



# Component-by-Component Digit-by-Digit Construction of Good Polynomial Lattice Rules in Weighted Walsh Spaces

Adrian Ebert<sup>1</sup> · Peter Kritzer<sup>1</sup> · Onyekachi Osisioigu<sup>1</sup> · Tetiana Stepaniuk<sup>2</sup>

Received: 20 August 2020 / Revised: 13 April 2021 / Accepted: 19 April 2021 /  
Published online: 25 August 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

## Abstract

We consider the efficient construction of polynomial lattice rules, which are special cases of so-called quasi-Monte Carlo (QMC) rules. These are of particular interest for the approximate computation of multivariate integrals where the dimension  $d$  may be in the hundreds or thousands. We study a construction method that assembles the generating vector, which is in this case a vector of polynomials over a finite field, of the polynomial lattice rule in a digit-by-digit (or, equivalently, coefficient-by-coefficient) fashion. As we will show, the integration error of the corresponding QMC rules achieves excellent convergence order, and, under suitable conditions, we can vanquish the curse of dimensionality by considering function spaces equipped with coordinate weights. The construction algorithm is based on a quality measure that is independent of the underlying smoothness of the function space and can be implemented in a fast manner (without the use of fast Fourier transformations). Furthermore, we illustrate our findings with extensive numerical results.

---

Communicated by Wolfgang Dahmen.

---

A. Ebert, P. Kritzer, and O. Osisioigu are supported by the Austrian Science Fund (FWF): Project F5506, which is part of the Special Research Program “Quasi-Monte Carlo Methods: Theory and Applications”. T. Stepaniuk is supported by the Alexander von Humboldt Foundation.

---

✉ Adrian Ebert  
adrian.ebert@oeaw.ac.at

Peter Kritzer  
peter.kritzer@oeaw.ac.at

Onyekachi Osisioigu  
onyekachi.osisioigu@oeaw.ac.at

Tetiana Stepaniuk  
stepaniuk@math.uni-luebeck.de

<sup>1</sup> Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstr. 69, 4040 Linz, Austria

<sup>2</sup> Institute of Mathematics, University of Lübeck, Ratzeburger Allee 160, 23562 Lübeck, Germany

**Keywords** Numerical integration · Polynomial lattice points · Quasi-Monte Carlo methods · Weighted function spaces · Digit-by-digit construction · Component-by-component construction · Fast implementations

**Mathematics Subject Classification** 65D30 · 65D32 · 41A55 · 41A63

## 1 Introduction

In this article, we study the problem of multivariate numerical integration for a subclass of square-integrable functions  $f \in L^2([0, 1]^d)$ . We consider special instances of so-called quasi-Monte Carlo (QMC) rules, which are methods to approximate integrals

$$I_d(f) = \int_{[0,1]^d} f(\mathbf{x}) \, d\mathbf{x}$$

by equal-weight quadrature rules,

$$Q_{N,d}(f) = \frac{1}{N} \sum_{n=0}^{N-1} f(\mathbf{x}_n),$$

where the integration nodes  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{N-1}$  are deterministically chosen in  $[0, 1]^d$ . This is in contrast to Monte Carlo rules, where the integration nodes are chosen randomly; with QMC rules, we try to make a deliberate and sophisticated choice of the points  $\mathbf{x}_n$  with the aim of obtaining better error bounds than for Monte Carlo. The crucial challenge is to find integration nodes yielding a low approximation error simultaneously for a large class of functions that may depend on many variables. This means that, usually, one needs to be able to find millions of good integration nodes in very high dimensions which is a considerable computational challenge.

In the literature on QMC methods, there are two main concepts that are commonly made use of when trying to find sets of integration nodes with good properties. These are, on the one hand, lattice point sets, as introduced independently by Korobov (see [10]) and Hlawka (see [9]). For more recent introductions to lattice rules, we refer to [16,21]. The other class of commonly used QMC integration nodes is that of (digital)  $(t, m, d)$ -nets and  $(t, d)$ -sequences, as introduced by Niederreiter, building up on ideas by Sobol' and Faure (see [14,16]). A special case of  $(t, m, d)$ -nets, namely so-called polynomial lattice point sets, is the focus of the present paper. These point sets were introduced in [15] and have their name since their structure can be viewed as analogous to (ordinary) lattice point sets.

While the construction principle of lattice point sets is based on integer arithmetic, polynomial lattice point sets are based on polynomial arithmetic over finite fields. To be more precise, we will fix a prime  $b$  and consider the finite field  $\mathbb{F}_b$  with  $b$  elements. A polynomial lattice point set with  $b^m$  points in  $[0, 1]^d$  is constructed by means of a modulus  $p \in \mathbb{F}_b[x]$  with  $\deg(p) = m$ , and a generating vector  $\mathbf{g} \in (\mathbb{F}_b[x])^d$  (we refer to Sect. 2.2 for the precise definition). The QMC rule using the polynomial lattice point

set as integration nodes is then called a polynomial lattice rule. It will be convenient in this paper to assume that the modulus has the form  $p_m(x) = x^m$ . However, it is crucial to note that not every choice of the generating vector  $\mathbf{g}$  yields a polynomial lattice point set that has good properties, in the sense that the integration error of the corresponding polynomial lattice rule is sufficiently low. On the contrary, it is usually highly non-trivial to find good generating vectors of polynomial lattice rules, and there are (except for special cases) no explicit constructions of such good generating vectors known. Hence, one has to resort to computer search algorithms for finding generating vectors of polynomial lattice point sets of high quality. Regarding the error measure, we consider in this paper the worst-case setting, i.e., we consider a particular normed function space and the supremum of the integration error over the unit ball of the space.

It is known that (ordinary) lattice rules are well suited for the numerical integration of functions with pointwise convergent Fourier series (see again, e.g., [16] or [21]). On the other hand, polynomial lattice rules are usually applied for the numerical integration of functions that can be represented by Walsh series (cf. [2,4,5]). We will therefore define a reproducing kernel Hilbert space based on Walsh functions in Sect. 2.1, which will be considered throughout the paper. The function space under consideration will be characterized by a smoothness parameter  $\alpha$  (in some publications, this parameter is also referred to as “digital smoothness parameter” in the context of Walsh series). Indeed, the parameter  $\alpha$  is linked to the speed of decay of the Walsh coefficients of the functions in our space, but there is also a connection to the number of derivatives that exist for the elements of the space (we refer to [5] and the references therein for details).

The function space considered here is closely related to other function spaces considered in the literature, such as in [2,4,5]; indeed, results that we show for the space considered in the present paper immediately imply corresponding results for some of the Walsh spaces considered in these references. Furthermore, our Hilbert space will be a “weighted” function space in the sense of Sloan and Woźniakowski (cf. [23]). This means that we assign non-negative real numbers (weights) to the coordinates, or groups of coordinates, of the integration problem, in order to model the different influence of the coordinates on the problem. As pointed out in [23] and numerous other papers, this method is justified by practical high-dimensional problems in which different coordinates may indeed have a very different degree of influence on the value of an integral. The weights will be incorporated in the inner product and norm of the function space in a suitable way. Using this setting, it is plausible that a nominally very high-dimensional problem may have a rather low “effective dimension,” i.e., only a certain, possibly small, part of the components has a significant influence on the integration problem and the error made by approximative algorithms. This may then yield situations where a curse of dimensionality can be avoided.

In the present paper, we will restrict ourselves, for technical reasons, to considering the most common choice of weights, the so-called product weights, but we suspect that the construction of QMC rules presented here could also work for other choices of weights. We refer to Remark 5 in Sect. 3.3 for further comments on this question.

The first efficient construction of good generating vectors of polynomial lattice point sets was done in [2]. In that paper, the authors considered the so-called component-by-

component (CBC) approach, which is a greedy algorithm to construct one component of the generating vector at a time. CBC algorithms were first considered for ordinary lattice point sets, with the first examples in the literature going back to Korobov (cf. [11]), and later a rediscovery by Sloan and Reztsov (cf. [22]). The fast CBC construction, which is due to Cools and Nuyens (see, e.g., [18–20]), makes the CBC construction computationally competitive and is currently the standard method to construct high-dimensional lattice point sets of good quality. It is well known (see, e.g., [2] and again [18]) that CBC constructions also work for the efficient search for generating vectors of polynomial lattice point sets; and also in this case, a fast algorithm is available.

### 1.1 Overview and Main Results

In the present paper, we present another, different algorithm to construct generating vectors of polynomial lattice point sets in an efficient way. This construction is also based on a component-by-component approach. However, as opposed to the CBC algorithms for polynomial lattice point sets currently available in the literature, our new approach constructs the single components of the generating vector  $\mathbf{g}$  “digit-by-digit” and the used search criterion is independent of the smoothness parameter  $\alpha$ . Actually, the term “digit-by-digit” is based on a similar approach that exists for ordinary lattice point sets (see [12,13], and for similar results in a more up-to-date setting, [7]). In the context of polynomial lattice point sets, the generating vector  $\mathbf{g}$  consists of polynomials, so it would be more appropriate to speak of a “coefficient-by-coefficient” instead of a “digit-by-digit” construction. However, to stay consistent regarding the name of the method, and to avoid confusion with the “component-by-component” approach, we keep the name “digit-by-digit” construction also for polynomial lattice rules. In fact, the algorithm which we will present in Sect. 3.2 contains two loops. An outer loop in which the different components are constructed, and an inner loop in which the coefficients (digits) of each component of the generating vector are constructed. Both loops can be regarded as greedy, i.e., choices that have been made in previous steps are kept fixed.

We will show that the polynomial lattice rules obtained by our new construction method (Algorithm 1) satisfy upper error bounds that are arbitrarily close to the optimal convergence rate. In particular, we will prove the following main result regarding the behavior of the worst-case error  $e_{b^m, d, \alpha, \boldsymbol{\gamma}}$  of integration in the weighted function space under consideration.

**Theorem 1** *Let  $b$  be prime, let  $m, d \in \mathbb{N}$  with  $m \geq 4$ , let  $N = b^m$ , let  $\alpha > 1$ , and let  $(\gamma_j)_{j \geq 1}$  be positive product weights satisfying*

$$\sum_{j \geq 1} \gamma_j < \infty.$$

*Then, for any  $\delta > 0$  and each  $\alpha > 1$ , the worst-case error of the polynomial lattice rule with generating vector  $\mathbf{g}$  constructed by Algorithm 1 (with weight sequence  $\boldsymbol{\gamma} = (\gamma_j)_{j \geq 1}$ ) satisfies*

$$e_{b^m,d,\alpha,\boldsymbol{\gamma}^\alpha}(\mathbf{g}) \leq \frac{1}{N^\alpha} (C(\boldsymbol{\gamma}^\alpha) + \bar{C}(\boldsymbol{\gamma}, \delta) N^{\alpha\delta}),$$

with positive constants  $C(\boldsymbol{\gamma}^\alpha)$  and  $\bar{C}(\boldsymbol{\gamma}, \delta)$ , that are independent of  $d$  and  $N$ .

Theorem 1 shows that under suitable conditions on the coordinate weights, we can vanquish the curse of dimensionality, i.e., avoid exponential dependence of the error on the dimension  $d$  of the integration problem, or even obtain error bounds that are independent of the dimension. Furthermore, the devised algorithm (when run with weights  $\boldsymbol{\gamma}$ ) constructs good polynomial lattice rules for which the proven bounds on the worst-case error hold simultaneously for all  $\alpha > 1$ . This can be an advantage over standard construction algorithms which are often tailored to a particular function space, and thus, the corresponding error bounds only hold for a particular  $\alpha$ . We will discuss this issue in more detail in the remainder of the article.

We note that Theorem 1 is a simplified version of Theorem 8, and the latter is proven in Sect. 3. As the proof of Theorem 8, which is essential to the derivation of our algorithm, is fairly technical, we will briefly outline the proof strategy that is used in Sect. 3.

The main idea is to relate our error measure, which is related to the dual net of the polynomial lattice rule, to an auxiliary quantity  $T_{\alpha,\boldsymbol{\gamma}}$  that only depends on a subset of the dual net. The difference of these two quantities will be shown to be of an order of magnitude that is at least as good as what we may expect for the error itself, which justifies concentrating on  $T_{\alpha,\boldsymbol{\gamma}}$  only. (see Proposition 1). As a next step, we show in Theorem 4 that  $T_{\boldsymbol{\gamma}} := T_{1,\boldsymbol{\gamma}}$  can be bounded by the quality function  $H_{d,m,\boldsymbol{\gamma}}$  (modulo some additive terms that can be controlled), and thus that it suffices to minimize this quality function in order to device good polynomial lattice rules. Analyzing the average of  $H_{d,m,\boldsymbol{\gamma}}$  with respect to the choices of extending the degree of the components of the generating vector (Lemma 7) then leads to the formulation of the component-by-component digit-by-digit construction in Algorithm 1. The remaining step is to show that  $H_{d,m,\boldsymbol{\gamma}}$  for the constructed polynomial lattice rules is of order  $\mathcal{O}(b^{-m})$ , which is done via an inductive argument over the dimension in Theorems 5 and 6.

Due to its structure, the obtained component-by-component digit-by-digit (CBC-DBD) algorithm can be implemented in a fast manner that only requires  $\mathcal{O}(d m b^m)$  operations. This is competitive with state-of-the-art CBC constructions; however, our algorithm does not require the use of fast Fourier transformations as well as the underlying knowledge regarding circulant matrices. We see this as a benefit of our construction.

This paper is an extension of the research on the CBC-DBD construction of lattice point sets (cf. [7]) to a special construction method for polynomial lattice point sets. As the names of the point sets already suggest, there are certain similarities in the two approaches, but instead of working with integer arithmetic, as it is the case for lattice point sets, we now need to work with polynomial arithmetic over a finite field, which makes the result technically more demanding. In particular, we would like to stress that—while in the earlier paper [7] on lattice rules we restricted ourselves to  $N$  of the form  $2^m$  for technical reasons—we were able to generalize this assumption to arbitrary prime bases  $b$  instead of  $b = 2$ , i.e., we allow  $N$  to be any prime power. This implies that we can cover a more general class of point sets as compared to [7]. The

main work underlying this more general result is to be found in Lemma 1, in which we analyze a Walsh-Dirichlet kernel.

The fact that our new algorithm can be used for the efficient construction of good polynomial lattice point sets, with a running time competitive with other common construction methods, but without the need of using the fast Fourier transform, is an advantage of the method presented. This is even more important as it is known from the theory developed by Dick (see, e.g., the monograph [5]) that certain variants of polynomial lattice rules (so-called higher-order polynomial lattice rules) allow for improved convergence rates for smoother functions, and it would be of great interest to make the method presented here also applicable to that case, for which we hope to lay the basis in this paper. We stress that analogous results for lattice rules regarding the integration of smooth functions are not known, which is why it is even more important to make the CBC-DBD construction available for polynomial lattice rules.

The rest of the paper is structured as follows. In Sect. 2, we introduce the function space setting as well as polynomial lattice rules and analyze the corresponding worst-case error expression. In Sect. 3, we derive the component-by-component digit-by-digit (or, for short, CBC-DBD) construction algorithm for polynomial lattice rules following the proof strategy outlined above. In particular, we analyze the worst-case error behavior of the resulting integration rules. In Sect. 4, we show that the introduced construction method can be implemented in a fast manner, competitive with state-of-the-art construction algorithms. Finally, the article is concluded in Sect. 5, where we illustrate our main results by extensive numerical experiments.

To conclude this introductory section, we fix some notation. In what follows, we denote the set of positive integers by  $\mathbb{N}$  and the set of non-negative integers by  $\mathbb{N}_0$ . To denote subsets of components, we use fraktur font, e.g.,  $u \subset \mathbb{N}$  and additionally write shorthand  $\{1:d\} := \{1, \dots, d\}$ . For the projection of a vector  $x \in [0, 1]^d$  or  $k \in \mathbb{N}^d$  onto the components in a set  $u \subseteq \{1:d\}$ , we write  $x_u = (x_j)_{j \in u}$  or  $k_u = (k_j)_{j \in u}$ , respectively. With a slight abuse of notation, we will frequently identify elements of the finite field  $\mathbb{F}_b$  of prime cardinality  $b$  with elements of the group of integers modulo  $b$  denoted by  $\mathbb{Z}_b$ .

## 2 Polynomial Lattice Rules in Weighted Walsh Spaces

In this article, we consider numerical integration of a sub-class of the square-integrable functions  $f \in L^2([0, 1]^d)$  which can be represented in terms of their Walsh series. This particular series representation of a function is based on the so-called Walsh functions, which are defined as follows.

**Definition 1** Let  $b \geq 2$  be an integer. For a non-negative integer  $k$ , we define the  $k$ -th Walsh function  ${}_b \text{wal}_k : [0, 1) \rightarrow \mathbb{C}$  by

$${}_b \text{wal}_k(x) := e^{2\pi i(\kappa_0 \xi_1 + \kappa_1 \xi_2 + \dots + \kappa_{a-1} \xi_a)/b}$$

with  $x \in [0, 1)$  and base  $b$  representations  $k = \kappa_0 + \kappa_1 b + \dots + \kappa_{a-1} b^{a-1}$  and  $x = \xi_1 b^{-1} + \xi_2 b^{-2} + \dots$  (unique in the sense that infinitely many of the  $\xi_i$  must

be different from  $b - 1$ ) with coefficients  $\kappa_i, \xi_i \in \{0, 1, \dots, b - 1\}$ . The imposed restriction that infinitely many of the  $\xi_i$  must be different from  $b - 1$  assures that no ambiguous representations arise.

For  $d \in \mathbb{N}$ , an integer vector  $\mathbf{k} = (k_1, \dots, k_d) \in \mathbb{N}_0^d$  and  $\mathbf{x} = (x_1, \dots, x_d) \in [0, 1)^d$ , we define the  $\mathbf{k}$ -th ( $d$ -variate) Walsh function  ${}_b\text{wal}_{\mathbf{k}} : [0, 1)^d \rightarrow \mathbb{C}$  by

$${}_b\text{wal}_{\mathbf{k}}(\mathbf{x}) := \prod_{j=1}^d {}_b\text{wal}_{k_j}(x_j).$$

In the following, we will consider the base  $b \geq 2$  as fixed (for the sake of simplicity, we will assume that  $b$  is prime), and then simply write  $\text{wal}_{\mathbf{k}}$  or  $\text{wal}_{\mathbf{k}}$  instead of  ${}_b\text{wal}_{\mathbf{k}}$  or  ${}_b\text{wal}_{\mathbf{k}}$ , respectively. It is known (see, e.g., [5]) that the Walsh functions in any fixed base  $b$  form an orthonormal basis of  $L^2([0, 1]^d)$ .

As indicated, we consider a class of square-integrable functions that can be represented in terms of their Walsh series, that is,

$$f(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{N}_0^d} \hat{f}(\mathbf{k}) \text{wal}_{\mathbf{k}}(\mathbf{x}) \quad \text{with} \quad \hat{f}(\mathbf{k}) := \int_{[0,1]^d} f(\mathbf{x}) \overline{\text{wal}_{\mathbf{k}}(\mathbf{x})} \, d\mathbf{x}, \quad (1)$$

where we call  $\hat{f}(\mathbf{k})$  the  $\mathbf{k}$ -th Walsh coefficient of  $f$ .

