Approximation of Functions Using Digital Nets

Josef Dick[1], Peter Kritzer[2], and Frances Y. Kuo[3]

[1] UNSW Asia, 1 Kay Siang Road, Singapore 248922. E-mail: `j.dick@unswasia.edu.sg`
[2] Fachbereich Mathematik, Universität Salzburg, Hellbrunnerstraße 34, A-5020 Salzburg, Austria. E-mail: `peter.kritzer@sbg.ac.at`
[3] School of Mathematics and Statistics, University of New South Wales, Sydney NSW 2052, Australia. E-mail: `f.kuo@unsw.edu.au`

Summary. In analogy to the recent paper [11] which studied lattice rule algorithms for approximation in weighted Korobov spaces, we consider the approximation problem in a weighted Hilbert space of Walsh series. Our approximation uses a truncated Walsh series with Walsh coefficients approximated by numerical integration using digital nets. We show that digital nets (or more precisely, polynomial lattices) tailored specially for the approximation problem lead to better error bounds. The error bounds can be independent of the dimension s, or depend only polynomially on s, under certain conditions on the weights defining the function space.

1 Introduction

We introduce an algorithm to approximate functions $f : [0,1]^s \to \mathbb{R}$ in certain Hilbert spaces. These spaces are in analogy to *weighted Korobov spaces* (see [24]), but instead of trigonometric functions we use *Walsh functions*, see Section 2. Recently, the approximation problem has been studied in [11] where a function from the weighted Korobov space is approximated by a truncated Fourier series, with the remaining Fourier coefficients approximated using *lattice rules*. Here, in analogy, we want to approximate functions from a Hilbert space of *Walsh series* using *digital nets* (see [19] or Section 4 below).

More precisely, every function f in our Hilbert space H_s is given by its Walsh-series representation

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

where $\mathbb{N}_0 = \{0,1,2,\ldots\}$ denotes the set of nonnegative integers, and $\hat{f}(\boldsymbol{k})$ are the *Walsh coefficients* associated with the Walsh functions $\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})$ (see (3) and (4) below). For functions $f \in H_s$, the values of $|\hat{f}(\boldsymbol{k})|$ are larger for $\boldsymbol{k}$

"closer" to $\mathbf{0}$. We introduce a set $\mathcal{A}_s$ of vectors $\boldsymbol{k} \in \mathbb{N}_0^s$ that are close to $\mathbf{0}$ in some sense, and we approximate f by the Walsh polynomial

$$F(\boldsymbol{x}) := \sum_{\boldsymbol{k} \in \mathcal{A}_s} \hat{F}(\boldsymbol{k}) \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}), \quad \text{with} \quad \hat{F}(\boldsymbol{k}) := \frac{1}{N} \sum_{n=0}^{N-1} f(\boldsymbol{x}_n) \overline{\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}_n)}, \quad (2)$$

where $\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}\} \in [0,1)^s$ is a digital net. A similar algorithm for lattice rules was proposed in [10] and has recently been studied in [11] (see also [12, 15, 30]).

It is natural to use digital nets for the approximation of the integrals arising from the Walsh coefficients, since Walsh functions are characters over the group formed by digital nets (see [7] or (13) below), which implies that the Walsh coefficients are aliased via the so-called *dual net* $\mathcal{D}$ (see [7] or (12) below), i.e., it can be shown that

$$\hat{F}(\boldsymbol{k}) = \hat{f}(\boldsymbol{k}) + \sum_{\boldsymbol{h} \in \mathcal{D}} \hat{f}(\boldsymbol{h} \oplus \boldsymbol{k}),$$

where $\oplus$ denotes digit-wise addition modulo b, and it is to act on the vectors component-wise. If the dual net $\mathcal{D}$ contains only elements $\boldsymbol{k}$ which are in some sense large and $\hat{f}(\boldsymbol{k})$ is small for large $\boldsymbol{k}$, then $\hat{F}(\boldsymbol{k})$ will be a good approximation of $\hat{f}(\boldsymbol{k})$, as $\sum_{\boldsymbol{h} \in \mathcal{D}} \hat{f}(\boldsymbol{h} \oplus \boldsymbol{k})$ is small in this case compared to $\hat{f}(\boldsymbol{k})$. Hence the Walsh polynomial $F(\boldsymbol{x})$ will give a good approximation to $f(\boldsymbol{x})$.

There are several ways of finding suitable digital nets. One choice is to construct *polynomial lattices* which are suitable for integration in the space H_s (see [5]). This way one can make use of the *weights* (see [23]), which are introduced to moderate the importance of successive variables. Another way is to use existing digital nets, say, obtained from Sobol$'$ sequences or Niederreiter sequences. The third method is to construct polynomial lattices for approximation directly. This construction is similar to the one considered in [5], but with a different quality measure which appears in the upper bound on the approximation error and, at least theoretically, yields a better approximation algorithm. This is also in analogy to the results for lattice rule algorithms in [11] for approximation in weighted Korobov spaces.

We also study *tractability* and *strong tractability* of the approximation problem in H_s. Strong tractability means that the error converges to zero with increasing N independently of the dimension s whereas tractability means that the error converges with N with at most a polynomial dependence on s. We show that our approximation algorithms based on digital nets achieve tractability or strong tractability error bounds under certain conditions on the weights. These results are again analogous to the results in [11].

This paper is organized as follows. We introduce the weighted Hilbert space of Walsh series in Section 2, and we discuss the approximation problem in Section 3. In Section 4 we review and develop results on digital nets for the integration problem that are relevant to the approximation problem. The final section, Section 5, contains the main results of this paper as discussed above.

2 Weighted Hilbert Spaces of Walsh Series

Let $b \geq 2$ be an integer – the *base*. (Later we will restrict ourselves to a *prime* base b for simplicity.) Let $\mathbb{N}_0$ denote the set of nonnegative integers.

Each $k \in \mathbb{N}_0$ has a b-adic representation $k = \sum_{i=0}^{\infty} \kappa_i b^i$, $\kappa_i \in \{0, \ldots, b-1\}$. Each $x \in [0, 1)$ has a b-adic representation $x = \sum_{i=1}^{\infty} \chi_i b^{-i}$, $\chi_i \in \{0, \ldots, b-1\}$, which is unique in the sense that infinitely many of the χ_i must differ from $b-1$. If $\kappa_a \neq 0$ is the highest nonzero digit of k, we define the *Walsh function* $\mathrm{wal}_k : [0, 1) \longrightarrow \mathbb{C}$ by

$$\mathrm{wal}_k(x) := e^{2\pi \mathtt{i}(\chi_1 \kappa_0 + \cdots + \chi_{a+1} \kappa_a)/b}. \tag{3}$$

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

$$\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}) := \prod_{j=1}^{s} \mathrm{wal}_{k_j}(x_j). \tag{4}$$

It follows from the definition above that Walsh functions are piecewise constant functions. For more information on Walsh functions, see, e.g., [2, 27].

We consider functions in a weighted Hilbert space of Walsh series. This function space was considered in [5, 6, 7, 8]; the notion of weights was first introduced in [23].

Let $\alpha > 1$, $s \geq 1$, and $b \geq 2$ be fixed. Let $\boldsymbol{\gamma} = (\gamma_j)_{j=1}^{\infty}$ be a sequence of non-increasing weights, with $0 < \gamma_j \leq 1$ for all j. The weighted Hilbert space $H_s = H_{\mathrm{wal},b,s,\alpha,\boldsymbol{\gamma}}$ is a tensor product of s one-dimensional Hilbert spaces of univariate functions, each with weight γ_j. Every function f in H_s can be written in a Walsh-series representation (1).

The inner product and norm in H_s are defined by

$$\langle f, g \rangle_{H_s} := \sum_{\boldsymbol{k} \in \mathbb{N}_0^s} r(\alpha, \boldsymbol{\gamma}, \boldsymbol{k})^{-1} \hat{f}(\boldsymbol{k}) \, \overline{\hat{g}(\boldsymbol{k})},$$

and $\|f\|_{H_s} := \langle f, f \rangle_{H_s}^{1/2}$, where $r(\alpha, \boldsymbol{\gamma}, \boldsymbol{k}) := \prod_{j=1}^{s} r(\alpha, \gamma_j, k_j)$, with

$$r(\alpha, \gamma, k) := \begin{cases} 1 & \text{if } k = 0, \\ \gamma \, b^{-\alpha \psi_b(k)} & \text{if } k \neq 0, \end{cases} \qquad \text{and} \qquad \psi_b(k) := \lfloor \log_b(k) \rfloor. \tag{5}$$

(Equivalently, $\psi_b(k) = a$ iff $\kappa_a \neq 0$ is the highest nonzero digit in the b-adic representation of $k = \sum_{i=0}^{\infty} \kappa_i b^i$.) For $x > 1$ we define

$$\mu(x) := \sum_{k=1}^{\infty} b^{-x\psi_b(k)} = (b-1)\sum_{a=0}^{\infty} b^{-(x-1)a} = \frac{b^x(b-1)}{b^x-b}. \tag{6}$$

(The equalities hold since for any $a \geq 0$ there are $b^a(b-1)$ values of $k \geq 1$ for which $\psi_b(k) = a$.) Note that $\mu(x)$ is defined in analogy to the Riemann zeta function $\zeta(x) := \sum_{k=1}^{\infty} k^{-x}$, $x > 1$.

The space H_s is a Hilbert space with the reproducing kernel (see [1, 7])

$$K(\boldsymbol{x},\boldsymbol{y}) = \sum_{\boldsymbol{k}\in\mathbb{N}_0^s} r(\alpha,\boldsymbol{\gamma},\boldsymbol{k})\,\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})\,\overline{\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{y})}.$$

The kernel satisfies the reproducing property $\langle f, K(\cdot,\boldsymbol{y})\rangle_{H_s} = f(\boldsymbol{y})$ for all $f \in H_s$ and all $\boldsymbol{y} \in [0,1)^s$.

