1},k_{d+1}) \sum_{\substack{\boldsymbol{k} \in \mathbb{N}_0^d \\ (\boldsymbol{k},k_{d+1}) \in \mathscr{D}_p(\boldsymbol{q}_d^*,q_{d+1})}} r_\alpha(\boldsymbol{\gamma},\boldsymbol{k}).$$

We see from Algorithm 1 that q_{d+1}^* is chosen in such a way that the worst-case error $e_{b^m,\alpha}((\boldsymbol{q}_d^*,q_{d+1}),p)$ is minimized. Since the only dependence on

q_{d+1} is in $\theta(\boldsymbol{q}_d^*, q_{d+1})$ we have $\theta(\boldsymbol{q}_d^*, q_{d+1}^*) \le \theta(\boldsymbol{q}_d^*, q_{d+1})$ for all $q_{d+1} \in G_{b,n}$. This implies that for all $1/\alpha < \lambda \le 1$ we have

$$\begin{aligned}
\theta(\boldsymbol{q}_d^*, q_{d+1}^*)^\lambda &\le \frac{1}{b^n} \sum_{q_{d+1} \in G_{b,n}} \theta(\boldsymbol{q}_d^*, q_{d+1})^\lambda \\
&= \frac{1}{b^n} \sum_{q_{d+1} \in G_{b,n}} \left(\sum_{k_{d+1}=1}^{\infty} r_\alpha(\gamma_{d+1}, k_{d+1}) \sum_{\substack{\boldsymbol{k} \in \mathbb{N}_0^d \\ (\boldsymbol{k}, k_{d+1}) \in \mathscr{D}_p(\boldsymbol{q}_d^*, q_{d+1})}} r_\alpha(\boldsymbol{\gamma}, \boldsymbol{k}) \right)^\lambda \\
&\le \frac{1}{b^n} \sum_{q_{d+1} \in G_{b,n}} \sum_{k_{d+1}=1}^{\infty} r_\alpha^\lambda(\gamma_{d+1}, k_{d+1}) \left(\sum_{\substack{\boldsymbol{k} \in \mathbb{N}_0^d \\ (\boldsymbol{k}, k_{d+1}) \in \mathscr{D}_p(\boldsymbol{q}_d^*, q_{d+1})}} r_\alpha(\boldsymbol{\gamma}, \boldsymbol{k}) \right)^\lambda \\
&\le \sum_{\substack{k_{d+1}=1 \\ p | \overline{k}_{d+1}}}^{\infty} r_\alpha^\lambda(\gamma_{d+1}, k_{d+1}) \left(\sum_{\substack{\boldsymbol{k} \in \mathbb{N}_0^d \\ \overline{\boldsymbol{k}} \cdot \boldsymbol{q}_d^* \equiv u \pmod p \\ \deg(u) < n-m}} r_\alpha(\boldsymbol{\gamma}, \boldsymbol{k}) \right)^\lambda \qquad (7) \\
&\quad + \frac{1}{b^n} \sum_{\substack{k_{d+1}=1 \\ p \nmid \overline{k}_{d+1}}}^{\infty} r_\alpha^\lambda(\gamma_{d+1}, k_{d+1}) \sum_{q_{d+1} \in G_{b,n}} \sum_{\substack{\boldsymbol{k} \in \mathbb{N}_0^d \\ \overline{\boldsymbol{k}} \cdot \boldsymbol{q}_d^* + \overline{k}_{d+1} q_{d+1} \equiv u \pmod p \\ \deg(u) < n-m}} r_\alpha^\lambda(\boldsymbol{\gamma}, \boldsymbol{k}), \qquad (8)
\end{aligned}$$

where we used Jensen's inequality, which states that for a sequence (a_k) of nonnegative reals we have $\left(\sum a_k\right)^\lambda \le \sum a_k^\lambda$ for any $0 < \lambda \le 1$. We now prove bounds for the terms in Equations (7) and (8). First we consider the term in Equation (7). We have

$$\sum_{\substack{k=1 \\ p | \overline{k}}}^{\infty} r_\alpha^\lambda(\gamma, k) = \sum_{l=1}^{\infty} r_\alpha^\lambda(\gamma, b^n l) + \sum_{l=0}^{\infty} \sum_{\substack{k=1 \\ p | \overline{k}}}^{b^n-1} r_\alpha^\lambda(\gamma, k + b^n l).$$

For $l > 0$ we have $r_\alpha(\gamma, b^n l) \le b^{-n} r_\alpha(\gamma, l)$. Further for $1 \le k < b^n$ the polynomial p never divides $\overline{k}$ since $\deg(p) = n$. Hence

$$\sum_{\substack{k=1\\ p|\overline{k}}}^{\infty} r_\alpha^\lambda(\gamma, k) = \sum_{l=1}^{\infty} r_\alpha^\lambda(\gamma, b^n l) \le b^{-\lambda n} \sum_{l=1}^{\infty} r_\alpha^\lambda(\gamma, l) \le \frac{\gamma^\lambda C_{b,\alpha,\lambda}}{b^{\lambda n}}.$$

Therefore we can bound the term in Equation (7) by

$$\begin{aligned}
& \sum_{\substack{k_{d+1}=1\\ p|\overline{k}_{d+1}}}^{\infty} r_\alpha^\lambda(\gamma_{d+1}, k_{d+1}) \left(\sum_{\substack{\boldsymbol{k}\in\mathbb{N}_0^d\\ \overline{\boldsymbol{k}}\cdot\boldsymbol{q}_d^*\equiv u \pmod p\\ \deg(u)<n-m}} r_\alpha(\boldsymbol{\gamma}, \boldsymbol{k}) \right)^\lambda \\
\le\ & \frac{\gamma_{d+1}^\lambda C_{b,\alpha,\lambda}}{b^{\lambda n}} \left(1 + \sum_{\substack{\boldsymbol{k}\in\mathbb{N}_0^d\setminus\{\boldsymbol{0}\}\\ \overline{\boldsymbol{k}}\cdot\boldsymbol{q}_d^*\equiv u \pmod p\\ \deg(u)<n-m}} r_\alpha(\boldsymbol{\gamma}, \boldsymbol{k}) \right)^\lambda \\
\le\ & \frac{\gamma_{d+1}^\lambda C_{b,\alpha,\lambda}}{b^{\lambda n}} \left(1 + e_{b^m,\alpha}(\boldsymbol{q}_d^*, p)^\lambda \right). \qquad (9)
\end{aligned}$$