It is known from the literature on QMC methods in the past decades that it is advantageous to choose the integration nodes of a QMC rule such that there exists an efficient way of expressing the integration error for elements in the function class under consideration. In the case where the integrand  $f$  can be represented in terms of Walsh series as in (1), it is common to consider quasi-Monte Carlo rules which are based on so-called digital nets and sequences. Digital  $(t, m, d)$ -nets are point sets consisting of  $b^m$  elements in  $[0, 1]^d$  that satisfy certain regular distribution properties and were in their most general form introduced in [15] (see also [16]). These point sets are generated by using  $d$  generating matrices  $C_1, \dots, C_d$  over a finite field or ring. In particular, for a digital  $(t, m, d)$ -net  $P = \{\mathbf{x}_0, \dots, \mathbf{x}_{b^m-1}\} \subset [0, 1]^d$  constructed over  $\mathbb{Z}_b = \{0, 1, \dots, b - 1\}$  with generating matrices  $C_1, \dots, C_d \in \mathbb{Z}_b^{m \times m}$  the integration error of a QMC rule based on  $P$  takes a special form. It is commonly known, see, e.g., [3, Theorem 6.4], that approximating the integral  $I_d(f)$  of a  $d$ -variate function  $f$  using a QMC rule  $Q_{b^m,d}(f; P)$ , that is,

$$Q_{b^m,d}(f) = Q_{b^m,d}(f; P) := \frac{1}{b^m} \sum_{n=0}^{b^m-1} f(\mathbf{x}_n) \approx \int_{[0,1]^d} f(\mathbf{x}) \, d\mathbf{x} =: I_d(f),$$

leads to an integration error of the form

$$Q_{b^m,d}(f; P) - I_d(f) = \sum_{\mathbf{0} \neq \mathbf{k} \in \mathcal{D}} \hat{f}(\mathbf{k}) \quad (2)$$

with the dual net  $\mathcal{D} = \mathcal{D}(C_1, \dots, C_d) := \{\mathbf{k} \in \mathbb{N}_0^d \mid C_1^\top \mathbf{k}_1^m + \dots + C_d^\top \mathbf{k}_d^m = \bar{\mathbf{0}}\}$ , where for  $k \in \mathbb{N}_0$  with base  $b$  expansion  $k = \kappa_0 + \kappa_1 b + \dots + \kappa_a b^a$  we define the vector  $\mathbf{k}^m = (\kappa_0, \kappa_1, \dots, \kappa_{m-1}) \in \mathbb{Z}_b^m$ , and where we denote by  $\bar{\mathbf{0}}$  the zero vector in  $\mathbb{Z}_b^m$ . Equation (2) is a consequence of the following character property of Walsh functions,

$$\frac{1}{b^m} \sum_{n=0}^{b^m-1} \text{wal}_k(x_n) = \begin{cases} 1, & \text{if } C_1^\top \mathbf{k}_1^m + \dots + C_d^\top \mathbf{k}_d^m = \bar{\mathbf{0}}, \\ 0, & \text{otherwise.} \end{cases}$$

We will also use this property in the subsequent analysis.

### 2.1 The Weighted Walsh Space

Based on the decay of the Walsh coefficients  $\hat{f}(\mathbf{k})$  in (1), we will define a function space for the integrands considered in this paper. As mentioned in the introduction, this space will be equipped with weights to model the varying influence of the coordinates. To this end, let  $\boldsymbol{\gamma} = (\gamma_j)_{j \geq 1}$  be a non-increasing sequence of positive real numbers. The weights  $\gamma_j$  will appear in the definition of the inner product and norm of the function space defined below. Intuitively, we can think of the weight  $\gamma_j$  describing the degree of influence of the  $j$ -th variable on the integration problem. Hence, we assume (w.l.o.g.) that the coordinates are ordered according to their influence. It will also be convenient to define

$$\gamma_u := \prod_{j \in u} \gamma_j$$

for a subset  $u \subseteq \{1:d\}$ , and to additionally set  $\gamma_\emptyset$  to equal 1. The weights  $\gamma_u$  are (for obvious reasons) called product weights. In the recent literature on QMC rules, also other types of weights have been considered, but we will restrict ourselves to product weights here. We refer to [3] for further information on this subject.

For prime base  $b \geq 2$  and given smoothness parameter  $\alpha > 1$ , we set  $\psi_b(k) := \lfloor \log_b(k) \rfloor$  for  $k \in \mathbb{N}$  and define the decay function  $r_\alpha : \mathbb{N}_0 \rightarrow \mathbb{R}$  by

$$r_\alpha(k) = r_\alpha(b, k) := \begin{cases} 1, & \text{if } k = 0, \\ b^{\alpha \psi_b(k)}, & \text{if } k \neq 0, \end{cases}$$

with  $k \in \mathbb{N}_0$ . It is also convenient to define the quantity

$$\mu_b(\alpha) := \sum_{k=1}^{\infty} (r_\alpha(k))^{-1} = \sum_{a=0}^{\infty} \frac{1}{b^{a\alpha}} \sum_{k=b^a}^{b^{a+1}-1} 1 = \sum_{a=0}^{\infty} \frac{(b-1)b^a}{b^{a\alpha}} = \frac{b^\alpha(b-1)}{b^\alpha - b}. \tag{3}$$



For the multivariate case with dimension  $d \in \mathbb{N}$ , integer vector  $\mathbf{k} = (k_1, \dots, k_d) \in \mathbb{N}_0^d$ , and a sequence of weights  $\boldsymbol{\gamma} = (\gamma_j)_{j \geq 1}$ , we define the weighted decay functions

$$r_\alpha(\mathbf{k}) := \prod_{j=1}^d r_\alpha(k_j) \quad \text{and} \quad r_{\alpha, \boldsymbol{\gamma}}(\mathbf{k}) := \gamma_{\text{supp}(\mathbf{k})}^{-1} r_\alpha(\mathbf{k}) = \gamma_{\text{supp}(\mathbf{k})}^{-1} \prod_{j \in \text{supp}(\mathbf{k})} b^{\alpha \psi_b(k_j)}$$

with  $\text{supp}(\mathbf{k}) := \{j \in \{1:d\} \mid k_j \neq 0\}$ .

Using this decay function, we can estimate the integration error obtained in (2) by

$$\begin{aligned} |Q_{b^m, d}(f; P) - I_d(f)| &= \left| \sum_{\mathbf{0} \neq \mathbf{k} \in \mathcal{D}} \hat{f}(\mathbf{k}) \right| = \left| \sum_{\mathbf{0} \neq \mathbf{k} \in \mathbb{N}_0^d} \hat{f}(\mathbf{k}) r_{\alpha, \boldsymbol{\gamma}}(\mathbf{k}) (r_{\alpha, \boldsymbol{\gamma}}(\mathbf{k}))^{-1} \mathbf{1}_{\mathcal{D}}(\mathbf{k}) \right| \\ &\leq \left( \sup_{\mathbf{k} \in \mathbb{N}_0^d} |\hat{f}(\mathbf{k})| r_{\alpha, \boldsymbol{\gamma}}(\mathbf{k}) \right) \left( \sum_{\mathbf{0} \neq \mathbf{k} \in \mathcal{D}} (r_{\alpha, \boldsymbol{\gamma}}(\mathbf{k}))^{-1} \right) \end{aligned} \tag{4}$$

with  $\mathbf{1}_{\mathcal{D}}$  denoting the indicator function of the dual net  $\mathcal{D}$ . Based on this estimate, we define, for real  $\alpha > 1$  and a sequence of strictly positive weights  $\boldsymbol{\gamma} = (\gamma_j)_{j \geq 1}$ , the weighted Walsh space as

$$W_{d, \boldsymbol{\gamma}}^\alpha := \{f \in L^2([0, 1]^d) \mid \|f\|_{W_{d, \boldsymbol{\gamma}}^\alpha} < \infty\}$$

with corresponding norm  $\|\cdot\|_{W_{d, \boldsymbol{\gamma}}^\alpha}$  given by

$$\|f\|_{W_{d, \boldsymbol{\gamma}}^\alpha} := \sup_{\mathbf{k} \in \mathbb{N}_0^d} |\hat{f}(\mathbf{k})| r_{\alpha, \boldsymbol{\gamma}}(\mathbf{k}). \tag{5}$$

**Remark 1** We remark that the definition of the norm implies that functions in  $W_{d, \boldsymbol{\gamma}}^\alpha$  have an absolutely convergent Walsh series which converges pointwise (see, e.g., [5]).

**Remark 2** We would like to note here that in many recent papers (e.g., [2,4]), a slightly different function space  $\widetilde{W}_{d, \boldsymbol{\gamma}}^\alpha$  based on Walsh functions has been studied. In  $\widetilde{W}_{d, \boldsymbol{\gamma}}^\alpha$ , the norm is not given as an  $\infty$ -norm as in (5), but in the  $L^2$ -sense, i.e.,

$$\|f\|_{\widetilde{W}_{d, \boldsymbol{\gamma}}^\alpha} := \left( \sum_{\mathbf{k} \in \mathbb{N}_0^d} |\hat{f}(\mathbf{k})|^2 r_{\alpha, \boldsymbol{\gamma}}(\mathbf{k}) \right)^{\frac{1}{2}}.$$

This definition of the norm corresponds to alternatively applying Hölder’s inequality with  $p = q = 2$  in the bound on the integration error that led to (4). As we will see below, the worst-case error expressions for  $W_{d, \boldsymbol{\gamma}}^\alpha$  and  $\widetilde{W}_{d, \boldsymbol{\gamma}}^\alpha$  are closely related to each other.

In order to assess the quality of the QMC methods constructed later on, we will use the worst-case error in the weighted Walsh space as the error criterion. Indeed, the worst-case error for the QMC rule  $Q_{b^m,d}(\cdot; P)$  in the space  $W_{d,\gamma}^\alpha$  is defined as

$$e_{b^m,d,\alpha,\gamma}(P) := \sup_{\substack{f \in W_{d,\gamma}^\alpha \\ \|f\|_{W_{d,\gamma}^\alpha} \leq 1}} |I_d(f) - Q_{b^m,d}(f; P)|.$$

A useful formula for the worst-case error for  $(t, m, d)$ -nets in the function space  $W_{d,\gamma}^\alpha$  is given in the following theorem.

**Theorem 2** *Let  $m, d \in \mathbb{N}$ ,  $\alpha > 1$ , a prime  $b$ , and a sequence of positive weights  $\gamma = (\gamma_j)_{j \geq 1}$  be given. Then, the worst-case error  $e_{b^m,d,\alpha,\gamma}(P)$  of the QMC rule  $Q_{b^m,d}(\cdot; P)$  based on the digital  $(t, m, d)$ -net  $P = \{\mathbf{x}_0, \dots, \mathbf{x}_{b^m-1}\}$  with generating matrices  $C_1, \dots, C_d$  in the space  $W_{d,\gamma}^\alpha$  satisfies*

$$e_{b^m,d,\alpha,\gamma}(P) = \sum_{\mathbf{0} \neq \mathbf{k} \in \mathcal{D}} (r_{\alpha,\gamma}(\mathbf{k}))^{-1}. \tag{6}$$

**Proof** Recalling the definition of the worst-case error of the QMC rule  $Q_{b^m,d}(\cdot; P)$ , the combination of (4) and the definition of  $\|\cdot\|_{W_{d,\gamma}^\alpha}$  leads to the estimate

$$e_{b^m,d,\alpha,\gamma}(P) \leq \sup_{\substack{f \in W_{d,\gamma}^\alpha \\ \|f\|_{W_{d,\gamma}^\alpha} \leq 1}} \|f\|_{W_{d,\gamma}^\alpha} \sum_{\mathbf{0} \neq \mathbf{k} \in \mathcal{D}} (r_{\alpha,\gamma}(\mathbf{k}))^{-1} \leq \sum_{\mathbf{0} \neq \mathbf{k} \in \mathcal{D}} (r_{\alpha,\gamma}(\mathbf{k}))^{-1}.$$

Observing that the function  $f_0$  with Walsh coefficients  $\hat{f}_0(\mathbf{k}) = (r_{\alpha,\gamma}(\mathbf{k}))^{-1}$  has norm  $\|f_0\|_{W_{d,\gamma}^\alpha} = 1$  and that its integration error equals

$$Q_{b^m,d}(f_0, P) - I_d(f_0) = \sum_{\mathbf{0} \neq \mathbf{k} \in \mathcal{D}} (r_{\alpha,\gamma}(\mathbf{k}))^{-1},$$

we obtain that the previous upper bound is attained such that the claimed identity follows. □

**Remark 3** Returning to the alternative Walsh space  $\tilde{W}_{d,\gamma}^\alpha$  once again, it is known from [4] that the worst-case error in this space equals

$$\left( \sum_{\mathbf{0} \neq \mathbf{k} \in \mathcal{D}} (r_{\alpha,\gamma}(\mathbf{k}))^{-1} \right)^{1/2},$$

which is just the square root of the worst-case error in  $W_{d,\gamma}^\alpha$ , as outlined in Theorem 2. Therefore, we see that the worst-case errors in these Walsh spaces are intimately related to each other, and all results shown here for  $W_{d,\gamma}^\alpha$  immediately yield corresponding results for  $\tilde{W}_{d,\gamma}^\alpha$ .

### 2.2 Polynomial Lattice Rules

While Theorem 2 is a very useful result, the question of how to find and construct  $(t, m, d)$ -nets with a low integration error for practical purposes remains. One of the most powerful ways of obtaining nets is to consider a special case, namely so-called polynomial lattice point sets, as introduced by Niederreiter in [15]. The name “polynomial lattice point sets” is due to the fact that the structure of polynomial lattice point sets is similar to that of ordinary lattice point sets as introduced by Korobov [10] and Hlawka [9]. However, while lattice point sets are based on integer arithmetic, polynomial lattice point sets are obtained by using polynomial arithmetic over finite fields. We also point out that there are nowadays variants of polynomial lattice point sets which are especially suited for integrating functions with higher smoothness (see, e.g., [5]). However, we will not consider higher order polynomial lattices here, but restrict ourselves to the more classical construction scheme. We point out that polynomial lattice point sets are actually a special case of so-called digital  $(t, m, d)$ -nets, which can be constructed using generating matrices  $C_1, \dots, C_d$  over a finite field. For our purposes, though, it is more convenient to define these point sets in an alternative way. Before we give the precise definition, we need to introduce some notation.

As before and for the remainder of this article, we will assume that the base  $b$  is prime, even if not specifically mentioned. Let  $\mathbb{F}_b((x^{-1}))$  be the field of formal Laurent series over  $\mathbb{F}_b$  with elements of the form

$$L = \sum_{\ell=w}^{\infty} t_{\ell} x^{-\ell},$$

where  $w$  is an arbitrary integer and all  $t_{\ell} \in \mathbb{F}_b$ . We further denote by  $\mathbb{F}_b[x]$  the set of all polynomials over  $\mathbb{F}_b$  and define the map  $v_m : \mathbb{F}_b((x^{-1})) \rightarrow [0, 1)$  by

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

There is a close connection between the base  $b$  expansions of natural numbers and the polynomial ring  $\mathbb{F}_b[x]$ . For  $n \in \mathbb{N}_0$  with base  $b$  expansion  $n = n_0 + n_1 b + \dots + n_a b^a$ , we associate  $n$  with the polynomial

$$n(x) := \sum_{k=0}^a n_k x^k \in \mathbb{F}_b[x].$$

Furthermore, we define the greatest common divisor of two polynomials  $f, g \in \mathbb{F}_b[x]$ , denoted by  $\gcd(f, g)$ , as the monic polynomial of highest degree that divides both  $f$  and  $g$ .

The definition of a polynomial lattice point set is then given as follows. We note that here and in the following we consider the zero polynomial to have degree  $-\infty$ , hence, the case  $n = 0$  is included in the following definition. Furthermore, we remark

that the occurring fractions of the form  $q(x)/p(x)$  with  $p, q \in \mathbb{F}_b[x]$  are elements of the field  $\mathbb{F}_b((x^{-1}))$ .

**Definition 2** (*Polynomial lattice*) Let  $b$  be prime and let  $m, d \in \mathbb{N}$  be given. Furthermore, choose  $p \in \mathbb{F}_b[x]$  with  $\deg(p) = m$ , and let  $g_1, \dots, g_d \in \mathbb{F}_b[x]$ . Then, the point set  $P(\mathbf{g}, p)$ , defined as the collection of the  $b^m$  points

$$\mathbf{x}_n := \left( v_m \left( \frac{n(x) g_1(x)}{p(x)} \right), \dots, v_m \left( \frac{n(x) g_d(x)}{p(x)} \right) \right) \in [0, 1)^d$$

for  $n \in \mathbb{F}_b[x]$  with  $\deg(n) < m$ , is called a polynomial lattice point set (we sometimes also refer to the point set as polynomial lattice for short), where the vector  $\mathbf{g} = (g_1, \dots, g_d) \in (\mathbb{F}_b[x])^d$  is called the generating vector.

As pointed out above, due to the construction principle and the similarities to the construction of (rank-1) lattices,  $P(\mathbf{g}, p)$  is often called a (rank-1) polynomial lattice and a QMC rule using the point set  $P(\mathbf{g}, p)$  is referred to as a polynomial lattice rule (modulo  $p$ ). Furthermore, note that one can restrict the choice of the components  $g_j$  of  $\mathbf{g}$  to the sets

$$G_{b,m} := \{g \in \mathbb{F}_b[x] \mid \deg(g) < m\} \quad \text{or} \quad G_{b,m}^* := \{g \in \mathbb{F}_b[x] \setminus \{0\} \mid \deg(g) < m\}.$$

We also add that it is known from the literature on polynomial lattice point sets that it is desirable to have  $\gcd(g_j, p) = 1$  for the components  $g_j$  of  $\mathbf{g}$ , as this guarantees certain regularity properties.

For prime  $b$ , the generating matrices  $C_1, \dots, C_d \in \mathbb{F}_b^{m \times m}$  of a polynomial lattice point set  $P(\mathbf{g}, p)$  can be obtained from the generating vector  $\mathbf{g}$  and  $p$ , cf. [5, Theorem 10.5]. To this end, define for  $k \in \mathbb{N}_0$  with  $b$ -adic expansion  $k = \kappa_0 + \kappa_1 b + \dots + \kappa_{a-1} b^{a-1}$  the truncation map  $\text{tr}_m : \mathbb{N}_0 \rightarrow G_{b,m}$  via

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

where we consider  $\kappa_j$  as 0 if  $j \geq a$ . If we apply  $\text{tr}_m$  to a  $d$ -dimensional vector, we define its  $d$ -variate generalization  $\text{tr}_m(\mathbf{k})$  to be applied componentwise. It then follows that the dual net  $\mathcal{D}(\mathbf{g}, p)$  of a polynomial lattice with generating vector  $\mathbf{g}$ , modulus  $p$  with  $\deg(p) = m$ , and generating matrices  $C_1, \dots, C_d$  equals (see, e.g., [16, Lemma 4.40])

$$\mathcal{D}(\mathbf{g}, p) = \{\mathbf{k} \in \mathbb{N}_0^d \mid C_1^\top \mathbf{k}_1^m + \dots + C_d^\top \mathbf{k}_d^m = \bar{\mathbf{0}}\} = \{\mathbf{k} \in \mathbb{N}_0^d \mid \text{tr}_m(\mathbf{k}) \cdot \mathbf{g} \equiv 0 \pmod{p}\},$$

where for two vectors  $\mathbf{u}, \mathbf{v} \in (\mathbb{F}_b[x])^d$  we define the vector dot product  $\mathbf{u} \cdot \mathbf{v} = \sum_{j=1}^d u_j v_j$ . Furthermore, for a subset  $u \subseteq \{1:d\}$  we introduce the notation

$$\mathcal{D}_u = \mathcal{D}_u(\mathbf{g}, p) = \mathcal{D}_u(\mathbf{g}_u) := \{\mathbf{k}_u \in \mathbb{N}^{|u|} \mid \text{tr}_m(\mathbf{k}_u) \cdot \mathbf{g}_u \equiv 0 \pmod{p}\}.$$

Due to the obtained equivalence for the dual net of a polynomial lattice, the result in Theorem 2 also applies to polynomial lattice rules with  $\mathcal{D}(C_1, \dots, C_d)$  replaced by

$\mathcal{D}(\mathbf{g}, p)$ . Furthermore, we will henceforth denote the worst-case error of a QMC rule based on the polynomial lattice point set  $P(\mathbf{g}, p)$  in the space  $W_{d,\boldsymbol{\gamma}}^\alpha$  by  $e_{b^m,d,\alpha,\boldsymbol{\gamma}}(\mathbf{g})$ .