As we have said in the introduction, we approximate functions from H_s by truncated Walsh series, see (2). Now we define precisely the set of Walsh terms to remain in the truncated Walsh series. In analogy to [11], let $M > 0$ and define

$$\mathcal{A}_s(M) := \{\boldsymbol{k} \in \mathbb{N}_0^s : r(\alpha,\boldsymbol{\gamma},\boldsymbol{k})^{-1} \leq M\}. \tag{7}$$

Following [11, Lemma 1] and its proof, we can derive a number of properties for our set $\mathcal{A}_s(M)$ here; the most important one is an upper bound on the cardinality of the set, which we state as a lemma below.

Lemma 1. (cf. [11, Lemma 1(d)]) *For any $M > 0$ we have*

$$|\mathcal{A}_s(M)| \leq M^q \prod_{j=1}^{s} \left(1 + \mu(\alpha q)\gamma_j^q\right)$$

for all $q > 1/\alpha$, where the function μ is defined in (6).

We end this section with a useful property that will be needed later. For $k, h \in \mathbb{N}_0$ with b-adic representations $k = \sum_{i=0}^{\infty} \kappa_i b^i$ and $h = \sum_{i=0}^{\infty} \hbar_i b^i$, let $\oplus$ and $\ominus$ denote digit-wise addition and subtraction modulo b, i.e.,

$$k \oplus h := \sum_{i=0}^{\infty} ((\kappa_i + \hbar_i) \bmod b)\, b^i \quad \text{and} \quad k \ominus h := \sum_{i=0}^{\infty} ((\kappa_i - \hbar_i) \bmod b)\, b^i.$$

For vectors $\boldsymbol{h}, \boldsymbol{k} \in \mathbb{N}_0^s$, the operations are defined component-wise.

Lemma 2. (cf. [21, Formula (23)]) *For any $\boldsymbol{h}, \boldsymbol{k} \in \mathbb{N}_0^s$, we have*

$$r(\alpha,\boldsymbol{\gamma},\boldsymbol{h}\oplus\boldsymbol{k}) \leq r(\alpha,\boldsymbol{\gamma},\boldsymbol{k})\,r(\alpha,\boldsymbol{\gamma},\boldsymbol{h})^{-1}.$$

Proof. It is sufficient to prove the result in one dimension, i.e., $r(\alpha,\gamma,h\oplus k) \leq r(\alpha,\gamma,k)\,r(\alpha,\gamma,h)^{-1}$. Clearly this holds when $h = 0$ or $k = 0$. When $h \neq 0$ and $k \neq 0$, we have

$$r(\alpha,\gamma,h\oplus k)\frac{r(\alpha,\gamma,h)}{r(\alpha,\gamma,k)} = \gamma\left(\frac{b^{\psi_b(k)-\psi_b(h)}}{b^{\psi_b(k\oplus h)}}\right)^{\alpha} \leq 1,$$

because $\psi_b(k) - \psi_b(h) \leq \psi_b(k \oplus h)$. This completes the proof. □

3 Approximation in the Weighted Hilbert Space H_s

We now discuss the approximation problem in the weighted Hilbert space H_s following closely the discussions from [11, 21] for the weighted Korobov space, see also [25, 28, 29] for general results.

Without loss of generality (see, e.g., [25]), we approximate f by a linear algorithm of the form

$$A_{N,s}(f) = \sum_{n=0}^{N-1} a_n L_n(f),$$

where each a_n is a function from $L_2([0,1]^s)$ and each L_n is a continuous linear functional defined on H_s from a permissible class Λ of information. We consider two classes: Λ^{all} is the class of all continuous linear functionals, while Λ^{std} is the class of standard information consisting only of function evaluations. In other words, $L_n \in \Lambda^{\mathrm{std}}$ iff there exists $\boldsymbol{x}_n \in [0,1]^s$ such that $L_n(f) = f(\boldsymbol{x}_n)$ for all $f \in H_s$. (The approximation (2) in the introduction is of the linear form above and uses standard information from Λ^{std}.)

The *worst case error* of the algorithm $A_{N,s}$ is defined as

$$e^{\mathrm{wor-app}}_{N,s}(A_{N,s}) := \sup_{\|f\|_{H_s} \le 1} \|f - A_{N,s}(f)\|_{L_2([0,1]^s)} .$$

The *initial error* associated with $A_{0,s} \equiv 0$ is

$$e^{\mathrm{wor-app}}_{0,s} := \sup_{\|f\|_{H_s} \le 1} \|f\|_{L_2([0,1]^s)} = 1,$$

where the exact value 1 is obtained by considering $f \equiv 1$. Since the initial error is conveniently 1, from this point on we omit the initial error from our discussion.

For $\varepsilon \in (0,1)$, $s \ge 1$, and $\Lambda \in \{\Lambda^{\mathrm{all}}, \Lambda^{\mathrm{std}}\}$, we define

$$N^{\mathrm{wor}}(\varepsilon, s, \Lambda) := \min \left\{ N : \exists A_{N,s} \text{ with } L_n \in \Lambda \text{ so that } e^{\mathrm{wor-app}}_{N,s}(A_{N,s}) \le \varepsilon \right\} .$$

We say that the approximation problem for the space H_s is *tractable* in the class Λ iff there are nonnegative numbers C, p, and a such that

$$N^{\mathrm{wor}}(\varepsilon, s, \Lambda) \le C \varepsilon^{-p} s^a \qquad \forall \varepsilon \in (0,1) \quad \text{and} \quad \forall s \ge 1. \tag{8}$$

The approximation problem is *strongly tractable* in the class Λ iff (8) holds with $a = 0$. In this case, the infimum of the numbers p is called the *exponent of strong tractability*, and is denoted by $p^{\mathrm{wor-app}}(\Lambda)$.

It is known from classical results (see, e.g., [25]) that the optimal algorithm in the class Λ^{all} is the truncated Walsh series

$$A^{(\mathrm{opt})}_{N,s}(f)(\boldsymbol{x}) := \sum_{\boldsymbol{k} \in \mathcal{A}_s(\varepsilon^{-2})} \hat{f}(\boldsymbol{k}) \, \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}), \qquad N = \left|\mathcal{A}_s(\varepsilon^{-2})\right| ,$$

where we have taken $M = \varepsilon^{-2}$ in (7), which ensures that the worst case error satisfies $e_{N,s}^{\mathrm{wor-app}}(A_{N,s}^{(\mathrm{opt})}) \le \varepsilon$. In fact, it is known from the general result in [28] that strong tractability and tractability in the class Λ^{all} are equivalent, and they hold iff $s_{\boldsymbol{\gamma}} < \infty$, where

$$s_{\boldsymbol{\gamma}} \;:=\; \inf\left\{\lambda > 0 : \sum_{j=1}^{s} \gamma_j^{\lambda} < \infty\right\} \tag{9}$$

is known as the *sum exponent* of the weights $\boldsymbol{\gamma} = (\gamma_j)_{j=1}^{\infty}$. Furthermore, the exponent of strong tractability is $p^{\mathrm{wor-app}}(\Lambda^{\mathrm{all}}) = 2\max(1/\alpha, s_{\boldsymbol{\gamma}})$.

For the class Λ^{std}, which is the focus of this paper, a lower bound on the worst case error for any algorithm $A_{N,s}(f) = \sum_{n=0}^{N-1} a_n f(\boldsymbol{x}_n)$ can be obtained following the argument in [21], i.e.,

$$e_{N,s}^{\mathrm{wor-app}}(A_{N,s}) \;\ge\; \sup_{\|f\|_{H_s} \le 1} \left| \int_{[0,1]^s} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \sum_{n=0}^{N-1} b_n f(\boldsymbol{x}_n) \right|,$$

where $b_n := \int_{[0,1]^s} a_n(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$. This lower bound is exactly the worst case integration error in H_s for the linear integration rule $\sum_{n=0}^{N-1} b_n f(\boldsymbol{x}_n)$. Hence the approximation problem is no easier than the integration problem in H_s, and thus the necessary condition for (strong) tractability for the integration problem in H_s is also necessary for the approximation problem.

(Strong) tractability in the weighted Hilbert space H_s for the family of equal-weight integration rules have been analyzed in [7], where it is shown that strong tractability holds iff $\sum_{j=1}^{\infty} \gamma_j < \infty$, and tractability holds iff $\limsup_{s\to\infty} \sum_{j=1}^{s} \gamma_j / \ln(s+1) < \infty$. The same conditions can be obtained for the family of linear integration rules following the argument used in [24] for the weighted Korobov space. Hence, the same conditions are necessary for (strong) tractability of the approximation problem in the class Λ^{std}. It follows from [29] that these conditions are also sufficient for (strong) tractability of approximation. Moreover, if $\sum_{j=1}^{\infty} \gamma_j < \infty$ then the exponent of strong tractability satisfies $p^{\mathrm{wor-app}}(\Lambda^{\mathrm{std}}) \in [p^{\mathrm{wor-app}}(\Lambda^{\mathrm{all}}), p^{\mathrm{wor-app}}(\Lambda^{\mathrm{all}}) + 2]$, see [29, Corollary 2(i)].

We summarize this discussion in the following theorem.

Theorem 1. *Consider the approximation problem in the worst case setting in the weighted Hilbert space H_s.*

- *Strong tractability and tractability in the class Λ^{all} are equivalent, and they hold iff $s_{\boldsymbol{\gamma}} < \infty$, where $s_{\boldsymbol{\gamma}}$ is defined in* (9)*. When this holds, the exponent of strong tractability is*

$$p^{\mathrm{wor-app}}(\Lambda^{\mathrm{all}}) \;=\; 2\max\left(\tfrac{1}{\alpha}, s_{\boldsymbol{\gamma}}\right).$$

- *The problem is strongly tractable in the class $\Lambda^{\rm std}$ iff*

$$\sum_{j=1}^{\infty} \gamma_j \;<\; \infty. \tag{10}$$

When this holds, the exponent of strong tractability satisfies

$$p^{\rm wor-app}(\Lambda^{\rm std}) \;\in\; \left[p^{\rm wor-app}(\Lambda^{\rm all}), p^{\rm wor-app}(\Lambda^{\rm all})+2\right].$$

- *The problem is tractable in the class $\Lambda^{\rm std}$ iff*

$$\ell \;:=\; \limsup_{s\to\infty} \frac{\sum_{j=1}^{s}\gamma_j}{\ln(s+1)} \;<\; \infty. \tag{11}$$

Note that when (10) holds, we have $s_{\boldsymbol{\gamma}} \le 1$. When (10) does not hold but (11) holds, we have $s_{\boldsymbol{\gamma}} = 1$.