Next we provide a bound for the term in Equation (8). We have

$$\begin{aligned}
& \frac{1}{b^n} \sum_{q_{d+1}\in G_{b,n}} \sum_{\substack{k_{d+1}=1\\ p\nmid\overline{k}_{d+1}}}^{\infty} r_\alpha^\lambda(\gamma_{d+1}, k_{d+1}) \left(\sum_{(\boldsymbol{k},k_{d+1})\in\mathcal{D}_p(\boldsymbol{q}_d^*,q_{d+1})} r_\alpha(\boldsymbol{\gamma}, \boldsymbol{k}) \right)^\lambda \\
\le\ & \frac{1}{b^n} \sum_{\substack{k_{d+1}=1\\ p\nmid\overline{k}_{d+1}}}^{\infty} r_\alpha^\lambda(\gamma_{d+1}, k_{d+1}) \sum_{q_{d+1}\in G_{b,n}} \sum_{\substack{\boldsymbol{k}\in\mathbb{N}_0^d\\ \overline{\boldsymbol{k}}\cdot\boldsymbol{q}_d^*+\overline{k}_{d+1}q_{d+1}\equiv u \pmod p\\ \deg(u)<n-m}} r_\alpha^\lambda(\boldsymbol{\gamma}, \boldsymbol{k}).
\end{aligned}$$

Now we have

$$\sum_{q_{d+1}\in G_{b,n}} \sum_{\substack{\boldsymbol{k}\in\mathbb{N}_0^d\\ \overline{\boldsymbol{k}}\cdot\boldsymbol{q}_d^*+\overline{k}_{d+1}q_{d+1}\equiv u \pmod p\\ \deg(u)<n-m}} r_\alpha^\lambda(\boldsymbol{\gamma}, \boldsymbol{k})$$

$$
\begin{aligned}
&= \sum_{\boldsymbol{k}\in\mathbb{N}_0^d} r_\alpha^\lambda(\boldsymbol{\gamma},\boldsymbol{k}) \sum_{\substack{u\in\mathbb{Z}_b[x]\\ \deg(u)<n-m}} \sum_{\substack{q_{d+1}\in G_{b,n}\\ \overline{\boldsymbol{k}}\cdot\boldsymbol{q}_d^*+\overline{k}_{d+1}q_{d+1}\equiv u \pmod p}} 1 \\
&\le \sum_{\boldsymbol{k}\in\mathbb{N}_0^d} r_\alpha^\lambda(\boldsymbol{\gamma},\boldsymbol{k}) b^{n-m} \\
&= b^{n-m}\prod_{j=1}^d (1+C_{b,\alpha,\lambda}\gamma_j^\lambda).
\end{aligned}
$$

Hence

$$
\begin{aligned}
&\frac{1}{b^n}\sum_{\substack{k_{d+1}=1\\ p\nmid\overline{k}_{d+1}}}^{\infty} r_\alpha^\lambda(\gamma_{d+1},k_{d+1}) \sum_{q_{d+1}\in G_{b,n}} \sum_{\substack{\boldsymbol{k}\in\mathbb{N}_0^d\\ \overline{\boldsymbol{k}}\cdot\boldsymbol{q}_d^*+\overline{k}_{d+1}q_{d+1}\equiv u \pmod p\\ \deg(u)<n-m}} r_\alpha^\lambda(\boldsymbol{\gamma},\boldsymbol{k}) \\
&\quad\le \frac{1}{b^n}\sum_{k_{d+1}=1}^{\infty} r_\alpha^\lambda(\gamma_{d+1},k_{d+1}) b^{n-m}\prod_{j=1}^d(1+C_{b,\alpha,\lambda}\gamma_j^\lambda) \\
&\quad\le \frac{1}{b^m} C_{b,\alpha,\lambda}\gamma_{d+1}^\lambda \prod_{j=1}^d (1+C_{b,\alpha,\lambda}\gamma_j^\lambda)\,. \qquad (10)
\end{aligned}
$$

Now, from Equations (9) and (10) it follows that

$$
\begin{aligned}
\theta(\boldsymbol{q}_d^*, q_{d+1}^*) &\le \Bigg(\frac{\gamma_{d+1}^\lambda C_{b,\alpha,\lambda}}{b^{\lambda n}}(1+e_{b^m,\alpha}(\boldsymbol{q}_d^*,p)^\lambda) \\
&\qquad + \frac{1}{b^m}C_{b,\alpha,\lambda}\gamma_{d+1}^\lambda\prod_{j=1}^d(1+C_{b,\alpha,\lambda}\gamma_j^\lambda)\Bigg)^{1/\lambda} \\
&\le \gamma_{d+1}C_{b,\alpha,\lambda}^{1/\lambda}\left[\frac{1}{b^{\lambda n}} + e_{b^m,\alpha}(\boldsymbol{q}_d^*,p)^\lambda + \frac{1}{b^m}\prod_{j=1}^d(1+C_{b,\alpha,\lambda}\gamma_j^\lambda)\right]^{1/\lambda}.
\end{aligned}
$$

We now set $\tau = 1/\lambda$ and use the induction hypothesis to obtain

$$
\theta(\boldsymbol{q}_d^*, q_{d+1}^*) \le \gamma_{d+1}C_{b,\alpha,1/\tau}^{\tau}\left(\frac{1}{b^{n/\tau}} + e_{b^m,\alpha}(\boldsymbol{q}_d^*,p)^{1/\tau} + \frac{1}{b^m}\prod_{j=1}^d(1+C_{b,\alpha,1/\tau}\gamma_j^{1/\tau})\right)^{\tau}
$$

$$\le \gamma_{d+1} C_{b,\alpha,1/\tau}^{\tau} \left(\frac{3}{b^{\min(m,n/\tau)}} \prod_{j=1}^{d} (1 + 3\gamma_j^{1/\tau} C_{b,\alpha,1/\tau}) \right)^{\tau}$$

$$= \frac{3^{\tau}}{b^{\min(\tau m,n)}} \gamma_{d+1} C_{b,\alpha,1/\tau}^{\tau} \prod_{j=1}^{d} (1 + 3\gamma_j^{1/\tau} C_{b,\alpha,1/\tau})^{\tau}.$$

Finally, we have

$$\begin{aligned} e_{b^m,\alpha}(\boldsymbol{q}_{d+1}^*, p) &= e_{b^m,\alpha}(\boldsymbol{q}_d^*, p) + \theta(\boldsymbol{q}_d^*, q_{d+1}^*) \\ &\le \frac{1}{b^{\min(\tau m,n)}} \prod_{j=1}^{d} (1 + 3\gamma_j^{1/\tau} C_{b,\alpha,1/\tau})^{\tau} \\ &\quad + \frac{3^{\tau}}{b^{\min(\tau m,n)}} \gamma_{d+1} C_{b,\alpha,1/\tau}^{\tau} \prod_{j=1}^{d} (1 + 3\gamma_j^{1/\tau} C_{b,\alpha,1/\tau})^{\tau} \\ &= \frac{1}{b^{\min(\tau m,n)}} (1 + 3^{\tau} \gamma_{d+1} C_{b,\alpha,1/\tau}^{\tau}) \prod_{j=1}^{d} (1 + 3\gamma_j^{1/\tau} C_{b,\alpha,1/\tau})^{\tau} \\ &\le \frac{1}{b^{\min(\tau m,n)}} \prod_{j=1}^{d+1} (1 + 3\gamma_j^{1/\tau} C_{b,\alpha,1/\tau})^{\tau}, \end{aligned}$$

where we again used Jensen's inequality. □

From Theorem 3.1 we obtain the following corollary.

