Discrepancy theory and quasi-Monte Carlo integration

Josef Dick and Friedrich Pillichshammer

[1] Josef Dick School of Mathematics and Statistics, The University of New South Wales, Sydney, NSW 2052, Australia, josef.dick@unsw.edu.au
[2] Friedrich Pillichshammer Institute for Financial Mathematics, University of Linz, Altenbergerstraße 69, 4040 Linz, Austria, friedrich.pillichshammer@jku.at

Summary. * In this article we show the deep connections between discrepancy theory on the one hand and quasi-Monte Carlo integration on the other. Discrepancy theory was established as an area of research going back to the seminal paper by Weyl (1916), whereas Monte Carlo (and later quasi-Monte Carlo) was invented in the 1940s by John von Neumann and Stanislaw Ulam to solve practical problems. The connection between these areas is well understood and will be presented here. We further include state of the art methods for quasi-Monte Carlo integration.

Summary. In this article we show the deep connections between discrepancy theory on the one hand and quasi-Monte Carlo integration on the other. Discrepancy theory was established as an area of research going back to the seminal paper by Weyl (1916), whereas Monte Carlo (and later quasi-Monte Carlo) was invented in the 1940s by John von Neumann and Stanislaw Ulam to solve practical problems. The connection between these areas is well understood and will be presented here. We further include state of the art methods for quasi-Monte Carlo integration.

1 Introduction - the connection between discrepancy theory and quasi-Monte Carlo integration

Let us start with introducing the concepts of discrepancy and quasi-Monte Carlo (QMC) for the domain $[0,1]^s$ and for a point set $\mathscr{P} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$. To define discrepancy, we define a set of 'test sets' $\mathcal{B}$. For instance, a common choice is the set of all intervals anchored at $\boldsymbol{0} = (0, \ldots, 0)$, denoted by $[\boldsymbol{0}, \boldsymbol{t}) = \prod_{i=1}^s [0, t_i)$, where $\boldsymbol{t} = (t_1, \ldots, t_s)$. The *local discrepancy* then is

$$\Delta_{\mathscr{P}}(\boldsymbol{t}) = \frac{1}{N} \sum_{n=1}^{N} 1_{[\boldsymbol{0},\boldsymbol{t})}(\boldsymbol{x}_n) - \prod_{i=1}^{s} t_i,$$

where $1_{[\boldsymbol{0},\boldsymbol{t})}$ denotes the *characteristic function* of the interval $[\boldsymbol{0}, \boldsymbol{t})$, i.e. $1_{[\boldsymbol{0},\boldsymbol{t})}(\boldsymbol{x})$ is one if $\boldsymbol{x}$ belongs to $[\boldsymbol{0}, \boldsymbol{t})$ and zero otherwise; see Fig. 1.

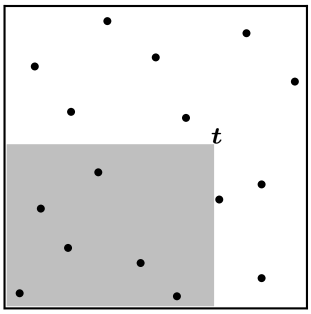

Fig. 1: The local discrepancy $\Delta_{\mathscr{P}}(\boldsymbol{t})$ measures the difference between the relative number of points that belong to the interval $[\boldsymbol{0}, \boldsymbol{t})$ and its volume.

The L_p-*discrepancy* of $\mathscr{P}$ is then the L_p norm $1 \le p \le \infty$ of $\Delta_{\mathscr{P}}$ given by

$$L_p(\mathscr{P}) = \|\Delta_{\mathscr{P}}\|_{L_p} = \left(\int_{[0,1]^s} |\Delta_{\mathscr{P}}(\boldsymbol{t})|^p \,\mathrm{d}\boldsymbol{t}\right)^{1/p}.$$

The L_∞-norm of the discrepancy function is also called the *star-discrepancy*, which is denoted by $D_N^*(\mathscr{P})$, i.e., $D_N^*(\mathscr{P}) = \|\Delta_{\mathscr{P}}\|_{L_\infty} = \sup |\Delta_{\mathscr{P}}(\boldsymbol{t})|$, where the supremum is extended over all $\boldsymbol{t} \in [0,1]^s$.

A *quasi-Monte Carlo rule* based on a point set $\mathscr{P} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ is an equal weight quadrature rule

$$Q_{\mathscr{P}}(f) := \frac{1}{N} \sum_{n=1}^{N} f(\boldsymbol{x}_n),$$

which can be used to approximate the integral $\int_{[0,1]^s} f(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}$. It is assumed that the quadrature points are chosen in some deterministic way which yields a small integration error for certain function classes.

To illustrate the connection between discrepancy and integration error we first consider discrepancy and numerical integration on the unit interval $[0,1]$.

1.1 An elementary approach

Central to showing the connection between the discrepancy of a point set and the integration error is the characteristic function of an interval. For numbers $x \in \mathbb{R}$ the characteristic function of an interval I is given by

$$1_I(x) = \begin{cases} 1 \text{ if } x \in I, \\ 0 \text{ otherwise.} \end{cases}$$

To give a glimpse of this connection between discrepancy and integration error, note the following two properties of the characteristic function:

1. Let $\mathscr{P} = \{x_1, \ldots, x_N\} \subseteq [0,1]$ be a point set. Then
$$\Delta_{\mathscr{P}}(t) := \frac{1}{N}\sum_{n=1}^{N} 1_{[0,t)}(x_n) - t$$
measures the *discrepancy* between the proportion of the points in the interval $[0,t)$ and the length of the interval.
2. Let $f : [0,1] \to \mathbb{R}$ be continuously differentiable. Then
$$f(x) = f(1) - \int_0^1 f'(t) 1_{(x,1]}(t)\,\mathrm{d}t = f(1) - \int_0^1 f'(t) 1_{[0,t)}(x)\,\mathrm{d}t. \tag{1}$$

These two properties can now be connected naturally in the following way. Let $f : [0,1] \to \mathbb{R}$ be continuously differentiable. Consider the *integration error* of f using a quasi-Monte Carlo rule $Q_{\mathscr{P}}(f) = \frac{1}{N}\sum_{n=1}^{N} f(x_n)$, given by
$$e(f;\mathscr{P}) = \int_0^1 f(x)\,\mathrm{d}x - \frac{1}{N}\sum_{n=1}^{N} f(x_n).$$
Then, using (1), we obtain
$$\begin{aligned}
&e(f;\mathscr{P}) \\
&= \int_0^1 \left[f(1) - \int_0^1 f'(t)1_{[0,t)}(x)\,\mathrm{d}t\right]\mathrm{d}x - \frac{1}{N}\sum_{n=1}^{N}\left[f(1) - \int_0^1 f'(t)1_{[0,t)}(x_n)\,\mathrm{d}t\right] \\
&= f(1) - \int_0^1\int_0^1 f'(t)1_{[0,t)}(x)\,\mathrm{d}t\,\mathrm{d}x - f(1) + \frac{1}{N}\sum_{n=1}^{N}\int_0^1 f'(t)1_{[0,t)}(x_n)\,\mathrm{d}t \\
&= \int_0^1 f'(t)\left[\frac{1}{N}\sum_{n=1}^{N} 1_{[0,t)}(x_n) - \int_0^1 1_{[0,t)}(x)\,\mathrm{d}x\right]\mathrm{d}t \\
&= \int_0^1 f'(t)\Delta_{\mathscr{P}}(t)\,\mathrm{d}t.
\end{aligned}$$
By using Hölder's inequality we obtain
$$|e(f;\mathscr{P})| \le L_p(\mathscr{P})\|f'\|_{L_q}, \tag{2}$$
where $1 \le p,q \le \infty$, $1/p + 1/q = 1$ and $\|g\|_{L_q} = \left(\int_0^1 |g(t)|^q\,\mathrm{d}t\right)^{1/q}$ with the obvious modification for $q = \infty$. For $p = \infty$, inequality (2) is a simplified version of Koksma's inequality (see Kuipers and Niederreiter [Chapter 2, Theorem 5.1]kuinie).

Some remarks regarding the last inequality are in order. We have obtained an upper bound on the integration error which is a product of two factors,

- one of which, $\|f'\|_{L_q}$ depends only on the integrand f; it is a semi-norm of f, and
- one of which, the L_p-discrepancy $L_p(\mathscr{P})$ of $\mathscr{P}$, depends only on the point set $\mathscr{P}$.

Thus (2) shows that quadrature points with small

L_p-discrepancy will yield a small integration error for functions with finite semi-norm $\|f'\|_{L_q}$.

Notice that there is a simple reason why a semi-norm rather than a norm is sufficient in (2): for any constant $c \in \mathbb{R}$ we have $e(f+c;\mathscr{P}) = e(f;\mathscr{P})$, i.e., constant functions are integrated exactly by the quasi-Monte Carlo rule $Q_{\mathscr{P}}$.

1.2 A reproducing kernel approach

Until now we conveniently assumed that f is continuously differentiable. However, there is a practical framework called reproducing kernel Hilbert spaces by Aronszajn [3], which defines a whole class of functions. On the domain $[0,1]$, a *reproducing kernel* is a function $K:[0,1]\times[0,1]\to\mathbb{R}$ which is

- symmetric: $K(x,y)=K(y,x)$ for all $x,y\in[0,1]$, and
- positive definite, that is,

$$\sum_{k,l=1}^{N} a_k \overline{a}_l K(x_k,x_l)\ge 0$$

  for all $a_1,\dots,a_N\in\mathbb{C}$ and $x_1,\dots,x_N\in[0,1]$. (Here, $\overline{a}_l$ denotes the conjugate complex of a_l.)

A reproducing kernel can naturally be defined using the characteristic function $1_{[0,x)}$ by setting

$$K(x,y)=1+\int_0^1 1_{(x,1]}(t)1_{(y,1]}(t)\,\mathrm{d}t=1+\min(1-x,1-y).$$

The function K such defined is symmetric and positive definite, and thus a reproducing kernel. Associated with this reproducing kernel is a set $\mathcal{H}(K)$ of functions $f:[0,1]\to\mathbb{R}$ and an inner product $\langle\cdot,\cdot\rangle_{\mathcal{H}(K)}$ on $\mathcal{H}(K)$ such that

- $K(\cdot,y)\in\mathcal{H}(K)$ for all $y\in[0,1]$, and
- $f(y)=\langle f,K(\cdot,y)\rangle_{\mathcal{H}(K)}$ for all $y\in[0,1]$ and $f\in\mathcal{H}(K)$.

From Aronszajn [3] it is known that the function space $\mathcal{H}(K)$ is a Hilbert space with an inner product which is uniquely defined. For functions f,g, which can be represented in the form (1), the inner product is given by

$$\langle f,g\rangle_{\mathcal{H}(K)}=f(1)g(1)+\int_0^1 f'(x)g'(x)\,\mathrm{d}x.$$

These functions f and g are absolutely continuous and $f',g'\in L_2([0,1])$, the space of square integrable functions defined on $[0,1]$.

Let $y\in[0,1]$ be fixed. Then $k(x):=K(x,y)$ has the representation

$$k(x)=1-\int_0^1 1_{(x,1]}(t)[-1_{(y,1]}(t)]\,\mathrm{d}t.$$

Thus, by matching it with the pattern from (1), $k(1)=1$ and $k'(x)=-1_{[y,1]}(x)$. Thus

$$\langle f,K(\cdot,y)\rangle_{\mathcal{H}(K)}=f(1)1-\int_0^1 f'(x)1_{[y,1]}(x)\,\mathrm{d}x=f(y).$$

Then the integration error of f using a quasi-Monte Carlo rule based on $\mathscr{P}=\{x_1,\dots,x_n\}$ is given by

$$e(f;\mathscr{P})=\int_0^1 f(x)\,\mathrm{d}x-\frac{1}{N}\sum_{n=1}^{N} f(x_n)$$

$$= \int_0^1 \langle f, K(\cdot, x)\rangle_{\mathcal{H}} \,\mathrm{d}x - \frac{1}{N}\sum_{n=1}^{N} \langle f, K(\cdot, x_n)\rangle_{\mathcal{H}(K)}$$
$$= \left\langle f, \int_0^1 K(\cdot, x)\,\mathrm{d}x - \frac{1}{N}\sum_{n=1}^{N} K(\cdot, x_n)\right\rangle_{\mathcal{H}(K)}.$$

We have

$$\begin{aligned} h(z) &:= \int_0^1 K(z,x)\,\mathrm{d}x - \frac{1}{N}\sum_{n=1}^{N} K(z,x_n) \\ &= -\int_0^1 1_{(z,1]}(t)\left[\frac{1}{N}\sum_{n=1}^{N} 1_{(x_n,1]}(t) - \int_0^1 1_{(x,1]}(t)\,\mathrm{d}x\right]\mathrm{d}t \\ &= -\int_0^1 1_{(z,1]}(t)\left[\frac{1}{N}\sum_{n=1}^{N} 1_{[0,t)}(x_n) - t\right]\mathrm{d}t \\ &= -\int_0^1 1_{(z,1]}(t)\Delta_{\mathscr{P}}(t)\,\mathrm{d}t. \end{aligned}$$

Thus, matching the representation of h given above with the pattern from (1), we obtain $h(1) = 0$ and $h'(x) = \Delta_{\mathscr{P}}(x)$. Hence we have

$$e(f;\mathscr{P}) = \langle f, h\rangle_{\mathcal{H}(K)} = f(1)0 + \int_0^1 f'(x)\Delta_{\mathscr{P}}(x)\,\mathrm{d}x = \int_0^1 f'(x)\Delta_{\mathscr{P}}(x)\,\mathrm{d}x.$$

Thus, taking the absolute value and using Hölder's inequality we again obtain (2).

So far we have considered the integration error for a particular function f. Since we have now a function space $\mathcal{H}(K)$ with inner product $\langle\cdot,\cdot\rangle_{\mathcal{H}(K)}$, we can define the corresponding norm by $\|\cdot\|_{\mathcal{H}(K)} = \langle\cdot,\cdot\rangle_{\mathcal{H}(K)}^{1/2}$. Then it is meaningful to define the worst-case error in the unit ball of $\mathcal{H}(K)$ by

$$e(\mathcal{H}(K);\mathscr{P}) = \sup_{f\in\mathcal{H}(K), \|f\|_{\mathcal{H}(K)}\le 1} |e(f;\mathscr{P})|.$$

Since $e(f;\mathscr{P}) = \langle f, h\rangle_{\mathcal{H}(K)}$ we obtain

$$\begin{aligned} e(\mathcal{H}(K);\mathscr{P}) &= \sup_{f\in\mathcal{H}(K), \|f\|_{\mathcal{H}(K)}\le 1} |\langle f, h\rangle_{\mathcal{H}(K)}| \\ &\le \sup_{f\in\mathcal{H}(K), \|f\|_{\mathcal{H}(K)}\le 1} \|f\|_{\mathcal{H}(K)}\|h\|_{\mathcal{H}(K)} = \|h\|_{\mathcal{H}(K)}. \end{aligned}$$

On the other hand, we have $h \in \mathcal{H}(K)$ and by choosing $f = h/\|h\|_{\mathcal{H}(K)}$ we obtain that

$$e(\mathcal{H}(K);\mathscr{P}) = \|h\|_{\mathcal{H}(K)}.$$

This yields the formula

$$\begin{aligned} &e^2(\mathcal{H}(K);\mathscr{P}) = \langle h, h\rangle_{\mathcal{H}(K)} \qquad (3) \\ &= \int_0^1\int_0^1 K(x,y)\,\mathrm{d}x\,\mathrm{d}y - \frac{2}{N}\sum_{n=1}^{N}\int_0^1 K(x,x_n)\,\mathrm{d}x + \frac{1}{N^2}\sum_{n,m=1}^{N} K(x_n,x_m). \end{aligned}$$

As $h(1) = 0$ and $h' = \Delta_{\mathscr{P}}$ we have $\|h\|_{\mathcal{H}(K)} = \|\Delta_{\mathscr{P}}\|_{L_2} = L_2(\mathscr{P})$. Thus

$$e(\mathcal{H}(K); \mathscr{P}) = L_2(\mathscr{P})$$

and so (3) yields an explicit expression for the L_2-discrepancy (which is the one-dimensional version of a formula that is sometimes attributed to Warnock; see Matoušek [72, Lemma 2.14])

$$(L_2(\mathscr{P}))^2 = \frac{4}{3} - \frac{2}{N}\sum_{n=1}^{N} \frac{3 - x_n^2}{2} + \frac{1}{N^2}\sum_{n,m=1}^{N} [1 + \min(1 - x_n, 1 - x_m)].$$

Since the reproducing kernel function K has a closed form, the worst-case error can be computed for given point sets $\mathscr{P}$.

In the following we also consider another, related, reproducing kernel, namely

$$K(x, y) = \min(1 - x, 1 - y).$$

The corresponding reproducing kernel Hilbert space consists of the same functions f as in the reproducing kernel Hilbert space as above with the restriction that $f(1) = 0$. The corresponding inner product is then simply $\int_0^1 f'(x)g'(x)\,\mathrm{d}x$.

1.3 Discrepancy and numerical integration in arbitrary dimension

The step from $[0, 1]$ to $[0, 1]^s$ for some $s \geq 1$ is achieved by considering tensor product function spaces. Let now $\mathscr{P} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subseteq [0, 1]^s$. The reproducing kernel for functions on $[0, 1]^s$ is simply given by

$$K(\boldsymbol{x}, \boldsymbol{y}) = \int_{[0,1]^s} 1_{(\boldsymbol{x},\boldsymbol{1}]}(\boldsymbol{t}) 1_{(\boldsymbol{y},\boldsymbol{1}]}(\boldsymbol{t})\,\mathrm{d}t = \prod_{i=1}^{s} \min(1 - x_i, 1 - y_i), \tag{4}$$

where $\boldsymbol{x} = (x_1, \ldots, x_s)$, $\boldsymbol{y} = (y_1, \ldots y_s)$, and $(\boldsymbol{x}, \boldsymbol{1}] = \prod_{i=1}^{s}(x_i, 1]$. The corresponding Hilbert space $\mathcal{H}(K)$ is the s-fold tensor product of the one dimensional reproducing kernel Hilbert spaces with reproducing kernel $K(x, y) = \min(1-x, 1-y)$. In particular, if $f \in \mathcal{H}(K)$, then $\|\partial^s f/\partial \boldsymbol{x}\|_{L_2} < \infty$ and $\frac{\partial^{|\mathfrak{u}|} f}{\partial \boldsymbol{x}_{\mathfrak{u}}}(\boldsymbol{z}_{\mathfrak{u}}, \boldsymbol{1}) = 0$ for $u \subseteq \{1, \ldots, s\}$, where $\partial \boldsymbol{x}_{\mathfrak{u}} = \prod_{i \in \mathfrak{u}} \partial x_i$ and where $(\boldsymbol{z}_{\mathfrak{u}}, \boldsymbol{1})$ stands for the vector whose ith component is z_i if $i \in \mathfrak{u}$ and 1 otherwise. Further, for $\mathfrak{u} = \emptyset$ we have $\frac{\partial^{|\mathfrak{u}|} f}{\partial \boldsymbol{x}_{\mathfrak{u}}}(\boldsymbol{z}_{\mathfrak{u}}, \boldsymbol{1}) := f(\boldsymbol{1})$. The inner product is given by

$$\langle f, g \rangle_{\mathcal{H}(K)} = \int_{[0,1]^s} \frac{\partial^s f}{\partial \boldsymbol{x}}(\boldsymbol{t}) \frac{\partial^s g}{\partial \boldsymbol{x}}(\boldsymbol{t})\,\mathrm{d}\boldsymbol{t}.$$

The same steps as in the previous two subsections can be carried out to obtain the discrepancy function

$$\Delta_{\mathscr{P}}(\boldsymbol{t}) = \frac{1}{N}\sum_{n=1}^{N} 1_{[\boldsymbol{0},\boldsymbol{t})}(\boldsymbol{x}_n) - \prod_{i=1}^{s} t_i.$$

Again, an analogue of (2) holds, namely for the integration error of a function $f \in \mathcal{H}(K)$ using a quasi-Monte Carlo rule based on $\mathscr{P}$ we have

$$|e(f;\mathscr{P})| \le L_p(\mathscr{P}) \|\partial^s f/\partial \boldsymbol{x}\|_{L_q}.$$

Again, for $p = \infty$, this is a simplified version of the Koksma-Hlawka inequality (see Kuipers and Niederreiter [63, Chapter 2, Theorem 5.5]).

The worst-case error is again given by

$$e(\mathcal{H}(K);\mathscr{P}) = \sup_{f \in \mathcal{H}(K), \|f\|_{\mathcal{H}(K)} \le 1} |e(f;\mathscr{P})| = \|h\|_{\mathcal{H}(K)},$$

where $h(\boldsymbol{z}) = \int_{[0,1]^s} K(\boldsymbol{z},\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} - \frac{1}{N}\sum_{n=1}^N K(\boldsymbol{z},\boldsymbol{x}_n)$. Again we have $\|h\|_{\mathcal{H}(K)} = \|\Delta_{\mathscr{P}}\|_{L_2} = L_2(\mathscr{P})$, therefore we obtain

$$e(\mathcal{H}(K);\mathscr{P}) = L_2(\mathscr{P}).$$

The analogue of (3) yields

$$\begin{aligned} e^2(\mathcal{H}(K);\mathscr{P}) &= \langle h,h\rangle_{\mathcal{H}(K)} \\ &= \int_{[0,1]^s}\int_{[0,1]^s} K(\boldsymbol{x},\boldsymbol{y}) \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y} - \frac{2}{N}\sum_{n=1}^N \int_{[0,1]^s} K(\boldsymbol{x},\boldsymbol{x}_n)\,\mathrm{d}\boldsymbol{x} \\ &\quad + \frac{1}{N^2}\sum_{n,m=1}^N K(\boldsymbol{x}_n,\boldsymbol{x}_m). \end{aligned}$$

Since there is an explicit expression for the reproducing kernel (4) we obtain an explicit expression for the L_2-discrepancy of $\mathscr{P} = \{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N\}$, sometimes called Warnock's formula

$$(L_2(\mathscr{P}))^2 = \frac{1}{3^s} - \frac{2}{N}\sum_{n=1}^N \prod_{i=1}^s \frac{1-x_{n,i}^2}{2} + \frac{1}{N^2}\sum_{n,m=1}^N \prod_{i=1}^s \min(1-x_{n,i}, 1-x_{m,i}), \quad (5)$$

where $x_{n,i}$ denotes the ith component of the point $\boldsymbol{x}_n$.

The discrepancy defined this way does not take lower order projections into account. To include also lower dimensional projections we use the reproducing kernel

$$\begin{aligned} K(\boldsymbol{x},\boldsymbol{y}) &= \int_{[0,1]^s} \prod_{i=1}^s \left[1 + 1_{(x_i,1]}(t_i) 1_{(y_i,1]}(t_i)\right] \,\mathrm{d}\boldsymbol{t} \\ &= \prod_{i=1}^s \left[1 + \min(1-x_i, 1-y_i)\right]. \end{aligned}$$

In this case the inner product is given by

$$\langle f,g\rangle_{\mathcal{H}(K)} = \sum_{\mathfrak{u}\subseteq[s]} \int_{[0,1]^{|\mathfrak{u}|}} \frac{\partial^{|\mathfrak{u}|} f}{\partial \boldsymbol{x}_{\mathfrak{u}}}(\boldsymbol{t}_{\mathfrak{u}},\boldsymbol{1}) \frac{\partial^{|\mathfrak{u}|} g}{\partial \boldsymbol{x}_{\mathfrak{u}}}(\boldsymbol{t}_{\mathfrak{u}},\boldsymbol{1}) \,\mathrm{d}\boldsymbol{t}_{\mathfrak{u}},$$

where $[s] = \{1,\ldots,s\}$ and where for $\mathfrak{u} \subseteq [s]$ and $\boldsymbol{x} = (x_1,\ldots,x_s)$ we write $\boldsymbol{x}_{\mathfrak{u}}$ for the $|\mathfrak{u}|$-dimensional projection of $\boldsymbol{x}$ onto the coordinates given by $\mathfrak{u}$ and where $(\boldsymbol{x}_{\mathfrak{u}},\boldsymbol{1})$ is the s-dimensional vector whose ith component is x_i if $i \in \mathfrak{u}$ and 1 otherwise. Further we have

$$e^2(\mathcal{H}(K);\mathscr{P}) = \sum_{\emptyset \ne \mathfrak{u}\subseteq[s]} (L_2(\mathscr{P}_{\mathfrak{u}}))^2$$

$$= \frac{4^s}{3^s} - \frac{2}{N}\sum_{n=1}^{N}\prod_{i=1}^{s}\frac{3-x_{n,i}^2}{2} + \frac{1}{N^2}\sum_{n,m=1}^{N}\prod_{i=1}^{s}[1+\min(1-x_{n,i},1-x_{m,i})],$$

where $\mathscr{P}_{\mathfrak{u}}$ stands for the projection of the points in $\mathscr{P}$ onto the coordinates in $\mathfrak{u}$ and $L_2(\mathscr{P}_{\mathfrak{u}})$ stands for the L_2 discrepancy of $\mathscr{P}_{\mathfrak{u}}$.

1.4 Integration in weighted function spaces

Sloan and Woźniakowski [107] (see also D., Sloan, Wang, Woźniakowski [34]) introduced a weighted discrepancy. The idea is that in many applications some projections are more important than others and that this should also be reflected in the quality measure of the point set.

The difference in the importance of projections is usually modelled by introducing so-called weights. Here we restrict ourselves to product-weights. Let $\boldsymbol{\gamma} = (\gamma_i)_{i\ge 1}$ be a sequence of weights in $\mathbb{R}^+$. We use then the reproducing kernel

$$K_{\boldsymbol{\gamma}}(\boldsymbol{x},\boldsymbol{y}) = \int_{[0,1]^s}\prod_{i=1}^{s}\left[1+\gamma_i 1_{(x_i,1]}(t_i)1_{(y_i,1]}(t_i)\right]\,\mathrm{d}\boldsymbol{t}$$
$$= \prod_{i=1}^{s}\left[1+\gamma_i\min(1-x_i,1-y_i)\right].$$

In this case the inner product is given by

$$\langle f,g\rangle_{\mathcal{H}(K_{\boldsymbol{\gamma}})} = \sum_{\mathfrak{u}\subseteq[s]}\boldsymbol{\gamma}_{\mathfrak{u}}^{-1}\int_{[0,1]^{|\mathfrak{u}|}}\frac{\partial^{|\mathfrak{u}|}f}{\partial\boldsymbol{x}_{\mathfrak{u}}}(\boldsymbol{t}_{\mathfrak{u}},\boldsymbol{1})\frac{\partial^{|\mathfrak{u}|}g}{\partial\boldsymbol{x}_{\mathfrak{u}}}(\boldsymbol{t}_{\mathfrak{u}},\boldsymbol{1})\,\mathrm{d}\boldsymbol{t}_{\mathfrak{u}},$$

where for $\mathfrak{u}\subseteq[s]$ we write $\boldsymbol{\gamma}_{\mathfrak{u}} = \prod_{i\in\mathfrak{u}}\gamma_i$ and for $\mathfrak{u}=\emptyset$ we have $\boldsymbol{\gamma}_{\emptyset}=1$ and $\frac{\partial^{|\mathfrak{u}|}f}{\partial\boldsymbol{x}_{\mathfrak{u}}}(\boldsymbol{t}_{\mathfrak{u}},\boldsymbol{1}) := f(\boldsymbol{1})$.

With

$$h(\boldsymbol{z}) = \int_{[0,1]^s}K_{\boldsymbol{\gamma}}(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y} - \frac{1}{N}\sum_{n=1}^{N}K_{\boldsymbol{\gamma}}(\boldsymbol{x},\boldsymbol{x}_n)$$
$$= \prod_{i=1}^{s}\left(1+\frac{\gamma_i}{2}(1-x_i^2)\right) - \frac{1}{N}\sum_{n=1}^{N}\prod_{i=1}^{s}[1+\gamma_i\min(1-z_i,1-x_{n,i})]$$

and

$$\frac{\partial^{|\mathfrak{u}|}h}{\partial\boldsymbol{z}_{\mathfrak{u}}}(\boldsymbol{t}_u,\boldsymbol{1}) = (-1)^{|\mathfrak{u}|+1}\boldsymbol{\gamma}_{\mathfrak{u}}\Delta_{\mathscr{P}}(\boldsymbol{t}_{\mathfrak{u}})$$

we obtain for the integration error of a function $f\in\mathcal{H}(K_{\boldsymbol{\gamma}})$,

$$e(f;\mathscr{P}) = \langle f,h\rangle_{\mathcal{H}(K_{\boldsymbol{\gamma}})}$$
$$= \sum_{\mathfrak{u}\subseteq[s]}\boldsymbol{\gamma}_{\mathfrak{u}}^{-1}(-1)^{|\mathfrak{u}|+1}\boldsymbol{\gamma}_{\mathfrak{u}}\int_{[0,1]^{|\mathfrak{u}|}}\frac{\partial^{|\mathfrak{u}|}f}{\partial\boldsymbol{x}_{\mathfrak{u}}}(\boldsymbol{t}_{\mathfrak{u}},\boldsymbol{1})\Delta_{\mathscr{P}}(\boldsymbol{t}_{\mathfrak{u}},\boldsymbol{1})\,\mathrm{d}\boldsymbol{t}_{\mathfrak{u}}.$$

The unweighted version of this formula is due to Hlawka [51] and Zaremba [114] and is called *Hlawka-Zaremba identity.* Applying Hölder's inequality for integrals and sums we obtain

$$|e(f;\mathscr{P})| \le \|f\|_{\mathcal{H}(K_{\boldsymbol{\gamma}}),q} L_{p,\boldsymbol{\gamma}}(\mathscr{P}),$$

where $1 \le p, q \le \infty$ and $1/p + 1/q = 1$,

$$\|f\|_{\mathcal{H}(K_{\boldsymbol{\gamma}}),q} = \left(\sum_{\mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}}^{-q} \int_{[0,1]^{|\mathfrak{u}|}} \left| \frac{\partial^{|\mathfrak{u}|} f}{\partial \boldsymbol{x}_{\mathfrak{u}}} (\boldsymbol{t}_{\mathfrak{u}}, \mathbf{1}) \right|^q \mathrm{d}\boldsymbol{t}_{\mathfrak{u}} \right)^{1/q}$$

and the so-called *weighted L_p-discrepancy* is given by

$$\begin{aligned} L_{p,\boldsymbol{\gamma}}(\mathscr{P}) &= \left(\sum_{\emptyset \ne \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}}^p (L_p(\mathscr{P}_{\mathfrak{u}}))^p \right)^{1/p} \\ &= \left(\sum_{\emptyset \ne \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}}^p \int_{[0,1]^{|\mathfrak{u}|}} |\Delta_{\mathscr{P}}((\boldsymbol{t}_{\mathfrak{u}}, \mathbf{1}))|^p \, \mathrm{d}\boldsymbol{t}_{\mathfrak{u}} \right)^{1/p}, \end{aligned}$$

where $\Delta_{\mathscr{P}}$ and L_p denote the usual local and L_p-discrepancy, respectively. In the case $p = \infty$ we also write

$$D_{N,\boldsymbol{\gamma}}^*(\mathscr{P}) = \max_{\emptyset \ne \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}} D_N^*(\mathscr{P}_{\mathfrak{u}}).$$

We call $D_{N,\boldsymbol{\gamma}}^*$ the *weighted star-discrepancy* of $\mathscr{P}$. Note that $D_{N,\mathbf{1}}^* = D_N^*$, where $\mathbf{1} = (1)_{i \ge 1}$, the sequence of weights where every weight is equal to one.

We also obtain a weighted discrepancy and Warnock-type formula, given by

$$\begin{aligned} e^2(\mathcal{H}(K_{\boldsymbol{\gamma}}); \mathscr{P}) = (L_{2,\boldsymbol{\gamma}}(\mathscr{P}))^2 &= \sum_{\emptyset \ne \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}}^2 (L_2(\mathscr{P}_{\mathfrak{u}}))^2 \\ &= \prod_{i=1}^s \left(1 + \frac{\gamma_i^2}{3}\right) - \frac{2}{N} \sum_{n=1}^N \prod_{i=1}^s \left(1 + \gamma_i^2 \frac{1 - x_{n,i}^2}{2}\right) \\ &\quad + \frac{1}{N^2} \sum_{n,m=1}^N \prod_{i=1}^s [1 + \gamma_i^2 \min(1 - x_{n,i}, 1 - x_{m,i})]. \end{aligned} \tag{6}$$

We remark that the assumption that $s < \infty$ can also be removed. In particular, Gnewuch [43] considered numerical integration in infinite dimensional reproducing kernel Hilbert spaces. Further, integration over $\mathbb{R}^s$ rather than $[0,1]^s$ has for instance been considered in [21].

1.5 Discrepancy and quasi-Monte Carlo on the sphere

The above approach can be generalised in various ways, see for instance Gnewuch [42, 85]. In the following we illustrate the above approach on a different domain, see Brauchart and D. [9]. Consider the sphere $\mathbb{S}^s = \{(x_1, \ldots, x_{s+1}) \in \mathbb{R}^{s+1} : x_1^2 + \cdots + x_{s+1}^2 = 1\}$. As test sets we use spherical caps

$$C(t, \boldsymbol{x}) = \{\boldsymbol{z} \in \mathbb{S}^s : \langle \boldsymbol{z}, \boldsymbol{x} \rangle \ge t\}, \quad \boldsymbol{x} \in \mathbb{S}^s, -1 \le t \le 1,$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product in $\mathbb{R}^{s+1}$. We use the same approach as above. Let now $\mathscr{P} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subseteq \mathbb{S}^s$.