The results for the class $\Lambda^{\rm std}$ are non-constructive. In this paper we obtain constructive algorithms based on digital nets, and we reduce the upper bound on the exponent of strong tractability to $p^{\rm wor-app}(\Lambda^{\rm std}) \le 2\,p^{\rm wor-app}(\Lambda^{\rm all})$.

4 Integration Using Digital Nets

In this section we introduce nets and review results on numerical integration rules using those point sets.

A detailed theory of (t,m,s)-nets and (t,s)-sequences was developed in [16] (see also [19, Chapter 4] and [20] for a recent survey). The (t,m,s)-nets in base b provide sets of b^m points in the s-dimensional unit cube $[0,1)^s$ which are well distributed if the quality parameter t is small.

Definition 1. *Let $b \ge 2$, $s \ge 1$ and $0 \le t \le m$ be integers. A point set P consisting of b^m points in $[0,1)^s$ forms a (t,m,s)-net in base b if every subinterval $J = \prod_{j=1}^{s}[a_j b^{-d_j}, (a_j+1)b^{-d_j}) \subseteq [0,1)^s$ of volume b^{t-m}, with integers $d_j \ge 0$ and integers $0 \le a_j < b^{d_j}$ for $1 \le j \le s$, contains exactly b^t points of P.*

In practice, all concrete constructions of (t,m,s)-nets are based on the general construction scheme of *digital nets*. To avoid too many technical notions we restrict ourselves to digital point sets defined over the finite field $\mathbb{Z}_b = \{0,1,\ldots,b-1\}$ with b *prime*. For a more general definition, see, e.g., [13, 14, 19]. Throughout the paper, $\top$ means the transpose of a vector or matrix.

Definition 2. *Let b be a prime and let $s \ge 1$ and $m \ge 1$ be integers. Let $C_1,\ldots,C_s$ be $m \times m$ matrices over the finite field $\mathbb{Z}_b$. For each $0 \le n < b^m$ with b-adic representation $n = \sum_{i=0}^{m-1}\eta_i b^i$, and each $1 \le j \le s$, we multiply the matrix C_j by the vector $(\eta_0,\ldots,\eta_{m-1})^\top \in \mathbb{Z}_b^m$, i.e.,*

$$C_j\,(\eta_0,\ldots,\eta_{m-1})^\top \;=:\; (\chi_{n,j,1},\ldots,\chi_{n,j,m})^\top \;\in\; \mathbb{Z}_b^m,$$

and set

$$x_{n,j} \;:=\; \frac{\chi_{n,j,1}}{b}+\cdots+\frac{\chi_{n,j,m}}{b^m}.$$

If the point set $\{\boldsymbol{x}_n=(x_{n,1},\ldots,x_{n,s}):0\le n<b^m\}$ *is a* (t,m,s)*-net in base* b *for some integer* t *with* $0\le t\le m$*, then it is called a digital* (t,m,s)*-net over* $\mathbb{Z}_b$*.*

See [19, Theorem 4.28] and [22] for results concerning the determination of the quality parameter t of digital nets.

Niederreiter introduced in [18] (see also [19, Section 4.4]) a special family of digital nets known now as *polynomial lattices*. In the following, let $\mathbb{Z}_b((x^{-1}))$ be the field of formal Laurent series over $\mathbb{Z}_b$, $\sum_{l=w}^{\infty} t_l x^{-l}$, where w is an arbitrary integer and all $t_l\in\mathbb{Z}_b$. Further, let $\mathbb{Z}_b[x]$ be the set of all polynomials over $\mathbb{Z}_b$, and let

$$R_{b,m} \;:=\; \{q\in\mathbb{Z}_b[x]:\deg(q)<m \text{ and } q\neq 0\}.$$

Definition 3. *Let* b *be a prime and let* $s\ge 1$ *and* $m\ge 1$ *be integers. Let* υ_m *be the map from* $\mathbb{Z}_b((x^{-1}))$ *to the interval* $[0,1)$ *defined by*

$$\upsilon_m\left(\sum_{l=w}^{\infty} t_l x^{-l}\right) \;:=\; \sum_{l=\max(1,w)}^{m} t_l b^{-l}.$$

Choose polynomials $p\in\mathbb{Z}_b[x]$ *with* $\deg(p)=m$ *and* $\boldsymbol{q}:=(q_1,\ldots,q_s)\in R_{b,m}^s$*. For each* $0\le n<b^m$ *with* b*-adic representation* $n=\sum_{i=0}^{m-1}\eta_i b^i$*, we associate* n *with the polynomial* $n(x)=\sum_{i=0}^{m-1}\eta_i x^i\in\mathbb{Z}_b[x]$*. Then the point set*

$$P_{\mathrm{PL}} \;:=\; \left\{\boldsymbol{x}_n=\left(\upsilon_m\left(\frac{n(x)q_1(x)}{p(x)}\right),\ldots,\upsilon_m\left(\frac{n(x)q_s(x)}{p(x)}\right)\right):0\le n<b^m\right\}$$

is a polynomial lattice.

We are ready to review known results on digital nets for integration. Let $P=\{\boldsymbol{x}_0,\ldots,\boldsymbol{x}_{N-1}\}$ denote a digital (t,m,s)-net over $\mathbb{Z}_b$ consisting of $N=b^m$ points. For $f\in H_s$, we approximate the integral of f by an equal-weight integration rule using the point set P. The *worst case error* of the point set P (or more precisely, of the equal-weight integration rule using the point set) for integration in the space H_s is defined by

$$e_{N,s}^{\mathrm{wor-int}}(P) \;:=\; \sup_{\|f\|_{H_s}\le 1}\left|\int_{[0,1]^s} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}-\frac{1}{N}\sum_{n=0}^{N-1} f(\boldsymbol{x}_n)\right|.$$

First we discuss the results from [7]. For $k\in\mathbb{N}_0$ with b-adic representation $k=\sum_{i=0}^{\infty}\kappa_i b^i$, we write

$$\mathrm{tr}_m(k) := (\kappa_0, \dots, \kappa_{m-1})^\top \in \mathbb{Z}_b^m$$

to denote the truncated digit vector of k. For a digital net P over $\mathbb{Z}_b$ generated by matrices $C_1, \dots, C_s$, we define the *dual net* $\mathcal{D}$ by

$$\mathcal{D} := \{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\} : C_1^\top \mathrm{tr}_m(k_1) + \cdots + C_s^\top \mathrm{tr}_m(k_s) = \boldsymbol{0}\}, \tag{12}$$

where the matrix-vector multiplications and vector additions are to be carried out in $\mathbb{Z}_b$. It is well known that Walsh functions are characters over the group formed by digital nets (see, e.g., [7]), i.e.,

$$\frac{1}{N} \sum_{n=0}^{N-1} \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}_n) = \begin{cases} 1 & \text{if } \boldsymbol{k} \in \mathcal{D} \cup \{\boldsymbol{0}\}, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

It follows easily from the character property that for any $f \in H_s$,

$$\int_{[0,1]^s} f(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} - \frac{1}{N} \sum_{n=0}^{N-1} f(\boldsymbol{x}_n) = - \sum_{\boldsymbol{k} \in \mathcal{D}} \hat{f}(\boldsymbol{k}), \tag{14}$$

and hence (see [7])

$$[e_{N,s}^{\mathrm{wor-int}}(P)]^2 = \sum_{\boldsymbol{k} \in \mathcal{D}} r(\alpha, \boldsymbol{\gamma}, \boldsymbol{k}). \tag{15}$$

4.1 Results for Polynomial Lattices

Now we discuss the results from [5] concerning polynomial lattices. We need some further notation: for every nonnegative integer $k = \sum_{i=0}^\infty \kappa_i b^i$ we define the polynomial

$$\widetilde{\mathrm{tr}}_m(k)(x) := \kappa_0 + \kappa_1 x + \cdots + \kappa_{m-1} x^{m-1} \in \mathbb{Z}_b[x],$$

and for the vector $\boldsymbol{k} = (k_1, \dots, k_s) \in \mathbb{N}_0^s$ we consider

$$\widetilde{\mathrm{tr}}_m(\boldsymbol{k}) := (\widetilde{\mathrm{tr}}_m(k_1), \dots, \widetilde{\mathrm{tr}}_m(k_s))^\top \in \mathbb{Z}_b[x]^s$$

to be a vector of polynomials. It is shown in [5] that the dual net for the polynomial lattice P_{PL}, with polynomials $p \in \mathbb{Z}_b[x]$ and $\boldsymbol{q} = (q_1, \dots, q_s) \in R_{b,m}^s$, can be expressed as

$$\mathcal{D}_{\mathrm{PL}} := \left\{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\} : \widetilde{\mathrm{tr}}_m(\boldsymbol{k}) \cdot \boldsymbol{q} \equiv 0 \pmod{p}\right\}, \tag{16}$$

where $\widetilde{\mathrm{tr}}_m(\boldsymbol{k}) \cdot \boldsymbol{q} \equiv 0 \pmod{p}$ means that the polynomial p divides the polynomial

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

The main result of [5] is summarized in the following lemma.

Lemma 3. (cf. [5, Algorithm 4.3 and Theorem 4.4]) *Given prime $b \geq 2$, positive integer m, and irreducible polynomial $p \in \mathbb{Z}_b[x]$, a vector of polynomials $\boldsymbol{q} = (q_1, \ldots, q_s) \in R_{b,m}^s$ for a polynomial lattice P_{PL} with $N = b^m$ points can be constructed by a component-by-component algorithm such that*

$$[e_{N,s}^{\mathrm{wor-int}}(P_{\mathrm{PL}})]^2 \;\leq\; (b^m - 1)^{-1/\lambda} \prod_{j=1}^{s} \left(1 + \mu(\alpha\lambda)\gamma_j^{\lambda}\right)^{1/\lambda}$$

for all $\lambda \in (1/\alpha, 1]$, where the function μ is defined in (6).