Corollary 3.1 *Let b be prime, let $s, n, m, \alpha \in \mathbb{N}$, $m \le n$ and $\alpha \ge 2$. Suppose $\boldsymbol{q}^* \in G_{b,n}^s$ is constructed using Algorithm 1 and p is chosen by Algorithm 1.*

- *We have*

$$e_{b^m,\alpha}(\boldsymbol{q}^*, p) \le \frac{c_{s,\alpha,\boldsymbol{\gamma},\delta}}{b^{\min((\alpha-\delta)m,n)}} \quad \forall 0 < \delta \le \alpha - 1,$$

  where

$$c_{s,\alpha,\boldsymbol{\gamma},\delta} := \prod_{j=1}^{s} \left(1 + 3\gamma_j^{\frac{1}{\alpha-\delta}} C_{b,\alpha,\frac{1}{\alpha-\delta}} \right)^{\alpha-\delta}.$$

- *Suppose $\sum_{j=1}^{\infty} \gamma_j^{\frac{1}{\alpha-\delta}} < \infty$, then $c_{s,\alpha,\boldsymbol{\gamma},\delta} \le c_{\infty,\alpha,\boldsymbol{\gamma},\delta} < \infty$ and we have*

$$e_{b^m,\alpha}(\boldsymbol{q}^*, p) \le \frac{c_{\infty,\alpha,\boldsymbol{\gamma},\delta}}{b^{\min((\alpha-\delta)m,n)}} \quad \forall 0 < \delta \le \alpha - 1.$$

  Thus the worst-case error is bounded independently of the dimension.

- *Under the assumption* $A := \limsup_{s\to\infty} \sum_{j=1}^{s} \gamma_j/(\log s) < \infty$ *we obtain* $c_{s,\alpha,\boldsymbol{\gamma},(\alpha-1)} \le \widetilde{c}_\eta s^{2C_{b,\alpha,1}(A+\eta)}$ *and therefore*

$$e_{b^m,\alpha}(\boldsymbol{q}^*, p) \le \frac{\widetilde{c}_\eta s^{2C_{b,\alpha,1}(A+\eta)}}{b^m} \quad \forall \eta > 0,$$

where $\widetilde{c}_\eta$ *depends only on* η. *Thus the worst-case error satisfies a bound which depends only polynomially on the dimension.*

Proof. The first part follows from Theorem 3.1 by setting $\tau = \alpha - \delta$. The second and the third part follow from the first part in exactly the same way as in the proof of [8, Corollary 4.5]. □

The above result shows that higher order polynomial lattice rules can achieve a worst-case error satisfying at the same time the almost optimal convergence rate and a bound which depends only polynomially (or even does not depend) on the dimension s (the technical term for such a behavior is (strong) polynomial tractability). Until now it is not known whether this is possible for ordinary lattice rules.

4 Optimal convergence rates for a range of smoothness parameters

In this section, we construct polynomial lattices which are optimal for a range of smoothness parameters; we use α and τ_α to denote the smoothness, where $2 \le \alpha \le \beta$, $1 \le \tau_\alpha < \alpha$.

We set

$$A_{m,n,s,\alpha,p}(\lambda) := \frac{1}{b^{sn}} \sum_{\boldsymbol{q}_s \in G_{b,n}^s} e^{\lambda}_{b^m,\alpha}(\boldsymbol{q}_s, p).$$

Proposition 4.1 *For* $\alpha \ge 2$ *and* $1/\alpha < \lambda \le 1$ *we have*

$$A_{m,n,s,\alpha,p}(\lambda) \le \frac{2}{b^{\min(m,\lambda n)}} \left(-1 + \prod_{j=1}^{s} (1 + \gamma_j^\lambda C_{b,\alpha,\lambda}) \right).$$

Proof. Using Proposition 2.1 and Jensen's inequality we obtain,

$$A_{m,n,s,\alpha,p}(\lambda) \quad \le \quad \frac{1}{b^{sn}} \sum_{\boldsymbol{q} \in G_{b,n}^s} \sum_{\boldsymbol{k} \in \mathscr{D}_p(\boldsymbol{q})} r_\alpha^\lambda(\boldsymbol{\gamma}, \boldsymbol{k})$$

$$= \sum_{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}} r_\alpha^\lambda(\boldsymbol{\gamma}, \boldsymbol{k}) \frac{1}{b^{sn}} \sum_{\substack{\boldsymbol{q} \in G_{b,n}^s \\ \overline{\boldsymbol{k}} \cdot \boldsymbol{q} \equiv u \pmod{p} \\ \deg(u) < n-m}} 1. \tag{11}$$

In the case where all components of $\overline{\boldsymbol{k}}$ are multiples of p every $\boldsymbol{q}$ satisfies the equation $\overline{\boldsymbol{k}} \cdot \boldsymbol{q} \equiv 0 \pmod{p}$ and hence we have

$$\frac{1}{b^{sn}} \sum_{\substack{\boldsymbol{q} \in G_{b,n}^s \\ \overline{\boldsymbol{k}} \cdot \boldsymbol{q} \equiv u \pmod{p} \\ \deg(u) < n-m}} 1 = 1$$

and the sum over all $\overline{\boldsymbol{k}}$ which satisfy this condition equals

$$\sum_{\substack{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\} \\ \overline{\boldsymbol{k}} \equiv \boldsymbol{0} \pmod{p}}} r_\alpha^\lambda(\boldsymbol{\gamma}, \boldsymbol{k}) = -1 + \prod_{j=1}^s \sum_{\substack{k=0 \\ p | \overline{k}}}^\infty r_\alpha^\lambda(\gamma_j, k).$$

Now we have

$$\sum_{\substack{k=0 \\ p | \overline{k}}}^\infty r_\alpha^\lambda(\gamma_j, k) = \sum_{l=0}^\infty r_\alpha^\lambda(\gamma_j, b^n l) + \sum_{l=0}^\infty \sum_{\substack{k=1 \\ p | \overline{k}}}^{b^n - 1} r_\alpha^\lambda(\gamma_j, k + b^n l).$$

For $l > 0$ we have $r_\alpha(\gamma_j, b^n l) \leq b^{-n} r_\alpha(\gamma_j, l)$ and further for $1 \leq k < b^n$ the polynomial p never divides $\overline{k}$ since $\deg(p) = n$. Hence

$$\sum_{\substack{k=0 \\ p | \overline{k}}}^\infty r_\alpha^\lambda(\gamma_j, k) = 1 + \sum_{l=1}^\infty r_\alpha^\lambda(\gamma_j, b^n l) \leq 1 + \frac{1}{b^{\lambda n}} \sum_{l=1}^\infty r_\alpha^\lambda(\gamma_j, l)\,.$$

