The tent transformation can improve the convergence rate of quasi-Monte Carlo algorithms using digital nets

Ligia L. Cristea*, Josef Dick†, Gunther Leobacher
and Friedrich Pillichshammer‡

Abstract

In this paper we investigate multivariate integration in reproducing kernel Sobolev spaces for which the second partial derivatives are square integrable. As quadrature points for our quasi-Monte Carlo algorithm we use digital (t, m, s)-nets over $\mathbb{Z}_2$ which are randomly digitally shifted and then folded using the tent transformation. For this QMC algorithm we show that the root mean square worst-case error converges with order $2^{m(-2+\varepsilon)}$ for any $\varepsilon > 0$, where 2^m is the number of points. A similar result for lattice rules has previously been shown by Hickernell.

Keywords: Digital nets, tent transformation, randomized QMC.
MSC 2000: 11K38, 11K06

1 Introduction

In order to approximate integrals of the form $I(h) = \int_{[0,1]^s} h(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ one often employs a quasi-Monte Carlo (QMC) algorithm $Q_{N,s}(h) = N^{-1}\sum_{n=0}^{N-1} h(\boldsymbol{x}_n)$, where the points $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1} \in [0,1]^s$ are carefully chosen quadrature points. Especially in higher dimensions one either chooses the quadrature points randomly (in which case $Q_{N,s}$ is called Monte Carlo algorithm) or deterministically. The two most prominent QMC rules are lattice rules (see [10, 16]) and QMC rules using digital (t, m, s)-nets as point sets (see [10]). Here we will focus on the latter one.

The advantage of using deterministic point sets lies in the fact that if the function satisfies some smoothness conditions one can obtain a better convergence rate of the integration error. It is known that Monte Carlo algorithms typically achieve a convergence rate of the integration error of order $N^{-1/2}$ (see [10] for example). In this case one only needs to assume that the function has finite variance. But by considering smoother function classes and using deterministic point sets one can obtain better convergence rates. For example if we consider periodic functions for which all the partial derivatives up to order α in each variable exist and are square integrable we can obtain a convergence order of $N^{-\alpha+\varepsilon}$ for any $\varepsilon > 0$, see [16]. This convergence rate is known to be best possible. The quadrature rules used in this case are lattice rules and the analysis is based on the

*Supported by the Austrian Research Foundation (FWF), Projects P 17022-N12 and S9609.
†Supported by the Australian Research Council under its Center of Excellence Program.
‡Supported by the Austrian Research Foundation (FWF), Projects S8305, P 17022-N12 and S9609.

Fourier series of the functions. (A lattice rule is a QMC rule which uses the point set $\{n\boldsymbol{g}/N\}$ for $n = 0, \ldots, N-1$, where $\boldsymbol{g} \in \{1, \ldots, N-1\}^s$ is the generating vector and for a real number x we define $\{x\} = x - \lfloor x \rfloor = x \pmod 1$.)

For function spaces which do not carry the assumption that the functions must be periodic the problem is much more difficult. If we assume that all partial derivatives of order up to one in each variable exist and are square integrable then one can use randomly shifted lattice rules to do the job, that is, to obtain a convergence rate of order $N^{-1+\varepsilon}$ for any $\varepsilon > 0$, see [18]. (If $\boldsymbol{x}_n$, $n = 0, \ldots, N-1$, is the point set used in a lattice rule, then a shifted lattice rule uses the point set $\{\boldsymbol{x} + \boldsymbol{\Delta}\}$ where $\boldsymbol{\Delta} \in [0,1)^s$. We speak of a randomly shifted lattice rule if the shift $\boldsymbol{\Delta}$ is chosen randomly.)

If we assume that all mixed partial derivatives of order up to 2 in each variable exist and are square integrable then randomly shifted lattice rules do not achieve a convergence order $N^{-2+\varepsilon}$, instead one obtains only a convergence order of $N^{-1+\varepsilon}$, see [6]. For such function classes one can of course use some transformation to make the functions periodic and then apply the theorems for periodic functions which then yields the better convergence rate of $N^{-2+\varepsilon}$, but as described in [6], this can increase the norm of the function. Hickernell [6] suggested to look at this problem differently. Namely, instead of periodizing the function one can transform the point set, that is, change the algorithm, and analyse the discrepancy of the new point set. This way, Hickernell showed that lattice rules achieve the best convergence order of $N^{-2+\varepsilon}$. The point set he used can be obtained in the following way: first one applies a random shift modulo 1 to the lattice rules and then uses the so-called tent transformation. This transformation is given by the function $\phi(x) = 1 - |2x - 1|$ (in [6] Hickernell called this transformation "baker's transformation"). So in effect the point set employed in his QMC algorithm is given by $\phi(\{\boldsymbol{x}_n + \boldsymbol{\Delta}\})$ for $n = 0, \ldots, N-1$, where $\boldsymbol{x}_n = \{n\boldsymbol{g}/N\}$ and $\boldsymbol{\Delta} \in [0,1)^s$ is chosen i.i.d. and the function ϕ is applied to each coordinate, see [6].

In the lattice rule case the shift-and-fold-invariant kernel [6] can be related to a reproducing kernel for a space of twice differentiable periodic functions. This then permits to obtain the improved convergence rate. When one uses digital nets on the other hand, no such relationship can be expected and hence the results in the following come as a surprise. The secret lies in the behavior of the Walsh coefficients of the digitally shifted-and-folded kernel, which is different from the one for the Fourier coefficients in the lattice rule case, but it still allows us to obtain the higher convergence rate. More precisely, we will show that similar results (compared to the one for lattice rules) hold for digital nets and polynomial lattice rules over $\mathbb{Z}_2$. If the functions have partial derivatives up to order 2 which are square integrable then we can obtain a convergence rate of order $2^{m(-2+\varepsilon)}$ for any $\varepsilon > 0$. As for lattice rules, our result is obtained by averaging over all generating matrices of digital nets or generating vectors of polynomial lattice rules. As opposed to lattice rules, for digital nets a priori constructions with good distribution properties are known, for example digital nets stemming from Sobol sequences, Niederreiter sequences or Niederreiter-Xing sequences. If such point sets satisfy similar upper bounds is not known and remains an interesting open problem.

We conclude the introduction with a brief outline of the paper. In the next section we introduce digital nets, polynomial lattice rules, digital shifts and Walsh functions. These definitions and some basic results provide us with the main tools to analyze and prove upper bounds for the integration error in certain reproducing kernel Sobolev spaces. In Section 3 we analyse the worst-case error for multivariate integration in reproducing kernel

Hilbert spaces using randomly digitally shifted and then folded point sets. We prove a formula for the mean square worst-case error in such spaces using the Walsh-coefficients of a certain related reproducing kernel. In Section 4 we use those results for the specific case of certain Sobolev spaces and we obtain an upper bound on the average of the mean square worst-case error over all digital nets and all polynomial lattice rules. In Section 5 we present numerical experiments. Appendix A contains the calculation of the shifted and folded kernel, whereas in Appendix B we show a closed form of the shift invariant kernel, which is related to the shifted and folded kernel. Appendix C contains several useful lemmas and finally Appendix D contains several pictures of digitally shifted and then folded digital nets.

2 Preliminary definitions and results

2.1 Digital nets

The construction method considered here builds on the concept of (t,m,s)-nets. A detailed theory on this topic was developed in Niederreiter [9] (see also [10, Chapter 4], for a recent survey see [13]). Those (t,m,s)-nets in base b provide sets of b^m points in the half open s-dimensional unit-cube, which are extremely well distributed if the quality parameter t is "small". The details are given in the following definition.

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

Note that any point set consisting of b^m points in $[0,1)^s$ is at least a (m,m,s)-net in base b. Of course, we wish to have a small value of the quality parameter t. Unfortunately the optimal value $t=0$ is not possible for all parameters $s \geq 1$ and $b \geq 2$. Niederreiter [9] proved that if a $(0,m,s)$-net in base b exists, then we have $s-1 \leq b$. Faure [5] provided a construction of $(0,m,s)$-nets in prime base $p \geq s-1$ and Niederreiter [9] extended Faure's construction to prime power bases $p^r \geq s-1$. So for example a $(0,m,s)$-net in base 2 only exists if $s=1$, $s=2$ or $s=3$.

In practice all concrete constructions of (t,m,s)-nets in a base b are based on a general construction scheme which is the concept of digital point sets. Here in this paper we only deal with the case $b=2$, i.e., we only consider (t,m,s)-nets in base 2 and hence we introduce the digital construction only for this special case. For a general definition see for example [10]. (It has been observed that a small base b and higher t-value yields better point sets than choosing a high base b such that we can achieve $t=0$. It appears therefore that the case $b=2$ might actually be the most important one.) In the following let $\mathbb{Z}_2$ denote the finite field with two elements.

Definition 2 *Let $s,m \geq 1$ be integers and choose s $m \times m$ matrices $C_1,\ldots,C_s$ over $\mathbb{Z}_2$. Consider the following construction principle for point sets consisting of 2^m points in $[0,1)^s$: represent n, $0 \leq n < 2^m$, in base 2, $n = n_0 + n_1 2 + \cdots + n_{m-1}2^{m-1}$, and multiply the matrix C_j, $1 \leq j \leq s$, with the vector $\vec{n} = (n_0,\ldots,n_{m-1})^\top$ of digits of n in $\mathbb{Z}_2$,*

$$C_j \vec{n} =: (x_{j,n,1},\ldots,x_{j,n,m})^\top.$$

Now we set

$$x_{j,n} := \frac{x_{j,n,1}}{2} + \cdots + \frac{x_{j,n,m}}{2^m}$$

and

$$\boldsymbol{x}_n = (x_{1,n}, \ldots, x_{s,n}).$$

If for some integer t with $0 \leq t \leq m$ the point set $\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{2^m-1}\}$ is a (t,m,s)-net in base 2, then it is called a digital (t,m,s)-net over $\mathbb{Z}_2$ (or briefly a digital net) and the matrices $C_1, \ldots, C_s$ are called the generating matrices of the digital net.

Concerning the determination of the quality parameter t of digital nets we refer to Niederreiter [10, Theorem 4.28]. It is well known that any d-dimensional projection of a digital (t,m,s)-net over $\mathbb{Z}_2$ is a digital (t,m,d)-net over $\mathbb{Z}_2$.

In the following we show that digital nets also have a group structure. Let $\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{2^m-1}\}$ be a digital net over $\mathbb{Z}_2$ generated by the $m \times m$ matrices $C_1, \ldots, C_s$ over $\mathbb{Z}_2$. For $\boldsymbol{x}_n = (x_{1,n}, \ldots, x_{s,n})$ and $x_{j,n} = x_{j,n,1}2^{-1} + \cdots + x_{j,n,m}2^{-m}$, $1 \leq j \leq s$, $0 \leq n < 2^m$, we identify $\boldsymbol{x}_n$ with

$$(x_{1,n,1}, \ldots, x_{1,n,m}, \ldots, x_{s,n,1}, \ldots, x_{s,n,m}) \in \mathbb{Z}_2^{ms}$$

and we define

$$\boldsymbol{x}_n \oplus \boldsymbol{x}_h := (x_{1,n,1} + x_{1,h,1}, \ldots, x_{s,n,m} + x_{s,h,m}) \in \mathbb{Z}_2^{ms}. \tag{1}$$

The subsequent lemma follows easily from the construction of digital nets.

Lemma 1 *Any digital net $\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{2^m-1}\}$ over $\mathbb{Z}_2$ is homomorph to a subgroup of $(\mathbb{Z}_2^{ms}, \oplus)$.*

2.2 Polynomial lattice rules

In [11] (see also [10]) Niederreiter introduced a special construction of digital nets. This construction is based on rational functions over finite fields. Again we restrict ourselves to digital nets over $\mathbb{Z}_2$.

Let $\mathbb{Z}_2((x^{-1}))$ be the field of formal Laurent series over $\mathbb{Z}_2$. Thus elements of $\mathbb{Z}_2((x^{-1}))$ are of the form

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

where w is an arbitrary integer and all $t_l \in \mathbb{Z}_2$. Note that $\mathbb{Z}_2((x^{-1}))$ contains the field of rational functions over $\mathbb{Z}_2$ as a subfield. Further let $\mathbb{Z}_2[x]$ be the set of all polynomials over $\mathbb{Z}_2$ and let $m \geq 1$ be an integer. For a given dimension $s \geq 2$, choose $f \in \mathbb{Z}_2[x]$, with $\deg(f) = m$, and let $g_1, \ldots, g_s \in \mathbb{Z}_2[x]$. Let φ_m be the map from $\mathbb{Z}_2((x^{-1}))$ to the interval $[0,1)$ defined by

$$\varphi_m\left(\sum_{l=w}^{\infty} t_l x^{-l}\right) = \sum_{l=\max(1,w)}^{m} t_l 2^{-l}.$$

For $0 \leq n < 2^m$ let $n = n_0 + n_1 2 + \cdots + n_{m-1}2^{m-1}$ be the base 2 expansion of n. With each such n we associate the polynomial

$$n(x) = \sum_{r=0}^{m-1} n_r x^r \in \mathbb{Z}_2[x].$$

Then $P(\boldsymbol{g}, f)$ is the point set consisting of the 2^m points

$$\boldsymbol{x}_n = \left(\varphi_m\left(\frac{n(x)g_1(x)}{f(x)}\right), \ldots, \varphi_m\left(\frac{n(x)g_s(x)}{f(x)}\right)\right) \in [0,1)^s,$$

for $0 \le n \le 2^m - 1$. Due to the construction principle, a QMC rule using the point set $P(\boldsymbol{g}, f)$ is often called a polynomial lattice rule. The vector $\boldsymbol{g}$ is called the generating vector of $P(\boldsymbol{g}, f)$ or the generating vector of the polynomial lattice rule, depending on the context. For $1 \le j \le s$ consider the expansion

$$\frac{g_j(x)}{f(x)} = \sum_{l=w_j}^{\infty} u_l^{(j)} x^{-l} \in \mathbb{Z}_2((x^{-1})),$$

where $w_j \le 1$. Then it can be shown (see [12]) that the point set $P(\boldsymbol{g}, f)$ is a digital net with generating matrices $C_1, \ldots, C_s$ where the elements $c_{i,r}^{(j)}$ of the matrix C_j are given by

$$c_{i,r}^{(j)} = u_{r+i}^{(j)} \in \mathbb{Z}_2$$

for $1 \le j \le s$, $1 \le i \le m$ and $0 \le r \le m-1$

2.3 Digital shift

For practical applications it is often useful to have a random element in the point set used (see [8]). On the other hand we wish to preserve the structure which a point set already has. That is, in this case we wish to randomize a (t, m, s)-net such that the resulting point set is again a (t, m, s)-net with the same quality parameter t. Several randomization methods for (t, m, s)-nets have been introduced (see [8], [14], [20]).

In this paper we consider point sets which are digitally shifted. First we introduce some notation which we will use frequently throughout the paper. By $\oplus$ we denote the digit-wise addition modulo 2 , i.e., for $x = \sum_{i=w}^{\infty} \frac{x_i}{2^i}$ and $y = \sum_{i=w}^{\infty} \frac{y_i}{2^i}$ we have

$$x \oplus y := \sum_{i=w}^{\infty} \frac{z_i}{2^i}, \quad \text{where} \quad z_i := x_i + y_i \pmod 2.$$

For vectors $\boldsymbol{x} = (x_1, \ldots, x_s)$ and $\boldsymbol{y} = (y_1, \ldots, y_s)$ we define the digit-wise addition modulo 2 coordinate-wise, i.e. $\boldsymbol{x} \oplus \boldsymbol{y} = (x_1 \oplus y_1, \ldots, x_s \oplus y_s)$.

Let $\boldsymbol{\sigma} \in [0,1)^s$ and let $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1} \in [0,1)^s$. The digitally shifted point set $\boldsymbol{y}_0, \ldots, \boldsymbol{y}_{N-1} \in [0,1)^s$ is then given by $\boldsymbol{y}_n = \boldsymbol{x}_n \oplus \boldsymbol{\sigma}$ for $n = 0, \ldots, N-1$. This randomization method is the digital analogue of the shift used for lattice rules.

The point set which we consider in this paper are digital nets over $\mathbb{Z}_2$, which are randomly shifted by a digital shift, with the shift $\boldsymbol{\sigma}$ i.i.d. in $[0,1)^s$, and then folded using the tent transformation. More precisely, let $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{2^m-1}$ be a digital net over $\mathbb{Z}_2$ and $\boldsymbol{\sigma} \in [0,1)^s$ be an i.i.d. random number then the randomly digitally shifted and then folded point set is given by

$$\boldsymbol{y}_n = \phi(\boldsymbol{x}_n \oplus \boldsymbol{\sigma}) \quad \text{for } n = 0, \ldots, 2^m - 1,$$

where $\phi(\boldsymbol{x}_n \oplus \boldsymbol{\sigma}) = (\phi(x_{1,n} \oplus \sigma_1), \ldots, \phi(x_{s,n} \oplus \sigma_s))$.

2.4 Walsh functions

In this section we introduce Walsh functions, which will be the main tool in our analysis. Again we confine ourselves to base 2 (for more information see [1], [19]). In the following let $\mathbb{N}_0$ denote the set of non-negative integers.

Definition 3 For a non-negative integer k with base 2 representation

$$k = \kappa_{a-1} 2^{a-1} + \cdots + \kappa_1 2 + \kappa_0,$$

with $\kappa_i \in \{0,1\}$, we define the Walsh function $\mathrm{wal}_k : [0,1) \longrightarrow \{-1,1\}$ by

$$\mathrm{wal}_k(x) := (-1)^{x_1 \kappa_0 + \cdots + x_a \kappa_{a-1}},$$

for $x \in [0,1)$ with base 2 representation $x = \frac{x_1}{2} + \frac{x_2}{2^2} + \cdots$ (unique in the sense that infinitely many of the x_i must be zero).

Definition 4 For dimension $s \geq 2$, $x_1, \ldots, x_s \in [0,1)$ and $k_1, \ldots, k_s \in \mathbb{N}_0$ we define $\mathrm{wal}_{k_1,\ldots,k_s} : [0,1)^s \longrightarrow \{-1,1\}$ by

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

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

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

We call $x \in [0,1)$ a dyadic rational if x can be represented by a finite base 2 expansion. In the following proposition we summarize some basic properties of Walsh functions.

