ACTA ARITHMETICA
117.4 (2005)

On the mean square weighted $\mathcal{L}_2$ discrepancy of randomized digital (t, m, s)-nets over $\mathbb{Z}_2$

by

JOSEF DICK (Sydney) and FRIEDRICH PILLICHSHAMMER (Linz)

1. Introduction. We study distribution properties of point sets in the s-dimensional unit cube $[0,1)^s$. There are various measures for the equidistribution of such point sets (see for example [7, 10, 11, 16, 19]). The one we consider here is based on the following function. For a set $P_N = \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{N-1}\}$ of points in the s-dimensional unit cube $[0,1)^s$ the discrepancy function is defined as

$$\Delta(t_1, \ldots, t_s) = \frac{A_N([0,t_1) \times \cdots \times [0,t_s))}{N} - t_1 \cdots t_s,$$

where $0 \le t_j \le 1$ and $A_N([0,t_1) \times \cdots \times [0,t_s))$ denotes the number of indices n with $\boldsymbol{x}_n \in [0,t_1) \times \cdots \times [0,t_s)$.

The discrepancy function measures the difference of the portion of points in an axis parallel box containing the origin and the volume of this box. Hence it is a measure of the irregularity of distribution of a point set in $[0,1)^s$. There are of course other functions which serve a comparable purpose, though this function has drawn a great deal of attention as various connections with applications have been pointed out, notably in numerical integration of functions (see for example [19, 29]). Further, we can use different norms of the discrepancy function, again yielding different quality measures. Amongst those norms especially the $\mathcal{L}_2$ norm and the $\mathcal{L}_\infty$ norm are of considerable interest and have been studied extensively (see for example [19, 29]). In the following we introduce some notation and define the weighted $\mathcal{L}_2$ discrepancy of a point set, which will be the focus of this paper.

Let $D = \{1, \ldots, s\}$. For $\mathfrak{u} \subseteq D$ let $\gamma_{\mathfrak{u}}$ be a non-negative real number, $|\mathfrak{u}|$ the cardinality of $\mathfrak{u}$ and for a vector $\boldsymbol{x} \in [0,1)^s$ let $\boldsymbol{x}_{\mathfrak{u}}$ denote the vector from

2000 *Mathematics Subject Classification*: 11K38, 11K06.

Key words and phrases: digital nets, weighted $\mathcal{L}_2$ discrepancy, Walsh series analysis.

The first author is supported by the Australian Research Council under its Center of Excellence Program.

The second author is supported by the Austrian Research Foundation (FWF), Project S8305 and Project P17022-N12.

$[0,1)^{|\mathfrak{u}|}$ containing all components of $\boldsymbol{x}$ whose indices are in $\mathfrak{u}$. Further let $d\boldsymbol{x}_{\mathfrak{u}} = \prod_{j\in\mathfrak{u}} dx_j$ and let $(\boldsymbol{x}_{\mathfrak{u}},1)$ be the vector from $[0,1)^s$ with all components whose indices are not in $\mathfrak{u}$ replaced by 1. Then the *weighted $\mathcal{L}_2$ discrepancy* of P_N is defined as (see [29])

$$\mathcal{L}_{2,N,\gamma}(P_N) = \Big(\sum_{\substack{\mathfrak{u}\subseteq D\\ \mathfrak{u}\neq\emptyset}} \gamma_{\mathfrak{u}} \int_{[0,1]^{|\mathfrak{u}|}} \Delta((\boldsymbol{x}_{\mathfrak{u}},1))^2 \, d\boldsymbol{x}_{\mathfrak{u}}\Big)^{1/2}. \tag{1}$$

This is a generalization of the classical $\mathcal{L}_2$ discrepancy. By choosing $\gamma_D = 1$ and $\gamma_{\mathfrak{u}} = 0$ for all $\mathfrak{u} \subset D$ we obtain the classical $\mathcal{L}_2$ discrepancy and if we choose $\gamma_{\mathfrak{u}} = 1$ for all $\mathfrak{u} \subseteq D$ we obtain the unweighted $\mathcal{L}_2$ discrepancy. Note that in this definition we also include the lower dimensional projections (see [9]). In [17] it has been pointed out that the classical $\mathcal{L}_2$ discrepancy of N copies of the point $(1,\dots,1)$ can almost yield the best value if the dimension is high compared to N. (Note that the $\mathcal{L}_2$ discrepancy does not change by considering point sets in $[0,1]^s$ rather than $[0,1)^s$.) Such a point set is obviously not well distributed in an intuitive sense. Including the lower dimensional projections much reduces this effect. The weights $\gamma_{\mathfrak{u}}$ are then introduced to modify the importance of the discrepancy of the projections, with the intention to adjust the measure to the usage of the point set (see [6, 29]). For example it has been observed that in many applications the higher dimensional projections are considerably less important than the lower dimensional ones.

There is a well known formula for the classical $\mathcal{L}_2$ discrepancy of a point set by Warnock (see for example [16]), which can easily be generalized to a formula for the weighted $\mathcal{L}_2$ discrepancy. This formula is given in the following proposition (for a hint on how to prove this formula see for example [15] or [16]).

Proposition 1. *Let $P_N = \{\boldsymbol{x}_0,\dots,\boldsymbol{x}_{N-1}\}$ be a point set in $[0,1)^s$. Then*

$$\mathcal{L}^2_{2,N,\gamma}(P_N) = \sum_{\substack{\mathfrak{u}\subseteq D\\ \mathfrak{u}\neq\emptyset}} \gamma_{\mathfrak{u}} \Big[\frac{1}{3^{|\mathfrak{u}|}} - \frac{2}{N}\sum_{n=0}^{N-1}\prod_{j\in\mathfrak{u}} \frac{1-x_{n,j}^2}{2} + \frac{1}{N^2}\sum_{n,m=0}^{N-1}\prod_{j\in\mathfrak{u}} \min(1-x_{n,j}, 1-x_{m,j})\Big],$$

where $x_{n,j}$ is the jth component of $\boldsymbol{x}_n$.

As the choice of weights is determined by the task (for example, approximating the integral of a function) and therefore not known a priori, we wish to find point sets which "work well" for many (if not all) choices of weights.

Several construction methods for point sets in the unit cube with good distribution properties are known. The method considered here builds on the concept of (t, m, s)-nets. A detailed theory of such nets was developed in Niederreiter [18] (see also [19, Chapter 4] for a survey). The (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 set P 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.

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 [18] proved that if a $(0, m, s)$-net in base b exists, then $s-1 \leq b$. Faure [8] provided a construction of $(0, m, s)$-nets in prime base $p \geq s-1$ and Niederreiter [18] 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 $s = 2$ 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 based on the concept of digital point sets. 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 [12, 13] or [19]. (It has been observed that a small base b and higher t-value yield 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 $s \geq 1$, $m \geq 1$ and $0 \leq t \leq m$ be integers. Choose s $m \times m$ matrices $C_1, \ldots, C_s$ over $\mathbb{Z}_2$ with the following property: for any integers $d_1, \ldots, d_s \geq 0$ with $d_1 + \cdots + d_s = m - t$ the system of

$$\begin{aligned}
&\text{the first } d_1 \text{ rows of } C_1, \text{ together with}\\
&\qquad\qquad\vdots\\
&\text{the first } d_{s-1} \text{ rows of } C_{s-1}, \text{ together with}\\
&\text{the first } d_s \text{ rows of } C_s
\end{aligned}$$

is linearly independent over $\mathbb{Z}_2$.

Consider the following construction principle for sets 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 $\vec{n} = (n_0, \ldots, n_{m-1})^T$ of digits of n in $\mathbb{Z}_2$,

$$C_j \vec{n} =: (y_1^{(j)}, \ldots, y_m^{(j)})^T.$$

Now we set

$$x_n^{(j)} := \frac{y_1^{(j)}}{2} + \cdots + \frac{y_m^{(j)}}{2^m}$$

and

$$\boldsymbol{x}_n = (x_n^{(1)}, \ldots, x_n^{(s)}).$$

The point set $\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{2^m-1}\}$ is called a *digital* (t, m, s)*-net over* $\mathbb{Z}_2$ and the matrices $C_1, \ldots, C_s$ are called the *generating matrices* of the digital net.

Note that any digital (t, m, s)-net over $\mathbb{Z}_2$ is a (t, m, s)-net in base 2 as shown by Niederreiter [19]. Further it follows from Definition 2 that any d-dimensional projection of a digital (t, m, s)-net over $\mathbb{Z}_2$ is a digital (t, m, d)-net over $\mathbb{Z}_2$.

For practical applications it is often useful to have a random element in the point set used (see [16]). 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 so 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 [16, 22, 32]). The method considered in this paper is a digital shift of depth m (see also [16]). The aim of the paper is then to analyze the expected value of the weighted $\mathcal{L}_2$ discrepancy of digitally shifted digital (t, m, s)-nets.

In the following we introduce the digital shift of depth m for the one-dimensional case. For higher dimensions each coordinate is randomized independently and therefore one just needs to apply the one-dimensional randomization method to each coordinate independently.

Let $P_{2^m} = \{x_0, \ldots, x_{2^m-1}\}$ be a digital $(t, m, 1)$-net over $\mathbb{Z}_2$ generated by the matrix C. Let

$$x_n = \frac{x_{n,1}}{2} + \frac{x_{n,2}}{2^2} + \cdots$$

be the dyadic digit expansion of x_n. In [5] a randomization method was considered which uses a digital shift $\sigma = \sigma_1/2 + \sigma_2/2^2 + \cdots$, where $\sigma \in [0, 1)$ was chosen randomly. Here we modify this method in the following way: first we choose the digits $\sigma_1, \ldots, \sigma_m \in \{0, 1\}$ i.i.d. Then we define

$$z_{n,i} \equiv x_{n,i} + \sigma_i \pmod 2 \qquad \text{for } i = 1, \ldots, m$$

with $z_{n,i} \in \{0, 1\}$. Further, for $n = 0, \ldots, 2^m - 1$, we choose $\delta_n \in [0, 1/2^m)$ i.i.d. Then the randomized point set $\widetilde{P}_{2^m} = \{z_0, \ldots, z_{2^m-1}\}$ is given by

$$z_n = \frac{z_{n,1}}{2} + \cdots + \frac{z_{n,m}}{2^m} + \delta_n.$$

This means we apply the same digital shift to the first m digits, whereas the following digits are shifted independently for each x_n. Therefore we call it a *digital shift of depth* m (see again [16]).

Sometimes we will write "digital shift" or simply "shift" instead of "digital shift of depth m". When we use a digital shift of depth m' in conjunction with digital (t,m,s)-nets we always assume that $m' = m$.

For arbitrary $s \geq 1$ it can be shown that a (t,m,s)-net in base 2 randomized by a digital shift of depth m independently in each coordinate is again a (t,m,s)-net in base 2 with the same quality parameter t. As this result is not essential for the following we omit the proof. Similar results have been shown before (see for example [5, 22]).

We conclude this section with an outline of the paper. In the subsequent section we introduce Walsh functions, which will be the main tool for the analysis of the $\mathcal{L}_2$ discrepancy. Several useful properties of these functions will be recalled.

In Section 3 we prove three main results. The first one is a formula for the mean square weighted $\mathcal{L}_2$ discrepancy of randomized digital nets. The formula is exact and involves a function of the generating matrices of the digital net. We then use this result to derive an exact formula for the mean square weighted $\mathcal{L}_2$ discrepancy of randomized digital $(0,m,s)$-nets in dimensions 2 and 3, which is in this case independent of the generating matrices. The convergence order is best possible and we compare the constant of the leading term with the lower bound given by Roth [27]. The third result is an upper bound on the mean square weighted $\mathcal{L}_2$ discrepancy of randomized digital nets in dimension $s > 3$. The difference between $s > 3$ and $s = 2,3$ is that $s > 3$ implies that $t > 0$. Therefore the exact value of the mean square weighted $\mathcal{L}_2$ discrepancy depends on the generating matrices. Still we can obtain an upper bound for this case which is independent of the generating matrices and only depends on the weights, the t-value, the number of points and the dimension s. This bound is of a simple form and easily computable. Again, the convergence order is best possible.