Therefore,

$$\begin{aligned} \sum_{\substack{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\} \\ \overline{\boldsymbol{k}} \equiv \boldsymbol{0} \pmod{p}}} r_\alpha^\lambda(\boldsymbol{\gamma}, \boldsymbol{k}) &\leq -1 + \prod_{j=1}^s (1 + b^{-\lambda n} \gamma_j^\lambda C_{b,\alpha,\lambda}) \\ &= \sum_{\emptyset \neq \mathfrak{u} \subseteq [s]} b^{-|\mathfrak{u}| \lambda n} \gamma_{\mathfrak{u}}^\lambda C_{b,\alpha,\lambda}^{|\mathfrak{u}|} \end{aligned}$$

$$\leq \frac{1}{b^{\lambda n}} \left(-1 + \prod_{j=1}^{s} (1 + \gamma_j^{\lambda} C_{b,\alpha,\lambda}) \right).$$

In the case where there is at least one component of $\overline{\boldsymbol{k}}$ which is not a multiple of p we have

$$\frac{1}{b^{sn}} \sum_{\substack{\boldsymbol{q} \in G_{b,n}^s \\ \overline{\boldsymbol{k}} \cdot \boldsymbol{q} \equiv u \pmod p \\ \deg(u) < n-m}} 1 = \frac{1}{b^m}$$

and therefore this part of Equation (11) is bounded by

$$\begin{aligned} \frac{1}{b^m} \sum_{\substack{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\} \\ \overline{\boldsymbol{k}} \not\equiv \boldsymbol{0} \pmod p}} r_\alpha^{\lambda}(\boldsymbol{\gamma}, \boldsymbol{k}) &\leq \frac{1}{b^m} \sum_{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}} r_\alpha^{\lambda}(\boldsymbol{\gamma}, \boldsymbol{k}) \\ &\leq \frac{1}{b^m} \left(-1 + \prod_{j=1}^{s} (1 + \gamma_j^{\lambda} C_{b,\alpha,\lambda}) \right). \end{aligned}$$

Altogether we now obtain that

$$\begin{aligned} A_{m,n,s,\alpha,p}(\lambda) &\leq \left(\frac{1}{b^m} + \frac{1}{b^{\lambda n}} \right) \left(-1 + \prod_{j=1}^{s} (1 + \gamma_j^{\lambda} C_{b,\alpha,\lambda}) \right) \\ &\leq \frac{2}{b^{\min(m,\lambda n)}} \left(-1 + \prod_{j=1}^{s} (1 + \gamma_j^{\lambda} C_{b,\alpha,\lambda}) \right) \end{aligned}$$

as required. □

Let $\alpha \leq \beta$ and set $n = \beta m$. Let ν denote the equiprobable measure on $G_{b,\beta m}^s$. For $c \geq 1$ and $1 \leq \tau < \alpha \leq \beta$ the following set is introduced:

$$\mathscr{C}_{b,\alpha}(c, \tau) := \left\{ \boldsymbol{q} \in G_{b,\beta m}^s : e_{b^m,\alpha}(\boldsymbol{q}, p) \leq E_{b,\alpha,\boldsymbol{\gamma},s,m}(c, \tau) \right\}, \tag{12}$$

where

$$E_{b,\alpha,\boldsymbol{\gamma},s,m}(c, \tau) := \frac{2^\tau c^\tau}{b^{\tau m}} \left(-1 + \prod_{j=1}^{s} (1 + \gamma_j^{1/\tau} C_{b,\alpha,1/\tau}) \right)^{\tau}.$$

Furthermore, let

$$\mathscr{C}_{b,\alpha}(c) := \bigcap_{1 \leq \tau < \alpha} \mathscr{C}_{b,\alpha}(c, \tau)$$

$$= \quad \left\{ \boldsymbol{q} \in G^s_{b,\beta m} : e_{b^m,\alpha}(\boldsymbol{q}, p) \leq E_{b,\alpha,\boldsymbol{\gamma},s,m}(c,\tau) \ \forall 1 \leq \tau < \alpha \right\}. \quad (13)$$

(Note that the intersection $\bigcap_{1\leq\tau<\alpha} \mathscr{C}_{b,\alpha}(c,\tau)$ is finite since $\mathscr{C}_{b,\alpha}(c,\tau)$ has only finitely many elements.)

Lemma 4.1 *Let $c \geq 1$ and $1 \leq \tau < \alpha \leq \beta$, then we have*

$$\nu(\mathscr{C}_{b,\alpha}(c,\tau)) > 1 - c^{-1}.$$

Proof. We denote $\overline{\mathscr{C}}_{b,\alpha}(c,\tau) := G^s_{b,\beta m} \setminus \mathscr{C}_{b,\alpha}(c,\tau)$. Then for all $1 \leq \tau < \alpha$ we have

$$\begin{aligned} A_{m,\beta m,s,\alpha,p}(1/\tau) &= \frac{1}{b^{s\beta m}} \sum_{\boldsymbol{q} \in G^s_{b,\beta m}} e^{1/\tau}_{b^m,\alpha}(\boldsymbol{q}, p) \\ &> \nu(\overline{\mathscr{C}}_{b,\alpha}(c,\tau)) \frac{2c}{b^m} \left(-1 + \prod_{j=1}^{s} (1 + \gamma_j^{1/\tau} C_{b,\alpha,1/\tau}) \right). \end{aligned}$$

Now using Proposition 4.1 we obtain $\nu(\overline{\mathscr{C}}_{b,\alpha}(c,\tau)) < c^{-1}$ and the result follows. □

Lemma 4.2 *Let $c \geq 1$, then we have*

$$\nu(\mathscr{C}_{b,\alpha}(c)) > 1 - c^{-1}.$$

Proof. Let $1 \leq \tau_* < \alpha$ be such that

$$E_{b,\alpha,\boldsymbol{\gamma},s,m}(c,\tau_*) = \inf_{1\leq\tau<\alpha} E_{b,\alpha,\boldsymbol{\gamma},s,m}(c,\tau)$$

(note that by Lemma 3.1 we have $E_{b,\alpha,\boldsymbol{\gamma},s,m}(c,\tau) \to \infty$ whenever $\tau \to \alpha^-$ and hence we can find τ_* with the demanded property). Then we have

$$\mathscr{C}_{b,\alpha}(c,\tau_*) \subseteq \bigcap_{1\leq\tau<\alpha} \mathscr{C}_{b,\alpha}(c,\tau) = \mathscr{C}_{b,\alpha}(c)$$

and hence the result follows from Lemma 4.1. □

If we choose $c = \beta$ in Lemma 4.2, then we obtain $\nu(\mathscr{C}_{b,\alpha}(\beta)) > 1 - \beta^{-1}$ and consequently we have

$$\nu\left(\bigcap_{\alpha=2}^{\beta} \mathscr{C}_{b,\alpha}(\beta) \right) = 1 - \nu\left(\bigcup_{\alpha=2}^{\beta} \overline{\mathscr{C}}_{b,\alpha}(\beta) \right) \geq 1 - \sum_{\alpha=2}^{\beta} \nu(\overline{\mathscr{C}}_{b,\alpha}(\beta)) > 0.$$

Hence we obtain the following theorem which establishes the existence of a $\boldsymbol{q}^* \in G^s_{b,\beta m}$ which achieves the optimal convergence rate for a range of α's.