We define a reproducing kernel

$$K(\boldsymbol{x},\boldsymbol{y}) = \int_{-1}^{1}\int_{\mathbb{S}^s} 1_{C(\boldsymbol{x},t)}(\boldsymbol{z}) 1_{C(\boldsymbol{y},t)}(\boldsymbol{z}) \,\mathrm{d}\mu(\boldsymbol{z})\,\mathrm{d}t,$$

where μ is the Lebesgue measure on the sphere $\mathbb{S}^s$ normalised to a probability measure.

The corresponding reproducing kernel Hilbert space $\mathcal{H}(K)$ then includes functions of the form

$$f(\boldsymbol{x}) = \int_{-1}^{1}\int_{\mathbb{S}^s} 1_{C(\boldsymbol{x},t)}(\boldsymbol{z}) f_0(\boldsymbol{z},t) \,\mathrm{d}\mu(\boldsymbol{z})\,\mathrm{d}t, \tag{7}$$

where $f_0 \in L_2(\mathbb{S}^s \times [-1,1])$, see Brauchart and D. [9].

Notice that in this case f_0 is not related to any classical derivative of f. We only assume that there exists a function $f_0 \in L_2(\mathbb{S}^s \times [-1,1])$ such that (7) holds. Notice further that for our purposes it is not necessary to be able to obtain f_0 from some given f (for the cube $[0,1]^s$ the function f_0 can be obtain via differentiation, but that fact was not used).

For functions $f,g : \mathbb{S}^s \to \mathbb{R}$ with representation of the form (7) we can define the inner product

$$\langle f,g\rangle_{\mathcal{H}(K)} = \int_{-1}^{1}\int_{\mathbb{S}^2} f_0(\boldsymbol{z},t) g_0(\boldsymbol{z},t)\,\mathrm{d}\mu(\boldsymbol{z})\,\mathrm{d}t.$$

Again, going through the same steps as above we obtain the discrepancy function

$$\Delta_{\mathscr{P}}(\boldsymbol{z},t) = \frac{1}{N}\sum_{n=1}^{N} 1_{C(\boldsymbol{z},t)}(\boldsymbol{x}_n) - \mu(C(\boldsymbol{z},t)).$$

Again, we obtain a Koksma-Hlawka type inequality of the form

$$\left| \int_{\mathbb{S}^s} f(\boldsymbol{x})\,\mathrm{d}\mu(\boldsymbol{x}) - \frac{1}{N}\sum_{n=1}^{N} f(\boldsymbol{x}_n)\right| \le \|\Delta_{\mathscr{P}}\|_{L_p} \|f_0\|_{L_q}.$$

We call $\|\Delta_{\mathscr{P}}\|_{L_p}$ the L_p spherical cap discrepancy

$$\|\Delta_{\mathscr{P}}\|_{L_p}^p = \int_{-1}^{1}\int_{\mathbb{S}^s} \left| \frac{1}{N}\sum_{n=1}^{N} 1_{C(\boldsymbol{z},t)}(\boldsymbol{x}_n) - \mu(C(\boldsymbol{z},t))\right|^p \,\mathrm{d}\mu(\boldsymbol{z})\,\mathrm{d}t.$$

Again, the L_2-discrepancy is related to the worst-case integration error

$$e(\mathcal{H}(K);\mathscr{P}) = \sup_{f\in\mathcal{H}(K), \|f\|_{\mathcal{H}(K)}\le 1} |e(f;\mathscr{P})| = \|\Delta_{\mathscr{P}}\|_{L_2}.$$

We also obtain

$$\begin{aligned} e^2(\mathcal{H}(K);\mathscr{P}) &= \|\Delta_{\mathscr{P}}\|_{L_2}^2 \\ &= \int_{\mathbb{S}^s}\int_{\mathbb{S}^s} K(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\mu(\boldsymbol{x})\,\mathrm{d}\mu(\boldsymbol{y}) - \frac{2}{N}\sum_{n=1}^{N}\int_{\mathbb{S}^s} K(\boldsymbol{x},\boldsymbol{x}_n)\,\mathrm{d}\mu(\boldsymbol{x}) \\ &\quad + \frac{1}{N^2}\sum_{n,m=1}^{N} K(\boldsymbol{x}_n,\boldsymbol{x}_m) \end{aligned} \tag{8}$$

$$= \frac{1}{N^2} \sum_{n,m=1}^{N} K(\boldsymbol{x}_n, \boldsymbol{x}_m) - \int_{\mathbb{S}^s} \int_{\mathbb{S}^s} K(\boldsymbol{x}, \boldsymbol{y}) \,\mathrm{d}\mu(\boldsymbol{x}) \,\mathrm{d}\mu(\boldsymbol{y}). \qquad (9)$$

The reproducing kernel for the sphere $\mathbb{S}^s$ even has a concise form, see Brauchart and D. [9], given by

$$K(\boldsymbol{x}, \boldsymbol{y}) = 1 - \frac{\Gamma(\frac{s+1}{2})}{s\sqrt{\pi}\Gamma(\frac{s}{2})} \|\boldsymbol{x} - \boldsymbol{y}\|, \qquad (10)$$

where $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^{s+1}$ and $\Gamma > 0$ is the Gamma function. Thus, using (8) and (10) we also obtain a Warnock-type formula

$$\|\Delta_{\mathscr{P}}\|_{L_2}^2 = \frac{\Gamma(\frac{s+1}{2})}{s\sqrt{\pi}\Gamma(\frac{s}{2})} \left[\int_{\mathbb{S}^s} \int_{\mathbb{S}^s} \|\boldsymbol{x} - \boldsymbol{y}\| \,\mathrm{d}\mu(\boldsymbol{x}) \,\mathrm{d}\mu(\boldsymbol{y}) - \frac{1}{N^2} \sum_{n,m=1}^{N} \|\boldsymbol{x}_n - \boldsymbol{x}_m\| \right].$$

This equality is known as *Stolarsky's invariance principle*, see Stolarsky [110]. The value of the distance integral is known explicityl and is given by

$$\int_{\mathbb{S}^s} \int_{\mathbb{S}^s} \|\boldsymbol{x} - \boldsymbol{y}\| \,\mathrm{d}\mu(\boldsymbol{x}) \,\mathrm{d}\mu(\boldsymbol{y}) = 2^s \frac{\Gamma((s+1)/2)\Gamma((s+1)/2)}{\sqrt{\pi}\Gamma(s-1/2)},$$

where Γ is the Gamma function.

It would be interesting to find generalisations of the geometric discrepancy defined above for other domains (manifolds) where the reproducing kernel also has a concise form to obtain analogues of Stolarsky's invariance principle for other domains.

2 Bounds on the discrepancy

In this section we discuss some bounds on the L_p-discrepancy. For $s, N \in \mathbb{N}$ and $1 \le p \le \infty$ let

$$\mathrm{disc}_p(N, s) = \inf_{\substack{\mathscr{P} \subseteq [0,1)^s \\ |\mathscr{P}| = N}} L_p(\mathscr{P})$$

denote the minimal L_p-discrepancy that can be achieved by point sets consisting of N points in $[0,1)^s$. Note that for any $1 \le p_1 \le p_2 \le \infty$ we have

$$\mathrm{disc}_{p_1}(N, s) \le \mathrm{disc}_{p_2}(N, s).$$

2.1 Asymptotic bounds

In the case $p = \infty$ it is known that for any fixed $s \in \mathbb{N}$ there exist constants $0 < c_s \le C_s$ such that

$$c_s \frac{(\log N)^{\kappa_s}}{N} \le \mathrm{disc}_\infty(N, s) \le C_s \frac{(\log N)^{s-1}}{N}, \qquad (11)$$

where $\kappa_2 = 1$ (see Bejian [7] and Schmidt [98]) and $\kappa_s \ge (s-1)/2$ for $s \ge 3$, which follows from a result of Roth [94]. For $s \ge 3$ the lower bound on κ_s has recently been improved to $\kappa_s \ge (s-1)/2 + \delta_s$ for some unknown $0 < \delta_s < 1/2$; see Bilyk, Lacey and Vagharshakyan [8]. The upper bound can even be achieved constructively.

A point set $\mathscr{P}$ whose star-discrepancy satisfies an upper bound of the form $D_N^*(\mathscr{P}) = O((\log N)^{\alpha_s}/N)$ as $N \to \infty$, where $\alpha_s \geq 0$, is sometimes called a *low discrepancy point set*. There are several methods to construct low discrepancy point sets. Examples of such point sets include:

- Hammersley point sets which are based on the infinite van der Corput sequence (see, e.g., [33] and Niederreiter [78]) achieving $\alpha_s = s - 1$.
- Lattice point sets (or, more general, integration lattices) which were introduced independently by Korobov [57] and Hlawka [52] and which are well explained in the books of Niederreiter [78] and of Sloan and Joe [103]. Here it is known that one can achieve $\alpha_2 = 1$ and $\alpha_s = s$ for $s \geq 3$. Lattice point sets will be discussed in Section 4.
- (t, m, s)-nets in base b which were introduced by Niederreiter [76, 78] and which are the main topic of the recent book [33]. Precursors of such nets go back to constructions of Sobol' [109] and Faure [39]. With nets one can achieve $\alpha_s = s-1$ for all $s \geq 1$. (t, m, s)-nets will be discussed in Section 3.1.

For $1 < p < \infty$ and for any fixed $s \in \mathbb{N}$ it is known that

$$\mathrm{disc}_p(N, s) \asymp_{s,p} \frac{(\log N)^{(s-1)/2}}{N} \quad \text{as} \quad N \to \infty, \tag{12}$$

where $A \asymp_{s,p} B$ means that there are constants $c_{s,p}, C_{s,p} > 0$ depending only on s, p such that $c_{s,p} B \leq A \leq C_{s,p} B$. Here the lower bound is due to Roth [94] for $p \geq 2$ and Schmidt [99] for $1 < p < 2$. The upper bound was shown first for the L_2-discrepancy by Davenport [15] for $s = 2$, by Roth [95, 96] and Frolov [40] for arbitrary dimensions $s \in \mathbb{N}$ and by Chen [10] for the general L_p case. But we know even more. For any $p > 1$, any dimension $s \in \mathbb{N}$ and any integer $N \geq 2$ there is an explicit construction of a point set $\mathscr{P}$ consisting of N points in the s-dimensional unit cube such that

$$L_p(\mathscr{P}) \ll_{s,p} \frac{(\log N)^{(s-1)/2}}{N},$$

where $A \ll_{s,p} B$ means that there is a constant $c'_{s,p} > 0$ depending only on s and p, such that $A \leq c'_{s,p} B$. Such a construction was first given by Davenport for $p = s = 2$ and by Chen and Skriganov [11] for the case $p = 2$ and arbitrary dimension s. Later Skriganov [102] generalised this construction to the L_p case with arbitrary $p > 1$. This construction is also explained in Chen and Skriganov [12] and in [33, Chapter 16].

2.2 Discrepancy and tractability

In many applications the dimension s can be rather large. But in this case, the asymptotically almost optimal bounds on the discrepancy given, e.g., in (11), are even not useful for a modest number N of points. For example, assume that for every $s, N \in \mathbb{N}$ we have a point set $\mathscr{P}_{s,N}$ in the s-dimensional unit cube of cardinality N with star-discrepancy of at most

$$D_N^*(\mathscr{P}_{s,N}) \ll_s \frac{(\log N)^s}{N}.$$

Hence for any $\varepsilon > 0$ the star-discrepancy behaves asymptotically like $N^{-1+\varepsilon}$, which is the optimal rate of convergence since for dimension $s = 1$ we already have

$D_N^*(\mathscr{P}_{1,N}) \geq 1/(2N)$. However, the function $N \to (\log N)^s/N$ decreases to zero not until $N \geq \mathrm{e}^s$. For $N \leq \mathrm{e}^s$ this function is increasing which means that for cardinality N in this range our discrepancy bounds are useless. But even for moderately large dimension s, the value of e^s is huge, such that point sets with cardinality $N \geq \mathrm{e}^s$ cannot be used for practical applications. Therefore, the bound (11) is only useful if N is large compared to the dimension s.

Hence we are interested in the discrepancy of point sets with not too large cardinality N (compared to s). To analyse this problem systematically one considers the following quantity. For $\varepsilon > 0$ let

$$N_\infty(s,\varepsilon) = \min\{N \in \mathbb{N} \,:\, \mathrm{disc}_\infty(N,s) \leq \varepsilon\},$$

the so-called *inverse of the L_∞-discrepancy.* This is the minimal cardinality N of a point set in $[0,1)^s$ such that we can achieve a star-discrepancy not larger than ε.

It is known that

$$\mathrm{disc}_\infty(N,s) \leq c\sqrt{\frac{s}{N}} \tag{13}$$

for all $N, s \in \mathbb{N}$ from which it follows that

$$N_\infty(s,\varepsilon) \leq Cs\varepsilon^{-2} \tag{14}$$

for some positive constants c and C. This was shown first by Heinrich, Novak, Wasilkowski and Woźniakowski [48] by using deep results from probability theory. Later, Aistleitner [2] showed by a simplified argument that in (13) one can even choose $c = 10$.

Hence, the inverse of star-discrepancy depends only polynomially on s and ε^{-1}. In Information-based Complexity (IBC) theory such a behaviour is called *polynomial tractability.*

Furthermore, it is known that the dependence on the dimension s of the upper bound on the Nth minimal star-discrepancy in (14) cannot be improved. It was shown by Hinrichs [49, Theorem 1] that there exist constants $c, \varepsilon_0 > 0$ such that

$$N_\infty(s,\varepsilon) \geq cs/\varepsilon$$

for $0 < \varepsilon < \varepsilon_0$ and $\mathrm{disc}_\infty(N,s) \geq \min(\varepsilon_0, cs/n)$.

The bound (13) is only an existence result. Until now no explicit construction of a point set $\mathscr{P}$ of cardinality N in $[0,1)^s$ for which $D_N^*(\mathscr{P})$ satisfies (13) is known. A first constructive approach of such points for which the bound (13) is nearly achieved is given in Doerr, Gnewuch and Srivastav [37] which is further improved in Doerr and Gnewuch [36]. There, a deterministic algorithm is presented that constructs point sets $\mathscr{P}_{N,s}$ consisting of N points in $[0,1)^s$ satisfying

$$D_N^*(\mathscr{P}_{N,s}) = O\left(\frac{s^{1/2}}{N^{1/2}}(\log(N+1))^{1/2}\right)$$

in run-time $O(s\log(sN)(\sigma N)^s)$, where $\sigma = \sigma(s) = O((\log s)^2/(s\log\log s)) \to 0$ as $s \to \infty$ and where the implied constants in the O-notations are independent of s and N. However, this is still by far too expensive to obtain point sets for high dimensional applications. A small improvement for the run time is presented in Doerr, Gnewuch, Kritzer and P. [38]. However, this improvement has to be payed with a worse dependence of the bound for the star-discrepancy on the dimension s.

Let us now turn our attention to the analogue problem for the L_2-discrepancy instead of star-discrepancy. Contrary to the star-discrepancy here it makes little sense to ask for the smallest cardinality of a point set with L_2-discrepancy of at most some $\varepsilon > 0$. The reason for this is that the L_2-discrepancy of the empty point set in the s-dimensional unit cube is exactly $3^{-s/2}$, which follows from (5), or in other words, $\mathrm{disc}_2(0,s) = 3^{-s/2}$. Thus for s large enough, the empty set has always L_2-discrepancy smaller than ε. (This is not the case for the star-discrepancy which is always one for the empty set.) This may suggest that for large s, the L_2-discrepancy is not properly scaled. In the following we therefore use the L_2-discrepancy of the empty point set $\mathrm{disc}_2(0,s)$ as a reference.

In general, for $1 \le p \le \infty$, $s \in \mathbb{N}$ and $\varepsilon > 0$ the *inverse of the L_p-discrepancy* is hence defined as

$$N_p(s,\varepsilon) = \min\left\{N \in \mathbb{N} \,:\, \mathrm{disc}_p(N,s) \le \varepsilon\, \mathrm{disc}_p(0,s)\right\}.$$

For $N_2(s,\varepsilon)$ the situation is quite different compared to $N_\infty(s,\varepsilon)$. It was shown in Sloan and Woźniakowski [107] and Woźniakowski [113] (in a much more general setting) that for $\varepsilon \in (0,1)$ we have

$$N_2(s,\varepsilon) \ge (1-\varepsilon^2)\left(\frac{9}{8}\right)^s. \tag{15}$$

Hence $N_2(s,\varepsilon)$ grows exponentially in dimension s. A direct proof of (15) is also presented in [33, Proof of Proposition 3.58]. For more general results see Novak and Woźniakowski [85, Chapter 11]. Hence the inverse of the L_2-discrepancy depends at least exponentially on the dimension s. In IBC theory this exponential dependence on the dimension is called *intractability* or the *curse of dimensionality*. For a more detailed discussion of tractability of various notions of discrepancy we refer to the work of Novak and Woźniakowski [82, 83, 84, 85].

2.3 Weighted discrepancy and strong tractability

One of the reasons for introducing a weighted discrepancy in Section 1.4 is that with this concept one can overcome the curse of dimensionality for the L_2-discrepancy under suitable conditions on the weights $\boldsymbol{\gamma}$. Also for the weighted star-discrepancy one can obtain a weaker dependence on the dimension for suitable choices of weights.

For $s, N \in \mathbb{N}$ and $1 \le p \le \infty$ and for a sequence $\boldsymbol{\gamma}$ of weights we define

$$\mathrm{disc}_{p,\boldsymbol{\gamma}}(N,s) = \inf_{\substack{\mathscr{P} \subseteq [0,1)^s \\ |\mathscr{P}| = N}} L_{p,\boldsymbol{\gamma}}(\mathscr{P}).$$

For $s \in \mathbb{N}$ and $\varepsilon > 0$ the *inverse of the weighted L_p-discrepancy* is defined as

$$N_{p,\boldsymbol{\gamma}}(s,\varepsilon) = \min\left\{N \in \mathbb{N} \,:\, \mathrm{disc}_{p,\boldsymbol{\gamma}}(N,s) \le \varepsilon\, \mathrm{disc}_{p,\boldsymbol{\gamma}}(0,s)\right\}.$$

In Hinrichs, P., and Schmid [50, Theorem 1] it has been shown that there exists a constant $C > 0$ such that

$$\mathrm{disc}_{\infty,\boldsymbol{\gamma}}(N,s) \le C \frac{1+\sqrt{\log s}}{\sqrt{N}} \max_{\emptyset \ne \mathfrak{u} \subseteq I_s} \gamma_{\mathfrak{u}} \sqrt{|\mathfrak{u}|}. \tag{16}$$

Hence, if

$$\sup_{s=1,2,\ldots} \max_{\emptyset \neq \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}} \sqrt{|\mathfrak{u}|} < \infty, \tag{17}$$

then there exists a $C_{\boldsymbol{\gamma}} > 0$ such that

$$\mathrm{disc}_{\infty,\boldsymbol{\gamma}}(N,s) \leq C_{\boldsymbol{\gamma}} \frac{1+\sqrt{\log s}}{\sqrt{N}},$$

and therefore

$$N_{\infty,\boldsymbol{\gamma}}(\varepsilon,s) \leq \left\lceil \widetilde{C}_{\boldsymbol{\gamma}} \left(1+\sqrt{\log s}\right)^2 \varepsilon^{-2} \right\rceil.$$

for some $\widetilde{C}_{\boldsymbol{\gamma}} > 0$. This means that the weighted star-discrepancy is polynomially tractable whenever the weights satisfy condition (17). Compared to the usual star-discrepancy, see (14), here we have a much weaker dependence on the dimension s. Note that (17) is a very mild condition on the weights. It is enough that the weights γ_i are decreasing and that $\gamma_i < 1$ for an index $i \in \mathbb{N}$. Under a stronger condition on the weights one can even obtain the following property.

If $\sum_{i \geq 1} \gamma_i < \infty$, then for any $\delta > 0$ there exists a $C_{\delta,\boldsymbol{\gamma}} > 0$ such that

$$\mathrm{disc}_{\infty,\boldsymbol{\gamma}}(N,s) \leq \frac{C_{\delta,\boldsymbol{\gamma}}}{N^{1-\delta}} \tag{18}$$

and hence

$$N_{\infty,\boldsymbol{\gamma}}(\varepsilon,s) \leq \left\lceil \widetilde{C}_{\delta,\boldsymbol{\gamma}} \varepsilon^{-\frac{1}{1-\delta}} \right\rceil \tag{19}$$

for some $\widetilde{C}_{\delta,\boldsymbol{\gamma}} > 0$. Since this bound is even independent of the dimension one says that the weighted star-discrepancy is *strongly* polynomially tractable. The bound in (18) can be achieved with a superposition of polynomial lattice point sets as discussed in Section 5; see [33, Corollary 10.30].

We know from Section 2.2 that the classical L_2-discrepancy is subject to the curse of dimensionality. This disadvantage can be overcome when we change to the weighted setting.

Averaging the squared weighted L_2-discrepancy yields

$$\begin{aligned} &\int_{[0,1]^{sN}} (L_{2,\boldsymbol{\gamma}}(\{\boldsymbol{\tau}_1,\ldots,\boldsymbol{\tau}_N\}))^2 \,\mathrm{d}\boldsymbol{\tau}_1 \cdots \mathrm{d}\boldsymbol{\tau}_N \\ &= \frac{1}{N}\left(\prod_{i=1}^s \left(1+\frac{\gamma_i^2}{2}\right) - \prod_{i=1}^s \left(1+\frac{\gamma_i^2}{3}\right)\right) \end{aligned}$$

and hence

$$\mathrm{disc}_{2,\boldsymbol{\gamma}}(N,s) \leq \frac{1}{N^{1/2}}\left(\prod_{i=1}^s \left(1+\frac{\gamma_i^2}{2}\right) - \prod_{i=1}^s \left(1+\frac{\gamma_i^2}{3}\right)\right)^{1/2}.$$

Note that $\mathrm{disc}_{2,\boldsymbol{\gamma}}(0,s) = \left(-1+\prod_{i=1}^s \left(1+\frac{\gamma_i^2}{3}\right)\right)^{1/2}$. Therefore, we obtain

$$\frac{\mathrm{disc}_{2,\boldsymbol{\gamma}}(N,s)}{\mathrm{disc}_{2,\boldsymbol{\gamma}}(0,s)} \leq \frac{1}{N^{1/2}} \exp\left(\frac{1}{6}\sum_{i=1}^s \gamma_i^2\right) \tag{20}$$

(for details we refer to Sloan and Woźniakowski [107], see also [33, Proof of Theorem 3.64]). Hence if $\sum_{i\geq 1} \gamma_i^2 < \infty$ then there exists a $C_{\boldsymbol{\gamma}} > 0$ such that

$$N_{2,\boldsymbol{\gamma}}(\varepsilon, s) \le C_{\boldsymbol{\gamma}} \varepsilon^{-2}.$$

Again this bound is independent of the dimension s and hence the weighted L_2-discrepancy is strongly tractable as long as the squared weights γ_i^2, $i \ge 1$, are summable. On the other hand, this condition is also necessary for strong tractability which follows from (15) (see again [33, Proof of Theorem 3.64] for details).

If we only would have $\limsup_{s\to\infty} \sum_{i=1}^s \gamma_i^2/(\log s)$, then we still obtain from (20) that the weighted L_2-discrepancy is polynomially tractable.

Further results on the tractability of weighted discrepancy can be found in [33], Hinrichs, P. and Schmid [50], Leobacher and P. [68] and Novak and Woźniakowski [85].

2.4 Definition of tractability for the worst-case integration error

Let us return to the integration problem for functions from a reproducing kernel Hilbert space $\mathcal{H}(K)$. By $e(\mathcal{H}(K); \mathscr{P})$ we denote the worst-case error of a quasi-Monte Carlo rule based on the point set $\mathscr{P}$. The initial error is defined by

$$e(\mathcal{H}(K); \emptyset) = \sup_{f \in \mathcal{H}(K), \|f\|_{\mathcal{H}(K)} \le 1} \left| \int_{[0,1]^s} f(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} \right|.$$

For $\varepsilon > 0$ let $N_{\mathcal{H}(K)}(\varepsilon, s)$ denote the minimal number of nodes that are required to reduce the initial error by a factor of ε, i.e.,

$$N_{\mathcal{H}(K)}(\varepsilon, s) = \min\{N \in \mathbb{N} : \exists \mathscr{P} \subseteq [0,1)^s, |\mathscr{P}| = N \text{ and } e(\mathcal{H}(K); \mathscr{P}) \le \varepsilon e(\mathcal{H}(K); \emptyset)\}.$$

This number is called the *information complexity* of QMC integration in $\mathcal{H}(K)$.

Now one says that multivariate integration in the space $\mathcal{H}(K)$ is *polynomially (QMC) tractable*, if there exist non-negative C, α, β such that

$$N_{\mathcal{H}(K)}(\varepsilon, s) \le C s^\alpha \varepsilon^{-\beta}$$

holds for all dimensions $s \in \mathbb{N}$ and for all $\varepsilon > 0$. If this inequality holds with $\alpha = 0$, then one says that multivariate integration in the space $\mathcal{H}(K)$ is *strongly (polynomially) (QMC) tractable*. The infima α and β are called the *s-exponent* and the *ε-exponent* of (strong) polynomial (QMC) tractability.

We remark that there are further notions of tractability such as, e.g., weak tractability or T-tractability. For more information we refer to the books by Novak and Woźniakowski [83, 85].

3 Low discrepancy point sets and sequences

As stated at the beginning, quasi-Monte Carlo rules use deterministic constructions of quadrature points which yield small integration errors. For the reproducing kernel Hilbert spaces on $[0,1]^s$, we know from Section 1 that this amounts to constructing point sets with small discrepancy. Explicit constructions of low discrepancy sequences where given by Sobol [109], Faure [39], Niederreiter [76] and Niederreiter-Xing [79]. The following section gives an introduction to the underlying ideas.

3.1 Nets and sequences

The aim is to construct a point set $\mathscr{P}_N = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}\}$ (in this context it is convenient to index the points from 0 rather than 1) such that the discrepancy $\|\Delta_{\mathscr{P}_N}\|_{L_p}$ converges with the (almost) optimal order. To do so, we discretise the problem by choosing the point set $\mathscr{P}_N$ such that the local discrepancy $\Delta_{\mathscr{P}_N}(\boldsymbol{z}) = 0$ for certain $\boldsymbol{z} \in [0,1]^s$ (those $\boldsymbol{z}$ in turn are chosen such that the discrepancy of $\mathscr{P}_N$ is small, as we explain below).

It turns out that, if one chooses a base $b \ge 2$ and $N = b^m$, for every dimension $s \ge 1$ there exists a nonnegative integer t such that for all positive integers m there exists a point set $\mathscr{P}_{b^m} = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{b^m-1}\}$ such that $\Delta_{\mathscr{P}_{b^m}}(\boldsymbol{z}) = 0$ for all $\boldsymbol{z} = (z_1, \ldots, z_s)$ of the form

$$z_i = \frac{a_i}{b^{d_i}} \quad \text{for } 1 \le i \le s,$$

where $0 < a_i \le b^{d_i}$ is an integer and $d_1 + \cdots + d_s \le m - t$ with $d_1, \ldots, d_s \in \mathbb{N}_0$. We stress that the value of t can be chosen independently of m (but has to dependent on s). A point set $\mathscr{P}_{b^m}$ which satisfies this property is called a (t,m,s)-net in base b. An equivalent description of (t,m,s)-nets in base b is given in the following definition.

Definition 1. *Let $b \ge 2$, $m, s \ge 1$ and $0 \le t \le m$ be integers. A point set $\mathscr{P}_{b^m} = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{b^m-1}\} \subseteq [0,1)^s$ is called a (t,m,s)-net in base b, if for all $d_1, \ldots, d_s \in \mathbb{N}_0$ with $d_1 + \cdots + d_s = m - t$, the elementary interval*

$$\prod_{i=1}^{s} \left[\frac{a_i}{b^{d_i}}, \frac{a_i+1}{b^{d_i}}\right)$$

contains exactly b^t points of $\mathscr{P}_{b^m}$ for all integers $0 \le a_i < b^{d_i}$.

A sequence of points $S = (\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots) \subseteq [0,1)^s$ is called a (t,s)-sequence in base b, if for all $k \ge 1$ and $m > t$ the point set

$$\{\boldsymbol{x}_{(k-1)b^m}, \ldots \boldsymbol{x}_{kb^m-1}\}$$

is a (t,m,s)-net in base b.

Clearly, every point set $\mathscr{P}_{b^m}$ in $[0,1)^s$ is a (t,m,s)-net in base b with $t = m$. Smaller values of t imply a stronger condition on the point set since elementary intervals of higher resolution are considered. This implies better distribution properties of the point set. However, a necessary condition such that a $(0,m,s)$-net in base b exists is $s \le b+1$, and a necessary condition such that a $(0,s)$-sequence in base b exists is $s \le b$. On the other hand, for fixed base b, for a (t,s)-sequence to exist we must have $t \ge c_b s + d_b$ for some constants $c_b > 0$ and d_b which depend on b but not on s. The parameter t is often referred to as the *quality parameter* of the net. An introduction into the theory of (t,m,s)-nets and (t,s)-sequences can be found in [33] and in Niederreiter [78, Chapter 4].

As an example, Fig. 2 shows a $(0,4,2)$-net in base 2 which is a $2^4 = 16$ element point set in $[0,1)^2$ where every elementary interval

$$\left[\frac{A}{2^d}, \frac{A+1}{2^d}\right) \times \left[\frac{B}{2^{4-d}}, \frac{B+1}{2^{4-d}}\right)$$

for $d \in \{0,1,2,3,4\}$, $A \in \{0, \ldots, 2^d - 1\}$ and $B \in \{0, \ldots, 2^{4-d}\}$ contains exactly one point.

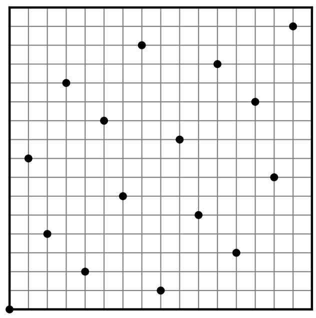

Fig. 2: A $(0,4,2)$-net in base 2.

3.2 Digital nets and sequences

Explicit constructions of (t,m,s)-nets can be obtained using the digital construction scheme. Such point sets are then called *digital nets* (or *digital* (t,m,s)*-nets* if the point set is a (t,m,s)-net).

To describe the digital construction scheme, let b be a prime number and let $\mathbb{Z}_b$ be the finite field of order b (a prime power and the finite field $\mathbb{F}_b$ could be used as well) and let $d,m \in \mathbb{N}$. Let $C_1,\ldots,C_s \in \mathbb{Z}_b^{dm\times m}$ be s matrices of size $dm \times m$ with elements in $\mathbb{Z}_b$ (the so-called *generating matrices*). The ith coordinate $x_{n,i}$ of the nth point $\boldsymbol{x}_n = (x_{n,1},\ldots,x_{n,s})$, $0 \le n < b^m$ and $1 \le i \le s$, of the digital net is obtained in the following way:

Digital construction scheme

- For $0 \le n < b^m$ let $n = n_0 + n_1 b + \cdots + n_{m-1} b^{m-1}$ be the base b representation of n and let $\mathbf{n} = (n_0,\ldots,n_{m-1})^\top \in \mathbb{Z}_b^m$ be the digit vector of n.
- Let
$$\mathbf{y}_{n,i} = C_i \mathbf{n}.$$
- For $\mathbf{y}_{n,i} = (y_{n,i,1},\ldots,y_{n,i,dm})^\top \in \mathbb{Z}_b^{dm}$ set
$$x_{n,i} = \frac{y_{n,i,1}}{b} + \cdots + \frac{y_{n,i,dm}}{b^{dm}}.$$

In order to obtain a sequence of points $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots$ one uses generating matrices of size $\infty \times \infty$, that is, $C_1,\ldots,C_s \in \mathbb{Z}_b^{\infty\times\infty}$. Such a sequence is then called a *digital sequence* (or *digital* (t,s)*-sequence* if the sequence is a (t,s)-sequence).

The classical construction of digital nets proposed by Niederreiter [78] uses $d = 1$.

The search for (t,m,s)-nets and (t,s)-sequences has now been reduced to finding suitable matrices $C_1,\ldots,C_s$. The geometric property of (t,m,s)-nets can also be translated into an algebraic property for the generating matrices.

Definition 2. *Let b be prime and $m,s \ge 1$ be a integers. Then the point set generated by the matrices $C_1,\ldots,C_s \in \mathbb{Z}_b^{m\times m}$ is called a digital (t,m,s)-net over $\mathbb{Z}_b$ if for all $d_1,\ldots,d_s \in \mathbb{N}_0$ with $\sum_{i=1}^s d_i \le m-t$ the system of vectors*

$$\mathbf{c}_{1,1},\ldots,\mathbf{c}_{1,d_1},\ldots,\mathbf{c}_{s,1},\ldots,\mathbf{c}_{s,d_s}\in\mathbb{Z}_b^m,$$

where $\mathbf{c}_{i,k}$ *denotes the* k*th row of* C_i*, is linearly independent over* $\mathbb{Z}_b$*.*

The sequence generated by the matrices $C_1,\ldots,C_s\in\mathbb{Z}_b^{\infty\times\infty}$ *is called a digital* (t,s)*-sequence over* $\mathbb{Z}_b$ *if for all* $m\ge t$ *the left-upper* $m\times m$ *submatrices* $C_1^{(m)},\ldots,C_s^{(m)}$ *of* $C_1,\ldots,C_s$ *generate a digital* (t,m,s)*-net over* $\mathbb{Z}_b$*.*

In Niederreiter [78] it has been shown that a digital (t,m,s)-net over $\mathbb{Z}_b$ is a (t,m,s)-net in base b and that a digital (t,s)-sequence over $\mathbb{Z}_b$ is a (t,s)-sequence in base b.

Explicit constructions of suitable generating matrices are available, see [33] and Niederreiter [78]. We describe the construction by Niederreiter as an example.