Using the property $\prod_{j=1}^s (1 + x_j) = \exp(\sum_{j=1}^s \ln(1 + x_j)) \leq \exp(\sum_{j=1}^s x_j)$ for all nonnegative x_j, we see from Lemma 3 that if $s_{\boldsymbol{\gamma}} \leq 1/\alpha$ then

$$e_{N,s}^{\mathrm{wor-int}}(P_{\mathrm{PL}}) \;=\; \mathcal{O}(N^{-\alpha/2+\delta}), \quad \delta > 0,$$

with the implied factor in the big-$\mathcal{O}$ notation is independent of N and s. This is the optimal rate of convergence for integration in H_s.

4.2 Results for General Digital Nets

For any digital (t, m, s)-net with regular generating matrices, we can obtain a worst case error bound in terms of its t-value. This is in analogy to results obtained in [3, 6, 8].

Lemma 4. *Let P be a digital (t, m, s)-net over $\mathbb{Z}_b$ with non-singular generating matrices. For each $\emptyset \neq \mathfrak{u} \subseteq \{1, \ldots, s\}$, suppose that the projection of P onto the coordinates in $\mathfrak{u}$ is a $(t_{\mathfrak{u}}, m, |\mathfrak{u}|)$-net. Then we have*

$$[e_{N,s}^{\mathrm{wor-int}}(P)]^2 \;\leq\; \frac{1}{b^{\alpha m}} \left(1 + \sum_{\emptyset \neq \mathfrak{u} \subseteq \{1,\ldots,s\}} b^{\alpha t_{\mathfrak{u}}} \prod_{j \in \mathfrak{u}} \left(b^{\alpha+1}(m+2)\mu(\alpha)\gamma_j\right) \right).$$

Proof. We start with (15) and consider all vectors $\boldsymbol{k}$ in the dual net $\mathcal{D}$ given by (12). If $\boldsymbol{k} = b^m \boldsymbol{l}$ with $\boldsymbol{l} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}$, then $\mathrm{tr}_m(k_j) = \boldsymbol{0}$ for $1 \leq j \leq s$. Otherwise we can write $\boldsymbol{k} = \boldsymbol{k}^* + b^m \boldsymbol{l}$ with $\boldsymbol{l} \in \mathbb{N}_0^s$, $\boldsymbol{k}^* = (k_1^*, \ldots, k_s^*) \neq \boldsymbol{0}$ and $0 \leq k_j^* < b^m$ for all $1 \leq j \leq s$. In the latter case we have $\mathrm{tr}_m(k_j) = \mathrm{tr}_m(k_j^*)$ for all $1 \leq j \leq s$. Thus we have (after renaming $\boldsymbol{k}^*$ to $\boldsymbol{k}$)

$$[e_{N,s}^{\mathrm{wor-int}}(P)]^2 \;=\; \sum_{\boldsymbol{l} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}} r(\alpha, \boldsymbol{\gamma}, b^m \boldsymbol{l}) + \sum_{\boldsymbol{k} \in \mathcal{D}^*} \sum_{\boldsymbol{l} \in \mathbb{N}_0^s} r(\alpha, \boldsymbol{\gamma}, \boldsymbol{k} + b^m \boldsymbol{l}) \;=:\; \Sigma_1^* + \Sigma_2^*,$$

where

$$\mathcal{D}^* \;:=\; \left\{ \boldsymbol{k} \in \{0, \ldots, b^m - 1\}^s \setminus \{\boldsymbol{0}\} : C_1^\top \mathrm{tr}_m(k_1) + \cdots + C_s^\top \mathrm{tr}_m(k_s) = \boldsymbol{0} \right\}.$$

It follows from the definition (5) that for $0 \leq k_j < b^m$ we have

$$\sum_{l=0}^{\infty} r(\alpha,\gamma_j,k_j+b^m l) = r(\alpha,\gamma_j,k_j) + \sum_{l=1}^{\infty} r(\alpha,\gamma_j,b^m l) = r(\alpha,\gamma_j,k_j) + \frac{\mu(\alpha)}{b^{m\alpha}}\gamma_j.$$

Thus

$$\Sigma_1^* = \prod_{j=1}^{s}\left(1+\frac{\mu(\alpha)}{b^{m\alpha}}\gamma_j\right) - 1$$

and

$$\begin{aligned}\Sigma_2^* &= \sum_{\boldsymbol{k}\in\mathcal{D}^*}\prod_{j=1}^{s}\left(r(\alpha,\gamma_j,k_j)+\frac{\mu(\alpha)}{b^{m\alpha}}\gamma_j\right)\\ &= \sum_{\boldsymbol{k}\in\mathcal{D}^*} r(\alpha,\boldsymbol{\gamma},\boldsymbol{k}) + \sum_{\mathfrak{u}\subsetneq\{1,\ldots,s\}}\left[\left(\sum_{\boldsymbol{k}\in\mathcal{D}^*}\prod_{j\in\mathfrak{u}} r(\alpha,\gamma_j,k_j)\right)\prod_{j\notin\mathfrak{u}}\left(\frac{\mu(\alpha)\gamma_j}{b^{m\alpha}}\right)\right]. \end{aligned} \quad (17)$$

First we investigate the sum $\sum_{\boldsymbol{k}\in\mathcal{D}^*}\prod_{j\in\mathfrak{u}} r(\alpha,\gamma_j,k_j)$ where $\mathfrak{u}$ is a proper subset of $\{1,\ldots,s\}$. Let $\boldsymbol{k}=(k_1,\ldots,k_s)\in\{0,\ldots,b^m-1\}^s\setminus\{\boldsymbol{0}\}$ and $j_0\notin\mathfrak{u}$. Since the generating matrices $C_1,\ldots,C_s$ are non-singular, for any combination of the $s-1$ components $k_j\in\{0,\ldots,b^m-1\}$ with $j\neq j_0$, there is exactly one value of $k_{j_0}\in\{0,\ldots,b^m-1\}$ which ensures that $\boldsymbol{k}\in\mathcal{D}^*$. Hence we have

$$\begin{aligned}\sum_{\boldsymbol{k}\in\mathcal{D}^*}\prod_{j\in\mathfrak{u}} r(\alpha,\gamma_j,k_j) &= b^{m(s-|\mathfrak{u}|-1)}\prod_{j\in\mathfrak{u}}\left(\sum_{k=0}^{b^m-1} r(\alpha,\gamma_j,k)\right)-1\\ &\le b^{m(s-|\mathfrak{u}|-1)}\prod_{j\in\mathfrak{u}}(1+\mu(\alpha)\gamma_j)-1,\end{aligned}$$

from which we can show that the second term in (17) is bounded by

$$\frac{1}{b^{\alpha m}}\prod_{j=1}^{s}(1+2\mu(\alpha)\gamma_j) - \Sigma_1^*.$$

It remains to obtain a bound on the first term in (17). Here we only outline the most important steps; the details follow closely the proofs of [3, Lemma 7] and [8, Lemma 7], see also [6, Lemma 3].

We have

$$\sum_{\boldsymbol{k}\in\mathcal{D}^*} r(\alpha,\boldsymbol{\gamma},\boldsymbol{k}) = \sum_{\substack{\emptyset\neq\mathfrak{u}\subseteq\{1,\ldots,s\}\\ \mathfrak{u}=\{u_1,\ldots,u_e\}}} \sum_{\substack{k_{u_1},\ldots,k_{u_e}=1\\ C_{u_1}^\top \mathrm{tr}_m(k_{u_1})+\cdots+C_{u_e}^\top \mathrm{tr}_m(k_{u_e})=\boldsymbol{0}}}^{b^m-1} \prod_{j\in\mathfrak{u}} r(\alpha,\gamma_j,k_j). \quad (18)$$

The $\mathfrak{u}=\{1,\ldots,s\}$ term in (18) is

$$\sum_{\substack{k_1,\ldots,k_s=1 \\ C_1^\top \mathrm{tr}_m(k_1)+\cdots+C_s^\top \mathrm{tr}_m(k_s)=\mathbf{0}}}^{b^m-1} \prod_{j=1}^{s} r(\alpha,\gamma_j,k_j)$$

$$= \sum_{v_1,\ldots,v_s=0}^{m-1} \frac{\prod_{j=1}^s \gamma_j}{b^{\alpha(v_1+\cdots+v_s)}} \sum_{l_1,\ldots,l_s=1}^{b-1} \underbrace{\sum_{k_1=l_1 b^{v_1}}^{(l_1+1)b^{v_1}-1} \cdots \sum_{k_s=l_s b^{v_s}}^{(l_s+1)b^{v_s}-1}}_{C_1^\top \mathrm{tr}_m(k_1)+\cdots+C_s^\top \mathrm{tr}_m(k_s)=\mathbf{0}} 1.$$

Using the fact that P is a digital (t,m,s)-net, it can be shown that

$$\underbrace{\sum_{k_1=l_1 b^{v_1}}^{(l_1+1)b^{v_1}-1} \cdots \sum_{k_s=l_s b^{v_s}}^{(l_s+1)b^{v_s}-1}}_{C_1^\top \mathrm{tr}_m(k_1)+\cdots+C_s^\top \mathrm{tr}_m(k_s)=\mathbf{0}} 1 \le (b-1)^s \sum_{\substack{v_1,\ldots,v_s=0 \\ m-t-s+1\le v_1+\cdots+v_s\le m-t}}^{m-1} \frac{\prod_{j=1}^s \gamma_j}{b^{\alpha(v_1+\cdots+v_s)}}$$

$$+ (b-1)^s \sum_{\substack{v_1,\ldots,v_s=0 \\ v_1+\cdots+v_s>m-t}}^{m-1} \frac{\prod_{j=1}^s \gamma_j}{b^{\alpha(v_1+\cdots+v_s)}} b^{v_1+\cdots+v_s-m+t}.$$