In Section 4 we deal with the classical $\mathcal{L}_2$ discrepancy. In 1954 Roth [27] proved that the $\mathcal{L}_2$ discrepancy of any point set in $[0,1)^s$ consisting of N elements is at least $c_1(s)(\log N)^{(s-1)/2}N^{-1}$, with a constant $c_1(s) = 2^{-2s-4}((s-1)!)^{-1/2}$. In 1980 Roth [26] also showed that there exists a set of N elements in $[0,1)^s$ with $\mathcal{L}_2$ discrepancy of at most $c_2(s)(\log N)^{(s-1)/2}N^{-1}$, with a constant $c_2(s)$ depending only on the dimension s. (See also [1] for variations of Roth's result. For dimension $s = 2$ this was proven by Davenport [4] already in 1956. Quite recently, Chen and Skriganov [2] gave concrete examples—not only existence results as Roth did—of point sets in arbitrary dimensions which achieve the minimal order of the $\mathcal{L}_2$ discrepancy.) Therefore the exact dependence on N is known. Here we are interested in the constant $c_1(s)$ of the lower bound of Roth. In a first result we extract the constant of the leading term from the previous calculations in Section 3. By a construction of Niederreiter–Xing [21] we know that for any m and s there always exists a digital $(5s, m, s)$-net over $\mathbb{Z}_2$. With an appropriate shift we find that such point sets have an $\mathcal{L}_2$ discrepancy of at most

$$\frac{(\log N)^{(s-1)/2}}{N} \frac{22^s}{(\log 2)^{(s-1)/2}((s-1)!)^{1/2}} + \mathcal{O}\left(\frac{(\log N)^{(s-2)/2}}{N}\right).$$

The constant $22^s(\log 2)^{-(s-1)/2}((s-1)!)^{-1/2}$ improves a result by Hickernell [9] considerably and seems to be the best known constant of this kind.

Secondly, we prove an upper bound on the classical $\mathcal{L}_2$ discrepancy of shifted Niederreiter–Xing nets (see [20]). We consider a sequence of shifted digital nets, where the number N of points is relatively small compared to the dimension. For this sequence of shifted digital nets we obtain an upper bound which shows that Roth's [27] lower bound is also best possible in the dimension s.

2. Walsh functions. In this section we introduce Walsh functions, which will be the main tool in our analysis of the mean square weighted $\mathcal{L}_2$ discrepancy. Again we confine ourselves to base 2 (for more information see [3, 24, 25, 30]). 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* $\mathrm{wal}_k : [0,1) \to \{-1,1\}$ 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 $x = x_1/2 + x_2/2^2 + \cdots$ (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 $\mathrm{wal}_{k_1,\ldots,k_s} : [0,1)^s \to \{-1,1\}$ 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 introduce some notation. By $\oplus$ we denote the digitwise addition modulo 2, i.e., for $x = \sum_{i=w}^{\infty} x_i/2^i$ and $y = \sum_{i=w}^{\infty} y_i/2^i$ we have

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

In the following proposition we summarize some basic properties of Walsh functions.

PROPOSITION 2. (a) *For all $k, l \in \mathbb{N}_0$ and all $x, y \in [0,1)$ we have*

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

(b) *We have*

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

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

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

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

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

(e) *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 (a)–(c) are straightforward ([24]). For (d) see [3, Lemma 1] or [24, Corollary 4] and for (e) see [3] or [24, Satz 1]. ∎

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 $\boldsymbol{x}_n = (x_{n,1}, \ldots, x_{n,s})$ and $x_{n,j} = x_{n,j,1}/2 + \cdots + x_{n,j,m}/2^m$, $1 \leq j \leq s$, $0 \leq n < 2^m$, we identify $\boldsymbol{x}_n$ with

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

and we define

$$\boldsymbol{x}_n \oplus \boldsymbol{x}_h := (x_{n,1,1} + x_{h,1,1}, \ldots, x_{n,s,m} + x_{h,s,m}) \in \mathbb{Z}_2^{ms}. \tag{2}$$

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 a subgroup of $(\mathbb{Z}_2^{ms}, \oplus)$.*

The following lemma will be very useful for our investigation. It was already shown in [5], but for completeness we include a proof.

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 $0 \leq k_1, \ldots, k_s < 2^m$ we have*

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

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

Proof. By (2) we have

$$\mathrm{wal}_{k_1,\dots,k_s}(\boldsymbol{x}_n \oplus \boldsymbol{x}_h) = \mathrm{wal}_{k_1,\dots,k_s}(\boldsymbol{x}_n)\,\mathrm{wal}_{k_1,\dots,k_s}(\boldsymbol{x}_h)$$

and hence $\mathrm{wal}_{k_1,\dots,k_s}$ is a character on $(\mathbb{Z}_2^{ms},\oplus)$. From Lemma 1 we know that the digital net $\{\boldsymbol{x}_0,\dots,\boldsymbol{x}_{2^m-1}\}$ is a subgroup of $(\mathbb{Z}_2^{ms},\oplus)$ and so

$$\sum_{n=0}^{2^m-1} \mathrm{wal}_{k_1,\dots,k_s}(\boldsymbol{x}_n) = \begin{cases} 2^m & \text{if } \mathrm{wal}_{k_1,\dots,k_s}(\boldsymbol{x}_n) = 1\ \forall n = 0,\dots,2^m-1, \\ 0 & \text{otherwise.} \end{cases}$$

Now $\mathrm{wal}_{k_1,\dots,k_s}(\boldsymbol{x}_n) = 1$ for all $n = 0,\dots,2^m-1$ iff

$$\sum_{j=1}^{s} (\vec{k}_j | \vec{x}_{n,j}) = 0 \qquad \forall n = 0,\dots,2^m-1$$

(here $(\cdot|\cdot)$ denotes the usual inner product in $\mathbb{Z}_2^m$). This means by the definition of the net that

$$\sum_{j=1}^{s} (\vec{k}_j | C_j \vec{n}) = 0 \qquad \forall n = 0,\dots,2^m-1$$

and this is satisfied iff $C_1^T \vec{k_1} + \dots + C_s^T \vec{k_s} = \vec{0}$, as claimed. ∎

3. On the mean square weighted $\mathcal{L}_2$ discrepancy of randomized nets. In the following subsection we prove a formula for the mean square weighted $\mathcal{L}_2$ discrepancy of randomized digital nets. This formula depends on the generating matrices of the digital net. Subsequently we use this formula to derive the exact value of the mean square weighted $\mathcal{L}_2$ discrepancy for digital $(0,m,s)$-nets over $\mathbb{Z}_2$ for $s = 2,3$. Next we obtain a bound for the general case, that is, for the mean square weighted $\mathcal{L}_2$ discrepancy of digital (t,m,s)-nets over $\mathbb{Z}_2$.

3.1. *A formula for the mean square weighted $\mathcal{L}_2$ discrepancy of randomized nets.* The aim of this subsection is to prove the following theorem.

THEOREM 1. *Let P_{2^m} be a digital (t,m,s)-net over $\mathbb{Z}_2$ with generating matrices $C_1,\dots,C_s$. Let $\widetilde{P}_{2^m}$ be the point set obtained after applying an i.i.d. random digital shift of depth m independently to each coordinate of each point of P_{2^m}. Then the mean square weighted $\mathcal{L}_2$ discrepancy of $\widetilde{P}_{2^m}$ is given by*

$$\mathbb{E}[\mathcal{L}^2_{2,2^m,\gamma}(\widetilde{P}_{2^m})] = \sum_{\substack{\mathfrak{u}\subseteq D\\ \mathfrak{u}\neq\emptyset}} \gamma_{\mathfrak{u}} \left[\frac{1}{2^{m+|\mathfrak{u}|}} \left(1 - \left(1 - \frac{1}{3\cdot 2^m}\right)^{|\mathfrak{u}|}\right) + \frac{1}{3^{|\mathfrak{u}|}} \sum_{\substack{\mathfrak{v}\subseteq\mathfrak{u}\\ \mathfrak{v}\neq\emptyset}} \left(\frac{3}{2}\right)^{|\mathfrak{v}|} \mathcal{B}(\mathfrak{v}) \right],$$

where for $\mathfrak{v} = \{v_1, \dots, v_e\}$ *we have*

$$\mathcal{B}(\mathfrak{v}) = \sum_{\substack{k_1,\dots,k_e=1 \\ C_{v_1}^T \vec{k}_1 + \cdots + C_{v_e}^T \vec{k}_e = \vec{0}}}^{2^m-1} \prod_{j=1}^{e} \psi(k_j),$$

with $\psi(k) = 1/(6 \cdot 4^{r(k)})$ *and* $r(k)$ *such that* $2^{r(k)} \le k < 2^{r(k)+1}$.

The proof of this theorem is based on the Walsh series representation of the formula for the $\mathcal{L}_2$ discrepancy given in Proposition 1. As we will see later, the function ψ in the theorem above is related to the Walsh coefficients of a certain function appearing in the formula for the $\mathcal{L}_2$ discrepancy. We need several lemmas.

LEMMA 3. *Let* $x_1, x_2 \in [0,1)$ *and let* $z_1, z_2 \in [0,1)$ *be the points obtained after applying an i.i.d. random digital shift of depth* m *to* x_1 *and* x_2. *Then*

$$\mathbb{E}[\mathrm{wal}_k(z_1)\,\mathrm{wal}_l(z_2)] = \begin{cases} \mathrm{wal}_k(x_1 \oplus x_2) & \text{if } 0 \le k = l < 2^m, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. Let $x_n = x_{n,1}/2 + x_{n,2}/2^2 + \cdots$ for $n = 1, 2$. Further let $\sigma_1, \dots, \sigma_m \in \{0,1\}$ be i.i.d. and for $n = 1,2$ let $\delta_n = \delta_{n,m+1}/2^{m+1} + \delta_{n,m+2}/2^{m+2} + \cdots \in [0, 1/2^m)$ be i.i.d. Then define $z_{n,i} \equiv x_{n,i} + \sigma_i \pmod 2$ for $i = 1, \dots, m$ and $z_n = z_{n,1}/2 + \cdots + z_{n,m}/2^m + \delta_n$ for $n = 1, 2$.

First let $k, l \in \mathbb{N}$, more precisely, let $k = k_u 2^u + \cdots + k_1 2 + k_0$ and $l = l_v 2^v + \cdots + l_1 2 + l_0$ be the dyadic expansions of k and l with $k_u = l_v = 1$. Further set $k_{u+1} = k_{u+2} = \cdots = 0$ and also $l_{v+1} = l_{v+2} = \cdots = 0$. Then

$$\begin{aligned}(3)\quad & \mathbb{E}[\mathrm{wal}_k(z_1)\,\mathrm{wal}_l(z_2)] \\ &= (-1)^{k_0 x_{1,1} + \cdots + k_{m-1} x_{1,m}} (-1)^{l_0 x_{2,1} + \cdots + l_{m-1} x_{2,m}} \\ &\quad \times \frac{1}{2} \sum_{\sigma_1=0}^{1} (-1)^{(k_0+l_0)\sigma_1} \cdots \frac{1}{2} \sum_{\sigma_m=0}^{1} (-1)^{(k_{m-1}+l_{m-1})\sigma_m} \\ &\quad \times \frac{1}{2} \sum_{\delta_{1,m+1}=0}^{1} (-1)^{k_m \delta_{1,m+1}} \frac{1}{2} \sum_{\delta_{1,m+2}=0}^{1} (-1)^{k_{m+1}\delta_{1,m+2}} \cdots \\ &\quad \times \frac{1}{2} \sum_{\delta_{2,m+1}=0}^{1} (-1)^{l_m \delta_{2,m+1}} \frac{1}{2} \sum_{\delta_{2,m+2}=0}^{1} (-1)^{l_{m+1}\delta_{2,m+2}} \cdots .\end{aligned}$$

(The product above consists only of finitely many factors as $k_{u+1} = k_{u+2} = \cdots = 0$ and for $\kappa \ge \max(m, u+1)$ we have $\frac{1}{2}\sum_{\delta_{1,\kappa+1}=0}^{1} (-1)^{k_\kappa \delta_{1,\kappa+1}} = 1$. The same argument holds for the last line in the equation above.)