Let $s\in\mathbb{N}$, b be a prime number and let $p_1,\ldots,p_s\in\mathbb{Z}_b[x]$ be distinct monic irreducible polynomials over $\mathbb{Z}_b$. Let $e_i=\deg(p_i)$ for $1\le i\le s$. For $1\le i\le s$, $j\ge 1$ and $0\le k<e_i$, consider the expansions

$$\frac{x^{e_i-1-k}}{p_i(x)^j}=\sum_{r=0}^{\infty}a^{(i)}(j,k,r)x^{-r-1}$$

over the field $\mathbb{Z}_b((x^{-1}))$ of formal Laurent series. Then we define the matrix $C_i=(c_{j,r}^{(i)})_{j\ge1,r\ge0}$ by

$$c_{j,r}^{(i)}=a^{(i)}(Q+1,k,r)\in\mathbb{Z}_b\quad\text{for } 1\le i\le s, j\ge 1, r\ge 0, \tag{21}$$

where $j-1=Qe_i+k$ with integers $Q=Q(i,j)$ and $k=k(i,j)$ satisfying $0\le k<e_i$. Digital sequences for which generating matrices are given by (21) are called *Niederreiter sequences.* The following result holds:

Theorem 1 (Niederreiter [76, Theorem 1], D. and Niederreiter [29]). *The digital sequence with generating matrices* $C_1,\ldots,C_s\in\mathbb{Z}_b^{\infty\times\infty}$ *given by* (21) *is a* (t,s)*-sequence in base* b *with*

$$t=\sum_{i=1}^{s}(e_i-1).$$

If $p_i(x)=x-i-1\in\mathbb{Z}_b[x]$ for $1\le i\le s$, then we obtain the digital $(0,s)$-sequence over $\mathbb{Z}_b$ which is known as *Faure sequence* Faure [39]. By setting $b=2$, $p_1(x)=x\in\mathbb{Z}_2[x]$ and $p_2,\ldots,p_s$ are distinct primitive polynomials over $\mathbb{Z}_2$, then we obtain *Sobol' sequences* Sobol' [109].

3.3 Discrepancy bounds

We have obtained constructions of (t,m,s)-nets and (t,s)-sequences which yield uniformly distributed point sets and sequences. These nets are designed such that the local discrepancy is 0 for many points. Since the discrepancy can only vary slowly, one can expect that the discrepancy of the net itself is small. This is indeed the case. For instance, the following classical result holds.

Theorem 2 (Niederreiter [78, Theorem 4.5 and Theorem 4.6]). *The star-discrepancy of a* (t,m,s)*-net* $\mathscr{P}$ *in base* b *is bounded by*

$$D^*_{b^m}(\mathscr{P}) \le b^{-(m-t)} \sum_{i=0}^{s-1} \binom{s-1}{i} \binom{m-t}{i} \left\lfloor \frac{b}{2} \right\rfloor^i$$

for $b \ge 3$, and for $b = 2$ we have

$$D^*_{b^m}(\mathscr{P}) \le 2^{-(m-t)} \sum_{i=0}^{s-1} \binom{m-t}{i}.$$

To illustrate the basic idea for the proof of this discrepancy bound we show the result in the most simple case $s = b = 2$ and $t = 0$. A proof for the general result can be found in [33, Proof of Corollary 5.3].

Proof. For a measurable set C let $A(C)$ denote the number of elements of $\mathscr{P}$ which belong to C.

We consider an interval $B = [0, \alpha) \times [0, \beta)$ where the dyadic digit expansion of α and β is given by

$$\alpha = \frac{a_1}{2} + \frac{a_2}{2^2} + \cdots + \frac{a_m}{2^m} + \cdots,$$
$$\beta = \frac{b_1}{2} + \frac{b_2}{2^2} + \cdots + \frac{b_m}{2^m} + \cdots.$$

The basic idea is to approximate the interval B from the interior and from the exterior with disjoint unions of elementary intervals. Let

$$I_1 := \left[0, \frac{a_1}{2}\right) \times \left[0, \frac{b_1}{2} + \cdots + \frac{b_{m-1}}{2^{m-1}}\right),$$
$$J_1 := \left[0, \frac{a_1}{2}\right) \times \left[0, \frac{b_1}{2} + \cdots + \frac{b_{m-1}}{2^{m-1}} + \frac{1}{2^{m-1}}\right).$$

Then we have $I_1 \subseteq B$ and

$$I_1 = \bigcup_{k=0}^{2^{m-1}b_1 + \cdots + b_{m-1} - 1} \left[0, \frac{a_1}{2}\right) \times \left[\frac{k}{2^{m-1}}, \frac{k+1}{2^{m-1}}\right)$$

is a disjoint union of two-dimensional elementary intervals of area 2^{-m}. By the $(0, m, 2)$-net property we know that each of these intervals contains exactly one element of $\mathscr{P}$. Hence it follows that $A(I_1) = 2^m \lambda(I_1)$. In the same way it follows that $A(J_1) = 2^m \lambda(J_1)$.

Let further

$$I_k := \left[\frac{a_1}{2} + \cdots + \frac{a_{k-1}}{2^{k-1}}, \frac{a_1}{2} + \cdots + \frac{a_k}{2^k}\right) \times \left[0, \frac{b_1}{2} + \cdots + \frac{b_{m-k}}{2^{m-k}}\right),$$
$$J_k := \left[\frac{a_1}{2} + \cdots + \frac{a_{k-1}}{2^{k-1}}, \frac{a_1}{2} + \cdots + \frac{a_k}{2^k}\right) \times \left[0, \frac{b_1}{2} + \cdots + \frac{b_{m-k}}{2^{m-k}} + \frac{1}{2^{m-k}}\right)$$

for $1 \le k \le m-1$ and put

$$I_m := \left[\frac{a_1}{2} + \cdots + \frac{a_{m-1}}{2^{m-1}}, \frac{a_1}{2} + \cdots + \frac{a_m}{2^m}\right) \times [0, 0) = \emptyset,$$
$$J_m := \left[\frac{a_1}{2} + \cdots + \frac{a_{m-1}}{2^{m-1}}, \frac{a_1}{2} + \cdots + \frac{a_m}{2^m}\right) \times [0, 1).$$

Using the $(0, m, 2)$-net property again, it follows, in the same way as for I_1 and J_1, that $A(I_k) = 2^m \lambda(I_k)$ and $A(J_k) = 2^m \lambda(J_k)$ for all $1 \le k \le m$. Furthermore, note that $\lambda_2(J_k \setminus I_k) \le 2^{-m}$ for all $1 \le k \le m$.

Putting

$$\underline{B} := \bigcup_{k=1}^{m} I_k,$$
$$\overline{B} := \bigcup_{k=1}^{m} J_k \cup \left(\left[\frac{a_1}{2} + \cdots + \frac{a_m}{2^m}, \frac{a_1}{2} + \cdots + \frac{a_m}{2^m} + \frac{1}{2^m} \right) \times [0,1) \right)$$

we have $\underline{B} \subseteq B \subseteq \overline{B}$, $A(\underline{B}) = 2^m \lambda_2(\underline{B})$ and, by using the $(0, m, 2)$-net property again, $A(\overline{B}) = 2^m \lambda_2(\overline{B})$. Hence

$$\lambda_2(\underline{B}) = 2^{-m} A(\underline{B}) \le 2^{-m} A(B) \le 2^{-m} A(\overline{B}) = \lambda_2(\overline{B})$$

and

$$-\lambda_2(\overline{B}) \le -\lambda_2(B) \le -\lambda_2(\underline{B}).$$

Therefore, we obtain

$$\lambda_2(\underline{B}) - \lambda_2(\overline{B}) \le 2^{-m} A(B) - \lambda_2(B) \le \lambda_2(\overline{B}) - \lambda_2(\underline{B}),$$

and hence

$$\left| 2^{-m} A(B) - \lambda_2(B) \right| \le \lambda_2(\overline{B} \setminus \underline{B}) \le \frac{m}{2^m} + \frac{1}{2^m} = \frac{1}{2^m} \sum_{i=0}^{1} \binom{m}{i},$$

independent of the choice of B. □

The until now best asymptotic result for the star-discrepancy of general (t, m, s)-nets in base b has been shown by Kritzer [59].

Theorem 3 (Kritzer [59]). *The star-discrepancy of a (t, m, s)-net $\mathscr{P}$ in base b with $m > 0$ satisfies*

$$b^m D^*_{b^m}(\mathscr{P}) \le B(s, b) b^t m^{s-1} + O(b^t m^{s-2}),$$

where the implied O-constant depends only on b and s and where

$$B(s, b) = \left\lfloor \frac{b}{2} \right\rfloor^s \frac{1}{(b + (-1)^b)(s-1)!(\log b)^{s-1}}.$$

Thus, (t, m, s)-nets achieve a convergence order of $N^{-1}(\log N)^{s-1}$. Notice that since $L_p(\mathscr{P}) \le D^*_N(\mathscr{P})$ for all $1 \le p \le \infty$, this bound also applies to the L_p-discrepancy. Apart from the power in the $\log N$ factor, it is known that this rate of convergence is best possible.

3.4 Randomised quasi-Monte Carlo

So far we have considered deterministic constructions of quadrature points in the unit cube. The advantage of quadrature algorithms based on deterministic constructions is that the convergence rate of the integration error improves for functions with integrable partial mixed derivatives of order up to 1, which is not the case for so-called standard Monte Carlo (MC). Standard MC approximates the integral $\int_{[0,1]^s} f(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}$ with $\frac{1}{N}\sum_{n=1}^N f(\boldsymbol{z}_n)$, where $\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N$ are uniformly i.i.d. in $[0,1]^s$. However, there is also some merit in choosing the quadrature points randomly in $[0,1]^s$ as in the standard MC algorithm. The most obvious case of the usefulness of this choice is if the integrand does not have sufficient smoothness for QMC, in fact, standard MC works for functions in $L_2([0,1]^s)$. Another advantage is that one can obtain a statistical estimation of the variance of the estimator. Let

$$\widehat{I}(f;\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N) = \int_{[0,1]^s} f(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} - \frac{1}{N}\sum_{n=1}^N f(\boldsymbol{z}_n).$$

Since the quadrature points $\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N$ are chosen randomly from the uniform distribution, for each given f, the quantity $\widehat{I}(f;\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N)$ is a random variable. The variance $\mathrm{Var}(\widehat{I}(f;\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N))$ of $\widehat{I}(f;\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N)$ satisfies

$$\begin{aligned}
\mathrm{Var}(\widehat{I}(f;\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N)) &= \mathbb{E}\left(\widehat{I}^2(f;\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N)\right) \\
&= \int_{[0,1]^s}\cdots\int_{[0,1]^s} \widehat{I}^2(f;\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N)\,\mathrm{d}\boldsymbol{z}_1\ldots\mathrm{d}\boldsymbol{z}_N \\
&= \frac{1}{N}\left(\int_{[0,1]^s} f^2(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \left(\int_{[0,1]^s} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\right)^2\right) \\
&= \frac{\mathrm{Var}(f)}{N},
\end{aligned}$$

which shows the convergence rate of order $N^{-1/2}$ of the standard deviation

$$\mathrm{Std}(\widehat{I}(f;\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N)) = \sqrt{\mathrm{Var}(\widehat{I}(f;\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N))}$$

to the correct value. For $f \in L_2([0,1]^s)$, the variance decays with order $N^{-1/2}$ for the standard MC method. Functions with higher order smoothness do not yield an improved rate of convergence for standard MC.

The aim of randomised QMC is to construct a hybrid of MC and QMC with 'the best of both worlds'. To define a setting to analyse the variance in this case, one considers the *randomised error*, which one can also call the *worst-case-root-mean-square error*. That is, let $\mathcal{B}$ be some Banach space with norm $\|\cdot\|_{\mathcal{B}}$. Then the randomised error is defined as

$$e_{\mathrm{ran}}(\mathcal{B};\widetilde{\mathscr{P}}) = \sup_{f\in\mathcal{B},\|f\|_{\mathcal{B}}\le 1} \sqrt{\mathrm{Var}(\widehat{I}(f))},$$

where $\widetilde{\mathscr{P}} = \{\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N\}$ is some randomised point set in $[0,1]^s$ (concrete examples of such point set are discussed below). In the remainder of this subsection we consider $\mathcal{B}$ to be the reproducing kernel Hilbert space with reproducing kernel $K(\boldsymbol{x},\boldsymbol{y}) = \prod_{i=1}^s (1+\gamma_i \min(1-x_i, 1-y_i))$, i.e., $\mathcal{B} = \mathcal{H}(K)$.

There are several ways of obtaining randomised (t,m,s)-nets and (t,s)-sequences, such that the (t,m,s)-net and (t,s)-sequences structure, respectively, are preserved. A simple way of doing so is by using a digital shift $\boldsymbol{\sigma} \in [0,1]^s$. Assume that $\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{b^m-1}\} \subseteq [0,1)^s$ forms a (t,m,s)-net in base b, where $\boldsymbol{x}_n = (x_{n,1}, \ldots, x_{n,s})$ and $x_{n,i} = x_{n,i,1} b^{-1} + x_{n,i,2} b^{-2} + \cdots$. Let $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_s) \in [0,1)^s$, with $\sigma_i = \sigma_{i,1} b^{-1} + \sigma_{i,2} b^{-2} + \cdots$, be i.i.d. uniformly distributed in $[0,1)^s$. In all the b-adic representations we assume that infinitely many digits are different from $b-1$.

We now define the randomised point set $\{\boldsymbol{z}_0, \ldots, \boldsymbol{z}_{b^m-1}\}$, where $\boldsymbol{z}_n = (z_{n,1}, \ldots, z_{n,s})$ and $z_{n,i} = z_{n,i,1} b^{-1} + z_{n,i,2} b^{-2} + \cdots$. This is done by defining the digits $z_{n,i,k} \in \{0, \ldots, b-1\}$ by

$$z_{n,i,k} \equiv x_{n,i,k} + \sigma_{i,k} \pmod b \quad \text{for all } 0 \le n < b^m, 1 \le i \le s, k \ge 1. \tag{22}$$

The point set $\boldsymbol{z}_0, \ldots, \boldsymbol{z}_{b^m-1}$ is called a *randomly digitally shifted* (t,m,s)*-net*. It can be shown that, with probability 1, the randomly digitally shifted (t,m,s)-nets in base b are again (t,m,s)-nets in base b.

There are also variations of this method. For instance, one can use (22) for $1 \le k \le m$ and set $z_{n,i,k} = 0$ for $k > m$. Or one can use (22) for $1 \le k \le m$ and choose $z_{n,i,k}$ uniformly i.i.d. in $\{0, \ldots, b-1\}$ for $k > m$. We call this method a *digital shift of depth* m. The convergence rate of the randomised error for functions from the reproducing kernel Hilbert spaces considered in Section 1.3 is of order $N^{-1} (\log N)^{(s-1)/2}$; see Chen and Skriganov [12], Cristea, D., and P. [14], and [31]. However, it is known that the best possible convergence rate of the randomised error for this function space is of order $N^{-3/2} (\log N)^{c_1(s)}$, again with $c_1(s) \asymp s$; see Bakhvalov [4] and also Novak [81]. We discuss in the following a randomisation method for (t,m,s)-nets and (t,s)-sequences which yields an improvement of the convergence rate of the randomised error for the reproducing kernel Hilbert space with kernel $K(\boldsymbol{x}, \boldsymbol{y}) = \prod_{i=1}^s (1 + \gamma_i \min(1 - x_i, 1 - y_i))$. This method goes back to Owen [89, 90, 91] and is called *Owen's scrambling*, see also [33, Section 13.5].

Owen's scrambling algorithm is best described for some generic point $\boldsymbol{x} \in [0,1)^s$, with $\boldsymbol{x} = (x_1, \ldots, x_s)$ and $x_i = \xi_{i,1} b^{-1} + \xi_{i,2} b^{-2} + \cdots$. The scrambled point shall be denoted by $\boldsymbol{y} \in [0,1)^s$, where $\boldsymbol{y} = (y_1, \ldots, y_s)$ and $y_i = \eta_{i,1} b^{-1} + \eta_{i,2} b^{-2} + \cdots$. The point $\boldsymbol{y}$ is obtained by applying permutations to each digit of each coordinate of $\boldsymbol{x}$. The permutation applied to $\xi_{i,l}$ depends on $\xi_{i,k}$ for $1 \le k < l$. Specifically, $\eta_{i,1} = \pi_i(\xi_{i,1})$, $\eta_{i,2} = \pi_{i,\xi_{i,1}}(\xi_{i,2})$, $\eta_{i,3} = \pi_{i,\xi_{i,1},\xi_{i,2}}(\xi_{i,3})$, and in general

$$\eta_{i,k} = \pi_{i,\xi_{i,1},\ldots,\xi_{i,k-1}}(\xi_{i,k}), \tag{23}$$

where $\pi_{i,\xi_{i,1},\ldots,\xi_{i,k-1}}$ is a random permutation of $\{0, \ldots, b-1\}$. We assume that permutations with different indices are chosen mutually independent from each other and that each permutation is chosen with the same probability.

To describe Owen's scrambling, for $1 \le i \le s$ let

$$\varPi_i = \{\pi_{i,\xi_{i,1},\ldots\xi_{i,k-1}} : k \in \mathbb{N}, \xi_{i,1}, \ldots, \xi_{i,k-1} \in \{0, \ldots, b-1\}\}$$

be a given set of permutations, where for $k = 1$ we set $\pi_{i,\xi_{i,1},\ldots,\xi_{i,k-1}} = \pi_i$, and let $\boldsymbol{\varPi} = (\varPi_1, \ldots, \varPi_s)$. Then, on applying Owen's scrambling using these permutations to some point $\boldsymbol{x} \in [0,1)^s$, we write $\boldsymbol{y} = \boldsymbol{\varPi}(\boldsymbol{x})$ and $y_i = \varPi_i(x_i)$ for $1 \le i \le s$, where $\boldsymbol{y}$ is the point obtained by applying Owen's scrambling to $\boldsymbol{x}$ using the set of permutations $\boldsymbol{\varPi} = (\varPi_1, \ldots, \varPi_s)$.

For a (t,s)-sequence $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots$, the Owen scrambled sequence $\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots$ is then given by $\boldsymbol{y}_n = \boldsymbol{\Pi}(\boldsymbol{x}_n)$ for $n \ge 0$ (for (t,m,s)-nets one just uses $0 \le n < b^m$). The convergence rate of the random case error for Owen scrambled (t,m,s)-nets is then

$$e_{\mathrm{ran}}(\mathcal{H}(K); \widetilde{\mathscr{P}}) \ll_s \frac{(\log N)^{(s-1)/2}}{N^{3/2}}. \tag{24}$$

Further, Loh [69] even proved a central limit theorem for Owen scrambled $(0,m,s)$ nets.

Note that Owen's scrambling is complicated to implement, since all the randomly chosen permutations need to be stored. Therefore, several simplifications have been introduced which simplify the randomisation but still achieve (24). The main idea is to design randomisations such that Owen's lemma [90, Lemma 2] still holds. This then implies that also (24) still holds.

For instance, the following properties are sufficient for scrambling, see Hong and Hickernell [53] and Matoušek [71]:

A. Each of the sets of permutations Π_i is sampled from the same distribution $\mathcal{D}$ and these sampling are mutually independent.
B. If $x_i \in [0,1)$ is any real number and Π_i is drawn from the distribution $\mathcal{D}$, then $\Pi_i(x_i)$ is uniformly distributed in $[0,1)$.
C. Let $a_u = a_{u,1}b^{-1} + a_{u,2}b^{-2} + \cdots$ for $u = 1,2$ and $c_u = \Pi(a_u) = c_{u,1}b^{-1} + c_{u,2}b^{-2} + \cdots$. Assume that $a_{1,k} = a_{2,k}$ for $1 \le k < r$ and $a_{1,r} \ne a_{2,r}$. Then
   a. $c_{1,k} = c_{2,k}$ for $1 \le k < r$;
   b. $(c_{1,r}, c_{2,r})$ is uniformly distributed on the set $\{(d,e) \in \mathbb{Z}_b^2 : d \ne e\}$;
   c. $c_{u,k}$ are independent for $k > r$ and $u = 1,2$.

Further simplifications are possible, since often the precise distribution does not have to be known, only their first and second moments, see Matoušek [71].

A scrambling of digital nets which satisfies the conditions above and is easier to implement than Owen's scrambling is the following: Let $C_1, \ldots, C_s \in \mathbb{Z}_b^{\infty \times \infty}$ be generating matrices of a digital (t,s)-sequence over $\mathbb{Z}_b$ and let $L_1, \ldots, L_s \in \mathbb{Z}_b^{\infty \times \infty}$ be non-singular, lower triangular matrices. We choose those matrices randomly such that for $L_i = (\lambda_{u,v})_{u,v \ge 1}$, where $\lambda_{u,v} \in U\{0,\ldots,b-1\}$ i.i.d. for $u > v$, $\lambda_{u,u} \in U\{1,\ldots,b-1\}$ i.i.d. and $\lambda_{u,v} = 0$ for $u < v$. Further, let $\boldsymbol{e}_i \in \mathbb{Z}_b^{\infty}$ be chosen i.i.d. randomly in $U\{0,\ldots,b-1\}^{\infty}$. Then we obtain a scrambled sequence by setting

$$\mathbf{y}_{n,i} = L_i C_i \mathbf{n} + e_i \pmod b \quad \text{for } 1 \le i \le s,$$

and for $\mathbf{y}_{n,i} = (y_{n,i,1}, y_{n,i,2}, \ldots)^\top \in \mathbb{Z}_b^{\infty}$ we set

$$z_{n,i} = y_{n,i,1}b^{-1} + y_{n,i,2}b^{-2} + \cdots$$

and $\boldsymbol{z}_n = (z_{n,1}, \ldots, z_{n,s}) \in [0,1)^s$.

The sequence $(\boldsymbol{z}_0, \boldsymbol{z}_1, \ldots)$ is again a (t,s)-sequence in base b with probability 1 and also satisfies the properties A, B, C. Therefore, the randomised error for such a sequence is bounded by

$$e_{\mathrm{ran}}(\mathcal{H}(K); \widetilde{\mathscr{P}}) \ll_s \frac{(\log N)^{(s-1)/2}}{N^{3/2}}.$$

3.5 Higher order nets and sequences

Quasi-Monte Carlo rules using digital nets as quadrature points, as described above, achieve a convergence rate of the integration error of order $N^{-1}(\log N)^s$ for functions of bounded variation. If the integrand has more smoothness, then the above result does not yield a better rate of convergence. In [17, 18] it was shown how to construct QMC rules which can achieve a convergence rate of order $N^{-\alpha}(\log N)^{\alpha s}$ for integrands with square integrable partial mixed derivatives up to order α in each variable (we say that such functions have smoothness α in the following).

Let us now consider digital nets and digital sequences. Above we have seen that an algebraic property of the generating matrices ensures that the corresponding digital net has small discrepancy. These digital nets are therefore useful as quadrature points in a QMC algorithm. Similarly, we now explain the algebraic properties of the generating matrices of the digital nets necessary such that the corresponding QMC rules achieve the almost optimal rate of convergence for integrands of smoothness α. The following definition is a special case of [18, Definition 4.3 and Definition 4.8].

Definition 3. *Let $\alpha, m \in \mathbb{N}$. Let $C_1, \ldots, C_s \in \mathbb{Z}_b^{\alpha m \times m}$ be generating matrices of a digital net and let $\mathbf{c}_{i,k}$ denote the kth row of C_i. Then the point set generated by $C_1, \ldots, C_s$ is a digital $(t, \alpha, \alpha m \times m, s)$-net over $\mathbb{Z}_b$ if for all integers $1 \le j_{i,1} < \cdots < j_{i,\nu_i} \le \alpha m$, $1 \le i \le s$, such that*

$$\sum_{i=1}^{s} \sum_{k=1}^{\min(\alpha,\nu_i)} j_{i,k} \le \alpha m - t,$$

the row vectors

$$\mathbf{c}_{1,j_{1,1}}, \ldots, \mathbf{c}_{1,j_{1,\nu_i}}, \ldots, \mathbf{c}_{s,j_{s,1}}, \ldots, \mathbf{c}_{s,j_{s,\nu_s}}$$

are linearly independent over $\mathbb{Z}_b$.

Let $C_1, \ldots, C_s \in \mathbb{Z}_b^{\infty \times \infty}$ be the generating matrices of a digital sequence. If for all $m \ge t/\alpha$ the left-upper $\alpha m \times m$ submatrices $C_1^{(\alpha m, m)}, \ldots, C_m^{(\alpha m, m)}$ of $C_1, \ldots, C_s$ generate a digital $(t, \alpha, \alpha m \times m, s)$-net, then the sequence generated by $C_1, \ldots, C_s$ is a digital (t, α, s)-sequence over $\mathbb{Z}_b$.

There is an explicit construction method for such higher order nets and sequences which works the following way.

Higher order net construction

- Choose a $(t, m, \alpha s)$-net in base b whose elements are of the form $\boldsymbol{x}_n = (x_{n,1}, \ldots, x_{n,\alpha s}) \in [0,1]^{\alpha s}$ and $x_{n,i} = x_{n,i,1} b^{-1} + \cdots + x_{n,i,m} b^{-m} + \cdots$ for $1 \le i \le \alpha s$ and $0 \le n < b^m$.
- For $0 \le n < b^m$ define $\boldsymbol{y}_n = (y_{n,1}, \ldots, y_{n,s}) \in [0,1]^s$ by

$$y_{n,i} = \sum_{j=1}^{m} \sum_{k=1}^{\alpha} x_{n,(i-1)\alpha+k,j} b^{-k-(j-1)\alpha} \quad \text{for } 1 \le i \le s.$$

The net $\{\boldsymbol{y}_0, \ldots, \boldsymbol{y}_{b^m-1}\}$ is called a *higher order net.* This construction can easily be extended to higher order sequences; see [33, Section 15.2]. Furthermore, one can also apply the construction method to the generating matrices directly.

The following explicit construction method of suitable generating matrices was introduced in [17, 18].

Let $\alpha \geq 1$ and let $C_1, \ldots, C_{s\alpha}$ be the generating matrices of a digital $(t', m, \alpha s)$-net. As we will see later, the choice of the underlying digital $(t', m, \alpha s)$-net has a direct impact on the bound of the t-value of the digital $(t, \alpha, \alpha m \times m, s)$-net, which was proven in [17, 18]. Let $C_j = (\mathbf{c}_{j,1}, \ldots, \mathbf{c}_{j,m})^\top$ for $1 \leq j \leq \alpha s$; i.e., $\mathbf{c}_{j,l}$ are the row vectors of C_j. Now let the matrix $C_j^{(\alpha)}$ consist of the first rows of the matrices $C_{(j-1)\alpha+1}, \ldots, C_{j\alpha}$, then the second rows of $C_{(j-1)\alpha+1}, \ldots, C_{j\alpha}$, and so on, in the order described in the following: The matrix $C_j^{(\alpha)}$ is a $\alpha m \times m$ matrix; i.e., $C_j^{(\alpha)} = (\mathbf{c}_{j,1}^{(\alpha)}, \ldots, \mathbf{c}_{j,\alpha m}^{(\alpha)})^\top$, where $\mathbf{c}_{j,l}^{(\alpha)} = \mathbf{c}_{u,v}$ with $l = (v-j)\alpha + u$, $1 \leq v \leq m$, and $(j-1)\alpha < u \leq j\alpha$ for $1 \leq l \leq \alpha m$ and $1 \leq j \leq s$.

We remark that this construction can be extended to digital (t, α, s)-sequences by letting $C_j = (\mathbf{c}_{j,1}, \mathbf{c}_{j,2}, \ldots)^\top$, for $1 \leq j \leq \alpha s$, denote the generating matrices of a digital $(t', \alpha s)$-sequence; the resulting matrices $C_j^{(\alpha)}$, $1 \leq j \leq s$, are now $\infty \times \infty$ matrices, where again we have $C_j^{(\alpha)} = (\mathbf{c}_{j,1}^{(\alpha)}, \mathbf{c}_{j,2}^{(\alpha)}, \ldots)^\top$, where $\mathbf{c}_{j,l}^{(\alpha)} = \mathbf{c}_{u,v}$ with $l = (v-j)\alpha + u$, $v \geq 1$, and $(j-1)\alpha < u \leq j\alpha$ for $l \geq 1$ and $1 \leq j \leq s$. We have the following result on the quality parameter:

Theorem 4 ([18]). *Let $\alpha \in \mathbb{N}$ and let $C_1, \ldots, C_{\alpha s}$ be the generating matrices of a digital $(t', m, \alpha s)$-net over $\mathbb{Z}_b$ of prime order b. Let $C_1^{(\alpha)}, \ldots, C_s^{(\alpha)}$ be defined as above. Then the matrices $C_1^{(\alpha)}, \ldots, C_s^{(\alpha)}$ are the generating matrices of a digital $(t, \alpha, \alpha m \times m, s)$-net over $\mathbb{Z}_b$ with*

$$t = \alpha \min\left(m, t' + \left\lfloor \frac{s(\alpha-1)}{2} \right\rfloor\right).$$

Furthermore, for $m = \infty$, the matrices $C_1^{(\alpha)}, \ldots, C_s^{(\alpha)}$ obtained from the generating matrices $C_1, \ldots, C_{\alpha s}$ of a $(t', \alpha s)$-sequence over $\mathbb{Z}_b$ are the generating matrices of a digital (t, α, s)-sequence over $\mathbb{Z}_b$ with

$$t = \alpha\left(t' + \left\lfloor \frac{s(\alpha-1)}{2} \right\rfloor\right).$$

A slight improvement of this result for some cases can be found in D. and Kritzer [23].

We have the following result on the absolute error for the integration of function with smoothness α with QMC rules based on higher order nets.

Theorem 5 ([18], [33, Chapter 15]). *Let $\{\boldsymbol{y}_0, \ldots, \boldsymbol{y}_{b^m-1}\}$ be a higher order net constructed from a digital $(t', m, \alpha s)$-net in base b. Assume that the integrand f has smoothness α. Then the absolute error converges with order*

$$\left| \int_{[0,1]^s} f(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} - \frac{1}{b^m} \sum_{n=0}^{b^m-1} f(\boldsymbol{y}_n) \right| \ll_{s,b} \frac{m^{\alpha s}}{b^{m\alpha - t}},$$

where $t = \alpha \min(t' + \lfloor s(\alpha-1)/2 \rfloor, m)$.

Furthermore, it has been shown that, asymptotically, the t-value achieved in the above construction is optimal in the following sense.

Theorem 6 (D. and Baldeaux [22, Theorem 5]). *Assume that $t, \alpha, s, b \in \mathbb{N}$, b prime, are such that there exists a (t, α, s)-sequence over $\mathbb{Z}_b$. Then*

$$t > s\frac{\alpha(\alpha-1)}{2} - \alpha.$$

Since explicit constructions of (t, s)-sequences with $t = \mathcal{O}(s)$ are known, see Niederreiter and Xing [80], it follows that the asymptotic behaviour of the t-value of digital (t, α, s)-sequences is

$$t \asymp_b s\alpha^2.$$

Furthermore, explicit constructions can be obtained using the method from [17, 18] introduced above. However, it would be interesting to find other explicit constructions of higher order nets and sequences which can achieve smaller values of t for small values of m and s.

3.6 Scrambled higher order nets

In the previous section we have seen how QMC rules can be constructed such that the integration error converges with order $N^{-\alpha}(\log N)^{\alpha s}$. Then, the question arises whether there is also a generalisation of Owen's scrambling such that the randomised error achieves a higher rate of convergence as well. An affirmative answer can be given with the following construction of 'higher order scrambled' nets.

Scrambled higher order nets

- Choose a $(t, m, \alpha s)$-net in base b whose elements are of the form $\boldsymbol{x}_n = (x_{n,1}, \ldots, x_{n,\alpha s}) \in [0,1]^{\alpha s}$ and $x_{n,i} = x_{n,i,1}b^{-1} + \cdots + x_{n,i,m}b^{-m} + \cdots$ for $1 \le i \le \alpha s$ and $0 \le n < b^m$.
- Apply Owen's scrambling or one of its simplifications to the digital net to obtain randomised point set $\{\boldsymbol{z}_0, \ldots, \boldsymbol{z}_{b^m-1}\}$ where $\boldsymbol{z}_n = (z_{n,1}, \ldots, z_{n,\alpha s}) \in [0,1]^{\alpha s}$ and $z_{n,i} = z_{n,i,1}b^{-1} + \cdots + z_{n,i,m}b^{-m} + \cdots$ for $1 \le i \le \alpha s$ and $0 \le n < b^m$.
- For $0 \le n < b^m$ define $\boldsymbol{y}_n = (y_{n,1}, \ldots, y_{n,s}) \in [0,1]^s$ by

$$y_{n,i} = \sum_{j=1}^{\infty} \sum_{k=1}^{\alpha} z_{n,(i-1)\alpha+k,j} b^{-k-(j-1)\alpha} \quad \text{for } 1 \le i \le s.$$

The net $\{\boldsymbol{y}_0, \ldots, \boldsymbol{y}_{b^m-1}\}$ is called a *scrambled higher order net*. Again, this construction can easily be extended to higher order sequences.

The points $\boldsymbol{y}_0, \boldsymbol{y}_1, \ldots$ can be used in a QMC rule to obtain an improved convergence rate for the randomised error for functions with smoothness α.

Theorem 7 ([20, Theorem 10]). *Let $\{\boldsymbol{y}_0, \ldots, \boldsymbol{y}_{b^m-1}\}$ be a scrambled higher order net constructed from a digital $(t, m, \alpha s)$-net in base b. Assume that the integrand f has smoothness α. Then the randomised error converges with order*

$$\sqrt{\mathrm{Var}(\widehat{I}(f))} \ll_{s,b} \frac{m^{(\alpha+1)s/2}}{b^{(m-t)(\alpha+1/2)}}.$$

Apart from the power of the $\log_b N(= m)$ factor, this convergence order cannot be improved; see Bakhvalov [4] and also Novak [81].