The $\emptyset \ne \mathfrak{u} \subsetneq \{1,\ldots,s\}$ terms in (18) can be estimated in a similar way by making use of the fact that the projection of P onto the coordinates in $\mathfrak{u}$ is a digital $(t_{\mathfrak{u}},m,|\mathfrak{u}|)$-net. Combining all the terms together, we finally obtain

$$\sum_{\boldsymbol{k}\in\mathcal{D}^*} r(\alpha,\boldsymbol{\gamma},\boldsymbol{k}) \le \sum_{\emptyset\ne\mathfrak{u}\subseteq\{1,\ldots,s\}} \left(\frac{b-1}{b^{\alpha-1}-1}\right)^{|\mathfrak{u}|} \frac{2(m-t_{\mathfrak{u}}+2)^{|\mathfrak{u}|-1}}{b^{\alpha(m-t_{\mathfrak{u}}+1-2|\mathfrak{u}|)}} \prod_{j\in\mathfrak{u}} \gamma_j$$

$$\le \frac{1}{b^{\alpha(m+1)}} \sum_{\emptyset\ne\mathfrak{u}\subseteq\{1,\ldots,s\}} b^{\alpha t_{\mathfrak{u}}} \prod_{j\in\mathfrak{u}} \left(b^{\alpha+1}(m+2)\mu(\alpha)\gamma_j\right),$$

from which the result can be derived. □

We now give two examples of digital nets for which explicit bounds on the values of $t_{\mathfrak{u}}$ are known. Let P_{Sob} and P_{Nie} denote the digital net generated by the modified (as discussed below) left upper $m \times m$ sub-matrices of the generating matrices of *Sobol′ sequences* and *Niederreiter sequences* (which are examples of *digital* (t,s)*-sequences*, see, e.g., [18]), respectively. We need to modify the generating matrices to make them regular; this can be achieved by changing the least significant rows of the matrices without influencing the digital net property nor the quality parameter of the net and its projections.

Lemma 5. *Let* $P \in \{P_{\mathrm{Sob}}, P_{\mathrm{Nie}}\}$ *be a digital* (t,m,s)*-net over* $\mathbb{Z}_b$ *obtained from either a Sobol′ sequence* $(b=2)$ *or a Niederreiter sequence. We have*

$$[e_{N,s}^{\mathrm{wor-int}}(P_{\mathrm{Sob}})]^2$$
$$\le \frac{1}{2^{\alpha m}} \prod_{j=1}^{s} \left(2^{\alpha c+1} \left(j \log_2(j+1) \log_2 \log_2(j+3)\right)^{\alpha} (m+2)\mu(\alpha)\gamma_j\right),$$

where c is some constant independent of all parameters, and

$$[e^{\mathrm{wor-int}}_{N,s}(P_{\mathrm{Nie}})]^2 \;\le\; \frac{1}{b^{\alpha m}} \prod_{j=1}^{s} \left(b^{2\alpha+1} \left(j \log_b(j+b)\right)^{\alpha} (m+2)\mu(\alpha)\gamma_j \right).$$

If

$$\begin{cases} \sum_{j=1}^{\infty} (j \ln j \, \ln\ln j)^{\alpha}\gamma_j \;<\; \infty & \textit{when } P = P_{\mathrm{Sob}}, \\ \sum_{j=1}^{\infty} (j \ln j)^{\alpha}\gamma_j \;<\; \infty & \textit{when } P = P_{\mathrm{Nie}}, \end{cases} \tag{19}$$

then

$$[e^{\mathrm{wor-int}}_{N,s}(P)]^2 \;\le\; C_\delta \, N^{-\alpha+\delta}, \quad C_\delta \in \{C_{\mathrm{Sob},\delta}, C_{\mathrm{Nie},\delta}\}, \quad \delta > 0,$$

where $C_{\mathrm{Sob},\delta}$ and $C_{\mathrm{Nie},\delta}$ are independent of m and s but depend on δ, b, α, and $\boldsymbol{\gamma}$.

Proof. The construction of Sobol′ sequences makes use of primitive polynomials in base $b = 2$, one polynomial p_j for each dimension j, with non-decreasing degrees as the dimension increases. It is known (see e.g. [26]) that $t_{\mathfrak{u}} = \sum_{j\in\mathfrak{u}} (\deg(p_j) - 1)$ and $\deg(p_j) \le \log_2 j + \log_2 \log_2(j+1) + \log_2\log_2\log_2(j+3) + c$, where c is a constant independent of j. (Note that the above formula for $t_{\mathfrak{u}}$ is associated with the whole sequence; for a net of b^m points $t_{\mathfrak{u}}$ is bounded by the minimum of m and the given formula.) Thus we have

$$b^{t_{\mathfrak{u}}} \;=\; 2^{t_{\mathfrak{u}}} \;\le\; \prod_{j\in\mathfrak{u}} \left(2^{c-1} j \log_2(j+1) \log_2\log_2(j+3) \right).$$

On the other hand, the construction of Niederreiter sequences makes use of monic irreducible polynomials, and it is known (see [26, Lemma 2]) that $\deg(p_j) \le \log_b j + \log_b \log_b(j+b) + 2$. Thus in this case

$$b^{t_{\mathfrak{u}}} \;\le\; \prod_{j\in\mathfrak{u}} \left(bj \log_b(j+b) \right).$$

Substituting these bounds on $b^{t_{\mathfrak{u}}}$ into Lemma 4 proves the first part of this lemma.

To prove the second part of this lemma, we follow closely the proof of [9, Lemma 3]. Consider first the Niederreiter sequence and suppose that $\sum_{j=1}^{\infty} (j \ln j)^{\alpha}\gamma_j < \infty$. For $k \ge 0$, define $\sigma_k := b^{2\alpha+1}\mu(\alpha) \sum_{j=k+1}^{\infty} (j \log_b(j+b))^{\alpha}\gamma_j$. Then we have

$$\prod_{j=1}^{s} \left(b^{2\alpha+1} \left(j \log_b(j+b)\right)^{\alpha} (m+2)\mu(\alpha)\gamma_j \right) \;\le\; (1+\sigma_k^{-1})^k \, b^{(m+2)\sigma_k(\sigma_0+1)}.$$

Let $\delta > 0$ and choose k_δ such that $\sigma_{k_\delta}(\sigma_0 + 1) \le \delta$. The desired result is obtained with $C_{\mathrm{Nie},\delta} := b^{2\delta}(1+\sigma_{k_\delta}^{-1})^{k_\delta}$. The result for the Sobol′ sequence can be obtained in the same way. □

5 Approximation Using Digital Nets

Now we formalize the approximation algorithm (2). For $M > 0$ and $P = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}\}$ a digital net with $N = b^m$ points, we define

$$A_{N,s,M}(f) := \sum_{\boldsymbol{h} \in \mathcal{A}_s(M)} \left(\frac{1}{N} \sum_{n=0}^{N-1} f(\boldsymbol{x}_n) \overline{\mathrm{wal}_{\boldsymbol{h}}(\boldsymbol{x}_n)} \right) \mathrm{wal}_{\boldsymbol{h}}(\boldsymbol{x}).$$

Recall that the *worst case error* for the algorithm $A_{N,s,M}$ using the point set P is defined by

$$\mathrm{e}^{\mathrm{wor-app}}_{N,s,M}(P) = \mathrm{e}^{\mathrm{wor-app}}_{N,s,M}(A_{N,s,M}) := \sup_{\|f\|_{H_s} \le 1} \|f - A_{N,s,M}(f)\|_{L_2([0,1]^s)} .$$

We have

$$\begin{aligned}
&\|f - A_{N,s,M}(f)\|^2_{L_2([0,1]^s)} \\
&= \sum_{\boldsymbol{h} \notin \mathcal{A}_s(M)} |\hat{f}(\boldsymbol{h})|^2 + \sum_{\boldsymbol{h} \in \mathcal{A}_s(M)} \left| \int_{[0,1]^s} f(\boldsymbol{x}) \overline{\mathrm{wal}_{\boldsymbol{h}}(\boldsymbol{x})} \,\mathrm{d}\boldsymbol{x} - \frac{1}{N} \sum_{n=0}^{N-1} f(\boldsymbol{x}_n) \overline{\mathrm{wal}_{\boldsymbol{h}}(\boldsymbol{x}_n)} \right|^2 \\
&\le \frac{1}{M} \|f\|^2_{H_s} + \sum_{\boldsymbol{h} \in \mathcal{A}_s(M)} |\langle f, \tau_{\boldsymbol{h}} \rangle_{H_s}|^2 ,
\end{aligned}$$

where

$$\tau_{\boldsymbol{h}}(\boldsymbol{t}) := \int_{[0,1]^s} K(\boldsymbol{t}, \boldsymbol{x}) \mathrm{wal}_{\boldsymbol{h}}(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} - \frac{1}{N} \sum_{n=0}^{N-1} K(\boldsymbol{t}, \boldsymbol{x}_n) \mathrm{wal}_{\boldsymbol{h}}(\boldsymbol{x}_n).$$

Hence

$$\mathrm{e}^{\mathrm{wor-app}}_{N,s,M}(P) = \left(\frac{\beta}{M} + \sup_{\|f\|_{H_s} \le 1} \sum_{\boldsymbol{h} \in \mathcal{A}_s(M)} |\langle f, \tau_{\boldsymbol{h}} \rangle_{H_s}|^2 \right)^{1/2}$$

for some $\beta \in [0,1]$. Moreover, it can be shown that the second term involving the supremum is essentially the spectral radius ρ of the matrix T_P whose entries are given by $\langle \tau_{\boldsymbol{h}}, \tau_{\boldsymbol{p}} \rangle_{H_s}$.