First we consider the case where $u \geq m$. We have

$$\frac{1}{2} \sum_{\delta_{1,u+1}=0}^{1} (-1)^{k_u \delta_{1,u+1}} = 0$$

and therefore $\mathbb{E}[\mathrm{wal}_k(z_1)\,\mathrm{wal}_l(z_2)] = 0$. The same holds if $v \geq m$. Now assume that there is an $\omega \in \{0, \ldots, m-1\}$ such that $k_\omega \neq l_\omega$. Then $k_\omega + l_\omega \equiv 1 \pmod 2$ and

$$\frac{1}{2} \sum_{\sigma_{\omega+1}=0}^{1} (-1)^{(k_\omega + l_\omega)\sigma_{\omega+1}} = 0.$$

Therefore in this case $\mathbb{E}[\mathrm{wal}_k(z_1)\,\mathrm{wal}_l(z_2)] = 0$. Now let $k = l$ and $k \in \{0, \ldots, 2^m - 1\}$. It follows from (3) that

$$\mathbb{E}[\mathrm{wal}_k(z_1)\,\mathrm{wal}_k(z_2)] = (-1)^{k_0(x_{1,1}+x_{2,1})+\cdots+k_{m-1}(x_{1,m}+x_{2,m})}$$

and the result follows. ∎

In the following lemma we calculate the Walsh coefficients of the function $|z_1 - z_2|$. This function appears in the formula for the $\mathcal{L}_2$ discrepancy through the equation $\min(z_1, z_2) = \frac{1}{2}(z_1 + z_2 - |z_1 - z_2|)$.

Lemma 4. *Let $z_1, z_2 \in [0, 1)$. Then*

$$|z_1 - z_2| = \sum_{k,l=0}^{\infty} \tau(k, l)\,\mathrm{wal}_k(z_1)\,\mathrm{wal}_l(z_2),$$

where $\tau(0) := \tau(0, 0) = 1/3$ and $\tau(k) := \tau(k, k) = -1/(6 \cdot 4^{r(k)})$ for $k > 0$. For $k > 0$, $r(k)$ denotes the unique integer r such that $2^r \leq k < 2^{r+1}$.

Proof. As $|z_1 - z_2| \in \mathcal{L}_2([0, 1)^2)$ it follows from Proposition 2 that the function $|z_1 - z_2|$ can be represented by Walsh functions. We have

$$\tau(k, l) = \int_0^1\!\!\int_0^1 |z_1 - z_2|\,\mathrm{wal}_k(z_1)\,\mathrm{wal}_l(z_2)\,dz_1\,dz_2.$$

As $\mathrm{wal}_0(z) = 1$ for all $z \in [0, 1)$, we have

$$\tau(0, 0) = \int_0^1\!\!\int_0^1 |z_1 - z_2|\,dz_1\,dz_2 = \frac{1}{3}.$$

Let now $k = l > 0$ and $k = k_r 2^r + \cdots + k_1 2 + k_0$, where r is such that $k_r = 1$, $u = u_r 2^r + \cdots + u_1 2 + u_0$ and $v = v_r 2^r + \cdots + v_1 2 + v_0$. Then

$$\tau(k, k) = \int_0^1\!\!\int_0^1 |z_1 - z_2|\,\mathrm{wal}_k(z_1 \oplus z_2)\,dz_1\,dz_2$$

$$= \sum_{u=0}^{2^{r+1}-1} \sum_{v=0}^{2^{r+1}-1} (-1)^{k_0(u_r+v_r)+\cdots+k_r(u_0+v_0)} \times \int_{u/2^{r+1}}^{(u+1)/2^{r+1}} \int_{v/2^{r+1}}^{(v+1)/2^{r+1}} |z_1 - z_2| \, dz_1 \, dz_2.$$

We have the following equalities: if $0 \le u < 2^{r+1}$, then

$$\int_{u/2^{r+1}}^{(u+1)/2^{r+1}} \int_{u/2^{r+1}}^{(u+1)/2^{r+1}} |z_1 - z_2| \, dz_1 \, dz_2 = \frac{1}{3 \cdot 2^{3(r+1)}},$$

and for $0 \le u, v < 2^{r+1}$, $u \ne v$, we have

$$\int_{u/2^{r+1}}^{(u+1)/2^{r+1}} \int_{v/2^{r+1}}^{(v+1)/2^{r+1}} |z_1 - z_2| \, dz_1 \, dz_2 = \frac{|u-v|}{2^{3(r+1)}}.$$

Thus

$$\begin{aligned} \tau(k,k) &= \sum_{u=0}^{2^{r+1}-1} \frac{1}{3 \cdot 2^{3(r+1)}} + \sum_{u=0}^{2^{r+1}-1} \sum_{\substack{v=0 \\ u \ne v}}^{2^{r+1}-1} (-1)^{k_0(u_r+v_r)+\cdots+k_r(u_0+v_0)} \frac{|u-v|}{2^{3(r+1)}} \\ &= \frac{1}{3 \cdot 2^{2(r+1)}} + \frac{1}{2^{3r+2}} \sum_{u=0}^{2^{r+1}-2} \sum_{v=u+1}^{2^{r+1}-1} (-1)^{k_0(u_r+v_r)+\cdots+k_r(u_0+v_0)} (v-u). \end{aligned}$$

We define

$$\theta(u,v) = (-1)^{k_0(u_r+v_r)+\cdots+k_r(u_0+v_0)} (v-u).$$

In order to find the value of the double sum in the expression for $\tau(k,k)$ let $u = u_r 2^r + \cdots + u_1 2$ and $v = v_r 2^r + \cdots + v_1 2$, where $v > u$. We now consider the sum of $\theta(u,v), \theta(u+1,v), \theta(u,v+1)$ and $\theta(u+1,v+1)$. Observe that u and v are even, that is, $u_0 = v_0 = 0$, and $k = k_r 2^r + \cdots + k_1 2 + k_0$, where r is such that $k_r = 1$. We obtain

$$\begin{aligned} &|\theta(u,v) + \theta(u+1,v) + \theta(u,v+1) + \theta(u+1,v+1)| \\ &\quad = (v-u) - ((v+1)-u) - (v-(u+1)) + ((v+1)-(u+1)) \\ &\quad = 0. \end{aligned}$$

After applying this procedure we are left with the following terms:

$$\theta(0,1), \theta(2,3), \ldots, \theta(2^{r+1}-2, 2^{r+1}-1).$$

Observe that in all cases we have $v - u = 1$, hence $u_i = v_i$ for $i = 1, \ldots, r$ and $u_0 = 0$ and $v_0 = 1$. Therefore

$$(-1)^{k_0(u_r+v_r)+\cdots+k_{r-1}(u_1+v_1)+k_r(u_0+v_0)} = -1.$$

Thus we obtain

$$\tau(k,k)=\frac{1}{3\cdot 2^{2(r+1)}}+\frac{1}{2^{3r+2}}\sum_{\substack{u=0\\ 2|u}}^{2^{r+1}-2}(-1)=\frac{1}{3\cdot 2^{2(r+1)}}-\frac{2^r}{2^{3r+2}}=-\frac{1}{6\cdot 2^{2r}}. \blacksquare$$

LEMMA 5. *Let $x_1, x_2 \in [0,1)$ and let $z_1, z_2 \in [0,1)$ be the points obtained after applying an i.i.d. random digital shift of depth m to x_1 and x_2.*

(a) *We have*

$$\mathbb{E}[z_1]=\frac{1}{2},\qquad \mathbb{E}[z_1^2]=\frac{1}{3}.$$

(b) *We have*

$$\mathbb{E}[|z_1-z_2|]=\sum_{k=0}^{2^m-1}\tau(k)\,\mathrm{wal}_k(x_1\oplus x_2),$$

where $\tau(0)=1/3$ and $\tau(k)=-1/(6\cdot 4^{r(k)})$ for $k>0$. For $k>0$, $r(k)$ denotes the unique integer r such that $2^r\le k<2^{r+1}$.

(c) *We have*

$$\mathbb{E}[\min(1-z_1,1-z_2)]=\frac{1}{2}\Big(1-\sum_{k=0}^{2^m-1}\tau(k)\,\mathrm{wal}_k(x_1\oplus x_2)\Big).$$

Proof. (a) The proof of these two formulae is straightforward.

(b) In Lemma 4 it was shown that

$$|z_1-z_2|=\sum_{k,l=0}^{\infty}\tau(k,l)\,\mathrm{wal}_k(z_1)\,\mathrm{wal}_l(z_2),$$

where $\tau(k)=\tau(k,k)=-1/(6\cdot 4^{r(k)})$ for $k>0$ and $\tau(0,0)=1/3$. (We do not need to know $\tau(k,l)$ for $k\ne l$ for our purposes here.) The result now follows from the linearity of the expectation value and Lemma 3.

(c) This result follows from (a) and (b) together with the formula

$$\min(z_1,z_2)=\frac{1}{2}(z_1+z_2-|z_1-z_2|). \blacksquare$$

We are now ready to prove Theorem 1.

Proof of Theorem 1. Let $\widetilde{P}_{2^m}=\{\boldsymbol{z}_0,\dots,\boldsymbol{z}_{2^m-1}\}$ and $\boldsymbol{z}_n=(z_{n,1},\dots,z_{n,s})$. From Proposition 1, Lemma 5 and the linearity of expectation we get

$$\mathbb{E}[\mathcal{L}^2_{2,2^m,\gamma}(\widetilde{P}_{2^m})]=\sum_{\substack{\mathfrak{u}\subseteq D\\ \mathfrak{u}\ne\emptyset}}\gamma_{\mathfrak{u}}\bigg[\frac{1}{3^{|\mathfrak{u}|}}-\frac{2}{2^m}\sum_{n=0}^{2^m-1}\prod_{j\in\mathfrak{u}}\frac{1-\mathbb{E}[z_{n,j}^2]}{2}$$
$$+\frac{1}{2^{2m}}\sum_{n,h=0}^{2^m-1}\prod_{j\in\mathfrak{u}}\mathbb{E}[\min(1-z_{n,j},1-z_{h,j})]\bigg]$$

$$= \sum_{\substack{\mathfrak{u}\subseteq D\\ \mathfrak{u}\neq\emptyset}} \gamma_{\mathfrak{u}} \Bigg[-\frac{1}{3^{|\mathfrak{u}|}} + \frac{1}{2^{2m}} \sum_{n=0}^{2^m-1} \prod_{j\in\mathfrak{u}} \mathbb{E}[1-z_{n,j}] + \frac{1}{2^{2m}} \sum_{\substack{n,h=0\\ n\neq h}}^{2^m-1} \prod_{j\in\mathfrak{u}} \mathbb{E}[\min(1-z_{n,j}, 1-z_{h,j})]\Bigg].$$