Proposition 1

1. *For all $k, l \in \mathbb{N}_0$ and all $x, y \in [0,1)$, with the restriction that if x, y are not dyadic rationals then $x \oplus y$ is not allowed to be a dyadic rational, we have*

$$\mathrm{wal}_k(x) \cdot \mathrm{wal}_l(x) = \mathrm{wal}_{k \oplus l}(x), \quad \mathrm{wal}_k(x) \cdot \mathrm{wal}_k(y) = \mathrm{wal}_k(x \oplus y).$$

2. *We have*

$$\int_0^1 \mathrm{wal}_0(x)\,\mathrm{d}x = 1 \quad \text{and} \quad \int_0^1 \mathrm{wal}_k(x)\,\mathrm{d}x = 0 \text{ if } k > 0.$$

3. *For all $\boldsymbol{k}, \boldsymbol{l} \in \mathbb{N}_0^s$ we have the following orthogonality properties:*

$$\int_{[0,1]^s} \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}) \mathrm{wal}_{\boldsymbol{l}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \begin{cases} 1 & \text{if } \boldsymbol{k} = \boldsymbol{l}, \\ 0 & \text{otherwise.} \end{cases}$$

4. *For any $h \in \mathcal{L}_2([0,1)^s)$ and any $\boldsymbol{\sigma} \in [0,1)^s$ we have*

$$\int_{[0,1)^s} h(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{[0,1)^s} h(\boldsymbol{x} \oplus \boldsymbol{\sigma})\,\mathrm{d}\boldsymbol{x}.$$

5. *For any integer $s \geq 1$ the system $\{\mathrm{wal}_{k_1,\ldots,k_s} : k_1, \ldots, k_s \geq 0\}$ is a complete orthonormal system in $\mathcal{L}_2([0,1)^s)$.*

Proof. The proofs of 1.-3. are straightforward. For item 4. see [1, Lemma 1] and for item 5. see [1]. $\square$

The following lemma will be very useful for our investigation. A proof of this result can be found in [3, 4, 15].

Lemma 2 *Let $\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{2^m-1}\}$ be a digital (t, m, s)-net over $\mathbb{Z}_2$ generated by the $m \times m$ matrices $C_1, \ldots, C_s$ over $\mathbb{Z}_2$. Then for all integers $0 \le k_1, \ldots, k_s < 2^m$ we have*

$$\sum_{n=0}^{2^m-1} \mathrm{wal}_{k_1,\ldots,k_s}(\boldsymbol{x}_n) = \begin{cases} 2^m & \text{if } C_1^\top \vec{k}_1 + \cdots + C_s^\top \vec{k}_s = \vec{0}, \\ 0 & \text{otherwise,} \end{cases}$$

where for $0 \le k < 2^m$ with $k = \kappa_0 + \kappa_1 2 + \cdots + \kappa_{m-1} 2^{m-1}$ we write $\vec{k} = (\kappa_0, \ldots, \kappa_{m-1})^\top \in \mathbb{Z}_2^m$ and $\vec{0}$ denotes the zero vector in $\mathbb{Z}_2^m$.

3 Worst-case error in reproducing kernel Hilbert spaces using randomly shifted and then folded point sets

In general we are interested in approximating the integrals of functions h from a reproducing kernel Hilbert space H,

$$I_s(h) = \int_{[0,1]^s} h(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}.$$

We approximate the integral $I_s(h)$ by QMC algorithms, which are equal weight quadrature rules of the form

$$Q_{N,s}(h) = \frac{1}{N} \sum_{n=0}^{N-1} h(\boldsymbol{x}_n), \tag{2}$$

with a deterministically chosen point set $P_N = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}\} \subseteq [0,1]^s$.

The worst-case error of a QMC rule $Q_{N,s}$ for integration in the space H with reproducing kernel K and norm $\|\cdot\|$ is defined by

$$e(P_N, K) := \sup_{h \in H, \|h\| \le 1} |I_s(h) - Q_{N,s}(h)|,$$

and the initial error is

$$e_{0,s} := \sup_{h \in H, \|h\| \le 1} |I_s(h)|.$$

It is known that (see for example [17]) the worst-case error for multivariate integration in a reproducing kernel Hilbert space with reproducing kernel K using a point set P_N is given by

$$e^2(P_N, K) = \int_{[0,1]^{2s}} K(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{y} - \frac{2}{N} \sum_{\boldsymbol{x} \in P_N} \int_{[0,1]^s} K(\boldsymbol{x}, \boldsymbol{y}) \, \mathrm{d}\boldsymbol{y} + \frac{1}{N^2} \sum_{\boldsymbol{x}, \boldsymbol{y} \in P_N} K(\boldsymbol{x}, \boldsymbol{y}). \tag{3}$$

Let P_N be a point set consisting of N points in $[0,1]^s$ and let $\widetilde{P}_{N,\boldsymbol{\sigma},\phi}$ denote the point set P_N which is first randomized by a digital shift $\boldsymbol{\sigma}$ and then folded component wise by

ϕ. Define the mean square worst-case error for integration in a reproducing kernel Hilbert space with reproducing kernel K by

$$\widehat{e}^2_{\mathrm{sh},\phi}(P_N, K) := \mathbb{E}[e^2(\widetilde{P}_{N,\boldsymbol{\sigma},\phi}, K)], \tag{4}$$

where the expectation value is with respect to the random digital shift $\boldsymbol{\sigma}$. Further, for a reproducing kernel $K \in \mathcal{L}_2([0,1]^{2s})$ let the digitally shifted and folded kernel $K_{\mathrm{sh},\phi}$ be given by

$$K_{\mathrm{sh},\phi}(\boldsymbol{x}, \boldsymbol{y}) = \int_{[0,1]^s} K(\phi(\boldsymbol{x} \oplus \boldsymbol{\sigma}), \phi(\boldsymbol{y} \oplus \boldsymbol{\sigma}))\,\mathrm{d}\boldsymbol{\sigma},$$

where for vectors $\boldsymbol{x} \in \mathbb{R}^s$ the tent transformation $\phi(\boldsymbol{x})$ is defined component wise.

As done by Hickernell in [6] we can apply the random digital shift and the tent transformation to the reproducing kernel and then use (3) to obtain the mean square worst-case error as shown in the following theorem.

Theorem 1 *Let $K \in \mathcal{L}_2([0,1]^{2s})$ be a reproducing kernel, let $P_N \subseteq [0,1]^s$ be a point set consisting of N points and let $\widetilde{P}_{N,\boldsymbol{\sigma},\phi}$ be the point set P_N first randomized by a digital shift and then transformed by ϕ. Then we have*

$$\widehat{e}^2_{\mathrm{sh},\phi}(P_N, K) = e^2(P_N, K_{\mathrm{sh},\phi}).$$

Proof. In order to calculate $\widehat{e}^2_{\mathrm{sh},\phi}(P_N, K)$ we use (3) and

$$\mathbb{E}[e^2(\widetilde{P}_{N,\boldsymbol{\sigma},\phi})] = \int_{[0,1]^s} e^2(\widetilde{P}_{N,\boldsymbol{\sigma},\phi}, K)\,\mathrm{d}\boldsymbol{\sigma}.$$

Note that for any $h \in \mathcal{L}_2([0,1])$ we have

$$\int_0^1 h(\phi(x))\,\mathrm{d}x = \int_0^1 h(x)\,\mathrm{d}x.$$

Hence together with Proposition 1 it follows that

$$\begin{aligned}\int_{[0,1]^s}\int_{[0,1]^{2s}} K(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{\sigma} &= \int_{[0,1]^s}\int_{[0,1]^{2s}} K(\phi(\boldsymbol{x}\oplus\boldsymbol{\sigma}),\phi(\boldsymbol{y}\oplus\boldsymbol{\sigma}))\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{\sigma} \\ &= \int_{[0,1]^{2s}} K_{\mathrm{sh},\phi}(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y}.\end{aligned}$$

In the same way, we obtain

$$\begin{aligned}\int_{[0,1]^s}\frac{2}{N}\sum_{\boldsymbol{x}\in\widetilde{P}_{N,\boldsymbol{\sigma},\phi}}\int_{[0,1]^s} K(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{\sigma} &= \int_{[0,1]^s}\frac{2}{N}\sum_{\boldsymbol{x}\in P_N}\int_{[0,1]^s} K(\phi(\boldsymbol{x}\oplus\boldsymbol{\sigma}),\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{\sigma} \\ &= \frac{2}{N}\sum_{\boldsymbol{x}\in P_N}\int_{[0,1]^s} K_{\mathrm{sh},\phi}(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}.\end{aligned}$$

For the right-hand sum in (3) we obtain

$$\begin{aligned}\int_{[0,1]^s}\frac{1}{N^2}\sum_{\boldsymbol{x},\boldsymbol{y}\in\widetilde{P}_{N,\boldsymbol{\sigma},\phi}} K(\boldsymbol{x},\boldsymbol{y})\,\mathrm{d}\boldsymbol{\sigma} &= \frac{1}{N^2}\sum_{\boldsymbol{x},\boldsymbol{y}\in P_N}\int_{[0,1]^s} K(\phi(\boldsymbol{x}\oplus\boldsymbol{\sigma}),\phi(\boldsymbol{y}\oplus\boldsymbol{\sigma}))\,\mathrm{d}\boldsymbol{\sigma} \\ &= \frac{1}{N^2}\sum_{\boldsymbol{x},\boldsymbol{y}\in P_N} K_{\mathrm{sh},\phi}(\boldsymbol{x},\boldsymbol{y}).\end{aligned}$$

The result now follows from (3) and (4). □

The above theorem shows that the mean square worst-case error can be obtained via (3) and the digitally shifted and folded kernel. We will show a theorem how this kernel can be calculated using Walsh functions. Before we state this result we introduce the following function: for $k = \kappa_0 + \kappa_1 2 + \kappa_2 2^2 + \cdots$ denote the sum-of-digits function as $\sigma(k) = \kappa_0 + \kappa_1 + \kappa_2 + \cdots$. Further, the following lemma will be required for the promised result.

Lemma 3 *Let $k = \kappa_0 + \kappa_1 2 + \cdots + \kappa_{a-1} 2^{a-1}$ with $\kappa_{a-1} \neq 0$. Let the base 2 representation of $x \in [0,1)$ be given by $x = \frac{x_1}{2} + \frac{x_2}{2^2} + \cdots$ where we assume that $x_i = 0$ for infinitely many $i \in \mathbb{N}$. Further let $x_e = 1$ for some $e \geq a+1$. Then we have*

$$\mathrm{wal}_k(1-x) = (-1)^{\sigma(k)} \mathrm{wal}_k(x).$$

Proof. We have

$$1 - x = \frac{1-x_1}{2} + \cdots + \frac{1-x_a}{2^a} + \cdots$$

and therefore

$$\begin{aligned} \mathrm{wal}_k(1-x) &= (-1)^{\kappa_0(1-x_1)+\cdots+\kappa_{a-1}(1-x_a)} \\ &= (-1)^{\kappa_0+\cdots+\kappa_{a-1}} (-1)^{\kappa_0 x_1+\cdots+\kappa_{a-1}x_a} = (-1)^{\sigma(k)} \mathrm{wal}_k(x). \end{aligned}$$

□

The following theorem is useful for the calculation of the digitally shifted and folded kernel.

Theorem 2 *Let $K \in \mathcal{L}_2([0,1]^{2s})$ be a reproducing kernel. Then the digitally shifted and then folded kernel $K_{\mathrm{sh},\phi}$ is given by*

$$K_{\mathrm{sh},\phi}(\boldsymbol{x},\boldsymbol{y}) = \sum_{\boldsymbol{k}\in\mathcal{E}} \hat{K}_{\mathrm{sh},\phi}(\boldsymbol{k},\boldsymbol{k}) \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}) \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{y}),$$

where $\mathcal{E} = \{\boldsymbol{k} = (k_1,\ldots,k_s) \in \mathbb{N}_0^s : \sigma(k_j) \equiv 0 \pmod 2$ for all $j = 1,\ldots,s\}$ and

$$\hat{K}_{\mathrm{sh},\phi}(\boldsymbol{k},\boldsymbol{k}) = \int_{[0,1]^s}\int_{[0,1]^s} K(\boldsymbol{x},\boldsymbol{y}) \mathrm{wal}_{\lfloor \boldsymbol{k}/2 \rfloor}(\boldsymbol{x}) \mathrm{wal}_{\lfloor \boldsymbol{k}/2 \rfloor}(\boldsymbol{y}) \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y}$$

with $\lfloor \boldsymbol{k}/2 \rfloor = (\lfloor k_1/2 \rfloor,\ldots,\lfloor k_s/2 \rfloor)$ and for $k_j = \kappa_{j,0} + 2\kappa_{j,1} + \cdots$ we have $\lfloor k_j/2 \rfloor = \kappa_{j,1} + 2\kappa_{j,1} + \cdots$.

Remark 1 Note that $\lfloor k_j/2 \rfloor$ is independent of $\kappa_{j,0}$, but the condition $\sigma(k_j) \equiv 0 \pmod 2$ implies that for given $\kappa_{j,1}, \kappa_{j,2}, \ldots$ there is only one choice for $\kappa_{j,0}$ such that $\sigma(k_j) \equiv 0 \pmod 2$. Further it is obvious that for any choice of $\kappa_{j,1}, \kappa_{j,2}, \ldots$ there is exactly one $\kappa_{j,0} \in \{0,1\}$ such that $\sigma(k_j) \equiv 0 \pmod 2$.

Remark 2 Let the digital shift invariant kernel, as defined in [3], be given by

$$K_{\mathrm{sh}}(\boldsymbol{x},\boldsymbol{y}) = \int_{[0,1]^s} K(\boldsymbol{x} \oplus \boldsymbol{\sigma}, \boldsymbol{y} \oplus \boldsymbol{\sigma}) \,\mathrm{d}\boldsymbol{\sigma},$$

then we can write

$$K_{\rm sh}(\boldsymbol{x},\boldsymbol{y}) = \sum_{\boldsymbol{k}\in\mathbb{N}_0^s} \hat{K}_{\rm sh}(\boldsymbol{k},\boldsymbol{k}) {\rm wal}_{\boldsymbol{k}}(\boldsymbol{x}) {\rm wal}_{\boldsymbol{k}}(\boldsymbol{y}),$$

where

$$\hat{K}_{\rm sh}(\boldsymbol{k},\boldsymbol{k}) = \int_{[0,1]^s}\int_{[0,1]^s} K(\boldsymbol{x},\boldsymbol{y}) {\rm wal}_{\boldsymbol{k}}(\boldsymbol{x}) {\rm wal}_{\boldsymbol{k}}(\boldsymbol{y}) \,{\rm d}\boldsymbol{x}\,{\rm d}\boldsymbol{y}.$$

Let now $\nu(\boldsymbol{k}) = (\nu(k_1),\ldots,\nu(k_s)) \in \{0,1\}^s$ be defined by $\nu(k_j) \equiv \kappa_{j,0}+\kappa_{j,1}+\cdots \pmod 2$ for $j=1,\ldots,s$. It then follows from the above theorem that $\hat{K}_{\rm sh}(\boldsymbol{k},\boldsymbol{k}) = \hat{K}_{{\rm sh},\phi}(2\boldsymbol{k}+\nu(\boldsymbol{k}), 2\boldsymbol{k}+\nu(\boldsymbol{k}))$ and hence we can write

$$K_{\rm sh}(\boldsymbol{x},\boldsymbol{y}) = \sum_{\boldsymbol{k}\in\mathbb{N}_0^s} \hat{K}_{{\rm sh},\phi}(2\boldsymbol{k}+\nu(\boldsymbol{k}), 2\boldsymbol{k}+\nu(\boldsymbol{k})) {\rm wal}_{\boldsymbol{k}}(\boldsymbol{x}) {\rm wal}_{\boldsymbol{k}}(\boldsymbol{y}).$$

We can of course also do the opposite and write

$$\begin{aligned} K_{{\rm sh},\phi}(\boldsymbol{x},\boldsymbol{y}) &= \sum_{\boldsymbol{k}\in\mathbb{N}_0^s} \hat{K}_{\rm sh}(\boldsymbol{k},\boldsymbol{k}) {\rm wal}_{2\boldsymbol{k}+\nu(\boldsymbol{k})}(\boldsymbol{x}) {\rm wal}_{2\boldsymbol{k}+\nu(\boldsymbol{k})}(\boldsymbol{y}) \\ &= \sum_{\boldsymbol{k}\in\mathcal{E}} \hat{K}_{\rm sh}(\lfloor \boldsymbol{k}/2\rfloor, \lfloor \boldsymbol{k}/2\rfloor) {\rm wal}_{\boldsymbol{k}}(\boldsymbol{x}) {\rm wal}_{\boldsymbol{k}}(\boldsymbol{y}). \end{aligned}$$

Proof. Let $K_\phi(\boldsymbol{x},\boldsymbol{y}) = K(\phi(\boldsymbol{x}),\phi(\boldsymbol{y}))$. Then as $K \in \mathcal{L}_2([0,1]^{2s})$ it follows that $K_\phi \in \mathcal{L}_2([0,1]^{2s})$ and hence by Proposition 1 we can write

$$K_\phi(\boldsymbol{x},\boldsymbol{y}) = \sum_{\boldsymbol{k}\in\mathbb{N}_0^s}\sum_{\boldsymbol{k}'\in\mathbb{N}_0^s} \hat{K}_\phi(\boldsymbol{k},\boldsymbol{k}') {\rm wal}_{\boldsymbol{k}}(\boldsymbol{x}) {\rm wal}_{\boldsymbol{k}'}(\boldsymbol{y}), \tag{5}$$

with

$$\hat{K}_\phi(\boldsymbol{k},\boldsymbol{k}') = \int_{[0,1]^s}\int_{[0,1]^s} K_\phi(\boldsymbol{x},\boldsymbol{y}) {\rm wal}_{\boldsymbol{k}}(\boldsymbol{x}) {\rm wal}_{\boldsymbol{k}'}(\boldsymbol{y}) \,{\rm d}\boldsymbol{x}\,{\rm d}\boldsymbol{y}.$$

Note that

$$\begin{aligned} \int_{[0,1]^s} {\rm wal}_{\boldsymbol{k}}(\boldsymbol{x}\oplus\boldsymbol{\sigma}) {\rm wal}_{\boldsymbol{k}'}(\boldsymbol{y}\oplus\boldsymbol{\sigma}) \,{\rm d}\boldsymbol{\sigma} &= {\rm wal}_{\boldsymbol{k}}(\boldsymbol{x}) {\rm wal}_{\boldsymbol{k}'}(\boldsymbol{y}) \int_{[0,1]^s} {\rm wal}_{\boldsymbol{k}}(\boldsymbol{\sigma}) {\rm wal}_{\boldsymbol{k}'}(\boldsymbol{\sigma}) \,{\rm d}\boldsymbol{\sigma} \\ &= \begin{cases} 0 & \text{if } \boldsymbol{k}\neq\boldsymbol{k}', \\ {\rm wal}_{\boldsymbol{k}}(\boldsymbol{x}) {\rm wal}_{\boldsymbol{k}'}(\boldsymbol{y}) & \text{if } \boldsymbol{k}=\boldsymbol{k}'. \end{cases} \end{aligned} \tag{6}$$