Using (14), it can be shown that

$$\begin{aligned}
\tau_{\boldsymbol{h}}(\boldsymbol{t}) &= - \sum_{\boldsymbol{k} \in \mathcal{D}} r(\alpha, \boldsymbol{\gamma}, \boldsymbol{h} \ominus \boldsymbol{k}) \, \mathrm{wal}_{\boldsymbol{h} \ominus \boldsymbol{k}}(\boldsymbol{t}) \\
&= - \sum_{\substack{\boldsymbol{q} \in \mathbb{N}_0^s \setminus \{\boldsymbol{h}\} \\ C_1^\top \mathrm{tr}_m(h_1 \ominus q_1) + \cdots + C_s^\top \mathrm{tr}_m(h_s \ominus q_s) = \boldsymbol{0}}} r(\alpha, \boldsymbol{\gamma}, \boldsymbol{q}) \, \mathrm{wal}_{\boldsymbol{q}}(\boldsymbol{t}).
\end{aligned}$$

Consequently,

$$\langle \tau_{\boldsymbol{h}}, \tau_{\boldsymbol{p}} \rangle_{H_s} = \begin{cases} 0 & \text{if } C_1^\top \mathrm{tr}_m(h_1 \ominus p_1) + \cdots + C_s^\top \mathrm{tr}_m(h_s \ominus p_s) \neq \boldsymbol{0}, \\ \displaystyle\sum_{\substack{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}, \boldsymbol{p} \ominus \boldsymbol{h}\} \\ C_1^\top \mathrm{tr}_m(k_1) + \cdots + C_s^\top \mathrm{tr}_m(k_s) = \boldsymbol{0}}} r(\alpha, \boldsymbol{\gamma}, \boldsymbol{h} \oplus \boldsymbol{k}) & \text{otherwise.} \end{cases} \tag{20}$$

We state the result in the following lemma.

Lemma 6. (cf. [11, Lemma 2]) *The worst case error for the approximation algorithm $A_{N,s,M}$ using a digital net P satisfies*

$$e_{N,s,M}^{\mathrm{wor-app}}(P) = \left(\frac{\beta}{M} + \rho(T_P) \right)^{1/2} \quad \textit{for some} \quad \beta \in [0,1],$$

where T_P is a nonnegative-definite symmetric $|\mathcal{A}_s(M)| \times |\mathcal{A}_s(M)|$ matrix with entries given by $\langle \tau_{\boldsymbol{h}}, \tau_{\boldsymbol{p}} \rangle_{H_s}$ in (20) for $\boldsymbol{h}, \boldsymbol{p} \in \mathcal{A}_s(M)$.

Unfortunately we do not have a computable expression for the spectral radius $\rho(T_P)$. Therefore we consider its upper bound, the trace of T_P,

$$\rho(T_P) \le \mathrm{trace}(T_P) = \sum_{\boldsymbol{h} \in \mathcal{A}_s(M)} \sum_{\boldsymbol{k} \in \mathcal{D}} r(\alpha, \boldsymbol{\gamma}, \boldsymbol{h} \oplus \boldsymbol{k}). \tag{21}$$

5.1 Nets Constructed for Integration

A natural question to ask is: how good are the nets constructed for integration when they are used for approximation? To relate the worst case error for approximation $e_{N,s}^{\mathrm{wor-app}}(P)$ to the worst case error for integration $e_{N,s}^{\mathrm{wor-int}}(P)$, we apply Lemma 2 to (21) and obtain

$$\rho(T_P) \le \sum_{\boldsymbol{h} \in \mathcal{A}_s(M)} \frac{1}{r(\alpha, \boldsymbol{\gamma}, \boldsymbol{h})} \sum_{\boldsymbol{k} \in \mathcal{D}} r(\alpha, \boldsymbol{\gamma}, \boldsymbol{k}) \le M |\mathcal{A}_s(M)| [e_{N,s}^{\mathrm{wor-int}}(P)]^2.$$

Hence it follows from Lemma 6 that

$$e_{N,s,M}^{\mathrm{wor-app}}(P) \le \left(\frac{1}{M} + M |\mathcal{A}_s(M)| [e_{N,s}^{\mathrm{wor-int}}(P)]^2 \right)^{1/2}. \tag{22}$$

Applying Lemmas 1 and 3 to (22), we obtain the following result for polynomial lattices constructed for the integration problem.

Lemma 7. (cf. [11, Lemma 3]) *Let P_{PL} be a polynomial lattice constructed component-by-component for integration. Then the worst case error for the approximation algorithm $A_{N,s,M}$ using P_{PL} satisfies*

$$e_{N,s,M}^{\mathrm{wor-app}}(P_{\mathrm{PL}}) \le \left(\frac{1}{M} + \frac{C_{s,q,\lambda} M^{q+1}}{(N-1)^{1/\lambda}} \right)^{1/2}$$

for all $q > 1/\alpha$ *and* $\lambda \in (1/\alpha, 1]$, *where*

$$C_{s,q,\lambda} := \prod_{j=1}^{s} \left(1 + \mu(\alpha\lambda)\gamma_j^{\lambda}\right)^{1/\lambda} \left(1 + \mu(\alpha q)\gamma_j^{q}\right).$$

Given $\varepsilon \in (0,1)$, we want to find small M and $N = b^m$ for which the error bound in Lemma 7 is at most ε. To ensure that the two terms in the error bound are of the same order, we first choose $M = 2\varepsilon^{-2}$, and then choose N such that the second term is no more than the first term. Hence it is sufficient that we take $N = b^m$ with

$$m = \left\lceil \log_b \left(\left(C_{s,q,\lambda} M^{q+2} \right)^{\lambda} + 1 \right) \right\rceil. \tag{23}$$

Using the property $\prod_{j=1}^{s}(1 + x_j) = \exp(\sum_{j=1}^{s} \log(1 + x_j)) \le \exp(\sum_{j=1}^{s} x_j)$ for nonnegative x_j, we can write

$$C_{s,q,\lambda} \le \exp\left(\frac{\mu(\alpha\lambda)}{\lambda} \sum_{j=1}^{s} \gamma_j^{\lambda} + \mu(\alpha q) \sum_{j=1}^{s} \gamma_j^{q} \right) \tag{24}$$

$$= (s+1)^{\mu(\alpha\lambda)\lambda^{-1} \sum_{j=1}^{s} \gamma_j^{\lambda} / \ln(s+1) + \mu(\alpha q) \sum_{j=1}^{s} \gamma_j^{q} / \ln(s+1)}. \tag{25}$$

Let $p^* = 2\max(1/\alpha, s_{\boldsymbol{\gamma}})$. When (10) holds but $s_{\boldsymbol{\gamma}} = 1$, we have $p^* = 2$ and we take $q = \lambda = p^*/2 = 1$. Then we see from (24) that $\sup_{s\ge1} C_{s,q,\lambda} < \infty$. When (10) holds and $s_{\boldsymbol{\gamma}} < 1$, we have $p^* < 2$ and we choose $q = \lambda = p^*/2 + \delta$ for some $\delta > 0$. Then $\mu(\alpha\lambda) < \infty$ and $\sum_{j=1}^{\infty} \gamma_j^{\lambda} < \infty$, and once again we see from (24) that $\sup_{s\ge1} C_{s,q,\lambda} < \infty$. In both cases, we see from (23) that $N = \mathcal{O}(\varepsilon^{-p})$, with p equal to or arbitrarily close to $2p^* + {p^*}^2/2$ as δ goes to 0.

When (11) holds but not (10), we have $s_{\boldsymbol{\gamma}} = 1$. We take $q = \lambda = 1$ and it follows from (23) and (25) that $N = \mathcal{O}(\varepsilon^{-6})$ and $C_{s,q,\lambda} = \mathcal{O}(s^a)$, with a arbitrarily close to $2\mu(\alpha)\ell$.

We summarize the analysis in the following theorem.

Theorem 2. (cf. [11, Theorem 1]) *Let* P_{PL} *be a polynomial lattice constructed component-by-component for integration. For* $\varepsilon \in (0,1)$ *set* $M = 2\varepsilon^{-2}$. *If* (10) *holds, then the approximation algorithm* $A_{N,s,M}$ *using* P_{PL} *achieves the worst case error bound* $e^{\mathrm{wor-app}}_{N,s,M}(P_{\mathrm{PL}}) \le \varepsilon$ *using* $N = \mathcal{O}(\varepsilon^{-p})$ *function values, with* p *equal to or arbitrarily close to*

$$2\, p^{\mathrm{wor-app}}(\Lambda^{\mathrm{all}}) + \frac{[p^{\mathrm{wor-app}}(\Lambda^{\mathrm{all}})]^2}{2}.$$

If (10) *does not hold but* (11) *holds, then the error bound* $e^{\mathrm{wor-app}}_{N,s,M}(P_{\mathrm{PL}}) \le \varepsilon$ *is achieved using* $N = \mathcal{O}(s^a \varepsilon^{-6})$ *function values, with* a *arbitrarily close to* $2\mu(\alpha)\ell$. *The implied factors in the big* $\mathcal{O}$*-notations are independent of* ε *and* s.

Now we use Lemmas 1 and 5 in (22) to derive results for digital nets obtained from Sobol′ sequence or Niederreiter sequence.

Lemma 8. *Let $P \in \{P_{\mathrm{Sob}}, P_{\mathrm{Nie}}\}$ be a digital net obtained from either a Sobol′ or a Niederreiter sequence. If* (19) *holds, then the worst case error for the approximation algorithm $A_{N,s,M}$ using P satisfies*

$$e_{N,s,M}^{\mathrm{wor-app}}(P) \le \left(\frac{1}{M} + \frac{\bar{C}_{s,q,\delta} M^{q+1}}{N^{\alpha-\delta}}\right)^{1/2}, \quad \bar{C}_{s,q,\delta} := C_\delta \prod_{j=1}^{s} \left(1 + \mu(\alpha q)\gamma_j^q\right),$$

for all $q > 1/\alpha$ and $\delta > 0$, with $C_\delta \in \{C_{\mathrm{Sob},\delta}, C_{\mathrm{Nie},\delta}\}$ given in Lemma 5.

Note that both conditions on the weights in (19) imply (10) as well as $s_{\boldsymbol{\gamma}} \le 1/\alpha$. For $\varepsilon \in (0,1)$ we take $q = 1/\alpha + \delta$, $M = 2\varepsilon^{-2}$ and $N = b^m$ with

$$m = \left\lceil \log_b \left(\left(\bar{C}_{s,q,\delta} M^{q+2}\right)^{1/(\alpha-\delta)} \right) \right\rceil.$$

Then we have $\sup_{s \ge 1} \bar{C}_{s,q,\delta} < \infty$ and $N = \mathcal{O}(\varepsilon^{-p})$, with p arbitrarily close to $4/\alpha + 2/\alpha^2$ as δ goes to 0. This is summarized in the theorem below.