### 2.3 The Quality Measure

In this section, we introduce an alternative quality measure which, opposed to the worst-case error expression  $e_{b^m,d,\alpha,\boldsymbol{\gamma}}$  in (6), is independent of the parameter  $\alpha$ .

For  $\alpha \geq 1$ , given weight sequence  $\boldsymbol{\gamma} = (\gamma_j)_{j \geq 1}$ ,  $m \in \mathbb{N}$ , modulus  $p \in \mathbb{F}_b[x]$  with  $\deg(p) = m$ , and  $\mathbf{g} \in (\mathbb{F}_b[x])^d$ , we define the quantities

$$T_{\boldsymbol{\gamma}}(\mathbf{g}, p) := \sum_{\mathbf{0} \neq \mathbf{k} \in A_p(\mathbf{g})} (r_{1,\boldsymbol{\gamma}}(\mathbf{k}))^{-1}, \quad T_{\alpha,\boldsymbol{\gamma}}(\mathbf{g}, p) := \sum_{\mathbf{0} \neq \mathbf{k} \in A_p(\mathbf{g})} (r_{\alpha,\boldsymbol{\gamma}}(\mathbf{k}))^{-1} \quad (7)$$

with index set given by

$$A_p(\mathbf{g}) := \{\mathbf{k} \in \{0, 1, \dots, b^m - 1\}^d \mid \mathbf{k} \in \mathcal{D}(\mathbf{g}, p)\}.$$

Furthermore, for a subset  $\emptyset \neq \mathbf{u} \subseteq \{1:d\}$ , we introduce the sets

$$A_{\mathbf{u}} = A_{p,\mathbf{u}}(\mathbf{g}_{\mathbf{u}}) = A_{p,\mathbf{u}}(\mathbf{g}) := \{\mathbf{k}_{\mathbf{u}} \in \{0, 1, \dots, b^m - 1\}^{|\mathbf{u}|} \mid \mathbf{k}_{\mathbf{u}} \in \mathcal{D}_{\mathbf{u}}(\mathbf{g}, p)\},$$

$$A_{\mathbf{u}}^* = A_{p,\mathbf{u}}^*(\mathbf{g}_{\mathbf{u}}) = A_{p,\mathbf{u}}^*(\mathbf{g}) := \{\mathbf{k}_{\mathbf{u}} \in \{1, \dots, b^m - 1\}^{|\mathbf{u}|} \mid \mathbf{k}_{\mathbf{u}} \in \mathcal{D}_{\mathbf{u}}(\mathbf{g}, p)\},$$

and for a polynomial  $p \in \mathbb{F}_b[x]$  define the indicator function  $\delta_p : \mathbb{F}_b[x] \rightarrow \{0, 1\}$  by

$$\delta_p(q) := \begin{cases} 1, & \text{if } q \equiv 0 \pmod{p}, \\ 0, & \text{if } q \not\equiv 0 \pmod{p}. \end{cases}$$

In the following proposition, we estimate the difference between the worst-case error  $e_{b^m,d,\alpha,\boldsymbol{\gamma}}(\mathbf{g})$  and the truncated quality measure  $T_{\alpha,\boldsymbol{\gamma}}(\mathbf{g}, p)$  of a polynomial lattice rule with generator  $\mathbf{g}$  and modulus  $p \in \mathbb{F}_b[x]$  with  $\deg(p) = m$ .

**Proposition 1** *Let  $\boldsymbol{\gamma} = (\gamma_j)_{j \geq 1}$  be a sequence of positive weights, let  $p \in \mathbb{F}_b[x]$  with  $\deg(p) = m$ , and let  $\mathbf{g} = (g_1, \dots, g_d) \in G_{b,m}^d$  such that  $\gcd(g_j, p) = 1$  for all  $j = 1, \dots, d$ . Then, for any  $\alpha > 1$  and  $N = b^m$ , we have*

$$e_{b^m,d,\alpha,\boldsymbol{\gamma}}(\mathbf{g}) - T_{\alpha,\boldsymbol{\gamma}}(\mathbf{g}, p) \leq \frac{1}{N^\alpha} \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:d\}} \gamma_{\mathbf{u}}(2\mu_b(\alpha))^{|\mathbf{u}|}.$$

**Proof** For a non-empty subset  $\emptyset \neq \mathbf{u} \subseteq \{1:d\}$  and  $i \in \{1:d\}$ , we write for short  $\mathbf{k}_{\mathbf{u} \setminus \{i\}} \in \mathbb{N}_0^{|\mathbf{u}|-1}$  and  $\mathbf{g}_{\mathbf{u} \setminus \{i\}} \in G_{b,m}^{|\mathbf{u}|-1}$  to denote the projections on the components in  $\mathbf{u} \setminus \{i\}$ . The difference can then be rewritten as

$$e_{b^m, d, \alpha, \gamma}(\mathbf{g}) - T_{\alpha, \gamma}(\mathbf{g}, p) = \sum_{\emptyset \neq u \subseteq \{1:d\}} \left( \sum_{\mathbf{k}_u \in \mathcal{D}_u(\mathbf{g}_u)} (r_{\alpha, \gamma}(\mathbf{k}_u))^{-1} - \sum_{\mathbf{k}_u \in A_{p, u}^*(\mathbf{g}_u)} (r_{\alpha, \gamma}(\mathbf{k}_u))^{-1} \right),$$

motivating us to define the quantity

$$S_{\alpha, \gamma, u} := \sum_{\mathbf{k}_u \in \mathcal{D}_u(\mathbf{g}_u)} (r_{\alpha, \gamma}(\mathbf{k}_u))^{-1} - \sum_{\mathbf{k}_u \in A_{p, u}^*(\mathbf{g}_u)} (r_{\alpha, \gamma}(\mathbf{k}_u))^{-1}$$

for  $\emptyset \neq u \subseteq \{1 : d\}$ . In the following, we distinguish two cases.

**Case 1:** Suppose that  $|u| = 1$  such that  $u = \{j\}$  for some  $j \in \{1:d\}$ . Then, we have

$$\begin{aligned} S_{\alpha, \gamma, \{j\}} &= \sum_{\substack{k \in \mathbb{N} \\ \text{tr}_m(k) g_j \equiv 0 \pmod{p}}} (r_{\alpha, \gamma_j}(k))^{-1} - \sum_{\substack{k \in \{1, \dots, b^m - 1\} \\ \text{tr}_m(k) g_j \equiv 0 \pmod{p}}} (r_{\alpha, \gamma_j}(k))^{-1} \\ &= \sum_{\substack{k \geq b^m \\ \text{tr}_m(k) g_j \equiv 0 \pmod{p}}} (r_{\alpha, \gamma_j}(k))^{-1}. \end{aligned}$$

Note that  $\text{tr}_m(k) g_j \equiv 0 \pmod{p}$  if and only if there is a  $c \in \mathbb{F}_b[x]$  such that  $\text{tr}_m(k) g_j = cp$  and thus, since  $\text{gcd}(g_j, p) = 1$ , we have that  $\text{tr}_m(k) = ap$  for some  $a \in \mathbb{F}_b[x]$ . But  $\text{deg}(\text{tr}_m(k)) < m$  while  $\text{deg}(p) = m$ , which implies that  $\text{tr}_m(k) = 0$  and thus  $k = t b^m$  for some  $t \in \mathbb{N}$ . This yields

$$\begin{aligned} S_{\alpha, \gamma, \{j\}} &= \sum_{t=1}^{\infty} (r_{\alpha, \gamma_j}(t b^m))^{-1} = \gamma_j \sum_{t=1}^{\infty} b^{-\alpha \lfloor \log_b t b^m \rfloor} = \gamma_j \sum_{t=1}^{\infty} b^{-\alpha \lfloor m + \log_b t \rfloor} \\ &= \gamma_j \sum_{t=1}^{\infty} b^{-\alpha m} b^{-\alpha \lfloor \log_b t \rfloor} = \frac{\gamma_j}{b^{\alpha m}} \sum_{t=1}^{\infty} b^{-\alpha \psi_b(t)} = \gamma_j \frac{\mu_b(\alpha)}{b^{\alpha m}}. \end{aligned}$$

**Case 2:** Suppose that  $|u| \geq 2$ . In this case, we find that

$$S_{\alpha, \gamma, u} \leq \sum_{i \in u} \sum_{\mathbf{k}_{u \setminus \{i\}} \in \mathbb{N}^{|u|-1}} \sum_{k_i \geq b^m} \frac{\delta_p(\text{tr}_m(k_i) g_i + \text{tr}_m(\mathbf{k}_{u \setminus \{i\}}) \cdot \mathbf{g}_{u \setminus \{i\}})}{r_{\alpha, \gamma}(\mathbf{k}_u)}.$$

Then, for  $\mathbf{k}_{u \setminus \{i\}} \in \mathbb{N}^{|\mathbf{u}|-1}$ , we write  $q = \text{tr}_m(\mathbf{k}_{u \setminus \{i\}}) \cdot \mathbf{g}_{u \setminus \{i\}}$  and estimate the expression

$$\begin{aligned} & \sum_{k_i \geq b^m} \frac{\delta_p(\text{tr}_m(k_i)g_i + q)}{r_{\alpha, \mathbf{y}}(\mathbf{k}_u)} \\ &= \gamma_u \sum_{k_i \geq b^m} \frac{\delta_p(\text{tr}_m(k_i)g_i + q)}{\prod_{j \in u} b^{\alpha \lfloor \log_b k_j \rfloor}} \\ &= \gamma_u \prod_{\substack{j \in u \\ j \neq i}} b^{-\alpha \lfloor \log_b k_j \rfloor} \sum_{k_i \geq b^m} \frac{\delta_p(\text{tr}_m(k_i)g_i + q)}{b^{\alpha \lfloor \log_b k_i \rfloor}} \\ &= \gamma_u \prod_{\substack{j \in u \\ j \neq i}} b^{-\alpha \lfloor \log_b k_j \rfloor} \sum_{t=1}^{\infty} \sum_{k_i = tb^m}^{(t+1)b^m - 1} \frac{\delta_p(\text{tr}_m(k_i)g_i + q)}{b^{\alpha \lfloor \log_b k_i \rfloor}} \\ &\leq \gamma_u \prod_{\substack{j \in u \\ j \neq i}} b^{-\alpha \lfloor \log_b k_j \rfloor} \sum_{t=1}^{\infty} b^{-\alpha \lfloor \log_b tb^m \rfloor} \underbrace{\sum_{k_i = tb^m}^{(t+1)b^m - 1} \delta_p(\text{tr}_m(k_i)g_i + q)}_{=1} \\ &= \gamma_u \prod_{\substack{j \in u \\ j \neq i}} b^{-\alpha \lfloor \log_b k_j \rfloor} \sum_{t=1}^{\infty} b^{-\alpha \lfloor m + \log_b t \rfloor} = \gamma_u \frac{\mu_b(\alpha)}{b^{\alpha m}} \prod_{\substack{j \in u \\ j \neq i}} b^{-\alpha \lfloor \log_b k_j \rfloor}, \end{aligned}$$

where the penultimate equality follows since if  $\text{gcd}(g_i, p) = 1$  then for each  $t$  and each  $q \in \mathbb{F}_p[x]$  there exists exactly one  $k \in \{tb^m, \dots, (t + 1)b^m - 1\}$  such that  $\text{tr}_m(k)g_i + q \equiv 0 \pmod{p}$ .

Hence, we can estimate  $S_{\alpha, \mathbf{y}, u}$ , for  $|\mathbf{u}| \geq 2$ , by

$$\begin{aligned} S_{\alpha, \mathbf{y}, u} &\leq \sum_{i \in u} \sum_{\mathbf{k}_{u \setminus \{i\}} \in \mathbb{N}^{|\mathbf{u}|-1}} \gamma_u \frac{\mu_b(\alpha)}{b^{\alpha m}} \prod_{\substack{j \in u \\ j \neq i}} b^{-\alpha \lfloor \log_b k_j \rfloor} \\ &= \gamma_u \frac{\mu_b(\alpha)}{b^{\alpha m}} \sum_{i \in u} \left( \sum_{k=1}^{\infty} b^{-\alpha \lfloor \log_b k \rfloor} \right)^{|\mathbf{u}|-1} \\ &= \gamma_u \frac{\mu_b(\alpha)}{b^{\alpha m}} \sum_{i \in u} \mu_b(\alpha)^{|\mathbf{u}|-1} = \gamma_u \frac{\mu_b(\alpha)^{|\mathbf{u}|}}{b^{\alpha m}} |\mathbf{u}| \leq \gamma_u \frac{1}{N^\alpha} (2\mu_b(\alpha))^{|\mathbf{u}|}. \end{aligned}$$

In summary, we obtain, using the results for both cases from above,

$$\sum_{\emptyset \neq u \subseteq \{1:d\}} S_{\alpha, \mathbf{y}, u} \leq \frac{1}{N^\alpha} \sum_{\emptyset \neq u \subseteq \{1:d\}} \gamma_u (2\mu_b(\alpha))^{|\mathbf{u}|},$$

which is the claimed upper estimate. □

Based on the previous result, it is straightforward to show the existence of good polynomial lattice rules with respect to the worst-case error in the weighted Walsh space, if one assumes the modulus  $p$  to be irreducible. We omit the proof, which uses standard methods, and instead refer the reader to [7, Theorems 9 and 10], where an analogous result for lattice rules has been proven.

**Theorem 3** *Let  $p \in \mathbb{F}_b[x]$  be an irreducible polynomial with  $\deg(p) = m$ , let  $N = b^m$ , and let  $\boldsymbol{\gamma} = (\gamma_j)_{j \geq 1}$  be positive weights. Then, there exists a  $\mathbf{g} \in G_{b,m}^d$  such that, for all  $\alpha > 1$ , the worst-case error  $e_{b^m,d,\alpha,\boldsymbol{\gamma}}(\mathbf{g})$  satisfies*

$$e_{b^m,d,\alpha,\boldsymbol{\gamma}}(\mathbf{g}) \leq \frac{1}{N^\alpha} \left( \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:d\}} \gamma_{\mathbf{u}} (2\mu_b(\alpha))^{|\mathbf{u}|} + \left( \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:d\}} \gamma_{\mathbf{u}}^{1/\alpha} (m(b-1))^{|\mathbf{u}|} \right)^\alpha \right).$$

Even though the result in Theorem 3 assures us that there always exist generating vectors of polynomial lattice point sets which are in a certain sense good, the result is not constructive. The road which we will take in the present paper is slightly different. Instead of assuming an irreducible modulus  $p$ , we will assume that  $p$  has the special form  $p(x) = p_m(x) = x^m$  and show a constructive approach to find generating vectors of good polynomial lattice rules. This will be the main result of our paper, which is stated in Theorem 8.

### 3 The CBC-DBD Construction for Polynomial Lattice Rules

In this section, we formulate and analyze a method for the construction of good polynomial lattice rules. In contrast to the existence result in Theorem 3, our construction method yields polynomial lattice rules with modulus  $p(x) = x^m$ . At first, we prove some auxiliary statements which will be needed in the further analysis.

#### 3.1 Preliminary Results

We consider the following Walsh series for  $x \in (0, 1)$ , based on the decay function  $r_1$ ,

$$\sum_{k=0}^{\infty} \frac{\text{wal}_k(x)}{r_1(k)} = 1 + \sum_{k=1}^{\infty} \frac{e^{2\pi i(\kappa_0 \xi_1 + \kappa_1 \xi_2 + \dots) / b}}{b^{\lfloor \log_b(k) \rfloor}},$$

which, as we will see, is closely related to our quality criterion  $T_{\boldsymbol{\gamma}}$  introduced in (7). To this end, we define, for  $n \in \mathbb{N}$ , the  $n$ -th Walsh-Dirichlet kernel by

$$D_n(x) = \sum_{k=0}^{n-1} \text{wal}_k(x).$$



From [5, Lemma A.17], it then follows that, for  $x \in (0, 1)$ ,

$$D_{b^t}(x) = \begin{cases} b^t, & \text{if } x \in (0, \frac{1}{b^t}), \\ 0, & \text{if } x \in [\frac{1}{b^t}, 1). \end{cases} \tag{8}$$

We can then prove the following identity.

**Lemma 1** *For base  $b \geq 2$ , the Walsh series of  $-(b - 1)(\lfloor \log_b(x) \rfloor + 1)$  equals, pointwise for  $x \in (0, 1)$ ,*

$$-(b - 1)(\lfloor \log_b(x) \rfloor + 1) = 1 + \sum_{k=1}^{\infty} \frac{e^{2\pi i(\kappa_0 \xi_1 + \kappa_1 \xi_2 + \dots) / b}}{b^{\lfloor \log_b(k) \rfloor}} = \sum_{k=0}^{\infty} \frac{\text{wal}_k(x)}{r_1(k)}.$$

**Proof** Using the definition of the Walsh-Dirichlet kernel, we obtain

$$\sum_{k=1}^{\infty} \frac{\text{wal}_k(x)}{b^{\lfloor \log_b(k) \rfloor}} = \sum_{t=1}^{\infty} \sum_{k=b^{t-1}}^{b^t-1} \frac{\text{wal}_k(x)}{b^{t-1}} = \sum_{t=1}^{\infty} \frac{D_{b^t}(x) - D_{b^{t-1}}(x)}{b^{t-1}},$$

and from (8) we find that for  $t \geq 1$  we have

$$D_{b^t}(x) - D_{b^{t-1}}(x) = \begin{cases} (b - 1)b^{t-1}, & \text{if } x \in (0, \frac{1}{b^t}), \\ -b^{t-1}, & \text{if } x \in [\frac{1}{b^t}, \frac{1}{b^{t-1}}), \\ 0, & \text{if } x \in [\frac{1}{b^{t-1}}, 1). \end{cases} \tag{9}$$

Applied inductively, the relation in (9) yields that for  $x \in [\frac{1}{b^t}, \frac{1}{b^{t-1}})$  we have

$$\sum_{\ell=1}^{\infty} \frac{D_{b^\ell}(x) - D_{b^{\ell-1}}(x)}{b^{\ell-1}} = \sum_{\ell=1}^{t-1} \frac{(b - 1)b^{\ell-1}}{b^{\ell-1}} - \frac{b^{t-1}}{b^{t-1}} = (t - 1)(b - 1) - 1$$

for all  $t \geq 1$ , which is equivalent to

$$1 + \sum_{k=1}^{\infty} \frac{\text{wal}_k(x)}{b^{\lfloor \log_b(k) \rfloor}} = (b - 1)(t - 1) = -(b - 1)(-t + 1) = -(b - 1)(\lfloor \log_b(x) \rfloor + 1)$$

for  $x \in [\frac{1}{b^t}, \frac{1}{b^{t-1}})$  and for all  $t \in \mathbb{N}$ . This proves the claimed identity. □

Based on the previous result in Lemma 1, we show that the function  $-(b - 1)(\lfloor \log_b(x) \rfloor + 1)$  can be written in terms of its truncated Walsh series with uniformly bounded remainder term.