As $K_{{\rm sh},\phi}(\boldsymbol{x},\boldsymbol{y}) = \int_{[0,1]^s} K_\phi(\boldsymbol{x}\oplus\boldsymbol{\sigma},\boldsymbol{y}\oplus\boldsymbol{\sigma}) \,{\rm d}\boldsymbol{\sigma}$ it follows from Proposition 1, (5) and (6) that

$$K_{{\rm sh},\phi}(\boldsymbol{x},\boldsymbol{y}) = \sum_{\boldsymbol{k}\in\mathbb{N}_0^s} \hat{K}_\phi(\boldsymbol{k},\boldsymbol{k}) {\rm wal}_{\boldsymbol{k}}(\boldsymbol{x}) {\rm wal}_{\boldsymbol{k}}(\boldsymbol{y}).$$

Hence it remains to show that $\hat{K}_\phi(\boldsymbol{k},\boldsymbol{k}) = \hat{K}_{{\rm sh},\phi}(\boldsymbol{k},\boldsymbol{k})$ for all $\boldsymbol{k}\in\mathcal{E}$ and $\hat{K}_\phi(\boldsymbol{k},\boldsymbol{k}) = 0$ for all $\boldsymbol{k}\in\mathbb{N}_0^s\setminus\mathcal{E}$. As

$$\hat{K}_\phi(\boldsymbol{k},\boldsymbol{k}) = \int_{[0,1]^s}\int_{[0,1]^s} K(\phi(\boldsymbol{x}),\phi(\boldsymbol{y})) {\rm wal}_{\boldsymbol{k}}(\boldsymbol{x}) {\rm wal}_{\boldsymbol{k}}(\boldsymbol{y}) \,{\rm d}\boldsymbol{x}\,{\rm d}\boldsymbol{y}$$

we can divide each integral from 0 to 1 into $\int_0^{1/2} + \int_{1/2}^1$. Note that $\phi(x) = 2x$ for $x \in [0, 1/2]$ and $\phi(x) = 2 - 2x$ for $x \in [1/2, 1]$ and hence by a transformation of variables we obtain two integrals from 0 to 1. Doing this we get

$$\hat{K}_\phi(\boldsymbol{k}, \boldsymbol{k}) = \int_{[0,1]^s} \int_{[0,1]^s} K(\boldsymbol{x}, \boldsymbol{y}) \prod_{j=1}^s A_{k_j}(x_j, y_j) \,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y},$$

where

$$\begin{aligned} A_{k_j}(x_j, y_j) &= \frac{1}{4}\Bigg[\mathrm{wal}_{k_j}\left(\frac{x_j}{2}\right)\mathrm{wal}_{k_j}\left(\frac{y_j}{2}\right) + \mathrm{wal}_{k_j}\left(1 - \frac{x_j}{2}\right)\mathrm{wal}_{k_j}\left(\frac{y_j}{2}\right) \\ &\quad + \mathrm{wal}_{k_j}\left(\frac{x_j}{2}\right)\mathrm{wal}_{k_j}\left(1 - \frac{y_j}{2}\right) + \mathrm{wal}_{k_j}\left(1 - \frac{x_j}{2}\right)\mathrm{wal}_{k_j}\left(1 - \frac{y_j}{2}\right)\Bigg]. \end{aligned}$$

To simplify the above expression we use Lemma 3 and obtain for almost all $x, y \in [0, 1)$ that

$$\begin{aligned} &A_k(x, y) \\ &= \frac{1}{4}\left[\mathrm{wal}_k\left(\frac{x}{2}\right)\mathrm{wal}_k\left(\frac{y}{2}\right)(1 + (-1)^{\sigma(k)}) + \mathrm{wal}_k\left(1 - \frac{x}{2}\right)\mathrm{wal}_k\left(1 - \frac{y}{2}\right)(1 + (-1)^{\sigma(k)})\right] \\ &= \frac{1 + (-1)^{\sigma(k)}}{2}\mathrm{wal}_k\left(\frac{x}{2}\right)\mathrm{wal}_k\left(\frac{y}{2}\right). \end{aligned}$$

As $A_k(x, y) = 0$ if $\sigma(k) \equiv 1 \pmod 2$, $A_k(x, y) = \mathrm{wal}_k\left(\frac{x}{2}\right)\mathrm{wal}_k\left(\frac{y}{2}\right)$ if $\sigma(k) \equiv 0 \pmod 2$ and $\mathrm{wal}_k\left(\frac{x}{2}\right) = \mathrm{wal}_{\lfloor k/2 \rfloor}(x)$ the result follows. □

In order to obtain a formula for the worst-case error we can now use (3), Theorem 1 and Theorem 2.

Theorem 3 *Let $K \in \mathcal{L}_2([0, 1]^{2s})$ be a reproducing kernel and let $K_{\mathrm{sh},\phi}$ be the associated shifted and folded kernel.*

1. *Let $P_N \subseteq [0, 1]^s$ be a point set consisting of N points and let $\widetilde{P}_{N,\boldsymbol{\sigma},\phi}$ be the randomly digitally shifted and then folded version of P_N. The mean square worst-case error in a reproducing kernel Hilbert space with kernel $K \in \mathcal{L}_2([0, 1]^{2s})$ using the point set $\widetilde{P}_{N,\boldsymbol{\sigma},\phi}$ is given by*

$$\widehat{e}^2_{\mathrm{sh},\phi}(P_N, K) = -\hat{K}_{\mathrm{sh},\phi}(\mathbf{0}, \mathbf{0}) + \frac{1}{N^2}\sum_{\boldsymbol{x},\boldsymbol{y} \in P_N} K_{\mathrm{sh},\phi}(\boldsymbol{x}, \boldsymbol{y}).$$

2. *For the case when P_{2^m} is a digital (t, m, s)-net over $\mathbb{Z}_2$ we obtain*

$$\widehat{e}^2_{\mathrm{sh},\phi}(P_{2^m}, K) = -\hat{K}_{\mathrm{sh},\phi}(\mathbf{0}, \mathbf{0}) + \frac{1}{2^m}\sum_{\boldsymbol{x} \in P_{2^m}} K_{\mathrm{sh},\phi}(\boldsymbol{x}, \mathbf{0}).$$

Proof. From Proposition 1 and Theorem 2 follows that $\int_{[0,1]^s} K_{\mathrm{sh},\phi}(\boldsymbol{x}, \boldsymbol{y})\,\mathrm{d}\boldsymbol{y} = \hat{K}_{\mathrm{sh},\phi}(\mathbf{0}, \mathbf{0})$. Hence the first part follows now from (3). Further, it follows from Proposition 1 and

Theorem 2 that for $\boldsymbol{x}, \boldsymbol{y} \in [0,1)^s$, where all coordinates of $\boldsymbol{x}$ and $\boldsymbol{y}$ are dyadic rationals, $K_{\mathrm{sh},\phi}(\boldsymbol{x}, \boldsymbol{y}) = K_{\mathrm{sh},\phi}(\boldsymbol{x} \ominus \boldsymbol{y}, \boldsymbol{0})$. Hence Lemma 1 implies that

$$\sum_{\boldsymbol{x},\boldsymbol{y} \in P_{2^m}} K_{\mathrm{sh},\phi}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{\boldsymbol{x},\boldsymbol{y} \in P_{2^m}} K_{\mathrm{sh},\phi}(\boldsymbol{x} \ominus \boldsymbol{y}, \boldsymbol{0}) = \sum_{\boldsymbol{x} \in P_{2^m}} K_{\mathrm{sh},\phi}(\boldsymbol{x}, \boldsymbol{0}).$$

The second part now follows. □

In the following we will show how the mean square worst-case error in a reproducing kernel Hilbert space with kernel K using randomly digitally shifted and then folded digital nets or polynomial lattice rules can be written as a certain sum over $\hat{K}_{\mathrm{sh},\phi}(\boldsymbol{k}, \boldsymbol{k})$.

Before we state the result, we have to introduce some notation: for a non-negative integer k with base 2 representation $k = \sum_{i=0}^{\infty} \kappa_i 2^i$ (note that the sum is in fact finite) we write

$$\mathrm{tr}_m(k) := \kappa_0 + \kappa_1 2 + \cdots + \kappa_{m-1} 2^{m-1}$$

and

$$\mathrm{tr}_m(\vec{k}) := (\kappa_0, \ldots, \kappa_{m-1})^\top \in \mathbb{Z}_2^m.$$

Further we often associate k with the polynomial $k(x) = \sum_{i=0}^{\infty} \kappa_i x^i$ and vice versa. In this case we also write

$$\mathrm{tr}_m(k) := \kappa_0 + \kappa_1 x + \cdots + \kappa_m x^m.$$

Depending on the context it should always be clear which meaning $\mathrm{tr}_m(k)$ has. For a vector $\boldsymbol{k} \in \mathbb{N}_0^s$ or $\boldsymbol{k} \in \mathbb{Z}_2[x]^s$, $\mathrm{tr}_m(\boldsymbol{k})$ is defined component-wise.

For vectors $\boldsymbol{g} = (g_1, \ldots, g_s) \in \mathbb{Z}_2[x]^s$ and $\boldsymbol{k} = (k_1, \ldots, k_s) \in \mathbb{Z}_2[x]^s$ we define the "inner product"

$$\boldsymbol{k} \cdot \boldsymbol{g} = \sum_{i=1}^{s} k_i g_i$$

and we write $g \equiv 0 \pmod{f}$ if f divides g in $\mathbb{Z}_2[x]$.

Theorem 4 *Let the set $\mathcal{E}$ be defined as in Theorem 2.*

1. *Let $P_N \subseteq [0,1]^s$ be a point set consisting of N points and let $\widetilde{P}_{N,\boldsymbol{\sigma},\phi}$ be the randomly digitally shifted and then folded version of P_N. The mean square worst-case error in a reproducing kernel Hilbert space with kernel $K \in \mathcal{L}_2([0,1]^{2s})$ using the point set $\widetilde{P}_{N,\boldsymbol{\sigma},\phi}$ is given by*

$$\widehat{e}^2_{\mathrm{sh},\phi}(P_N, K) = \sum_{\boldsymbol{k} \in \mathcal{E} \setminus \{\boldsymbol{0}\}} \hat{K}_{\mathrm{sh},\phi}(\boldsymbol{k}, \boldsymbol{k}) \frac{1}{N^2} \sum_{\boldsymbol{x},\boldsymbol{y} \in P_N} \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}) \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{y}).$$

2. *For the case when P_{2^m} is a digital (t,m,s)-net over $\mathbb{Z}_2$ with generating matrices $C_1, \ldots, C_s$ we obtain*

$$\widehat{e}^2_{\mathrm{sh},\phi}(P_{2^m}, K) = \sum_{\boldsymbol{k} \in \mathcal{D} \cap \mathcal{E}} \hat{K}_{\mathrm{sh},\phi}(\boldsymbol{k}, \boldsymbol{k}),$$

where $\mathcal{D} = \{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\} : C_1^\top \mathrm{tr}_m(\vec{k}_1) + \cdots + C_s^\top \mathrm{tr}_m(\vec{k}_s) = \vec{0}\}$.

3. *For the case when P_{2^m} is the digital net $P(\boldsymbol{g}, f)$, with $f \in \mathbb{Z}_2[x]$ irreducible and* $\deg(f) = m$ *and generating vector $\boldsymbol{g}$, then*

$$\widehat{e}^2_{\mathrm{sh},\phi}(P_{2^m}, K) = \sum_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{g},f} \cap \mathcal{E}} \hat{K}_{\mathrm{sh},\phi}(\boldsymbol{k}, \boldsymbol{k}),$$

where $\mathcal{D}_{\boldsymbol{g},f} = \{\boldsymbol{k} \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\} : \mathrm{tr}_m(\boldsymbol{k}) \cdot \boldsymbol{g} \equiv 0 \,(\mathrm{mod}\, f)\}$.

Proof. The first part follows from Theorem 2 and Theorem 3.

For the second part observe that we obtain from Proposition 1 that

$$\frac{1}{2^{2m}} \sum_{\boldsymbol{x},\boldsymbol{y} \in P_{2^m}} \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{y}) = \frac{1}{2^{2m}} \sum_{\boldsymbol{y} \in P_{2^m}} \left(\sum_{\boldsymbol{x} \in P_{2^m}} \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x} \oplus \boldsymbol{y}) \right).$$

Due to the group structure of a digital net, see Lemma 1, each summand in the outermost sum has the same value and therefore we obtain

$$\frac{1}{2^{2m}} \sum_{\boldsymbol{x},\boldsymbol{y} \in P_{2^m}} \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x})\mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{y}) = \frac{1}{2^m} \sum_{\boldsymbol{x} \in P_{2^m}} \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}) = \begin{cases} 1 & \text{if } \boldsymbol{k} \in \mathcal{D}, \\ 0 & \text{otherwise}, \end{cases}$$

where the last equality follows from Lemma 2. Hence the second part of the theorem follows.

The third part of the theorem follows from the second part together with the fact that if $C_1, \ldots, C_s$ are generating matrices for the point set $P(\boldsymbol{g}, f)$, then for any $\boldsymbol{k} = (k_1, \ldots, k_s) \in \mathbb{N}_0^s \setminus \{\boldsymbol{0}\}$ we have

$$C_1^\top \mathrm{tr}_m(\vec{k}_1) + \cdots + C_s^\top \mathrm{tr}_m(\vec{k}_s) = \vec{0} \quad \text{iff} \quad \mathrm{tr}_m(\boldsymbol{k}) \cdot \boldsymbol{g} \equiv 0 \,(\mathrm{mod}\, f).$$

This was first proved in [10, Lemma 4.40]. Hence also the third part follows. □

4 The mean square worst-case error in weighted Sobolev spaces

Before we consider the Sobolev spaces we introduce some notation. Let $S = \{1, \ldots, s\}$ be the set of coordinate indices. For $u \subseteq S$, let $|u|$ denote the cardinality of u and let $\boldsymbol{x}_u$ denote the vector of elements of $\boldsymbol{x}$ whose coordinate indices are contained in u. For a sequence $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots)$ of non-negative integers (the so-called weights) let $\gamma_u = \prod_{j \in u} \gamma_j$ (see [17] for more information on the weights).

First let us consider the reproducing kernel Sobolev space which contains functions where the first partial derivatives of order up to one are square integrable. The norm in this space is given by

$$\|h\|^2 = \sum_{u \subseteq S} \frac{1}{\gamma_u} \int_{[0,1]^{|u|}} \left[\int_{[0,1]^{s-|u|}} \frac{\partial^{|u|} h}{\partial \boldsymbol{x}_u} \, \mathrm{d}\boldsymbol{x}_{S \setminus u} \right]^2 \mathrm{d}\boldsymbol{x}_u.$$

This space has previously been considered in [2] and [3]. Therein it was shown that randomly digitally shifted digital nets achieve a convergence order of $N^{-1+\varepsilon}$ for any $\varepsilon > 0$.

Further it is known that this convergence rate is best possible. Hence it cannot be expected that using the tent transformation can significantly improve the result. On the other hand it is also important to know if using the tent transformation yields a slower convergence. That this is not the case can be checked using Theorem 2, Theorem 4 and Remark 2. Indeed, if one uses randomly digitally shifted and then folded point sets instead of just randomly digitally shifted point sets one can obtain almost the same results as shown in [2] and [3]. A similar result holds also for lattice rules as shown in [6].

Hence in the following we turn our attention to the Sobolev space $H_{s,\boldsymbol{\gamma}}$, for which we require a stronger smoothness assumption. More precisely, we now assume that the partial derivatives up to order two have to be square integrable. This space is now defined by (see also [6])

$$H_{s,\boldsymbol{\gamma}} := \{h : \|h\| < \infty\},$$

where

$$\|h\|^2 = \sum_{u \subseteq S} \sum_{v \subseteq u} \frac{1}{\gamma_u \gamma_v} \int_{[0,1]^{|v|}} \left[\int_{[0,1]^{s-|v|}} \frac{\partial^{|u|+|v|} h}{\partial \boldsymbol{x}_u \partial \boldsymbol{x}_v} \, \mathrm{d}\boldsymbol{x}_{S \setminus v} \right]^2 \mathrm{d}\boldsymbol{x}_v.$$

The reproducing kernel of this Sobolev space is given by

$$K_{s,\boldsymbol{\gamma}}(\boldsymbol{x}, \boldsymbol{y}) = \prod_{j=1}^{s} K_{\gamma_j}(x_j, y_j),$$

where

$$K_\gamma(x, y) = 1 + \gamma B_1(x) B_1(y) + \frac{\gamma^2}{4} B_2(x) B_2(y) - \frac{\gamma^2}{24} B_4(|x - y|). \tag{7}$$

Here B_k denotes the k-th Bernoulli polynomial, i.e., $B_1(x) = x - \frac{1}{2}$, $B_2(x) = x^2 - x + \frac{1}{6}$ and $B_4(x) = x^4 - 2x^3 + x^2 - \frac{1}{30}$.

For this reproducing kernel it is shown in Appendix A that the digitally shifted and then folded kernel is given by

$$K_{s,\boldsymbol{\gamma},\mathrm{sh},\phi}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{\boldsymbol{k} \in \mathbb{N}_0^s} r(\boldsymbol{k}, \boldsymbol{\gamma}) \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{x}) \mathrm{wal}_{\boldsymbol{k}}(\boldsymbol{y}), \tag{8}$$

where for $\boldsymbol{k} = (k_1, \ldots, k_s) \in \mathbb{N}_0^s$, $r(\boldsymbol{k}, \boldsymbol{\gamma}) = \prod_{j=1}^{s} r(k_j, \gamma_j)$ and $r(k, \gamma) = 0$ if $\sigma(k) \equiv 1 \pmod 2$ and for $\sigma(k) \equiv 0 \pmod 2$ we define

$$r(k, \gamma) = \begin{cases} 1 & \text{if } k = 0, \\ \frac{\gamma}{4 \cdot 2^{2a}} + \frac{\gamma^2}{120 \cdot 2^{4a}} & \text{if } k = 2^a + 1 \text{ with } a \geq 1, \\ \frac{\gamma^2}{12 \cdot 2^{2(a+j)}} + \frac{\gamma^2}{20 \cdot 2^{4a}} & \text{if } k = 2^a + 2^j + l, \text{ where } 1 \leq j < a \text{ and } 0 \leq l < 2^j \\ & \text{with } \ \sigma(l) \equiv 0 \pmod 2). \end{cases}$$

Note that $r(\boldsymbol{k}, \boldsymbol{\gamma})$ is zero for $\boldsymbol{k} \in \mathbb{N}_0^s \setminus \mathcal{E}$, hence in (8) we can also sum over all $\boldsymbol{k}$ in $\mathbb{N}_0^s$ rather than $\mathcal{E}$.