A few remarks are in order. One cannot change the order in the construction of higher order scrambled nets and sequences. That is, if one applies the higher order construction first, and the scrambling method afterwards, one does not obtain an improved rate of convergence. Further, currently there is no known scrambling method for general higher order digital nets and sequences (only for higher order digital nets or sequences which have been obtain using the higher order construction above). Thus, for instance, the scrambling method above cannot be applied to higher order polynomial lattice rules, which will be defined below.

3.7 Digitally Shifted Nets

Digital shifts of depth m of a (t, m, s)-net have already been introduced in Section 3.4. Here we consider a so-called simplified shift of depth m.

Assume that $\mathscr{P} = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{b^m-1}\}$ forms a (t, m, s)-net in base b where $\boldsymbol{x}_n = (x_{n,1}, \ldots, x_{n,s})$ and $x_{n,i} = x_{n,i,1}b^{-1} + x_{n,i,2}b^{-2} + \cdots$. Choose $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_s)$ where $\sigma_i = \sigma_{i,1}b^{-1} + \cdots + \sigma_{i,m}b^{-m}$ and $\sigma_{i,j}$ are independent and uniformly distributed in $\{0, \ldots, b-1\}$.

We now define the randomised point set $\{\boldsymbol{z}_0, \ldots, \boldsymbol{z}_{b^m-1}\}$, where $\boldsymbol{z}_n = (z_{n,1}, \ldots, z_{n,s})$ and

$$z_{n,i} = \frac{z_{n,i,1}}{b} + \cdots + \frac{z_{n,i,m}}{b^m} + \frac{1}{2b^m},$$

where

$$z_{n,i,k} \equiv x_{n,i,k} + \sigma_{i,k} \pmod b \quad \text{for all } 0 \le n < b^m, 1 \le i \le s, 1 \le k \le m.$$

Such a digital shift is called a *simplified digital shift of depth m*. We denote a point set $\mathscr{P}$ that is digitally shifted by a simplified digital shift of depth m by $\widehat{\mathscr{P}_{\boldsymbol{\sigma}}}$.

Note that for the simplified digital shift, we only have b^{sm} possibilities, which means only m digits per dimension need to be selected for performing a simplified digital shift.

A simplified digital shift (of depth m) preserves the (t, m, s)-net structure; see [33, Section 4.4.4].

Theorem 8 (Cristea, D., and P. [14, Theorem 1]). *Let $\mathscr{P}$ be a digital (t, m, s)-net over $\mathbb{Z}_b$ with generating matrices $C_1, \ldots, C_s$. Then the mean square weighted L_2-discrepancy of $\widehat{\mathscr{P}_{\boldsymbol{\sigma}}}$ is given by*

$$\begin{aligned}
&\mathbb{E}[L_{2,\boldsymbol{\gamma}}^2(\widehat{\mathscr{P}_{\boldsymbol{\sigma}}})] \\
&= \sum_{\emptyset \ne \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}}^2 \left[2\left(\frac{1}{3^{|\mathfrak{u}|}} - \left(\frac{1}{3} + \frac{1}{24b^{2m}} \right)^{|\mathfrak{u}|} \right) + \frac{1}{b^m 2^{|\mathfrak{u}|}} \left(1 - \left(1 - \frac{1}{3b^m} \right)^{|\mathfrak{u}|} \right) \right] \\
&\quad + \sum_{\substack{\emptyset \ne \mathfrak{u} \subseteq [s] \\ \mathfrak{u} = \{u_1, \ldots, u_e\}}} \frac{\boldsymbol{\gamma}_{\mathfrak{u}}^2}{3^{|\mathfrak{u}|}} \sum_{\substack{k_1, \ldots, k_e = 0 \\ (k_1, \ldots, k_e) \ne (0, \ldots, 0) \\ C_{u_1}^\top \mathbf{k}_1 + \cdots + C_{u_e}^\top \mathbf{k}_e = \mathbf{0}}}^{b^m - 1} \prod_{i=1}^{e} \psi(k_i),
\end{aligned}$$

where $\psi(0) = 1$ and

$$\psi(k) = \frac{3}{2b^{2(r+1)}}\left(\frac{1}{\sin^2(\kappa_r \pi/p)} - \frac{1}{3}\right)$$

if $k = \kappa_0 + \kappa_1 b + \cdots + \kappa_r b^r$ *with* $\kappa_i \in \{0, \ldots, b-1\}$ *and* $\kappa_r \neq 0$.

The proof of this result is based on a Walsh expansion of the Warnock-type formula (6) and on the orthogonality properties of Walsh functions. A proof for the unweighted case can also be found in [33, Section 16.5]. An estimate of the sums involved in the above formula yields the following result.

Corollary 1 (Cristea, D., and P. [14, Theorem 2]). *Let $\mathscr{P}$ be a digital (t, m, s)-net over $\mathbb{Z}_b$ with $t < m$. Then the mean square weighted L_2-discrepancy of $\widehat{\mathscr{P}_{\boldsymbol{\sigma}}}$ is bounded by*

$$\mathbb{E}[L_{2,\boldsymbol{\gamma}}^2(\widehat{\mathscr{P}_{\boldsymbol{\sigma}}})] \leq \frac{1}{b^{2m}} \sum_{\emptyset \neq \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}}^2 \left[\frac{1}{6} + b^{2t}\left(\frac{b^2 - b + 3}{6}\right)^{|\mathfrak{u}|} (m-t)^{|\mathfrak{u}|-1}\right].$$

For example, in the unweighted case, i.e., $\boldsymbol{\gamma} = \mathbf{1} = (1, 1, \ldots)$ we obtain

$$\mathbb{E}[L_{2,\mathbf{1}}^2(\widehat{\mathscr{P}_{\boldsymbol{\sigma}}})] \ll_{s,b} \frac{(m-t)^{s-1}}{b^{2(m-t)}}.$$

In particular, for every digital (t, m, s) net $\mathscr{P}$ over $\mathbb{Z}_b$ there exists a simplified digital shift $\boldsymbol{\sigma}^* \in \{0, 1/b^m, \ldots, (b^m - 1)/b^m\}^s$ of depth m such that

$$L_{2,\mathbf{1}}(\widehat{\mathscr{P}_{\boldsymbol{\sigma}^*}}) \ll_{s,b} \frac{(m-t)^{\frac{s-1}{2}}}{b^{m-t}}.$$

According to Roth's lower bound (12), for a $(0, m, s)$-net this bound is best possible in the order of magnitude in m.

Let $M_{b,m}$ be the set of all $m \times m$ matrices with entries over $\mathbb{Z}_b$ and let $\mathcal{C}_{s,b} := \{(C_1, \ldots, C_s) : C_i \in M_{b,m} \text{ for } 1 \leq i \leq s\}$. Let $1/2 < \lambda \leq 1$. Then consider the average

$$A_{b^m,s,\lambda} := \frac{1}{b^{m^2 s}} \sum_{(C_1,\ldots,C_s) \in \mathcal{C}_{s,b}} \left(\mathbb{E}[L_{2,\boldsymbol{\gamma}}^2(\widehat{\mathscr{P}_{\boldsymbol{\sigma}}})]\right)^{\lambda}. \tag{25}$$

Using Theorem 8 we have

$$\begin{aligned} A_{b^m,s,\lambda} \leq{}& \frac{1}{b^{2\lambda m}} \sum_{\emptyset \neq \mathfrak{u} \subseteq [s]} \frac{\boldsymbol{\gamma}_{\mathfrak{u}}^{2\lambda}}{3^{\lambda} 2^{\lambda|\mathfrak{u}|}} \left(1 + \frac{1}{3b^m}\right)^{\lambda|\mathfrak{u}|} \\ &+ \frac{1}{b^{m^2|\mathfrak{u}|}} \sum_{\substack{\emptyset \neq \mathfrak{u} \subseteq [s] \\ \mathfrak{u} = \{u_1,\ldots,u_e\}}} \frac{\boldsymbol{\gamma}_{\mathfrak{u}}^{2\lambda}}{3^{\lambda|\mathfrak{u}|}} \sum_{\substack{k_1,\ldots,k_e=0 \\ (k_1,\ldots,k_e) \neq (0,\ldots,0)}}^{b^m-1} \prod_{i=1}^{e} \psi(k_i)^{\lambda} \sum_{\substack{C_{u_1},\ldots,C_{u_e} \in M_{b,m} \\ C_{u_1}^\top \mathbf{k}_1 + \cdots + C_{u_e}^\top \mathbf{k}_e = \mathbf{0}}} 1 \end{aligned}$$

Let $k_i = \kappa_{i,0} + \kappa_{i,1} b + \ldots + \kappa_{i,m-1} b^{m-1}$ and let $\boldsymbol{c}_{i,j}$ be the jth row vector of the matrix C_i. Since at least one $k_i \neq 0$ it follows that there is a $\kappa_{i,j} \neq 0$. First, assume that $\kappa_{1,0} \neq 0$. Then for any choice of

$$\boldsymbol{c}_{u_1,2}, \ldots, \boldsymbol{c}_{u_1,m}, \boldsymbol{c}_{u_2,1}, \ldots, \boldsymbol{c}_{u_2,m}, \ldots, \boldsymbol{c}_{u_e,1}, \ldots, \boldsymbol{c}_{u_e,m} \in \mathbb{Z}_b^m$$

we can find exactly one vector $\boldsymbol{c}_{u_1,1} \in \mathbb{Z}_b^m$ such that $C_{u_1}^\top \mathbf{k}_1 + \cdots + C_{u_e}^\top \mathbf{k}_e = \mathbf{0}$ is fulfilled. The same argument holds with $\kappa_{1,0}$ replaced by $\kappa_{i,j}$ and $\boldsymbol{c}_{u_1,1}$ replaced by $\boldsymbol{c}_{u_i,j+1}$. Therefore, we have

$$\sum_{\substack{C_{u_1},\ldots,C_{u_e} \in M_{b,m} \\ C_{u_1}^\top \mathbf{k}_1 + \cdots + C_{u_e}^\top \mathbf{k}_e = \mathbf{0}}} 1 = b^{m^2|\mathfrak{u}|-m}$$

and hence, after some elementary algebra, we obtain for $1/2 < \lambda \le 1$,

$$\begin{aligned} A_{b^m,s} &\le \frac{1}{b^{2\lambda m}} \sum_{\emptyset \ne \mathfrak{u} \subseteq [s]} \frac{\boldsymbol{\gamma}_{\mathfrak{u}}^{2\lambda}}{3^\lambda} \left(\frac{2}{3}\right)^{\lambda|\mathfrak{u}|} + \frac{1}{b^m} \sum_{\emptyset \ne \mathfrak{u} \subseteq [s]} \frac{\boldsymbol{\gamma}_{\mathfrak{u}}^{2\lambda}}{3^{\lambda|\mathfrak{u}|}} \left(1 + \sum_{k=1}^{b^m-1} \psi(k)^\lambda\right)^{|\mathfrak{u}|} \\ &\le \frac{1}{b^m} \sum_{\emptyset \ne \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}}^{2\lambda} c_{b,\lambda}^{|\mathfrak{u}|} \\ &\le \frac{1}{b^m} \prod_{i=1}^s (1 + c_{b,\lambda} \gamma_i^{2\lambda}), \end{aligned}$$

for some $c_{b,\lambda} > 0$.

Thus, for $1/2 < \lambda \le 1$, there exists a sequence $(\widehat{\mathscr{P}}_{\boldsymbol{\sigma}^*})_{m \ge 1}$ of simplified digitally shifted digital nets over $\mathbb{Z}_b$ for which we have

$$L_{2,\boldsymbol{\gamma}}(\widehat{\mathscr{P}}_{\boldsymbol{\sigma}^*}) \le \frac{1}{b^{m/(2\lambda)}} \prod_{i=1}^s (1 + c_{b,\lambda} \gamma_i^{2\lambda})^{1/(2\lambda)} \quad \text{for all} \quad m \in \mathbb{N}.$$

If $\sum_{i \ge 1} \gamma_i^{2\lambda} < \infty$, then we obtain

$$\begin{aligned} \frac{\mathrm{disc}_{2,\boldsymbol{\gamma}}(b^m, s)}{\mathrm{disc}_{2,\boldsymbol{\gamma}}(0, s)} &\le \frac{L_{2,\boldsymbol{\gamma}}(\mathscr{P}_{\boldsymbol{\sigma}_m^*})}{\mathrm{disc}_{2,\boldsymbol{\gamma}}(0,1)} \\ &\le \frac{1}{b^{m/(2\lambda)}} \exp\left(\frac{1}{2\lambda} \sum_{i=1}^s \log(1 + c_{b,\lambda} \gamma_i^{2\lambda})\right) \mathrm{disc}_{2,\boldsymbol{\gamma}}(0,1)^{-1} \\ &\le \frac{1}{b^{m/(2\lambda)}} \exp\left(\frac{c_{b,\lambda}}{2\lambda} \sum_{i=1}^\infty \gamma_i^{2\lambda}\right) \mathrm{disc}_{2,\boldsymbol{\gamma}}(0,1)^{-1} =: \frac{C_{b,\lambda,\boldsymbol{\gamma}}}{b^{m/(2\lambda)}}, \end{aligned}$$

and this bound is independent of the dimension s. For $\varepsilon > 0$ choose $m \in \mathbb{N}$ such that $b^{m-1} < \lceil (C_{b,\lambda,\boldsymbol{\gamma}} \varepsilon^{-1})^{2\lambda} \rceil =: N \le b^m$. Then we have

$$\frac{\mathrm{disc}_{2,\boldsymbol{\gamma}}(b^m, s)}{\mathrm{disc}_{2,\boldsymbol{\gamma}}(0, s)} \le \varepsilon$$

and hence

$$N_{2,\boldsymbol{\gamma}}(\varepsilon, s) \le b^m < bN = b\lceil (C_{b,\lambda,\boldsymbol{\gamma}} \varepsilon^{-1})^{2\lambda} \rceil.$$

This means that the weighted L_2-discrepancy is strongly tractable with ε-exponent at most 2λ whenever $\sum_{i \ge 1} \gamma_i^{2\lambda} < \infty$ for some $\lambda \in (1/2, 1]$, and the corresponding bounds can be achieved with digitally shifted digital nets.

4 Lattice rules

In this section we present another construction method for low-discrepancy point sets in $[0,1]^s$ which can be used for QMC algorithms. In the following we write $\{x\} = x - \lfloor x \rfloor$ for the fractional part of a nonnegative real number. For vectors $\boldsymbol{x} = (x_1, \ldots, x_s) \in \mathbb{R}^s$ we set $\{\boldsymbol{x}\} := (\{x_1\}, \ldots, \{x_s\})$.

Definition 4. *For an integer $N \geq 2$ and for $\boldsymbol{g} \in \mathbb{Z}^s$ the point set $\mathscr{P}(\boldsymbol{g}, N)$ consisting of the N elements*

$$\boldsymbol{x}_n = \left\{ \frac{n}{N} \boldsymbol{g} \right\} \quad \text{for all } 0 \leq n < N$$

is called a lattice point set. A QMC rule using $\mathscr{P}(\boldsymbol{g}, N)$ as underlying node set is called a lattice rule. Hence a lattice rule is of the form

$$Q_{N,\boldsymbol{g}}(f) = \frac{1}{N} \sum_{n=0}^{N-1} f\left(\left\{ \frac{n}{N} \boldsymbol{g} \right\} \right).$$

Lattice point sets were introduced independently by Korobov [57] and Hlawka [52]. A detailed treatise can be found in the books of Niederreiter [78] and of Sloan and Joe [103].

An important property of a lattice point set $\mathscr{P}(\boldsymbol{g}, N)$ is that for all $\boldsymbol{h} \in \mathbb{Z}^s$ we have

$$\sum_{n=0}^{N-1} \exp\left(2\pi \mathtt{i} \frac{n}{N} \boldsymbol{g} \cdot \boldsymbol{h} \right) = \begin{cases} N & \text{if } \boldsymbol{g} \cdot \boldsymbol{h} = 0 \pmod{N}, \\ 0 & \text{otherwise} \end{cases} \tag{26}$$

This property motivates the following definition.

Definition 5. *The dual lattice of the lattice point set $\mathscr{P}(\boldsymbol{g}, N)$ from Definition 4 is defined as*

$$\mathcal{L}_{\boldsymbol{g},N} = \{ \boldsymbol{h} \in \mathbb{Z}^s \, : \, \boldsymbol{h} \cdot \boldsymbol{g} \equiv 0 \pmod{N} \}.$$

Property (26) is the reason why it is most convenient to consider one-periodic functions for the analysis of the integration error of lattice rules. This analysis can again be described in terms of a reproducing kernel as explained in Section 1.

4.1 The worst-case error of lattice rules in weighted Korobov spaces

We consider a reproducing kernel of the form

$$K_{\mathrm{Kor}}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{\boldsymbol{h} \in \mathbb{Z}^s} \omega_{\boldsymbol{h}} \exp(2\pi \mathtt{i} \boldsymbol{h} \cdot (\boldsymbol{x} - \boldsymbol{y}))$$

for all $\boldsymbol{x}$ and $\boldsymbol{y}$ in $[0,1]^s$ with some weights $\omega_{\boldsymbol{h}} \in \mathbb{R}^+$ for all $\boldsymbol{h} \in \mathbb{Z}^s$ such that $\sum_{\boldsymbol{h} \in \mathbb{Z}^s} \omega_{\boldsymbol{h}} < \infty$, which may also depend on other parameters. This choice guarantees that the kernel is well defined, since

$$|K_{\mathrm{Kor}}(\boldsymbol{x}, \boldsymbol{y})| \leq K_{\mathrm{Kor}}(\boldsymbol{x}, \boldsymbol{x}) = \sum_{\boldsymbol{h} \in \mathbb{Z}^s} \omega_{\boldsymbol{h}} < \infty.$$

Obviously, the function $K_{\mathrm{Kor}}(\boldsymbol{x}, \boldsymbol{y})$ is symmetric in $\boldsymbol{x}$ and $\boldsymbol{y}$ and it is easy to show that it is also positive definite. Therefore, $K_{\mathrm{Kor}}(\boldsymbol{x}, \boldsymbol{y})$ is indeed a reproducing kernel.

Common examples for $\omega_{\boldsymbol{h}}$ are the following ones:

- $\omega_{\boldsymbol{h}} = r_{\alpha,\boldsymbol{\gamma}}(\boldsymbol{h})^{-1}$ where $\alpha > 1$ is a real, $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots)$ is a sequence of positive reals, for $\boldsymbol{h} = (h_1, \ldots, h_s)$ we put $r_{\alpha,\boldsymbol{\gamma}}(\boldsymbol{h}) = \prod_{i=1}^{s} r_{\alpha,\gamma_i}(h_i)$, and for $h \in \mathbb{Z}$ and $\gamma > 0$ we put
$$r_{\alpha,\gamma}(h) = \begin{cases} 1 & \text{if } h = 0, \\ \gamma^{-1}|h|^{\alpha} & \text{if } h \neq 0. \end{cases}$$
In this case we will write $K_{\mathrm{Kor}} = K_{s,\alpha,\boldsymbol{\gamma}}$. In the unweighted case, i.e., $\boldsymbol{\gamma} = (1, 1, \ldots)$, we simply write r_α instead of $r_{\alpha,\boldsymbol{\gamma}}$ and $K_{\mathrm{Kor}} = K_{s,\alpha}$.
- $\omega_{\boldsymbol{h}} = \omega^{\|\boldsymbol{h}\|_1}$ with some $\omega \in (0, 1)$, where $\|\boldsymbol{h}\|_1 := |h_1| + \cdots + |h_s|$ for $\boldsymbol{h} = (h_1, \ldots, h_s)$. In this case we will write $K_{\mathrm{Kor}} = K'_{s,\omega}$.

Associated with this reproducing kernel is now the Hilbert space $\mathcal{H}(K_{\mathrm{Kor}})$ of functions $f : [0, 1]^s \to \mathbb{R}$ which are one-periodic in each variable. The corresponding inner product is given by

$$\langle f, g\rangle_{\mathcal{H}(K_{\mathrm{Kor}})} = \sum_{\boldsymbol{h}\in\mathbb{Z}^s} \omega_{\boldsymbol{h}}^{-1} \widehat{f}(\boldsymbol{h})\overline{\widehat{g}(\boldsymbol{h})}, \tag{27}$$

where $\widehat{f}(\boldsymbol{h}) = \int_{[0,1]^s} f(\boldsymbol{x}) \exp(-2\pi \mathtt{i} \boldsymbol{h} \cdot \boldsymbol{x})\, \mathrm{d}\boldsymbol{x}$ is the $\boldsymbol{h}$th Fourier coefficient of f. As usual, the norm in $\mathcal{H}(K_{\mathrm{Kor}})$ is defined by

$$\| \cdot \|^2_{\mathcal{H}(K_{\mathrm{Kor}})} = \langle f, f\rangle_{\mathcal{H}(K_{\mathrm{Kor}})} = \sum_{\boldsymbol{h}\in\mathbb{Z}^s} \omega_{\boldsymbol{h}}^{-1} |\widehat{f}(\boldsymbol{h})|^2.$$

The function space $\mathcal{H}(K_{\mathrm{Kor}})$ is called a *Korobov space*.

Using the approach from Section 1 it follows that the worst-case error for integration in $\mathcal{H}(K_{\mathrm{Kor}})$ using a quasi-Monte Carlo rule based on a point set $\mathscr{P} = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}\}$ is given by

$$\begin{aligned} e^2(\mathcal{H}(K_{\mathrm{Kor}}); \mathscr{P}) &= \sup_{\substack{f\in\mathcal{H}(K_{\mathrm{Kor}}) \\ \|f\|_{\mathcal{H}(K_{\mathrm{Kor}})}\le 1}} e^2(f; \mathscr{P}) \\ &= \int_{[0,1]^s} \int_{[0,1]^s} K_{\mathrm{Kor}}(\boldsymbol{x}, \boldsymbol{y})\, \mathrm{d}\boldsymbol{x}\, \mathrm{d}\boldsymbol{y} \\ &\quad - \frac{2}{N} \sum_{n=0}^{N-1} \int_{[0,1]^s} K_{\mathrm{Kor}}(\boldsymbol{x}, \boldsymbol{x}_n)\, \mathrm{d}\boldsymbol{x} + \frac{1}{N^2} \sum_{n,m=1}^{N} K_{\mathrm{Kor}}(\boldsymbol{x}_n, \boldsymbol{x}_m) \\ &= \sum_{\boldsymbol{h}\in\mathbb{Z}^s\setminus\{\boldsymbol{0}\}} \omega_{\boldsymbol{h}} \left| \frac{1}{N} \sum_{n=0}^{N-1} \exp(2\pi \mathtt{i} \boldsymbol{h} \cdot \boldsymbol{x}_n) \right|^2. \end{aligned}$$

If the QMC rule is a lattice rule, then, using (26) and the notation from Definition 5, we obtain the following simplified expression for the worst-case error.

Theorem 9. *The worst-case error of a lattice rule for integration in the Korobov space $\mathcal{H}(K_{\mathrm{Kor}})$ is given by*

$$e^2(\mathcal{H}(K_{\mathrm{Kor}}); \mathscr{P}(\boldsymbol{g}, N)) = \sum_{\boldsymbol{h}\in\mathcal{L}_{\boldsymbol{g},N}\setminus\{\boldsymbol{0}\}} \omega_{\boldsymbol{h}}.$$

Remark 1. If $\alpha \geq 2$ is an even integer, then the Bernoulli polynomial B_α of degree α has the Fourier expansion

$$B_\alpha(x) = \frac{(-1)^{(\alpha+2)/2}\alpha!}{(2\pi)^\alpha} \sum_{\substack{h\in\mathbb{Z}\\ h\neq 0}} \frac{\exp(2\pi \mathtt{i} hx)}{|h|^\alpha} \quad \text{for all } x \in [0,1);$$

see for example Sloan and Joe [103, Appendix C]. Hence in this case we obtain

$$\begin{aligned} &e^2(\mathcal{H}(K_{s,\alpha,\boldsymbol{\gamma}}); \mathscr{P}(\boldsymbol{g}, N)) \\ &= -1 + \frac{1}{N}\sum_{k=0}^{N-1}\prod_{i=1}^{s}\left(1 + \gamma_i \frac{(-1)^{(\alpha+2)/2}(2\pi)^\alpha}{\alpha!} B_\alpha\left(\left\{\frac{kg_i}{N}\right\}\right)\right), \end{aligned}$$

so that $e^2(\mathcal{H}(K_{s,\alpha,\boldsymbol{\gamma}}); \mathscr{P}(\boldsymbol{g}, N))$ can be calculated in $O(Ns)$ operations.

We present the following lower bound on the integration error for numerical integration in the Korobov space. We prove this lower bound for quadrature rules of the form

$$Q_{\mathscr{P},\boldsymbol{w}}(f) = \sum_{n=0}^{N-1} w_n f(\boldsymbol{x}_n),$$

where $\mathscr{P} = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}\} \subseteq [0,1]^s$ and $\boldsymbol{w} = (w_0, \ldots, w_{N-1}) \in \mathbb{R}^N$ are given. The result holds even for more general quadrature rules not considered here, see Bakhvalov [4]. Note that a QMC rule is obtained by choosing $\boldsymbol{w} = (N^{-1}, \ldots, N^{-1}) \in \mathbb{R}^N$. In this case we write $e(\mathcal{H}(K_{\mathrm{Kor}}); \mathscr{P}(\boldsymbol{g}, N); \boldsymbol{w})$ for the worst-case error of the quadrature rule $Q_{\mathscr{P},\boldsymbol{w}}$ in the Korobov space.

Theorem 10. *Let $\mathscr{P}$ be an arbitrary N-element point set in $[0,1]^s$ and let $\boldsymbol{w} = (w_0, \ldots, w_{N-1}) \in \mathbb{R}^N$.*

1. *(Bakhvalov [4], Temlyakov [111]) If $K_{\mathrm{Kor}} = K_{s,\alpha}$ for an $\alpha > 1$, then*

$$e(\mathcal{H}(K_{s,\alpha}); \mathscr{P}; \boldsymbol{w}) \geq C(s,\alpha)\frac{(\log N)^{(s-1)/2}}{N^{\alpha/2}},$$

   where $C(s,\alpha) > 0$ depends on α and s, but not on N and $\boldsymbol{w}$.

2. *(Šarygin [97]) If $K_{\mathrm{Kor}} = K'_{s,\omega}$ for an $\omega \in (0,1)$, then*

$$e(\mathcal{H}(K'_{s,\omega}); \mathscr{P}; \boldsymbol{w}) \geq \left(1 + \frac{2}{1-\omega}\right)^{-s/2} \omega^{(s!(N+1))^{1/s}}.$$

Proof. 1. We follow the proof of Temlyakov [111, Lemma 3.1]. In the same way as was shown above, the worst-case error in the Korobov space for an arbitrary quadrature rule can be written as

$$e^2(\mathcal{H}(K_{s,\alpha}); \mathscr{P}; \boldsymbol{w}) = |1-\beta|^2 + \sum_{\boldsymbol{h}\in\mathbb{Z}^s\setminus\{\boldsymbol{0}\}} r_\alpha^{-1}(\boldsymbol{h}) \left|\sum_{n=0}^{N-1} w_n \exp(2\pi \mathtt{i} \boldsymbol{h}\cdot\boldsymbol{x}_n)\right|^2,$$

where $\beta = \sum_{n=0}^{N-1} w_n$.

Let $f : \mathbb{R} \to \mathbb{R}$ be an infinitely times differentiable function such that $f(x) > 0$ for $x \in (0,1)$ and $f^{(r)}(x) = 0$ for $x \in \mathbb{R} \setminus (0,1)$ for all $0 \leq r \leq a := \lceil \alpha/2 \rceil + 1$. For instance, let

$$f(x) = \begin{cases} x^{a+1}(1-x)^{a+1} & \text{for } x \in (0,1) \\ 0 & \text{otherwise.} \end{cases} \tag{28}$$

For $m \in \mathbb{N}_0$ let $f_m(x) = f(2^{m+2}x)$ and for $\boldsymbol{m} = (m_1, \ldots, m_s) \in \mathbb{N}_0^s$ let

$$f_{\boldsymbol{m}}(\boldsymbol{x}) = \prod_{i=1}^s f_{m_i}(x_i),$$

where $\boldsymbol{x} = (x_1, \ldots, x_s)$. Let $\|\boldsymbol{m}\|_1 = m_1 + \cdots + m_s$. We obtain

$$\widehat{f}_{\boldsymbol{m}}(\boldsymbol{0}) = \prod_{i=1}^s \int_0^1 f(2^{m_i+2}x)\,\mathrm{d}x = \prod_{i=1}^s \frac{1}{2^{m_i+2}} \int_0^1 f(y)\,\mathrm{d}y = \frac{1}{2^{\|\boldsymbol{m}\|_1+2s}} I^s(f),$$

where $I(f) = \int_0^1 f(y)\,\mathrm{d}y$. For instance, by choosing f according to (28) we obtain

$$I(f) = B(a+2, a+2) = \frac{((a+1)!)^2}{(2a+3)!},$$

where B denotes the beta function.
Let t be such that

$$2N \le 2^t < 4N.$$

Let

$$B_{\boldsymbol{m}} = \left\{ \boldsymbol{y} \in [0,1]^s : \sum_{n=0}^{N-1} w_n f_{\boldsymbol{m}}(\boldsymbol{x}_n - \boldsymbol{y}) = 0 \right\}.$$

Notice that the support of $f_{\boldsymbol{m}}(\boldsymbol{x}_n - \boldsymbol{y})$ (as a function of $\boldsymbol{y}$) is contained in the interval $\prod_{i=1}^s (x_{i,n} - 2^{-m_i-2}, x_{i,n})$ and hence the support of $F(\boldsymbol{y}) = \frac{1}{N}\sum_{n=0}^{N-1} f_{\boldsymbol{m}}(\boldsymbol{x}_n - \boldsymbol{y})$ is contained in $\bigcup_{n=0}^{N-1} \prod_{i=1}^s (x_{i,n} - 2^{-m_i-2}, x_{i,n})$. Therefore the area of the support of F is at most $N2^{-\|\boldsymbol{m}\|_1}$. Thus for all $\boldsymbol{m}$ such that $\|\boldsymbol{m}\|_1 = t$ we have

$$\lambda_s(B_{\boldsymbol{m}}) \ge 1 - N2^{-\|\boldsymbol{m}\|_1} = 1 - N2^{-t} > 1/4,$$

where λ_s denotes the s dimensional Lebesgue measure.
Let $\beta = \sum_{n=0}^{N-1} w_n$. Thus we have

$$\begin{aligned} \lambda_s(B_{\boldsymbol{m}})|\widehat{f}_{\boldsymbol{m}}(\boldsymbol{0})|^2|\beta|^2 &= \int_{B_{\boldsymbol{m}}} |Q_{\boldsymbol{w}}(f_{\boldsymbol{m}}(\cdot - \boldsymbol{y})) - \widehat{f}_{\boldsymbol{m}}(\boldsymbol{0})\beta|^2\,\mathrm{d}\boldsymbol{y} \\ &\le \int_{[0,1]^s} |Q_{\boldsymbol{w}}(f_{\boldsymbol{m}}(\cdot - \boldsymbol{y})) - \widehat{f}_{\boldsymbol{m}}(\boldsymbol{0})\beta|^2\,\mathrm{d}\boldsymbol{y} \\ &= \sum_{\boldsymbol{h} \in \mathbb{Z}^s \setminus \{\boldsymbol{0}\}} |\widehat{f}_{\boldsymbol{m}}(\boldsymbol{h})|^2 \left| \sum_{n=0}^{N-1} w_n \exp(2\pi \mathrm{i} \boldsymbol{h} \cdot \boldsymbol{x}_n) \right|^2. \end{aligned}$$

We have

$$\begin{aligned} \widehat{f}_m(h) &= \int_0^1 f(2^{m+2}x) \exp(-2\pi \mathrm{i} h x)\,\mathrm{d}x \\ &= \frac{1}{2^{m+2}} \int_0^1 f(y) \exp(-2\pi \mathrm{i} h 2^{-m-2} y)\,\mathrm{d}y = \frac{1}{2^{m+2}} \widehat{f}(h 2^{-m-2}), \end{aligned}$$

where $\widehat{f}$ denotes the Fourier transform of f. Since, by assumption, f is infinitely times differentiable, integration by parts shows that for any $m \in \mathbb{N}_0$ we have

$$|\widehat{f}_m(h)| \le 2^{-m-2}|\widehat{f}(h2^{-m-2})| \le C_a 2^{-m-2}\min(1,(h2^{-m-2})^{-a}),$$

where the constant $C_a > 0$ depends only on a and f. Then for $\boldsymbol{m}$ with $\|\boldsymbol{m}\|_1 = t$ we have

$$\begin{aligned}|\widehat{f}_{\boldsymbol{m}}(\boldsymbol{h})| &\le C(a,s)\prod_{i=1}^{s}(2^{-m_i}\min(1,2^{am_i}r_a^{-1}(h_i)))\\ &= C(a,s)2^{(\alpha/2-1)t}\prod_{i=1}^{s}(2^{-\alpha m_i/2}\min(1,2^{am_i}r_a^{-1}(h_i))).\end{aligned}$$

By summing over all choices of $\boldsymbol{m}$ with $\|\boldsymbol{m}\|_1 = t$ we obtain

$$\begin{aligned}\sum_{\substack{\boldsymbol{m}\in\mathbb{N}_0^s\\ \|\boldsymbol{m}\|_1=t}}|\widehat{f}_{\boldsymbol{m}}(\boldsymbol{h})|^2 &\le 2^{(\alpha-2)t}C^2(a,s)\sum_{\substack{\boldsymbol{m}\in\mathbb{N}_0^s\\ \|\boldsymbol{m}\|_1=t}}\prod_{i=1}^{s}(2^{-\alpha m_i}\min(1,2^{2am_i}r_a^{-2}(h_i)))\\ &\le 2^{(\alpha-2)t}C^2(a,s)\prod_{i=1}^{s}\sum_{m=0}^{\infty}2^{-\alpha m}\min(1,2^{2am}r_a^{-2}(h_i)).\end{aligned}$$

The last sum can now be estimated by

$$\begin{aligned}&\sum_{m=0}^{\infty}\frac{1}{2^{\alpha m}}\min(1,2^{2am}r_a^{-2}(h_i))\\ &=\sum_{0\le m\le(\log_2 r_a(h_i))/a}2^{(2a-\alpha)m}r_a^{-2}(h_i)+\sum_{m>(\log_2 r_a(h_i))/a}2^{-\alpha m}\\ &\le\frac{r_{2a-\alpha}(h_i)2^{2a-\alpha}-1}{2^{2a-\alpha}-1}r_{2a}^{-1}(h_i)+\frac{r_\alpha^{-1}(h_i)2^\alpha}{2^\alpha-1}\\ &\le r_\alpha^{-1}(h_i)\left(1+\frac{2^\alpha}{2^\alpha-1}\right)\le 3r_\alpha^{-1}(h_i).\end{aligned}$$

Thus we have

$$C_1(a,s)2^{-(\alpha-2)t}\sum_{\substack{\boldsymbol{m}\in\mathbb{N}_0^s\\ \|\boldsymbol{m}\|_1=t}}|\widehat{f}_{\boldsymbol{m}}(\boldsymbol{h})|^2 \le r_\alpha^{-1}(\boldsymbol{h}).$$

We obtain

$$\begin{aligned}&e^2(\mathcal{H}(K_{s,\alpha});\mathscr{P};\boldsymbol{w}) = |1-\beta|^2+\sum_{\boldsymbol{h}\in\mathbb{Z}^s\setminus\{\boldsymbol{0}\}}r_\alpha^{-1}(\boldsymbol{h})\left|\sum_{n=0}^{N-1}w_n\exp(2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x}_n)\right|^2\\ &\ge|1-\beta|^2+C_1(a,s)\frac{1}{2^{(\alpha-2)t}}\sum_{\substack{\boldsymbol{m}\in\mathbb{N}_0^s\\ \|\boldsymbol{m}\|_1=t}}\sum_{\boldsymbol{h}\in\mathbb{Z}^s\setminus\{\boldsymbol{0}\}}|\widehat{f}_{\boldsymbol{m}}(\boldsymbol{h})|^2\left|\sum_{n=0}^{N-1}w_n\exp(2\pi\mathrm{i}\boldsymbol{h}\cdot\boldsymbol{x}_n)\right|^2\\ &\ge|1-\beta|^2+C_1(a,s)\frac{1}{2^{(\alpha-2)t}}\sum_{\substack{\boldsymbol{m}\in\mathbb{N}_0^s\\ \|\boldsymbol{m}\|_1=t}}\lambda_s(B_{\boldsymbol{m}})|\widehat{f}_{\boldsymbol{m}}(\boldsymbol{0})|^2|\beta|^2\end{aligned}$$

$$\geq |1-\beta|^2 + C_2(a,s)|\beta|^2 \frac{2^{2t}}{N^\alpha} \sum_{\substack{\boldsymbol{m}\in\mathbb{N}_0^s \\ \|\boldsymbol{m}\|_1=t}} 2^{-2t-4s} I^{2s}(f)$$

$$\geq |1-\beta|^2 + C_3(a,s)|\beta|^2 N^{-\alpha} \binom{t+s-1}{s-1}.$$

Set $A = C_3(a,s)N^{-\alpha}\binom{t+s-1}{s-1}$. Then the last expression can be written as $|1-\beta|^2 + A|\beta|^2$, which satisfies

$$e^2(\mathcal{H}(K_{s,\alpha});\mathscr{P};\boldsymbol{w}) \geq |1-\beta|^2 + A|\beta|^2 \geq \frac{A}{1+A} \geq C_4(a,s)N^{-\alpha}\binom{t+s-1}{s-1},$$

which implies the result, since $t \geq \log_2(N)$.