Now we use Lemma 5 again to obtain

$$\mathbb{E}[\mathcal{L}^2_{2,2^m,\gamma}(\widetilde{P}_{2^m})] = \sum_{\substack{\mathfrak{u}\subseteq D\\ \mathfrak{u}\neq\emptyset}} \gamma_{\mathfrak{u}} \Bigg[-\frac{1}{3^{|\mathfrak{u}|}} + \frac{1}{2^m}\frac{1}{2^{|\mathfrak{u}|}} + \frac{1}{2^{2m}} \sum_{\substack{n,h=0\\ n\neq h}}^{2^m-1} \prod_{j\in\mathfrak{u}} \frac{1}{2}\Big(1 - \sum_{k=0}^{2^m-1} \tau(k)\,\mathrm{wal}_k(x_{n,j}\oplus x_{h,j})\Big)\Bigg].$$

We have

$$\prod_{j\in\mathfrak{u}} \Big(1 - \sum_{k=0}^{2^m-1} \tau(k)\,\mathrm{wal}_k(x_{n,j}\oplus x_{h,j})\Big) = 1 + \sum_{\substack{\mathfrak{w}\subseteq\mathfrak{u}\\ \mathfrak{w}\neq\emptyset\\ \mathfrak{w}=\{w_1,\dots,w_d\}}} (-1)^{|\mathfrak{w}|}$$

$$\times \sum_{k_1=0}^{2^m-1} \cdots \sum_{k_d=0}^{2^m-1} \tau(k_1)\cdots\tau(k_d)\,\mathrm{wal}_{k_1,\dots,k_d}(x_{n,w_1}\oplus x_{h,w_1},\dots,x_{n,w_d}\oplus x_{h,w_d}).$$

Thus

$$\mathbb{E}[\mathcal{L}^2_{2,2^m,\gamma}(\widetilde{P}_{2^m})] = \sum_{\substack{\mathfrak{u}\subseteq D\\ \mathfrak{u}\neq\emptyset}} \gamma_{\mathfrak{u}} \Bigg[-\frac{1}{3^{|\mathfrak{u}|}} + \frac{1}{2^m}\frac{1}{2^{|\mathfrak{u}|}} + \frac{1}{2^{2m}} \sum_{\substack{n,h=0\\ n\neq h}}^{2^m-1} \frac{1}{2^{|\mathfrak{u}|}} + \frac{1}{2^{|\mathfrak{u}|}}\frac{1}{2^{2m}} \sum_{\substack{n,h=0\\ n\neq h}}^{2^m-1} \sum_{\substack{\mathfrak{w}\subseteq\mathfrak{u}\\ \mathfrak{w}\neq\emptyset\\ \mathfrak{w}=\{w_1,\dots,w_d\}}} (-1)^d \times \sum_{k_1=0}^{2^m-1} \cdots \sum_{k_d=0}^{2^m-1} \prod_{i=1}^{d} \tau(k_i)\,\mathrm{wal}_{k_i}(x_{n,w_i}\oplus x_{h,w_i})\Bigg].$$

We have

$$\sum_{k=0}^{2^m-1} \tau(k) = \frac{1}{3} - \sum_{r=0}^{m-1} \sum_{k=2^r}^{2^{r+1}-1} \frac{1}{6\cdot 4^r} = \frac{1}{3\cdot 2^m},$$

and therefore

$$\sum_{\substack{\mathfrak{w}\subseteq\mathfrak{u}\\ \mathfrak{w}\neq\emptyset}} (-1)^{|\mathfrak{w}|} \sum_{k_1,\ldots,k_{|\mathfrak{w}|}=0}^{2^m-1} \prod_{i=1}^{|\mathfrak{w}|} \tau(k_i) = \sum_{\substack{\mathfrak{w}\subseteq\mathfrak{u}\\ \mathfrak{w}\neq\emptyset}} \left(-\frac{1}{3\cdot 2^m}\right)^{|\mathfrak{w}|}$$
$$= \sum_{r=1}^{|\mathfrak{u}|} \binom{|\mathfrak{u}|}{r} \left(-\frac{1}{3\cdot 2^m}\right)^{r}$$
$$= \left(1-\frac{1}{3\cdot 2^m}\right)^{|\mathfrak{u}|} - 1.$$

By adding and subtracting this in the above expression we obtain

$$\mathbb{E}[\mathcal{L}^2_{2,2^m,\gamma}(\widetilde{P}_{2^m})] = \sum_{\substack{\mathfrak{u}\subseteq D\\ \mathfrak{u}\neq\emptyset}} \gamma_{\mathfrak{u}} \Bigg[\frac{1}{2^{|\mathfrak{u}|}} - \frac{1}{3^{|\mathfrak{u}|}} + \left(1-\left(1-\frac{1}{3\cdot 2^m}\right)^{|\mathfrak{u}|}\right)\frac{1}{2^m}\,\frac{1}{2^{|\mathfrak{u}|}}$$
$$+ \frac{1}{2^{|\mathfrak{u}|}}\,\frac{1}{2^{2m}} \sum_{n,h=0}^{2^m-1} \sum_{\substack{\mathfrak{w}\subseteq\mathfrak{u}\\ \mathfrak{w}\neq\emptyset\\ \mathfrak{w}=\{w_1,\ldots,w_d\}}} (-1)^d$$
$$\times \sum_{k_1,\ldots,k_d=0}^{2^m-1} \prod_{i=1}^{d} \tau(k_i)\,\mathrm{wal}_{k_i}(x_{n,w_i}\oplus x_{h,w_i})\Bigg].$$

Since $\tau(0)=1/3$ we have

$$\frac{1}{2^{2m}}\,\frac{1}{2^{|\mathfrak{u}|}} \sum_{n,h=0}^{2^m-1} \sum_{\substack{\mathfrak{w}\subseteq\mathfrak{u}\\ \mathfrak{w}\neq\emptyset}} (-1)^{|\mathfrak{w}|}\tau(0)^{|\mathfrak{w}|} = \frac{1}{3^{|\mathfrak{u}|}} - \frac{1}{2^{|\mathfrak{u}|}}.$$

Hence

$$\mathbb{E}[\mathcal{L}^2_{2,2^m,\gamma}(\widetilde{P}_{2^m})]$$
$$= \sum_{\substack{\mathfrak{u}\subseteq D\\ \mathfrak{u}\neq\emptyset}} \gamma_{\mathfrak{u}} \Bigg[\frac{1}{2^{|\mathfrak{u}|}} - \frac{1}{3^{|\mathfrak{u}|}} + \left(1-\left(1-\frac{1}{3\cdot 2^m}\right)^{|\mathfrak{u}|}\right)\frac{1}{2^m}\,\frac{1}{2^{|\mathfrak{u}|}} + \frac{1}{3^{|\mathfrak{u}|}} - \frac{1}{2^{|\mathfrak{u}|}}$$
$$+ \frac{1}{2^{|\mathfrak{u}|}}\,\frac{1}{2^{2m}} \sum_{\substack{\mathfrak{w}\subseteq\mathfrak{u}\\ \mathfrak{w}\neq\emptyset\\ \mathfrak{w}=\{w_1,\ldots,w_d\}}} (-1)^d$$
$$\times \sum_{\substack{k_1,\ldots,k_d=0\\ (k_1,\ldots,k_d)\neq(0,\ldots,0)}}^{2^m-1} \sum_{n,h=0}^{2^m-1} \prod_{i=1}^{d} \tau(k_i)\,\mathrm{wal}_{k_i}(x_{n,w_i}\oplus x_{h,w_i})\Bigg].$$

From the group structure of digital nets (see Lemma 1) and from Lemma 2 it follows that for any digital net $\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{2^m-1}\}$ generated by the $m \times m$ matrices $C_1, \ldots, C_s$, we have

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

Since the d-dimensional projection of a digital (t, m, s)-net is again a digital (t, m, d)-net (see Introduction) we get (with $\mathfrak{w} = \{w_1, \ldots, w_d\}$)

$$\begin{aligned}
&\sum_{\substack{k_1,\ldots,k_d=0 \\ (k_1,\ldots,k_d) \neq (0,\ldots,0)}}^{2^m-1} \sum_{n,h=0}^{2^m-1} \prod_{j=1}^{d} \tau(k_j)\, \mathrm{wal}_{k_j}(x_{n,w_j} \oplus x_{h,w_j}) \\
&\qquad = 2^{2m} \sum_{\substack{k_1,\ldots,k_d=0 \\ (k_1,\ldots,k_d) \neq (0,\ldots,0) \\ C_{w_1}^T \vec{k}_1 + \cdots + C_{w_d}^T \vec{k}_d = \vec{0}}}^{2^m-1} \prod_{i=1}^{d} \tau(k_i) \\
&\qquad = 2^{2m} \sum_{\substack{\mathfrak{v} \subseteq \mathfrak{w} \\ \mathfrak{v} \neq \emptyset \\ \mathfrak{v} = \{v_1,\ldots,v_e\}}} \frac{1}{3^{|\mathfrak{w}|-|\mathfrak{v}|}} \sum_{\substack{k_1,\ldots,k_e=1 \\ C_{v_1}^T \vec{k}_1 + \cdots + C_{v_e}^T \vec{k}_e = \vec{0}}}^{2^m-1} \prod_{j=1}^{e} \tau(k_j).
\end{aligned}$$

As $\prod_{j=1}^{e} \tau(k_j) = (-1)^e \prod_{j=1}^{e} \psi(k_j)$ we have

$$\sum_{\substack{k_1,\ldots,k_d=0 \\ (k_1,\ldots,k_d) \neq (0,\ldots,0)}}^{2^m-1} \sum_{n,h=0}^{2^m-1} \prod_{j=1}^{d} \tau(k_j)\, \mathrm{wal}_{k_j}(x_{n,w_j} \oplus x_{h,w_j}) = \frac{2^{2m}}{3^{|\mathfrak{w}|}} \sum_{\substack{\mathfrak{v} \subseteq \mathfrak{w} \\ \mathfrak{v} \neq \emptyset}} (-3)^{|\mathfrak{v}|} \mathcal{B}(\mathfrak{v}).$$

Thus we obtain

$$\begin{aligned}
\mathbb{E}[\mathcal{L}_{2,2^m,\gamma}^2(\widetilde{P}_{2^m})] = \sum_{\substack{\mathfrak{u} \subseteq D \\ \mathfrak{u} \neq \emptyset}} \gamma_{\mathfrak{u}} \Bigg[& \frac{1}{2^{m+|\mathfrak{u}|}} \left(1 - \left(1 - \frac{1}{3 \cdot 2^m}\right)^{|\mathfrak{u}|}\right) \\
& + \frac{1}{2^{|\mathfrak{u}|}} \sum_{\substack{\mathfrak{w} \subseteq \mathfrak{u} \\ \mathfrak{w} \neq \emptyset}} \left(-\frac{1}{3}\right)^{|\mathfrak{w}|} \sum_{\substack{\mathfrak{v} \subseteq \mathfrak{w} \\ \mathfrak{v} \neq \emptyset}} (-3)^{|\mathfrak{v}|} \mathcal{B}(\mathfrak{v}) \Bigg].
\end{aligned}$$