Theorem 4.1 *Let $\beta, m, s \in \mathbb{N}$, $\beta \geq 2$ and let $p \in \mathbb{Z}_b[x]$ with $\deg(p) = \beta m$. Then there exists a $\boldsymbol{q}^* \in G_{b,\beta m}^s$ such that*

$$e_{b^m,\alpha}(\boldsymbol{q}^*, p) \leq \frac{2^{\tau_\alpha}\beta^{\tau_\alpha}}{b^{\tau_\alpha m}}\left(-1+\prod_{j=1}^{s}(1+\gamma_j^{1/\tau_\alpha}C_{b,\alpha,1/\tau_\alpha})\right)^{\tau_\alpha} \tag{14}$$

for all $2 \leq \alpha \leq \beta$ and for all $1 \leq \tau_\alpha < \alpha$.

The proof of Theorem 4.1 suggests that in principle we can find $\boldsymbol{q}^*$ which satisfies Equation (14) for all $2 \leq \alpha \leq \beta$ and all $1 \leq \tau_\alpha < \alpha$ by using a so-called "sieve algorithm" which will be explained in the following.

Use a computer search to find $\lfloor (1-\beta^{-1})b^{\beta ms} \rfloor + 1$ of the $b^{\beta ms}$ vectors $\boldsymbol{q}$ in $G_{b,\beta m}^s$ which satisfy

$$e_{b^m,2}(\boldsymbol{q}, p) \leq E_{b,2,\boldsymbol{\gamma},s,m}(\beta, \tau_2) \quad \forall 1 \leq \tau_2 < 2,$$

and label this set $\mathscr{T}_2$. By Lemma 4.2 we know that at least such a number of vectors exists.

Then proceed by using a computer search to find $\lfloor (1-2\beta^{-1})b^{\beta ms} \rfloor + 1$ vectors $\boldsymbol{q}$ in $\mathscr{T}_2$ which satisfy

$$e_{b^m,3}(\boldsymbol{q}, p) \leq E_{b,3,\boldsymbol{\gamma},s,m}(\beta, \tau_3) \quad \forall 1 \leq \tau_3 < 3$$

and label this set $\mathscr{T}_3$. Since

$$\nu\left(\bigcap_{\alpha=2}^{3}\mathscr{C}_{b,\alpha}(\beta)\right) = 1-\nu\left(\bigcup_{\alpha=2}^{3}\overline{\mathscr{C}}_{b,\alpha}(\beta)\right) \geq 1-\sum_{\alpha=2}^{3}\nu(\overline{\mathscr{C}}_{b,\alpha}(\beta)) > 1-\frac{2}{\beta},$$

we know that there are at least $\lfloor (1-2\beta^{-1})b^{\beta ms} \rfloor + 1$ values in $\mathscr{T}_2$ to populate the set $\mathscr{T}_3$.

In the same way we proceed to construct the sets $\mathscr{T}_4, \ldots, \mathscr{T}_\beta$. Theorem 4.1 guarantees that $\mathscr{T}_\beta$ is not empty and we may select $\boldsymbol{q}^*$ to be any vector from $\mathscr{T}_\beta$. This vector satisfies Equation (14) for all $2 \leq \alpha \leq \beta$ and all $1 \leq \tau_\alpha < \alpha$.

However, in practice such a search algorithm would not be applicable since it is much too time consuming. For this reason we show in the following how the sieve algorithm may be combined with the component-by-component (CBC) algorithm; the resulting algorithm, referred to as "CBC sieve algorithm" is presented in Algorithm 2 and its computational complexity is feasible. For its statement we use the following notation:

For $2 \le \alpha \le \beta$ and $p \in \mathbb{Z}_b[x]$ with $\deg(p) = \beta m$ we define the following: for $d = 0$ and $q_1 \in G_{b,\beta m}$ we set

$$\theta_\alpha(0, q_1) := e_{b^m,\alpha}(q_1, p),$$

and for $d \in \mathbb{N}$, $\boldsymbol{q}_d \in G_{b,\beta m}^d$ and $q_{d+1} \in G_{b,\beta m}$ we set

$$\theta_\alpha(\boldsymbol{q}_d, q_{d+1}) := e_{b^m,\alpha}((\boldsymbol{q}_d, q_{d+1}), p) - e_{b^m,\alpha}(\boldsymbol{q}_d, p).$$

Furthermore, for short we use the notation

$$M_{d,\alpha}(\tau) := \frac{1}{b^m} \prod_{j=1}^{d} (1 + 3\beta \gamma_j^{1/\tau} C_{b,\alpha,1/\tau}). \tag{15}$$

Now we prove the following result.

Theorem 4.2 *Let $s, m, \beta \in \mathbb{N}$, $\beta \ge 2$, then Algorithm 2 constructs a vector $\boldsymbol{q}_d^* \in G_{b,\beta m}^d$ such that*

$$e_{b^m,\alpha}(\boldsymbol{q}_d^*, p) \le \frac{1}{b^{\tau_\alpha m}} \prod_{j=1}^{d} (1 + 3\beta \gamma_j^{1/\tau_\alpha} C_{b,\alpha,1/\tau_\alpha})^{\tau_\alpha}$$

for all $1 \le \tau_\alpha < \alpha$ and for all $2 \le \alpha \le \beta$.

To prove Theorem 4.2 we introduce the following set: for $\boldsymbol{q}_d \in G_{b,\beta m}^d$ let $\mathscr{F}_\alpha(c, \boldsymbol{q}_d)$ be the set of all $q_{d+1} \in G_{b,\beta m}$ such that

$$\theta_\alpha(\boldsymbol{q}_d, q_{d+1}) \le \left(3c\gamma_{d+1}^{1/\tau_\alpha} C_{b,\alpha,1/\tau_\alpha} M_{d,\alpha}(\tau_\alpha)\right)^{\tau_\alpha} \tag{16}$$

for all $1 \le \tau_\alpha < \alpha$.

Lemma 4.3 *Let $2 \le \alpha \le \beta$ and let $c \ge 1$. Assume that there exists a $\boldsymbol{q}_d \in G_{b,\beta m}^d$ such that*

$$e_{b^m,\alpha}(\boldsymbol{q}_d, p) \le M_{d,\alpha}(\tau_\alpha)^{\tau_\alpha} \tag{17}$$

for all $1 \le \tau_\alpha < \alpha$ and for all $2 \le \alpha \le \beta$. Then

$$\nu\left(\mathscr{F}_\alpha(c, \boldsymbol{q}_d)\right) > 1 - c^{-1}.$$

Algorithm 2 CBC sieve algorithm for $2 \le \alpha \le \beta$

Require: b a prime, $s, m, \beta \in \mathbb{N}$, $\beta \ge 2$, and $p \in \mathbb{Z}_b[x]$ with $\deg(p) = \beta m$.