2. The lower bound in the case that $K_{\mathrm{Kor}} = K'_{s,\omega}$ for some $\omega \in (0,1)$ can be deduced from the following result due to Šarygin [97]: For any point set $\mathscr{P} = \{\boldsymbol{x}_0,\ldots,\boldsymbol{x}_{N-1}\}$ in $[0,1)^s$ one can find a periodic function $f : [0,1]^s \to \mathbb{R}$ with the following properties:
   - the Fourier coefficients of f satisfy $|\widehat{f}(\boldsymbol{h})| \leq \omega^{|\boldsymbol{h}|}$;
   - $f(\boldsymbol{x}_n) = 0$ for all $0 \leq n < N$;
   - $\int_{[0,1]^s} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \geq \omega^{(s!(N+1))^{1/s}}$.

   Then the function $g(\boldsymbol{x}) = (1+\frac{2}{1-\omega})^{-s/2} f(\boldsymbol{x})$ is contained in the unit ball of the Korobov space $\mathcal{H}(K'_{s,\omega})$ and hence the result follows from the properties of the function f.□

Remark 2. For the worst-case error of lattice rules for integration in the Korobov spaces $\mathcal{H}(K_{s,\alpha,\boldsymbol{\gamma}})$ and $\mathcal{H}(K'_{s,\omega})$ we have the following lower bounds:

For any integer $N \geq 2$ and $\boldsymbol{g} \in \{0,1,\ldots,N-1\}^s$ we have

$$e(\mathcal{H}(K_{s,\alpha,\boldsymbol{\gamma}});\mathscr{P}(\boldsymbol{g},N)) \geq \left(2\zeta(\alpha)\sum_{i=1}^s \gamma_i\right)^{1/2} \frac{1}{N^{\alpha/2}},$$

where $\zeta(\alpha) = \sum_{h\geq 1} h^{-\alpha}$, and

$$e(\mathcal{H}(K'_{s,\omega});\mathscr{P}(\boldsymbol{g},N)) \geq \omega^{\frac{1}{2}(s!N)^{1/s}}.$$

Proof. We have

$$e^2(\mathcal{H}(K_{s,\alpha,\boldsymbol{\gamma}});\mathscr{P}(\boldsymbol{g},N)) = \sum_{\boldsymbol{h}\in\mathcal{L}_{\boldsymbol{g},N}\setminus\{\boldsymbol{0}\}} \frac{1}{r_{\alpha,\boldsymbol{\gamma}}(\boldsymbol{h})} \geq \sum_{i=1}^s \sum_{\substack{h_i\in\mathbb{Z}\setminus\{0\} \\ h_i g_i \equiv 0 \pmod N}} \frac{\gamma_i}{|h_i|^\alpha}$$

$$\geq 2\sum_{i=1}^s \gamma_i \sum_{h=1}^\infty \frac{1}{(Nh)^\alpha} = 2\zeta(\alpha)\left(\sum_{i=1}^s \gamma_i\right)\frac{1}{N^\alpha}.$$

Define $\rho(\boldsymbol{g}) = \min_{\boldsymbol{h}\in\mathcal{L}_{\boldsymbol{g},N}\setminus\{\boldsymbol{0}\}} \|\boldsymbol{h}\|_1$. Then we have

$$e^2(\mathcal{H}(K'_{s,\omega});\mathscr{P}(\boldsymbol{g},N)) = \sum_{\boldsymbol{h}\in\mathcal{L}_{\boldsymbol{g},N}\setminus\{\boldsymbol{0}\}} \omega^{\|\boldsymbol{h}\|_1} = \sum_{k=\rho(\boldsymbol{g})}^\infty \omega^k \sum_{\substack{\boldsymbol{h}\in\mathcal{L}_{\boldsymbol{g},N}\setminus\{\boldsymbol{0}\} \\ \|\boldsymbol{h}\|_1=k}} 1 \geq \omega^{\rho(\boldsymbol{h})}.$$

In Lyness [70, Section 5] it is shown that for any $N \in \mathbb{N}$ and any $\boldsymbol{g} \in \{0, 1, \ldots, N-1\}^s$ we have $\rho(\boldsymbol{g}) \le (s!N)^{1/s}$. Hence we have

$$e^2(\mathcal{H}(K'_{s,\omega}); \mathscr{P}(\boldsymbol{g}, N)) \ge \omega^{(s!N)^{1/s}}. \square$$

Remark 3. The quantity $\rho(\boldsymbol{g})$ used above is the enhanced trigonometric degree of a lattice rule Cools and Lyness [13] and Lyness [70]. A cubature rule of enhanced trigonometric degree δ is one that integrates all trigonometric polynomials of degree less then δ exactly.

There are also existence results for lattice point sets. In the case of $K_{\mathrm{Kor}} = K_{s,\alpha,\boldsymbol{\gamma}}$ these results are mainly based on averaging arguments over all lattice points $\boldsymbol{g} \in \{0, \ldots, N-1\}^s$ for given $N \ge 2$. These methods are nowadays quite standard and there are even constructions of such lattice points based on a component-by-component approach. Here one constructs successive the components of $\boldsymbol{g}$. This approach was introduced by Korobov [58] and later re-invented by Sloan and Reztsov [106]. In the following let $\mathbb{P} = \{2, 3, 5, \ldots\}$ denote the set of prime numbers.

Algorithm 1 *Let $N \in \mathbb{P}$ and let $s \ge 2$.*

1. *Choose $g_1 = 1$.*
2. *For $d > 1$, assume we have already constructed $g_1, \ldots, g_{d-1}$. Then find $g_d \in \{1, \ldots, N-1\}$ which minimises $e^2(\mathcal{H}(K_{s,\alpha,\boldsymbol{\gamma}}); \mathscr{P}((g_1, \ldots, g_{d-1}, z), N))$ as a function of $z \in \{1, \ldots, N-1\}$.*

Variations of this algorithm for shifted and randomly shifted lattice rules were analysed by Sloan, Kuo and Joe [104, 105]. The case where N is not necessarily a prime number was first considered in [16].

A straightforward implementation of Algorithm 1 would require $O(N^2 s^2)$ operations for the construction of a lattice point set $\mathscr{P}(\boldsymbol{g}, N)$ in dimension s. Using the so-called *fast component-by-component algorithm* due to Nuyens and Cools the construction costs can be reduced to $O(sN \log N)$ operations; see Nuyens and Cools [86, 87, 88] and the references therein for more detailed information.

In the case of $K_{\mathrm{Kor}} = K'_{s,\omega}$ there is, until now, only one existence result for lattice rules. Note, however, that a (modified) regular grid can be used to obtain an exponential rate of convergence with polynomial tractability; see D., Larcher, P. and Woźniakowski [27]. Since a regular grid can be obtained using the digital construction scheme, it also follows that there are digital nets for which one can obtain an exponential rate of convergence for integrands from $\mathcal{H}(K'_{s,\omega})$.

Theorem 11 (Kuo [62] and D., Larcher, P., and Woźniakowski [27]).

1. *For any prime number N, a vector $\boldsymbol{g} \in \{0, \ldots, N-1\}^s$ can be found using Algorithm 1 such that*

$$e^2(\mathcal{H}(K_{s,\alpha,\boldsymbol{\gamma}}); \mathscr{P}(\boldsymbol{g}, N)) \le \frac{2^{1/\lambda}}{N^{1/\lambda}} \prod_{i=1}^{s} (1 + 2\gamma_i^\lambda \zeta(\alpha\lambda))^{1/\lambda}$$

for $1/\alpha < \lambda \le 1$.

2. For any prime number N, there exists a $\boldsymbol{g} \in \{0, \ldots, N-1\}^s$ such that

$$e^2(\mathcal{H}(K'_{s,\omega}); \mathscr{P}(\boldsymbol{g}, N)) \le \omega^{2^{-1}(s!\,N)^{1/s}} \left(\frac{4\mathrm{e}}{\omega - \omega^2}\right)^s N.$$

Proof. 1. A proof of this result can be found in Kuo [62].
2. We have

$$\begin{aligned} e^2(\mathcal{H}(K'_{s,\omega}); \mathscr{P}(\boldsymbol{g}, N)) &= \sum_{\boldsymbol{h} \in \mathcal{L}_{\boldsymbol{g},N} \setminus \{\boldsymbol{0}\}} \omega^{\|\boldsymbol{h}\|_1} \\ &= \sum_{k=\rho(\boldsymbol{g})}^{\infty} \omega^k \sum_{\substack{\boldsymbol{h} \in \mathcal{L}_{\boldsymbol{g},N} \setminus \{\boldsymbol{0}\} \\ \|\boldsymbol{h}\|_1 = k}} 1 \qquad (29) \\ &\le \sum_{k=\rho(\boldsymbol{g})}^{\infty} \omega^k 2^s \binom{k+s-1}{s-1} \\ &\le \omega^{\rho(\boldsymbol{g})} 2^s (1-\omega)^{-s} \binom{\rho(\boldsymbol{g}) + s - 1}{s-1}, \qquad (30) \end{aligned}$$

where we used

$$\sum_{k=\rho}^{\infty} \binom{k+r-1}{r-1} \omega^k \le \omega^{\rho} \binom{\rho+r-1}{r-1} (1-\omega)^{-r}, \tag{31}$$

which can be shown using the binomial theorem; see Matoušek [71, Lemma 2.18] or [31, Lemma 6].
Now we show that for a prime number N, there exists a $\boldsymbol{g} \in \{0, 1, \ldots, N-1\}^s$ such that

$$\rho(\boldsymbol{g}) \ge \lceil 2^{-1}(s!\,N)^{1/s} \rceil - s. \tag{32}$$

For a given $\boldsymbol{h} = (h_1, \ldots, h_s) \in \mathbb{Z}^s \setminus \{\boldsymbol{0}\}$ with $|h_i| < N$ for $1 \le i \le s$, there are N^{s-1} choices of $\boldsymbol{g} \in \{0, 1, \ldots, N-1\}^s$ such that $\boldsymbol{g} \cdot \boldsymbol{h} \equiv 0 \pmod{N}$. Furthermore,

$$|\{\boldsymbol{h} \in \mathbb{Z}^s : \|\boldsymbol{h}\|_1 = \ell\}| \le 2^s \binom{\ell+s-1}{s-1}.$$

Let $\rho < N$ be a given positive integer (note that $\rho(\boldsymbol{g}) < N$ always). Then,

$$|\{\boldsymbol{h} \in \mathbb{Z}^s : \|\boldsymbol{h}\|_1 \le \rho\}| \le 2^s \sum_{\ell=0}^{\rho} \binom{\ell+s-1}{s-1} = 2^s \binom{\rho+s}{s}.$$

Therefore,

$$|\{\boldsymbol{g} \in \{0, 1, \ldots, N-1\}^s : \ \rho(\boldsymbol{g}) \le \rho\}| \le N^{s-1} 2^s \binom{\rho+s}{s}.$$

Note that the total number of possible generators $\boldsymbol{g} \in \{0, 1, \ldots, N-1\}^s$ is N^s. Thus, if

$$N^{s-1} 2^s \binom{\rho+s}{s} < N^s, \tag{33}$$

there exists a $\boldsymbol{g} \in \{0,1,\ldots,N-1\}^s$ such that $\rho(\boldsymbol{g}) > \rho$. We estimate

$$2^s \binom{\rho+s}{s} \leq 2^s(\rho+s)^s(s!)^{-1}.$$

This means that (33) is satisfied if $2^s(\rho+s)^s(s!)^{-1} < N$, i.e., for $\rho = \lceil 2^{-1}(s!\,N)^{1/s}\rceil - s - 1$. Hence (32) is shown. Combining (30) and (32) yields the desired result.□

Let us discuss some tractability issues for the Korobov space $K_{s,\alpha,\boldsymbol{\gamma}}$ where $\alpha > 1$. Assume that $\sum_{i\geq 1}\gamma_i^\lambda < \infty$ for some $\lambda \in (1/\alpha, 1]$. Then we obtain from Theorem 11 that

$$\begin{aligned} e(\mathcal{H}(K_{s,\alpha,\boldsymbol{\gamma}}); \mathscr{P}(\boldsymbol{g},N)) &\leq \frac{2^{1/(2\lambda)}}{N^{1/(2\lambda)}} \exp\left(\frac{1}{2\lambda}\sum_{i=1}^{s}\log(1+2\gamma_i^\lambda\zeta(\alpha\lambda))\right) \\ &\leq \frac{2^{1/(2\lambda)}}{N^{1/(2\lambda)}} \exp\left(\frac{\zeta(\alpha\lambda)}{\lambda}\sum_{i\geq 1}\gamma_i^\lambda\right) =: \frac{C_{\alpha,\boldsymbol{\gamma}}}{N^{1/(2\lambda)}} \end{aligned}$$

and this bound is independent of the dimension s. Note that for the initial error we have $e(\mathcal{H}(K_{s,\alpha,\boldsymbol{\gamma}});\emptyset) = 1$; see Sloan and Woźniakowski [108].

For $\varepsilon > 0$ let N be the smallest prime number that is larger or equal $\lceil (C_{\alpha,\boldsymbol{\gamma}}\varepsilon^{-1})^{2\lambda}\rceil =: M$. Then we have $e(\mathcal{H}(K_{s,\alpha,\boldsymbol{\gamma}}); \mathscr{P}(\boldsymbol{g},N)) \leq \varepsilon$ and hence

$$N_{\mathcal{H}(K_{s,\alpha,\boldsymbol{\gamma}})}(\varepsilon,s) \leq N < 2M = 2\lceil (C_{\alpha,\boldsymbol{\gamma}}\varepsilon^{-1})^{2\lambda}\rceil,$$

where we used Bertand's postulate which tells us that $M \leq N < 2M$. Hence multivariate integration in $\mathcal{H}(K_{s,\alpha,\boldsymbol{\gamma}})$ is strongly tractable with ε-exponent at most 2λ whenever $\sum_{i\geq 1}\gamma_i^\lambda < \infty$ for some $\lambda \in (1/\alpha, 1]$. The corresponding bounds can be achieved with lattice point sets. Furthermore, in Sloan and Woźniakowski [108, Theorem 5] it was shown that the condition $\sum_{i\geq 1}\gamma_i < \infty$ is also necessary for strong tractability.

Under weaker assumptions on the weights one can still obtain polynomial tractability. For more results in this direction we refer to Sloan and Woźniakowski [108] or to Novak and Woźniakowski [85, Chapter 16].

A discussion of tractability issues for the Korobov space $\mathcal{H}(K'_{s,\omega})$ can be found in D., Larcher, P., and Woźniakowski [27].

4.2 Star-discrepancy of lattice point sets

For a lattice point set $\mathscr{P}(\boldsymbol{g},N)$ each point $\boldsymbol{x}_n$ is of the form $\boldsymbol{x}_n = \{\boldsymbol{y}_n/N\}$ with $\boldsymbol{y}_n = n\boldsymbol{g} \in \mathbb{Z}^s$. In particular, the elements of a lattice point set have always rational components. For such point sets there is a variant of the inequality of Erdős-Turán-Koksma due to Niederreiter [78, Theorem 3.10], i.e., a general upper bound for the discrepancy in terms of exponential sums. To formulate this result we need some notation.

For an integer $M \geq 2$, let

$$C(M) = (-M/2, M/2] \cap \mathbb{Z}$$

and let $C_s(M)$ be the Cartesian product of s copies of $C(M)$. Furthermore, let

$$C_s^*(M) = C_s(M) \setminus \{\mathbf{0}\}.$$

For $h \in C(M)$ put

$$r(h,M) = \begin{cases} M \sin(\pi |h|/M) & \text{if } h \neq 0, \\ 1 & \text{if } h = 0. \end{cases}$$

For $\boldsymbol{h} = (h_1, \ldots, h_s) \in C_s(M)$, put $r(\boldsymbol{h}, M) = \prod_{i=1}^s r(h_i, M)$.

Proposition 1 (Niederreiter [78, Theorem 3.10]). *Let $M \geq 2$ be an integer and let $\mathscr{P} = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}\}$ be a point set in the s-dimensional unit cube where $\boldsymbol{x}_n$ is of the form $\boldsymbol{x}_n = \{\boldsymbol{y}_n/M\}$ with $\boldsymbol{y}_n \in \mathbb{Z}^s$ for all $0 \leq n < N$. Then we have*

$$D_N^*(\mathscr{P}) \leq 1 - \left(1 - \frac{1}{M}\right)^s + \sum_{\boldsymbol{h} \in C_s^*(M)} \frac{1}{r(\boldsymbol{h}, M)} \left| \frac{1}{N} \sum_{n=0}^{N-1} \exp(2\pi \mathtt{i} \boldsymbol{h} \cdot \boldsymbol{y}_n / M) \right|.$$

Applying this result to a lattice point set $\mathscr{P}(\boldsymbol{g}, N)$, i.e., $M = N$, and using (26) and the fact that for $\boldsymbol{h} \in C_s^*(N)$ we have $r(h, N) \geq 2r(h)$, where $r(h) := r_{1,1}(h) = \max(1, |h|)$, we obtain the following result.

Theorem 12 (Niederreiter [78, Theorem 5.6]). *For the star-discrepancy of a lattice point set $\mathscr{P}(\boldsymbol{g}, N)$ we have*

$$D_N^*(\mathscr{P}(\boldsymbol{g}, N)) \leq 1 - \left(1 - \frac{1}{N}\right)^s + \frac{1}{2} R(\boldsymbol{g}, N) \leq \frac{s}{N} + \frac{1}{2} R(\boldsymbol{g}, N),$$

where

$$R(\boldsymbol{g}, N) := \sum_{\boldsymbol{h} \in C_s^*(N) \cap \mathcal{L}_{\boldsymbol{g},N}} \frac{1}{r(\boldsymbol{h})}.$$

Theorem 12 gives a bound on the star-discrepancy of lattice point sets which is much easier to handle than D_N^* itself (note that the exact computation of the star-discrepancy of a given point set is an NP-hard problem; see Gnewuch, Srivastav, and Winzen [44]). Using (26) again, the quantity $R(\boldsymbol{g}, N)$ can be written as

$$R(\boldsymbol{g}, N) = -1 + \frac{1}{N} \sum_{n=0}^{N-1} \prod_{i=1}^{s} \left(1 + \sum_{h \in C^*(N)} \frac{\exp(2\pi \mathtt{i} h n g_i / N)}{|h|}\right), \tag{34}$$

and hence its calculation requires $O(N^2 s)$ operations. This can be reduced to $O(Ns)$ operations by using an asymptotic expansion; see Joe and Sloan [56].

It has been shown by Larcher [64] that for any dimension $s \geq 2$ there exists some $c_s > 0$ such that for all $N \geq 2$ and all $\boldsymbol{g} \in \mathbb{Z}^s$ we have

$$R(\boldsymbol{g}, N) \geq c_s \frac{(\log N)^s}{N}.$$

On the other hand, it has been shown by Niederreiter [78, Theorem 5.10] that for any integers $s \geq 2$ and $N \geq 2$ we have

$$\frac{1}{|G_s(N)|}\sum_{\boldsymbol{g}\in G_s(N)} R(\boldsymbol{g},N) = \frac{1}{N}(2\log N + c)^s - \frac{2s\log N}{N} + O\left(\frac{(\log\log N)^2}{N}\right), \quad (35)$$

with $c = 2\gamma - 1 + 1 = 0.768\ldots$, where $\gamma = 0.577\ldots$ is the Euler constant and where $G_s(N) = \{\boldsymbol{g} = (g_1,\ldots,g_s) \in C_s(N) : \gcd(g_i,N) = 1 \text{ for } 1 \le i \le s\}$. In particular, we have the following result.

Theorem 13 (Larcher [64] and Niederreiter [78]). *For any integers $s \ge 2$ and $N \ge 2$ there exist $\boldsymbol{g} \in G_s(N)$ such that*

$$R(\boldsymbol{g},N) \asymp_s \frac{(\log N)^s}{N}$$

and this order of magnitude is best possible.

If N is a prime we can use the following component-by-component algorithm for the construction of a 'good' lattice point. For a construction for composite N we refer to Sinescu and Joe [100].

Algorithm 2 *Let $N \in \mathbb{P}$ and let $s \ge 2$.*

1. *Choose $g_1 = 1$.*
2. *For $d > 1$, assume we have already constructed $g_1,\ldots,g_{d-1}$. Then find $g_d \in \{1,\ldots,N-1\}$ which minimises $R_N((g_1,\ldots,g_{d-1},z))$ as a function of $z \in \{1,\ldots,N-1\}$.*

Again, a straightforward implementation of Algorithm 2 would require $O(N^2s^2)$ operations for the construction of a lattice point set $\mathscr{P}(\boldsymbol{g},N)$ in dimension s. Using the so-called *fast component-by-component algorithm* due to Nuyens and Cools the construction costs can be reduced to $O(sN\log N)$ operations; see, again, Nuyens and Cools [86, 87, 88].

The following result shows that Algorithm 2 provides an optimal lattice point with respect to the order of magnitude of $R(\boldsymbol{g},N)$.

Theorem 14 (Joe [54]). *Let $N \in \mathbb{P}$ and suppose that $\boldsymbol{g} = (g_1,\ldots,g_s)$ is constructed according to Algorithm 2. Then for all $1 \le d \le s$ we have*

$$R(\boldsymbol{g}^{(d)},N) \le \frac{1}{N-1}(1+S_N)^d,$$

where $\boldsymbol{g}^{(d)} = (g_1,\ldots,g_d)$ and where $S_N = \sum_{h\in C^*(N)} |h|^{-1}$.

Proof. Since $N \in \mathbb{P}$ it follows that $R(g_1,N) = 0$ for all $g_1 \in \{1,\ldots,N-1\}$. Let $d \ge 1$ and assume that we have

$$R(\boldsymbol{g},N) \le \frac{1}{N-1}(1+S_N)^d,$$

where $\boldsymbol{g} = (g_1,\ldots,g_d)$. Now we consider $(\boldsymbol{g},g_{d+1}) := (g_1,\ldots,g_d,g_{d+1})$.

As g_{d+1} minimises $R((\boldsymbol{g},\cdot),N)$ over $\{1,\ldots,N-1\}$ we obtain

$$R((\boldsymbol{g}, g_{d+1}), N) \le \frac{1}{N-1} \sum_{g_{d+1}=1}^{N-1} \sum_{\substack{(\boldsymbol{h}, h_{d+1}) \in C^*_{d+1}(N) \\ \boldsymbol{h}\cdot\boldsymbol{g} + h_{d+1} g_{d+1} \equiv 0 \pmod N}} \frac{1}{r(\boldsymbol{h})} \frac{1}{r(h_{d+1})}$$
$$= \sum_{(\boldsymbol{h}, h_{d+1}) \in C^*_{d+1}(N)} \frac{1}{r(\boldsymbol{h})} \frac{1}{r(h_{d+1})} \frac{1}{N-1} \sum_{\substack{g_{d+1}=1 \\ h_{d+1} g_{d+1} \equiv -\boldsymbol{h}\cdot\boldsymbol{g} \pmod N}}^{N-1} 1,$$

where we just changed the order of summation. Separating out the term where $h_{d+1} = 0$ we obtain

$$R((\boldsymbol{g}, g_{d+1}), N)$$
$$\le R(\boldsymbol{g}, N) + \sum_{\boldsymbol{h} \in C_d(N)} \frac{1}{r(\boldsymbol{h})} \sum_{h_{d+1} \in C^*(N)} \frac{1}{r(h_{d+1})} \frac{1}{N-1} \sum_{\substack{g_{d+1}=1 \\ h_{d+1} g_{d+1} \equiv -\boldsymbol{h}\cdot\boldsymbol{g} \pmod N}}^{N-1} 1.$$

Since $N \in \mathbb{P}$, the congruence $h_{d+1} g_{d+1} \equiv -\boldsymbol{h}\cdot\boldsymbol{g} \pmod N$ has exactly one solution $g_{d+1} \in \{1, \ldots, N-1\}$ if $\boldsymbol{h}\cdot\boldsymbol{g} \not\equiv 0 \pmod N$ and no solution in $\{1, \ldots, N-1\}$ if $\boldsymbol{h}\cdot\boldsymbol{g} \equiv 0 \pmod N$. From this insight it follows that

$$\begin{aligned}
R((\boldsymbol{g}, g_{d+1}), N) &\le R(\boldsymbol{g}, N) + \frac{1}{N-1} \sum_{\boldsymbol{h} \in C_d(N)} \frac{1}{r(\boldsymbol{h})} \sum_{h_{d+1} \in C^*(N)} \frac{1}{r(h_{d+1})} \\
&= R(\boldsymbol{g}, N) + \frac{S_N}{N-1} \sum_{\boldsymbol{h} \in C_d(N)} \frac{1}{r(\boldsymbol{h})} \\
&= R(\boldsymbol{g}, N) + \frac{S_N}{N-1} (1 + S_N)^d \\
&\le \frac{1}{N-1} (1 + S_N)^d + \frac{S_N}{N-1} (1 + S_N)^d \\
&= \frac{1}{N-1} (1 + S_N)^{d+1},
\end{aligned}$$

where we used the induction hypotheses to bound $R(\boldsymbol{g}, N)$.□

It can be shown that $S_N \le 2 \log N$ (for a proof of this fact see Niederreiter [73, Lemmas 1 and 2]). Therefore, from Theorem 12 and Theorem 14 we obtain the following bound on the star-discrepancy of the lattice point set whose generating vector is constructed with Algorithm 2.

Corollary 2 (Joe [54]). *Let $N \in \mathbb{P}$ and suppose that $\boldsymbol{g} = (g_1, \ldots, g_s)$ is constructed according to Algorithm 2. Then for all $1 \le d \le s$ we have*

$$D^*_N(\mathscr{P}(\boldsymbol{g}^{(d)}, N)) \le \frac{d + (2 \log N)^d}{N},$$

where $\boldsymbol{g}^{(d)} = (g_1, \ldots, g_d)$.

Korobov suggested the use lattice points of the form $\boldsymbol{g} = (1, g, g^2, \ldots, g^{s-1})$ with $g \in \mathbb{Z}$ to restrict the number of candidates that must be inspected. At least for $N \in \mathbb{P}$ there is a result in the vein of (35) for such so-called *Korobov lattice points*.

Niederreiter [78, Theorem 5.18] showed that for any $N \in \mathbb{P}$ and any integer $s \geq 2$ we have

$$\frac{1}{N}\sum_{g=0}^{N-1} R((1,g,g^2,\ldots,g^{s-1}),N) < \frac{s-1}{N}(2\log N+1)^s.$$

Combined with Theorem 12 this leads to the following result.

Corollary 3. *Let $N \in \mathbb{P}$ and let $s \geq 2$ be an integer. For any real $0 < \varepsilon \leq 1$ there exist more than $(1-\varepsilon)N$ elements $g \in \{0,\ldots,N-1\}$ such that*

$$D_N^*(\mathscr{P}((1,g,g^2,\ldots,g^{s-1}),N)) \leq \frac{s}{N} + \frac{1}{\varepsilon}\frac{s-1}{2N}(2\log N+1)^s.$$

4.3 Weighted star-discrepancy of lattice point sets

For the weighted star-discrepancy of a lattice point set $\mathscr{P}(\boldsymbol{g},N)$ we obtain from Theorem 12

$$\begin{aligned} D_{N,\boldsymbol{\gamma}}^*(\mathscr{P}(\boldsymbol{g},N)) &= \max_{\emptyset\neq\mathfrak{u}\subseteq[s]} \boldsymbol{\gamma}_{\mathfrak{u}} D_N^*(\mathscr{P}(\boldsymbol{g}_{\mathfrak{u}},N)) \\ &\leq \sum_{\emptyset\neq\mathfrak{u}\subseteq[s]} \boldsymbol{\gamma}_{\mathfrak{u}} D_N^*(\mathscr{P}(\boldsymbol{g}_{\mathfrak{u}},N)) \\ &\leq \sum_{\emptyset\neq\mathfrak{u}\subseteq[s]} \boldsymbol{\gamma}_{\mathfrak{u}} \left(1-\left(1-\frac{1}{N}\right)^{|\mathfrak{u}|}\right) + \frac{1}{2}\sum_{\emptyset\neq\mathfrak{u}\subseteq[s]} \boldsymbol{\gamma}_{\mathfrak{u}} R(\boldsymbol{g}_{\mathfrak{u}},N), \end{aligned} \tag{36}$$

where $\boldsymbol{g}_{\mathfrak{u}}$ denotes the projection of $\boldsymbol{g}$ onto the components given by $\mathfrak{u}$. Hence $\mathscr{P}(\boldsymbol{g}_{\mathfrak{u}},N)$ is the $|\mathfrak{u}|$-dimensional lattice point set which is obtained by a projection of the points from $\mathscr{P}(\boldsymbol{g},N)$ onto the components given by $\mathfrak{u}$.