**Lemma 2** Let  $N = b^m$  with  $m \in \mathbb{N}$  and base  $b \geq 2$ . Then, for any  $x \in (0, 1)$  there exists a  $\tau = \tau(x) \in \mathbb{R}$  with  $|\tau(x)| < \frac{b}{b-1}$  such that

$$-(b - 1)(\lfloor \log_b(x) \rfloor + 1) = \sum_{k=0}^{N-1} \frac{\text{wal}_k(x)}{r_1(k)} + \frac{\tau(x)}{Nx}. \tag{10}$$

**Proof** The expansion in Lemma 1 allows us to write

$$-(b - 1)(\lfloor \log_b(x) \rfloor + 1) = \sum_{k=0}^{N-1} \frac{\text{wal}_k(x)}{r_1(k)} + R_N(x),$$

where the remainder  $R_N(x)$  has the form

$$R_N(x) = \sum_{k=b^m}^{\infty} \frac{\text{wal}_k(x)}{r_1(k)} = \sum_{k=b^m}^{\infty} \frac{\text{wal}_k(x)}{b^{\lfloor \log_b(k) \rfloor}} = \sum_{t=m}^{\infty} \frac{D_{b^{t+1}}(x) - D_{b^t}(x)}{b^t}.$$

From (9), we then see that the following inequality holds,

$$|D_{b^{t+1}}(x) - D_{b^t}(x)| < \frac{1}{x}, \quad t \in \mathbb{N}, \quad x \in (0, 1),$$

and thus, we obtain

$$|R_N(x)| = \left| \sum_{t=m}^{\infty} \frac{D_{b^{t+1}}(x) - D_{b^t}(x)}{b^t} \right| < \frac{1}{x} \sum_{t=m}^{\infty} \frac{1}{b^t} = \frac{b}{(b - 1)b^m} \frac{1}{x} = \frac{b}{(b - 1)Nx},$$

which implies the existence of a  $\tau(x) \in \mathbb{R}$  with  $|\tau(x)| < \frac{b}{b-1}$  such that the identity (10) holds. □

**Remark 4** Using a more involved argument, the result in Lemma 2 can also be extended to general  $N \in \mathbb{N}$ . In particular, we obtain that for any  $x \in (0, 1)$  there exists a  $\tau = \tau(x) \in \mathbb{R}$  such that

$$-(b - 1)(\lfloor \log_b(x) \rfloor + 1) = \sum_{k=0}^{N-1} \frac{\text{wal}_k(x)}{r_1(k)} + \frac{\tau}{Nx}$$

with  $|\tau| < b \left( \frac{1}{b-1} + 2 \right)$  for  $b = 2$  and with  $|\tau| < b \left( \frac{1}{b-1} + 2b \right)$  for  $b > 2$ .

We will also make use of the following lemma, which was proved in [7, Lemma 3].

**Lemma 3** For  $j \in \{1:d\}$ , let  $u_j, w_j$ , and  $\rho_j$  be real numbers satisfying

$$(a) \quad u_j = w_j + \rho_j, \quad (b) \quad |u_j| \leq \bar{u}_j, \quad (c) \quad \bar{u}_j \geq 1,$$

for all  $j \in \{1:d\}$ . Then, for any subset  $\emptyset \neq u \subseteq \{1:d\}$  there exists a  $\theta_u$  with  $|\theta_u| \leq 1$  such that

$$\prod_{j \in u} u_j = \prod_{j \in u} w_j + \theta_u \left( \prod_{j \in u} (\bar{u}_j + |\rho_j|) \right) \sum_{j \in u} |\rho_j|.$$

Furthermore, we recall the character property of Walsh functions for polynomial lattice rules with prime base  $b$ . Let  $P(\mathbf{g}, p) = \{\mathbf{x}_0, \dots, \mathbf{x}_{b^m-1}\}$  be a polynomial lattice with generating vector  $\mathbf{g} \in (\mathbb{F}_b[x])^d$  and modulus  $p \in \mathbb{F}_b[x]$  with  $\deg(p) = m$ . Then, for any integer vector  $\mathbf{k} \in \mathbb{N}_0^d$  the following identity holds,

$$\frac{1}{b^m} \sum_{n=0}^{b^m-1} \text{wal}_{\mathbf{k}}(\mathbf{x}_n) = \delta_p(\text{tr}_m(\mathbf{k}) \cdot \mathbf{g}) = \begin{cases} 1, & \text{if } \text{tr}_m(\mathbf{k}) \cdot \mathbf{g} \equiv 0 \pmod{p}, \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

We remark that an analogous result to (11) also holds if we only consider projections of the polynomial lattice and the generating vector onto a non-empty subset of  $\{1:d\}$ , as also the projection of a polynomial lattice is a polynomial lattice that is generated by the corresponding projection of the generating vector.

We now state an auxiliary result that will be useful at several instances in this paper.

**Lemma 4** *Let  $P(\mathbf{g}, p)$  be a polynomial lattice with modulus  $p \in \mathbb{F}_b[x]$  with  $\deg(p) = m$  and generating vector  $\mathbf{g} = (g_1, \dots, g_d) \in (\mathbb{F}_b[x])^d$  such that  $\gcd(g_j, p) = 1$  for  $1 \leq j \leq d$ . Then, each one-dimensional projection of  $P(\mathbf{g}, p)$  is the full grid*

$$\left\{ 0, \frac{1}{b^m}, \dots, \frac{b^m - 1}{b^m} \right\},$$

and in particular the projection of the point with index 0 is always 0.

**Proof** The result follows from Definition 2 and [5, Remark 10.3]. □

Additionally, we will need the following result.

**Lemma 5** *Let  $P(\mathbf{g}, p) = \{\mathbf{x}_0, \dots, \mathbf{x}_{b^m-1}\}$  be a polynomial lattice point set with modulus  $p \in \mathbb{F}_b[x]$  with  $\deg(p) = m$  and generating vector  $\mathbf{g} \in (\mathbb{F}_b[x])^d$  such that  $\gcd(g_j, p) = 1$  for  $1 \leq j \leq d$ . Furthermore, let  $m \geq 4$ . For a point  $\mathbf{x}_n$  with  $n \in \{0, 1, \dots, b^m - 1\}$ , we denote its coordinates via  $\mathbf{x}_n = (x_{n,1}, \dots, x_{n,d})$ . Then, for any  $j \in \{1, \dots, d\}$ , it is true that*

$$\frac{1}{b^m} \sum_{n=1}^{b^m-1} \frac{1}{x_{n,j}} < 1 + m \ln b \leq m(b - 1).$$

**Proof** We recall that the point set  $P(\mathbf{g}, p)$  is defined as the collection of the  $b^m$  points of the form

$$\mathbf{x}_n = \left( v_m \left( \frac{n(x) g_1(x)}{p(x)} \right), \dots, v_m \left( \frac{n(x) g_d(x)}{p(x)} \right) \right)$$

for  $n \in \mathbb{F}_b[x]$  with  $\deg(n) < m$ . Due to Lemma 4, we know that  $\{x_{1,j}, \dots, x_{b^m-1,j}\}$  equals the set  $\left\{\frac{1}{b^m}, \dots, \frac{b^m-1}{b^m}\right\}$  for each  $j \in \{1, \dots, d\}$ . Thus, we can estimate

$$\begin{aligned} \frac{1}{b^m} \sum_{n=1}^{b^m-1} \frac{1}{x_{n,j}} &= \frac{1}{b^m} \sum_{n=1}^{b^m-1} \frac{b^m}{n} = \sum_{n=1}^{b^m-1} \frac{1}{n} \leq 1 + \int_1^{b^m-1} \frac{1}{x} \, dx \\ &= 1 + \ln(b^m - 1) < 1 + \ln(b^m) = 1 + m \ln b \leq m(b - 1), \end{aligned}$$

which yields the claimed result, where the last estimate follows from the assumption  $m \geq 4$ . □

### 3.2 The CBC-DBD Construction Algorithm

We are now ready to study the component-by-component digit-by-digit (CBC-DBD) construction for polynomial lattice rules, see also [7], where such an algorithm was analyzed for ordinary lattice rules. In particular, we will assume throughout this section that our modulus polynomial is of the form  $p_m(x) = x^m$  for  $m \in \mathbb{N}$ .

Concerning the weights, the algorithm can, as indicated in our main result (Theorem 8), be run with respect to the weights  $\boldsymbol{\gamma}^{1/\alpha} = (\gamma_j^{1/\alpha})_{j \geq 1}$  to obtain a polynomial lattice rule that yields a low worst-case error in the Walsh space  $W_{d,\boldsymbol{\gamma}^\alpha}$ , or, alternatively, with respect to the weights  $\boldsymbol{\gamma}$  to obtain good polynomial lattice rules in the space  $W_{d,\boldsymbol{\gamma}^\alpha}$ . In the latter case, the construction algorithm is independent of the smoothness parameter  $\alpha$  and we obtain worst-case error bounds that hold for all  $\alpha > 1$  simultaneously.

In order to avoid confusion, we will therefore denote the weights in this section by  $\boldsymbol{\eta}$  instead of  $\boldsymbol{\gamma}$  and outline the algorithm based on  $\boldsymbol{\eta}$ . In Theorem 8, we will then choose  $\boldsymbol{\eta}$  equal to  $\boldsymbol{\gamma}^{1/\alpha}$  or  $\boldsymbol{\gamma}$ , respectively. For technical reasons (see Remark 5), it will be necessary to assume that the positive weights  $\boldsymbol{\eta}$  are of product structure, that is,

$$\eta_u = \prod_{j \in u} \eta_j$$

for  $u \subseteq \{1:d\}$ , with a sequence of positive reals  $(\eta_j)_{j \geq 1}$ . However, we point out that the following theorem, which is crucial for the proposed construction method, also holds for general weights  $\boldsymbol{\eta} = (\eta_u)_{u \subseteq \{1:d\}}$ .

**Theorem 4** *Let  $b$  be prime, let  $m, d \in \mathbb{N}$  with  $m \geq 4$ , let  $p_m(x) = x^m \in \mathbb{F}_b[x]$ , and let  $\boldsymbol{\eta} = (\eta_u)_{u \subseteq \{1:d\}}$  be positive weights with  $\eta_\emptyset = 1$ . Furthermore, let  $\mathbf{g} = (g_1, \dots, g_d) \in (\mathbb{F}_b[x])^d$  with  $\deg(g_j) < m$  and  $\gcd(g_j, p_m) = 1$  for  $1 \leq j \leq d$ . Then,*

$$\begin{aligned} T_{\boldsymbol{\eta}}(\mathbf{g}, p_m) &\leq \frac{1}{b^m} H_{d,m,\boldsymbol{\eta}}(\mathbf{g}) - \sum_{\emptyset \neq u \subseteq \{1:d\}} \eta_u + \sum_{\emptyset \neq u \subseteq \{1:d\}} \frac{\eta_u}{b^m} ((b-1)m+1)^{|u|} \\ &\quad + \sum_{\emptyset \neq u \subseteq \{1:d\}} \frac{\eta_u}{b^m} (bm^{|u|}) \left( (b-1)m + \frac{b}{b-1} \right)^{|u|}, \end{aligned} \tag{12}$$

where we define the function  $H_{d,m,\eta} : (\mathbb{F}_b[x])^d \rightarrow \mathbb{R}$  as

$$H_{d,m,\eta}(\mathbf{g}) := \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:d\}} \eta_{\mathbf{u}} (1-b)^{|\mathbf{u}|} \sum_{n=1}^{b^m-1} \prod_{j \in \mathbf{u}} \left( \left\lfloor \log_b \left( v_m \left( \frac{n(x) g_j(x)}{x^m} \right) \right) \right\rfloor + 1 \right). \tag{13}$$

**Proof** We use the character property of Walsh functions in (11) to rewrite  $T_{\eta}(\mathbf{g}, p_m)$  with the help of the identity in Lemma 2. First, we recall that for  $k \in \mathbb{N}_0$  we have

$$r_1(k) = r_1(b, k) = \begin{cases} 1, & \text{for } k = 0, \\ b^{\lfloor \log_b(k) \rfloor}, & \text{for } k \neq 0. \end{cases}$$

Using this definition, we obtain that

$$\begin{aligned} & T_{\eta}(\mathbf{g}, p_m) \\ &= \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:d\}} \eta_{\mathbf{u}} \sum_{\mathbf{k}_{\mathbf{u}} \in \{1, \dots, b^m-1\}^{|\mathbf{u}|}} \frac{\delta_{p_m}(\text{tr}_m(\mathbf{k}_{\mathbf{u}}) \cdot \mathbf{g}_{\mathbf{u}})}{\prod_{j \in \mathbf{u}} r_1(k_j)} \\ &\leq \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:d\}} \eta_{\mathbf{u}} \sum_{\mathbf{0} \neq \mathbf{k}_{\mathbf{u}} \in \{0, \dots, b^m-1\}^{|\mathbf{u}|}} \frac{\delta_{p_m}(\text{tr}_m(\mathbf{k}_{\mathbf{u}}) \cdot \mathbf{g}_{\mathbf{u}})}{\prod_{j \in \mathbf{u}} r_1(k_j)} \\ &= \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:d\}} \frac{\eta_{\mathbf{u}}}{b^m} \sum_{n=0}^{b^m-1} \left[ \sum_{\mathbf{k}_{\mathbf{u}} \in \{0, 1, \dots, b^m-1\}^{|\mathbf{u}|}} \frac{\text{wal}_{\mathbf{k}_{\mathbf{u}}}(\mathbf{x}_{n,\mathbf{u}})}{\prod_{j \in \mathbf{u}} r_1(k_j)} - 1 \right] \\ &= \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:d\}} \frac{\eta_{\mathbf{u}}}{b^m} \left[ \sum_{\mathbf{k}_{\mathbf{u}} \in \{0, \dots, b^m-1\}^{|\mathbf{u}|}} \frac{1}{\prod_{j \in \mathbf{u}} r_1(k_j)} \right. \\ &\quad \left. + \sum_{n=1}^{b^m-1} \prod_{j \in \mathbf{u}} \left( 1 + \sum_{k=1}^{b^m-1} \frac{\text{wal}_k(x_{n,j})}{b^{\lfloor \log_b(k) \rfloor}} \right) \right] - \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:d\}} \eta_{\mathbf{u}} \\ &= \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:d\}} \frac{\eta_{\mathbf{u}}}{b^m} \left[ \sum_{\mathbf{k}_{\mathbf{u}} \in \{0, \dots, b^m-1\}^{|\mathbf{u}|}} \frac{1}{\prod_{j \in \mathbf{u}} r_1(k_j)} \right. \\ &\quad \left. + \sum_{n=1}^{b^m-1} \left[ \prod_{j \in \mathbf{u}} w_j(n) - \prod_{j \in \mathbf{u}} u_j(n) + \prod_{j \in \mathbf{u}} u_j(n) \right] \right] - \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:d\}} \eta_{\mathbf{u}} \\ &= \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:d\}} \frac{\eta_{\mathbf{u}}}{b^m} \sum_{\mathbf{k}_{\mathbf{u}} \in \{0, 1, \dots, b^m-1\}^{|\mathbf{u}|}} \frac{1}{\prod_{j \in \mathbf{u}} r_1(k_j)} + \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:d\}} \frac{\eta_{\mathbf{u}}}{b^m} \sum_{n=1}^{b^m-1} \prod_{j \in \mathbf{u}} u_j(n) \\ &\quad - \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:d\}} \frac{\eta_{\mathbf{u}}}{b^m} \sum_{n=1}^{b^m-1} \theta_{\mathbf{u}}(n) \left( \prod_{j \in \mathbf{u}} (\bar{u}_j + |\rho_j(n)|) \right) \sum_{j \in \mathbf{u}} |\rho_j(n)| - \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:d\}} \eta_{\mathbf{u}}, \tag{14} \end{aligned}$$

where we used Lemma 3 with

$$u_j = u_j(n) := -(b - 1)(\lfloor \log_b(x_{n,j}) \rfloor + 1), \quad \bar{u}_j = \bar{u}_j(n) := (b - 1)m,$$

$$w_j = w_j(n) := 1 + \sum_{k=1}^{b^m-1} \frac{\text{wal}_k(x_{n,j})}{b^{\lfloor \log_b(k) \rfloor}}, \quad \rho_j = \rho_j(n) := \frac{\tau_j(n)}{x_{n,j} b^m},$$

and all  $|\theta_u(n)| \leq 1$  and  $|\tau_j(n)| < \frac{b}{b-1}$ . Due to Lemma 2, Condition (a) of Lemma 3 is fulfilled. Furthermore, we see that by Lemma 4 we have for each  $j \in \{1:d\}$  that  $x_{n,j} \geq b^{-m}$  for every  $1 \leq n < b^m$ , and hence

$$|u_j(n)| \leq -(b - 1)(\lfloor \log_b(b^{-m}) \rfloor + 1) = -(b - 1)(-m + 1) < (b - 1)m = \bar{u}_j$$

with  $\bar{u}_j \geq 1$  such that also Conditions (b) and (c) of Lemma 3 are satisfied.

By simple calculations, the first sum in (14) can be shown to equal

$$\begin{aligned} & \sum_{\emptyset \neq u \subseteq \{1:d\}} \frac{\eta_u}{b^m} \sum_{k_u \in \{0,1,\dots,b^m-1\}^{|u|}} \frac{1}{\prod_{j \in u} r_1(k_j)} \\ &= \frac{1}{b^m} \sum_{\emptyset \neq u \subseteq \{1:d\}} \eta_u \left( 1 + \sum_{k=1}^{b^m-1} \frac{1}{b^{\lfloor \log_b(k) \rfloor}} \right)^{|u|} \\ &= \frac{1}{b^m} \sum_{\emptyset \neq u \subseteq \{1:d\}} \eta_u \left( 1 + \sum_{t=0}^{m-1} \sum_{k=b^t}^{b^{t+1}-1} \frac{1}{b^{\lfloor \log_b(k) \rfloor}} \right)^{|u|} \\ &= \frac{1}{b^m} \sum_{\emptyset \neq u \subseteq \{1:d\}} \eta_u ((b - 1)m + 1)^{|u|}, \end{aligned}$$

while the third sum in (14) can be bounded by

$$\begin{aligned} & - \sum_{\emptyset \neq u \subseteq \{1:d\}} \frac{\eta_u}{b^m} \sum_{n=1}^{b^m-1} \theta_u(n) \left( \prod_{j \in u} (\bar{u}_j + |\rho_j(n)|) \right) \sum_{j \in u} |\rho_j(n)| \\ &= - \sum_{\emptyset \neq u \subseteq \{1:d\}} \frac{\eta_u}{b^m} \sum_{n=1}^{b^m-1} \theta_u(n) \left( \prod_{j \in u} \left( (b - 1)m + \frac{|\tau_j(n)|}{x_{n,j} b^m} \right) \right) \sum_{j \in u} \frac{|\tau_j(n)|}{x_{n,j} b^m} \\ &\leq \sum_{\emptyset \neq u \subseteq \{1:d\}} \frac{\eta_u}{b^m} \sum_{n=1}^{b^m-1} |\theta_u(n)| \left( \prod_{j \in u} \left( (b - 1)m + \frac{b}{b - 1} \right) \right) \sum_{j \in u} \frac{|\tau_j(n)|}{x_{n,j} b^m} \\ &\leq \sum_{\emptyset \neq u \subseteq \{1:d\}} \frac{\eta_u}{b^m} \left( \prod_{j \in u} \left( (b - 1)m + \frac{b}{b - 1} \right) \right) \sum_{j \in u} \sum_{n=1}^{b^m-1} \frac{b}{(b - 1)b^m} \frac{1}{x_{n,j}} \end{aligned}$$

$$\begin{aligned} &\leq \sum_{\emptyset \neq u \subseteq \{1:d\}} \frac{\eta_u}{b^m} \left( \prod_{j \in u} \left( (b-1)m + \frac{b}{b-1} \right) \right) \frac{b}{b-1} \sum_{j \in u} m(b-1) \\ &= \frac{1}{b^m} \sum_{\emptyset \neq u \subseteq \{1:d\}} \eta_u (b^m |u|) \left( (b-1)m + \frac{b}{b-1} \right)^{|u|}, \end{aligned}$$

where we used Lemma 5 and the fact that  $x_{n,j} \geq b^{-m}$  for each  $j$  and all  $1 \leq n < b^m$ . Combining these results with (14) yields the claimed result.  $\square$

We observe that the latter terms in (12) are of order  $\mathcal{O}(b^{-m})$  and the implied constants can be bounded independently of the dimension provided that certain conditions on the weight sequence  $\eta$  hold (see the proof of Theorem 8 for further details). As this is the desired order of  $T_\eta(\mathbf{g}, p_m)$ , Theorem 4 implies that it essentially suffices to find a generating vector  $\mathbf{g} \in (\mathbb{F}_b[x])^d$  such that  $H_{d,m,\eta}(\mathbf{g})$  is small, which then implies that also a good bound on  $T_\eta(\mathbf{g}, p_m)$  holds. We will therefore consider the quantity  $H_{d,m,\eta}$  as a search criterion for good generating vectors.