Using this result and Theorem 4 we can obtain a formula for the mean square worst-case error in $H_{s,\boldsymbol{\gamma}}$ for randomly digitally shifted and then folded point sets. One just has to set $\hat{K}_{\mathrm{sh},\phi}(\boldsymbol{k}, \boldsymbol{k}) = \prod_{j=1}^{s} r(k_j, \gamma_j)$ in Theorem 4 to obtain the mean square worst-case error in the space $H_{s,\boldsymbol{\gamma}}$. In the following we will prove some existence results for digital nets and polynomial lattice rules.

4.1 An existence result for digital nets

Let M_m be the set of all $m \times m$ matrices with entries in $\{0,1\}$ and let $\mathcal{C} := \{(C_1, \ldots, C_s) : C_j \in M_m \text{ for } j = 1, \ldots, s\}$, (note that $|\mathcal{C}| = 2^{m^2 s}$). Then we define

$$A_{2^m,s,\lambda} \;:=\; \frac{1}{2^{m^2 s}} \sum_{(C_1,\ldots,C_s)\in\mathcal{C}} \sum_{\boldsymbol{k}\in\mathcal{D}} r^{\lambda}(\boldsymbol{k},\boldsymbol{\gamma}). \tag{9}$$

In the following lemma we obtain a bound on $A_{2^m,s,\lambda}$.

Lemma 4 *For $\frac{1}{4} < \lambda \le 1$ we have*

$$A_{2^m,s,\lambda} \le \frac{2}{2^m} \prod_{j=1}^{s} (1 + \gamma_j^{2\lambda} \tau_\lambda + \gamma_j^{\lambda} \zeta_\lambda),$$

where

$$\zeta_\lambda = \frac{1}{2^{2\lambda}(2^{2\lambda}-1)}$$

and

$$\tau_\lambda = \frac{3^\lambda + 20^\lambda + 80^\lambda}{960^\lambda(16^\lambda - 1)} + \frac{3^\lambda + 5^\lambda + 20^\lambda}{60^\lambda(2 - 3\cdot 16^\lambda + 256^\lambda)}.$$

Proof. As in [3, Proof of Lemma 4] we find that

$$\begin{aligned}
A_{2^m,s,\lambda} &= -1 + \left(1 - \frac{1}{2^m}\right) \sum_{\boldsymbol{l}\in\mathbb{N}_0^s} r^{\lambda}(2^m \boldsymbol{l}, \boldsymbol{\gamma}) + \frac{1}{2^m} \sum_{\boldsymbol{l}\in\mathbb{N}_0^s} r^{\lambda}(\boldsymbol{l},\boldsymbol{\gamma}) \\
&\le -1 + \sum_{\boldsymbol{l}\in\mathbb{N}_0^s} r^{\lambda}(2^m \boldsymbol{l}, \boldsymbol{\gamma}) + \frac{1}{2^m} \sum_{\boldsymbol{l}\in\mathbb{N}_0^s} r^{\lambda}(\boldsymbol{l},\boldsymbol{\gamma}) \\
&= -1 + \prod_{j=1}^{s} \left(\sum_{l=0}^{\infty} r^{\lambda}(2^m l, \gamma_j) \right) + \frac{1}{2^m} \prod_{j=1}^{s} \left(\sum_{l=0}^{\infty} r^{\lambda}(l, \gamma_j) \right).
\end{aligned}$$

With Lemma 9 from Appendix C we obtain

$$\begin{aligned}
-1 + \prod_{j=1}^{s} \left(\sum_{l=0}^{\infty} r^{\lambda}(2^m l, \gamma_j) \right) &\le -1 + \prod_{j=1}^{s} \left(1 + \frac{\gamma_j^{2\lambda}}{2^{4\lambda m}} \tau_\lambda \right) = \sum_{\substack{u\subseteq\{1,\ldots,s\}\\ u\neq\emptyset}} \prod_{j\in u} \frac{\gamma_j^{2\lambda}}{2^{4\lambda m}} \tau_\lambda \\
&\le \frac{1}{2^m} \prod_{j=1}^{s} (1 + \gamma_j^{2\lambda} \tau_\lambda).
\end{aligned}$$

Therefore, with Lemma 8 from Appendix C and since $\widehat{\tau}_\lambda \le \tau_\lambda$ we obtain

$$\begin{aligned}
A_{2^m,s,\lambda} &\le \frac{1}{2^m} \prod_{j=1}^{s} (1 + \gamma_j^{2\lambda} \tau_\lambda) + \frac{1}{2^m} \prod_{j=1}^{s} (1 + \gamma_j^{2\lambda} \widehat{\tau}_\lambda + \gamma_j^{\lambda} \zeta_\lambda) \\
&\le \frac{2}{2^m} \prod_{j=1}^{s} (1 + \gamma_j^{2\lambda} \tau_\lambda + \gamma_j^{\lambda} \zeta_\lambda)
\end{aligned}$$

and hence the result follows. $\square$

Remark 3 The restriction $\lambda > 1/4$ in the lemma above is needed as otherwise $\sum_{l=0}^{\infty} r^{\lambda}(l,\gamma)$ is not finite anymore.

Remark 4 In Appendix A we also calculated the Walsh coefficients of the shift invariant kernel $K_{s,\boldsymbol{\gamma},\mathrm{sh}}$. It can be shown that if one replaces $r(\boldsymbol{k},\boldsymbol{\gamma})$ in (9), (10), (11), Lemma 4, Lemma 5 and Lemma 6 with $K_{s,\boldsymbol{\gamma},\mathrm{sh}}(\boldsymbol{k},\boldsymbol{k})$ we can obtain a similar result to Lemma 4, Lemma 5 and Lemma 6 only for $\frac{1}{2} < \lambda \leq 1$. Hence in the existence results above we would only obtain a convergence rate of $\mathcal{O}(2^{m(-1+\varepsilon)})$ for any $\varepsilon > 0$.

From the average result obtained above we can now obtain an existence result for digital nets. We will make use of Jensen's inequality, which states that for any sequence of non-negative real numbers $(a_k)_{k\geq 1}$ and any $0 < \lambda \leq 1$ we have

$$\left(\sum_k a_k\right)^{\lambda} \leq \sum_k a_k^{\lambda}.$$

We have the following theorem.

Theorem 5 *There exists a digital (t,m,s)-net P_{2^m} over $\mathbb{Z}_2$ such that for any $\frac{1}{4} < \lambda \leq 1$ the mean square worst-case error is bounded by*

$$\widehat{e}^2_{\mathrm{sh},\phi}(P_{2^m},K_{s,\boldsymbol{\gamma}}) \leq c_{s,\boldsymbol{\gamma},\lambda} 2^{-m/\lambda},$$

where

$$c_{s,\boldsymbol{\gamma},\lambda} := 2^{1/\lambda} \prod_{j=1}^{s} (1+\gamma_j^{2\lambda}\tau_\lambda + \gamma_j^{\lambda}\zeta_\lambda)^{1/\lambda}.$$

Here τ_λ and ζ_λ are defined as in Lemma 4.

Proof. Let $\widehat{e}_{\mathrm{sh},\phi}(C_1,\ldots,C_s) = \widehat{e}_{\mathrm{sh},\phi}(P_{2^m},K_{s,\boldsymbol{\gamma}})$, where P_{2^m} is the digital net generated by the matrices $C_1,\ldots,C_s$. Then for any $0 < \lambda \leq 1$ there exists a choice of $C_1',\ldots,C_s'$ such that

$$\widehat{e}^{2\lambda}_{\mathrm{sh},\phi}(C_1',\ldots,C_s') \leq \frac{1}{2^{m^2 s}} \sum_{(C_1,\ldots,C_s)\in\mathcal{C}} \widehat{e}^{2\lambda}_{\mathrm{sh},\phi}(C_1,\ldots,C_s).$$

From Jensen's inequality it now follows that

$$\frac{1}{2^{m^2 s}} \sum_{(C_1,\ldots,C_s)\in\mathcal{C}} \widehat{e}^{2\lambda}_{\mathrm{sh},\phi}(C_1,\ldots,C_s) \leq \frac{1}{2^{m^2 s}} \sum_{(C_1,\ldots,C_s)\in\mathcal{C}} \sum_{\boldsymbol{k}\in\mathcal{D}} r^{\lambda}(\boldsymbol{k},\boldsymbol{\gamma}) = A_{2^m,s,\lambda}.$$

The result now follows from Lemma 4. □

Theorem 5 shows that there exists a randomly digitally shifted digital net which is then folded using the tent transformation which achieves a convergence order of $\mathcal{O}(2^{m(-2+\varepsilon)})$ for any $\varepsilon > 0$. A comparable result holds for lattice rules, see [6]. For digital nets explicit constructions do exist [10, 12, 13] (in contrast to lattice rules), but at present it is not clear if those constructions in conjunction with the tent transformation can yield such a convergence.

Further, Remark 4 implies that the arguments used above would not yield a convergence order of $\mathcal{O}(2^{m(-2+\varepsilon)})$ for any $\varepsilon > 0$ if one uses only a random digital shift but not the tent transformation. Instead the argument would only yield a convergence order of $\mathcal{O}(2^{m(-1+\varepsilon)})$ for any $\varepsilon > 0$.

4.2 An existence result for polynomial lattice rules

On the other hand we can show that a similar result as for the average over all digital nets holds for the average over all digital nets $P(\boldsymbol{g}, f)$, where f is an irreducible polynomial. Note that the number of digital nets of the form $P(\boldsymbol{g}, f)$ is $(2^m - 1)^s$, whereas the number of all digital nets is $2^{m^2 s}$, which is of course much larger.

By G_m we denote the set consisting of all nonzero polynomials from $\mathbb{Z}_2[x]$ with degree smaller than m, i.e.,

$$G_m := \{h \in \mathbb{Z}_2[x] \setminus \{0\} : \deg(h) < m\}.$$

For a polynomial $f \in \mathbb{Z}_2[x]$ we define

$$\widehat{A}_{f,s,\lambda} := \frac{1}{(2^m - 1)^s} \sum_{\boldsymbol{g} \in G_m} \sum_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{g},f}} r^{\lambda}(\boldsymbol{k}, \boldsymbol{\gamma}). \tag{10}$$

Lemma 5 *Let $f \in \mathbb{Z}_2[x]$ be irreducible with $\deg(f) = m$. For $\frac{1}{4} < \lambda \leq 1$ we have*

$$\widehat{A}_{f,s,\lambda} \leq \frac{3}{2^m} \prod_{j=1}^{s} (1 + \gamma_j^{2\lambda} \tau_\lambda + \gamma_j^{\lambda} \zeta_\lambda),$$

where τ_λ and ζ_λ are defined as in Lemma 4.

Proof. We have

$$\widehat{A}_{f,s,\lambda} = \frac{1}{(2^m - 1)^s} \sum_{\substack{\boldsymbol{k} \in \mathbb{Z}_2[x]^s \\ \boldsymbol{k} \neq \boldsymbol{0}}} r^{\lambda}(\boldsymbol{k}, \boldsymbol{\gamma}) \sum_{\substack{\boldsymbol{g} \in G_m^s \\ \mathrm{tr}_m(\boldsymbol{k}) \cdot \boldsymbol{g} \equiv 0 \,(\mathrm{mod}\ f)}} 1.$$

Now we consider two cases:

1. If $\boldsymbol{k} = x^m \boldsymbol{l}$ with $\boldsymbol{l} \in \mathbb{Z}_2[x]^s$, $\boldsymbol{l} \neq \boldsymbol{0}$, then we have $\mathrm{tr}_m(\boldsymbol{k}) = \boldsymbol{0}$ and therefore

$$\sum_{\substack{\boldsymbol{g} \in G_m^s \\ \mathrm{tr}_m(\boldsymbol{k}) \cdot \boldsymbol{g} \equiv 0 \,(\mathrm{mod}\ f)}} 1 = (2^m - 1)^s.$$

2. If $\boldsymbol{k} = x^m \boldsymbol{l} + \boldsymbol{k}^*$ with $\boldsymbol{l} \in \mathbb{Z}_2[x]^s$ and $\boldsymbol{k}^* = (k_1^*, \ldots, k_s^*) \in \mathbb{Z}_2[x]^s$, $\boldsymbol{k}^* \neq \boldsymbol{0}$ and $\deg(k_i^*) < m$. Then we have $\mathrm{tr}_m(\boldsymbol{k}) = \boldsymbol{k}^*$ and hence

$$\sum_{\substack{\boldsymbol{g} \in G_m^s \\ \mathrm{tr}_m(\boldsymbol{k}) \cdot \boldsymbol{g} \equiv 0 \,(\mathrm{mod}\ f)}} 1 = \sum_{\substack{\boldsymbol{g} \in G_m^s \\ \boldsymbol{k}^* \cdot \boldsymbol{g} \equiv 0 \,(\mathrm{mod}\ f)}} 1.$$

   If $\boldsymbol{k}^* = (0, \ldots, 0, k_i^*, 0, \ldots, 0)$ with $k_i^* \neq 0$ then there is no polynomial $g_i \in G_m$ such that $\boldsymbol{k}^* \cdot \boldsymbol{g} = k_i^* g_i \equiv 0 \,(\mathrm{mod}\, f)$, since f is irreducible. Otherwise we have

$$\sum_{\substack{\boldsymbol{g} \in G_m^s \\ \boldsymbol{k}^* \cdot \boldsymbol{g} \equiv 0 \,(\mathrm{mod}\ f)}} 1 = (2^m - 1)^{s-1}.$$

Together we obtain

$$\begin{aligned}
\widehat{A}_{f,s,\lambda} &\leq \sum_{\substack{\boldsymbol{l}\in\mathbb{Z}_2[x]^s\\ \boldsymbol{l}\neq\boldsymbol{0}}} r^\lambda(2^m\boldsymbol{l},\boldsymbol{\gamma}) + \frac{1}{2^m-1}\sum_{\boldsymbol{l}\in\mathbb{Z}_2[x]^s}\sum_{\substack{\boldsymbol{k}^*\in\mathbb{Z}_2[x]^s\setminus\{\boldsymbol{0}\}\\ \deg(k_i^*)<m\ \forall i}} r^\lambda(2^m\boldsymbol{l}+\boldsymbol{k}^*,\boldsymbol{\gamma})\\
&\leq -1+\sum_{\boldsymbol{l}\in\mathbb{N}_0^s} r^\lambda(2^m\boldsymbol{l},\boldsymbol{\gamma}) + \frac{2}{2^m}\sum_{\boldsymbol{k}\in\mathbb{N}_0^s} r^\lambda(\boldsymbol{k},\boldsymbol{\gamma}).
\end{aligned}$$

Now the result follows as in the proof of Lemma 4. □

Theorem 6 *Let $f\in\mathbb{Z}_2[x]$ be irreducible with $\deg(f)=m\geq 1$. Then there exists a vector $\boldsymbol{g}\in G_m^s$ such that for any $\frac{1}{4}<\lambda\leq 1$ the mean square worst-case error is bounded by*

$$\widehat{e}^2_{\mathrm{sh},\phi}(P(\boldsymbol{g},f),K_{s,\boldsymbol{\gamma}}) \leq \widehat{c}_{s,\boldsymbol{\gamma},\lambda}2^{-m/\lambda},$$

where

$$\widehat{c}_{s,\boldsymbol{\gamma},\lambda} := 3^{1/\lambda}\prod_{j=1}^s(1+\gamma_j^{2\lambda}\tau_\lambda+\gamma_j^\lambda\zeta_\lambda)^{1/\lambda}.$$

Here τ_λ and ζ_λ are defined as in Lemma 4.

Proof. The result is proved in the same way as Theorem 5 with Lemma 4 replaced by Lemma 5. □

4.3 A component-by-component construction of polynomial lattice rules

In this subsection we show, how digital nets of the form $P(\boldsymbol{g},f)$ for which we obtain a bound for the mean square worst-case error as in Theorem 6 can be found by computer search.

Algorithm 1 *Given a dimension s, an integer $m\geq 1$ and weights $\boldsymbol{\gamma}=(\gamma_j)_{j\geq 1}$.*

1. *Choose an irreducible polynomial $f\in\mathbb{Z}_2[x]$ with $\deg(p)=m$.*
2. *Set $g_1^*=1$.*
3. *For $d=2,3,\ldots,s$, find $g_d^*\in G_m$ by minimizing the square worst-case error*

   $$\widehat{e}^2_{\mathrm{sh},\phi}(P((g_1^*,\ldots,g_{d-1}^*,g_d),f),K_{s,\boldsymbol{\gamma}})$$

   as a function in g_d.

Theorem 7 *Let $f\in\mathbb{Z}_2[x]$ be irreducible, with $\deg(f)=m\in\mathbb{N}$. Suppose that $(g_1^*,\ldots,g_s^*)\in G_m^s$ is constructed according to Algorithm 1. Then for all $d=1,2,\ldots,s$ we have*

$$\widehat{e}^2_{\mathrm{sh},\phi}(P((g_1^*,\ldots,g_d^*),f),K_{d,\boldsymbol{\gamma}}) \leq \widehat{c}_{d,\boldsymbol{\gamma},\lambda}2^{-m/\lambda},$$

where

$$\widehat{c}_{s,\boldsymbol{\gamma},\lambda} := 2^{1/\lambda}\prod_{j=1}^d(1+\gamma_j^{2\lambda}\tau_\lambda+\gamma_j^\lambda\zeta_\lambda)^{1/\lambda},$$

and where $\frac{1}{4}<\lambda\leq 1$. Here τ_λ and ζ_λ are defined as in Lemma 4.

Proof. The proof of this result follows exactly the lines of the proof of [2, Theorem 4.4]. □

4.4 An existence result for Korobov polynomial lattice rules

Let now

$$\widehat{\widehat{A}}_{f,s,\lambda} := \frac{1}{2^m - 1} \sum_{g \in G_m} \sum_{\boldsymbol{k} \in \mathcal{D}_{\boldsymbol{v}_s(g),f}} r^\lambda(\boldsymbol{k}, \boldsymbol{\gamma}), \tag{11}$$

where for $g \in G_m$ we use the notation $\boldsymbol{v}_s(g) = (1, g, g^2, \ldots, g^{s-1}) \pmod{f}$. Note that this corresponds to averaging over all polynomial lattice rules of Korobov form, see [2] where similar calculations have been carried out.