Set $\widetilde{r}(h,\gamma) = 1+\gamma$ if $h=0$, and $\gamma r(h)$ if $h \neq 0$ and set $\widetilde{r}(\boldsymbol{h},\boldsymbol{\gamma}) = \prod_{i=1}^s \widetilde{r}(h_i,\gamma_i)$. Then it follows from (34) that

$$\begin{aligned} &\sum_{\emptyset\neq\mathfrak{u}\subseteq[s]} \boldsymbol{\gamma}_{\mathfrak{u}} R(\boldsymbol{g}_{\mathfrak{u}},N) \\ &= -\sum_{\emptyset\neq\mathfrak{u}\subseteq[s]} \boldsymbol{\gamma}_{\mathfrak{u}} + \sum_{\emptyset\neq\mathfrak{u}\subseteq[s]} \frac{1}{N}\sum_{n=0}^{N-1}\prod_{i\in\mathfrak{u}} \gamma_i \left(1+\sum_{h\in C^*(N)} \frac{\exp(2\pi \mathtt{i} hng_i/N)}{|h|}\right) \\ &= -\prod_{i=1}^s(1+\gamma_i) + \frac{1}{N}\sum_{n=0}^{N-1}\prod_{i=1}^s \left(1+\gamma_i+\gamma_i\sum_{h\in C^*(N)} \frac{\exp(2\pi \mathtt{i} hng_i/N)}{|h|}\right) \\ &= -\prod_{i=1}^s(1+\gamma_i) + \frac{1}{N}\sum_{n=0}^{N-1}\prod_{i=1}^s \sum_{h\in C(N)} \widetilde{r}(h,\gamma_i)\exp(2\pi\mathtt{i} hng_i/N) && (37) \\ &= -\prod_{i=1}^s(1+\gamma_i) + \sum_{\boldsymbol{h}\in C_s(N)} \widetilde{r}(\boldsymbol{h},\boldsymbol{\gamma})\frac{1}{N}\sum_{n=0}^{N-1}\exp(2\pi\mathtt{i}\boldsymbol{h}\cdot\boldsymbol{g}n/N) \\ &= \sum_{\boldsymbol{h}\in C_s^*(N)\cap\mathcal{L}_{\boldsymbol{g},N}} \widetilde{r}(\boldsymbol{h},\boldsymbol{\gamma}) =: \widetilde{R}_{\boldsymbol{\gamma}}(\boldsymbol{g},N). && (38) \end{aligned}$$

From (37) we see that $\widetilde{R}_{\boldsymbol{\gamma}}(\boldsymbol{g},N)$ can be computed in $O(N^2 s)$ operations. Again this can be reduced to $O(Ns)$ operations by using an asymptotic expansion; for details see Joe [55, Appendix A].

Hence for the weighted star-discrepancy of a lattice point set $\mathscr{P}(\boldsymbol{g},N)$ we obtain

$$D^*_{N,\boldsymbol{\gamma}}(\mathscr{P}(\boldsymbol{g},N)) \le \sum_{\emptyset \neq \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}} \left(1 - \left(1 - \frac{1}{N}\right)^{|\mathfrak{u}|}\right) + \frac{1}{2}\widetilde{R}_{\boldsymbol{\gamma}}(\boldsymbol{g},N).$$

If the weights γ_i, $i \ge 1$, are summable, then it has been shown by Joe [55] that

$$\sum_{\emptyset \neq \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}} \left(1 - \left(1 - \frac{1}{N}\right)^{|\mathfrak{u}|}\right) \le \frac{\max(1,\Gamma)\exp(\sum_{i\ge 1}\gamma_i)}{N} \tag{39}$$

where $\Gamma = \sum_{i \ge 1} \gamma_i/(1+\gamma_i) < \infty$.

If $N \in \mathbb{P}$ one can again use the component-by-component algorithm (Algorithm 2) with R replaced by $\widetilde{R}_{\boldsymbol{\gamma}}$ for the construction of a 'good' lattice point.

Theorem 15 (Joe [55]). *Let $N \in \mathbb{P}$ and suppose that $\boldsymbol{g} = (g_1,\ldots,g_s)$ is constructed according to Algorithm 2 (with R replaced by $\widetilde{R}_{\boldsymbol{\gamma}}$). Then for all $1 \le d \le s$ we have*

$$\widetilde{R}_{\boldsymbol{\gamma}}(\boldsymbol{g}^{(d)},N) \le \frac{1}{N-1}\prod_{i=1}^{d}(1+\gamma_i(1+S_N))^d,$$

where $\boldsymbol{g}^{(d)} = (g_1,\ldots,g_d)$ and $S_N = \sum_{h \in C^*(N)} |h|^{-1}$.

Using the estimate $S_N \le 2\log N$ from the previous section we obtain the following result.

Corollary 4 (Joe [55]). *Let $N \in \mathbb{P}$ and suppose that $\boldsymbol{g} = (g_1,\ldots,g_s)$ is constructed according to Algorithm 2 (with R replaced by $\widetilde{R}_{\boldsymbol{\gamma}}$). Then for all $1 \le d \le s$ we have*

$$D^*_{N,\boldsymbol{\gamma}}(\mathscr{P}(\boldsymbol{g}^{(d)},N)) \le \sum_{\emptyset \neq \mathfrak{u} \subseteq [d]} \boldsymbol{\gamma}_{\mathfrak{u}} \left(1 - \left(1 - \frac{1}{N}\right)^{|\mathfrak{u}|}\right) + \frac{1}{N}\prod_{i=1}^{d}(1+2\gamma_i \log N)\,,$$

where $\boldsymbol{g}^{(d)} = (g_1,\ldots,g_d)$ and $[d] = \{1,\ldots,d\}$.

Assume that $\sum_{i\ge 1}\gamma_i < \infty$, then we obtain from Hickernell and Niederreiter [46, Lemma 3] that for any $\delta > 0$ there exists a $c_{\boldsymbol{\gamma},\delta} > 0$, independent of s and $\boldsymbol{\gamma}$, such that

$$\prod_{i=1}^{s}(1+2\gamma_i\log N) \le c_{\boldsymbol{\gamma},\delta} N^{\delta} \text{ for any } s \in \mathbb{N}.$$

Using this, (39) and Corollary 4 we obtain the following result.

Corollary 5. *Let $N \in \mathbb{P}$ and suppose that $\boldsymbol{g}$ is constructed according to Algorithm 2 (with R replaced by $\widetilde{R}_{\boldsymbol{\gamma}}$).*

If $\sum_{i \ge 1}\gamma_i < \infty$, then for any $\delta > 0$ there exists a $c_{\boldsymbol{\gamma},\delta} > 0$, independent of s and N, such that the weighted star-discrepancy of $\mathscr{P}(\boldsymbol{g},N)$ satisfies

$$D^*_{N,\boldsymbol{\gamma}}(\mathscr{P}(\boldsymbol{g},N)) \le \frac{c_{\boldsymbol{\gamma},\delta}}{N^{1-\delta}}.$$

In particular, if $\sum_{i \ge 1} \gamma_i < \infty$, then for any prime number N it follows from Corollary 5 that

$$\mathrm{disc}_{\infty,\boldsymbol{\gamma}}(N,s) \le \frac{c_{\boldsymbol{\gamma},\delta}}{N^{1-\delta}}$$

and that the bound can be achieved by a lattice point set.

Note that $\mathrm{disc}_{\infty,\boldsymbol{\gamma}}(0,s) = \max_{\emptyset \ne \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}} \ge \gamma_1 > 0$.

For $\varepsilon > 0$ and $\delta > 0$ let N be the smallest prime number that is larger or equal to $\lceil (c_{\boldsymbol{\gamma},\delta} \gamma_1^{-1} \varepsilon^{-1})^{1/(1-\delta)} \rceil =: M$. Then we have $\mathrm{disc}_{\infty,\boldsymbol{\gamma}}(N,s) \le \varepsilon \mathrm{disc}_{\infty,\boldsymbol{\gamma}}(0,s)$ and hence

$$N_{\infty,\boldsymbol{\gamma}}(\varepsilon,s) \le N < 2M = 2\lceil (c_{\boldsymbol{\gamma},\delta} \gamma_1^{-1} \varepsilon^{-1})^{1/(1-\delta)} \rceil,$$

where we used Bertand's postulate which tells us that $M \le N < 2M$. This bound, which is independent of the dimension s, was already presented in (19) and shows that the weighted star-discrepancy is strongly tractable with ε-exponent equal to one whenever the weights γ_i, $i \ge 1$, are summable.

For the weighted L_p-discrepancy of a point set $\mathscr{P}$ we have

$$L_{p,\boldsymbol{\gamma}}(\mathscr{P}) \le \left(\sum_{\emptyset \ne \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}}^p (D_N^*(\mathscr{P}_{\mathfrak{u}}))^p \right)^{1/p} \le \sum_{\emptyset \ne \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}} D_N^*(\mathscr{P}_{\mathfrak{u}}),$$

where we used Jensen's inequality, which states that $\sum_k a_k^p \le \left(\sum_k a_k\right)^p$ for any $p \ge 1$ and non-negative reals a_k. Hence, for a lattice point set we obtain from (36) and (38) that

$$L_{p,\boldsymbol{\gamma}}(\mathscr{P}(\boldsymbol{g},N)) \le \sum_{\emptyset \ne \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}} \left(1 - \left(1 - \frac{1}{N}\right)^{|\mathfrak{u}|} \right) + \frac{1}{2} \widetilde{R}_{\boldsymbol{\gamma}}(\boldsymbol{g},N).$$

This means that the results for the weighted star-discrepancy apply also for the weighted L_p-discrepancy. In particular, if $\sum_{i \ge 1} \gamma_i < \infty$, then the weighted L_p-discrepancy is strongly tractable with ε-exponent equal to one. (Note that $L_{p,\boldsymbol{\gamma}}^p(\emptyset) = -1 + \prod_{i=1}^s (1 + \frac{\gamma_i^p}{p+1})$.)

5 Polynomial lattice rules

There is also an algebraic analogue of lattice rules which is based on arithmetic of polynomials over finite fields. This construction has first been introduced by Niederreiter [77] as special construction of digital nets over a finite field $\mathbb{F}_q$ where q is a prime power. For simplicity, here we only consider prime bases b and the finite field $\mathbb{Z}_b$ of order b. On the other hand, here we generalize Niederreiter's approach to get a construction for higher order nets. This was first considered in [32].

Let $b \in \mathbb{P}$, let $\mathbb{Z}_b[x]$ denote the set of polynomials in x with coefficients in $\mathbb{Z}_b$ and let $\mathbb{Z}_b((x^{-1}))$ denote the set of formal Laurent series $\sum_{l=w}^\infty u_l x^{-l}$ where $w \in \mathbb{Z}$ and $u_l \in \mathbb{Z}_b$. For $n \in \mathbb{N}$ let $G_{b,n} := \{q \in \mathbb{Z}_b[x] : \deg(q) < n\}$ and $G_{b,n}^* = G_{b,n} \setminus \{0\}$. Note that $|G_n| = b^n$.

Definition 6. *Let $\alpha, m, s \in \mathbb{N}$ and choose an irreducible polynomial $p \in \mathbb{Z}_b[x]$ with $\deg(p) = \alpha m$. Further let $\boldsymbol{q} = (q_1, \ldots, q_s) \in G_{b,\alpha m}^s$ and consider the expansions*

$$\frac{q_i(x)}{p(x)} = \sum_{l=w_i}^{\infty} u_l^{(i)} x^{-l} \in \mathbb{Z}_b((x^{-1})),$$

where $w_i \le 1$. Define the $\alpha m \times m$ matrices $C_1, \ldots, C_s$ over $\mathbb{Z}_b$, $C_i = (c_{k,l}^{(i)})$, by

$$c_{k,l}^{(i)} = u_{k+l-1}^{(i)} \in \mathbb{Z}_b \quad \text{for } 1 \le i \le s, 1 \le k \le \alpha m, 1 \le l \le m.$$

The matrices $C_1, \ldots, C_s$ generate a digital (t,m,s)-net $\mathscr{P}_\alpha(\boldsymbol{q}, p)$ over $\mathbb{Z}_b$ which is called a polynomial lattice point set. The quadrature rule $Q_{\mathscr{P}_\alpha(\boldsymbol{q},p)}$ is called a polynomial lattice rule.

Polynomial lattice point sets can also be introduced independent from digital net theory. To this end let υ_n, $n \in \mathbb{N}$, be the map from $\mathbb{Z}_b((x^{-1}))$ to the interval $[0,1)$ defined by

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

Then the following construction is equivalent to Definition 6; see [33, Chapter 10]. For a given dimension $s \ge 1$, choose $p \in \mathbb{Z}_b[x]$ with $\deg(p) = \alpha m$ and let $q_1, \ldots, q_s \in \mathbb{Z}_b[x]$. Then $\mathscr{P}_\alpha(\boldsymbol{q}, p)$ is the point set consisting of the b^m points

$$\boldsymbol{x}_h = \left(\upsilon_{\alpha m}\left(\frac{h(x)q_1(x)}{p(x)}\right), \ldots, \upsilon_{\alpha m}\left(\frac{h(x)q_s(x)}{p(x)}\right)\right) \quad \text{for } h \in G_{b,m}.$$

Observe the similarity of this construction with the definition of ordinary lattice point sets from Definition 4, which is the reason for calling $\mathscr{P}_\alpha(\boldsymbol{q}, p)$ a polynomial lattice point set.

Classical polynomial lattice point sets assume that $\alpha = 1$, which also corresponds to the case of classical digital nets; see Niederreiter [77, 78], [33, Chapter 10] or [92]. In this case we simply omit the index α and write $\mathscr{P}(\boldsymbol{q}, p)$.

The dual polynomial lattice is of importance for studying polynomial lattice rules.

Definition 7. *Let $\alpha, m \in \mathbb{N}$. The dual polynomial lattice of a polynomial lattice point set $\mathscr{P}_\alpha(\boldsymbol{q}, p)$ with generating vector $\boldsymbol{q} \in G_{b,\alpha m}^s$ and $p \in \mathbb{Z}_b[x]$ with $\deg(p) = \alpha m$, is given by*

$$\mathcal{D}_\alpha(\boldsymbol{q}, p) = \{\boldsymbol{k} \in G_{b,\alpha m}^s : \boldsymbol{k} \cdot \boldsymbol{q} \equiv a \pmod{p} \text{ where } \deg(a) < (\alpha - 1)m\}.$$

For convenience we also write $\mathcal{D}_\alpha^*(\boldsymbol{q}, p) = \mathcal{D}_\alpha(\boldsymbol{q}, p) \setminus \{\boldsymbol{0}\}$. If $\alpha = 1$ we again omit the index α for the sake of simplicity.

For the following we will need the concept of b-adic Walsh functions which we introduce now. For $b \ge 2$ we denote by ω_b the bth primitive root of unity $\exp(2\pi \mathtt{i}/b)$.

Definition 8. *Let $k \in \mathbb{N}_0$ with b-adic expansion $k = \kappa_0 + \kappa_1 b + \kappa_2 b^2 + \cdots$. The kth b-adic Walsh function ${}_b\mathrm{wal}_k : \mathbb{R} \to \mathbb{C}$, periodic with period one, is defined as*

$${}_b\mathrm{wal}_k(x) = \omega_b^{\kappa_0 \xi_1 + \kappa_1 \xi_2 + \kappa_2 \xi_3 + \cdots},$$

for $x \in [0,1)$ with b-adic expansion $x = \xi_1 b^{-1} + \xi_2 b^{-2} + \xi_3 b^{-3} + \cdots$ (unique in the sense that infinitely many of the digits ξ_i must be different from $b-1$). We call the system $\{{}_b\mathrm{wal}_k : k \in \mathbb{N}_0\}$ the b-adic Walsh function system.

Now we generalise the definition of Walsh functions to higher dimensions.

Definition 9. *For dimension $s \geq 2$, and $k_1, \ldots, k_s \in \mathbb{N}_0$ we define the s-dimensional b-adic Walsh function ${}_b\mathrm{wal}_{k_1,\ldots,k_s} : \mathbb{R}^s \to \mathbb{C}$ by*

$${}_b\mathrm{wal}_{k_1,\ldots,k_s}(x_1, \ldots, x_s) := \prod_{j=1}^{s} {}_b\mathrm{wal}_{k_j}(x_j).$$

For 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 write, with some abuse of notation,

$${}_b\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}) := {}_b\mathrm{wal}_{k_1,\ldots,k_s}(x_1, \ldots, x_s).$$

The system $\{{}_b\mathrm{wal}_{\boldsymbol{k}} : \boldsymbol{k} \in \mathbb{N}_0^s\}$ is called the s-dimensional b-adic Walsh function system.

Basic properties of Walsh functions are summarised in [33, Appendix A].

An important property of polynomial lattice point sets is that

$$\sum_{\boldsymbol{x} \in \mathscr{P}_\alpha(\boldsymbol{q},p)} {}_b\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}) = \begin{cases} b^m & \text{if } \boldsymbol{k} \in \mathcal{D}_\alpha(\boldsymbol{q},p), \\ 0 & \text{otherwise.} \end{cases} \tag{40}$$

For a proof of this property we refer to [33, Lemmas 4.75 and 10.6] for $\alpha = 1$ and [33, Lemmas 4.75 and 15.25] for $\alpha \geq 1$.

Based on the definition of polynomial lattice point sets in terms of digital nets, the dual polynomial lattice is related to the dual net of a digital net; see [33, Lemma 10.6] or Niederreiter [78, Lemma 4.40]. There is also a concept related to the t-value of digital nets which we introduce in the following.

Definition 10. *Let $\alpha, m \in \mathbb{N}$. The figure of merit ϱ_α of a polynomial lattice point set $\mathscr{P}_\alpha(\boldsymbol{q},p)$ with generating vector $\boldsymbol{q} \in G_{b,\alpha m}^s$ and modulus $p \in \mathbb{Z}_b[x]$ with $\deg(p) = \alpha m$ is given by*

$$\varrho_\alpha(\boldsymbol{q},p) = -1 + \min_{\boldsymbol{k} \in \mathcal{D}_\alpha^*(\boldsymbol{q},p)} \deg_\alpha(\boldsymbol{k}),$$

where $\deg_\alpha(\boldsymbol{k}) = \sum_{i=1}^s \deg_\alpha(k_i)$ for $\boldsymbol{k} = (k_1, \ldots, k_s)$ and

$$\deg_\alpha(k_i) = \begin{cases} \sum_{u=1}^{\min(\alpha,\nu)} a_u & \text{for } k_i(x) = \kappa_1 x^{a_1 - 1} + \cdots + \kappa_\nu x^{a_\nu - 1}, \\ 0 & \text{otherwise,} \end{cases}$$

with $a_1 > a_2 > \cdots > a_\nu > 0$ and $\kappa_1, \ldots, \kappa_\nu \in \mathbb{Z}_b \setminus \{0\}$. For $\alpha = 1$ we again omit the index α and write $\varrho(\boldsymbol{q},p)$.

The following theorem connects the figure of merit of a polynomial lattice point set with the t-value.

Theorem 16 (Niederreiter [78, Theorem 4.42], D., Kritzer, P., and Schmid [25, Theorem 2]). *Let $\alpha, m \in \mathbb{N}$. Let $\mathscr{P}_\alpha(\boldsymbol{q},p)$ be a polynomial lattice point set with $\boldsymbol{q} \in G_{b,\alpha m}^s$ and modulus $p \in \mathbb{Z}_b[x]$ with $\deg(p) = \alpha m$. Then $\mathscr{P}_\alpha(\boldsymbol{q},p)$ is a digital $(t, \alpha, 1, \alpha m \times m, s)$-net in base b with*

$$t = \alpha m - \varrho_\alpha(\boldsymbol{q},p).$$

This result connects polynomial lattice point sets and digital nets. Hence the bounds on the discrepancy of digital nets, that is, Theorems 2, 3, 5, also apply to polynomial lattice point sets. Further results on the star-discrepancy of polynomial lattice point sets will be presented in Section 5.3.

The usefulness of polynomial lattice point sets for multivariate integration has been shown by connecting it to digital nets. However, as opposed to digital nets and sequences, no explicit constructions of polynomial lattice point sets are known except for dimension 2, see Niederreiter [78, pp. 87, 88]. For higher dimensions one relies on computer search algorithms to find good polynomial lattice rules. There are several methods of finding good polynomial lattice point sets, each based on a different criterion, which we consider in the following.

5.1 The worst-case error of polynomial lattice rules in weighted Walsh spaces

Similarly to lattice rules, we introduce a suitable space of functions for which the integration error of polynomial lattice rules for $\alpha = 1$ can be analysed (the case $\alpha > 1$ will be considered later in this section). This function space is based on Walsh functions as introduced in Definition 8 and 9, respectively.

We define a Walsh function space analogously to the Korobov space, essentially by replacing the exponential function $\exp(2\pi \mathrm{i} h x)$ by Walsh functions ${}_b\mathrm{wal}_k$.

Let $s \geq 1$ and $b \geq 2$ be integers, $\delta > 1$ a real and $\boldsymbol{\gamma} = (\gamma_i)_{i \geq 1}$ a sequence of nonnegative reals. The s-dimensional weighted Walsh space is the reproducing kernel Hilbert space of b-adic Walsh series $f(\boldsymbol{x}) = \sum_{\boldsymbol{k} \in \mathbb{N}_0^s} \widehat{f}(\boldsymbol{k})\, {}_b\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})$ with reproducing kernel defined by

$$K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}}(\boldsymbol{x},\boldsymbol{y}) = \sum_{\boldsymbol{k} \in \mathbb{N}_0^s} r_{\mathrm{wal},b,\delta}(\boldsymbol{k},\boldsymbol{\gamma})\, {}_b\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x} \ominus \boldsymbol{y}),$$

where for $\boldsymbol{k} = (k_1, \ldots, k_s)$ we put $r_{\mathrm{wal},b,\delta}(\boldsymbol{k},\boldsymbol{\gamma}) = \prod_{i=1}^s r_{\mathrm{wal},b,\delta}(k_i,\gamma_i)$ and for $k \in \mathbb{N}_0$ and $\gamma > 0$ we write

$$r_{\mathrm{wal},b,\delta}(k,\gamma) = \begin{cases} 1 & \text{if } k = 0, \\ \gamma b^{-\delta a} & \text{if } k = \kappa_0 + \kappa_1 b + \cdots + \kappa_a b^a \text{ and } \kappa_a \neq 0. \end{cases}$$

Furthermore, for $x = \sum_{i=w}^{\infty} \xi_i b^{-i}$ and $y = \sum_{i=w}^{\infty} \eta_i b^{-i}$ by $\ominus$ we denote the digit-wise subtraction modulo b, i.e.,

$$x \ominus y := \sum_{i=w}^{\infty} z_i b^{-i} \quad \text{where} \quad z_i := x_i - y_i \pmod{b}.$$

For vectors $\boldsymbol{x}, \boldsymbol{y}$ we apply $\ominus$ component wise.

The inner-product in $\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}})$ is given by

$$\langle f, g \rangle_{\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}})} = \sum_{\boldsymbol{k} \in \mathbb{N}_0^s} r_{\mathrm{wal},b,\delta}(\boldsymbol{k},\boldsymbol{\gamma})^{-1} \widehat{f}(\boldsymbol{k}) \overline{\widehat{g}(\boldsymbol{k})},$$

where the Walsh coefficients are given by $\widehat{f}(\boldsymbol{k}) = \int_{[0,1]^s} f(\boldsymbol{x})\, {}_b\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}$.

Similarly to the Korobov space, the worst-case integration error for a QMC rule in $\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}})$ using a polynomial lattice point set $\mathscr{P}(\boldsymbol{q},p)=\{\boldsymbol{x}_0,\ldots,\boldsymbol{x}_{N-1}\}$ is given by

$$e^2(\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}});\mathscr{P}(\boldsymbol{q},p))=\sum_{\boldsymbol{k}\in\mathbb{N}_0^s\setminus\{\boldsymbol{0}\}} r_{\mathrm{wal},b,\delta}(\boldsymbol{k},\boldsymbol{\gamma})\left|\frac{1}{N}\sum_{n=0}^{N-1} {}_b\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}_n)\right|^2.$$

There is also an analogue for Theorem 9.

Theorem 17 (D., Kuo, P., and Sloan [26, Lemma 4.1]). *The worst-case error of a polynomial lattice rule for integration in the Walsh space $\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}})$ is given by*

$$e^2(\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}});\mathscr{P}(\boldsymbol{q},p))=\sum_{\boldsymbol{h}\in\mathcal{D}^*(\boldsymbol{q},p)} r_{\mathrm{wal},b,\delta}(\boldsymbol{k},\boldsymbol{\gamma}).$$

In [30, Theorem 2] it was shown that there is a concise formula for the square worst-case error for a QMC rule based on a digital net. Applying this formula for a polynomial lattice point set $\mathscr{P}(\boldsymbol{q},p)=\{\boldsymbol{x}_0,\ldots,\boldsymbol{x}_{b^m-1}\}$ with $\boldsymbol{x}_n=(x_{n,1},\ldots,x_{n,s})$, we have

$$e^2(\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}});\mathscr{P}(\boldsymbol{q},p))=-1+\frac{1}{b^m}\sum_{n=0}^{b^m-1}\prod_{i=1}^{s}(1+\gamma_i\phi(x_{n,i})),$$

where for $x\in[0,1)$ we have

$$\phi(x)=\frac{b^\delta(b-1)}{b^\delta-b}-\begin{cases}0 & \text{if } x=0,\\ b^{\lfloor\log_b x\rfloor(\delta-1)}\frac{b^{2\delta}-b^\delta}{b^\delta-b} & \text{if } x>0.\end{cases}$$

This equation can now be used to obtain a construction algorithm for polynomial lattice point sets in the following way.

Algorithm 3 *Let $p\in\mathbb{Z}_b[x]$ be a polynomial of degree $m\in\mathbb{N}$ and let $s\geq 2$.*

1. *Choose $q_1=1\in\mathbb{Z}_b[x]$.*
2. *For $d>1$, assume we have already constructed $q_1,\ldots,q_{d-1}\in\mathbb{Z}_b[x]$. Then find $z\in G_{b,m}$ which minimises $e^2(\mathcal{H}(K_{\mathrm{wal},d,b,\delta,\boldsymbol{\gamma}});\mathscr{P}((q_1,\ldots,q_{d-1},z),p))$ as a function of $z\in G_{b,m}$.*

For polynomial lattice point sets constructed by Algorithm 3 we have the following result.

Theorem 18 (D., Kuo, P., and Sloan [26, Theorem 4.4]). *Let $p\in\mathbb{Z}_b[x]$ be irreducible, with $\deg(p)=m$. Suppose $\boldsymbol{q}=(q_1,\ldots,q_s)\in G_{b,m}^s$ is constructed by Algorithm 3. Then for all $1\leq d\leq s$ we have*

$$e(\mathcal{H}(K_{\mathrm{wal},d,b,\delta,\boldsymbol{\gamma}});\mathscr{P}(\boldsymbol{q}^{(d)},p))\leq\frac{2^{1/(2\lambda)}}{b^{m/(2\lambda)}}\prod_{i=1}^{d}(1+\mu(\delta\lambda)\gamma_i^\lambda)^{1/(2\lambda)},$$

for all $1/\delta<\lambda\leq 1$, where $\boldsymbol{q}^{(d)}=(q_1,\ldots,q_d)$ and where $\mu(\delta\lambda)=\frac{b^{\delta\lambda}(b-1)}{b^{\delta\lambda}-b}$.

Assume that $\sum_{i \ge 1} \gamma_i^\lambda < \infty$ for some $\lambda \in (1/\delta, 1]$. Then we obtain from Theorem 18 that

$$e(\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}}); \mathscr{P}(\boldsymbol{q}, p)) \le \frac{2^{1/(2\lambda)}}{b^{m/(2\lambda)}} \exp\left(\frac{1}{2\lambda} \sum_{i=1}^{s} \log(1 + \mu(\delta\lambda)\gamma_i^\lambda)\right)$$
$$\le \frac{2^{1/(2\lambda)}}{b^{m/(2\lambda)}} \exp\left(\frac{\mu(\delta\lambda)}{2\lambda} \sum_{i \ge 1} \gamma_i^\lambda\right) =: \frac{C_{b,\delta,\boldsymbol{\gamma}}}{b^{m/(2\lambda)}}$$

and this bound is independent of the dimension s. Note that for the initial error we have $e(\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}}); \emptyset) = 1$; see [30, p. 162].

For $\varepsilon > 0$ choose $m \in \mathbb{N}$ such that $b^{m-1} < \lceil (C_{b,\delta,\boldsymbol{\gamma}} \varepsilon^{-1})^{2\lambda} \rceil =: M \le b^m$. Then we have $e(\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}}); \mathscr{P}(\boldsymbol{q}, p)) \le \varepsilon$ and hence

$$N_{\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}})}(\varepsilon, s) \le b^m < bM = b\lceil (C_{b,\delta,\boldsymbol{\gamma}} \varepsilon^{-1})^{2\lambda} \rceil.$$

Thus multivariate integration in $\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}})$ is strongly tractable with ε-exponent at most 2λ whenever $\sum_{i \ge 1} \gamma_i^\lambda < \infty$ for some $\lambda \in (1/\delta, 1]$. The corresponding bounds can be achieved with polynomial lattice point sets. Furthermore, in [30, Corollary 1] it was shown that the condition $\sum_{i \ge 1} \gamma_i < \infty$ is also necessary for strong tractability.

Under weaker assumption on the weights one can still obtain polynomial tractability. For more results in this direction we refer to [30].

To put the result from Theorem 18 into context, we provide a lower bound on the integration error, which can be viewed as an analogue to Theorem 10. This result shows that the convergence rate in Theorem 18 is almost best possible.

Theorem 19 (Roth [94] and Heinrich, Hickernell, and Yue [47, Theorem 9]). *Let $\mathscr{P}$ be an arbitrary N-element point set in $[0,1]^s$ and let $\boldsymbol{w} = (w_0, \ldots, w_{N-1}) \in \mathbb{R}^N$. Then for any $\delta > 1$*

$$e(\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}}); \mathscr{P}; \boldsymbol{w}) \ge C_{s,\delta,\boldsymbol{\gamma}} \frac{(\log N)^{(s-1)/2}}{N^{\delta/2}},$$

where $C_{s,\delta,\boldsymbol{\gamma}} > 0$ depends on $\delta, s, \boldsymbol{\gamma}$ but not on N.

Proof. Choose $t \in \mathbb{N}_0$ such that

$$2N \le b^t < 4N.$$

Let $\boldsymbol{m} = (m_1, \ldots, m_s) \in \mathbb{N}_0^s$ with $\|\boldsymbol{m}\|_1 = m_1 + \cdots + m_s = t$ and let $\boldsymbol{l} = (l_1, \ldots, l_s) \in \mathbb{N}_0^s$ with $0 \le l_i < b^{m_i}$ for $1 \le i \le s$. Set

$$B_{\boldsymbol{l},\boldsymbol{m}} = \prod_{i=1}^{s} \left[\frac{l_i}{b^{m_i}}, \frac{l_i + 1}{b^{m_i}}\right).$$

For a given $\boldsymbol{x} \in [0,1)^s$ let $\boldsymbol{l}(\boldsymbol{x})$ be such that $\boldsymbol{x} \in B_{\boldsymbol{l}(\boldsymbol{x}),\boldsymbol{m}}$. Now we define

$$f_{\boldsymbol{m}}(\boldsymbol{x}) = \begin{cases} 0 & \text{if } B_{\boldsymbol{l}(\boldsymbol{x}),\boldsymbol{m}} \cap \mathscr{P} \ne \emptyset, \\ 1 & \text{otherwise.} \end{cases}$$

Then we have

$$\int_{[0,1]^s} f_{\boldsymbol{m}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \ge 1 - \frac{N}{b^t} \ge \frac{1}{2}.$$

We set

$$F(\boldsymbol{x}) = \sum_{\substack{\boldsymbol{m}\in\mathbb{N}_0^s\\ \|\boldsymbol{m}\|_1=t}} f_{\boldsymbol{m}}(\boldsymbol{x})$$

and therefore we have

$$\int_{[0,1]^s} F(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \sum_{\substack{\boldsymbol{m}\in\mathbb{N}_0^s\\ \|\boldsymbol{m}\|_1=t}} \int_{[0,1]^s} f_{\boldsymbol{m}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \ge \sum_{\substack{\boldsymbol{m}\in\mathbb{N}_0^s\\ \|\boldsymbol{m}\|_1=t}} \frac{1}{2} \ge \frac{1}{2}\binom{t+s-1}{s-1}.$$

Further, we have $F(\boldsymbol{x}_n) = 0$ for all $1 \le n \le N$ and thus

$$\left| \int_{[0,1]^s} F(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \sum_{n=1}^{N} w_n F(\boldsymbol{x}_n) \right| \ge \frac{1}{2}\binom{t+s-1}{s-1}.$$

We now estimate the norm of F. By Parseval's theorem we have

$$\sum_{\boldsymbol{k}\in\mathbb{N}_0^s} |\widehat{f}_{\boldsymbol{m}}(\boldsymbol{k})|^2 = \int_{[0,1]^s} |f_{\boldsymbol{m}}(\boldsymbol{x})|^2\,\mathrm{d}\boldsymbol{x} \le 1$$

and hence

$$\sum_{\boldsymbol{k}\in\mathbb{N}_0^s} |\widehat{f}_{\boldsymbol{m}}(\boldsymbol{k})\overline{\widehat{f}_{\boldsymbol{m}'}(\boldsymbol{k})}| \le \left(\sum_{\boldsymbol{k}\in\mathbb{N}_0^s} |\widehat{f}_{\boldsymbol{m}}(\boldsymbol{k})|^2\right)^{1/2} \left(\sum_{\boldsymbol{k}\in\mathbb{N}_0^s} |\widehat{f}_{\boldsymbol{m}'}(\boldsymbol{k})|^2\right)^{1/2} \le 1.$$

Further note that $\widehat{f}_{\boldsymbol{m}}(\boldsymbol{k}) = 0$ if there is an $1 \le i \le s$ such that $k_i \ge b^{m_i}$. Thus

$$\begin{aligned}
&\|F\|^2_{\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}})} \\
&= \sum_{\substack{\boldsymbol{m},\boldsymbol{m}'\in\mathbb{N}_0^s\\ \|\boldsymbol{m}\|_1=\|\boldsymbol{m}'\|_1=t}} \langle f_{\boldsymbol{m}}, f_{\boldsymbol{m}'}\rangle_{\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}})} \\
&\le \sum_{\substack{\boldsymbol{m},\boldsymbol{m}'\in\mathbb{N}_0^s\\ \|\boldsymbol{m}\|_1=\|\boldsymbol{m}'\|_1=t}} \sum_{u_1=0}^{\min(m_1,m_1')} \cdots \sum_{u_s=0}^{\min(m_s,m_s')} \\
&\qquad \times \sum_{k_1=\lfloor b^{u_1-1}\rfloor}^{b^{u_1}-1} \cdots \sum_{k_s=\lfloor b^{u_s-1}\rfloor}^{b^{u_s}-1} \frac{1}{r_{\mathrm{wal},b,\delta}(\boldsymbol{k},\boldsymbol{\gamma})} |\widehat{f}_{\boldsymbol{m}}(\boldsymbol{k})\overline{\widehat{f}_{\boldsymbol{m}'}(\boldsymbol{k})}| \\
&\le \sum_{\substack{\boldsymbol{m},\boldsymbol{m}'\in\mathbb{N}_0^s\\ \|\boldsymbol{m}\|_1=\|\boldsymbol{m}'\|_1=t}} \sum_{u_1=0}^{\min(m_1,m_1')} \cdots \sum_{u_s=0}^{\min(m_s,m_s')} \frac{1}{r_{\mathrm{wal},b,\delta}((b^{u_1}-1,\ldots,b^{u_s}-1),\boldsymbol{\gamma})} \\
&\le \sum_{\substack{\boldsymbol{m},\boldsymbol{m}'\in\mathbb{N}_0^s\\ \|\boldsymbol{m}\|_1=\|\boldsymbol{m}'\|_1=t}} \prod_{i=1}^{s} \left[1 + \gamma_i^{-1} \sum_{u_i=1}^{\min(m_i,m_i')} b^{\delta(u_i-1)}\right]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{\substack{\boldsymbol{m},\boldsymbol{m}' \in \mathbb{N}_0^s \\ \|\boldsymbol{m}\|_1 = \|\boldsymbol{m}'\|_1 = t}} \prod_{i=1}^s \left[1 + \gamma_i^{-1} \frac{b^{\delta \min(m_i, m_i')} - 1}{b^\delta - 1}\right] \\
&\le C'_{\delta,s,\boldsymbol{\gamma}} \sum_{\substack{\boldsymbol{m},\boldsymbol{m}' \in \mathbb{N}_0^s \\ \|\boldsymbol{m}\|_1 = \|\boldsymbol{m}'\|_1 = t}} b^{\delta \sum_{i=1}^s \min(m_i, m_i')},
\end{aligned}
$$

where $C'_{\delta,s,\boldsymbol{\gamma}} > 0$ is a constant which depends only on $\delta, s, \boldsymbol{\gamma}$.