Set $\mathscr{T}_{1,d} := G_{b,\beta m}$ for all $1 \le d \le s$ and $\boldsymbol{q}_0^* := 0$.
for $d = 0$ to $s - 1$ **do**
  for $\alpha = 2$ to β **do**
    perform a computer search to find $\lfloor (1-(\alpha-1)\beta^{-1}) b^{\beta m} \rfloor + 1$ elements q in $\mathscr{T}_{\alpha-1,d+1}$ to populate the set $\mathscr{T}_{\alpha,d+1}$, which is a subset of
    if $d = 0$ **then**

$$\left\{ q \in \mathscr{T}_{\alpha-1,d+1} : \theta_\alpha(0, q) \le \frac{1}{b^{\tau_\alpha m}} \left(1 + 3\gamma_1^{1/\tau_\alpha} C_{b,\alpha,1/\tau_\alpha} \right)^{\tau_\alpha} \ \forall 1 \le \tau_\alpha < \alpha \right\}$$

    else

$$\left\{ q \in \mathscr{T}_{\alpha-1,d+1} : \theta_\alpha(\boldsymbol{q}_d^*, q) \le \left(3\beta\gamma_{d+1}^{1/\tau_\alpha} C_{b,\alpha,1/\tau_\alpha} M_{d,\alpha,\boldsymbol{\gamma}}(\tau_\alpha) \right)^{\tau_\alpha} \ \forall 1 \le \tau_\alpha < \alpha \right\}$$

    end if
  end for
  Select $q^* \in \mathscr{T}_{\beta,d+1}$.
  Set $\boldsymbol{q}_{d+1}^* = (\boldsymbol{q}_d^*, q^*)$.
end for
return $\boldsymbol{q}^* = \boldsymbol{q}_s^*$.

Proof. From the proof of Theorem 3.1 and using Assumption (17) for all $1/\alpha < \lambda \le 1$ we have

$$\begin{aligned} & \frac{1}{b^{\beta m}} \sum_{q_{d+1} \in G_{b,\beta m}} \theta_\alpha(\boldsymbol{q}_d^*, q_{d+1})^\lambda \\ \le \ & \gamma_{d+1}^\lambda C_{b,\alpha,\lambda} \left(\frac{1}{b^{\lambda\alpha m}} + e_{b^m,\alpha}(\boldsymbol{q}_d^*, p)^\lambda + \frac{1}{b^m} \prod_{j=1}^{d} (1 + \gamma_j^\lambda C_{b,\alpha,\lambda}) \right) \\ \le \ & 3\gamma_{d+1}^\lambda C_{b,\alpha,\lambda} M_{d,\alpha}(1/\lambda). \end{aligned}$$

From this the result follows in the same way as in the proofs of Lemmas 4.1 and 4.2. $\square$

Now we give the proof of Theorem 4.2.

Proof. The proof is completed by double induction on d and α. We proceed by induction on d and firstly show the result for $d=1$, i.e. we need to prove that Algorithm 2 constructs a $q_1^* \in G_{b,\beta m}$ such that

$$e_{b^m,\alpha}(q_1^*, p) \leq \frac{1}{b^{\tau_\alpha m}}(1 + 3\beta\gamma_1^{1/\tau_\alpha} C_{b,\alpha,1/\tau_\alpha})^{\tau_\alpha} \tag{18}$$

for all $1 \leq \tau_\alpha < \alpha$ and for all $2 \leq \alpha \leq \beta$. We now proceed by induction on α: we wish to show that for $2 \leq \alpha \leq \beta$, we can find $\lfloor (1-(\alpha-1)\beta^{-1})b^{\beta m}\rfloor + 1$ elements $q \in \mathscr{T}_{\alpha-1,1}$ to populate $\mathscr{T}_{\alpha,1}$, see Algorithm 2 for the definition of $\mathscr{T}_{\alpha,1}$, in particular, we note that $\mathscr{T}_{1,1} = G_{b,\beta m}$. Consequently, Equation (18) will follow from the definition of $\mathscr{T}_{\alpha,1}$.

We firstly show the required for $\alpha = 2$: from the definition of $\mathscr{C}_{b,2}(\beta)$, see Equation (13), we have

$$e_{b^m,2}(q,p) \leq \frac{1}{b^{\tau_2 m}}(1 + 3\beta\gamma_1^{1/\tau_2} C_{b,2,1/\tau_2})^{\tau_2} \quad \forall 1 \leq \tau_2 < 2\,.$$

According to Lemma 4.2, $\nu(\mathscr{C}_{b,2}(\beta)) > 1 - \beta^{-1}$, hence there are $\lfloor (1-\beta^{-1})b^{\beta m}\rfloor + 1$ elements in $\mathscr{T}_{1,1}$ to populate $\mathscr{T}_{2,1}$. We now formulate the induction hypothesis that for $2 \leq \alpha < \beta$, there are $\lfloor (1-(\alpha-1)\beta^{-1})b^{\beta m}\rfloor + 1$ elements to populate $\mathscr{T}_{\alpha,1}$, hence $\nu(\mathscr{T}_{\alpha,1}) > 1-(\alpha-1)\beta^{-1}$. We want to show that

$$\begin{aligned}&\nu\left(\{q \in \mathscr{T}_{\alpha,1} : e_{b^m,\alpha+1}(q,p) \leq M_{1,\alpha+1}(\tau_{\alpha+1})^{\tau_{\alpha+1}}\, \forall 1 \leq \tau_{\alpha+1} < \alpha+1\}\right)\\ &> 1 - \alpha\beta^{-1},\end{aligned} \tag{19}$$

which implies that there are $\lfloor (1-\alpha\beta^{-1})b^{\beta m}\rfloor + 1$ elements in $\mathscr{T}_{\alpha,1}$ to populate $\mathscr{T}_{\alpha+1,1}$; we remind the reader that this would complete the induction over α. But

$$\begin{aligned}&\{q \in \mathscr{T}_{\alpha,1} : e_{b^m,\alpha+1}(q,p) \leq M_{1,\alpha+1}(\tau_{\alpha+1})^{\tau_{\alpha+1}}\, \forall 1 \leq \tau_{\alpha+1} < \alpha+1\}\\ &= \mathscr{T}_{\alpha,1} \cap \{q \in G_{b,\beta m} : e_{b^m,\alpha+1}(q,p) \leq M_{1,\alpha+1}(\tau_{\alpha+1})^{\tau_{\alpha+1}} \forall 1 \leq \tau_{\alpha+1} < \alpha+1\}\,,\end{aligned}$$

hence we get Equation (19) from the induction assumption and from Lemma 4.2. This completes the induction over α. As we have shown that for $2 \leq \alpha \leq \beta$ we can find $\lfloor (1-(\alpha-1)\beta^{-1})b^{\beta m}\rfloor + 1$ elements q in $\mathscr{T}_{\alpha-1,1}$ to populate $\mathscr{T}_{\alpha,1}$, it follows from the definition of $\mathscr{T}_{\alpha,1}$, see Algorithm 2, that Equation (18) holds.