At first, we prove the following result which will be needed in the further analysis and remind the reader that by  $p_m$  we denote the polynomial  $p_m \in \mathbb{F}_b[x]$  with  $p_m(x) = x^m$  for  $m \in \mathbb{N}$ .

**Lemma 6** *Let a prime  $b$ , an integer  $t \geq 2$ , and polynomials  $\ell, q \in \mathbb{F}_b[x]$  with  $\gcd(\ell, p_1) = \gcd(q, p_1) = 1$  be given. Then, the following identity holds:*

$$\sum_{g \in \mathbb{F}_b} \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x)(q(x) + x^{t-1}g)}{x^t} \right) \right) \right\rfloor + 1 \right) = \left\lfloor \log_b \left( v_{t-1} \left( \frac{\ell(x)q(x)}{x^{t-1}} \right) \right) \right\rfloor.$$

**Proof** Assume that the product of the polynomials  $\ell$  and  $q$  is given by

$$\ell(x)q(x) = \sum_{i=0}^r a_i x^i \quad \text{with } a_0, a_r \neq 0.$$

Let, furthermore,

$$\ell(x) = \sum_{k=0}^v \ell_k x^k,$$

where we note that  $v \leq r$ . Hence, we obtain that for  $g \in \mathbb{F}_b$

$$\frac{\ell(x)(q(x) + x^{t-1}g)}{x^t} = \sum_{i=t}^r a_i x^{i-t} + \sum_{k=1}^v \ell_k g x^{k-1} + (a_{t-1} + \ell_0 g)x^{-1} + \sum_{i=0}^{t-2} a_i x^{i-t}$$

and thus, we have that if  $a_{t-1} + \ell_0 g \not\equiv 0 \pmod{b}$ , then

$$\begin{aligned} & \left\lceil \log_b \left( v_t \left( \frac{\ell(x)(q(x) + x^{t-1}g)}{x^t} \right) \right) \right\rceil + 1 \\ &= \left\lceil \log_b \left( \frac{a_{t-1} + \ell_0 g}{b} + \sum_{i=0}^{t-2} a_i b^{i-t} \right) \right\rceil + 1 = -1 + 1 = 0. \end{aligned}$$

Otherwise, if  $a_{t-1} + \ell_0 g \equiv 0 \pmod{b}$ , then

$$\begin{aligned} v_t \left( \frac{\ell(x)(q(x) + x^{t-1}g)}{x^t} \right) &= \sum_{i=0}^{t-2} a_i b^{i-t} = \sum_{i=2}^t a_{t-i} b^{-i} \\ &= \frac{1}{b} \left( \sum_{i=1}^{t-1} a_{t-i-1} b^{-i} \right) = \frac{1}{b} v_{t-1} \left( \frac{\ell(x)q(x)}{x^{t-1}} \right), \end{aligned}$$

and therefore

$$\begin{aligned} \left\lceil \log_b \left( v_t \left( \frac{\ell(x)(q(x) + x^{t-1}g)}{x^t} \right) \right) \right\rceil + 1 &= \left\lceil \log_b \left( \frac{1}{b} v_{t-1} \left( \frac{\ell(x)q(x)}{x^{t-1}} \right) \right) \right\rceil + 1 \\ &= \left\lceil \log_b \left( v_{t-1} \left( \frac{\ell(x)q(x)}{x^{t-1}} \right) \right) \right\rceil. \end{aligned}$$

Observing that there exists exactly one  $g \in \mathbb{F}_b$  for which  $a_{t-1} + \ell_0 g \equiv 0 \pmod{b}$  and combining the two cases considered, we immediately obtain the claimed identity.  $\square$

With the help of Lemma 6, we can prove the following result which motivates the choice of our quality function for Algorithm 1.

**Lemma 7** *For integers  $m \in \mathbb{N}$  and  $w \in \{1, \dots, m\}$ , let  $b$  be prime,  $g \in \mathbb{F}_b$ , and  $\mathbf{g} \in (\mathbb{F}_b[x])^d$  with  $\gcd(g_j, p_1) = 1$  for all  $1 \leq j \leq d$ , where  $g_d \in G_{b,w-1}$ , and let  $\boldsymbol{\eta} = (\eta_u)_{u \subseteq \{1:d\}}$ , where  $\eta_u = \prod_{j \in u} \eta_j$  with positive reals  $(\eta_j)_{j \geq 1}$ , be product weights. Then, the average of  $H_{d,m,\boldsymbol{\eta}}$  with respect to the choices for extending the degree of  $g_d + g p_{w-1}$  up to  $m$  equals*

$$\begin{aligned} & \frac{1}{b^{m-w}} \sum_{\bar{g} \in G_{b,m-w}} H_{d,m,\boldsymbol{\eta}}(g_1, \dots, g_{d-1}, g_d + g p_{w-1} + \bar{g} p_w) \\ &= \sum_{t=w}^m \sum_{\substack{\ell=1 \\ \ell \not\equiv 0 \pmod{b}}}^{b^t-1} \eta_d (1-b) b^{w-t} \left( \left\lceil \log_b \left( v_w \left( \frac{\ell(x)(g_d(x) + g x^{w-1})}{x^w} \right) \right) \right\rceil + 1 \right) \\ & \quad \times \prod_{j=1}^{d-1} \left( 1 + \eta_j (1-b) \left( \left\lceil \log_b \left( v_t \left( \frac{\ell(x)g_j(x)}{x^t} \right) \right) \right\rceil + 1 \right) \right) \\ & \quad + S_{m,w,\boldsymbol{\eta}}(\mathbf{g}) - (b^m - 1), \end{aligned} \tag{15}$$



where the term  $S_{m,w,\eta}(\mathbf{g})$ , which does not depend on  $g$  and  $\bar{g}$ , is given by

$$\begin{aligned}
 S_{m,w,\eta}(\mathbf{g}) &= \sum_{t=1}^{w-1} \sum_{\substack{\ell=1 \\ \ell \not\equiv 0 \pmod{b}}}^{b^t-1} \prod_{j=1}^d \left( 1 + \eta_j(1-b) \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x) g_j(x)}{x^t} \right) \right) \right\rfloor + 1 \right) \right) \\
 &+ \sum_{t=w}^m \sum_{\substack{\ell=1 \\ \ell \not\equiv 0 \pmod{b}}}^{b^t-1} \prod_{j=1}^{d-1} \left( 1 + \eta_j(1-b) \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x) g_j(x)}{x^t} \right) \right) \right\rfloor + 1 \right) \right) \\
 &\times (1 + \eta_d(1 - b^{w-t})).
 \end{aligned}$$

**Proof** For product weights  $\eta_u = \prod_{j \in u} \eta_j$  and  $\tilde{\mathbf{g}} = (\tilde{g}_1, \dots, \tilde{g}_d) \in (\mathbb{F}_b[x])^d$  with  $\gcd(\tilde{g}_j, p_1) = 1$  for all  $1 \leq j \leq d$ , the quantity  $H_{d,m,\eta}(\tilde{\mathbf{g}})$  defined in (13) equals

$$\begin{aligned}
 H_{d,m,\eta}(\tilde{\mathbf{g}}) &= \sum_{n=1}^{b^m-1} \prod_{j=1}^d \left( 1 + \eta_j(1-b) \left( \left\lfloor \log_b \left( v_m \left( \frac{n(x) \tilde{g}_j(x)}{x^m} \right) \right) \right\rfloor + 1 \right) \right) \\
 &- (b^m - 1).
 \end{aligned}$$

We define  $\tilde{H}_{d,m,\eta}(\tilde{\mathbf{g}}) := H_{d,m,\eta}(\tilde{\mathbf{g}}) + (b^m - 1)$  which in turn can be rewritten as

$$\tilde{H}_{d,m,\eta}(\tilde{\mathbf{g}}) = \sum_{t=1}^m \sum_{\substack{\ell=1 \\ \ell \not\equiv 0 \pmod{b}}}^{b^t-1} \prod_{j=1}^d \left( 1 + \eta_j(1-b) \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x) \tilde{g}_j(x)}{x^t} \right) \right) \right\rfloor + 1 \right) \right). \tag{16}$$

In order to see this, note that for any function  $f : \mathbb{R} \rightarrow \mathbb{R}$  we have that

$$\begin{aligned}
 &\sum_{n=1}^{b^m-1} f \left( v_m \left( \frac{n(x) \tilde{g}_j(x)}{x^m} \right) \right) \\
 &= \sum_{n=1}^{b^m-1} f \left( v_m \left( \frac{n(x) \tilde{g}_j(x) \pmod{x^m}}{x^m} \right) \right) = \sum_{n=1}^{b^m-1} f \left( v_m \left( \frac{n(x)}{x^m} \right) \right)
 \end{aligned}$$

for each  $j \in \{1, \dots, d\}$  since  $\gcd(\tilde{g}_j, x) = 1$  and that  $v_m \left( \frac{n}{x^m} \right) = \frac{n}{b^m}$  for  $n \in \{1, \dots, b^m - 1\}$ . Then, using the general identity

$$\sum_{n=1}^{b^m-1} f(n/b^m) = \sum_{t=1}^m \sum_{\substack{\ell=1 \\ \ell \not\equiv 0 \pmod{b}}}^{b^t-1} f(\ell/b^t)$$

yields the rewritten formula for  $\tilde{H}_{d,m,\eta}(\tilde{\mathbf{g}})$  in (16).

Setting  $\tilde{g}_d = g_d + g p_{w-1} + \bar{g} p_w$  with  $\bar{g} \in G_{b,m-w}$  and  $\tilde{g}_j = g_j$  for  $j \in \{1:d-1\}$ , we can write

$$\begin{aligned} & \frac{1}{b^{m-w}} \sum_{\bar{g} \in G_{b,m-w}} H_{d,m,\eta}(g_1, \dots, g_{d-1}, g_d + g p_{w-1} + \bar{g} p_w) \\ &= \frac{1}{b^{m-w}} \sum_{\bar{g} \in G_{b,m-w}} \bar{H}_{d,m,\eta}(\tilde{\mathbf{g}}) - (b^m - 1) \\ &= \frac{1}{b^{m-w}} \sum_{\bar{g} \in G_{b,m-w}} \sum_{t=1}^m \sum_{\substack{\ell=1 \\ \ell \not\equiv 0 \pmod{b}}}^{b^t-1} \prod_{j=1}^d \\ & \quad \left( 1 + \eta_j(1-b) \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x) \tilde{g}_j(x)}{x^t} \right) \right) \right\rfloor + 1 \right) \right) - (b^m - 1) \\ &= \frac{1}{b^{m-w}} \sum_{\bar{g} \in G_{b,m-w}} \sum_{t=1}^{w-1} \sum_{\substack{\ell=1 \\ \ell \not\equiv 0 \pmod{b}}}^{b^t-1} \prod_{j=1}^d \\ & \quad \left( 1 + \eta_j(1-b) \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x) \tilde{g}_j(x)}{x^t} \right) \right) \right\rfloor + 1 \right) \right) - (b^m - 1) \\ & \quad + \frac{1}{b^{m-w}} \sum_{\bar{g} \in G_{b,m-w}} \sum_{t=w}^m \sum_{\substack{\ell=1 \\ \ell \not\equiv 0 \pmod{b}}}^{b^t-1} \prod_{j=1}^d \\ & \quad \left( 1 + \eta_j(1-b) \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x) \tilde{g}_j(x)}{x^t} \right) \right) \right\rfloor + 1 \right) \right). \end{aligned}$$

The term  $-(b^m - 1)$  in (15) is therefore accounted for. What is more, by the definition of  $v_t$  we have for any  $q \in \mathbb{F}_b[x]$  that

$$v_t \left( \frac{q(x)}{x^t} \right) = v_t \left( \frac{q(x) \bmod x^t}{x^t} \right), \tag{17}$$

and hence

$$\begin{aligned} & \frac{1}{b^{m-w}} \sum_{\bar{g} \in G_{b,m-w}} \sum_{t=1}^{w-1} \sum_{\substack{\ell=1 \\ \ell \not\equiv 0 \pmod{b}}}^{b^t-1} \prod_{j=1}^d \left( 1 + \eta_j(1-b) \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x) \tilde{g}_j(x)}{x^t} \right) \right) \right\rfloor + 1 \right) \right) \\ &= \sum_{t=1}^{w-1} \sum_{\substack{\ell=1 \\ \ell \not\equiv 0 \pmod{b}}}^{b^t-1} \prod_{j=1}^d \left( 1 + \eta_j(1-b) \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x) g_j}{x^t} \right) \right) \right\rfloor + 1 \right) \right), \end{aligned}$$

which is the first sum in  $S_{m,w,\eta}$ , and, in particular, is independent of  $g$  and all  $\bar{g} \in G_{b,m-w}$ .

The second sum in  $S_{m,w,\eta}$  and all remaining terms in identity (15) are obtained by considering

$$\begin{aligned} & \frac{1}{b^{m-w}} \sum_{\bar{g} \in G_{b,m-w}} \sum_{t=w}^m \sum_{\substack{\ell=1 \\ \ell \neq 0 \pmod{b}}}^{b^t-1} \prod_{j=1}^d \left( 1 + \eta_j(1-b) \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x) \tilde{g}_j(x)}{x^t} \right) \right) \right\rfloor + 1 \right) \right) \\ &= \sum_{t=w}^m \sum_{\substack{\ell=1 \\ \ell \neq 0 \pmod{b}}}^{b^t-1} \prod_{j=1}^{d-1} \left( 1 + \eta_j(1-b) \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x) g_j(x)}{x^t} \right) \right) \right\rfloor + 1 \right) \right) \\ & \quad \times \left( 1 + \eta_d(1-b) \frac{1}{b^{m-w}} \right. \\ & \quad \times \left. \sum_{\bar{g} \in G_{b,m-w}} \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x) (g_d(x) + g x^{w-1} + \bar{g}(x) x^w)}{x^t} \right) \right) \right\rfloor + 1 \right) \right), \end{aligned} \tag{18}$$

such that, with the help of (17) and under the repeated use of Lemma 6, we obtain for each  $t \in \{w + 1, \dots, m\}$  that

$$\begin{aligned} & \sum_{\bar{g} \in G_{b,m-w}} \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x) (g_d(x) + g x^{w-1} + \bar{g}(x) x^w)}{x^t} \right) \right) \right\rfloor + 1 \right) \\ &= b^{m-t} \sum_{\bar{g} \in G_{b,t-w}} \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x) (g_d(x) + g x^{w-1} + \bar{g}(x) x^w)}{x^t} \right) \right) \right\rfloor + 1 \right) \\ &= b^{m-t} \sum_{\bar{g} \in G_{b,t-w-1}} \left( \left\lfloor \log_b \left( v_{t-1} \left( \frac{\ell(x) (g_d(x) + g x^{w-1} + \bar{g}(x) x^w)}{x^{t-1}} \right) \right) \right\rfloor + 1 - 1 \right) \\ &= b^{m-t} \left( \left\lfloor \log_b \left( v_w \left( \frac{\ell(x) (g_d(x) + g x^{w-1})}{x^w} \right) \right) \right\rfloor + 1 \right) - b^m \sum_{r=w+1}^t b^{-r} \\ &= b^{m-w} b^{w-t} \left( \left\lfloor \log_b \left( v_w \left( \frac{\ell(x) (g_d(x) + g x^{w-1})}{x^w} \right) \right) \right\rfloor + 1 \right) - b^{m-w} \left( \frac{1 - b^{w-t}}{b - 1} \right). \end{aligned}$$

Combining this with the identity in (18) yields the remaining term of  $S_{m,w,\eta}$  and the first term in (15) such that the claimed result is proved. □

We note that only the first term of (15) in Lemma 7 depends on the  $(w - 1)$ -th order term  $g x^{w-1}$  of  $g_d$ . Therefore, we can introduce the quality function for our algorithm which is based on the first term of (15), yet slightly adjusted by an additional summand that is independent of  $g$  and  $\bar{g}$ .

**Definition 3** (Digit-wise quality function) Let  $q \in \mathbb{F}_b[x]$ , with prime  $b$ , let  $m, d \in \mathbb{N}$ , and let  $\eta = (\eta_u)_{u \subseteq \{1:d\}}$ , where  $\eta_u = \prod_{j \in u} \eta_j$  with positive reals  $(\eta_j)_{j \geq 1}$ , be product weights. For integers  $w \in \{1:m\}$ ,  $r \in \{1:d\}$ , and polynomials  $g_1, \dots, g_{r-1} \in \mathbb{F}_b[x]$  with  $\gcd(g_j, p_1) = 1$  for  $j = 1, \dots, r - 1$ , we define the quality function  $h_{r,w,m,\eta} : \mathbb{F}_b[x] \rightarrow \mathbb{R}$  as

$$h_{r,w,m,\eta}(q) := \sum_{t=w}^m \frac{1}{b^{t-w}} \sum_{\substack{\ell=1 \\ \ell \neq 0 \pmod{b}}}^{b^t-1} \left( 1 + \eta_r(1-b) \left( \left\lfloor \log_b \left( v_w \left( \frac{\ell(x) q(x)}{x^w} \right) \right) \right\rfloor + 1 \right) \right) \\ \times \prod_{j=1}^{r-1} \left( 1 + \eta_j(1-b) \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x) g_j(x)}{x^t} \right) \right) \right\rfloor + 1 \right) \right).$$

We remark that the function  $h_{r,w,m,\eta}$  directly depends on the polynomials  $g_1, \dots, g_{r-1}$  even though this is not visible in the notation. In the remainder of this section, however, these polynomials will always be the components of the generating vector which were selected in the previous steps of our algorithm. Based on the quality function  $h_{r,w,m,\eta}$ , we formulate the component-by-component digit-by-digit algorithm.

---

### Algorithm 1 Component-by-component digit-by-digit algorithm

---

**Input:** Prime number  $b \geq 2$ , integers  $m, d \in \mathbb{N}$ , and positive product weights  $\eta = (\eta_j)_{j \geq 1}$ .

Set  $g_1 = 1$  and  $g_{2,1} = \dots = g_{d,1} = 1$ .

**for**  $r = 2$  **to**  $d$  **do**

**for**  $w = 2$  **to**  $m$  **do**

$g^* = \operatorname{argmin}_{g \in \mathbb{F}_b} h_{r,w,m,\eta}(g_{r,w-1} + g p_{w-1})$

$g_{r,w} = g_{r,w-1} + g^* p_{w-1}$

**end for**

  Set  $g_r := g_{r,m}$ .

**end for**

Set  $\mathbf{g} = (g_1, \dots, g_d)$ .

**Return:** Generating vector  $\mathbf{g} = (g_1, \dots, g_d) \in (G_{b,m}^*)^d$ .

---

In the next section, we study the worst-case error behavior of polynomial lattice rules with generating vectors obtained by Algorithm 1.

### 3.3 Error Bounds for the Constructed Polynomial Lattice Rules

The following theorem shows that for the constructed polynomial lattice rules the quantity  $H_{d,m,\eta}(\mathbf{g})$ , which for product weights  $\eta_u = \prod_{j \in u} \eta_j$  equals

$$H_{d,m,\eta}(\mathbf{g}) = \sum_{n=1}^{b^m-1} \prod_{j=1}^d \left( 1 + \eta_j(1-b) \left( \left\lfloor \log_b \left( v_m \left( \frac{n(x) g_j(x)}{x^m} \right) \right) \right\rfloor + 1 \right) \right) \\ - (b^m - 1) \\ = \sum_{t=1}^m \sum_{\substack{\ell=1 \\ \ell \neq 0 \pmod{b}}}^{b^t-1} \prod_{j=1}^d \left( 1 + \eta_j(1-b) \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x) g_j(x)}{x^t} \right) \right) \right\rfloor + 1 \right) \right) \\ - (b^m - 1),$$

can be related to the quantity  $H_{d-1,m,\eta}(\mathbf{g}_{\{1:d-1\}})$ .