We now estimate the last sum. We have

$$
\sum_{\substack{\boldsymbol{m},\boldsymbol{m}' \in \mathbb{N}_0^s \\ \|\boldsymbol{m}\|_1 = \|\boldsymbol{m}'\|_1 = t}} b^{\delta \sum_{i=1}^s \min(m_i, m_i')} = \sum_{\rho=0}^t b^{\delta\rho} A(\rho, t),
$$

where $A(\rho, t)$ is the number of solutions to the system of equations

$$
\begin{aligned}
m_1 + \cdots + m_s &= t, \\
m_1' + \cdots + m_s' &= t, \\
\min(m_1, m_1') + \cdots + \min(m_s, m_s') &= \rho.
\end{aligned}
$$

Thus, for $0 \le \rho \le t$ we have

$$
\begin{aligned}
A(\rho, t) &\le \sum_{u=1}^{s-1} \binom{s}{u} \binom{\rho + s - 1}{s - 1} \binom{t - \rho + u - 1}{u - 1} \binom{t - \rho + s - u - 1}{s - u - 1} \\
&\le C_s \rho^{s-1} (t - \rho)^{s-2},
\end{aligned}
$$

for some constant $C_s > 0$ independent of t and ρ. Therefore we have

$$
\begin{aligned}
\sum_{\rho=0}^t b^{\delta\rho} A(\rho, t) &\le C_s \sum_{\rho=0}^t b^{\delta\rho} \rho^{s-1} (t - \rho)^{s-1} \\
&\le C_s' \int_0^t b^{\delta\rho} \rho^{s-1} (t - \rho)^{s-1} \, \mathrm{d}\rho \\
&\le C_s'' t^{s-1/2} b^{t\delta/2} I_{s-1/2}(\log b^{t\delta/2}),
\end{aligned}
$$

where we used Prudnikov, Brychkov, and Marichev [93, Subsection 2.3.6, Eq. (1)] and where $I_{s-1/2}$ denotes the modified Bessel function of the first kind. Since $I_{s-1/2}(z) \le c_s \frac{\mathrm{e}^z}{\sqrt{2\pi z}}$ (see for instance Abramowitz and Stegun [1, Eq. 9.7.1]), it follows that there is a constant $C > 0$ (depending only on s, δ but not on t) such that

$$
\sum_{\rho=0}^t b^{\delta\rho} A(\rho, t) \le C b^{t\delta} t^{s-1}.
$$

Thus, there is a constant $C_{\delta,s,\boldsymbol{\gamma}} > 0$, such that

$$
\|F\|_{\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}})} \le C_{\delta,s,\boldsymbol{\gamma}} b^{t\delta/2} t^{(s-1)/2}.
$$

Let $g(\boldsymbol{x}) = F(\boldsymbol{x}) / \|F\|_{\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}})}$. Then g is in the unit ball of $\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}})$ and we have

$$
\begin{aligned}
e(\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}});\mathscr{P};\boldsymbol{w}) &\ge \left|\int_{[0,1]^s} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \sum_{n=1}^{N} w_n g(\boldsymbol{x}_n)\right| \\
&\ge \frac{1}{2}\|F\|^{-1}\binom{t+s-1}{s-1} \\
&\ge C N^{-\delta/2}(\log N)^{(s-1)/2}
\end{aligned}
$$

for some constant $C > 0$ independent of N.□

Notice that in the above constructions it was sufficient to use polynomial lattice point sets $\mathscr{P}(\boldsymbol{q}, p)$. Generally, the space $\mathcal{H}(K_{\mathrm{wal},s,b,\delta,\boldsymbol{\gamma}})$ does not contain smooth functions for $\delta > 1$. Hence, if one wants to consider classical spaces of functions with smoothness α, the above results only work for $\alpha \le 1$.

To extend the construction of polynomial lattice rules to integrands of smoothness $\alpha > 1$, one needs to use polynomial lattice point sets $\mathscr{P}_\alpha(\boldsymbol{g}, p)$ with $\alpha > 1$. This has been shown in Baldeaux, D., Greslehner, and P. [5] and Baldeaux, D., Leobacher, Nuyens, and P. [6].

We introduce a space of functions of smoothness $\alpha \in \mathbb{N}$ in the following. For such α let $L_{s,\alpha,\boldsymbol{\gamma}} : [0,1]^s \times [0,1]^s \to \mathbb{R}$ be the reproducing kernel given by

$$
L_{s,\alpha,\boldsymbol{\gamma}}(\boldsymbol{x},\boldsymbol{y}) = \prod_{i=1}^{s}\left[1 + \gamma_i \sum_{a=1}^{\alpha} \frac{B_a(x_i)B_a(y_i)}{(a!)^2} - (-1)^\alpha \gamma_i \frac{B_{2\alpha}(|x_i - y_i|)}{(2\alpha)!}\right],
$$

where $\boldsymbol{x} = (x_1, \ldots, x_s)$, $\boldsymbol{y} = (y_1, \ldots, y_s)$, and $B_a(\cdot)$ denotes the Bernoulli polynomial of degree a. For instance we have $B_1(x) = x - 1/2$, $B_2(x) = x^2 - x + 1/6$, $B_3(x) = x^3 - 3x^2/2 + x/2$, and so on. For dimension $s = 1$ the inner product in the reproducing kernel Hilbert space $\mathcal{H}(L_{1,\alpha,\gamma})$ is given by

$$
\begin{aligned}
\langle f, g\rangle_{\mathcal{H}(L_{1,\alpha,\gamma})} = &\int_0^1 f(x)\,\mathrm{d}x \int_0^1 \overline{g(x)}\,\mathrm{d}x + \frac{1}{\gamma}\sum_{a=1}^{\alpha-1}\int_0^1 f^{(a)}(x)\,\mathrm{d}x \int_0^1 \overline{g^{(a)}(x)}\,\mathrm{d}x \\
&+ \frac{1}{\gamma}\int_0^1 f^{(\alpha)}(x)\overline{g^{(\alpha)}(x)}\,\mathrm{d}x, \qquad (41)
\end{aligned}
$$

where $f^{(a)}$ denotes the ath derivative of f. For $s > 1$ the space $\mathcal{H}(L_{s,\alpha,\boldsymbol{\gamma}})$ is the s-fold tensor product of the one-dimensional spaces $\mathcal{H}(L_{1,\alpha,\gamma_i})$, $1 \le i \le s$.

The extension of the construction algorithm in Baldeaux, D., Greslehner, and P. [5] and Baldeaux, D., Leobacher, Nuyens, and P. [6] for $\alpha > 1$ is based on a continuous embedding of $\mathcal{H}(L_{s,\alpha,\boldsymbol{\gamma}})$ into a space of Walsh series [19]. It is then shown that Algorithm 3 can be used with a generalised quality criterion of order $\alpha > 1$ stemming from the Walsh space together with polynomial lattice point sets $\mathscr{P}_\alpha(\boldsymbol{q}, p)$. The convergence rate obtained in this case is of the form

$$
e^2(\mathcal{H}(L_{s,\alpha,\boldsymbol{\gamma}}); \mathscr{P}_\alpha(\boldsymbol{q}, p)) \le b^{-\lambda m} \prod_{i=1}^{s} [1 + \gamma_i^{1/\lambda} C_{b,\alpha,\lambda}]^\lambda,
$$

where $C_{b,\alpha,\lambda} > 0$ and $1 \le \lambda < \alpha$.

Notice that the Korobov space $\mathcal{H}(K_{s,\alpha,\boldsymbol{\gamma}})$ is continuously embedded in the space $\mathcal{H}(L_{s,\alpha,\boldsymbol{\gamma}})$. In fact, the restriction of $\mathcal{H}(L_{s,\alpha,\boldsymbol{\gamma}})$ to functions with one-periodic partial mixed derivatives up to order $\alpha - 1$ in each variable, yields the space $\mathcal{H}(K_{s,\alpha,\boldsymbol{\gamma}})$.

This can be seen for instance by considering the one-dimensional inner product for such functions, in which case (41) reduces to

$$\langle f, g\rangle_{\mathcal{H}(L_{1,\alpha,\gamma})} = \int_0^1 f(x)\,\mathrm{d}x \int_0^1 \overline{g(x)}\,\mathrm{d}x + \frac{1}{\gamma}\int_0^1 f^{(\alpha)}(x)\overline{g^{(\alpha)}(x)}\,\mathrm{d}x,$$

which is equivalent to the inner product given in (27) (which can be shown by substituting the Fourier series for f and g in the inner product above).

Thus, Theorem 10 also applies to multivariate integration in the space $\mathcal{H}(L_{s,\alpha,\mathbf{1}})$ and we have

$$e(\mathcal{H}(L_{s,\alpha,\mathbf{1}}); \mathscr{P}; \boldsymbol{w}) \ge e(\mathcal{H}(K_{s,\alpha}); \mathscr{P}; \boldsymbol{w}) \ge C(s,\alpha,\beta)\frac{(\log N)^{(s-1)/2}}{N^{\alpha/2}}.$$

Thus, the construction algorithm for higher order polynomial lattice rules yields quadrature rules which are almost best possible in terms of their convergence rate.

In the following we present an alternative approach to constructing higher order polynomial lattice rules using some ideas from D., Sloan, Wang, and Woźniakowski [35] and Sinescu and L'Ecuyer [101]. First, notice that there is an explicit formula for the worst-case error in $\mathcal{H}(L_{s,\alpha,\boldsymbol{\gamma}})$ using (3), given by

$$\begin{aligned} e^2(\mathcal{H}(L_{s,\alpha,\boldsymbol{\gamma}}); \mathscr{P}) &= \int_{[0,1]^s}\int_{[0,1]^s} L_{s,\alpha,\boldsymbol{\gamma}}(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y} \\ &\quad - \frac{2}{N}\sum_{n=1}^N \int_{[0,1]^s} L_{s,\alpha,\boldsymbol{\gamma}}(\boldsymbol{x},\boldsymbol{x}_n)\,\mathrm{d}\boldsymbol{x} + \frac{1}{N^2}\sum_{n,n'=1}^N L_{s,\alpha,\boldsymbol{\gamma}}(\boldsymbol{x}_n,\boldsymbol{x}_{n'}) \\ &= -1 + \frac{1}{N^2}\sum_{n,n'=1}^N L_{s,\alpha,\boldsymbol{\gamma}}(\boldsymbol{x}_n,\boldsymbol{x}_{n'}), \end{aligned}$$

where $\mathscr{P} = \{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N\}$. Thus, for a given point set, the worst-case error can be computed in $O(N^2 s)$ operations. Thus, we can use the following algorithm to find good higher order polynomial lattice rules.

Algorithm 4 *Let $\alpha, m, s \ge 2$ be integers, let $p \in \mathbb{Z}_b[x]$ be a polynomial of degree αm.*

For $1 \le d \le s$, assume we have already constructed $q_1,\ldots,q_{d-1} \in G_{b,\alpha m}$. Then randomly choose c polynomials $h_1,\ldots,h_c \in G_{b,\alpha m}$, where $h_1,\ldots,h_c$ are uniformly i.i.d. Set $q_d = h_u$, where $1 \le u \le c$ is the value of w which minimises

$$e^2(\mathcal{H}(L_{d,\alpha,\boldsymbol{\gamma}}); \mathscr{P}_\alpha((q_1,\ldots,q_{d-1},h_w),p)).$$

The following result stems from Baldeaux, D., Greslehner, and P. [5]. Let $1/\alpha < \lambda \le 1$. Assume that $\boldsymbol{q} = (q_1,\ldots,q_{d-1}) \in G_{b,\alpha m}^{d-1}$ is such that

$$e^{2\lambda}(\mathcal{H}(L_{d-1,\alpha,\boldsymbol{\gamma}}); \mathscr{P}_\alpha(\boldsymbol{q},p)) \le \frac{1}{b^m}\prod_{i=1}^{d-1}(1+\gamma_i^\lambda C_{b,\alpha,\lambda}), \tag{42}$$

where the constant $C_{b,\alpha,\lambda} > 0$ depends only on b, α, λ. Then we have

$$\frac{1}{b^{\alpha m}} \sum_{q \in G_{b,\alpha m}} e^{2\lambda}(\mathcal{H}(L_{d,\alpha,\boldsymbol{\gamma}}); \mathscr{P}_\alpha((\boldsymbol{q}, q), p)) \leq \frac{1}{b^m} \prod_{i=1}^{d} (1 + \gamma_i^\lambda C_{b,\alpha,\lambda}),$$

where $(\boldsymbol{q}, q) = (q_1, \ldots, q_{d-1}, q)$. Using Markov's inequality we obtain, given (42) holds, that for all $t \geq 1$ we have

$$\# \left\{ q \in G_{b,\alpha m} : e^{2\lambda}(\mathcal{H}(L_{d,\alpha,\boldsymbol{\gamma}}); \mathscr{P}_\alpha((\boldsymbol{q}, q), p)) \leq \frac{t}{b^m} \prod_{i=1}^{d} (1 + \gamma_i^\lambda C_{b,\alpha,\lambda}) \right\}$$
$$> b^{\alpha m} \left(1 - \frac{1}{t}\right),$$

which can be written as

$$\# \left\{ q \in G_{b,\alpha m} : e(\mathcal{H}(L_{d,\alpha,\boldsymbol{\gamma}}); \mathscr{P}_\alpha((\boldsymbol{q}, q), p)) \leq \frac{t}{b^{\frac{m}{2\lambda}}} \prod_{i=1}^{d} (1 + \gamma_i^\lambda C_{b,\alpha,\lambda})^{\frac{1}{2\lambda}} \right\}$$
$$> b^{\alpha m} \left(1 - \frac{1}{t^{2\lambda}}\right).$$

Hence the probability that at least one of $h_1, \ldots, h_c$ satisfies

$$e(\mathcal{H}(L_{d,\alpha,\boldsymbol{\gamma}}); \mathscr{P}_\alpha((\boldsymbol{q}, h_w), p)) \leq \frac{t}{b^{\frac{m}{2\lambda}}} \prod_{i=1}^{d} (1 + \gamma_i^\lambda C_{b,\alpha,\lambda})^{\frac{1}{2\lambda}}$$

is at least $1 - t^{-c/(2\lambda)}$. Thus, we have the following theorem.

Theorem 20. *Let $1/\alpha < \lambda \leq 1$. The probability that the vector $\boldsymbol{q} = (q_1, \ldots, q_s) \in G_{b,\alpha m}^s$ constructed by Algorithm 4 satisfies*

$$e(\mathcal{H}(L_{d,\alpha,\boldsymbol{\gamma}}); \mathscr{P}_\alpha(\boldsymbol{q}^{(d)}, p)) \leq \frac{t}{b^{\frac{m}{2\lambda}}} \prod_{i=1}^{d} (1 + \gamma_i^\lambda C_{b,\alpha,\lambda})^{\frac{1}{2\lambda}}$$

for all $1 \leq d \leq s$, where $\boldsymbol{q}^{(d)} = (q_1, \ldots, q_d)$, is at least $(1 - t^{-c/(2\lambda)})^s \geq 1 - st^{-c/(2\lambda)}$.

5.2 The construction of polynomial lattice rules based on the figure of merit

Another way of constructing polynomial lattice point sets is based on the figure of merit ϱ_α from Definition 10. In Algorithms 3 and 4 we used the worst-case error in some function spaces to compare polynomial lattice rules. In the following we show how Theorems 3, 5 and 16 can also be used to obtain a construction of good polynomial lattice point sets.

The idea is to search for polynomial lattice point sets which maximise the figure of merit.

Algorithm 5 *Let $\alpha, b, m, s \in \mathbb{N}$, $b \geq 2$, be given and let $p \in \mathbb{Z}_b[x]$ with $\deg(p) = \alpha m$. Choose $\boldsymbol{q} \in G_{b,\alpha m}^s$ which maximises $\varrho_\alpha(\boldsymbol{q}, p)$.*

Since computing the value of $\varrho_\alpha(\boldsymbol{q}, p)$ for given polynomials $\boldsymbol{q}, p$ is computationally expensive, Algorithm 5 can only be used for small values of α, m and s. In order to reduce the size of the search space, one can also consider a simplification due to Korobov [58].

Algorithm 6 *Let $\alpha, b, m, s \in \mathbb{N}$, $b \geq 2$, be given and let $p \in \mathbb{Z}_b[x]$ with $\deg(p) = \alpha m$. Choose $q \in G_{b,\alpha m}$ which maximises $\varrho_\alpha((q, q^2, \ldots, q^s), p)$.*

If $\alpha = 1$, one can also consider the generating vector $(1, q, \ldots, q^{s-1})$ (as originally proposed by Korobov for lattice rules).

The following result for $\alpha = 1$ was first shown in Larcher, Lauss, Niederreiter, and Schmid [66].

Theorem 21 (Larcher, Lauss, Niederreiter, and Schmid [66]). *Let $b \in \mathbb{P}$, $m, s \in \mathbb{N}$ with $s \geq 2$ and let $p \in \mathbb{Z}_b[x]$ be irreducible over $\mathbb{Z}_b[x]$ with $\deg(p) = m$. For $\varrho > 0$ define*

$$\Delta(s, \varrho) = \sum_{d=0}^{s-1} \binom{s}{d} (b-1)^{s-d} \sum_{l=0}^{\varrho+d} \binom{s-d+l-1}{l} b^l + 1 - b^{\varrho+s}.$$

1. *If $\Delta(s, \varrho) < b^m$, there exists a $\boldsymbol{q} \in G_{b,m}^s$ with*
$$\varrho(\boldsymbol{q}, p) \geq \varrho + s.$$
2. *If $\Delta(s, \varrho) < \frac{b^m}{s-1}$, there exists a polynomial $q \in G_{b,m}$ such that $\boldsymbol{g} \equiv \left(1, q, \ldots, q^{s-1}\right) \pmod{p}$ satisfies*
$$\varrho(\boldsymbol{q}, p) \geq \varrho + s.$$

Corollary 6. *Let $b \in \mathbb{P}$, $m, s \in \mathbb{N}$ with $s \geq 2$ and with m sufficiently large. Let $p \in \mathbb{Z}_b[x]$ be irreducible with $\deg(p) = m$.*

1. *There exists a vector $\boldsymbol{q} \in G_{b,m}^s$ with*
$$\varrho(\boldsymbol{q}, p) \geq \left\lfloor m - (s-1)(\log_b m - 1) + \log_b \frac{(s-1)!}{(b-1)^{s-1}} \right\rfloor.$$
2. *There exists a polynomial $q \in G_{b,m}$ such that $\boldsymbol{q} \equiv (1, q, \ldots, q^{s-1}) \pmod{p}$ satisfies*
$$\varrho(\boldsymbol{q}, p) \geq \left\lfloor m - (s-1)(\log_b m - 1) + \log_b \frac{(s-2)!}{(b-1)^{s-1}} \right\rfloor.$$

Together with Theorem 16 and Theorem 3 this result shows the existence of polynomial lattice point sets $\mathscr{P}(\boldsymbol{q}, p)$ with star-discrepancy of order

$$D_{b^m}^*(\mathscr{P}(\boldsymbol{q}, p)) \ll_{s,b} \frac{m^{2s-2}}{b^m}.$$

More precise results on the star-discrepancy of polynomial lattice point sets will be presented in Section 5.3.

For $\alpha > 1$ we have the following result from D., Kritzer, P., and Schmid [25].

Theorem 22 (D., Kritzer, P., and Schmid [25, Theorem 3]). *Let $b \in \mathbb{P}$, $m, \alpha, s \in \mathbb{N}$, $\alpha, s \geq 2$, and $p \in \mathbb{Z}_b[x]$ with $\deg(p) = \alpha m$ be irreducible. For $\varrho > 0$ define*

$$\Delta(s, \varrho, \alpha) = \sum_{l=0}^{\varrho} \sum_{i=1}^{s} \binom{s}{i} \sum_{\substack{l_1, \ldots, l_i \geq 1 \\ l_1 + \cdots + l_i = l}} \prod_{z=1}^{i} C(\alpha, l_z),$$

where

$$C(\alpha, l) = \sum_{v=1}^{\alpha-1} (b-1)^v \binom{l - \frac{v(v-1)}{2} - 1}{v-1} + \sum_{i=1}^{\lfloor l/\alpha \rfloor} (b-1)^\alpha b^{i-1} \binom{l - \alpha i - \frac{\alpha(\alpha-3)}{2} - 2}{\alpha - 2}.$$

1. *If $\Delta(s, \varrho, \alpha) < b^m$, there exists a $\boldsymbol{q} \in G_{b,\alpha m}^s$ with*
$$\varrho_\alpha(\boldsymbol{q}, p) \geq \varrho.$$
2. *If $\Delta(s, \varrho, \alpha) < \frac{b^m}{s-1}$, there exists a polynomial $q \in G_{b,\alpha m}$ such that $\boldsymbol{q} \equiv (q, q^2, \ldots, q^s) \pmod p$ satisfies*
$$\varrho_\alpha(\boldsymbol{q}, p) \geq \varrho.$$

The proofs of Theorems 21 and 22 are based on the following idea applied to codes and going back to Gilbert [41] and Varshamov [112]. We illustrate this idea for Algorithm 5.

First, note that there are $|G_{b,\alpha m}^s| = |G_{b,\alpha m}|^s = b^{\alpha m s}$ vectors $\boldsymbol{q}$ to choose from. The idea is to estimate the number of vectors $\boldsymbol{q} \in G_{b,\alpha m}^s$ for which $\varrho_{\alpha,m}(\boldsymbol{q}, p) < \varrho$ for some chosen $\varrho \geq 0$. If this number is smaller than the total number of possible choices of vectors $\boldsymbol{q} \in G_{b,\alpha m}^s$, it follows that there is at least one vector $\boldsymbol{q}$ with $\varrho_\alpha(\boldsymbol{q}, p) \geq \varrho$. For details we refer to [33, Chapter 10] and [33, Section 15.7.1], or to D., Kritzer, P., and Schmid [25] and Larcher, Lauss, Niederreiter, and Schmid [66].

In D., Kritzer, P., and Schmid [25] it was shown that Theorem 22 sometimes yields higher order digital nets with parameters better than the ones obtained using the higher order construction of Section 3.5.

5.3 Star-discrepancy of polynomial lattice point sets

For a polynomial lattice point set $\mathscr{P}(\boldsymbol{g}, p)$ with $\alpha = 1$ each point $\boldsymbol{x}_n$ is of the form $\boldsymbol{x}_n = \{\boldsymbol{y}_n / b^m\}$ with $\boldsymbol{y}_n \in \mathbb{Z}^s$. In particular, the elements of a polynomial lattice point set always have a finite b-adic digit expansion. A bound similar to that of Proposition 1 on the star-discrepancy of such point sets was first given by Niederreiter [75, Satz 2] (see also Niederreiter [78, Theorem 3.12]). An approach to this result by means of Walsh functions was described by Hellekalek [45, Theorem 1]. To formulate the result of Hellekalek we again need some notation.

Let $b \geq 2$ be an integer. For a vector $\boldsymbol{k} = (k_1, \ldots, k_s) \in \mathbb{N}_0^s$ we put $\rho_b(\boldsymbol{k}) := \prod_{i=1}^s \rho_b(k_i)$ where $\rho_b(0) = 1$ and where for $k \in \mathbb{N}_0$ we set

$$\rho_b(k) = \frac{1}{b^{r+1} \sin^2(\pi \kappa_r / b)}$$

if $k = \kappa_r b^r + k'$, where $\kappa_r \in \{1, \ldots, b-1\}$ and $0 \leq k' < b^r$.

Proposition 2 (Hellekalek [45] and Niederreiter [75]). *Let $N \geq 1$ and let $\mathscr{P} = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}\}$ be a point set in the s-dimensional unit cube where $\boldsymbol{x}_n$ is of the form $\boldsymbol{x}_n = \{\boldsymbol{y}_n / b^m\}$ with $\boldsymbol{y}_n \in \mathbb{Z}^s$, and $m, b \in \mathbb{N}$, $b \geq 2$. Then we have*

$$D_N^*(\mathscr{P}) \leq 1 - \left(1 - \frac{1}{b^m}\right)^s + \sum_{\substack{\boldsymbol{k} \in \mathbb{N}_0^s \\ 0 < |\boldsymbol{k}|_\infty < b^m}} \rho_b(\boldsymbol{k}) \left| \frac{1}{N} \sum_{n=0}^{N-1} {}_b\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}_n) \right|.$$

A proof of this result can also be found in [33, Proof of Theorem 3.28].

Applying this result to a polynomial lattice point set $\mathscr{P}(\boldsymbol{q}, p)$, in particular $N = b^m$, and using (40), we obtain the following result.

Theorem 23 (D., Leobacher, and P. [28]). *For the star-discrepancy of a polynomial lattice point set $\mathscr{P}(\boldsymbol{q}, p)$ we have*

$$D_{b^m}^*(\mathscr{P}(\boldsymbol{q}, p)) \leq 1 - \left(1 - \frac{1}{b^m}\right)^s + R_b(\boldsymbol{q}, p) \leq \frac{s}{b^m} + R_b(\boldsymbol{q}, p),$$

where

$$R_b(\boldsymbol{q}, p) := \sum_{\boldsymbol{h} \in \mathcal{D}^*(\boldsymbol{q}, p)} \rho_b(\boldsymbol{h}).$$

In the original version of the above results on the star-discrepancy, the squared sine function in the definition of ρ_b can be replaced by the ordinary sine function. Here we deal with the slightly weaker bound since in this case the quantity R_b can be computed efficiently. Assume that $\mathscr{P}(\boldsymbol{q}, p) = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{b^m-1}\}$, where $\boldsymbol{x}_n = (x_{n,1}, \ldots, x_{n,s})$. Then, using (40) we can write $R_b(\boldsymbol{q}, p)$ as

$$R_b(\boldsymbol{q}, p) = -1 + \frac{1}{b^m} \sum_{n=0}^{b^m-1} \prod_{i=1}^{s} \left(1 + \sum_{h=1}^{b^m-1} \rho_b(h)\, {}_b\mathrm{wal}_h(x_{n,i})\right). \tag{43}$$

and from this one can deduce that

$$R_b(\boldsymbol{q}, p) = -1 + \frac{1}{b^m} \sum_{n=0}^{b^m-1} \prod_{i=1}^{s} \phi_{b,m}(x_{n,i}),$$

where for $x = \xi_1 b^{-1} + \cdots + \xi_m b^{-m}$ we have

$$\phi_{b,m}(x) = \begin{cases} 1 + i_0 \frac{b^2-1}{3b} + \frac{2}{b}\xi_{i_0}(\xi_{i_0} - b) & \text{if } \xi_1 = \cdots = \xi_{i_0-1} = 0 \text{ and} \\ & \xi_{i_0} \neq 0 \text{ with } 1 \leq i_0 \leq m, \\ 1 + m\frac{b^2-1}{3b} & \text{otherwise.} \end{cases}$$

In particular, $R_b(\boldsymbol{q}, p)$ can be computed in $O(b^m s)$ operations; for a proof we refer to [33, Section 10.2].

It has been shown in Kritzer and P. [60] that there exists a $c_{s,b} > 0$ such that for any $p \in \mathbb{Z}_b[x]$ with $\deg(p) = m$ and any $\boldsymbol{q} \in (G_{b,m}^*)^s$ we have

$$R_b(\boldsymbol{q}, p) \geq c_{s,b} b^{\deg(\delta_s)} \frac{(m - \deg(\delta_s))^s}{b^m}, \quad \text{where } \delta_s := \gcd(q_1, \ldots, q_s, p).$$

On the other hand, let $b \in \mathbb{P}$, $s, m \in \mathbb{N}$, and let $p \in \mathbb{Z}_b[x]$ be irreducible with $\deg(p) = m$. Then we have

$$\frac{1}{|G_{b,m}^*|^s} \sum_{\boldsymbol{q} \in (G_{b,m}^*)^s} R_b(\boldsymbol{q}, p) = \frac{1}{b^m - 1} \left(\left(1 + m\frac{b^2-1}{3b} \right)^s - 1 - sm\frac{b^2-1}{3b} \right),$$

see D., Leobacher, and P. [28, Theorem 2.3] or [33, Theorem 10.21] for a proof. See also Niederreiter [78, Theorem 4.43]. In particular, we have the following result.

Theorem 24 (D., Leobacher, and P. [28, Theorem 2.3] and Kritzer and P. [60, Theorem 1.1]). *Let $b \in \mathbb{P}$, $s, m \in \mathbb{N}$, $s \geq 2$. For any $p \in \mathbb{Z}_b[x]$ with* $\deg(p) = m$ *there exists a $\boldsymbol{q} \in (G_{b,m}^*)^s$ such that*

$$R_b(\boldsymbol{q}, p) \asymp_{b,s} \frac{m^s}{b^m}$$

and this order of magnitude is best possible.

If p is irreducible we can use the following component-by-component algorithm for the construction of $\mathscr{P}(\boldsymbol{g}, p)$.

Algorithm 7 *Given $b \in \mathbb{P}$, $s, m \in \mathbb{N}$, and a polynomial $p \in \mathbb{Z}_b[x]$, with* $\deg(p) = m$.

1. *Choose $q_1 = 1$.*
2. *For $d > 1$, assume we have already constructed $q_1, \ldots, q_{d-1} \in G_{b,m}^*$. Then find $q_d \in G_{b,m}^*$ which minimises the quantity $R_b((q_1, \ldots, q_{d-1}, z), p)$ as a function of $z \in G_{b,m}^*$.*

Since the quantity $R_b(\boldsymbol{q}, p)$ can be calculated in $O(b^m s)$ operations, the cost of Algorithm 7 is of $O(b^{2m} s^2)$ operations. Using the fast component-by-component algorithm due to Nuyens and Cools [88] this can be reduced to $O(smb^m)$ operations; see [33, Section 10.3].

Theorem 25 (D., Leobacher, and P. [28, Theorem 2.7]). *Let $b \in \mathbb{P}$, $s, m \in \mathbb{N}$, and let $p \in \mathbb{Z}_b[x]$ be irreducible with* $\deg(p) = m$. *Suppose $\boldsymbol{q} = (q_1, \ldots, q_s) \in (G_{b,m}^*)^s$ is constructed according to Algorithm 7. Then for all $1 \leq d \leq s$ we have*

$$R_b(\boldsymbol{q}^{(d)}, p) \leq \frac{1}{b^m - 1} \left(1 + m\frac{b^2-1}{3b} \right)^d,$$

where $\boldsymbol{q}^{(d)} = (q_1, \ldots, q_d)$.

The proof of the result relies on similar ideas to those of the corresponding result for lattice rules from Theorem 14. A detailed proof can also be found in [33, Section 10.2.2].

A similar result for not necessarily irreducible polynomials is proven, but with much more technical effort, in D., Kritzer, Leobacher, and P. [24, Theorem 2].

Corollary 7 (D., Leobacher, and P. [28, Corollary 2.8]). *Let $b \in \mathbb{P}$, $s, m \in \mathbb{N}$, and let $p \in \mathbb{Z}_b[x]$ be irreducible with* $\deg(p) = m$. *Suppose $\boldsymbol{q} \in (G_{b,m}^*)^s$ is constructed according to Algorithm 7. Then we have*

$$D_{b^m}^*(\mathscr{P}(\boldsymbol{q}, p)) \leq \frac{s}{b^m} + \frac{1}{b^m - 1} \left(1 + m\frac{b^2-1}{3b} \right)^s$$

This result is not quite as good as the best existence result for point sets with low star-discrepancy from (11). However, the result is in line with the analogous result from the theory of lattice point sets, cf. Corollary 2.