Let now $\mathfrak{u}, \mathfrak{v}$, with $\emptyset \neq \mathfrak{v} \subseteq \mathfrak{u} \subseteq D$, be fixed. Then $\mathfrak{v} \subseteq \mathfrak{w} \subseteq \mathfrak{u}$ is equivalent to $(\mathfrak{w} \setminus \mathfrak{v}) \subseteq (\mathfrak{u} \setminus \mathfrak{v})$, provided that $\mathfrak{v} \subseteq \mathfrak{w}$. Therefore, for $|\mathfrak{v}| \leq w \leq |\mathfrak{u}|$, there are $\binom{|\mathfrak{u}|-|\mathfrak{v}|}{w-|\mathfrak{v}|}$ sets $\mathfrak{w}$ such that $|\mathfrak{w}|=w$ and $\mathfrak{v}\subseteq\mathfrak{w}\subseteq\mathfrak{u}$. Hence

$$\begin{aligned}\sum_{\substack{\mathfrak{w}\subseteq\mathfrak{u}\\ \mathfrak{w}\neq\emptyset}}\left(-\frac{1}{3}\right)^{|\mathfrak{w}|}\sum_{\substack{\mathfrak{v}\subseteq\mathfrak{w}\\ \mathfrak{v}\neq\emptyset}}(-3)^{|\mathfrak{v}|}\mathcal{B}(\mathfrak{v}) &= \sum_{\substack{\mathfrak{v}\subseteq\mathfrak{u}\\ \mathfrak{v}\neq\emptyset}}\sum_{w=|\mathfrak{v}|}^{|\mathfrak{u}|}\binom{|\mathfrak{u}|-|\mathfrak{v}|}{w-|\mathfrak{v}|}\left(-\frac{1}{3}\right)^{w}(-3)^{|\mathfrak{v}|}\mathcal{B}(\mathfrak{v})\\ &= \sum_{\substack{\mathfrak{v}\subseteq\mathfrak{u}\\ \mathfrak{v}\neq\emptyset}}\sum_{w=0}^{|\mathfrak{u}|-|\mathfrak{v}|}\binom{|\mathfrak{u}|-|\mathfrak{v}|}{w}\left(-\frac{1}{3}\right)^{w}\mathcal{B}(\mathfrak{v})\\ &= \sum_{\substack{\mathfrak{v}\subseteq\mathfrak{u}\\ \mathfrak{v}\neq\emptyset}}\left(\frac{2}{3}\right)^{|\mathfrak{u}|-|\mathfrak{v}|}\mathcal{B}(\mathfrak{v}),\end{aligned}$$

and the result follows. ∎

3.2. *The mean square weighted $\mathcal{L}_2$ discrepancy of randomized digital $(0,m,s)$-nets over $\mathbb{Z}_2$ for $s=2,3$.* In this subsection we calculate the exact value of the mean square weighted $\mathcal{L}_2$ discrepancy of randomized digital $(0,m,s)$-nets for $s=2,3$. We have the following theorem.

THEOREM 2. *For $s=2,3$ let $P_{s,2^m}$ be a digital $(0,m,s)$-net over $\mathbb{Z}_2$. Let $\widetilde{P}_{s,2^m}$ be the point set obtained after applying an i.i.d. random digital shift of depth m independently to each coordinate of each point of $P_{s,2^m}$. Then the mean square weighted $\mathcal{L}_2$ discrepancy of $\widetilde{P}_{s,2^m}$ for $s=2$ is given by*

$$\mathbb{E}[\mathcal{L}^2_{2,2^m,\gamma}(\widetilde{P}_{2,2^m})]=\frac{\gamma_{\{1,2\}}}{24}\frac{m}{2^{2m}}+\frac{1}{2^{2m}}\left(\frac{\gamma_{\{1\}}}{6}+\frac{\gamma_{\{2\}}}{6}+\frac{5\gamma_{\{1,2\}}}{36}\right)$$

and for $s=3$ the mean square weighted $\mathcal{L}_2$ discrepancy is given by

$$\begin{aligned}\mathbb{E}[\mathcal{L}^2_{2,2^m,\gamma}(\widetilde{P}_{3,2^m})]=\gamma_{\{1,2,3\}}&\left(\frac{1}{192}\frac{m^2}{2^{2m}}+\frac{23}{576}\frac{m}{2^{2m}}+\frac{19}{216}\frac{1}{2^{2m}}\right)\\ &+(\gamma_{\{1,2\}}+\gamma_{\{1,3\}}+\gamma_{\{2,3\}})\left(\frac{1}{24}\frac{m}{2^{2m}}+\frac{5}{36}\frac{1}{2^{2m}}\right)\\ &+(\gamma_{\{1\}}+\gamma_{\{2\}}+\gamma_{\{3\}})\frac{1}{6}\frac{1}{2^{2m}}.\end{aligned}$$

Proof. Let C_1 and C_2 denote the generating matrices of the digital $(0,m,2)$-net over $\mathbb{Z}_2$. For $s=2$ we obtain from Theorem 1

$$\begin{aligned}&\mathbb{E}[\mathcal{L}^2_{2,2^m,\gamma}(\widetilde{P}_{2,2^m})] \qquad (4)\\ &\quad=\sum_{\substack{\mathfrak{u}\subseteq\{1,2\}\\ \mathfrak{u}\neq\emptyset}}\gamma_{\mathfrak{u}}\left[\frac{1}{2^{m+|\mathfrak{u}|}}\left(1-\left(1-\frac{1}{3\cdot 2^m}\right)^{|\mathfrak{u}|}\right)+\frac{1}{3^{|\mathfrak{u}|}}\sum_{\substack{\mathfrak{v}\subseteq\mathfrak{u}\\ \mathfrak{v}\neq\emptyset}}\left(\frac{3}{2}\right)^{|\mathfrak{v}|}\mathcal{B}(\mathfrak{v})\right],\end{aligned}$$

where for $\mathfrak{v} = \{v_1, \ldots, v_e\}$ we have

$$\mathcal{B}(\mathfrak{v}) = \sum_{\substack{k_1,\ldots,k_e=1 \\ C_{v_1}^T \vec{k}_1 + \cdots + C_{v_e}^T \vec{k}_e = \vec{0}}}^{2^m-1} \prod_{j=1}^{e} \psi(k_j),$$

with $\psi(k) = 1/(6 \cdot 4^{r(k)})$ and $r(k)$ such that $2^{r(k)} \le k < 2^{r(k)+1}$.

First we note that, as C_1 and C_2 generate a $(0, m, 2)$-net, they are regular. Therefore $\mathcal{B}(\mathfrak{v}) = 0$ for $|\mathfrak{v}| = 1$. Thus for $|\mathfrak{u}| = 1$ the addend in (4) is $\frac{1}{2^{2m}} \frac{\gamma_{\mathfrak{u}}}{6}$.

Now let $\mathfrak{u} = \{1, 2\}$. We have

$$\frac{\gamma_{\{1,2\}}}{2^{m+|\mathfrak{u}|}} \left(1 - \left(1 - \frac{1}{3 \cdot 2^m}\right)^{|\mathfrak{u}|}\right) = \frac{\gamma_{\{1,2\}}}{6 \cdot 2^{2m}} - \frac{\gamma_{\{1,2\}}}{36 \cdot 2^{3m}}.$$

In the following we calculate $\mathcal{B}(\{1, 2\})$. Since the generating matrices C_1 and C_2 of a digital $(0, m, 2)$-net over $\mathbb{Z}_2$ must be regular, and since multiplying C_1 and C_2 by a regular matrix A does not change the point set (only its order) we may assume in the following that C_1 is the $m \times m$ identity matrix. Hence

$$C_1^T \vec{k}_1 + C_2^T \vec{k}_2 = \vec{0} \quad \text{iff} \quad \vec{k}_1 = C_2^T \vec{k}_2 =: \vec{k}_1(k_2).$$

Now we use the definition of ψ to get

$$\begin{aligned} \sum_{\substack{k_1,k_2=1 \\ C_1^T \vec{k}_1 + C_2^T \vec{k}_2 = \vec{0}}}^{2^m-1} \psi(k_1)\psi(k_2) &= \frac{1}{36} \sum_{k_2=1}^{2^m-1} \frac{1}{4^{r(k_1(k_2))}} \frac{1}{4^{r(k_2)}} \\ &= \frac{1}{36} \sum_{u=0}^{m-1} \frac{1}{4^u} \sum_{k_2=2^u}^{2^{u+1}-1} \frac{1}{4^{r(k_1(k_2))}}. \end{aligned}$$

Consider the innermost sum in the above expression. We have

$$\Sigma(u) := \sum_{k_2=2^u}^{2^{u+1}-1} \frac{1}{4^{r(k_1(k_2))}} = \sum_{w=0}^{m-1} \frac{1}{4^w} \sum_{\substack{k_2=2^u \\ r(k_1(k_2))=w}}^{2^{u+1}-1} 1.$$

From [14, proof of Theorem 1] we find that

$$\sum_{\substack{k_2=2^u \\ r(k_1(k_2))=w}}^{2^{u+1}-1} 1 = \begin{cases} 0 & \text{if } u + w \le m - 2, \\ 1 & \text{if } u + w = m - 1, \\ 2^{u+w-m} & \text{if } u + w \ge m. \end{cases}$$

Thus

$$\Sigma(u) = \frac{1}{4^{m-1-u}} + \sum_{w=m-u}^{m-1} \frac{2^{u+w-m}}{4^w} = \frac{6 \cdot 4^u}{4^m} - \frac{2 \cdot 2^u}{4^m}$$

and therefore

$$\sum_{\substack{k_1,k_2=1\\ C_1^T\vec{k}_1+C_2^T\vec{k}_2=\vec{0}}}^{2^m-1}\psi(k_1)\psi(k_2)=\frac{1}{36}\sum_{u=0}^{m-1}\frac{1}{4^u}\left(\frac{6\cdot 4^u}{4^m}-\frac{2\cdot 2^u}{4^m}\right)$$

$$=\frac{m}{6\cdot 4^m}-\frac{1}{4^m}\frac{2}{36}2\left(1-\frac{1}{2^m}\right)$$

$$=\frac{m}{6\cdot 4^m}-\frac{1}{9\cdot 4^m}+\frac{4}{36\cdot 2^{3m}}.$$

Now we insert this result in equation (4) to get

$$\mathbb{E}[\mathcal{L}^2_{2,2^m,\gamma}(\widetilde{P}_{2,2^m})]=\frac{1}{2^{2m}}\left(\frac{\gamma_{\{1\}}}{6}+\frac{\gamma_{\{2\}}}{6}+\frac{\gamma_{\{1,2\}}}{6}\right)-\frac{\gamma_{\{1,2\}}}{36\cdot 2^{3m}}$$

$$+\frac{\gamma_{\{1,2\}}}{4}\left(\frac{m}{6\cdot 4^m}-\frac{1}{9\cdot 4^m}+\frac{4}{36\cdot 2^{3m}}\right)$$

$$=\frac{\gamma_{\{1,2\}}}{24}\frac{m}{2^{2m}}+\frac{1}{2^{2m}}\left(\frac{\gamma_{\{1\}}}{6}+\frac{\gamma_{\{2\}}}{6}+\frac{5\gamma_{\{1,2\}}}{36}\right),$$

which is the desired result for $s=2$.