**Theorem 5** *Let  $b$  be prime,  $m, d \in \mathbb{N}$  be integers with  $d \geq 2$ , and let  $\eta = (\eta_j)_{j \geq 1}$  be positive product weights. Furthermore, denote by  $\mathbf{g}$  the corresponding generating vector constructed by Algorithm 1. Then,  $\mathbf{g}$  satisfies*

$$H_{d,m,\eta}(\mathbf{g}) \leq (1 + \eta_d) H_{d-1,m,\eta}(\mathbf{g}_{\{1:d-1\}}) + \eta_d(b^m - 1). \tag{19}$$

**Proof** We will prove (19) by an inductive argument over the selection of the terms of order  $1 \leq t \leq m - 1$  of the polynomial  $g_d \in \mathbb{F}_b[x]$ . We start by considering the term of order  $m - 1$ . According to Algorithm 1, this term has been selected by minimizing  $h_{d,m,m,\eta}(g_{d,m-1} + g p_{m-1})$  over the choices  $g \in \mathbb{F}_b$ , and where  $g_{d,m-1} \in G_{b,m-1}$  has been determined in the previous steps of the algorithm. By Lemma 7 (with  $w = m$ ) and Definition 3, this is equivalent to minimizing

$$H_{d,m,\eta}(g_1, \dots, g_{d-1}, g_{d,m-1} + g p_{m-1})$$

with respect to  $g \in \mathbb{F}_b$ . By the standard averaging argument, this yields

$$\begin{aligned} H_{d,m,\eta}(\mathbf{g}) &= \min_{\bar{g} \in \mathbb{F}_b} H_{d,m,\eta}(\mathbf{g}_{\{1:d-1\}}, g_{d,m-1} + \bar{g} p_{m-1}) \\ &\leq \frac{1}{b} \sum_{\bar{g} \in \mathbb{F}_b} H_{d,m,\eta}(\mathbf{g}_{\{1:d-1\}}, g_{d,m-1} + \bar{g} p_{m-1}) \\ &= \frac{1}{b} \sum_{\bar{g} \in G_{b,1}} H_{d,m,\eta}(\mathbf{g}_{\{1:d-1\}}, g_{d,m-2} + g p_{m-2} + \bar{g} p_{m-1}), \end{aligned} \tag{20}$$

where  $g_{d,m-1}$  has been split up into  $g_{d,m-2}$  and  $g p_{m-2}$  in accordance with Algorithm 1 such that  $g$  has been selected in the previous step of the algorithm and we used that  $G_{b,1} \cong \mathbb{F}_b$ .

Similarly, we observe that the term of order  $m - 2$  has been selected by minimizing  $h_{d,m-1,m,\eta}(g_{d,m-2} + g p_{m-2})$  with respect to the choices  $g \in \mathbb{F}_b$ . Again, by Lemma 7 (with  $w = m - 1$ ) and Definition 3 this is equivalent to minimizing

$$\frac{1}{b} \sum_{\bar{g} \in G_{b,1}} H_{d,m,\eta}(\mathbf{g}_{\{1:d-1\}}, g_{d,m-2} + g p_{m-2} + \bar{g} p_{m-1})$$

with respect to  $g \in G_{b,1} \cong \mathbb{F}_b$ . By the standard averaging argument, we obtain that

$$\begin{aligned} &\min_{\bar{g} \in G_{b,1}} \frac{1}{b} \sum_{\bar{g} \in \mathbb{F}_b} H_{d,m,\eta}(\mathbf{g}_{\{1:d-1\}}, g_{d,m-2} + g p_{m-2} + \bar{g} p_{m-1}) \\ &\leq \frac{1}{b^2} \sum_{g \in \mathbb{F}_b} \sum_{\bar{g} \in G_{b,1}} H_{d,m,\eta}(\mathbf{g}_{\{1:d-1\}}, g_{d,m-2} + g p_{m-2} + \bar{g} p_{m-1}) \\ &= \frac{1}{b^2} \sum_{\bar{g} \in G_{b,2}} H_{d,m,\eta}(\mathbf{g}_{\{1:d-1\}}, g_{d,m-3} + g p_{m-3} + \bar{g} p_{m-2}), \end{aligned}$$

where again we split up  $g_{d,m-2} = g_{d,m-3} + g_{p_{m-3}}$  according to Algorithm 1. Inductively repeating this argument and combining the result with the estimate in (20), we obtain the inequality

$$H_{d,m,\eta}(\mathbf{g}) \leq \frac{1}{b^{m-1}} \sum_{\bar{g} \in G_{b,m-1}} H_{d,m,\eta}(\mathbf{g}_{\{1:d-1\}}, 1 + \bar{g}_{p_1}),$$

where we used that in Algorithm 1 we set  $g_{d,1} = 1$ . Then, using Lemma 7 with  $w = 1$ ,  $g_d = 1$ , and  $g = 0$  to equate the right-hand side of the previous estimate, we finally obtain

$$\begin{aligned} &H_{d,m,\eta}(\mathbf{g}) \\ &\leq \sum_{t=1}^m \sum_{\substack{\ell=1 \\ \ell \not\equiv 0 \pmod{b}}}^{b^t-1} \eta_d(1-b)b^{1-t} \left( \left\lfloor \log_b \left( v_1 \left( \frac{\ell(x)}{x} \right) \right) \right\rfloor + 1 \right) \\ &\quad \times \prod_{j=1}^{d-1} \left( 1 + \eta_j(1-b) \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x)g_j(x)}{x^t} \right) \right) \right\rfloor + 1 \right) \right) - (b^m - 1) \\ &\quad + \sum_{t=1}^m \sum_{\substack{\ell=1 \\ \ell \not\equiv 0 \pmod{b}}}^{b^t-1} \prod_{j=1}^{d-1} \left( 1 + \eta_j(1-b) \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x)g_j(x)}{x^t} \right) \right) \right\rfloor + 1 \right) \right) \\ &\quad \left( 1 + \eta_d(1-b^{1-t}) \right). \end{aligned}$$

For  $\ell$  with  $\ell \not\equiv 0 \pmod{b}$ , which is equivalent to  $\gcd(\ell, p_1) = 1$ , we have for some  $a \in \mathbb{F}_b \setminus \{0\}$  that  $\lfloor \log_b(v_1(\ell(x)/x)) \rfloor + 1 = \lfloor \log_b(a/b) \rfloor + 1 = -1 + 1 = 0$ . Hence, we get

$$\begin{aligned} H_{d,m,\eta}(\mathbf{g}) &\leq \sum_{t=1}^m \sum_{\substack{\ell=1 \\ \ell \not\equiv 0 \pmod{b}}}^{b^t-1} \prod_{j=1}^{d-1} \left( 1 + \eta_j(1-b) \left( \left\lfloor \log_b \left( v_t \left( \frac{\ell(x)g_j(x)}{x^t} \right) \right) \right\rfloor + 1 \right) \right) \\ &\quad \times \left( 1 + \eta_d(1-b^{1-t}) \right) - (b^m - 1) \\ &\leq (1 + \eta_d) (H_{d-1,m,\eta}(\mathbf{g}_{\{1:d-1\}}) + (b^m - 1)) - (b^m - 1) \\ &= (1 + \eta_d) H_{d-1,m,\eta}(\mathbf{g}_{\{1:d-1\}}) + \eta_d(b^m - 1), \end{aligned}$$

which is the claimed estimate. □

**Remark 5** The proof strategy in Theorem 5, in which we relate the quantities  $H_{d,m,\eta}(\mathbf{g})$  and  $H_{d-1,m,\eta}(\mathbf{g})$  with each other, is the main reason why we restricted our analysis to product weights. Without this assumption, the inductive argument over the dimension becomes so complicated that so far it has not been possible to prove the induction step. The same issue has also occurred for the analogous result for lattice rules, see [7, Theorems 3 and 4].

Based on the result in Theorem 5, we can use an inductive argument to show that the quantity  $H_{d,m,\eta}(\mathbf{g})$  is sufficiently small if  $\mathbf{g}$  has been constructed by Algorithm 1.

**Theorem 6** *Let  $b$  be prime, let  $m, d \in \mathbb{N}$  be positive integers and let  $\eta = (\eta_j)_{j \geq 1}$  be positive product weights. Then, the generating vector  $\mathbf{g}$  constructed by Algorithm 1 satisfies*

$$H_{d,m,\eta}(\mathbf{g}) \leq b^m \left[ -1 + \prod_{j=1}^d (1 + \eta_j) \right].$$

**Proof** Due to the formulation of Algorithm 1, the estimate (19) obtained in Theorem 5 holds if we replace  $d$  by  $r$  for any  $r \in \{2, \dots, d\}$ , such that we get a result for  $H_{r,m,\eta}(\mathbf{g})$  for any  $r \in \{2, \dots, d\}$ . Hence, we can use this estimate inductively to obtain

$$\begin{aligned} H_{d,m,\eta}(\mathbf{g}) &\leq (1 + \eta_d)H_{d-1,m,\eta}(\mathbf{g}_{\{1:d-1\}}) + \eta_d(b^m - 1) \\ &\leq (1 + \eta_d)(1 + \eta_{d-1})H_{d-2,m,\eta}(\mathbf{g}_{\{1:d-2\}}) + (1 + \eta_d)\eta_{d-1}(b^m - 1) + \eta_d(b^m - 1) \\ &= H_{d-2,m,\eta}(\mathbf{g}_{\{1:d-2\}}) \prod_{j=d-1}^d (1 + \eta_j) + (b^m - 1) \left[ -1 + \prod_{j=d-1}^d (1 + \eta_j) \right] \\ &\leq H_{1,m,\eta}(g_1) \prod_{j=2}^d (1 + \eta_j) + (b^m - 1) \left[ -1 + \prod_{j=2}^d (1 + \eta_j) \right]. \end{aligned} \tag{21}$$

Next, we observe that

$$\begin{aligned} H_{1,m,\eta}(g_1) &= H_{1,m,\eta}(1) \\ &= \sum_{n=1}^{b^m-1} \left( 1 + \eta_1(1 - b) \left( \left\lfloor \log_b \left( v_m \left( \frac{n(x)}{x^m} \right) \right) \right\rfloor + 1 \right) \right) - (b^m - 1) \\ &= -\eta_1 \sum_{n=1}^{b^m-1} (b - 1) \left( \left\lfloor \log_b \left( v_m \left( \frac{n(x)}{x^m} \right) \right) \right\rfloor + 1 \right) \\ &= -\eta_1 \sum_{t=1}^m \sum_{\substack{n=1 \\ n \not\equiv 0 \pmod{b}}}^{b^t-1} (b - 1) \left\lfloor \log_b \left( v_t \left( \frac{n(x)}{x^t} \right) \right) \right\rfloor + \eta_1(1 - b)(b^m - 1) \\ &= -\eta_1 \sum_{t=1}^m \sum_{r=0}^{t-1} \sum_{\substack{n=1 \\ n \not\equiv 0 \pmod{b} \\ \deg(n(x))=r}}^{b^t-1} (b - 1) \left\lfloor \log_b \left( v_t \left( \frac{n(x)}{x^t} \right) \right) \right\rfloor + \eta_1(1 - b)(b^m - 1). \end{aligned}$$

For any polynomial  $n(x) \in \mathbb{F}_b[x]$  of degree  $0 \leq r < t$  with  $\gcd(n, x) = 1$ , we have that

$$\left\lfloor \log_b \left( v_t \left( \frac{n(x)}{x^t} \right) \right) \right\rfloor = -(t - r)$$

such that we can further deduce that

$$\begin{aligned}
 H_{1,m,\eta}(g_1) &= \eta_1 \sum_{t=1}^m (b-1) \sum_{r=0}^{t-1} \sum_{\substack{n=1 \\ n \not\equiv 0 \pmod{b} \\ \deg(n(x))=r}}^{b^t-1} (t-r) + \eta_1(1-b)(b^m-1) \\
 &= \eta_1 \sum_{t=1}^m (b-1) \left( (b-1)t + \sum_{r=1}^{t-1} (b-1)^2 b^{r-1} (t-r) \right) + \eta_1(1-b)(b^m-1) \\
 &= \eta_1 \sum_{t=1}^m (b-1) ((b-1)t + b^t - bt + t - 1) + \eta_1(1-b)(b^m-1) \\
 &= \eta_1(b-1) \sum_{t=1}^m (b^t - 1) + \eta_1(1-b)(b^m-1) \\
 &= \eta_1(b^{m+1} - bm - b + m) + \eta_1(1-b)(b^m-1) = \eta_1(b^m - (b-1)m - 1).
 \end{aligned}$$

Combining this with the estimate in (21), we finally obtain

$$\begin{aligned}
 H_{d,m,\eta}(\mathbf{g}) &\leq \eta_1 (b^m - 1) \prod_{j=2}^d (1 + \eta_j) + (b^m - 1) \left[ -1 + \prod_{j=2}^d (1 + \eta_j) \right] \\
 &= (b^m - 1) \left[ -1 + \prod_{j=1}^d (1 + \eta_j) \right],
 \end{aligned}$$

which yields the claimed estimate. □

Theorem 6 allows us to prove the following result regarding the construction in Algorithm 1.

**Theorem 7** *Let  $b$  be prime, let  $m, d \in \mathbb{N}$  with  $m \geq 4$ , and let  $(\eta_j)_{j \geq 1}$  be positive product weights. Then, the generating vector  $\mathbf{g}$  constructed by Algorithm 1 satisfies*

$$\begin{aligned}
 &T_\eta(\mathbf{g}, p_m) \\
 &\leq \frac{1}{b^m} \left[ \prod_{j=1}^d (1 + \eta_j((b-1)m + 1)) + bm \prod_{j=1}^d \left( 1 + \eta_j \left( 2(b-1)m + \frac{2b}{b-1} \right) \right) \right].
 \end{aligned}$$

**Proof** We remark that for reals  $a_1, \dots, a_d \in \mathbb{R}$  the general identity

$$\sum_{\emptyset \neq u \subseteq \{1:d\}} \prod_{j \in u} a_j = -1 + \prod_{j=1}^d (1 + a_j)$$

holds. Using the bound on  $T_\eta(\mathbf{g}, p_m)$  in Theorem 4 and inserting for  $\mathbf{g}$  the generating vector obtained from Algorithm 1, for which the bound on  $H_{d,m,\eta}(\mathbf{g})$  from Theorem 6 holds, yields



$$\begin{aligned}
 T_\eta(\mathbf{g}, p_m) &\leq \left[ -1 + \prod_{j=1}^d (1 + \eta_j) \right] - \left[ -1 + \prod_{j=1}^d (1 + \eta_j) \right] \\
 &\quad + \sum_{\emptyset \neq u \subseteq \{1:d\}} \frac{\eta_u}{b^m} ((b-1)m + 1)^{|u|} \\
 &\quad + \sum_{\emptyset \neq u \subseteq \{1:d\}} \frac{\eta_u}{b^m} (bm |u|) \left( (b-1)m + \frac{b}{b-1} \right)^{|u|} \\
 &\leq \frac{1}{b^m} \left[ \prod_{j=1}^d (1 + \eta_j ((b-1)m + 1)) \right. \\
 &\quad \left. + bm \prod_{j=1}^d \left( 1 + \eta_j \left( 2(b-1)m + \frac{2b}{b-1} \right) \right) \right],
 \end{aligned}$$

where in the last step we used that  $|u| \leq 2^{|u|}$ . Note that by the formulation of Algorithm 1 we have that  $\gcd(g_j, p_m) = 1$  for  $1 \leq j \leq d$  such that the conditions of Theorem 4 are satisfied. □

The next theorem states the main result of this paper, implying that by the construction in Algorithm 1 we obtain an error convergence rate that is arbitrarily close to the optimal rate of  $N^{-\alpha}$ . We know that this order is optimal due to the relation between the worst-case errors in  $W_{d,\boldsymbol{\gamma}}^\alpha$  and  $\tilde{W}_{d,\boldsymbol{\gamma}}^\alpha$  (see Theorem 2 and Remark 3) and due to the fact that the rate  $N^{-\alpha/2}$  is optimal in  $\tilde{W}_{d,\boldsymbol{\gamma}}^\alpha$ , cf., e.g., [6, Theorem 41]. Additionally, under a summability condition on the weights that is common in the related literature, the error can be bounded independently of the dimension, by which we obtain what is known as strong polynomial tractability in the context of information-based complexity. For an overview of different notions of tractability and basic concepts of information-based complexity, we refer the interested reader to [17].

**Theorem 8** *Let  $b$  be prime, let  $m, d \in \mathbb{N}$  with  $m \geq 4$ , let  $N = b^m$ , and let  $(\gamma_j)_{j \geq 1}$  be positive product weights satisfying*

$$\sum_{j \geq 1} \gamma_j < \infty.$$

*Furthermore, denote by  $\mathbf{g}$  the generating vector obtained by Algorithm 1, run for the weight sequence  $\boldsymbol{\eta} = \boldsymbol{\gamma} = (\gamma_j)_{j \geq 1}$ . Then, for any  $\delta > 0$  and each  $\alpha > 1$ , the generating vector  $\mathbf{g}$  satisfies*

$$e_{b^m, d, \alpha, \boldsymbol{\gamma}^\alpha}(\mathbf{g}) \leq \frac{1}{N^\alpha} \left( C(\boldsymbol{\gamma}^\alpha) + \bar{C}(\boldsymbol{\gamma}, \delta) N^{\alpha\delta} \right),$$

*with positive constants  $C(\boldsymbol{\gamma}^\alpha)$  and  $\bar{C}(\boldsymbol{\gamma}, \delta)$ , which are independent of  $d$  and  $N$ .*

Additionally, if Algorithm 1 is run for the weights  $\eta = \boldsymbol{\gamma}^{1/\alpha}$  with  $\alpha > 1$ , which satisfy

$$\sum_{j \geq 1} \gamma_j^{1/\alpha} < \infty,$$

then, for any  $\delta > 0$ , the resulting generating vector  $\tilde{\mathbf{g}}$  satisfies the error bound

$$e_{b^m, d, \alpha, \boldsymbol{\gamma}}(\tilde{\mathbf{g}}) \leq \frac{1}{N^\alpha} \left( K(\boldsymbol{\gamma}) + \bar{K}(\boldsymbol{\gamma}^{1/\alpha}, \delta) N^{\alpha\delta} \right),$$

with positive constants  $K(\boldsymbol{\gamma})$  and  $\bar{K}(\boldsymbol{\gamma}^{1/\alpha}, \delta)$ , which are independent of  $d$  and  $N$ .