For polynomial lattice point sets one knows that they also have the digital net structure. Based on this property one can prove the following improved, but still not optimal in the sense of (11), existence result.

Theorem 26 (Kritzer and P. [61], Larcher [65]). *Let $b \in \mathbb{P}$ and $s \in \mathbb{N}$. Then for any polynomial $p \in \mathbb{Z}_b[x]$ of degree m with $\gcd(p,x)=1$ or $p(x)=x^m$ there exists a generating vector $\boldsymbol{q} \in G_{b,m}^s$ such that*

$$D_{b^m}^*(\mathscr{P}(\boldsymbol{q},p)) \ll_{s,b} \frac{m^{s-1}\log m}{b^m}.$$

This bound is excellent in an asymptotic sense if $m \to \infty$. The dependence on the dimension s is not known. A construction of polynomial lattice point sets $\mathscr{P}(\boldsymbol{q},p)$ whose star-discrepancy satisfy the bound from Theorem 26 is not known so far. Despite the similarity of lattice point sets and polynomial lattice point sets in many aspects, there is no corresponding result of Theorem 26 for lattice point sets so far.

In dimension $s=2$ we have an explicit construction due to Niederreiter of a generating vector $\boldsymbol{q}$ such that $\mathscr{P}(\boldsymbol{q},x^m)$ has small star-discrepancy. For $s=b=2$ and for any $m \in \mathbb{N}$ let

$$q_m(x) = \sum_{j=0}^{\lfloor \log_2 m \rfloor + 1} x^{m-\lfloor m/2^j \rfloor} \in \mathbb{Z}_2[x] \quad \text{and} \quad \boldsymbol{q}_m = (1,q_m) \in \mathbb{Z}_2[x]^2.$$

Theorem 27 (Niederreiter [74] and [78]). *For any $m \in \mathbb{N}$ we have*

$$D_{2^m}^*(\mathscr{P}(\boldsymbol{q}_m,x^m)) \le \left(\frac{m}{3}+\frac{9}{19}\right)\frac{1}{2^m}.$$

Proof. Consider the continued fraction expansion

$$\frac{q_m(x)}{x^m} = [A_1,\ldots,A_h],$$

where $A_i \in \mathbb{Z}_2[x]$ and $\deg(A_i) \ge 1$ for $1 \le i \le h$. Niederreiter [74] proved that $K(q_m(x)/x^m) := \max_{1\le i\le h}\deg(A_i) = 1$. From Theorem 16 and from Niederreiter [78, Theorem 4.46] it follows that the two-dimensional polynomial lattice point set $\mathscr{P}(\boldsymbol{q}_m,x^m)$ is a digital $(0,m,2)$-net over $\mathbb{Z}_2$. In Larcher and P. [67] it has been shown that the star-discrepancy of any digital $(0,m,2)$-net over $\mathbb{Z}_2$ is at most $\left(\frac{m}{3}+\frac{9}{19}\right)2^{-m}$ and hence the result follows.□

5.4 Weighted star-discrepancy of polynomial lattice point sets

For the weighted star-discrepancy of a polynomial lattice point set we obtain from Theorem 23

$$D_{N,\boldsymbol{\gamma}}^*(\mathscr{P}(\boldsymbol{q},p)) = \max_{\emptyset \ne \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}} D_{N,\boldsymbol{\gamma}}^*(\mathscr{P}(\boldsymbol{q}_{\mathfrak{u}},p))$$

$$\leq \sum_{\emptyset \neq \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}} \left(1 - \left(1 - \frac{1}{b^m}\right)^{|\mathfrak{u}|}\right) + \sum_{\emptyset \neq \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}} R_b(\boldsymbol{q}_{\mathfrak{u}}, p),$$

where $\boldsymbol{q}_{\mathfrak{u}}$ denotes the projection of $\boldsymbol{q}$ onto the components given by $\mathfrak{u}$. Hence $\mathscr{P}(\boldsymbol{q}_{\mathfrak{u}}, p)$ is the $|\mathfrak{u}|$-dimensional polynomial lattice point set which is obtained by a projection of the points from $\mathscr{P}(\boldsymbol{q}, p)$ onto the components given by $\mathfrak{u}$.

Set $\widetilde{\rho}_b(h, \gamma) = 1 + \gamma$ if $h = 0$ and $\gamma \rho_b(h)$ if $h \neq 0$, and set $\widetilde{\rho}_b(\boldsymbol{h}, \boldsymbol{\gamma}) = \prod_{i=1}^s \widetilde{\rho}_b(h_i, \gamma_i)$. Then it follows from (43) in the same way as for the corresponding result for lattice rules that

$$\begin{aligned}
\widetilde{R}_{b,\boldsymbol{\gamma}}(\boldsymbol{q}, p) &:= \sum_{\emptyset \neq \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}} R_b(\boldsymbol{q}_{\mathfrak{u}}, p) \\
&= \sum_{\boldsymbol{h} \in \mathcal{D}(\boldsymbol{g}, p)^*} \widetilde{\rho}_b(\boldsymbol{h}, \boldsymbol{\gamma}) \\
&= -\prod_{i=1}^{s} (1 + \gamma_i) + \frac{1}{b^m} \sum_{n=0}^{b^m - 1} \prod_{i=1}^{s} (1 + \gamma_i \phi_{b,m}(x_{n,i})). \qquad (44)
\end{aligned}$$

From (44) we see that $\widetilde{R}_{b,\boldsymbol{\gamma}}(\boldsymbol{q}, p)$ can be computed in $O(b^m s)$ operations.

Hence for the weighted star-discrepancy of a polynomial lattice point set $\mathscr{P}(\boldsymbol{q}, p)$ we obtain

$$D^*_{b^m, \boldsymbol{\gamma}}(\mathscr{P}(\boldsymbol{q}, p)) \leq \sum_{\emptyset \neq \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}} \left(1 - \left(1 - \frac{1}{b^m}\right)^{|\mathfrak{u}|}\right) + \widetilde{R}_{b,\boldsymbol{\gamma}}(\boldsymbol{q}, p).$$

If p is irreducible one can again use the component-by-component algorithm (Algorithm 7) with R_b replaced by $\widetilde{R}_{b,\boldsymbol{\gamma}}$ for the construction of a 'good' generating vector.

Theorem 28 (D., Leobacher, and P. [28, Theorem 3.7]). *Let $b \in \mathbb{P}$, $s, m \in \mathbb{N}$, let $p \in \mathbb{Z}_b[x]$ be irreducible with $\deg(p) = m$. Suppose $\boldsymbol{q} = (q_1, \ldots, q_s) \in (G^*_{b,m})^s$ is constructed according to Algorithm 7 (with R_b replaced by $\widetilde{R}_{b,\boldsymbol{\gamma}}$). Then for all $1 \leq d \leq s$ we have*

$$\widetilde{R}_{b,\boldsymbol{\gamma}}(\boldsymbol{q}^{(d)}, p) \leq \frac{1}{b^m - 1} \prod_{i=1}^{d} \left(1 + \gamma_i \left(1 + m \frac{b^2 - 1}{3b}\right)\right),$$

where $\boldsymbol{q}^{(d)} = (q_1, \ldots, q_d)$.

Corollary 8 (D., Leobacher, and P. [28, Corollary 3.8]). *Let $b \in \mathbb{P}$, $s, m \in \mathbb{N}$, and let $p \in \mathbb{Z}_b[x]$ be irreducible with $\deg(p) = m$. Suppose $\boldsymbol{q} = (q_1, \ldots, q_s) \in (G^*_{b,m})^s$ is constructed according to Algorithm 7 (with R_b replaced by $\widetilde{R}_{b,\boldsymbol{\gamma}}$). Then for all $1 \leq d \leq s$ we have*

$$\begin{aligned}
D^*_{b^m, \boldsymbol{\gamma}}(\mathscr{P}(\boldsymbol{q}^{(d)}, p)) \leq &\sum_{\emptyset \neq \mathfrak{u} \subseteq [d]} \boldsymbol{\gamma}_{\mathfrak{u}} \left(1 - \left(1 - \frac{1}{b^m}\right)^{|\mathfrak{u}|}\right) \\
&+ \frac{1}{b^m - 1} \prod_{i=1}^{d} \left(1 + \gamma_i \left(1 + m \frac{b^2 - 1}{3b}\right)\right),
\end{aligned}$$

where $\boldsymbol{q}^{(d)} = (q_1, \ldots, q_d)$ and $[d] = \{1, \ldots, d\}$.

Similarly as for lattice point sets we can now deduce the following result.

Corollary 9. *Let p be irreducible and suppose that $\boldsymbol{q}$ is constructed according to Algorithm 7 (with R_b replaced by $\widetilde{R}_{b,\boldsymbol{\gamma}}$).*

If $\sum_{i\geq 1}\gamma_i < \infty$, then for any $\delta > 0$ there exists a $c_{b,\boldsymbol{\gamma},\delta} > 0$, independent of s and m, such that the weighted star-discrepancy of $\mathscr{P}(\boldsymbol{q}, p)$ satisfies

$$D^*_{b^m,\boldsymbol{\gamma}}(\mathscr{P}(\boldsymbol{q},p)) \leq \frac{c_{b,\boldsymbol{\gamma},\delta}}{b^{m(1-\delta)}}.$$

Assume that $\sum_{i\geq 1}\gamma_i < \infty$. For simplicity, we consider the case $b = 2$ only. Let $\delta > 0$ and let $N \in \mathbb{N}$ with binary representation $N = 2^{m_1} + \cdots + 2^{m_k}$, where $0 \leq m_1 < m_2 < \cdots < m_k$, i.e., $m_k = \lfloor \log_2 N \rfloor$, where $\log_2$ denotes the logarithm in base 2. For each $1 \leq i \leq k$ choose an irreducible polynomial $p_i \in \mathbb{Z}_2[x]$ with $\deg(p_i) = m_i$ and construct a vector $\boldsymbol{q}_i$ according to Algorithm 7 (with $b = 2$ and with R_2 replaced by $\widetilde{R}_{2,\boldsymbol{\gamma}}$). Then for the resulting polynomial lattice point sets $\mathscr{P}(\boldsymbol{q}_i, p_i)$ we obtain from Corollary 9 that

$$D^*_{2^{m_i},\boldsymbol{\gamma}}(\mathscr{P}(\boldsymbol{q}_i,p_i)) \leq \frac{c_{\boldsymbol{\gamma},\delta}}{2^{m_i(1-\delta)}}$$

for all $1 \leq i \leq k$. Let $\mathscr{P}_N = \mathscr{P}(\boldsymbol{q}_1, p_1) \cup \ldots \cup \mathscr{P}(\boldsymbol{q}_k, p_k)$ (here we mean a superposition where the multiplicity of elements matters). Then it follows from the triangle inequality for the star-discrepancy (see [33, Proposition 3.16]) and the definition of the weighted star-discrepancy that

$$\begin{aligned} D^*_{N,\boldsymbol{\gamma}}(\mathscr{P}_N) &\leq \sum_{i=1}^{k} \frac{2^{m_i}}{N} D^*_{2^{m_i},\boldsymbol{\gamma}}(\mathscr{P}(\boldsymbol{q}_i,p_i)) \leq \frac{c_{\boldsymbol{\gamma},\delta}}{N}\sum_{i=1}^{k} 2^{m_i\delta} \\ &\leq \frac{c_{\boldsymbol{\gamma},\delta}}{N}\sum_{j=0}^{\lfloor \log_2 N\rfloor} 2^{j\delta} \leq \frac{\widetilde{c}_{\boldsymbol{\gamma},\delta}}{N^{1-\delta}}. \end{aligned}$$

Hence for each $s, N \in \mathbb{N}$ there exists an N-element point set $\mathscr{P}_N$ in $[0,1)^s$ with $D^*_{N,\boldsymbol{\gamma}}(\mathscr{P}_N) \leq \widetilde{c}_{\gamma,\delta} N^{-1+\delta}$ and this point set is a superposition of polynomial lattice point sets.

In particular, if $\sum_{i\geq 1}\gamma_i < \infty$, it follows that for any $s, N \in \mathbb{N}$ we have

$$\mathrm{disc}_{\infty,\boldsymbol{\gamma}}(N,s) \leq \frac{\widetilde{c}_{\boldsymbol{\gamma},\delta}}{N^{1-\delta}}$$

and that the bound can be achieved by a superposition of polynomial lattice point sets.

Recall that $\mathrm{disc}_{\infty,\boldsymbol{\gamma}}(0,s) = \max_{\emptyset\neq\mathfrak{u}\subseteq[s]} \boldsymbol{\gamma}_{\mathfrak{u}} \geq \gamma_1 > 0$.

For $\varepsilon > 0$ and $\delta > 0$ we obtain

$$N_{\infty,\boldsymbol{\gamma}}(\varepsilon,s) \leq \left\lceil (\widetilde{c}_{\boldsymbol{\gamma},\delta}\gamma_1^{-1}\varepsilon^{-1})^{1/(1-\delta)} \right\rceil.$$

This bound, which is independent of the dimension s, was already presented in (19) and shows again that the weighted star-discrepancy is strongly tractable with ε-exponent equal to one whenever the weights γ_i, $i \geq 1$, are summable.

As for lattice point sets, for the weighted L_p-discrepancy of a polynomial lattice point set we obtain

$$L_{p,\boldsymbol{\gamma}}(\mathscr{P}(\boldsymbol{q},p)) \le \sum_{\emptyset \neq \mathfrak{u} \subseteq [s]} \boldsymbol{\gamma}_{\mathfrak{u}} \left(1 - \left(1 - \frac{1}{b^m}\right)^{|\mathfrak{u}|}\right) + \widetilde{R}_{b,\boldsymbol{\gamma}}(\boldsymbol{g},p).$$

Again, this means that the results for the weighted star-discrepancy apply also for the weighted L_p-discrepancy.

Acknowledgement. The first author is supported by a Queen Elizabeth II Fellowship from the Australian Research Council. The second author is partially supported by the Austrian Science Foundation (FWF), Project S9609, that is part of the Austrian National Research Network "Analytic Combinatorics and Probabilistic Number Theory".

The authors thank Michaela Szölgyenyi and Henryk Woźniakowski for many helpful suggestions.

References

1. M. Abramowitz and I. A. Stegun: Handbook of mathematical functions with formulas, graphs, and mathematical tables. National Bureau of Standards Applied Mathematics Series, 55, U.S. Government Printing Office, Washington, D.C. 1964.
2. Ch. Aistleitner: Covering numbers, dyadic chaining and discrepancy. J. Complexity, 27, 531–540, 2011.
3. N. Aronszajn: Theory of reproducing kernels. Trans. Amer. Math. Soc., 68, 337–404, 1950.
4. N. S. Bakhvalov: Approximate computation of multiple integrals. Vestnik Moskov. Univ. Ser. Mat. Meh. Astr. Fiz. Him., 4, 3–18, 1959. (Russian)
5. J. Baldeaux, J. Dick, J. Greslehner and F. Pillichshammer: Construction algorithms for higher order polynomial lattice rules. J. Complexity, 27, 281–299, 2011.
6. J. Baldeaux, J. Dick, G. Leobacher, D. Nuyens and F. Pillichshammer: Fast construction of multivariate higher order quadrature rules using polynomial lattice rules. Numer. Algorithms, 59, 403–431, 2012.
7. R. Béjian: Minoration de la discrépance d'une suite quelconque sur T. Acta Arith., 41, 185–202, 1982. (French)
8. D. Bilyk, M. T. Lacey and A. Vagharshakyan: On the small ball inequality in all dimensions. J. Funct. Anal., 254, 2470–2502, 2008.
9. J. Brauchart and J. Dick: A simple proof of Stolarsky's invariance principle. Proc. Amer. Math. Soc., 2011. (to appear)
10. W. W. L. Chen: On irregularities of point distribution. Mathematika, 27, 153–170, 1980.
11. W. W. L. Chen and M. M. Skriganov: Explicit constructions in the classical mean squares problem in irregularities of point distribution. Reine Angew. Math., 545, 67–95, 2002.
12. W. W. L. Chen and M. M. Skriganov: Orthogonality and digit shifts in the classical mean squares problem in irregularities of point distribution. In: *Diophantine Approximation: Festschrift for Wolfgang Schmidt*, pp. 141–159. Springer, Berlin, 2008.

13. R. Cools and J. N. Lyness: Three- and four-dimensional K-optimal lattice rules of moderate trigonometric degree. Math. Comp., 70, 1549–1567, 2001.
14. L. L. Cristea, J. Dick and F. Pillichshammer: On the mean square weighted L_2 discrepancy of randomized digital nets in prime base. J. Complexity, 22, 605–629, 2006.
15. H. Davenport: Note on irregularities of distribution. Mathematika, 3, 131–135, 1956.
16. J. Dick: On the convergence rate of the component-by-component construction of good lattice rules. J. Complexity, 20, 493–522, 2004.
17. J. Dick: Explicit constructions of quasi-Monte Carlo rules for the numerical integration of high dimensional periodic functions. SIAM J. Numer. Anal., 45, 2141–2176, 2007.
18. J. Dick: Walsh spaces containing smooth functions and quasi-Monte Carlo rules of arbitrary high order. SIAM J. Numer. Anal., 46, 1519–1553, 2008.
19. J. Dick: The decay of the Walsh coefficients of smooth functions. Bull. Austral. Math. Soc., 80, 430–453, 2009.
20. J. Dick: Higher order scrambled digital nets achieve the optimal rate of the root mean square error for smooth integrands, Ann. Statist., 39, 1372–1398, 2011.
21. J. Dick: QuasiMonte Carlo Numerical Integration on $\mathbb{R}^s$: Digital Nets and Worst-Case Error. SIAM J. Numer. Anal., 49, 1661–1691, 2011.
22. J. Dick and J. Baldeaux: Equidistribution properties of generalized nets and sequences. In: Monte Carlo and quasi-Monte Carlo Methods 2008, P. L'Ecuyer and A. B. Owen (eds.), Springer Verlag, Berlin, 305–322, 2009.
23. J. Dick and P. Kritzer: Duality theory and propagation rules for generalized digital nets. Math. Comp., 79, 993–1017, 2010.
24. J. Dick, P. Kritzer, G. Leobacher and F. Pillichshammer: Constructions of general polynomial lattice rules based on the weighted star-discrepancy. Finite Fields Appl., 13, 1045–1070, 2007.
25. J. Dick, P. Kritzer, F. Pillichshammer and W. Ch. Schmid: On the existence of higher order polynomial lattices based on a generalized figure of merit. J. Complexity, 23, 581–593, 2007.
26. J. Dick, F. Y. Kuo, F. Pillichshammer and I. H. Sloan: Construction algorithms for polynomial lattice rules for multivariate integration. Math. Comp., 74, 1895–1921, 2005.
27. J. Dick, G. Larcher, F. Pillichshammer and H. Woźniakowski: Exponential convergence and tractability of multivariate integration for Korobov spaces. Math. Comp., 80, 905–930, 2011.
28. J. Dick, G. Leobacher and F. Pillichshammer: Construction algorithms for digital nets with low weighted star-discrepancy. SIAM J. Numer. Anal., 43, 76–95, 2005.
29. J. Dick and H. Niederreiter: On the exact t-value of Niederreiter and Sobol' sequences. J. Complexity, 24, 572–581, 2008.
30. J. Dick and F. Pillichshammer: Multivariate integration in weighted Hilbert spaces based on Walsh functions and weighted Sobolev spaces. J. Complexity, 21, 149–195, 2005.
31. J. Dick and F. Pillichshammer: On the mean square weighted $\mathcal{L}_2$ discrepancy of randomized digital (t, m, s)-nets over $\mathbb{Z}_2$. Acta Arith., 117, 371–403, 2005.

32. J. Dick and F. Pillichshammer: Strong tractability of multivariate integration of arbitrary high order using digitally shifted polynomial lattice rules. J. Complexity, 23, 436–453, 2007.
33. J. Dick and F. Pillichshammer: *Digital nets and sequences – Discrepancy theory and quasi-Monte Carlo integration.* Cambridge University Press, Cambridge, 2010.
34. J. Dick, I. H. Sloan, X. Wang and H. Woźniakowski: Liberating the weights. J. Complexity, 20, 593–623, 2004.
35. J. Dick, I. H. Sloan, X. Wang and H. Woźniakowski: Good lattice rules in weighted Korobov spaces with general weights. Numer. Math., 103, 63–97, 2006.
36. B. Doerr and M. Gnewuch: Construction of low-discrepancy point sets of small size by bracketing covers and dependent randomized rounding. In: *Monte Carlo and Quasi-Monte Carlo Methods 2006*, pp. 299–312. Springer, Berlin, 2007.
37. B. Doerr, M. Gnewuch, and A. Srivastav: Bounds and constructions for the star-discrepancy via δ-covers. J. Complexity, 21, 691–709, 2005.
38. B. Doerr, M. Gnewuch, P. Kritzer, and F. Pillichshammer: Component-by-component construction of low-discrepancy point sets of small size. Monte Carlo Methods Appl., 14, 129–149, 2008.
39. H. Faure: Discrépance de suites associées à un système de numération (en dimension s), Acta Arith., 41, 337–351, 1982.
40. K. K. Frolov: Upper bound of the discrepancy in metric L_p, $2 \leq p < \infty$. Dokl. Akad. Nauk SSSR, 252, 805–807, 1980.
41. E. N. Gilbert: A comparison of signalling alphabets. Bell System Technical Journal, 31, 504–522, 1952.
42. M. Gnewuch: Weighted geometric discrepancies and numerical integration on reproducing kernel Hilbert spaces. J. Complexity, 28, 2–17, 2012.
43. M. Gnewuch: Infinite-dimensional integration on weighted Hilbert spaces. Math. Comp., 2012. (to appear)
44. M. Gnewuch, M., A. Srivastav and C. Winzen: Finding optimal volume subintervals with k-points and calculating the star-discrepancy are NP-hard problems. J. Complexity, 25, 115–127, 2009.
45. P. Hellekalek: General discrepancy estimates: the Walsh function system. Acta Arith., 67, 209–218, 1994.
46. F. J. Hickernell and H. Niederreiter: The existence of good extensible rank-1 lattices. J. Complexity, 19, 286–300, 2003.
47. S. Heinrich, F. J. Hickernell and R.-X. Yue: Optimal quadrature for Haar wavelet spaces. Math. Comp., 73, 259–277, 2004.
48. S. Heinrich, E. Novak, G. Wasilikowski, and H. Woźniakowski: The inverse of the star-discrepancy depends linearly on the dimension. Acta Arith., 96, 279–302, 2001.
49. A. Hinrichs: Covering numbers, Vapnik-Červonenkis classes and bounds for the star-discrepancy. J. Complexity, 20, 477–483, 2004.
50. A. Hinrichs, F. Pillichshammer and W. Ch. Schmid: Tractability properties of the weighted star-discrepancy. J. Complexity, 24, 134–143, 2008.
51. E. Hlawka: Über die Diskrepanz mehrdimensionaler Folgen mod 1. Math. Z., 77, 273–284, 1961. (German).
52. E. Hlawka: Zur angenäherten Berechnung mehrfacher Integrale. Monatsh. Math., 66, 140–151, 1962. (German)

53. H. S. Hong and F. J. Hickernell: Algorithm 823: Implementing scrambled digital sequences. ACM Trans. Math. Softw., 29, 95–109, 2003.
54. S. Joe: Component by component construction of rank-1 lattice rules having $O(n^{-1}(\ln n))^d)$ star-discrepancy. In: *Monte Carlo and Quasi-Monte Carlo Methods 2002.* (H. Niederreiter, ed.), Springer, Berlin Heidelberg New York, 2004.
55. S. Joe: Construction of good rank-1 lattice rules based on the weighted star-discrepancy. In: *Monte Carlo and Quasi-Monte Carlo Methods 2004.* (H. Niederreiter and D. Talay, eds.), pp. 181–196. Springer, Berlin, 2006.
56. S. Joe and I. H. Sloan: On computing the lattice rule criterion R. Math. Comp., 59, 557–568, 1992.
57. N. M. Korobov: The approximate computation of multiple integrals. Dokl. Akad. Nauk SSSR, 124, 1207–1210, 1959. (Russian)
58. N. M. Korobov: *Number-theoretic methods in approximate analysis.* Gosudarstv. Izdat. Fiz.-Mat. Lit., Moscow, 1963. (Russian)
59. P. Kritzer: Improved upper bounds on the star-discrepancy of (t, m, s)-nets and (t, s)-sequences. J. Complexity, 22, 336–347, 2006.
60. P. Kritzer and F. Pillichshammer: A lower bound on a quantity related to the quality of polynomial lattices. Funct. Approx. Comment. Math., 45, 125–137, 2011.
61. P. Kritzer and F. Pillichshammer: Low Discrepancy Polynomial Lattice Point Sets. Submitted, 2011.
62. F. Y. Kuo: Component-by-component constructions achieve the optimal rate of convergence for multivariate integration in weighted Korobov and Sobolev spaces. J. Complexity, 19, 301–320, 2003.
63. L. Kuipers and H. Niederreiter: *Uniform Distribution of Sequences.* John Wiley, New York, 1974; reprint, Dover Publications, Mineola, NY, 2006.
64. G. Larcher: A best lower bound for good lattice points. Monatsh. Math., 104, 45–51, 1987.
65. G. Larcher: Nets obtained from rational functions over finite fields. Acta Arith., 63, 1–13, 1993.
66. Larcher, G.; Lauss, A.; Niederreiter, H.; Schmid, W. Ch. Optimal polynomials for (t, m, s)(t,m,s)-nets and numerical integration of multivariate Walsh series. SIAM J. Numer. Anal., 33, 2239–2253, 1996.
67. G. Larcher and F. Pillichshammer: Sums of distances to the nearest integer and the discrepancy of digital nets. Acta Arith., 106, 379–408, 2003.
68. G. Leobacher and F. Pillichshammer: Bounds for the weighted L^p discrepancy and tractability of integration. J. Complexity, 19, 529–547, 2003.
69. W. L. Loh: On the asymptotic distribution of scrambled net quadrature. Ann. Statist., 31, 1282–1324, 2003.
70. J. Lyness: Notes on lattice rules. Numerical integration and its complexity (Oberwolfach, 2001). J. Complexity, 19, 321–331, 2003.
71. J. Matoušek: On the L_2-discrepancy for anchored boxes, J. Complexity, 14, 527–556, 1998.
72. J. Matoušek: *Geometric discrepancy.* Springer, Berlin Heidelberg New York, 1999.
73. H. Niederreiter: Existence of good lattice points in the sense of Hlawka. Monatsh. Math., 86, 203–219, 1978.
74. H. Niederreiter: Rational functions with partial quotients of small degree in their continued fraction expansion. Monatsh. Math., 103, 269–288, 1987.

75. H. Niederreiter: Pseudozufallszahlen und die Theorie der Gleichverteilung. Österreich. Akad. Wiss. Math.-Natur. Kl. Sitzungsber. II 195: 109–138, 1986. (In German).
76. H. Niederreiter: Low-discrepancy and low-dispersion sequences, J. Number Th., 30, 51–70, 1987.
77. H. Niederreiter: Low-discrepancy point sets obtained by digital constructions over finite fields. Czechoslovak Math. J., 42, 143–166, 1992.
78. H. Niederreiter: *Random Number Generation and Quasi-Monte Carlo Methods.* SIAM, Philadelphia, 1992.
79. H. Niederreiter and C. P. Xing: *Low-discrepancy sequences and global function fields with many rational places*, Finite Fields Appl., 2, 241–273, 1996.
80. H. Niederreiter and C.P. Xing: Algebraic curves over finite fields with many rational points and their applications. Number theory, 287300, Trends Math., Birkhaeuser, Basel, 2000.
81. E. Novak: *Deterministic and stochastic error bounds in numerical analysis.* Lecture Notes in Mathematics, 1349. Springer-Verlag, Berlin, 1988.
82. E. Novak and H. Woźniakowski: When are integration and discrepancy tractable? In: *Foundations of computational mathematics (Oxford, 1999)*, volume 284 of *London Math. Soc. Lecture Note Ser.*, pp. 211–266. Cambridge Univ. Press, Cambridge, 2001.
83. E. Novak and H. Woźniakowski: *Tractability of Multivariate Problems. Volume I: Linear Information.* European Mathematical Society Publishing House, Zürch, 2008.
84. E. Novak and H. Woźniakowski: L_2 discrepancy and multivariate integration. In: *Analytic number theory*, pages 359–388. Cambridge Univ. Press, Cambridge, 2009.
85. E. Novak and H. Woźniakowski: *Tractability of Multivariate Problems. Volume II: Standard Information for Functionals*. European Mathematical Society Publishing House, Zürch, 2010.
86. D. Nuyens and R. Cools: Fast algorithms for component-by-component construction of rank-1 lattice rules in shift-invariant reproducing kernel Hilbert spaces. Math. Comp., 75 903–920, 2006.
87. D. Nuyens and R. Cools: Fast component-by-component construction of rank-1 lattice rules with a non-prime number of points. J. Complexity, 22, 4–28, 2006.
88. D. Nuyens and R. Cools: Fast component-by-component construction, a reprise for different kernels. In: *Monte Carlo and Quasi-Monte Carlo Methods 2004* (H. Niederreiter and D. Talay, eds.), pp. 373–387. Springer, Berlin, 2006.
89. A. B. Owen: Randomly permuted (t, m, s)-nets and (t, s)-sequences. In: *Monte Carlo and quasi-Monte Carlo methods in scientific computing* (H. Niederreiter and J.-S. Shiue, eds.), 299–317, Lecture Notes in Statist., 106, Springer, New York, 1995.
90. A. B. Owen: Monte Carlo variance of scrambled net quadrature. SIAM J. Numer. Anal., 34, 1884–1910, 1997.
91. A. B. Owen: Scrambled net variance for integrals of smooth functions. Ann. Statist., 25, 1541–1562, 1997.
92. F. Pillichshammer: Polynomial lattice point sets. In: *Monte Carlo and Quasi-Monte Carlo Methods 2010* (L. Plaskota and H. Woźniakowski, eds.), Springer, Berlin, 2011. (to appear)

93. A. P. Prudnikov, Yu. A. Brychkov and I. O. Marichev: Integrals and series. Vol. 1. Elementary functions. Translated from the Russian and with a preface by N. M. Queen. Gordon & Breach Science Publishers, New York, 1986.
94. K. F. Roth: On irregularities of distribution. Mathematika, 1, 73–79, 1954.
95. K. F. Roth: On irregularities of distribution III. Acta Arith., 35, 373–384, 1979.
96. K. F. Roth: On irregularities of distribution IV. Acta Arith., 37, 67–75, 1980.
97. I. F. Šarygin: A lower estimate for the error of quadrature formulas for certain classes of functions. Zh. Vychisl. Mat. i Mat. Fiz., 3, 370–376, 1963. (Russian).
98. W. M. Schmidt: Irregularities of distribution VII. Acta Arith., 21, 45–50, 1972.
99. W. M. Schmidt: Irregularities of distribution X. In: *Number theory and algebra*, pp. 311–329. Academic Press, New York, 1977.
100. V. Sinescu and S. Joe: Good lattice rules with a composite number of points based on the product weighted star-discrepancy. Monte Carlo and quasi-Monte Carlo methods 2006, 645–658, Springer, Berlin, 2008.
101. V. Sinescu and P. L'Ecuyer: On the behaviour of the weighted star-discrepancy bounds for shifted latice rules, in Monte Carlo and quasi-Monte Carlo Methods 2008, P. L'Ecuyer and A. B. Owen (eds.), Springer, Berlin, 603–616, 2009.
102. M. M. Skriganov: Harmonic analysis on totally disconnected groups and irregularities of point distributions. J. Reine Angew. Math., 600, 25–49, 2006.
103. I. H. Sloan and S. Joe: *Lattice Methods for Multiple Integration*. Clarendon Press, Oxford, 1994.
104. I. H. Sloan, F. Y. Kuo and S. Joe: On the step-by-step construction of quasi-Monte Carlo integration rules that achieve strong tractability error bounds in weighted Sobolev spaces. Math. Comp., 71, 240, 1609–1640, 2002.
105. I. H. Sloan, F. Y. Kuo, and S. Joe: Constructing randomly shifted lattice rules in weighted Sobolev spaces. SIAM J. Numer. Anal., 40, 1650–1665, 2002.
106. I. H. Sloan and A. V. Reztsov: Component-by-component construction of good lattice rules. Math. Comp., 71, 263–273, 2002.
107. I. H. Sloan and H. Woźniakowski: When are quasi-Monte Carlo algorithms efficient for high-dimensional integrals? J. Complexity, 14, 1–33, 1998.
108. I. H. Sloan and H. Woźniakowski: Tractability of multivariate integration for weighted Korobov classes. J. Complexity, 17, 697–721, 2001.
109. I. M. Sobol: Distribution of points in a cube and approximate evaluation of integrals, Z. Vyčisl. Mat. i Mat. Fiz., Akademija Nauk SSSR. Žurnal Vyčislitel noĭ Matematiki i Matematičeskoĭ Fiziki, 7, 784–802, 1967.
110. K. B. Stolarsky: Sums of distances between points on a sphere. II. Proc. Amer. Math. Soc., 41, 575–582, 1973.
111. V. N. Temlyakov: Cubature formulas, discrepancy, and nonlinear approximation. Numerical integration and its complexity. J. Complexity, 19, 352–391, 2003.
112. R. R. Varshamov: Estimate of the number of signals in error correcting codes. Dokl. Acad. Nauk SSSR, 117, 739–741, 1957.
113. H. Woźniakowski: Efficiency of quasi-Monte Carlo algorithms for high dimensional integrals. In: *Monte Carlo and Quasi-Monte Carlo Methods 1998* (H. Niederreiter and J. Spanier, eds.), pp. 114–136. Springer, Berlin, 2000.
114. S. K. Zaremba: Some applications of multidimensional integration by parts. Ann. Poln. Math., 21, 85–96, 1968.