We turn to the case where $s=3$. Let C_1,C_2 and C_3 denote the generating matrices of the digital $(0,m,3)$-net over $\mathbb{Z}_2$. As the quality parameter t is zero it is clear that C_1,C_2 and C_3 are regular. Hence $\mathcal{B}(\mathfrak{v})=0$ for $|\mathfrak{v}|=1$. Further, for $\mathfrak{v}\subseteq\{1,2,3\}$ with $|\mathfrak{v}|=2$ we obtain from the first part of the proof

$$\mathcal{B}(\mathfrak{v})=\frac{m}{6\cdot 4^m}-\frac{1}{9\cdot 4^m}+\frac{1}{9\cdot 2^{3m}}.$$

So it remains to calculate $\mathcal{B}(\{1,2,3\})$. As above we may assume in the following that C_1 is the $m\times m$ identity matrix. Hence

$$C_1^T\vec{k}_1+C_2^T\vec{k}_2+C_3^T\vec{k}_3=\vec{0}\quad\text{iff}\quad \vec{k}_1=C_2^T\vec{k}_2+C_3^T\vec{k}_3=:\vec{k}_1(k_2,k_3).$$

Now we get

$$\mathcal{B}(\{1,2,3\})=\sum_{\substack{k_1,k_2,k_3=1\\ C_1^T\vec{k}_1+C_2^T\vec{k}_2+C_3^T\vec{k}_3=\vec{0}}}^{2^m-1}\psi(k_1)\psi(k_2)\psi(k_3)$$

$$=\frac{1}{216}\sum_{\substack{k_2,k_3=1\\ k_1(k_2,k_3)\neq 0}}^{2^m-1}\frac{1}{4^{r(k_1(k_2,k_3))}}\frac{1}{4^{r(k_2)+r(k_3)}}$$

$$=\frac{1}{216}\sum_{u,v=0}^{m-1}\frac{1}{4^{u+v}}\underbrace{\sum_{k_2=2^u}^{2^{u+1}-1}\sum_{k_3=2^v}^{2^{v+1}-1}}_{k_1(k_2,k_3)\neq 0}\frac{1}{4^{r(k_1(k_2,k_3))}}.$$

The innermost double sum in the above expression equals

$$\Sigma(u,v) := \underbrace{\sum_{k_2=2^u}^{2^{u+1}-1} \sum_{k_3=2^v}^{2^{v+1}-1}}_{k_1(k_2,k_3)\neq 0} \frac{1}{4^{r(k_1(k_2,k_3))}} = \sum_{w=0}^{m-1} \frac{1}{4^w} \underbrace{\sum_{k_2=2^u}^{2^{u+1}-1} \sum_{k_3=2^v}^{2^{v+1}-1}}_{r(k_1(k_2,k_3))=w} 1.$$

From [23, proof of Theorem 1] we find that

$$\underbrace{\sum_{k_2=2^u}^{2^{u+1}-1} \sum_{k_3=2^v}^{2^{v+1}-1}}_{r(k_1(k_2,k_3))=w} 1 = \begin{cases} 0 & \text{if } u+v+w \leq m-3, \\ 1 & \text{if } u+v+w = m-2, \\ 0 & \text{if } u+v+w = m-1, \\ 2^{u+v+w-m} & \text{if } u+v+w \geq m. \end{cases}$$

Therefore we get

$$\mathcal{B}(\{1,2,3\}) = \frac{1}{216} \sum_{\substack{u,v,w=0 \\ u+v+w=m-2}}^{m-1} \frac{1}{4^{m-2}} + \frac{1}{216} \frac{1}{2^m} \sum_{\substack{u,v,w=0 \\ u+v+w\geq m}}^{m-1} \frac{1}{2^{u+v+w}}.$$

For the first sum we have

$$\sum_{\substack{u,v,w=0 \\ u+v+w=m-2}}^{m-1} \frac{1}{4^{m-2}} = \frac{1}{4^{m-2}} \binom{m}{2}.$$

The second sum can be written as

$$\sum_{\substack{u,v,w=0 \\ u+v+w\geq m}}^{m-1} \frac{1}{2^{u+v+w}} = \sum_{l=m}^{3m-3} \frac{1}{2^l} \sum_{\substack{u,v,w=0 \\ u+v+w=l}}^{m-1} 1.$$

Define

$$f(k) := \sum_{\substack{u,v=0 \\ u+v=k}}^{m-1} 1.$$

Then we have

$$f(k) = \begin{cases} k+1 & \text{if } 0 \leq k \leq m-1, \\ 2m-k-1 & \text{if } m \leq k \leq 2m-2, \\ 0 & \text{if } k \geq 2m-1. \end{cases}$$

Now we obtain

$$\sum_{\substack{u,v,w=0 \\ u+v+w=l}}^{m-1} 1 = \sum_{k=0}^{2m-2} f(k) \sum_{\substack{w=0 \\ k+w=l}}^{m-1} 1 = \sum_{\substack{k=0 \\ 0\leq l-k\leq m-1}}^{2m-2} f(k) = \sum_{k=\max(0,l-m+1)}^{\min(l,2m-2)} f(k).$$

Therefore

$$\sum_{l=m}^{3m-3} \frac{1}{2^l} \sum_{\substack{u,v,w=0\\u+v+w=l}}^{m-1} 1 = \sum_{l=m}^{3m-3} \frac{1}{2^l} \sum_{k=\max(0,l-m+1)}^{\min(l,2m-2)} f(k)$$

$$= \sum_{l=m}^{3m-3} \frac{1}{2^l} \sum_{k=l-m+1}^{\min(l,2m-2)} f(k)$$

$$= \sum_{l=m}^{2m-2} \frac{1}{2^l} \sum_{k=l-m+1}^{l} f(k) + \sum_{l=2m-1}^{3m-3} \frac{1}{2^l} \sum_{k=l-m+1}^{2m-2} f(k).$$

After some straightforward but tedious calculations we obtain the formula for $\mathcal{B}(\{1,2,3\})$. The result then follows by inserting the above results in the formula from Theorem 1. ∎

Note that the generating matrices $C_1, \ldots, C_s$ do not appear in our formula and therefore the mean square weighted $\mathcal{L}_2$ discrepancy is the same for any digital $(0,m,s)$-net over $\mathbb{Z}_2$, for $s = 2, 3$. This is also true for scrambled $(0,m,s)$-nets in a base b (see [9, 16]). Furthermore, the expected value of the $\mathcal{L}_2$ discrepancy of scrambled $(0,m,2)$-nets is the same as for $(0,m,2)$-nets which are randomized using a digital shift of depth m (see [16]).

In the following we consider the classical $\mathcal{L}_2$ discrepancy, that is, we choose $\gamma_D = 1$ and $\gamma_{\mathfrak{u}} = 0$ for $\mathfrak{u} \subset D$. We denote this choice of weights by $\boldsymbol{\gamma}_c$. Roth [27] proved that for any dimension $s \geq 2$ there exists a constant $c(s) > 0$ such that for any set P_N of N points in the s-dimensional unit cube we have

$$\int_{[0,1]^s} \Delta(\boldsymbol{x})^2 \, d\boldsymbol{x} \geq c(s) \, \frac{(\log N)^{s-1}}{N^2}. \tag{5}$$

Therefore we obtain

$$\mathcal{L}_{2,N,\boldsymbol{\gamma}_c}(P_N) \geq c(s)^{1/2} \, \frac{(\log N)^{(s-1)/2}}{N}. \tag{6}$$

(See Section 4 for more details.) Thus Theorem 2 shows that the mean square $\mathcal{L}_2$ discrepancy of randomized digital $(0,m,s)$-nets over $\mathbb{Z}_2$ achieves the best possible rate of convergence for $s = 2$ and 3.

In the following we compare the constants. For $N = 2^m$ the constant of [27] can be improved to (see also the subsequent inequality (17))

$$c(2)^{1/2} = \frac{3}{2^8\sqrt{\log 2}} = 0.01407\ldots, \qquad c(3)^{1/2} = \frac{3}{2^{10}\sqrt{2}\,\log 2} = 0.00298\ldots.$$

In the following, for $s = 2$ and 3 and each $m \in \mathbb{N}$ the set $P_{s,2^m,\boldsymbol{\sigma}_{m,s}}$ is a digital $(0,m,s)$-net over $\mathbb{Z}_2$ shifted by the digital shift $\boldsymbol{\sigma}_{m,s}$ of depth m. We obtain the following corollary.

COROLLARY 1. *For $s = 2, 3$ there exist sequences $(\boldsymbol{\sigma}_{m,s})_{m\geq 1}$ of digital shifts of depth m and digital $(0, m, s)$-nets $(P_{s,2^m})_{m\geq 1}$ over $\mathbb{Z}_2$ such that the sequences $(P_{s,2^m,\boldsymbol{\sigma}_{m,s}})_{m\geq 1}$ of shifted nets satisfy*

$$\limsup_{m\to\infty} \frac{2^m \mathcal{L}_{2,2^m,\gamma_c}(P_{2,2^m,\boldsymbol{\sigma}_{m,2}})}{\sqrt{\log 2^m}} \leq \frac{1}{\sqrt{24 \log 2}} = 0.24518\ldots$$

and

$$\limsup_{m\to\infty} \frac{2^m \mathcal{L}_{2,2^m,\gamma_c}(P_{3,2^m,\boldsymbol{\sigma}_{m,3}})}{\log 2^m} \leq \frac{1}{\sqrt{192} \log 2} = 0.10411\ldots.$$

We remark that it might be possible to improve the constant in Corollary 1 by finding the best shift for each digital $(0, m, s)$-net over $\mathbb{Z}_2$.

Note that one can also obtain the constants for the weighted $\mathcal{L}_2$ discrepancy: the constant of a weighted lower bound can be obtained from the definition of weighted $\mathcal{L}_2$ discrepancy (1) and (5), and the constant for the upper bound can be obtained from Theorem 2.

3.3. *An upper bound on the mean square weighted $\mathcal{L}_2$ discrepancy of randomized digital (t, m, s)-nets over $\mathbb{Z}_2$.* In this subsection we derive an upper bound on the formula of Theorem 1. We have

THEOREM 3. *Let P_{2^m} be a digital (t, m, s)-net over $\mathbb{Z}_2$ with $t < m$. Let $\widetilde{P}_{2^m}$ be the point set obtained after applying an i.i.d. random digital shift of depth m independently to each coordinate of each point of P_{2^m}. Then the mean square weighted $\mathcal{L}_2$ discrepancy of $\widetilde{P}_{2^m}$ is bounded as follows*:

$$\mathbb{E}[\mathcal{L}^2_{2,2^m,\gamma}(\widetilde{P}_{2^m})] \leq \frac{1}{2^{2(m-t)}} \sum_{\substack{\mathfrak{u}\subseteq D\\ \mathfrak{u}\neq\emptyset}} \gamma_{\mathfrak{u}} (m-t)^{|\mathfrak{u}|-1}.$$

As for the exact value of the mean square weighted $\mathcal{L}_2$ discrepancy for $(0, m, s)$-nets with $s = 2, 3$, the generating matrices $C_1, \ldots, C_s$ do not appear in the upper bound, which now depends on the quality parameter t only. For $t > 0$ the exact value of $\mathcal{B}(\mathfrak{v})$ (see Theorem 1) depends on the generating matrices and therefore we prove a bound.

For fixed $s \geq 1$ there is a $t \geq 0$ such that for every $m > t$ there is a digital (t, m, s)-net over $\mathbb{Z}_2$ (see for example [19]). Thus Theorem 3 shows that the convergence order of the mean square weighted $\mathcal{L}_2$ discrepancy is best possible by the lower bound by Roth [27] (see (6)).

We need two lemmas for the proof of the above theorem.

LEMMA 6. *For $b > 1$ and integers $k, t_0 > 0$, we have*

$$\sum_{t=t_0}^{\infty} \binom{t+k-1}{k-1} b^{-t} \leq b^{-t_0} \binom{t_0+k-1}{k-1} \left(1 - \frac{1}{b}\right)^{-k}.$$

Proof. For completeness we give a short proof of the lemma which is taken from Matoušek [16]. By the binomial theorem we have

$$\sum_{t=t_0}^{\infty} \binom{t-t_0+k-1}{k-1} b^{-t} = b^{-t_0}\left(1-\frac{1}{b}\right)^{-k}.$$

Now use the inequality

$$\binom{t+k-1}{k-1} \Big/ \binom{t-t_0+k-1}{k-1} = \frac{(t+k-1)(t+k-2)\cdots(t-t_0+k)}{t(t-1)\cdots(t-t_0+1)} \le \binom{t_0+k-1}{k-1}. \ \blacksquare$$

LEMMA 7. *Let $C_1,\ldots,C_s$ be the generating matrices of a digital (t,m,s)-net over $\mathbb{Z}_2$. Further define $\mathcal{B}$ as in Theorem 1. Then for any $\mathfrak{v} \subseteq D$ we have*

$$\mathcal{B}(\mathfrak{v}) \le \frac{2^{2t}}{2^{2m}}\left(\frac{8}{9}\right)^{|\mathfrak{v}|}\left(m-t+\frac{1}{8}\right)^{|\mathfrak{v}|-1}.$$

Proof. To simplify the notation we prove the result only for $\mathfrak{v}=\{1,\ldots,s\}$. The other cases follow by the same arguments. We have

$$\mathcal{B}(\{1,\ldots,s\}) = \frac{1}{6^s}\sum_{v_1,\ldots,v_s=0}^{m-1}\frac{1}{4^{v_1+\cdots+v_s}}\underbrace{\sum_{k_1=2^{v_1}}^{2^{v_1+1}-1}\cdots\sum_{k_s=2^{v_s}}^{2^{v_s+1}-1}}_{C_1^T\vec{k}_1+\cdots+C_s^T\vec{k}_s=\vec{0}} 1.$$

Now we write

$$\Sigma(v_1,\ldots,v_s) := \underbrace{\sum_{k_1=2^{v_1}}^{2^{v_1+1}-1}\cdots\sum_{k_s=2^{v_s}}^{2^{v_s+1}-1}}_{C_1^T\vec{k}_1+\cdots+C_s^T\vec{k}_s=\vec{0}} 1. \tag{7}$$

For $1 \le j \le s$ and $1 \le i \le m$ let $\vec{c}_{j,i}^{\,T}$ denote the ith row vector of the matrix C_j.