**Proof** We know from Proposition 1 that

$$e_{b^m, d, \alpha, \eta^\alpha}(\mathbf{g}) \leq \frac{1}{N^\alpha} \sum_{\emptyset \neq u \subseteq \{1:d\}} \eta_u^\alpha (2\mu_b(\alpha))^{|u|} + T_{\alpha, \eta^\alpha}(\mathbf{g}, p_m).$$

For the special case of product weights  $\eta_u = \prod_{j \in u} \eta_j$ ,  $u \subseteq \{1:d\}$ , this yields

$$e_{b^m, d, \alpha, \eta^\alpha}(\mathbf{g}) \leq \frac{1}{N^\alpha} \prod_{j=1}^d \left( 1 + 2\mu_b(\alpha)\eta_j^\alpha \right) + T_{\alpha, \eta^\alpha}(\mathbf{g}, p_m).$$

Since  $\alpha > 1$ , we can use an inequality, sometimes referred to as Jensen’s inequality, which states that  $\sum_{i=1}^M y_i \leq \left( \sum_{i=1}^M y_i^p \right)^{1/p}$  for non-negative  $y_1, \dots, y_M$  and  $0 \leq p \leq 1$ . This yields

$$\begin{aligned} T_{\alpha, \eta^\alpha}(\mathbf{g}, p_m) &= \sum_{\mathbf{0} \neq \mathbf{k} \in A_p(\mathbf{g})} (r_{\alpha, \eta^\alpha}(\mathbf{k}))^{-1} = \sum_{\mathbf{0} \neq \mathbf{k} \in A_p(\mathbf{g})} (r_{1, \eta}(\mathbf{k}))^{-\alpha} \\ &\leq \left( \sum_{\mathbf{0} \neq \mathbf{k} \in A_p(\mathbf{g})} (r_{1, \eta}(\mathbf{k}))^{-1} \right)^\alpha = (T_\eta(\mathbf{g}, p_m))^\alpha, \end{aligned}$$

and by Theorem 7 we know that Algorithm 1 run for weights  $\eta$  yields  $\mathbf{g}$  which satisfy

$$\begin{aligned} T_\eta(\mathbf{g}, p_m) &\leq \frac{1}{b^m} \left[ \prod_{j=1}^d (1 + \eta_j((b-1)m + 1)) \right. \\ &\quad \left. + b m \prod_{j=1}^d \left( 1 + \eta_j \left( 2(b-1)m + \frac{2b}{b-1} \right) \right) \right]. \end{aligned}$$

From this, we deduce, using either the weights  $\eta = \boldsymbol{\gamma}^{1/\alpha}$  or  $\eta = \boldsymbol{\gamma}$  for Algorithm 1, that

$$\begin{aligned} b^m T_\eta(\mathbf{g}, p_m) &\leq \prod_{j=1}^d (1 + \eta_j 4bm) + b m \prod_{j=1}^d (1 + \eta_j 4bm) = (1 + b m) \prod_{j=1}^d (1 + \eta_j 4bm) \\ &\leq \tilde{C}(\delta/2) b^{m\delta/2} \prod_{j=1}^d (1 + \eta_j 4bm) \leq \tilde{C}(\delta/2) b^{m\delta/2} \prod_{j=1}^\infty (1 + \eta_j 4bm) \end{aligned}$$

for arbitrary  $\delta > 0$ , where  $\tilde{C}(\delta/2)$  is a constant depending only on  $\delta$ . Due to the imposed condition on the weights, i.e.,  $\sum_{j \geq 1} \gamma_j < \infty$  or  $\sum_{j \geq 1} \gamma_j^{1/\alpha} < \infty$ , we can use the result in [8, Lemma 3] to see that the last product can be bounded by  $\widehat{C}(\boldsymbol{\gamma}, \delta) b^{m\delta/2}$  or  $\widehat{C}(\boldsymbol{\gamma}^{1/\alpha}, \delta) b^{m\delta/2}$ , respectively, where  $\widehat{C}(\boldsymbol{\gamma}, \delta)$  and  $\widehat{C}(\boldsymbol{\gamma}^{1/\alpha}, \delta)$  may depend on  $\delta$  and the weights  $\boldsymbol{\gamma}$  or  $\boldsymbol{\gamma}^{1/\alpha}$ , but are independent of the dimension. Choosing  $\eta = \boldsymbol{\gamma}$ , this yields that

$$(T_\eta(\mathbf{g}, p_m))^\alpha = (T_\boldsymbol{\gamma}(\mathbf{g}, p_m))^\alpha \leq \frac{1}{N^\alpha} (\tilde{C}(\delta/2))^\alpha (\widehat{C}(\boldsymbol{\gamma}, \delta))^\alpha N^{\alpha\delta},$$

and similarly, for  $\eta = \boldsymbol{\gamma}^{1/\alpha}$ ,

$$(T_\eta(\mathbf{g}, p_m))^\alpha = (T_{\boldsymbol{\gamma}^{1/\alpha}}(\mathbf{g}, p_m))^\alpha \leq \frac{1}{N^\alpha} (\tilde{C}(\delta/2))^\alpha (\widehat{C}(\boldsymbol{\gamma}^{1/\alpha}, \delta))^\alpha N^{\alpha\delta}.$$

Setting then  $C(\boldsymbol{\gamma}^\alpha) = \prod_{j=1}^d (1 + 2\mu_b(\alpha)\gamma_j^\alpha)$  and  $\bar{C}(\boldsymbol{\gamma}, \delta) = (\tilde{C}(\delta/2))^\alpha (\widehat{C}(\boldsymbol{\gamma}, \delta))^\alpha$ , and, similarly,  $K(\boldsymbol{\gamma}) = \prod_{j=1}^d (1 + 2\mu_b(\alpha)\gamma_j)$  and  $\bar{K}(\boldsymbol{\gamma}^{1/\alpha}, \delta) = (\tilde{C}(\delta/2))^\alpha (\widehat{C}(\boldsymbol{\gamma}^{1/\alpha}, \delta))^\alpha$ , we obtain the claimed error estimates, where the first stated bound holds simultaneously for all  $\alpha > 1$ .  $\square$

The result in Theorem 8 consists of two statements regarding the worst-case error behavior of generating vectors constructed by Algorithm 1. On the one hand, when run with weights  $\boldsymbol{\gamma}^{1/\alpha}$ , and hence depending on the parameter  $\alpha$ , the algorithm yields typical error bounds for the worst-case error in the space  $W_{d,\boldsymbol{\gamma}^\alpha}^\alpha$ . We emphasize that this type of result could also be obtained by formulating and using an analogous CBC-DBD algorithm which is instead directly based on the search criterion  $e_{b^m,d,\alpha,\boldsymbol{\gamma}}$ . On the other hand, when run with weights  $\boldsymbol{\gamma}$ , thus independently of  $\alpha$ , the algorithm produces generating vectors for which bounds on the worst-case errors in the spaces  $W_{d,\boldsymbol{\gamma}^\alpha}^\alpha$  hold simultaneously for all  $\alpha > 1$ .

### 4 Fast Implementation of the Construction Scheme

In this section, we discuss the efficient implementation of the introduced CBC-DBD algorithm and analyze its complexity. Throughout this section, we will consider the implementation for the special case of  $b = 2$  and product weights  $\gamma_u = \prod_{j \in u} \gamma_j$

for a sequence of positive reals  $(\gamma_j)_{j \geq 1}$ . Choosing the prime base as  $b = 2$  allows for the use of bitwise operations which facilitate an efficient implementation of the construction scheme. We remark that the major challenge for the implementation of the algorithm for  $b > 2$  is an efficient computation of the polynomial multiplication modulo  $b$ , all other steps of the algorithm can be implemented analogously.

### 4.1 Implementation and Cost Analysis of the CBC-DBD Algorithm

Let  $q \in \mathbb{F}_2[x]$ ,  $m, d \in \mathbb{N}$  be positive integers and let  $\boldsymbol{\gamma} = (\gamma_u)_{u \subseteq \{1:d\}}$ , where  $\gamma_u = \prod_{j \in u} \gamma_j$  with positive reals  $(\gamma_j)_{j \geq 1}$ . We recall that for  $b = 2$  and integers  $w \in \{1:m\}$ ,  $r \in \{1:d\}$  the digit-wise quality function  $h_{r,w,m,\boldsymbol{\gamma}}$  in Definition 3, which is used in Algorithm 1, is given by

$$h_{r,w,m,\boldsymbol{\gamma}}(q) = \sum_{t=w}^m \frac{1}{2^{t-w}} \sum_{\substack{\ell=1 \\ \ell \equiv 1 \pmod{2}}}^{2^t-1} \left( 1 - \gamma_r \left( \left\lfloor \log_2 \left( v_w \left( \frac{\ell(x) q(x)}{x^w} \right) \right) \right\rfloor + 1 \right) \right) \times \prod_{j=1}^{r-1} \left( 1 - \gamma_j \left( \left\lfloor \log_2 \left( v_t \left( \frac{\ell(x) g_j(x)}{x^t} \right) \right) \right\rfloor + 1 \right) \right),$$

where the polynomials  $g_1, \dots, g_{r-1} \in \mathbb{F}_2[x]$  have been determined in the previous steps of the algorithm. Since the cost of a single evaluation of the function  $h_{r,w,m,\boldsymbol{\gamma}}$  is crucial for the total cost of Algorithm 1, we are interested in an efficient evaluation procedure which will be discussed in the following paragraph.

For integers  $t \in \{2, \dots, m\}$  and odd  $\ell \in \{1, \dots, 2^t - 1\}$ , we define the term  $a(r, t, \ell)$  as

$$a(r, t, \ell) := \prod_{j=1}^r \left( 1 - \gamma_j \left( \left\lfloor \log_2 \left( v_t \left( \frac{\ell(x) g_j(x)}{x^t} \right) \right) \right\rfloor + 1 \right) \right)$$

and observe that for the evaluation of  $h_{r,w,m,\boldsymbol{\gamma}}(q)$  we can compute and store the term  $a(r - 1, t, \ell)$  since it is independent of  $w$  and  $q$ . This way we can rewrite  $h_{r,w,m,\boldsymbol{\gamma}}(q)$  as

$$h_{r,w,m,\boldsymbol{\gamma}}(q) = \sum_{t=w}^m \frac{1}{2^{t-w}} \sum_{\substack{\ell=1 \\ \ell \equiv 1 \pmod{2}}}^{2^t-1} a(r - 1, t, \ell) \left( 1 - \gamma_r \left( \left\lfloor \log_2 \left( v_w \left( \frac{\ell(x) q(x)}{x^w} \right) \right) \right\rfloor + 1 \right) \right), \tag{22}$$

where in Algorithm 1, after having determined  $g_{r,w}$ , the values of  $a(r, w, \ell)$  for odd integers  $\ell \in \{1, \dots, 2^w - 1\}$  are computed via the recurrence relation

$$a(r, w, \ell) = a(r - 1, w, \ell) \left( 1 - \gamma_r \left( \left\lfloor \log_2 \left( v_w \left( \frac{\ell(x) g_{r,w}(x)}{x^w} \right) \right) \right\rfloor + 1 \right) \right).$$

For an algorithmic implementation, we introduce the vector  $\mathbf{u} = (u(1), \dots, u(2^m - 1)) \in \mathbb{R}^{2^m - 1}$  whose components, for the current  $r \in \{1, \dots, d\}$ , are given by

$$u(\ell 2^{m-t}) = \prod_{j=1}^r \left( 1 - \gamma_j \left( \left\lfloor \log_2 \left( v_t \left( \frac{\ell(x) g_j(x)}{x^t} \right) \right) \right\rfloor + 1 \right) \right) = a(r, t, \ell)$$

for each  $t = 1, \dots, m$  and corresponding odd index  $\ell \in \{1, \dots, 2^t - 1\}$ . In Algorithm 2, the quantity  $\mathbf{u}$  stores the values of  $a(r - 1, t, \ell)$  for the current  $r$  and can therefore be used to evaluate  $h_{r,w,m,\boldsymbol{\gamma}}$  according to equation (22). Furthermore, note that for the evaluation of  $h_{r,w,m,\boldsymbol{\gamma}}$  we do not require the values of  $a(r, t, \ell)$  for  $t = 2, \dots, w - 1$ . Combining these findings leads to the following fast implementation of Algorithm 1.

---

**Algorithm 2** Fast component-by-component digit-by-digit algorithm

---

**Input:** Integers  $m, d \in \mathbb{N}$  and positive weights  $(\gamma_j)_{j=1}^d$ .

for  $\ell = 1$  to  $2^m - 1$  do

$$u(\ell) = 1 - \gamma_1 \left( \left\lfloor \log_2 \left( v_m \left( \frac{\ell}{x^m} \right) \right) \right\rfloor + 1 \right)$$

end for

Set  $g_1 = 1$  and  $g_{2,1} = \dots = g_{d,1} = 1$ .

for  $r = 2$  to  $d$  do

for  $w = 2$  to  $m$  do

$$g^* = \operatorname{argmin}_{g \in \mathbb{F}_2} h_{r,w,m,\boldsymbol{\gamma}}(g_{r,w-1} + g x^{w-1}) \text{ with } h_{r,w,m,\boldsymbol{\gamma}} \text{ evaluated using (22)}$$

$$g_{r,w} = g_{r,w-1} + g^* x^{w-1}$$

for  $\ell = 1$  to  $2^w - 1$  in steps of 2 do

$$u(\ell 2^{m-w}) = u(\ell 2^{m-w}) \left( 1 - \gamma_r \left( \left\lfloor \log_2 \left( v_w \left( \frac{\ell g_{r,w}}{x^w} \right) \right) \right\rfloor + 1 \right) \right)$$

end for

end for

Set  $g_r := g_{r,m}$ .

end for

Set  $\mathbf{g} = (g_1, \dots, g_d)$ .

**Return:** Generating vector  $\mathbf{g} = (g_1, \dots, g_d) \in (G_{2,m}^*)^d$  for  $N = 2^m$ .

---

The computational complexity of Algorithm 2 is then summarized in the following theorem.

**Theorem 9** Let  $m, d \in \mathbb{N}$  and let  $\boldsymbol{\gamma} = (\gamma_j)_{j=1}^d$  be a given sequence of positive weights. Then, Algorithm 2 constructs a generating vector  $\mathbf{g} = (g_1, \dots, g_d) \in (G_{2,m}^*)^d$  using  $\mathcal{O}(d m 2^m)$  operations and requiring  $\mathcal{O}(2^m)$  memory.

**Proof** Due to the relation in (22), the cost of evaluating  $h_{r,w,m,\gamma}(q)$  can be reduced to  $\mathcal{O}(\sum_{t=w}^m 2^{t-1})$  operations. Thus, the number of calculations in the inner loop over  $w = 2, \dots, m$  of Algorithm 2 is of order

$$\mathcal{O}\left(\sum_{w=2}^m 2 \sum_{t=w}^m 2^{t-1}\right) = \mathcal{O}\left(\sum_{w=2}^m \sum_{t=w}^m 2^t\right) = \mathcal{O}(m 2^m - 2(2^m - 1)) = \mathcal{O}(m 2^m).$$

Hence, the outer loop over  $r = 2, \dots, d$ , which is the main cost of Algorithm 2, can be executed in  $\mathcal{O}(d m 2^m)$  operations. Furthermore, we observe that initialization and updating of the vector  $\mathbf{u} \in \mathbb{R}^{2^m-1}$  for each  $r$  can both be executed in

$$\sum_{w=2}^m \mathcal{O}(2^{w-1}) = \sum_{w=1}^{m-1} \mathcal{O}(2^w) = \mathcal{O}(2^m) \text{ operations,}$$

such that a total cost of  $\mathcal{O}(d 2^m)$  operations for maintaining  $\mathbf{u}$  is required in Algorithm 2. Additionally, storing the vector  $\mathbf{u}$  requires  $\mathcal{O}(2^m)$  of memory.  $\square$

We note that for the modulus polynomial  $x^m$  the evaluation of  $v_m(\ell(x)/x^m)$  requires only one integer division since  $v_m(\ell(x)/x^m) = \ell/2^m$ . Furthermore, we remark that the running time of Algorithm 2 can be reduced further by precomputing and storing the  $2^m - 1$  values

$$\left( \left\lfloor \log_2 \left( v_m \left( \frac{\ell}{x^m} \right) \right) \right\rfloor + 1 \right) \text{ for } \ell = 1, \dots, 2^m - 1.$$

The derivation leading to the fast implementation in Algorithm 2 is using arguments that were used in [7], where a component-by-component digit-by-digit construction for lattice rules in weighted Korobov spaces has been studied. Theorem 9 shows that the fast implementation of the component-by-component digit-by-digit construction for polynomial lattice rules achieves the same computational complexity as state-of-the-art component-by-component methods, see, e.g., [3], which, for product weights, require  $\mathcal{O}(d m 2^m)$  operations. In these constructions, the speed-up of the algorithm is achieved by reordering the involved matrices to be of circulant structure and by then employing a fast matrix-vector product which uses fast Fourier transformations (FFTs). We refer to [19] for further details on an implementation for polynomial lattice rules. In contrast, our method does not rely on the use of FFTs and the low time complexity of the resulting algorithm is due to the smaller search space for the components  $g_j$  of the generating vector  $\mathbf{g}$ . Furthermore, we remark that the mentioned state-of-the-art CBC constructions mainly use a primitive or irreducible modulus  $p \in \mathbb{F}_2[x]$  since then the multiplicative group of  $\mathbb{F}_2[x]/(p)$  is cyclic. While for reducible polynomials, such as  $p(x) = x^m$ , a fast CBC construction is theoretically possible by using a similar strategy as for the fast CBC construction for lattice rules with a composite number of points, there are, to the best of our knowledge, no explicit implementations of such an algorithm known. On the other hand, the CBC-DBD construction considered in this article immediately yields a fast algorithm for the construction of polynomial lattice rules in  $\mathcal{O}(d m 2^m)$  operations for  $p(x) = x^m$ .

### 5 Numerical Results

In this section, we illustrate the error convergence behavior of the polynomial lattice rules constructed by the CBC-DBD algorithm and visualize the computational complexity of the construction by means of numerical experiments. As in the previous section, we consider polynomial lattice rules in the weighted Walsh space  $W_{d,\boldsymbol{\gamma}}^\alpha$  for prime base  $b = 2$  and product weights  $\gamma_u = \prod_{j \in u} \gamma_j$  given in terms of positive reals  $(\gamma_j)_{j \geq 1}$ .

In order to demonstrate the performance of the algorithm, we compare the worst-case errors of the constructed polynomial lattice rules as well as the algorithm’s computation times to the corresponding quantities obtained by a state-of-the-art component-by-component algorithm, see, e.g., [3]. As remarked in the previous section, no fast CBC construction is known for the case  $p(x) = x^m$  such that instead we compare our algorithm with a CBC construction with primitive polynomial  $p \in \mathbb{F}_2[x]$  of degree  $m$  as the modulus. Both constructions deliver polynomial lattice rules for the spaces  $W_{d,\boldsymbol{\gamma}}^\alpha$  consisting of  $2^m$  cubature points.

The different algorithms have been implemented in MATLAB R2019b and Python 3.6.3. In Python, the implementations are available in double-precision as well as arbitrary-precision floating-point arithmetic with the latter provided by the multi-precision Python library mpmath.

#### 5.1 Error Convergence Behavior

Let  $m, d \in \mathbb{N}, \alpha > 1$ , and a sequence of positive weights  $\boldsymbol{\gamma} = (\gamma_j)_{j \geq 1}$  be given. By Theorem 2, the worst-case error of a polynomial lattice point set  $P(\mathbf{g}, p) = \{\mathbf{x}_0, \dots, \mathbf{x}_{b^m-1}\}$  in base  $b = 2$  with generating vector  $\mathbf{g}$  and modulus  $p \in \mathbb{F}_2[x]$ , with  $\deg(p) = m$ , in the space  $W_{d,\boldsymbol{\gamma}}^\alpha$  is given by

$$e_{b^m,d,\alpha,\boldsymbol{\gamma}}(\mathbf{g}) = \sum_{\mathbf{0} \neq \mathbf{k} \in \mathcal{D}(\mathbf{g},p)} (r_{\alpha,\boldsymbol{\gamma}}(\mathbf{k}))^{-1} = \frac{1}{b^m} \sum_{n=0}^{b^m-1} \sum_{\mathbf{0} \neq \mathbf{k} \in \mathbb{N}_0^d} \gamma_{\text{supp}(\mathbf{k})} \frac{\text{wal}_{\mathbf{k}}(\mathbf{x}_n)}{r_\alpha(\mathbf{k})}.$$

For  $b = 2$  and product weights  $\gamma_u = \prod_{j \in u} \gamma_j$ , this expression then equals

$$e_{2^m,d,\alpha,\boldsymbol{\gamma}}(\mathbf{g}) = -1 + \frac{1}{2^m} \sum_{n=0}^{2^m-1} \prod_{j=1}^d (1 + \gamma_j \phi_\alpha(x_{n,j}))$$

with  $\phi_\alpha : [0, 1] \rightarrow \mathbb{R}$  given by

$$\phi_\alpha(x) = \begin{cases} \mu_2(\alpha), & \text{if } x = 0, \\ \mu_2(\alpha) - 2^{(1+t)(\alpha-1)}(\mu_2(\alpha) + 1), & \text{otherwise, with } t = \lfloor \log_2(x) \rfloor, \end{cases}$$

see, e.g., [4, Theorem 2], where  $\mu_b(\alpha)$  is defined as in (3). For the polynomial lattice rules constructed by the algorithms considered, we will use this worst-case error expression as a measure of quality.