Lemma 6 *Let $f \in \mathbb{Z}_2[x]$ be irreducible with $\deg(f) = m \geq 1$. For $\frac{1}{4} < \lambda \leq 1$ we have*

$$\widehat{\widehat{A}}_{f,s,\lambda} \leq \frac{2s-1}{2^m} \prod_{j=1}^{s} (1 + \gamma_j^{2\lambda} \tau_\lambda + \gamma_j^\lambda \zeta_\lambda),$$

where τ_λ and ζ_λ are defined as in Lemma 4.

Proof. As in the proof of Lemma 5 we obtain

$$\widehat{\widehat{A}}_{f,s,\lambda} \leq \sum_{\substack{\boldsymbol{l} \in \mathbb{Z}_2[x]^s \\ \boldsymbol{l} \neq \boldsymbol{0}}} r^\lambda(2^m \boldsymbol{l}, \boldsymbol{\gamma}) + \frac{1}{2^m - 1} \sum_{l \in \mathbb{Z}_2[x]^s} \sum_{\substack{\boldsymbol{k}^* \in \mathbb{Z}_2[x]^s \setminus \{\boldsymbol{0}\} \\ \deg(k_i^*) < m \ \forall i}} r^\lambda(2^m \boldsymbol{l} + \boldsymbol{k}^*, \boldsymbol{\gamma}) \sum_{\substack{\boldsymbol{g} \in G_m^s \\ \boldsymbol{v}_s(g) \cdot \boldsymbol{k}^* \equiv 0 \pmod{f}}} 1.$$

Now we recall that for an irreducible polynomial $f \in \mathbb{Z}_2[x]$, with $\deg(f) = m \geq 1$, and a nonzero $(h_1, \ldots, h_s) \in \mathbb{Z}_2[x]^s$ with $\deg(h_i) < m$, $i = 1, \ldots, s$, the congruence

$$h_1 + h_2 g + \cdots + h_s g^{s-1} \equiv 0 \pmod{f}$$

has at most $s-1$ solutions $g \in G_m$. Therefore

$$\widehat{\widehat{A}}_{f,s,\lambda} \leq -1 + \sum_{\boldsymbol{l} \in \mathbb{N}_0^s} r^\lambda(2^m \boldsymbol{l}, \boldsymbol{\gamma}) + \frac{s-1}{2^m - 1} \sum_{k \in \mathbb{N}_0^s} r^\lambda(\boldsymbol{k}, \boldsymbol{\gamma}).$$

Now the result follows as in the proof of Lemma 4. □

Again we can obtain an existence result from the above lemma.

Theorem 8 *Let $f \in \mathbb{Z}_2[x]$ be irreducible with $\deg(f) = m \geq 1$. Then there exists a polynomial $g \in G_m$ such that for any $\frac{1}{4} < \lambda \leq 1$ the mean square worst-case error is bounded by*

$$\widehat{e}^2(P(\boldsymbol{v}_s(g), f), K_{s,\boldsymbol{\gamma}}) \leq \widehat{\widehat{c}}_{s,\boldsymbol{\gamma},\lambda} 2^{-m/\lambda},$$

where

$$\widehat{\widehat{c}}_{s,\boldsymbol{\gamma},\lambda} := (2s-1)^{1/\lambda} \prod_{j=1}^{s} (1 + \gamma_j^{2\lambda} \tau_\lambda + \gamma_j^\lambda \zeta_\lambda)^{1/\lambda}.$$

Here τ_λ and ζ_λ are defined as in Lemma 4.

Proof. The result is proved in the same way as Theorem 5 with Lemma 4 replaced by Lemma 6. □

A polynomial g^* for which we obtain a bound for the mean square worst-case error as in Theorem 8 can be found by the following Algorithm 2.

Algorithm 2 *Given a dimension $s \geq 2$, an integer $m \geq 1$ and weights $\boldsymbol{\gamma} = (\gamma_j)_{j\geq 1}$.*

1. *Choose an irreducible polynomial $f \in \mathbb{Z}_2[x]$ with $\deg(f) = m$.*
2. *Find $g^* \in G_m$ by minimizing the square worst-case error*

$$\widehat{e}^2_{\mathrm{sh},\phi}(P(\boldsymbol{v}_s(g), f), K_{s,\boldsymbol{\gamma}})$$

as a function in g.

5 Numerical experiments

In Lemma 7 in Appendix B we gave a closed form of the digitally shifted and folded kernel $K_{\boldsymbol{\gamma},\mathrm{sh},\phi}$. This result can be used in conjunction with Theorem 3 to obtain a closed form of the mean square worst-case error, which can be used as a quality measure for choosing good polynomial lattice rules. As noted in Remark 5 $\widehat{e}_{\mathrm{sh},\phi}(P_{2^m}, K_{s,\boldsymbol{\gamma}})$ can be computed in $\mathcal{O}(s2^m m)$ operations for digital nets or polynomial lattice rules. Hence this formula can be used in a component-by-component (Algorithm 1) or Korobov construction algorithm (Algorithm 2). A similar approach has previously been used in [2].

In the following we present the results for a component-by-component construction of polynomial lattice rules, more precisely we compute the mean square worst case errors for different values of m (2^m integration nodes) and s. The choice of irreducible polynomials f_m needed in Algorithm 1 for each m is shown in the following table:

m	1	2	3	4	5	6	7	8	9	10	11	12	13
f_m	3	7	11	25	47	103	203	487	865	1933	2881	6923	15847

Here, 3 denotes the polynomial $1 + x$, 7 the polynomial $1 + x + x^2$ and so on.

For our computation we fix weights $\gamma_1, \gamma_2, \ldots$ and compute

$$\widehat{e}(P_{m,s}, f_m, K_{s,\gamma}),$$

where $P_{m,s}$ is the point set corresponding to the polynomial lattice rule we obtain from the component-by-component construction. On the x-axis we plotted m while the y-axis shows the logarithm to the base 10 of $\widehat{e}(P_{m,s}, f_m, K_{s,\gamma})$. As a benchmark we also include a solid line which corresponds to convergence at the speed of $N^{-2} = 2^{-2m}$.

Figure 1 shows the results for $\gamma_j = j^{-2}$. We see that for moderate dimensions the error sticks to quadratic convergence. For higher dimensions it takes higher values of m to actually see near quadratic convergence behaviour, but it still can be observed for practical values of m. Notice that due to the fast decreasing weights the difference in the error between $s = 16$ and $s = 32$ is quite small.

The same observations can be made for the weights $\gamma_j = 0.5^j$, Figure 2. Here even the difference in the error between $s = 5$ and $s = 16$ is almost negligible for practical purposes.

On the other hand it is not surprising that in the case where $\gamma_j = 1$ for all j, i.e., when all coordinates are equally important, the curse of dimensionality takes over and

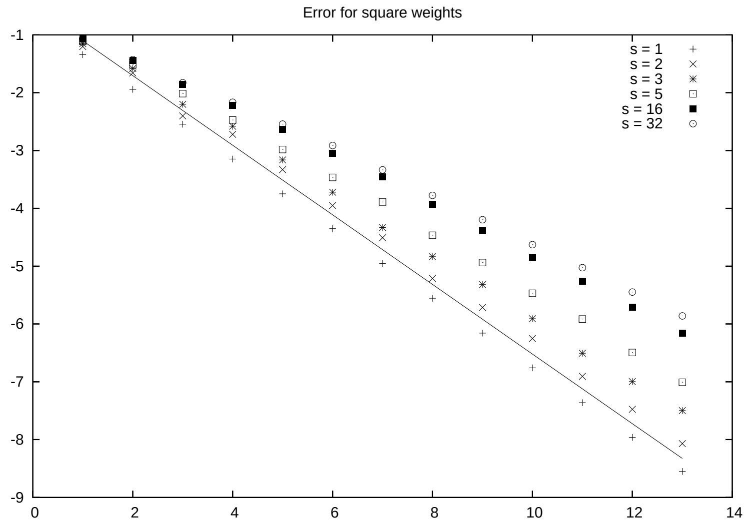

Figure 1: Root mean square worst case errors for $\gamma_j = \left(\frac{1}{j}\right)^2$

so the quadratic convergence for higher dimensions will be seen for very large values of m only. But still for dimension $s = 5$ we obtain a practically useful integration method. The results are shown in Figure 3.

It is instructive to compare the results for the folded points with the classical case. In Figure 4 we see that – in accordance with theory – the error of the non-folded point-set decreases like N^{-1}. We see that the method is less sensitive to increasing dimensionality and the graphs for $s = 5$ and $s = 16$ can hardly be distinguished in that case. But still the folded points for $s = 32$ do much better than the non-folded points.

Appendix A: Calculation of the shift invariant kernel $K_{s,\boldsymbol{\gamma},\mathrm{sh}}$ and the digitally shifted and folded kernel $K_{s,\boldsymbol{\gamma},\mathrm{sh},\phi}$

Here we compute the shift invariant kernel $K_{s,\boldsymbol{\gamma},\mathrm{sh}}(\boldsymbol{x},\boldsymbol{y})$ for the reproducing kernel

$$K_{s,\boldsymbol{\gamma}}(\boldsymbol{x},\boldsymbol{y}) = \prod_{j=1}^{s} K_{\gamma_j}(x_j, y_j),$$

where $K_\gamma(x,y)$ is given by (7). We have

$$\begin{aligned} K_{s,\boldsymbol{\gamma},\mathrm{sh}}(\boldsymbol{x},\boldsymbol{y}) &= \int_{[0,1]^s} K_{s,\boldsymbol{\gamma}}(\boldsymbol{x}\oplus\boldsymbol{\sigma}, \boldsymbol{y}\oplus\boldsymbol{\sigma})\,\mathrm{d}\boldsymbol{\sigma} \\ &= \prod_{j=1}^{s}\int_0^1 K_{\gamma_j}(x_j\oplus\sigma, y_j\oplus\sigma)\,\mathrm{d}\sigma. \end{aligned}$$

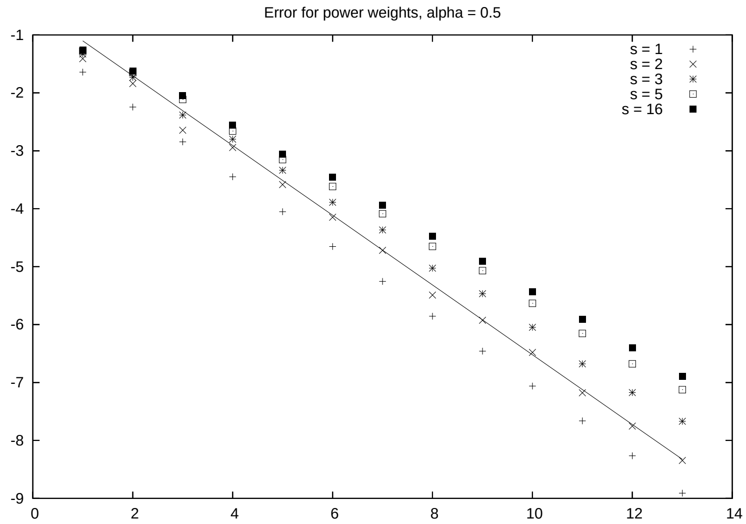

Figure 2: Root mean square worst case errors for $\gamma_j = (0.5)^j$

Hence it suffices to deal only with the one dimensional kernels. We only need to calculate the Walsh coefficients

$$\hat{K}_{\gamma_j,\mathrm{sh}}(k,k) = \int_0^1 \int_0^1 K_{\gamma_j}(x,y)\mathrm{wal}_k(x)\mathrm{wal}_k(y)\,\mathrm{d}x\,\mathrm{d}y. \tag{12}$$

First note that it is stated in Proposition 1 that

$$\int_0^1 \int_0^1 \mathrm{wal}_k(x)\mathrm{wal}_k(y)\,\mathrm{d}x\,\mathrm{d}y = \begin{cases} 1 & \text{if } k = 0, \\ 0 & \text{otherwise.} \end{cases}$$

It was shown in [3] that

$$x - \frac{1}{2} = -\frac{1}{2}\sum_{a=1}^{\infty} \frac{1}{2^a}\mathrm{wal}_{2^{a-1}}(x) \tag{13}$$

and as $B_1(x) = x - \frac{1}{2}$ we have

$$\int_0^1 \int_0^1 B_1(x)B_1(y)\mathrm{wal}_k(x)\mathrm{wal}_k(y)\,\mathrm{d}x\,\mathrm{d}y = \begin{cases} \frac{1}{4^{a+1}} & \text{if } k = 2^{a-1}, \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

We have $B_2(x) = x^2 - x + \frac{1}{6} = \left(x - \frac{1}{2}\right)^2 - \frac{1}{12}$ and thus

$$\begin{aligned} B_2(x) &= -\frac{1}{12} + \frac{1}{4}\sum_{a=1}^{\infty}\sum_{j=1}^{\infty} \frac{1}{2^{a+j}}\mathrm{wal}_{2^{a-1}}(x)\mathrm{wal}_{2^{j-1}}(x) \\ &= -\frac{1}{12} + \frac{1}{4}\sum_{a=1}^{\infty}\frac{1}{2^{2a}} + \frac{1}{2}\sum_{j=1}^{\infty}\sum_{a=j+1}^{\infty} \frac{1}{2^{a+j}}\mathrm{wal}_{2^{a-1}+2^{j-1}}(x) \\ &= \frac{1}{2}\sum_{j=1}^{\infty}\sum_{a=j+1}^{\infty} \frac{1}{2^{a+j}}\mathrm{wal}_{2^{a-1}+2^{j-1}}(x), \end{aligned}$$

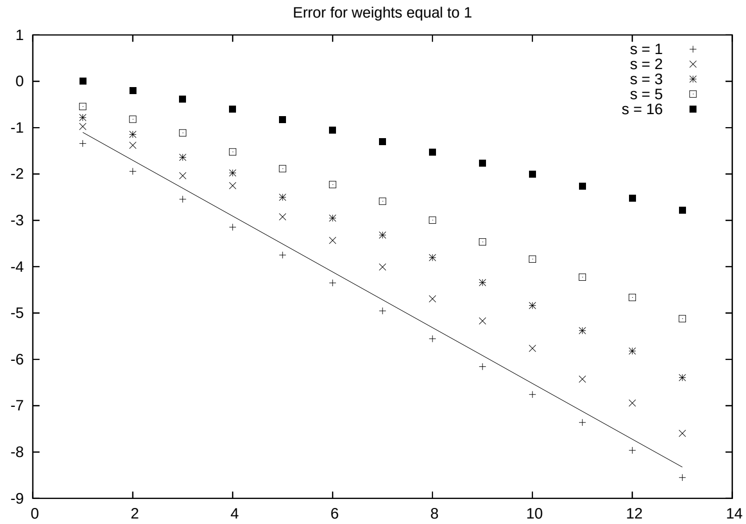

Figure 3: Root mean square worst case errors for $\gamma_j = 1$

which yields

$$\int_0^1 \int_0^1 B_2(x)B_2(y)\mathrm{wal}_k(x)\mathrm{wal}_k(y)\,\mathrm{d}x\,\mathrm{d}y = \begin{cases} \frac{1}{4^{a+j+1}} & \text{if } k = 2^{a-1} + 2^{j-1} \text{ for } 1 \le j \le a-1, \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

It remains to deal with $B_4(|x-y|)$. For $j \in \{0, 2, 3, 4\}$ define

$$I_j(k) := \int_0^1 \int_0^1 |x-y|^j \mathrm{wal}_k(x)\mathrm{wal}_k(y)\,\mathrm{d}x\,\mathrm{d}y.$$

Then we have

$$\int_0^1 \int_0^1 B_4(|x-y|)\mathrm{wal}_k(y)\mathrm{wal}_k(y)\,\mathrm{d}x\,\mathrm{d}y = I_4(k) - 2I_3(k) + I_2(k) - \frac{1}{30} I_0(k). \tag{16}$$

It now follows from Proposition 1 that

$$I_0(k) = \begin{cases} 1 & \text{if } k = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Note that for j even we do not need the to take the absolute value in the definition of $I_j(k)$ and hence we can use equation (13). First note that $\int_0^1 \int_0^1 (x-y)^2\,\mathrm{d}x\,\mathrm{d}y = \frac{1}{6}$. We have

$$\begin{aligned} (x-y)^2 &= \frac{1}{4}\left(\sum_{a=1}^\infty \frac{1}{2^a}[\mathrm{wal}_{2^{a-1}}(x) - \mathrm{wal}_{2^{a-1}}(y)]\right)^2 \\ &= \frac{1}{4}\sum_{a=1}^\infty \sum_{j=1}^\infty \frac{1}{2^{a+j}}[\mathrm{wal}_{2^{a-1}}(x) - \mathrm{wal}_{2^{a-1}}(y)][\mathrm{wal}_{2^{j-1}}(x) - \mathrm{wal}_{2^{j-1}}(y)]. \end{aligned}$$

Observe that for $k \in \mathbb{N}$ we have that

$$\int_0^1 \int_0^1 [\mathrm{wal}_{2^{a-1}}(x) - \mathrm{wal}_{2^{a-1}}(y)][\mathrm{wal}_{2^{j-1}}(x) - \mathrm{wal}_{2^{j-1}}(y)]\mathrm{wal}_k(x)\mathrm{wal}_k(y)\,\mathrm{d}x\,\mathrm{d}y$$

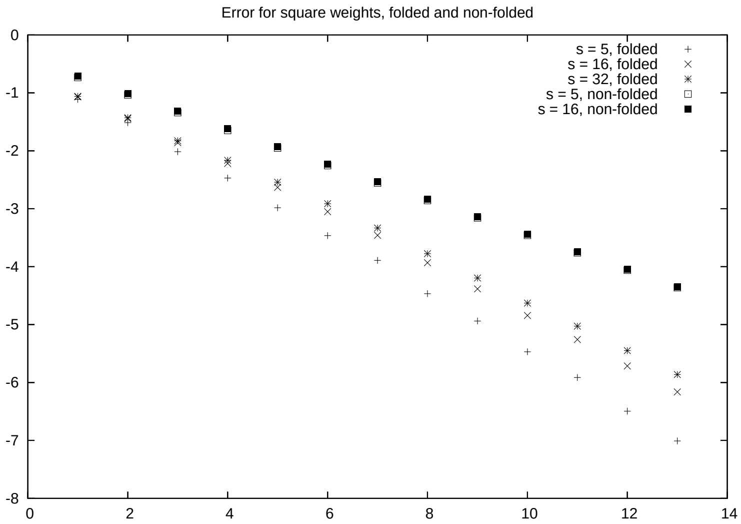

Figure 4: Root mean square worst case errors for $\gamma_j = (\frac{1}{j})^2$, folded and non-folded

is zero if $a \neq j$. If on the other hand $a = j$ and $k = 2^{a-1}$ the above integral yields -2 and for $k \neq 2^{a-1}$ the above integral is zero. Hence we obtain

$$I_2(k) = \begin{cases} \frac{1}{6} & \text{if } k = 0, \\ -\frac{1}{2\cdot 4^a} & \text{if } k = 2^{a-1}, \\ 0 & \text{otherwise.} \end{cases}$$