Theorem 3. *Let $P \in \{P_{\mathrm{Sob}}, P_{\mathrm{Nie}}\}$ be a digital net obtained from either a Sobol′ or a Niederreiter sequence. For $\varepsilon \in (0,1)$ set $M = 2\varepsilon^{-2}$. If* (19) *holds, then the approximation algorithm $A_{N,s,M}$ using P achieves the worst case error bound $e_{N,s,M}^{\mathrm{wor-app}}(P) \le \varepsilon$ using $N = \mathcal{O}(\varepsilon^{-p})$ function values, with p arbitrarily close to*

$$2\, p^{\mathrm{wor-app}}(\Lambda^{\mathrm{all}}) + \frac{[p^{\mathrm{wor-app}}(\Lambda^{\mathrm{all}})]^2}{2}.$$

The implied factor in the big $\mathcal{O}$-notation is independent of ε and s.

5.2 Polynomial Lattices Constructed for Approximation

In this section we study polynomial lattices with the generating polynomials specially constructed for the approximation problem. It is perhaps not surprising that such polynomial lattices yield smaller error bounds than those studied in the previous subsection.

Since $M\, r(\alpha, \boldsymbol{\gamma}, \boldsymbol{h}) \ge 1$ for all $\boldsymbol{h} \in \mathcal{A}_s(M)$, we have from (21) that $\rho(T_P) \le M\, S_{N,s}(P)$, where

$$S_{N,s}(P) := \sum_{\boldsymbol{h} \in \mathbb{N}_0^s} \sum_{\boldsymbol{k} \in \mathcal{D}} r(\alpha, \boldsymbol{\gamma}, \boldsymbol{h})\, r(\alpha, \boldsymbol{\gamma}, \boldsymbol{k} \oplus \boldsymbol{h}). \tag{26}$$

Thus it follows from Lemma 6 that

$$e_{N,s,M}^{\mathrm{wor-app}}(P) \le \left(\frac{1}{M} + M\, S_{N,s}(P)\right)^{1/2}. \tag{27}$$

An analogous expression to $S_{N,s}(P)$ for lattice rule algorithms in weighted Korobov spaces was considered in [4] for some integral equation problem. (It is advocated that the expression $S_{n,d}(\boldsymbol{z})$ in [4] should be considered instead of the quantity $E_{n,d}(\boldsymbol{z})$ in [11] for the approximation problem.) Observe that the quantity $S_{N,s}(P)$ depends only on the digital net $P = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}\}$ and does not depend on the value of M nor the set $\mathcal{A}_s(M)$. Following [7, Proof of Theorem 2]), we can rewrite $S_{N,s}(P)$ in an easily computable form

$$\begin{aligned} S_{N,s}(P) \;=\; &-\prod_{j=1}^{s}\left(1+\mu(2\alpha)\gamma_j^2\right) + \frac{1}{N}\prod_{j=1}^{s}\left(1+\mu(\alpha)\gamma_j\right)^2 \\ &+ \frac{1}{N}\sum_{n=1}^{N-1}\prod_{j=1}^{s}\left(1+\omega(x_{n,j})\gamma_j\right)^2, \end{aligned}$$

where $\omega(0) = \mu(\alpha)$, and $\omega(x) = \mu(\alpha) - b^{(a-1)(1-\alpha)}(\mu(\alpha)+1)$ if $x \neq 0$ and $\chi_a \neq 0$ is the first nonzero digit in the b-adic representation $x = \sum_{i=0}^{\infty} \chi_i b^{-i}$.

Let p be an irreducible polynomial of degree m. We wish to construct a vector of polynomials $\boldsymbol{q} = (q_1, \ldots, q_s)$ for a polynomial lattice P_{PL}, one polynomial at a time, such that the quantity $S_{N,s}(q_1, \ldots, q_s) = S_{N,s}(P)$ is as small as possible.

Algorithm 1 *Let $m \geq 1$ and $N = b^m$. Let p be an irreducible polynomial in $\mathbb{Z}_b[x]$ with $\deg(p) = m$.*

1. *Set $q_1 = 1$.*
2. *For $d = 2, 3, \ldots, s$ find q_d in $R_{b,m}$ to minimize $S_{N,s}(q_1, \ldots, q_{d-1}, q_d)$.*

Lemma 9. (cf. [11, Lemma 6]) *Let P^*_{PL} denote the polynomial lattice constructed by Algorithm 1. Then the worst case error for the approximation algorithm $A_{N,s,M}$ using P^*_{PL} satisfies*

$$e^{\mathrm{wor-app}}_{N,s,M}(P^*_{\mathrm{PL}}) \;\leq\; \left(\frac{1}{M} + \frac{\widetilde{C}_{s,\lambda,\delta} M}{N^{1/\lambda}}\right)^{1/2}$$

for all $\lambda \in (1/\alpha, 1]$ and $\delta > 0$, where

$$\widetilde{C}_{s,\lambda,\delta} \;:=\; \frac{1}{\delta}\prod_{j=1}^{s}\left(1+(1+\delta^\lambda)\mu(\alpha\lambda)\gamma_j^\lambda\right)^{2/\lambda}.$$

Proof. We prove by induction that the polynomials $q_1^*, \ldots, q_s^*$ for a polynomial lattice P^*_{PL} constructed by Algorithm 1 satisfy, for each $d = 1, \ldots, s$,

$$S_{N,d}(q_1^*, \ldots, q_d^*) \leq \widetilde{C}_{d,\lambda,\delta} N^{-1/\lambda} \tag{28}$$

for all $\lambda \in (1/\alpha, 1]$ and $\delta > 0$. Our proof follows the argument used in the proofs of [4, Lemma 4] and [11, Lemma 6]. We present here only a skeleton proof; the technical details can be verified in analogy to [4, 11].

The $d = 1$ case can easily be verified. Suppose now that $\boldsymbol{q}^* = (q_1^*, \ldots, q_d^*) \in R_{b,m}^d$ is chosen according to Algorithm 1 and that $S_{N,d}(\boldsymbol{q}^*)$ satisfies (28) for all $\lambda \in (1/\alpha, 1]$ and $\delta > 0$. By separating the $k_{d+1} = 0$ and $k_{d+1} \neq 0$ terms in (26) (with the dual net $\mathcal{D}$ replaced by $\mathcal{D}_{\mathrm{PL}}$ in (16)), we can write

$$S_{N,d+1}(\boldsymbol{q}^*, q_{d+1}) \;=\; \phi(\boldsymbol{q}^*) + \theta(\boldsymbol{q}^*, q_{d+1}),$$

where

$$\begin{aligned}\phi(\boldsymbol{q}^*) &= \sum_{h_{d+1}\in\mathbb{N}_0} r^2(\alpha, \gamma_{d+1}, h_{d+1}) \sum_{\boldsymbol{h}\in\mathbb{N}_0^d} \sum_{\substack{\boldsymbol{k}\in\mathbb{N}_0^d\setminus\{\boldsymbol{0}\} \\ \widetilde{\mathrm{tr}}_m(\boldsymbol{k})\cdot\boldsymbol{q}^*\equiv 0 \pmod p}} r(\alpha, \boldsymbol{\gamma}, \boldsymbol{h}) r(\alpha, \boldsymbol{\gamma}, \boldsymbol{k}\oplus\boldsymbol{h}) \\ &= (1 + \mu(2\alpha)\gamma_{d+1}^2)\, S_{N,d}(\boldsymbol{q}^*),\end{aligned}$$

and

$$\begin{aligned}\theta(\boldsymbol{q}^*, q_{d+1}) = \sum_{(\boldsymbol{h}, h_{d+1})\in\mathbb{N}_0^{d+1}} \sum_{k_{d+1}=1}^{\infty} & \sum_{\substack{\boldsymbol{k}\in\mathbb{N}_0^d \\ \widetilde{\mathrm{tr}}_m(\boldsymbol{k})\cdot\boldsymbol{q}^*\equiv -\widetilde{\mathrm{tr}}_m(k_{d+1})\cdot q_{d+1} \pmod p}} r(\alpha, \boldsymbol{\gamma}, \boldsymbol{h}) r(\alpha, \gamma_{d+1}, h_{d+1}) \\ & \times r(\alpha, \gamma_{d+1}, k_{d+1}\oplus h_{d+1}) r(\alpha, \boldsymbol{\gamma}, \boldsymbol{k}\oplus\boldsymbol{h}).\end{aligned}$$

We choose q_{d+1}^* to minimize $S_{N,d+1}(\boldsymbol{q}^*, q_{d+1})$. Then for any $\lambda \in (1/\alpha, 1]$ we have

$$\theta(\boldsymbol{q}^*, q_{d+1}^*) \;\le\; \left(\frac{1}{N-1} \sum_{q_{d+1}\in R_{b,m}} (\theta(\boldsymbol{q}^*, q_{d+1}))^\lambda \right)^{1/\lambda}.$$

After some very long and tedious calculations to estimate this average on the right hand side, with the aid of Jensen's inequality and the property $[r(\alpha, \gamma, h_j)]^\lambda = r(\alpha\lambda, \gamma^\lambda, h_j)$, we finally obtain

$$\theta(\boldsymbol{q}^*, q_{d+1}^*) \le \left(2\mu(\alpha\lambda)\gamma_{d+1}^\lambda + 4(\mu(\alpha\lambda))^2\gamma_{d+1}^{2\lambda}\right)^{1/\lambda} N^{-1/\lambda} \prod_{j=1}^d \left(1 + \mu(\alpha\lambda)\gamma_j^\lambda\right)^{2/\lambda}.$$

Hence it follows from the induction hypothesis that

$$\begin{aligned}S_{N,d+1}(\boldsymbol{q}^*, q_{d+1}^*) \;\le\; & \left(\left(1 + \mu(2\alpha)\gamma_{d+1}^2\right) + \delta \left(2\mu(\alpha\lambda)\gamma_{d+1}^\lambda + 4(\mu(\alpha\lambda))^2\gamma_{d+1}^{2\lambda}\right)^{1/\lambda} \right) \\ & \times \delta^{-1} N^{-1/\lambda} \prod_{j=1}^d \left(1 + (1+\delta^\lambda)\mu(\alpha\lambda)\gamma_j^\lambda\right)^{2/\lambda}.\end{aligned}$$