We now continue the induction on d, hence we assume that for $1 \leq d < s$ the algorithm has found $\boldsymbol{q}_d^*$ such that

$$e_{b^m,\alpha}(\boldsymbol{q}_d^*, p) \leq \frac{1}{b^{\tau_\alpha m}} \prod_{j=1}^{d} (1 + 3\beta\gamma_j^{1/\tau_\alpha} C_{b,\alpha,1/\tau_\alpha})^{\tau_\alpha} = M_{d,\alpha}(\tau_\alpha)^{\tau_\alpha} \tag{20}$$

for all $1 \leq \tau_\alpha < \alpha$ and for all $2 \leq \alpha \leq \beta$, see Equation (15) for the definition of $M_{d,\alpha}(\tau)$. Of course, this assumption is to be used to establish that the algorithm has found a $\boldsymbol{q}_{d+1}^* \in G_{b,\beta m}^{d+1}$ such that

$$e_{b^m,\alpha}(\boldsymbol{q}_{d+1}^*, p) \leq M_{d+1,\alpha}(\tau_\alpha)^{\tau_\alpha} \tag{21}$$

for all $1 \leq \tau_\alpha < \alpha$ and for all $2 \leq \alpha \leq \beta$. We prove Equation (21) by induction on α, as for the case $d = 1$. In particular, we will show that for $2 \leq \alpha \leq \beta$ we can find $\lfloor (1 - (\alpha - 1)\beta^{-1}) b^{\beta m} \rfloor + 1$ elements in $\mathscr{T}_{\alpha-1,d+1}$ to populate $\mathscr{T}_{\alpha,d+1}$, which means that for all $1 \leq \tau_\alpha < \alpha$ and for all $2 \leq \alpha \leq \beta$, Algorithm 2 finds a $q^* \in G_{b,\beta m}$ such that

$$\theta(\boldsymbol{q}_{d+1}^*, q^*) \leq \left(3\beta\gamma_{d+1}^{1/\tau_\alpha} C_{b,\alpha,1/\tau_\alpha} M_{d,\alpha}\right)^{\tau_\alpha}$$

for all $1 \leq \tau_\alpha < \alpha$ and for all $2 \leq \alpha \leq \beta$. A simple manipulation involving $e_{b^m,\alpha}(\boldsymbol{q}_d^*, p)$, $\theta(\boldsymbol{q}_{d+1}^*, q^*)$, and $e_{b^m,\alpha}(\boldsymbol{q}_{d+1}^*, p)$ will then complete the proof.

Let us now proceed with the induction on α, i.e. we show that we can find $\lfloor (1 - \beta^{-1}) b^{\beta m} \rfloor + 1$ elements in $\mathscr{T}_{1,d+1}$ to populate $\mathscr{T}_{2,d+1}$. According to Lemma 4.3, under the Assumption (20), we have

$$\nu\left(\mathscr{F}_\alpha(\beta, \boldsymbol{q}_d^*)\right) > 1 - \beta^{-1} \quad \forall 2 \leq \alpha \leq \beta,$$

hence there are $\lfloor (1 - \beta^{-1}) b^{\beta m} \rfloor + 1$ elements in $\mathscr{T}_{1,d+1}$ to populate $\mathscr{T}_{2,d+1}$. We now formulate the induction hypothesis that for $2 \leq \alpha < \beta$, there are $\lfloor (1 - (\alpha - 1)\beta^{-1}) b^{\beta m} \rfloor + 1$ elements to populate $\mathscr{T}_{\alpha,d+1}$, hence $\nu(\mathscr{T}_{\alpha,d+1}) > (1 - (\alpha - 1)\beta^{-1})$.

Since

$$\begin{aligned} &\Big\{ q \in \mathscr{T}_{\alpha,d+1} : \theta(\boldsymbol{q}_d^*, q) \leq \left(3\beta\gamma_{d+1}^{1/\tau_{\alpha+1}} C_{b,\alpha+1,1/\tau_{\alpha+1}} M_{d+1,\alpha+1}(\tau_{\alpha+1})\right)^{\tau_{\alpha+1}} \\ &\qquad \forall 1 \leq \tau_{\alpha+1} < \alpha + 1 \Big\} \\ &= \mathscr{T}_{\alpha,d+1} \cap \mathscr{F}_{\alpha+1}(\beta, \boldsymbol{q}_d^*) \end{aligned}$$

we obtain from the inductive hypothesis and from Lemma 4.3 that

$$\nu\Bigg(\Bigg\{q \in \mathscr{T}_{\alpha,d+1} : \theta(\boldsymbol{q}_d^*, q) \leq \Big(3\beta\gamma_{d+1}^{1/\tau_{\alpha+1}} C_{b,\alpha+1,1/\tau_{\alpha+1}} M_{d+1,\alpha+1,\boldsymbol{\gamma}}(\tau_{\alpha+1})\Big)^{\tau_{\alpha+1}} \\ \forall 1 \leq \tau_{\alpha+1} < \alpha+1\Bigg\}\Bigg) > 1 - \alpha\beta^{-1},$$

which implies that there are $\lfloor (1-\alpha\beta^{-1})b^{\beta m}\rfloor + 1$ elements in $\mathscr{T}_{\alpha,d+1}$ to populate $\mathscr{T}_{\alpha+1,d+1}$. This completes the induction on α, and we conclude that for $2 \leq \alpha \leq \beta$ there are $\lfloor (1-(\alpha-1)\beta^{-1})b^{\beta m}\rfloor + 1$ elements in $\mathscr{T}_{\alpha-1,d+1}$ to populate $\mathscr{T}_{\alpha,d+1}$. But from the definition of $\mathscr{T}_{\alpha,d+1}$, see Algorithm 2, this shows that Algorithm 2 finds a $q^* \in G_{b,\beta m}$ such that

$$\theta(\boldsymbol{q}_{d+1}^*, q^*) \leq \Big(3\beta\gamma_{d+1}^{1/\tau_\alpha} C_{b,\alpha,1/\tau_\alpha} M_{d,\alpha}\Big)^{\tau_\alpha}$$

for all $1 \leq \tau_\alpha < \alpha$ and for all $2 \leq \alpha \leq \beta$.

Using Equation (20) we obtain,

$$\begin{aligned} e_{b^m,\alpha}((\boldsymbol{q}_d^*, q^*), p) &= e_{b^m,\alpha}(\boldsymbol{q}_d^*, p) + \theta_\alpha((\boldsymbol{q}_d^*, q^*)) \\ &\leq M_{d,\alpha,\boldsymbol{\gamma}}(\tau_\alpha)^{\tau_\alpha}(1 + (3\beta\gamma_{d+1}^{1/\tau_\alpha} C_{b,\alpha,1/\tau_\alpha})^{\tau_\alpha}) \\ &\leq M_{d+1,\alpha,\boldsymbol{\gamma}}(\tau_\alpha)^{\tau_\alpha}, \end{aligned}$$

for all $1 \leq \tau_\alpha < \alpha$ and for all $2 \leq \alpha \leq \beta$, which completes the proof. $\square$

5 Optimal convergence rates for a range of smoothness parameters using Korobov polynomial lattice rules

In this section we study a special case of polynomial lattice rules, namely Korobov polynomial lattice rules. We present an algorithm which shows how to construct higher order Korobov polynomial lattice rules achieving optimal rates of convergence for a range of smoothness parameters. This algorithm is the same as the "sieve algorithm" presented in Section 4 (not to be confused with the CBC sieve algorithm, see Algorithm 2), but due to the structure of Korobov polynomial lattice rules, the cost of such an algorithm is feasible. Regarding notation, we use $\phi(q) := (q, q^2, \ldots, q^s) \pmod{p}$, $q \in G_{b,\beta m}$, to denote the generating vector of the higher order Korobov polynomial lattice rule $S_{p,m,\beta m}(\phi(q))$ and $e_{b^m,\alpha}(\phi(q),p)$ to denote the corresponding worst-case error, $2 \le \alpha \le \beta$; we recall that $\alpha \le \beta$ and $n = \beta m$. As in Section 3 we point out that we use generating vectors $\phi(q) := (q, q^2, \ldots, q^s) \pmod{p}$ instead of $(1, q, \ldots, q^{s-1})$ (see e.g. [8, Algorithm 4.6]), as otherwise the projection onto the first coordinate does not achieve a convergence rate of $b^{-\alpha m}$.