For $j = 4$ we can use the same method. Note that $\int_0^1 \int_0^1 (x-y)^4 \,\mathrm{d}x\,\mathrm{d}y = \frac{1}{15}$. Further we have

$$\begin{aligned} (x-y)^4 &= \frac{1}{16}\left(\sum_{a=1}^{\infty} \frac{1}{2^a}[\mathrm{wal}_{2^{a-1}}(x) - \mathrm{wal}_{2^{a-1}}(y)]\right)^4 \\ &= \frac{1}{16}\sum_{a_1=1}^{\infty}\sum_{a_2=1}^{\infty}\sum_{a_3=1}^{\infty}\sum_{a_4=1}^{\infty} \frac{1}{2^{a_1+a_2+a_3+a_4}} \prod_{i=1}^{4}[\mathrm{wal}_{2^{a_i-1}}(x) - \mathrm{wal}_{2^{a_i-1}}(y)]. \end{aligned} \quad (17)$$

For $k \in \mathbb{N}$ we need to consider

$$\int_0^1 \int_0^1 \prod_{i=1}^{4}[\mathrm{wal}_{2^{a_i-1}}(x) - \mathrm{wal}_{2^{a_i-1}}(y)]\mathrm{wal}_k(x)\mathrm{wal}_k(y)\,\mathrm{d}x\,\mathrm{d}y. \quad (18)$$

It follows from the orthogonality properties of the Walsh function system (see Proposition 1) that only if k is of the form $k = 2^{a-1}$ or $k = 2^{a-1} + 2^{j-1}$, $1 \le j < a$, the above integral is not zero. In order to obtain a non-zero value of (18) for $k = 2^{a-1}$ for some $a \ge 1$ we must have $a_{i_1} = a_{i_2} = a$ for some $1 \le i_1 < i_2 \le 4$ and for $i_3 \in \{1,2,3,4\} \setminus \{i_1, i_2\}$ and $i_4 \in \{1,2,3,4\} \setminus \{i_1,i_2,i_3\}$ we must have $a_{i_3} = a_{i_4}$ such that (18) becomes

$$-\int_0^1 \int_0^1 \mathrm{wal}_{2^{a_1-1}}(u)\mathrm{wal}_{2^{a_2-1}}(v)\mathrm{wal}_{2^{a_3-1}}(z)\mathrm{wal}_{2^{a_4-1}}(z)\mathrm{wal}_k(x)\mathrm{wal}_k(y)\,\mathrm{d}x\,\mathrm{d}y,$$

where z stands either for x or for y and we either have $u = x$ and $v = y$ or $u = y$ and $v = x$. Note that $\mathrm{wal}_{2^{a_3-1}}(z)\mathrm{wal}_{2^{a_4-1}}(z) = 1$ as $a_3 = a_4$. Further there are $\binom{4}{2}$ ways to choose the values of i_1, i_2. An exception is the case where $a_{i_1} = a_{i_2} = a_{i_3} = a_{i_4} = a$. In this case we have only $\binom{4}{3}$ ways to choose $\mathrm{wal}_{2^{a-1}}(u)\mathrm{wal}_{2^{a-1}}(v)\mathrm{wal}_{2^{a-1}}(v)\mathrm{wal}_{2^{a-1}}(v)$. Hence the value of $I_4(k)$ for $k = 2^{a-1}$ is given by

$$I_4(2^{a-1}) = -\frac{1}{16}\left(\frac{1}{2^{a+a}}2\cdot 2\binom{4}{2}\sum_{\substack{a_3=1\\a_3\neq a}}^{\infty}\frac{1}{2^{2a_3}} + \frac{8}{2^{a+a+a+a}}\right) = -\frac{1}{16}\left(\frac{24}{2^{2a}}\sum_{a_3=1}^{\infty}\frac{1}{2^{2a_3}} - \frac{16}{2^{4a}}\right)$$

and hence we have

$$I_4(2^{a-1}) = \frac{1}{2^{4a}} - \frac{1}{2\cdot 4^a}.$$

If $k = 2^{a-1} + 2^{j-1}$ for $1 \le j < a$ we have must have $a_{i_1} = a_{i_2} = a$ and $a_{i_3} = a_{i_4} = j$ for some i_1, i_2, i_3, i_4 such that $\{i_1, i_2, i_3, i_4\} = \{1, 2, 3, 4\}$ and (18) must become

$$\int_0^1\int_0^1 \mathrm{wal}_{2^{a_1-1}}(u)\mathrm{wal}_{2^{a_2-1}}(v)\mathrm{wal}_{2^{a_3-1}}(w)\mathrm{wal}_{2^{a_4-1}}(z)\mathrm{wal}_k(x)\mathrm{wal}_k(y)\,\mathrm{d}x\,\mathrm{d}y,$$

where now either $u = x$ and $v = y$ or $u = y$ and $v = x$ and further $w = x$ and $z = y$ or $w = y$ and $z = x$. Hence the value of $I_4(k)$ for $k = 2^{a-1} + 2^{j-1}$, $1 \le j < a$, is then given by

$$I_4(2^{a-1} + 2^{j-1}) = \frac{1}{16}\frac{1}{2^{a_1+a_2+a_3+a_4}}2\cdot 2\binom{4}{2} = \frac{3}{2\cdot 4^{a+j}}.$$

Hence we have

$$I_4(k) = \begin{cases} \frac{1}{15} & \text{if } k = 0,\\ \frac{1}{2^{4a}} - \frac{1}{2\cdot 4^a} & \text{if } k = 2^{a-1},\\ \frac{3}{2\cdot 4^{a+j}} & \text{if } k = 2^{a-1} + 2^{j-1} \text{ with } 1 \le j < a,\\ 0 & \text{otherwise.}\end{cases}$$

It remains to calculate $I_3(k)$. It can easily be checked that $I_3(0) = \frac{1}{10}$. For $k \ge 1$ let $k = \kappa_{a-1}2^{a-1} + \kappa_{a-2}2^{a-2} + \cdots + \kappa_0$ with $\kappa_{a-1} = 1$. We have

$$\begin{aligned} I_3(k) &= \int_0^1\int_0^1 |x-y|^3\mathrm{wal}_k(x)\mathrm{wal}_k(y)\,\mathrm{d}x\,\mathrm{d}y\\ &= \sum_{u=0}^{2^a-1}\sum_{v=0}^{2^a-1}(-1)^{\kappa_0(u_{a-1}+v_{a-1})+\cdots+\kappa_{a-1}(u_0+v_0)}\int_{u/2^a}^{(u+1)/2^a}\int_{v/2^a}^{(v+1)/2^a}|x-y|^3\,\mathrm{d}x\,\mathrm{d}y,\end{aligned}$$

where $u = u_{a-1}2^{a-1} + \cdots + u_0$ and v analogously.

For $0 \le u, v < 2^a$ it can be shown that

$$\int_{u/2^a}^{(u+1)/2^a}\int_{v/2^a}^{(v+1)/2^a}|x-y|^3\,\mathrm{d}x\,\mathrm{d}y = \begin{cases}\frac{1}{10\cdot 2^{5a}} & \text{if } u = v,\\ \frac{1}{2^{5a}}\left[|u-v|^3 + \frac{|u-v|}{2}\right] & \text{if } u \neq v.\end{cases}$$

Thus we have

$$\begin{aligned} I_3(k) &= \sum_{u=0}^{2^a-1}\frac{1}{10\cdot 2^{5a}} + \sum_{\substack{u,v=0\\u\neq v}}^{2^a-1}(-1)^{\kappa_0(u_{a-1}+v_{a-1})+\cdots+\kappa_{a-1}(u_0+v_0)}\frac{|u-v|}{2\cdot 2^{5a}}\\ &\quad + \sum_{\substack{u,v=0\\u\neq v}}^{2^a-1}(-1)^{\kappa_0(u_{a-1}+v_{a-1})+\cdots+\kappa_{a-1}(u_0+v_0)}\frac{|u-v|^3}{2^{5a}}. \end{aligned} \tag{19}$$

For the first sum we have $\sum_{u=0}^{2^a-1} \frac{1}{10\cdot 2^{5a}} = \frac{1}{10\cdot 2^{4a}}$. A sum very similar to the second sum was calculated in [3, Appendix A], we have

$$\sum_{\substack{u,v=0\\u\neq v}}^{2^a-1} (-1)^{\kappa_0(u_{a-1}+v_{a-1})+\cdots+\kappa_{a-1}(u_0+v_0)} \frac{|u-v|}{2\cdot 2^{5a}} = -\frac{1}{2\cdot 2^{4a}}.$$

Now we simplify the third sum. We have

$$\begin{aligned}&\sum_{\substack{u,v=0\\u\neq v}}^{2^a-1} (-1)^{\kappa_0(u_{a-1}+v_{a-1})+\cdots+\kappa_{a-1}(u_0+v_0)} \frac{|u-v|^3}{2^{5a}}\\ &= \frac{2}{2^{5a}} \sum_{u=0}^{2^a-2} \sum_{v=u+1}^{2^a-1} (v-u)^3 (-1)^{\kappa_0(u_{a-1}+v_{a-1})+\cdots+\kappa_{a-1}(u_0+v_0)}.\end{aligned}$$

Let now $u = u' + u_0$ and $v = v' + v_0$ with $u_0, v_0 \in \{0,1\}$ and u', v' even. Further let

$$\theta(u,v) = (v-u)^3 (-1)^{\kappa_0(u_{a-1}+v_{a-1})+\cdots+\kappa_{a-1}(u_0+v_0)}.$$

Let $v' > u'$, then we have

$$\sum_{u_0=0}^{1} \sum_{v_0=0}^{1} \theta(u'+u_0, v'+v_0) = -6(v'-u')(-1)^{\kappa_0(u_{a-1}+v_{a-1})+\cdots+\kappa_{a-2}(u_1+v_1)}.$$

For $v' = u'$ we have $v_0 = 1$ and $u_0 = 0$ as $v > u$ and $(-1)^{\kappa_0(u_{a-1}+v_{a-1})+\cdots+\kappa_{a-1}(u_0+v_0)} = -1$ as $\kappa_{a-1} = 1$ and $u_i = v_i$ for $i = 1, \ldots, a-1$. Hence

$$\begin{aligned}&\sum_{u=0}^{2^a-2} \sum_{v=u+1}^{2^a-1} (v-u)^3 (-1)^{\kappa_0(u_{a-1}+v_{a-1})+\cdots+\kappa_{a-1}(u_0+v_0)}\\ &= -2^{a-1} - 6 \sum_{\substack{u'=0\\2|u'}}^{2^a-4} \sum_{\substack{v'=u'+2\\2|v'}}^{2^a-2} (v'-u')(-1)^{\kappa_0(u_{a-1}+v_{a-1})+\cdots+\kappa_{a-2}(u_1+v_1)}.\end{aligned}$$

We have to distinguish two cases now, namely where $\kappa_0 = \cdots = \kappa_{a-2} = 0$, i.e. $k = 2^{a-1}$, and where this is not the case. If $k = 2^{a-1}$ we have

$$\begin{aligned}\sum_{u=0}^{2^a-2} \sum_{v=u+1}^{2^a-1} (v-u)^3 (-1)^{\kappa_0(u_{a-1}+v_{a-1})+\cdots+\kappa_{a-1}(u_0+v_0)} &= -2^{a-1} - 12 \sum_{u'=0}^{2^{a-1}-2} \sum_{v'=u'+1}^{2^{a-1}-1} (v'-u')\\ &= 2^{a-1} - 2^{3a-2}.\end{aligned}$$

Thus we have

$$I_3(2^{a-1}) = \frac{1}{10\cdot 2^{4a}} - \frac{1}{2\cdot 2^{4a}} + \frac{1}{2^{4a}} - \frac{1}{2\cdot 2^{2a}} = \frac{3}{5\cdot 2^{4a}} - \frac{1}{2\cdot 2^{2a}}.$$

Let now $0 \le j \le a-1$ be the largest number such that $\kappa_{j-1} = 1$, i.e., $k = 2^{a-1} + 2^{j-1} + \kappa_{j-2} 2^{j-2} + \cdots + \kappa_0$. Then we have

$$\begin{aligned}&\sum_{u=0}^{2^a-2} \sum_{v=u+1}^{2^a-1} (v-u)^3 (-1)^{\kappa_0(u_{a-1}+v_{a-1})+\cdots+\kappa_{a-1}(u_0+v_0)}\\ &= -2^{a-1} - 6 \sum_{\substack{u'=0\\2|u'}}^{2^a-4} \sum_{\substack{v'=u'+2\\2|v'}}^{2^a-2} (v'-u')(-1)^{\kappa_0(u_{a-1}+v_{a-1})+\cdots+\kappa_{j-1}(u_{a-j}+v_{a-j})}. \qquad (20)\end{aligned}$$

The last double sum can be written as

$$\sum_{\substack{u'=0\\2|u'}}^{2^a-4}\sum_{\substack{v'=u'+2\\2|v'}}^{2^a-2}(v'-u')(-1)^{\kappa_0(u_{a-1}+v_{a-1})+\cdots+\kappa_{j-1}(u_{a-j}+v_{a-j})}$$
$$= 2\sum_{n=0}^{2^{a-1}-2}\sum_{m=n+1}^{2^{a-1}-1}(m-n)(-1)^{\kappa_0(n_{a-2}+m_{a-2})+\cdots+\kappa_{j-1}(n_{a-j-1}+m_{a-j-1})}, \tag{21}$$

where $m = m_{a-2}2^{a-2}+\cdots+m_0$ and n analogously. Let $m' = m_{a-2}2^{a-2}+\cdots+m_{a-j}2^{a-j}$, $m'' = m_{a-j-2}2^{a-j-2}+\cdots+m_0$, $n' = n_{a-2}2^{a-2}+\cdots+n_{a-j}2^{a-j}$ and $n'' = n_{a-j-2}2^{a-j-2}+\cdots+n_0$.

First consider the case where $m' > n'$. We have

$$\sum_{n_{a-j-1}=0}^{1}\sum_{m_{a-j-1}=0}^{1}(m-n)(-1)^{\kappa_0(n_{a-2}+m_{a-2})+\cdots+\kappa_{j-1}(n_{a-j-1}+m_{a-j-1})}$$
$$= 2^{a-j-1}\sum_{n_{a-j-1}=0}^{1}\sum_{m_{a-j-1}=0}^{1}(m_{a-j-1}-n_{a-j-1})(-1)^{\kappa_0(n_{a-2}+m_{a-2})+\cdots+\kappa_{j-1}(n_{a-j-1}+m_{a-j-1})}$$
$$= 0.$$

Thus we are left with the case where $m' = n'$. Note that in this case

$$(-1)^{\kappa_0(n_{a-2}+m_{a-2})+\cdots+\kappa_{j-1}(n_{a-j-1}+m_{a-j-1})} = (-1)^{n_{a-j-1}+m_{a-j-1}}$$

as $\kappa_{j-1}=1$. Thus the double sum (21) is independent of $m'=n'$ and is equal to

$$2\cdot 2^{j-1}\sum_{n_{a-j-1}=0}^{1}\sum_{m_{a-j-1}=0}^{1}\sum_{n''=0}^{2^{a-j-1}-1}\sum_{m''=0}^{2^{a-j-1}-1}(m''-n''+2^{a-j-1}(m_{a-j-1}-n_{a-j-1}))(-1)^{m_{a-j-1}+n_{a-j-1}}, \tag{22}$$

where we have the additional assumption that $m_{a-j-1}2^{a-j-1}+m'' > n_{a-j-1}2^{a-j-1}+n''$. First consider the case where $m_{a-j-1}=1$ and $n_{a-j-1}=0$, then this part of (22) is

$$-2^j\sum_{n''=0}^{2^{a-j-1}-1}\sum_{m''=0}^{2^{a-j-1}-1}(2^{a-j-1}+m''-n'') = -2^j2^{3(a-j-1)}.$$

Now consider the case where $m_{a-j-1}=n_{a-j-1}$. In this case we have the assumption that $m'' > n''$ and hence this part of (22) is

$$2\cdot 2^j\sum_{n''=0}^{2^{a-j-1}-2}\sum_{m''=n''+1}^{2^{a-j-1}-1}(m''-n'') = 2^{j+1}\left(\frac{1}{48}2^{3(a-j)}-\frac{1}{12}2^{a-j}\right).$$

Thus (21) is given by

$$-2^j2^{3(a-j-1)}+2^{j+1}\left(\frac{1}{48}2^{3(a-j)}-\frac{1}{12}2^{a-j}\right) = -2^j\left(\frac{2^{a-j}}{6}+\frac{2^{3(a-j)}}{12}\right).$$

Using (19), (20), (21) and the above sum we obtain that for $k = 2^{a-1} + 2^{j-1} + \kappa_{j-2}2^{j-2} + \cdots + \kappa_0$ we have

$$I_3(k) = \frac{1}{10 \cdot 2^{4a}} - \frac{1}{2 \cdot 2^{4a}} - \frac{2}{2^{5a}} \left(2^{a-1} - 6 \cdot 2^j \left(\frac{2^{a-j}}{6} + \frac{2^{3(a-j)}}{12} \right) \right) = \frac{3}{5 \cdot 2^{4a}} + \frac{1}{2^{2(a+j)}}.$$

Thus we have

$$I_3(k) = \begin{cases} \frac{1}{10} & \text{if } k = 0, \\ \frac{3}{52^{4a}} - \frac{1}{2 \cdot 2^{2a}} & \text{if } k = 2^{a-1}, \\ \frac{3}{5 \cdot 2^{4a}} + \frac{1}{2^{2(a+j)}} & \text{if } k = 2^{a-1} + 2^{j-1} + \kappa_{j-2}2^{j-2} + \cdots + \kappa_0 \text{ with } 1 \le j \le a-1. \end{cases}$$

From (16) it now follows that

$$\begin{aligned} &\int_0^1 \int_0^1 B_4(|x-y|) \mathrm{wal}_k(y) \mathrm{wal}_k(y) \, \mathrm{d}x \, \mathrm{d}y \\ &= \begin{cases} 0 & \text{if } k = 0, \\ -\frac{1}{5 \cdot 2^{4a}} & \text{if } k = 2^{a-1} \text{ with } a \ge 1, \\ -\frac{6}{5 \cdot 2^{4a}} - \frac{1}{2 \cdot 2^{2(a+j)}} & \text{if } k = 2^{a-1} + 2^{j-1} \text{ where } 1 \le j < a, \\ -\frac{6}{5 \cdot 2^{4a}} - \frac{2}{2^{2(a+j)}} & \text{if } k = 2^{a-1} + 2^{j-1} + \kappa_{j-2}2^{j-2} + \cdots + \kappa_0 \text{ where } 1 \le j < a \\ & \text{and } \kappa_i = 1 \text{ for some } 0 \le i \le j-2. \end{cases} \end{aligned}$$

Using the above result together with (7), (12), (14) and (15) we obtain

$$\hat{K}_{\gamma_j,\mathrm{sh}}(k,k) = \begin{cases} 1 & \text{if } k = 0, \\ \frac{\gamma_j}{4 \cdot 2^{2a}} + \frac{\gamma_j^2}{120 \cdot 2^{4a}} & \text{if } k = 2^{a-1} \text{ with } a \ge 1, \\ \frac{\gamma_j^2}{12 \cdot 2^{2(a+j)}} + \frac{\gamma_j^2}{20 \cdot 2^{4a}} & \text{if } k = 2^{a-1} + 2^{j-1} + \kappa_{j-2}2^{j-2} + \cdots + \kappa_0 \text{ where } 1 \le j < a. \end{cases}$$

Note that the case $k = 2^{a-1} + 2^{j-1} + \kappa_{j-2}2^{j-2} + \cdots + \kappa_0$ in the equation for $\hat{K}^2_{\gamma_j,\mathrm{sh}}$ also includes the case where $k = 2^{a-1} + 2^{j-1}$.