With some elementary inequalities we can show that the multiplying factor in the expression above is bounded by $(1 + (1+\delta^\lambda)\mu(\alpha\lambda)\gamma_{d+1}^\lambda)^{2/\lambda}$. This completes the proof. □

For $\varepsilon \in (0,1)$, we choose $M = 2\varepsilon^{-2}$ and $N = b^m$ with

$$m = \left\lceil \log_b \left(\widetilde{C}_{s,\lambda,\delta} M^2 \right)^{\lambda} \right\rceil.$$

We have

$$\begin{aligned} \widetilde{C}_{s,\lambda,\delta} &\le \frac{1}{\delta} \exp\left(\frac{2(1+\delta^\lambda)\mu(\alpha\lambda)}{\lambda} \sum_{j=1}^{s} \gamma_j^\lambda \right) \\ &= \delta^{-1} (s+1)^{2(1+\delta^\lambda)\mu(\alpha\lambda)\lambda^{-1} \sum_{j=1}^{s} \gamma_j^\lambda / \ln(s+1)}. \end{aligned}$$

Let $p^* = 2\max(1/\alpha, s_{\boldsymbol{\gamma}})$. When (10) holds we take $\lambda = p^*/2 = 1$ if $s_{\boldsymbol{\gamma}} = 1$, and $\lambda = p^*/2 + \delta$ if $s_{\boldsymbol{\gamma}} < 1$. In both cases we have $\sup_{s \ge 1} \widetilde{C}_{s,\lambda,\delta} < \infty$ and $N = \mathcal{O}(\varepsilon^{-p})$, with p equal to or arbitrarily close to $2p^*$ as δ goes to 0. When (11) holds but not (10), we have $s_{\boldsymbol{\gamma}} = 1$ and we take $\lambda = 1$. Then $N = \mathcal{O}(\varepsilon^{-4})$ and $\widetilde{C}_{s,\lambda,\delta} = \mathcal{O}(s^a)$, with a arbitrarily close to $2\mu(\alpha)\ell$ as δ goes to 0. We summarize the analysis in the following theorem.

Theorem 4. (cf. [11, Theorem 3]) *Let P_{PL}^* be a polynomial lattice constructed component-by-component by Algorithm 1. For $\varepsilon \in (0,1)$ set $M = 2\varepsilon^{-2}$. If (10) holds, then the approximation algorithm $A_{N,s,M}$ using P_{PL}^* achieves the worst case error bound $e_{N,s,M}^{\mathrm{wor-app}}(P_{\mathrm{PL}}^*) \le \varepsilon$ using $N = \mathcal{O}(\varepsilon^{-p})$ function values, with p equal to or arbitrarily close to*

$$2\, p^{\mathrm{wor-app}}(\Lambda^{\mathrm{all}}).$$

If (10) does not hold but (11) holds, then the error bound $e_{N,s,M}^{\mathrm{wor-app}}(P_{\mathrm{PL}}^*) \le \varepsilon$ is achieved using $N = \mathcal{O}(s^a \varepsilon^{-4})$ function values, with a arbitrarily close to $2\mu(\alpha)\ell$. The implied factors in the big $\mathcal{O}$-notations are independent of ε and s.

Observe that when (10) holds we have $p^{\mathrm{wor-app}}(\Lambda^{\mathrm{all}}) \le 2$. Therefore $2\, p^{\mathrm{wor-app}}(\Lambda^{\mathrm{all}}) \le p^{\mathrm{wor-app}}(\Lambda^{\mathrm{all}}) + 2$, and we have improved the result in Theorem 1 using a fully constructive argument.

Remark. (cf. Theorem 1) *The exponent of strong tractability in the class Λ^{std} satisfies*

$$p^{\mathrm{wor-app}}(\Lambda^{\mathrm{std}}) \in [p^{\mathrm{wor-app}}(\Lambda^{\mathrm{all}}), 2\, p^{\mathrm{wor-app}}(\Lambda^{\mathrm{all}})].$$

Acknowledgments

The second author is supported by the Austrian Science Foundation (FWF), Project S9609, which is part of the Austrian National Research Network "Analytic Combinatorics and Probabilistic Number Theory". The third author is supported by a University of New South Wales Vice-chancellor's Postdoctoral Research Fellowship.

References

1. Aronszajn, N., Theory of reproducing kernels, Trans. Amer. Math. Soc. **68**, 337–404, 1950.
2. Chrestenson, H.E., A class of generalized Walsh functions. Pacific J. Math. **5**, 17–31, 1955.
3. Cristea, L.L., Dick, J., and Pillichshammer, F., On the mean square weighted $\mathcal{L}_2$ discrepancy of randomized digital nets in prime base. J. Complexity **22**, 605–629, 2006.
4. Dick, J., Kritzer, P., Kuo, F.Y., and Sloan, I.H., Lattice-Nyström method for Fredholm integral equation of the second kind. J. Complexity, submitted.
5. Dick, J., Kuo, F.Y., Pillichshammer, F., and Sloan, I.H., Construction algorithms for polynomial lattice rules for multivariate integration. Math. Comp. **74**, no. 252, 1895–1921, 2005.
6. Dick, J. and Pillichshammer, F., Dyadic Diaphony of Digital Nets over $\mathbb{Z}_2$. Monatsh. Math. **145**, 285–299, 2005.
7. Dick, J. and Pillichshammer, F., Multivariate integration in weighted Hilbert spaces based on Walsh functions and weighted Sobolev spaces. J. Complexity **21**, 149–195, 2005.
8. Dick, J. and Pillichshammer, F., On the mean square weighted L_2 discrepancy of randomized digital (t, m, s)-nets over $\mathbb{Z}_2$. Acta Arith. **117.4**, 371–403, 2005.
9. Hickernell, F.J. and Niederreiter, H., The existence of good extensible rank-1 lattices. J. Complexity **19**, 286–300, 2003.
10. Korobov, N.M., Number-theoretic Methods in Approximate Analysis, Fizmatgiz, Moscow, 1963.
11. Kuo, F.Y., Sloan, I.H., and Woźniakowski, H., Lattice Rules for Multivariate Approximation in the Worst Case Setting. In: Niederreiter, H. and Talay, D. (eds.), *Monte Carlo and Quasi-Monte Carlo Methods 2004*, pp. 289–330, Springer, Berlin, 2006.
12. Kuo, F.Y., Sloan, I.H., and Woźniakowski, H., Lattice Rule Algorithms for Multivariate Approximation in the Average Case Setting. J. Complexity, to appear.
13. Larcher, G., Digital point sets: analysis and application. In: Hellekalek, P. and Larcher, G. (eds.) *Random and Quasi-Random Point Sets*, pp. 167–222, Springer Lecture Notes in Statistics 138, 1998.
14. Larcher, G., Niederreiter, H., and Schmid, W.Ch., Digital nets and sequences constructed over finite rings and their application to quasi-Monte Carlo integration. Monatsh. Math. **121**, 231–253, 1996.
15. Li, D. and Hickernell, F.J., Trigonometric spectral collocation methods on lattices, In: Cheng, S.Y., Shu, C.-W., and Tang, T. (eds.), *Recent Advances in Scientific Computing and Partial Differential Equations*, AMS Series in Contemporary Mathematics, vol. 330, American Mathematical Society, Providence, Rhode Island, 121-132, 2003.
16. Niederreiter, H., Point Sets and Sequences with Small Discrepancy. Monatsh. Math. **104**, 273–337, 1987.
17. Niederreiter, H., Low-discrepancy and low-dispersion sequences. J. Number Theory **30**, 51–70, 1988.
18. Niederreiter, H., Low-discrepancy point sets obtained by digital constructions over finite fields. Czechoslovak Math. J. **42**, 143–166, 1992.
19. Niederreiter, H., *Random Number Generation and Quasi-Monte Carlo Methods*. CBMS-NSF Series in Applied Mathematics, **63**, SIAM, Philadelphia, 1992.

20. Niederreiter, H., Constructions of (t, m, s)-nets and (t, s)-sequences. Finite Fields Appl. **11**, 578–600, 2005.
21. Novak, E., Sloan, I. H., and Woźniakowski, H., Tractability of approximation for weighted Korobov spaces on classical and quantum computers, Found. Comput. Math. **4**, 121–156, 2004.
22. Pirsic, G., Schmid, W.Ch., Calculation of the quality parameter of digital nets and application to their construction. J. Complexity **17**: 827–839, 2001.
23. Sloan, I.H. and Woźniakowski, H., When are quasi-Monte Carlo algorithms efficient for high dimensional integrals?, J. Complexity, **14**, 1–33, 1998.
24. Sloan, I.H. and Woźniakowski, H., Tractability of multivariate integration for weighted Korobov classes, J. Complexity, **17**, 697–721, 2001.
25. Traub, J.F., Wasilkowski, G.W. and Woźniakowski, H., Information-Based Compelxity, Academic Press, New York, 1988.
26. Wang, X., Strong tractability of multivariate integration using quasi-Monte Carlo algorithms. Math. Comp. **72**, 823–838, 2002.
27. Walsh, J.L., A closed set of normal orthogonal functions. Amer. J. Math. **45**, 5–24, 1923.
28. Wasilkowski, G.W. and Woźniakowski, H., Weighted tensor product algorithms for linear multivariate problems, J. Complexity **15**, 402–447, 1999.
29. Wasilkowski, G.W. and Woźniakowski, H., On the power of standard information for weighted approximation. Found. Comput. Math. **1**, 417–434, 2001.
30. Zeng, X.Y., Leung, K.T., and Hickernell, F.J., Error analysis of splines for periodic problems using lattice designs, In: Niederreiter, H. and Talay, D. (eds.) *Monte Carlo and Quasi-Monte Carlo Methods 2004*, pp. 501–514, Springer, Berlin, 2006.