As in Section 4 we now introduce a "sieve algorithm" (see Algorithm 3) which shows how to obtain a generating vector for a higher order Korobov polynomial lattice rule, which achieves optimal convergence rates for a range of smoothness parameters, where we use the notation

$$\widetilde{E}_{b,\alpha,\boldsymbol{\gamma},s,m}(c,\tau) := \frac{c^\tau (s+1)^\tau}{b^{\tau m}} \left(-1 + \prod_{j=1}^{s} (1 + \gamma_j^{1/\tau} C_{b,\alpha,1/\tau}) \right)^{\tau}.$$

The next theorem shows that Algorithm 3 does indeed produce such a vector; as the proof is similar to the proof of Theorem 4.2, it is omitted.

Algorithm 3 Korobov sieve algorithm

Require: b a prime, $s, m, \beta \in \mathbb{N}$, $\beta \ge 2$, and $p \in \mathbb{Z}_b[x]$ with $\deg(p) = \beta m$.

1: Set $\mathscr{T}_1 := G_{b,\beta m}$.
2: **for** $\alpha = 2$ to β **do**
3: perform a computer search to find $\lfloor (1 - (\alpha - 1)\beta^{-1}) b^{\beta m} \rfloor + 1$ elements q in $\mathscr{T}_{\alpha-1}$ to populate the set $\mathscr{T}_\alpha$, which is a subset of

$$\left\{ q \in \mathscr{T}_{\alpha-1} : e_{b^m,\alpha}(\phi(q),p) \le \widetilde{E}_{b,\alpha,\boldsymbol{\gamma},s,m}(\beta,\tau_\alpha)\, \forall 1 \le \tau_\alpha < \alpha \right\}$$

4: **end for**
5: Select $q^* \in \mathscr{T}_\beta$
6: **return** q^*.

Theorem 5.1 *Let $s, m, \beta \in \mathbb{N}$, $\beta \ge 2$. Then Algorithm 3 finds an element $q \in G_{b,\beta m}$ such that*

$$e_{b^m,\alpha}(\phi(q),p) \le \frac{(s+1)^{\tau_\alpha} \beta^{\tau_\alpha}}{b^{\tau_\alpha m}} \left(-1 + \prod_{j=1}^{s} (1 + \gamma_j^{1/\tau_\alpha} C_{b,\alpha,1/\tau_\alpha}) \right)^{\tau_\alpha},$$

for all $1 \le \tau_\alpha < \alpha$, $2 \le \alpha \le \beta$.

Acknowledgements

The authors greatly appreciate the helpful comments and suggestions of an anonymous referee which helped to significantly improve the presentation of the paper.

References

[1] J. Baldeaux, J. Dick, G. Leobacher, D. Nuyens, F. Pillichshammer, Fast Construction of Higher Order Polynomial Lattice Rules, in preparation.

[2] H.E. Chrestenson, A class of generalized Walsh functions, Pacific J. Math. 5 (1955) 17–31.

[3] J. Dick, Explicit constructions of quasi-Monte Carlo rules for the numerical integration of high-dimensional periodic functions, SIAM J. Numer. Anal. 45 (2007) 2141–2176.

[4] J. Dick, The construction of extensible polynomial lattice rules with small weighted star discrepancy, Math. Comp., 76 (2007), 2077–2085.

[5] J. Dick, Walsh spaces containing smooth functions and quasi-Monte Carlo rules of arbitrary high order, SIAM J. Numer. Anal. 46 (2008) 1519–1553.

[6] J. Dick, The decay of the Walsh coefficients of smooth functions, Bull. Austral. Math. Soc. 80 (2009) 430–453.

[7] J. Dick, P. Kritzer, F. Pillichshammer, W. Ch. Schmid, On the existence of higher order polynomial lattices based on a generalized figure of merit, J. Complexity 23 (2007) 581–593.

[8] J. Dick, F. Kuo, F. Pillichshammer, I. Sloan, Construction algorithms for polynomial lattice rules for multivariate integration, Math. Comp. 74 (2005) 1895–1921.

[9] J. Dick, F. Pillichshammer, Multivariate integration in weighted Hilbert spaces based on Walsh functions and weighted Sobolev spaces, J. Complexity 21 (2005) 149–95.

[10] J. Dick, F. Pillichshammer, Strong tractability of multivariate integration of arbitrary high order using digitally shifted polynomial lattice rules, J. Complexity 23 (2007) 436–453.

[11] J. Dick, F. Pillichshammer, Digital Nets and Sequences, Cambridge University Press, Cambridge, 2010 (to appear).

[12] J. Dick, F. Pillichshammer, B. Waterhouse, The construction of good extensible rank-1 lattices, Math. Comp. 77 (2008) 2345–2373.

[13] N.M. Korobov, Properties and calculation of optimal coefficients, Dokl. Akad. Nauk SSSR 132 (1960) 1009-1012. (In Russian.)

[14] H. Niederreiter, Point sets and sequences with small discrepancy, Monatsh. Math. 104 (1987) 273-337.

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

[16] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, CBMS–NSF Series in Applied Mathematics 63, SIAM, Philadelphia, 1992.

[17] I. H. Sloan, S. Joe, Lattice methods for multiple integration, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1994.

[18] I. H. Sloan, A. Reztsov, Component-by-component construction of good lattice rules, Math. Comp. 71 (2002) 263–273.

[19] I.H. Sloan, H. Woźniakowski, When are quasi-Monte Carlo algorithms efficient for high dimensional integrals?, J. Complexity 14 (1998) 1–33.

[20] J.L. Walsh, A closed set of normal orthogonal functions, Amer. J. Math. 45 (1923) 5–24.

Authors' Addresses:

Jan Baldeaux and Josef Dick, School of Mathematics and Statistics, The University of New South Wales, Sydney, NSW 2052, Australia.

Email: janbaldeaux@gmail.com, josef.dick@unsw.edu.au

Julia Greslehner and Friedrich Pillichshammer, Institut für Finanzmathematik, Universität Linz, Altenbergstraße 69, A-4040 Linz, Austria. Email: julia.greslehner@gmx.at, friedrich.pillichshammer@jku.at