Thus we have

$$K_{\boldsymbol{\gamma},\mathrm{sh}}(\boldsymbol{x}, \boldsymbol{y}) = \prod_{j=1}^{s} K_{\gamma_j,\mathrm{sh}}(x_j, y_j),$$

where

$$K_{\gamma_j,\mathrm{sh}}(x,y) = \sum_{k=0}^{\infty} \hat{K}_{\gamma_j,\mathrm{sh}}(k,k) \mathrm{wal}_k(x) \mathrm{wal}_k(y).$$

We are now ready to calculate $\hat{K}_{\gamma_j,\mathrm{sh},\phi}$ using Remark 2. First note that $\hat{K}_{\gamma_j,\mathrm{sh},\phi}(k,k) = 0$ if $\sigma(k) \equiv 1 \pmod 2$. From Remark 2 we have that $\hat{K}_{\gamma_j,\mathrm{sh},\phi}(2k+\nu(k), 2k+\nu(k)) = \hat{K}_{\gamma_j,\mathrm{sh}}(k,k)$ and hence we have

$$\hat{K}_{\gamma_j,\mathrm{sh},\phi}(k,k) = \begin{cases} 1 & \text{if } k = 0, \\ \frac{\gamma_j}{4 \cdot 2^{2a}} + \frac{\gamma_j^2}{120 \cdot 2^{4a}} & \text{if } k = 2^a + 1 \text{ with } a \ge 1, \\ \frac{\gamma_j^2}{12 \cdot 2^{2(a+j)}} + \frac{\gamma_j^2}{20 \cdot 2^{4a}} & \text{if } k = 2^a + 2^j + \kappa_{j-1}2^{j-1} + \cdots + \kappa_0 \text{ where } 1 \le j < a \\ & \text{and } \kappa_{j-1} + \cdots + \kappa_0 \equiv 0 \pmod 2. \end{cases} \tag{23}$$

Thus we have

$$K_{\boldsymbol{\gamma},\mathrm{sh},\phi}(\boldsymbol{x},\boldsymbol{y}) = \prod_{j=1}^{s} K_{\gamma_j,\mathrm{sh},\phi}(x_j,y_j), \tag{24}$$

where

$$K_{\gamma_j,\mathrm{sh},\phi}(x,y) = \sum_{k=0}^{\infty} r(k,\gamma_j)\mathrm{wal}_k(x)\mathrm{wal}_k(y), \tag{25}$$

where, for simplicity, we define $r(k,\gamma) = \hat{K}_{\gamma,\mathrm{sh},\phi}(k,k)$.

Appendix B: Simplification of the digitally shifted and folded kernel $K_{\boldsymbol{\gamma},\mathrm{sh},\phi}$

In this section we prove a closed form of the digitally shifted and folded kernel $K_{\boldsymbol{\gamma},\mathrm{sh},\phi}$. As can be seen from (24) it is enough to consider the one dimensional case, indeed we only need to find a closed form of (25). We split up the sum in (25) into three parts according to (23), namely, $k=0$, k of the form 2^a+1 and k of the form $2^a+2^j+\kappa_{j-1}2^{j-1}+\cdots+\kappa_0$, where $\kappa_0+\cdots+\kappa_{j-1} \equiv 0 \,(\mathrm{mod}\, 2)$. For $k=0$ we have $\mathrm{wal}_0(x)\mathrm{wal}_0(y)=1$, hence this part gives 1.

Now we consider $k=2^a+1$ where we sum over all $a \geq 1$. Let $x = \frac{x_1}{2}+\frac{x_2}{2^2}+\cdots$ and y analogously. Then we have $\mathrm{wal}_{2^a+1}(x)\mathrm{wal}_{2^a+1}(y) = (-1)^{x_1+y_1+x_{a+1}+y_{a+1}}$ and hence

$$\begin{aligned}
&\sum_{a=1}^{\infty} \hat{K}_{\gamma_j,\mathrm{sh},\phi}(2^a+1,2^a+1)\mathrm{wal}_{2^a+1}(x)\mathrm{wal}_{2^a+1}(y) \qquad (26)\\
&= (-1)^{x_1+y_1}\sum_{a=1}^{\infty}\left(\frac{\gamma_j}{4\cdot 2^{2a}}+\frac{\gamma_j^2}{120\cdot 2^{4a}}\right)(-1)^{x_{a+1}+y_{a+1}}\\
&= \gamma_j(-1)^{x_1+y_1}\left(\frac{1}{12}-2\sum_{\substack{a=2\\ x_a\neq y_a}}^{\infty}\frac{1}{2^{2a}}\right)+\gamma_j^2(-1)^{x_1+y_1}\left(\frac{1}{1800}-\frac{4}{15}\sum_{\substack{a=2\\ x_a\neq y_a}}^{\infty}\frac{1}{2^{4a}}\right).
\end{aligned}$$

We consider the third part where $k=2^a+2^j+\kappa_{j-1}2^{j-1}+\cdots+\kappa_0$ with $1 \leq j < a$ and $\kappa_0+\cdots+\kappa_{j-1} \equiv 0 \,(\mathrm{mod}\, 2)$. We have

$$\begin{aligned}
&\sum_{a=2}^{\infty}\sum_{j=1}^{a-1}\sum_{\substack{k=0\\ \sigma(k)\equiv 0 \,(\bmod\, 2)}}^{2^j-1} \hat{K}_{\gamma_j,\mathrm{sh},\phi}(2^a+2^j+k,2^a+2^j+k)\mathrm{wal}_{2^a+2^j+k}(x)\mathrm{wal}_{2^a+2^j+k}(y)\\
&= \sum_{a=2}^{\infty}\sum_{j=1}^{a-1}\sum_{\substack{k=0\\ \sigma(k)\equiv 0 \,(\bmod\, 2)}}^{2^j-1}\left(\frac{\gamma_j^2}{12\cdot 2^{2(a+j)}}+\frac{\gamma_j^2}{20\cdot 2^{4a}}\right)(-1)^{x_{a+1}+y_{a+1}+x_{j+1}+y_{j+1}}\mathrm{wal}_k(x)\mathrm{wal}_k(y)\\
&= \sum_{a=2}^{\infty}(-1)^{x_{a+1}+y_{a+1}}\sum_{j=1}^{a-1}(-1)^{x_{j+1}+y_{j+1}}\left(\frac{\gamma_j^2}{12\cdot 2^{2(a+j)}}+\frac{\gamma_j^2}{20\cdot 2^{4a}}\right)\sum_{\substack{k=0\\ \sigma(k)\equiv 0 \,(\bmod\, 2)}}^{2^j-1}\mathrm{wal}_k(x)\mathrm{wal}_k(y).
\end{aligned} \tag{27}$$

First observe that

$$\sum_{\substack{k=0\\ \sigma(k)\equiv 0 \ (\bmod\ 2)}}^{2^j-1} \mathrm{wal}_k(x)\mathrm{wal}_k(y) = \sum_{\substack{\kappa_0,\ldots,\kappa_{j-1}=0\\ \kappa_0+\cdots+\kappa_{j-1}\equiv 0 \ (\bmod\ 2)}}^{1} (-1)^{\kappa_0(x_1+y_1)}\cdots(-1)^{\kappa_{j-1}(x_j+y_j)}$$
$$= \begin{cases} 2^{j-1} & \text{if either } x_i = y_i \text{ for all } i=1,\ldots,j, \\ & \text{or } x_i \neq y_i \text{ for all } i=1,\ldots,j, \\ 0 & \text{otherwise.} \end{cases}$$

If $x_1 = y_1$ then let i_0 be the smallest integer such that $x_{i_0} \neq y_{i_0}$ and if $x_1 \neq y_1$ then let i_0 be the smallest integer such that $x_{i_0} = y_{i_0}$. Hence $i_0 \geq 2$. Then we can write (27) as

$$\gamma_j^2 \sum_{a=2}^{\infty}(-1)^{x_{a+1}+y_{a+1}} \sum_{j=1}^{\min(a-1,i_0-1)} (-1)^{x_{j+1}+y_{j+1}} \left(\frac{1}{12\cdot 2^{2(a+j)}} + \frac{1}{20\cdot 2^{4a}}\right) 2^{j-1}$$
$$= \frac{\gamma_j^2}{40}\sum_{a=2}^{\infty}\frac{(-1)^{x_{a+1}+y_{a+1}}}{2^{4a}} \sum_{j=1}^{\min(a-1,i_0-1)} 2^j(-1)^{x_{j+1}+y_{j+1}}$$
$$+\frac{\gamma_j^2}{24}\sum_{a=2}^{\infty}\frac{(-1)^{x_{a+1}+y_{a+1}}}{2^{2a}} \sum_{j=1}^{\min(a-1,i_0-1)} \frac{(-1)^{x_{j+1}+y_{j+1}}}{2^j}. \tag{28}$$

We have for $a < i_0$ that $\sum_{j=1}^{\min(a-1,i_0-1)} 2^j(-1)^{x_{j+1}+y_{j+1}} = (-1)^{x_1+y_1}\sum_{j=1}^{a-1} 2^j = (-1)^{x_1+y_1}(2^a - 2)$ and for $a \geq i_0$ we have $\sum_{j=1}^{\min(a-1,i_0-1)} 2^j(-1)^{x_{j+1}+y_{j+1}} = (-1)^{x_1+y_1}(\sum_{j=1}^{i_0-2} 2^j - 2^{i_0-1}) = -(-1)^{x_1+y_1}2$. Hence we have

$$\sum_{a=2}^{\infty}\frac{(-1)^{x_{a+1}+y_{a+1}}}{2^{4a}} \sum_{j=1}^{\min(a-1,i_0-1)} 2^j(-1)^{x_{j+1}+y_{j+1}}$$
$$= (-1)^{x_1+y_1}\left(\sum_{a=2}^{i_0-1}\frac{(-1)^{x_{a+1}+y_{a+1}}(2^a-2)}{2^{4a}} - 2\sum_{a=i_0}^{\infty}\frac{(-1)^{x_{a+1}+y_{a+1}}}{2^{4a}}\right)$$
$$= \frac{1}{105} + \frac{992}{15\cdot 2^{4i_0}} - \frac{120}{7\cdot 2^{3i_0}} - 2(-1)^{x_1+y_1}\left(\sum_{a=i_0}^{\infty}\frac{1}{2^{4a}} - 2\sum_{\substack{a=i_0\\ x_{a+1}\neq y_{a+1}}}^{\infty}\frac{1}{2^{4a}}\right)$$
$$= \frac{1}{105} + \frac{992}{15\cdot 2^{4i_0}} - \frac{120}{7\cdot 2^{3i_0}} - (-1)^{x_1+y_1}\left(\frac{32}{15\cdot 2^{4i_0}} - 64\sum_{\substack{a=i_0+1\\ x_a\neq y_a}}^{\infty}\frac{1}{2^{4a}}\right), \tag{29}$$

where for $x = y$ the last sum is defined as 0.

Further, if $a < i_0$ we have $\sum_{j=1}^{\min(a-1,i_0-1)}\frac{(-1)^{x_{j+1}+y_{j+1}}}{2^j} = (-1)^{x_1+y_1}\sum_{j=1}^{a-1}\frac{1}{2^j} = (-1)^{x_1+y_1}(1-\frac{1}{2^{a-1}})$ and if $a \geq i_0$ we have $\sum_{j=1}^{\min(a-1,i_0-1)}\frac{(-1)^{x_{j+1}+y_{j+1}}}{2^j} = (-1)^{x_1+y_1}(\sum_{j=1}^{i_0-2}\frac{1}{2^j} - \frac{1}{2^{i_0-1}}) =$

$(-1)^{x_1+y_1}(1-\frac{6}{2^{i_0}})$. Hence we have

$$\begin{aligned}
&\sum_{a=2}^{\infty}\frac{(-1)^{x_{a+1}+y_{a+1}}}{2^{2a}}\sum_{j=1}^{\min(a-1,i_0-1)}\frac{(-1)^{x_{j+1}+y_{j+1}}}{2^j}\\
&= (-1)^{x_1+y_1}\left(\sum_{a=2}^{i_0-1}\frac{(-1)^{x_{a+1}+y_{a+1}}}{2^{2a}}\left(1-\frac{2}{2^a}\right)+\left(1-\frac{6}{2^{i_0}}\right)\sum_{a=i_0}^{\infty}\frac{(-1)^{x_{a+1}+y_{a+1}}}{2^{2a}}\right)\\
&= \frac{1}{21}+\frac{240}{7\cdot 2^{3i_0}}-\frac{28}{3\cdot 2^{2i_0}}+(-1)^{x_1+y_1}\left(1-\frac{6}{2^{i_0}}\right)\left(\sum_{a=i_0}^{\infty}\frac{1}{2^{2a}}-2\sum_{\substack{a=i_0\\ x_{a+1}\neq y_{a+1}}}^{\infty}\frac{1}{2^{2a}}\right)\\
&= \frac{1}{21}+\frac{240}{7\cdot 2^{3i_0}}-\frac{28}{3\cdot 2^{2i_0}}+(-1)^{x_1+y_1}\left(\frac{4}{3\cdot 2^{2i_0}}-\frac{8}{2^{3i_0}}-\left(8-\frac{48}{2^{i_0}}\right)\sum_{\substack{a=i_0+1\\ x_a\neq y_a}}^{\infty}\frac{1}{2^{2a}}\right) \quad (30)
\end{aligned}$$

where again the last sum is defined as 0 if $x=y$.

Thus we obtain from (28), (29) and (30) that (27) is equal to

$$\begin{aligned}
&\gamma_j^2\left(\frac{1}{450}-\frac{7}{18\cdot 2^{2i_0}}+\frac{1}{2^{3i_0}}+\frac{124}{75\cdot 2^{4i_0}}\right)\\
&\quad+\gamma_j^2(-1)^{x_1+y_1}\left(\frac{1}{18\cdot 2^{2i_0}}-\frac{1}{3\cdot 2^{3i_0}}-\frac{4}{75\cdot 2^{4i_0}}-\left(\frac{1}{3}-\frac{2}{2^{i_0}}\right)\sum_{\substack{a=i_0+1\\ x_a\neq y_a}}^{\infty}\frac{1}{2^{2a}}+\frac{8}{5}\sum_{\substack{a=i_0+1\\ x_a\neq y_a}}^{\infty}\frac{1}{2^{4a}}\right).
\end{aligned}$$

Hence the following result follows from (26) and (27).

Lemma 7 *Let $x=\frac{x_1}{2}+\frac{x_2}{2^2}+\cdots$ and $y=\frac{y_1}{2}+\frac{y_2}{2^2}+\cdots$ be the base 2 representation of x and y such that for infinitely many $i\in\mathbb{N}$ we have $x_i=0$ and analogously for y. If there is an integer $i>1$ such that $x_1+x_i+y_1+y_i\equiv 1 \pmod 2$ then let i_0 be the smallest such integer, otherwise set $i_0=\infty$. Then we have*

$$\begin{aligned}
&K_{\gamma_j,\mathrm{sh},\phi}(x,y)\\
&= 1+\gamma_j(-1)^{x_1+y_1}\left(\frac{1}{12}-2\sum_{\substack{a=2\\ x_a\neq y_a}}^{\infty}\frac{1}{2^{2a}}\right)+\gamma_j^2\left(\frac{1}{450}-\frac{7}{18\cdot 2^{2i_0}}+\frac{1}{2^{3i_0}}+\frac{124}{75\cdot 2^{4i_0}}\right)\\
&\quad+\gamma_j^2(-1)^{x_1+y_1}\left(\frac{1}{1800}+\frac{1}{18\cdot 2^{2i_0}}-\frac{1}{3\cdot 2^{3i_0}}-\frac{4}{75\cdot 2^{4i_0}}-\left(\frac{1}{3}-\frac{2}{2^{i_0}}\right)\sum_{\substack{a=i_0+1\\ x_a\neq y_a}}^{\infty}\frac{1}{2^{2a}}\right.\\
&\quad\left.+\frac{8}{5}\sum_{\substack{a=i_0+1\\ x_a\neq y_a}}^{\infty}\frac{1}{2^{4a}}-\frac{4}{15}\sum_{\substack{a=2\\ x_a\neq y_a}}^{\infty}\frac{1}{2^{4a}}\right),
\end{aligned}$$

where if $x_a=y_a$ for all $a\geq c$ or if $c=\infty$ we define $\sum_{\substack{a=c\\ x_a\neq y_a}}^{\infty}\frac{1}{b^a}=0$.