For $2^{v_j} \le k_j \le 2^{v_j+1}-1$, the binary digit expansion of k_j is of the form

$$k_j = k_{j,0} + k_{j,1}2 + \cdots + k_{j,v_j-1}2^{v_j-1} + 2^{v_j}.$$

Hence the condition in our sum (7) can be written as

$$\begin{aligned} &\vec{c}_{1,1}k_{1,0} + \cdots + \vec{c}_{1,v_1}k_{1,v_1-1} + \vec{c}_{1,v_1+1} \\ &+ \vec{c}_{2,1}k_{2,0} + \cdots + \vec{c}_{2,v_2}k_{2,v_2-1} + \vec{c}_{2,v_2+1} \\ &\ \ \vdots \\ &+ \vec{c}_{s,1}k_{s,0} + \cdots + \vec{c}_{s,v_s}k_{s,v_s-1} + \vec{c}_{s,v_s+1} = \vec{0}. \end{aligned} \tag{8}$$

Since by the digital (t,m,s)-net property (see Definition 2) the vectors

$$\vec{c}_{1,1},\ldots,\vec{c}_{1,v_1+1},\ldots,\vec{c}_{s,1},\ldots,\vec{c}_{s,v_s+1}$$

are linearly independent as long as $(v_1+1)+\cdots+(v_s+1) \le m-t$, we must have

$$v_1+\cdots+v_s \ge m-t-s+1. \tag{9}$$

Let now A denote the $m \times (v_1+\cdots+v_s)$ matrix with column vectors $\vec{c}_{1,1},\ldots,\vec{c}_{1,v_1},\ldots,\vec{c}_{s,1},\ldots,\vec{c}_{s,v_s}$, i.e.,

$$A := (\vec{c}_{1,1},\ldots,\vec{c}_{1,v_1},\ldots,\vec{c}_{s,1},\ldots,\vec{c}_{s,v_s}).$$

Further let

$$\begin{aligned}\vec{f} &:= \vec{c}_{1,v_1+1} \oplus \cdots \oplus \vec{c}_{s,v_s+1} \in \mathbb{Z}_2^m,\\ \vec{k} &:= (k_{1,0},\ldots,k_{1,v_1-1},\ldots,k_{s,0},\ldots,k_{s,v_s-1})^T \in \mathbb{Z}_2^{v_1+\cdots+v_s}.\end{aligned}$$

Then the linear system (8) can be written as

$$A\vec{k} = \vec{f} \tag{10}$$

and hence

$$\Sigma(v_1,\ldots,v_s) = \sum_{\substack{\vec{k}\in\mathbb{Z}_2^{v_1+\cdots+v_s}\\ A\vec{k}=\vec{f}}} 1 = \#\{\vec{k}\in\mathbb{Z}_2^{v_1+\cdots+v_s} : A\vec{k}=\vec{f}\}.$$

By the definition of the matrix A and since $C_1,\ldots,C_s$ are the generating matrices of a digital (t,m,s)-net over $\mathbb{Z}_2$ we have

$$\begin{aligned}&\operatorname{rank}(A) = v_1+\cdots+v_s \quad \text{if } v_1+\cdots+v_s \le m-t, \quad \text{and}\\ &\operatorname{rank}(A) \ge m-t \quad \text{else.}\end{aligned}$$

Let L denote the linear space of solutions of the homogeneous system $A\vec{k}=\vec{0}$ and let $\dim(L)$ denote the dimension of L. Then

$$\begin{aligned}&\dim(L) = 0 \quad \text{if } v_1+\cdots+v_s \le m-t, \quad \text{and}\\ &\dim(L) \le v_1+\cdots+v_s-m+t \quad \text{else.}\end{aligned}$$

Hence if $v_1+\cdots+v_s \le m-t$ we find that the system (10) has at most one solution and if $v_1+\cdots+v_s > m-t$ the system (10) has at most $2^{v_1+\cdots+v_s-m+t}$ solutions, i.e.,

$$\Sigma(v_1,\ldots,v_s) \le \begin{cases} 1 & \text{if } v_1+\cdots+v_s \le m-t,\\ 2^{v_1+\cdots+v_s-m+t} & \text{if } v_1+\cdots+v_s > m-t.\end{cases}$$

Together with condition (9) we obtain

$$\begin{aligned}\mathcal{B}(\{1,\ldots,s\}) \le \frac{1}{6^s} &\sum_{\substack{v_1,\ldots,v_s=0\\ m-t-s+1\le v_1+\cdots+v_s\le m-t}}^{m-1} \frac{1}{4^{v_1+\cdots+v_s}}\\ &+ \frac{1}{6^s}\sum_{\substack{v_1,\ldots,v_s=0\\ v_1+\cdots+v_s>m-t}}^{m-1} \frac{1}{4^{v_1+\cdots+v_s}}\,2^{v_1+\cdots+v_s-m+t}\end{aligned} \tag{11}$$

$$=: \Sigma_1 + \Sigma_2.$$

Now we have to estimate the sums Σ_1 and Σ_2. First we have

$$\Sigma_2 = \frac{1}{6^s}\,\frac{2^t}{2^m} \sum_{l=m-t+1}^{s(m-1)} \frac{1}{2^l} \sum_{\substack{v_1,\ldots,v_s=0\\ v_1+\cdots+v_s=l}}^{m-1} 1 \le \frac{1}{6^s}\,\frac{2^t}{2^m} \sum_{l=m-t+1}^{\infty} \binom{l+s-1}{s-1}\frac{1}{2^l},$$

where we used the fact that for fixed l the number of non-negative integer solutions of $v_1 + \cdots + v_s = l$ is $\binom{l+s-1}{s-1}$. Now we apply Lemma 6 to obtain

$$\Sigma_2 \le \frac{1}{6^s}\,\frac{2^t}{2^m}\,\frac{1}{2^{m-t+1}}\binom{m-t+s}{s-1}2^s = \frac{1}{3^s}\,\frac{4^t}{4^m}\,\frac{1}{2}\binom{m-t+s}{s-1}. \tag{12}$$

Finally, since

$$\binom{m-t+s}{s-1} = \frac{(m-t+2)(m-t+3)\cdots(m-t+s)}{1\cdot 2\cdots(s-1)} \le (m-t+2)^{s-1},$$

we obtain

$$\Sigma_2 \le \frac{1}{3^s}\,\frac{4^t}{4^m}\,\frac{1}{2}(m-t+2)^{s-1}.$$

Now we estimate Σ_1. If $m-t \ge s-1$ we proceed similarly to the above to obtain

$$\begin{aligned}\Sigma_1 &= \frac{1}{6^s} \sum_{l=m-t-s+1}^{m-t} \binom{l+s-1}{s-1}\frac{1}{4^l} \le \frac{1}{6^s}\,\frac{1}{4^{m-t-s+1}}\binom{m-t}{s-1}\left(\frac{3}{4}\right)^{-s} \\ &\le \frac{8^s}{9^s}\,\frac{4^t}{4^m}\,\frac{1}{4}\,\frac{(m-t)^{s-1}}{(s-1)!}.\end{aligned} \tag{13}$$

For this case we obtain

$$\begin{aligned}\mathcal{B}(\{1,\ldots,s\}) &\le \frac{8^s}{9^s}\,\frac{4^t}{4^m}\,\frac{1}{4}\,\frac{(m-t)^{s-1}}{(s-1)!} + \frac{1}{3^s}\,\frac{4^t}{4^m}\,\frac{1}{2}\,(m-t+2)^{s-1} \\ &\le \frac{8^s}{9^s}\,\frac{4^t}{4^m}\left(\frac{1}{4}\,\frac{(m-t)^{s-1}}{(s-1)!} + \frac{3}{8}\,\frac{1}{2}\left(\frac{3}{8}(m-t) + \frac{6}{8}\right)^{s-1}\right).\end{aligned}$$

As $\frac{3}{8}(m-t) + \frac{6}{8} \le m-t+\frac{1}{8}$ provided that $m-t>0$ we have

$$\mathcal{B}(\{1,\ldots,s\}) \le \frac{4^t}{4^m}\left(\frac{8}{9}\right)^s\left(m-t+\frac{1}{8}\right)^{s-1},$$

which is the desired bound.

Now we consider the case where $m-t < s-1$. We have

$$\begin{aligned}\Sigma_1 &= \frac{1}{6^s}\sum_{l=0}^{m-t}\binom{l+s-1}{s-1}\frac{1}{4^l} \le \frac{1}{6^s}\sum_{l=0}^{\infty}\binom{l+s-1}{s-1}\frac{1}{4^l} \\ &= \left(\frac{2}{9}\right)^s \le \frac{1}{16}\,\frac{8^s}{9^s}\,\frac{4^t}{4^m}.\end{aligned} \tag{14}$$

Thus we obtain

$$\mathcal{B}(\{1,\dots,s\}) \le \frac{1}{16}\frac{8^s}{9^s}\frac{4^t}{4^m} + \frac{1}{3^s}\frac{4^t}{4^m}\frac{1}{2}(m-t+2)^{s-1}$$
$$\le \frac{8^s}{9^s}\frac{4^t}{4^m}\left(\frac{1}{16} + \frac{3}{8}\frac{1}{2}\left(\frac{3}{8}(m-t) + \frac{6}{8}\right)^{s-1}\right).$$

The result now follows by using the same arguments as above. ∎

We are now ready to prove Theorem 3.