In particular, we consider the convergence behavior of the worst-case error  $e_{2^m, d, \alpha, \boldsymbol{\gamma}^\alpha}(\mathbf{g})$  for generating vectors  $\mathbf{g}$  obtained by the CBC-DBD algorithm (with modulus  $p(x) = x^m$ ) and compare it with the error rates for polynomial lattice rules constructed by the standard fast CBC algorithm (with primitive polynomial  $p \in \mathbb{F}_2[x]$  of degree  $m$ ) which uses the worst-case error  $e_{2^m, d, \alpha, \boldsymbol{\gamma}^\alpha}$  as the quality criterion. We display the computation results for dimension  $d = 100$  for different sequences of product weights  $\boldsymbol{\gamma} = (\gamma_j)_{j \geq 1}$ , different values of  $m$ , and different smoothness parameters  $\alpha$ . We stress that the almost optimal error rates of  $\mathcal{O}(N^{-\alpha+\delta})$ , guaranteed by Theorem 8, may not always be visible for the weights and ranges of  $N$  considered in our numerical experiments. The graphs shown are therefore to be understood as an illustration of the pre-asymptotic behavior of the worst-case error.

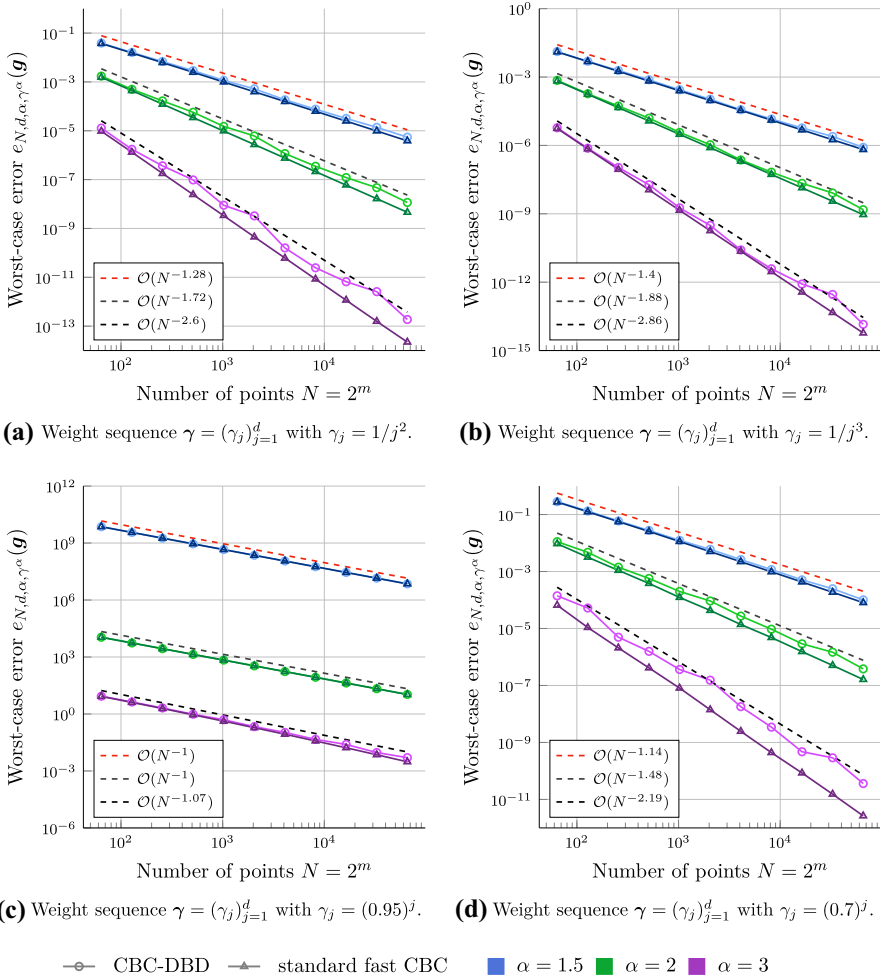
**Remark 6** We stress that in these numerical experiments we compare the CBC-DBD algorithm with modulus  $p(x) = x^m$  to the CBC construction with a primitive modulus polynomial. Both constructions yield polynomial lattices consisting of  $N = b^m$  points that have been constructed for the same function space  $W_{d, \boldsymbol{\gamma}}^\alpha$  such that the comparison is valid. To the best of our knowledge, there is no known implementation of the fast CBC algorithm for polynomial lattice rules based on the modulus  $p(x) = x^m$ . The reason for the elusiveness of such an implementation is the more involved structure of the group of units of the factor ring  $\mathbb{F}_b[x]/(x^m)$  when factored into cyclic groups, see, e.g., [24]. While for lattice rules the group of integer units modulo  $N = b^m$  is either cyclic (for odd  $b$ ) or can be factored into two cyclic subgroups (for  $b = 2$ ), which makes the corresponding generator easily computable, see, e.g., [19], the ring  $\mathbb{F}_b[x]/(x^m)$  factors into a larger number of cyclic subgroups (for sufficiently large  $m$ ) and their generating elements are less studied in the context of QMC methods.

The results in Fig. 1 show that the CBC-DBD algorithm constructs generating vectors of good polynomial lattice rules which have worst-case errors that are comparable to those of polynomial lattice rules obtained by the fast CBC algorithm. We observe identical asymptotic error rates for both algorithms considered, and also note that the CBC-DBD construction always delivers slightly higher error values. The latter behavior can easily be explained by the fact that the CBC construction is directly tailored to the space  $W_{d, \boldsymbol{\gamma}^\alpha}^\alpha$  for a particular  $\alpha$  since  $e_{b^m, d, \alpha, \boldsymbol{\gamma}^\alpha}$  is used as the quality measure. In contrast, the CBC-DBD construction is independent of the smoothness parameter  $\alpha$  and constructs polynomial lattices which have a good quality for all  $\alpha > 1$ . This in turn also means that, from a theoretical perspective, the CBC-DBD algorithm only needs to be executed once while the CBC construction has to be run for all considered  $\alpha$  in order to obtain theoretically assured error convergence rates. Additionally, we observe that the pre-asymptotic error decay is determined by the weight sequence  $\boldsymbol{\gamma} = (\gamma_j)_{j \geq 1}$ . The faster the weights  $\gamma_j$  decay, the closer the error rate is to the optimal rate of  $\mathcal{O}(N^{-\alpha})$  for the space  $W_{d, \boldsymbol{\gamma}^\alpha}^\alpha$ .

**Remark 7** In the recent article [1], it has been shown that polynomial lattice rules which were constructed for the weighted Walsh space  $W_{d, \boldsymbol{\gamma}}^\alpha$  can also achieve the



**Error convergence in the space  $W_{d,\gamma}^\alpha$  with  $d = 100, \alpha = 1.5, 2, 3$ .**



**Fig. 1** Convergence results of the worst-case error  $e_{2^m,d,\alpha,\gamma}^\alpha(\mathbf{g})$  in the weighted space  $W_{d,\gamma}^\alpha$  for smoothness parameters  $\alpha = 1.5, 2, 3$  with dimension  $d = 100$ . The generating vectors  $\mathbf{g}$  are constructed via the component-by-component digit-by-digit algorithm and the standard CBC construction for polynomial lattice rules for  $N = 2^m$ , respectively

almost optimal worst-case error convergence rate for the space  $W_{d,\gamma'}^{\alpha'}$  with different smoothness parameter  $\alpha' > 1$  and weight sequence  $\gamma'$ , provided that certain conditions on both weight sequences are satisfied. The result in [1], in particular Corollary 3, relies on a favorable relation between the different weight sequences and smoothness parameters and provides a theoretical foundation to use the standard CBC algorithm with quality measure  $e_{b^m,d,\alpha,\gamma}$  to obtain error bounds for related function space settings (and possibly different types of weights). We would like to point out that, in contrast,

**Table 1** Computation times (in s) for constructing the generating vector  $\mathbf{g}$  of a polynomial lattice rule with  $2^m$  points in  $d$  dimensions using the component-by-component digit-by-digit algorithm (**bold font**) and the standard fast CBC construction (normal font)

	$d = 50$	$d = 200$	$d = 500$	$d = 1000$	$d = 2000$
$m = 10$	0.005 <b>0.078</b>	0.017 <b>0.303</b>	0.04 <b>0.753</b>	0.075 <b>1.51</b>	0.149 <b>3.036</b>
$m = 12$	0.029 <b>0.121</b>	0.07 <b>0.482</b>	0.161 <b>1.207</b>	0.285 <b>2.388</b>	0.566 <b>4.769</b>
$m = 14$	0.101 <b>0.231</b>	0.279 <b>0.908</b>	0.658 <b>2.267</b>	1.295 <b>4.499</b>	2.519 <b>8.999</b>
$m = 16$	0.493 <b>0.641</b>	1.43 <b>2.563</b>	3.407 <b>6.549</b>	6.624 <b>12.875</b>	13.105 <b>25.628</b>
$m = 18$	2.542 <b>2.654</b>	7.617 <b>10.871</b>	17.469 <b>27.257</b>	34.07 <b>54.389</b>	67.235 <b>108.913</b>
$m = 20$	12.849 <b>19.48</b>	36.915 <b>78.477</b>	84.819 <b>195.5</b>	165.362 <b>394.627</b>	326.43 <b>784.482</b>

The highest computation time out of five independent runs is displayed. For both construction algorithms, the weight sequence  $(\gamma_j)_{j \geq 1}$  with  $\gamma_j = 1/j^2$  was used, the CBC algorithm used as additional input the smoothness parameter  $\alpha = 2$

our algorithm (when run with weights  $\boldsymbol{\gamma}$ ) is independent of  $\alpha$  and delivers QMC rules for which error bounds hold simultaneously for all  $\alpha > 1$ .

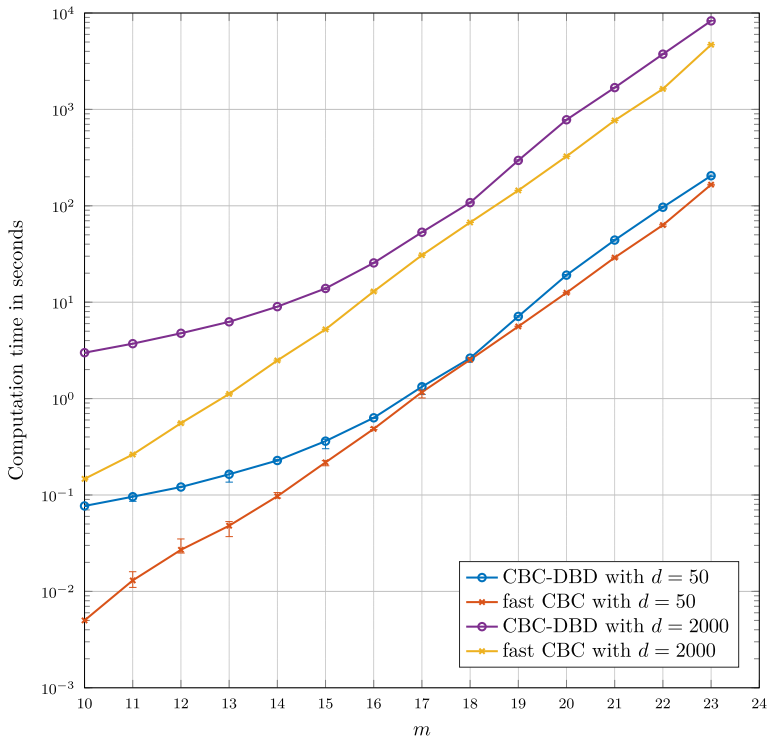
## 5.2 Computational Complexity

We demonstrate the computational complexity of Algorithm 2 which was proved in Theorem 9. For this purpose, we measure and compare the computation times of implementations of Algorithm 2 and the standard fast CBC algorithm for polynomial lattice rules with primitive modulus  $p \in \mathbb{F}_2[x]$ , cf., e.g., [19]. For all timings, we perform five independent measurements and then display either the highest time (Table 1) or the mean (Fig. 2) out of these five runs. We consider multiple values of  $m$ ,  $d \in \mathbb{N}$  and fix the positive weight sequence  $\boldsymbol{\gamma} = (\gamma_j)_{j \geq 1}$  with  $\gamma_j = 1/j^2$ . Note that the chosen weight sequence does not affect the computation times.

In Table 1, we display the timing results for the two considered algorithms. Furthermore, Fig. 2 provides a graphical illustration of the running times of both algorithms. We remark that the measured times only indicate the duration for the construction of the generating vectors but do not include the calculation of the corresponding worst-case error. All timings were performed on an Intel Core i5 CPU with 2.3 GHz using Python 3.6.3.

The timings displayed in Table 1 and Fig. 2 confirm that the computational complexity of both algorithms depends on  $m$  and  $d$  in a similar way and the measured times are in accordance with Proposition 9. Additionally, the linear dependence of the construction cost on the dimension  $d$  is well observable. The measured construction times for Algorithm 2 are slightly higher than for the fast CBC algorithm but

Mean computation times for CBC-DBD and fast CBC algorithm.



**Fig. 2** Mean computation times (in s) for constructing the generating vector  $\mathbf{g}$  of a polynomial lattice rule with  $2^m$  points in  $d \in \{50, 2000\}$  dimensions using the component-by-component digit-by-digit algorithm (circles) and the standard fast CBC construction (crosses). The vertical error bars indicate the spread between the independent timing runs

in general both algorithms can be executed in comparable time. This is especially remarkable since the fast CBC construction is based on fast Fourier transformations which rely on compiled and optimized code via Python’s Discrete Fourier Transform (numpy.fft) library while the CBC-DBD construction does not make use of any compiled libraries. Additionally, we remark that the spread between independent timing runs is neglectable, see also Fig. 2.

## 6 Conclusion

In this paper, we presented an algorithm for constructing good polynomial lattice rules for numerical integration in weighted Walsh spaces. In particular, we studied a component-by-component digit-by-digit (CBC-DBD) construction with quality measure independent of the smoothness parameter  $\alpha$ , similar to [7], where such an algorithm was analyzed for ordinary lattice rules. The construction algorithm is formulated for the special case of product weights and yields polynomial lattice rules which

admit error convergence rates that are arbitrarily close to the optimal convergence order. Furthermore, the proven error bounds become independent of the dimension if the weights satisfy suitable summability conditions. In addition to these theoretical results, we derived a fast implementation of the considered algorithm which exhibits the same computational complexity as the state-of-the-art fast CBC algorithm, but does not rely on the use of fast Fourier transformations (FFTs). The considered algorithm is, to the best of our knowledge, the first construction method for good polynomial lattice rules with modulus  $p(x) = x^m$  that requires only  $\mathcal{O}(dm2^m)$  operations. Extensive numerical experiments illustrated our findings and proved that the considered method is competitive with the standard fast CBC algorithm.

**Acknowledgements** The authors gratefully acknowledge the comments of two anonymous referees, which were very helpful in helping to improve the presentation of the results.

## References

1. Dick, J., Goda, T.: Stability of lattice rules and polynomial lattice rules constructed by the component-by-component algorithm. *J. Comput. Appl. Math.* **382**, 113062 (2021)
2. Dick, J., Kuo, F.Y., Pillichshammer, F., Sloan, I.H.: Construction algorithms for polynomial lattice rules for multivariate integration. *Math. Comput.* **74**, 1895–1921 (2005)
3. Dick, J., Kuo, F.Y., Sloan, I.H.: High-dimensional integration—the quasi-Monte Carlo way. *Acta Numer.* **22**, 133–288 (2013)
4. Dick, J., Pillichshammer, F.: Multivariate integration in weighted Hilbert spaces based on Walsh functions and weighted Sobolev spaces. *J. Complex.* **21**, 149–195 (2005)
5. Dick, J., Pillichshammer, F.: *Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration*. Cambridge University Press, Cambridge (2010)
6. Dick, J., Pillichshammer, F.: Discrepancy theory and quasi-Monte Carlo integration. In: Chen, W.W.L., Srivastav, A., Travaglino, G. (eds.) *A Panorama of Discrepancy Theory*, pp. 539–619. Springer, Cham (2014)
7. Ebert, A., Kritzer, P., Nuyens, D., Osisiogu, O.: Digit-by-digit and component-by-component constructions of lattice rules for periodic functions with unknown smoothness. *J. Complex.* (2021). **(to appear)**
8. Hickernell, F.J., Niederreiter, H.: The existence of good extensible rank-1 lattices. *J. Complex.* **19**, 286–300 (2003)
9. Hlawka, E.: Zur angenäherten Berechnung mehrfacher Integrale. *Monatshefte für Mathematik* **66**, 140–151 (1962)
10. Korobov, N.M.: Approximate evaluation of repeated integrals. *Dokl. Akad. Nauk SSSR* **124**, 1207–1210 (1959) **(in Russian)**
11. Korobov, N.M.: *Number-theoretic methods in approximate analysis*. Goz. Izdat. Fiz.-Math. (1963) **(in Russian)**
12. Korobov, N.M.: On the computation of optimal coefficients. *Dokl. Akad. Nauk SSSR* **267**, 289–292 (1982) **(in Russian)**
13. Korobov, N.M.: On the computation of optimal coefficients. *Dokl. Akad. Nauk SSSR* **26**, 590–593 (1982)
14. Niederreiter, H.: Point sets and sequences with small discrepancy. *Monatsh. Math.* **104**, 273–337 (1987)
15. Niederreiter, H.: Low-discrepancy point sets obtained by digital constructions over finite fields. *Czechoslovak Math. J.* **42**, 143–166 (1992)
16. Niederreiter, H.: *Random Number Generation and Quasi-Monte Carlo Methods*. Society for Industrial and Applied Mathematics, Philadelphia (1992)
17. Novak, E., Woźniakowski, H.: *Tractability of Multivariate Problems. Volume I: Linear Information*. EMS, Zurich (2008)

18. Nuyens, D.: The construction of good lattice rules and polynomial lattice rules. In: Kritzer, P., Niederreiter, H., Pillichshammer, F., Winterhof, A. (eds.) *Uniform Distribution and Quasi-Monte Carlo Methods: Discrepancy, Integration and Applications*, pp. 223–255. De Gruyter, Berlin (2014)
19. Nuyens, D., Cools, R.: Fast component-by-component construction, a reprise for different kernels. In: Niederreiter, H., Talay, D. (eds.) *Monte Carlo and Quasi-Monte Carlo Methods 2004*, pp. 373–387. Springer, Berlin (2006)
20. Nuyens, D., Cools, R.: Fast algorithms for component-by-component construction of rank-1 lattice rules in shift-invariant reproducing kernel Hilbert spaces. *Math. Comput.* **75**, 903–920 (2006)
21. Sloan, I.H., Joe, S.: *Lattice Methods for Multiple Integration*. Clarendon Press, Oxford (1994)
22. Sloan, I.H., Reztsov, V.A.: Component-by-component construction of good lattice rules. *Math. Comput.* **71**, 263–273 (2002)
23. Sloan, I.H., Woźniakowski, H.: When are quasi-Monte Carlo algorithms efficient for high-dimensional problems? *J. Complex.* **14**, 1–33 (1998)
24. Smith, J.L., Gallian, J.A.: Factoring finite factor rings. *Math. Mag.* **58**, 93–95 (1985)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.