Remark 5 In practice we want to compute the value of the kernel at the quadrature points $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1} \in [0,1)^s$, especially where the quadrature points stem from a digital net. Let $\boldsymbol{x}_n = (x_{1,n}, \ldots, x_{s,n})$ with $x_{j,n} = \frac{x_{j,n,1}}{2} + \frac{x_{j,n,2}}{2^2} + \cdots$. For digital nets, as can readily be seen, we have $x_{j,n,i} = 0$ for all $i > m$. Hence, if we want to evaluate $K_{s,\boldsymbol{\gamma},\mathrm{sh},\phi}(\boldsymbol{x}_n, \boldsymbol{x}_h)$ for some $0 \le n, h \le 2^m - 1$ (in case the point set is a digital net over $\mathbb{Z}_2$) the sum $\sum_{\substack{a=c \\ x_{j,n,a} \neq x_{j,h,a}}}^{\infty} \frac{1}{b^a} = \sum_{\substack{a=c \\ x_{j,n,a} \neq x_{j,h,a}}}^{m} \frac{1}{b^a}$, that is, the sum is finite and can be evaluated in $\mathcal{O}(m)$ operations.

By the way, similar arguments as above can be used to find the following simplified version of the shift invariant kernel: for $x \neq y$ we have

$$\begin{aligned} & K_{\gamma_j,\mathrm{sh}}(x,y) \\ = \; & 1 + \gamma \left(\frac{1}{12} - \frac{1}{2} \sum_{\substack{a=1 \\ x_a \neq y_a}}^{\infty} \frac{1}{2^{2a}} \right) \\ & + \gamma^2 \left(\frac{1}{360} - \frac{4}{225 \cdot 2^{3i_0}} + \left(\frac{1}{12} - \frac{2^{i_0}}{20} \right) \sum_{\substack{a=1 \\ x_a \neq y_a}}^{\infty} \frac{1}{2^{4a}} + \frac{1}{12} \left(-1 + \frac{2}{2^{i_0}} \right) \sum_{\substack{a=1 \\ x_a \neq y_a}}^{\infty} \frac{1}{2^{2a}} \right) \end{aligned}$$

where i_0 is defined as the smallest integer such that $x_{i_0} \neq y_{i_0}$ and for $x = y$ we have

$$K_{\gamma_j,\mathrm{sh}}(x,x) = 1 + \frac{\gamma_j}{12} + \frac{491\gamma_j^2}{567000}.$$

Appendix C: Some useful lemmas

In the following two lemmas we prove upper bounds on the sum of all Walsh coefficients $r(k,\gamma)$.

Lemma 8 *For all* $\frac{1}{4} < \lambda \le 1$ *we have*

$$\sum_{k=1}^{\infty} r^{\lambda}(k,\gamma) \le \gamma^{\lambda}\zeta_{\lambda} + \gamma^{2\lambda}\widehat{\tau}_{\lambda},$$

where

$$\zeta_{\lambda} := \frac{1}{2^{2\lambda}(2^{2\lambda}-1)}$$

and

$$\widehat{\tau}_{\lambda} := \frac{3^{\lambda} + 5^{\lambda} + 20^{\lambda}}{60^{\lambda}(2 - 3 \cdot 16^{\lambda} + 256^{\lambda})} + \frac{1}{120^{\lambda}(2^{4\lambda}-1)}.$$

Further we have equality if $\lambda = 1$*, i.e.,*

$$\sum_{k=1}^{\infty} r(k,\gamma) = \frac{\gamma}{12} + \frac{\gamma^2}{360}.$$

Proof. We have

$$
\begin{aligned}
\sum_{k=1}^{\infty} r^{\lambda}(k,\gamma) &= \sum_{a=1}^{\infty} r^{\lambda}(2^a+1,\gamma) + \sum_{a=2}^{\infty}\sum_{j=1}^{a-1} \sum_{\substack{l=0 \\ \sigma(l)\equiv 0 \,(\bmod\, 2)}}^{2^j-1} r^{\lambda}(2^a+2^j+l,\gamma) \\
&= \sum_{a=1}^{\infty} r^{\lambda}(2^a+1,\gamma) + \sum_{a=2}^{\infty}\sum_{j=1}^{a-1} r^{\lambda}(2^a+2^j,\gamma) 2^{j-1} \\
&=: \Sigma_1 + \Sigma_2.
\end{aligned}
$$

By Jensen's inequality for any $0<\lambda\le 1$ we have

$$r^{\lambda}(2^a+1,\gamma) \le \frac{\gamma^{\lambda}}{4^{\lambda}2^{2\lambda a}} + \frac{\gamma^{2\lambda}}{120^{\lambda}2^{4\lambda a}}$$

and therefore

$$\Sigma_1 \le \frac{\gamma^{\lambda}}{4^{\lambda}}\sum_{a=1}^{\infty}\frac{1}{2^{2\lambda a}} + \frac{\gamma^{2\lambda}}{120^{\lambda}}\sum_{a=1}^{\infty}\frac{1}{2^{4\lambda a}} = \frac{\gamma^{\lambda}}{2^{2\lambda}(2^{2\lambda}-1)} + \frac{\gamma^{2\lambda}}{120^{\lambda}(2^{4\lambda}-1)}.$$

Again by Jensen's inequality for any $0<\lambda\le 1$ we have

$$r^{\lambda}(2^a+2^j,\gamma) \le \frac{\gamma^{2\lambda}}{12^{\lambda}2^{2\lambda(a+j)}} + \frac{\gamma^{2\lambda}}{20^{\lambda}2^{4\lambda a}}$$

and therefore we have for $\frac{1}{4}<\lambda\le 1$ and $\lambda\neq\frac{1}{2}$ that

$$
\begin{aligned}
\Sigma_2 &\le \gamma^{2\lambda}\sum_{a=2}^{\infty}\sum_{j=1}^{a-1} 2^{j-1}\left(\frac{1}{12^{\lambda}2^{2\lambda(a+j)}} + \frac{1}{20^{\lambda}2^{4\lambda a}}\right) \\
&= \frac{\gamma^{2\lambda}}{48^{\lambda}(2^{2\lambda}-2)}\left(\frac{1}{2^{2\lambda}-1} - \frac{2}{2^{4\lambda}-2}\right) + \frac{\gamma^{2\lambda}}{320^{\lambda}}\left(\frac{2^2}{2^{4\lambda}-2} - \frac{2}{2^{4\lambda}-1}\right) \\
&= \gamma^{2\lambda}\frac{3^{\lambda}+5^{\lambda}+20^{\lambda}}{60^{\lambda}(2-3\cdot 16^{\lambda}+256^{\lambda})}.
\end{aligned}
$$

It is easily verified that this also holds for $\lambda=\frac{1}{2}$. The result follows by adding the bounds on Σ_1 and Σ_2. □

Lemma 9 *For all* $\frac{1}{4}<\lambda\le 1$ *and* $m\in\mathbb{N}$ *we have*

$$\sum_{l=1}^{\infty} r^{\lambda}(2^m l,\gamma) \le \frac{1}{2^{4m\lambda}}\gamma^{2\lambda}\tau_{\lambda},$$

where

$$\tau_{\lambda} := \frac{1}{48^{\lambda}} + \frac{1}{320^{\lambda}} + \frac{3^{\lambda}+20^{\lambda}+80^{\lambda}}{960^{\lambda}(16^{\lambda}-1)} + \frac{3^{\lambda}+5^{\lambda}+20^{\lambda}}{60^{\lambda}(2-3\cdot 16^{\lambda}+256^{\lambda})}.$$

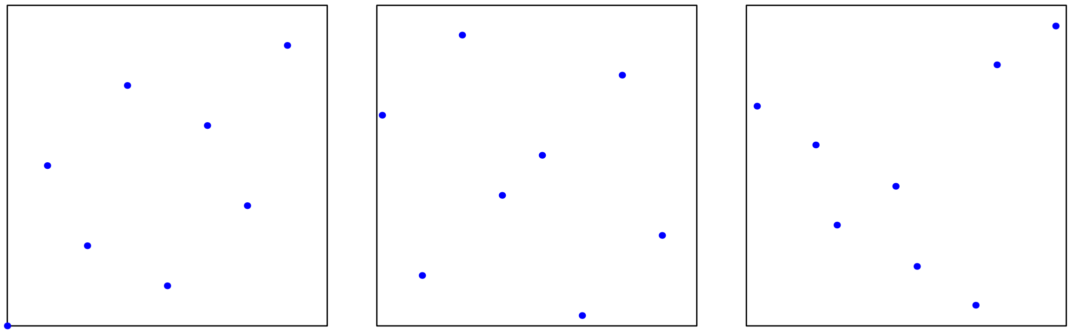

Figure 5: Hammersley net with 2^3 points, shifted version, shifted & folded version.

Proof. Note that in order for $r(k,\gamma)$ not to be zero k must either be of the form 2^a+1 or 2^a+2^j+u with $1 \le j < a$ and $0 \le u < 2^j$ with $\sigma(u) \equiv 0 \pmod 2$. As we now sum $r^\lambda(2^m l,\gamma)$ it follows that we only have to consider the latter case. Note that for $l \ge 1$ it follows from Jensen's inequality that $r^\lambda(2^m l,\gamma) \le \frac{\gamma^{2\lambda}}{2^{4m\lambda}}(\frac{1}{12^\lambda 2^{2\lambda(a+j)}} + \frac{1}{20^\lambda 2^{4\lambda a}})$. Hence we have

$$\begin{aligned}
\sum_{l=1}^\infty r^\lambda(2^m l,\gamma) &\le \frac{\gamma^{2\lambda}}{2^{4m\lambda}} \sum_{a=2}^\infty \sum_{j=0}^{a-1} \sum_{\substack{u=0 \\ \sigma(u)\equiv 0 \pmod 2}}^{2^j-1} \left(\frac{1}{12^\lambda 2^{2\lambda(a+j)}} + \frac{1}{20^\lambda 2^{4\lambda a}}\right) \\
&= \frac{\gamma^{2\lambda}}{2^{4m\lambda}} \left(\sum_{a=2}^\infty \left[\frac{1}{12^\lambda 2^{2\lambda a}} + \frac{1}{20^\lambda 2^{4\lambda a}}\right] + \sum_{a=2}^\infty \sum_{j=1}^{a-1} \left[\frac{2^{j-1}}{12^\lambda 2^{2\lambda(a+j)}} + \frac{2^{j-1}}{20^\lambda 2^{4\lambda a}}\right]\right) \\
&= \frac{\gamma^{2\lambda}}{2^{4m\lambda}} \left(\frac{3^\lambda + 20^\lambda + 80^\lambda}{960^\lambda(16^\lambda - 1)} + \frac{3^\lambda + 5^\lambda + 20^\lambda}{60^\lambda(2 - 3\cdot 16^\lambda + 256^\lambda)}\right).
\end{aligned}$$

The result follows. □

Appendix D: Some pictures of digitally shifted and then folded two dimensional digital nets

In the following we present some pictures which show how the original digital net is changed by first digitally shifting and then folding using the tent transformation.

References

[1] Chrestenson, H.E., A class of generalized Walsh functions. Pacific J. Math., **5**: 17–31, 1955.

[2] Dick, J., Kuo, F.Y., Pillichshammer, F., and Sloan, I.H., Construction algorithms for polynomial lattice rules for multivariate integration. Math. Comp., **74**: 1895–1921, 2005.

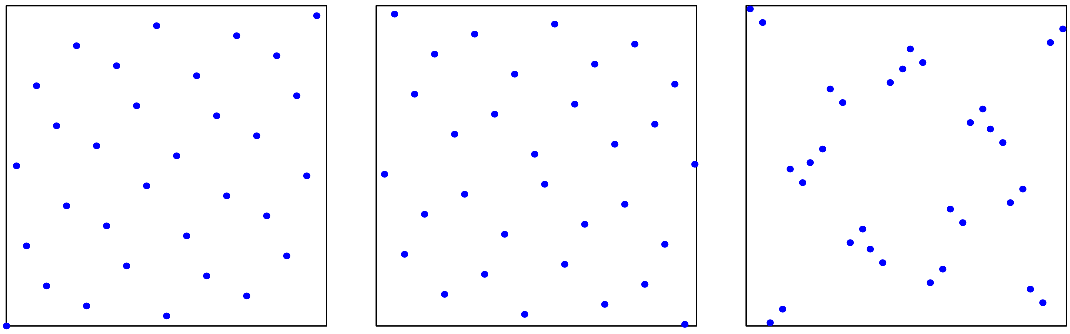

Figure 6: Hammersley net with 2^5 points, shifted version, shifted & folded version.

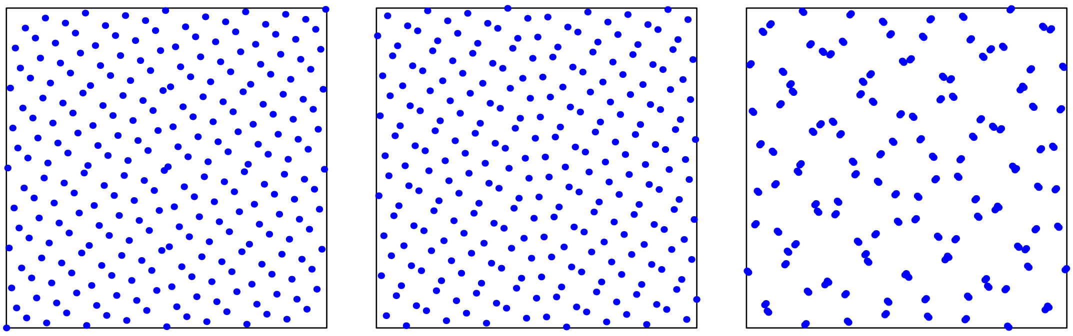

Figure 7: Hammersley net with 2^8 points, shifted version, shifted & folded version.

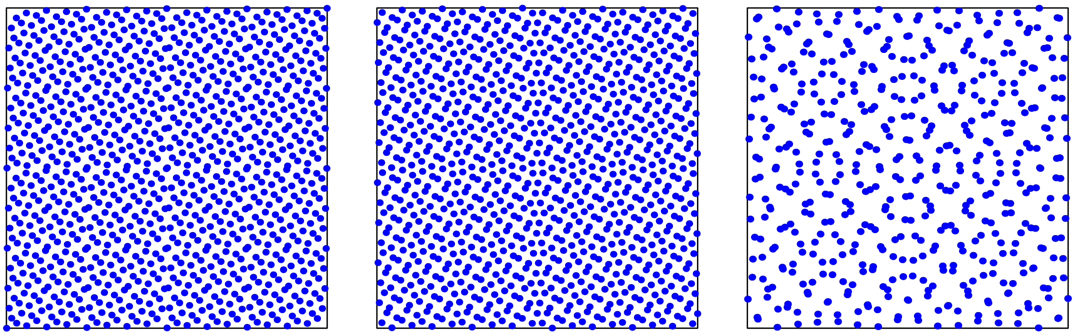

Figure 8: Hammersley net with 2^{10} points, shifted version, shifted & folded version.

[3] Dick, J., Pillichshammer, F., Multivariate integration in weighted Hilbert spaces based on Walsh functions and weighted Sobolev spaces. J. Complexity, **21**: 149–195, 2005.

[4] Dick, J., Pillichshammer, F., On the mean square weighted L_2 discrepancy of randomized digital (t, m, s)-nets over $\mathbb{Z}_2$. Acta Arith., **117**: 371–403, 2005.

[5] Faure H., Discrépance de Suites Associées à un Système de Numération (en Dimension s). Acta Arith., **41**: 337–351, 1982.

[6] Hickernell, F.J., Obtaining $O(N^{-2+\epsilon})$ convergence for lattice quadrature rules. In: Fang, K.T., Hickernell, K.J. and Niederreiter, H., eds., *Monte Carlo and Quasi-Monte Carlo Methods 2000.* Springer-Verlag, Berlin, 274–289, 2002.

[7] Larcher, G., Pillichshammer, F., Walsh series analysis of the L_2-discrepancy of symmetrisized point sets. Monatsh. Math., **132**: 1–18, 2001.

[8] Matoušek, J., *Geometric Discrepancy.* Algorithms and Combinatorics 18, Springer Verlag, Berlin, 1999.

[9] Niederreiter, H., Point sets and sequences with small discrepancy. Monatsh. Math., **104**: 273–337, 1987.

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

[11] Niederreiter, H., Low-discrepancy point sets obtained by digital constructions over finite fields. Czechoslovak Math. J., **42**: 143–166, 1992.

[12] Niederreiter, H., Finite fields, pseudorandom numbers, and quasirandom points. In: *Finite Fields, Coding Theory, and Advances in Communications and Computing* (Mullen, G.L., and Shiue, P.J.-S., eds.), pp. 375–394, Dekker, New York, 1992.

[13] Niederreiter, H., Constructions of (t, m, s)-nets and (t, s)-sequences. Finite Fields Appl., **11**: 578–600, 2005.

[14] Owen, A.B., Randomly permuted (t, m, s)-nets and (t, s)-sequences. Monte Carlo and quasi-Monte Carlo methods in scientific computing (Las Vegas, NV, 1994), Lecture Notes in Statist., **106**: 299–317, Springer, New York, 1995.

[15] Pirsic, G., Dick, J. and Pillichshammer, F., Cyclic digital nets, hyperplane nets and multivariate integration in Sobolev spaces. SIAM J. Numer. Anal., **44**, 385–411, 2006.

[16] Sloan, I.H., Joe, S., *Lattice Methods for Multiple Integration.* Clarendon Press, Oxford, 1994.

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

[18] Sloan, I.H., Woźniakowski, H., Tractability of multivariate integration for weighted Korobov classes. J. Complexity, **17**: 697–721, 2001.

[19] Walsh, J.L., A closed set of normal orthogonal functions. Amer. J. Math., **55**: 5–24, 1923.

[20] Yue, R.X., Hickernell, F.J., The discrepancy and gain coefficients of scrambled digital nets. J. Complexity, **18**: 135–151, 2002.

Author's Addresses:

Ligia L. Cristea, Institut für Finanzmathematik, Universität Linz, Altenbergstraße 69, A-4040 Linz, Austria. Email: ligia-loretta.cristea@jku.at

Josef Dick, School of Mathematics, University of New South Wales, Sydney 2052, Australia. Email: josi@maths.unsw.edu.au

Gunther Leobacher, Institut für Finanzmathematik, Universität Linz, Altenbergstraße 69, A-4040 Linz, Austria. Email: gunther.leobacher@jku.at

Friedrich Pillichshammer, Institut für Finanzmathematik, Universität Linz, Altenbergstraße 69, A-4040 Linz, Austria. Email: friedrich.pillichshammer@jku.at