Proof of Theorem 3. We use the formula of Theorem 1 together with Lemma 7 to obtain

$$\mathbb{E}[\mathcal{L}^2_{2,2^m,\gamma}(\widetilde{P}_{2^m})] \le \sum_{\substack{\mathfrak{u}\subseteq D\\ \mathfrak{u}\neq\emptyset}} \gamma_{\mathfrak{u}} \left[\frac{1}{2^{m+|\mathfrak{u}|}}\left(1-\left(1-\frac{1}{3\cdot 2^m}\right)^{|\mathfrak{u}|}\right) + \frac{1}{3^{|\mathfrak{u}|}}\frac{2^{2t}}{2^{2m}}\sum_{\substack{\mathfrak{v}\subseteq\mathfrak{u}\\ \mathfrak{v}\neq\emptyset}}\left(\frac{4}{3}\right)^{|\mathfrak{v}|}\left(m-t+\frac{1}{8}\right)^{|\mathfrak{v}|-1}\right].$$

We have

$$\frac{1}{3^{|\mathfrak{u}|}}\sum_{\substack{\mathfrak{v}\subseteq\mathfrak{u}\\ \mathfrak{v}\neq\emptyset}}\left(\frac{4}{3}\right)^{|\mathfrak{v}|}\left(m-t+\frac{1}{8}\right)^{|\mathfrak{v}|-1} \le (m-t)^{-1}\left(\frac{1}{3}+\frac{4}{9}\left(m-t+\frac{1}{8}\right)\right)^{|\mathfrak{u}|}$$
$$\le \left(\frac{5}{6}\right)^{|\mathfrak{u}|}(m-t)^{|\mathfrak{u}|-1},$$

provided that $m-t>0$. Since for $x<y$ we have $y^s-x^s = s\zeta^{s-1}(y-x)$ for a $x<\zeta<y$, we have

$$1-\left(1-\frac{1}{3\cdot 2^m}\right)^{|\mathfrak{u}|} \le \frac{|\mathfrak{u}|}{3\cdot 2^m}.$$

As $|\mathfrak{u}|/2^{|\mathfrak{u}|}\le 1/2$ for $|\mathfrak{u}|\ge 1$, we obtain

$$\frac{1}{2^{m+|\mathfrak{u}|}}\left(1-\left(1-\frac{1}{3\cdot 2^m}\right)^{|\mathfrak{u}|}\right) + \frac{1}{3^{|\mathfrak{u}|}}\frac{2^{2t}}{2^{2m}}\sum_{\substack{\mathfrak{v}\subseteq\mathfrak{u}\\ \mathfrak{v}\neq\emptyset}}\left(\frac{4}{3}\right)^{|\mathfrak{v}|}\left(m-t+\frac{1}{8}\right)^{|\mathfrak{v}|-1} \tag{15}$$
$$\le \frac{1}{2^{2m}}\left[\frac{1}{6}+2^{2t}\left(\frac{5}{6}\right)^{|\mathfrak{u}|}(m-t)^{|\mathfrak{u}|-1}\right] \le \frac{1}{2^{2(m-t)}}(m-t)^{|\mathfrak{u}|-1},$$

and the result follows. ∎

In the following corollary we refine the bound of Theorem 3 by including the t-values of the lower dimensional projections. Observe that it follows easily from Definition 2 that any projection of a digital (t,m,s)-net on the coordinates of $\emptyset\neq\mathfrak{u}\subseteq D$ is again a digital $(t_{\mathfrak{u}},m,|\mathfrak{u}|)$-net with some $t_{\mathfrak{u}}\le t$.

In the following we write "digital $((t_{\mathfrak{u}}), m, s)$-net" for a digital (t, m, s)-net where the projections on $\emptyset \neq \mathfrak{u} \subseteq D$ have quality parameter $t_{\mathfrak{u}}$. The subsequent corollary can be obtained by using (15).

COROLLARY 2. *Let P_{2^m} be a digital $((t_{\mathfrak{u}}), m, s)$-net over $\mathbb{Z}_2$ with*

$$\max_{\emptyset \neq \mathfrak{u} \subseteq D} t_{\mathfrak{u}} < m.$$

Let $\widetilde{P}_{2^m}$ be the point set obtained after applying an i.i.d. random digital shift of depth m independently to each coordinate of each point of P_{2^m}. Then the mean square weighted $\mathcal{L}_2$ discrepancy of $\widetilde{P}_{2^m}$ is bounded as follows:

$$\mathbb{E}[\mathcal{L}^2_{2,2^m,\gamma}(\widetilde{P}_{2^m})] \leq \frac{1}{2^{2m}} \sum_{\substack{\mathfrak{u} \subseteq D \\ \mathfrak{u} \neq \emptyset}} \gamma_{\mathfrak{u}} 2^{2t_{\mathfrak{u}}} (m - t_{\mathfrak{u}})^{|\mathfrak{u}|-1}.$$

4. Asymptotics. In this section we investigate the asymptotic behaviour of the $\mathcal{L}_2$ discrepancy. We consider the classical $\mathcal{L}_2$ discrepancy, that is, $\gamma_D = 1$ and $\gamma_{\mathfrak{u}} = 0$ for $\mathfrak{u} \subset D$. As before we denote these weights with $\boldsymbol{\gamma}_c$. (We remark that the results in this section, except Subsection 4.2, can be generalized to arbitrary weights.)

Let $\log_2$ denote the logarithm to base 2. By an extension of the result of Roth [27] to dimension s we find that for any set P_N of N points in the s-dimensional unit cube,

$$\mathcal{L}_{2,N,\gamma_c}(P_N) \geq \frac{1}{N} \sqrt{\binom{\lfloor \log_2 N \rfloor + s + 1}{s-1} \frac{1}{2^{2s+4}}}.$$

For sets of $N = 2^m$ points the result can be slightly improved:

$$\mathcal{L}_{2,2^m,\gamma_c}(P_{2^m}) \geq \frac{1}{2^m} \sqrt{\binom{m + s + 1}{s-1} \frac{3}{2^{2s+4}}}. \tag{16}$$

From

$$\binom{m+s+1}{s-1} \geq \frac{m^{s-1}}{(s-1)!}$$

and $m = \log N / \log 2$ it follows that

$$\mathcal{L}_{2,2^m,\gamma_c}(P_{2^m}) \geq \frac{(\log N)^{(s-1)/2}}{N} \frac{3}{2^{2s+4} (\log 2)^{(s-1)/2} \sqrt{(s-1)!}}. \tag{17}$$

In the following subsection we consider the asymptotic behaviour of the classical $\mathcal{L}_2$ discrepancy of certain shifted (t, m, s)-nets. In Subsection 4.2 we consider shifted Niederreiter–Xing nets and show that Roth's lower bound is essentially best possible in N and s.

4.1. *Asymptotics of the classical $\mathcal{L}_2$ discrepancy of shifted digital (t,m,s)-nets over $\mathbb{Z}_2$.* In the following, for $s \in \mathbb{N}$ and $m \in \mathbb{N}$ the set $P_{t,s,2^m,\boldsymbol{\sigma}_{m,s}}$ is a digital (t,m,s)-net over $\mathbb{Z}_2$ shifted by the digital shift $\boldsymbol{\sigma}_{m,s}$ of depth m. We obtain

Theorem 4. *Let $s > 3$, $0 \le t < m$ and $m - t \ge s$ be such that a digital (t,m,s)-net over $\mathbb{Z}_2$ exists. Then there exists a digital shift $\boldsymbol{\sigma}_{m,s}$ of depth m such that for the shifted net $P_{t,s,2^m,\boldsymbol{\sigma}_{m,s}}$ we have*

$$\mathcal{L}_{2,2^m,\gamma_c}(P_{t,s,2^m,\boldsymbol{\sigma}_{m,s}}) \le \frac{2^t}{2^m}\sqrt{\binom{m-t+s}{s-1}\left(\frac{2}{3}\right)^s} + \mathcal{O}\left(\frac{m^{(s-2)/2}}{2^m}\right).$$

Proof. From Theorem 1 we obtain

$$\begin{aligned}(18)\quad & \mathbb{E}[\mathcal{L}^2_{2,2^m,\gamma_c}(\widetilde{P}_{2^m})] \\ &= \frac{1}{2^{m+s}}\left(1 - \left(1 - \frac{1}{3\cdot 2^m}\right)^s\right) + \frac{1}{3^s}\sum_{\substack{\mathfrak{v}\subseteq D\\ \mathfrak{v}\ne\emptyset}}\left(\frac{3}{2}\right)^{|\mathfrak{v}|}\mathcal{B}(\mathfrak{v}).\end{aligned}$$

Lemma 7 shows that, in order to find the constant of the leading term, we only need to consider $\mathcal{B}(\{1,\dots,s\})$. From (11)–(13) we obtain

$$\mathcal{B}(\{1,\dots,s\}) \le \frac{2^{2t}}{2^{2m}}\left(\frac{1}{2}\,\frac{1}{3^s}\binom{m-t+s}{s-1} + \frac{1}{4}\,\frac{8^s}{9^s}\binom{m-t}{s-1}\right).$$

As the bound in Theorem 1 was obtained by averaging over all shifts it follows that there exists a shift which yields an $\mathcal{L}_2$ discrepancy smaller than or equal to this bound. The result follows. ■

Observe that for large m, apart from the t, the bound in Theorem 4 is similar to (16). We now consider (t,s)-sequences. A (t,s)-sequence in base 2 is a sequence of points $(\boldsymbol{x}_n)_{n\ge 0}$ such that for all $m > t$ and $l \ge 0$ the set $\{\boldsymbol{x}_n : l2^m \le n < (l+1)2^m\}$ is a (t,m,s)-net in base 2. A digital (t,s)-sequence over $\mathbb{Z}_2$ is obtained by using $\infty \times \infty$ generating matrices $C_1,\dots,C_s$ over $\mathbb{Z}_2$.

From [21] it follows that for every dimension s there exists a digital (t,s)-sequence over $\mathbb{Z}_2$ such that $t \le 5s$. Thus for all $s \ge 1$ and $m > 5s$ there is a digital $(5s,m,s)$-net over $\mathbb{Z}_2$. (Note that if there is a digital (t,m,s)-net then also a digital $(t+1,m,s)$-net exists.) Let $P_{s,5s,2^m,\boldsymbol{\sigma}_{m,s}}$ denote a digital $(5s,m,s)$-net over $\mathbb{Z}_2$ shifted by the digital shift $\boldsymbol{\sigma}_{m,s}$ of depth m. We are interested in the asymptotic behaviour of the $\mathcal{L}_2$ discrepancy. For m much larger than s and $t = 5s$, we have

$$\binom{m-t+s}{s-1} \le \binom{m}{s-1} \le \frac{m^{s-1}}{(s-1)!}.$$

Further, if $N = 2^m$, then $m = \log N/\log 2$. The following corollary now follows from Theorem 4.

COROLLARY 3. *For every* $s \geq 1$ *and* $m \geq 5s$ *there exists a shifted digital* $(5s, m, s)$*-net* $P_{5s,s,2^m,\boldsymbol{\sigma}_{m,s}}$ *over* $\mathbb{Z}_2$ *shifted by the digital shift* $\boldsymbol{\sigma}_{m,s}$ *of depth* m *such that*

$$\mathcal{L}_{2,2^m,\gamma_c}(P_{5s,s,2^m,\boldsymbol{\sigma}_{m,s}}) \leq \frac{(\log N)^{(s-1)/2}}{N} \frac{22^s}{(\log 2)^{(s-1)/2}\sqrt{(s-1)!}} + \mathcal{O}\left(\frac{(\log N)^{(s-2)/2}}{N}\right),$$

where $N = 2^m$.

We note that the convergence of $\mathcal{O}((\log N)^{(s-1)/2}N^{-1})$ is best possible (see (17)).

In the remaining part of this subsection we discuss the constant depending on s. Note that

$$C(s) := \frac{22^s}{(\log 2)^{(s-1)/2}\sqrt{(s-1)!}}$$

tends to zero faster than exponentially. The best constant of the leading term of an upper bound known to the authors was derived by Hickernell [9]. He showed that for scrambled $(0, m, s)$-nets in base $b \geq s - 1$, where b is a prime power, the constant of the leading term is

$$A(s) = \frac{(s - s^{-1})^{(s-1)/2}}{6^{s/2}\sqrt{(s-1)!}(\log s)^{(s-1)/2}} \approx \left(\frac{e^s}{\sqrt{6\pi s}\, 6^s(\log s)^{s-1}}\right)^{1/2} \quad \text{as } s \to \infty.$$

The right hand side is obtained by Stirling's formula. It can easily be checked that $C(s)$ tends to zero much faster than $A(s)$. Thus our result improves Hickernell's considerably.

Compared to (17) our constant $C(s)$ is not quite as good. It is known that for digital (t, s)-sequences over $\mathbb{Z}_2$ we always have

$$t > s \log_2 \frac{3}{2} - 4\log_2(s-2) - 23 \quad \text{for all } s \geq 3,$$

by a result of Schmid [28]. Hence for digital (t, m, s)-nets obtained from digital (t, s)-sequences Theorem 4 cannot yield a constant of the form $a^{-s/2}((s-1)!)^{-1/2}$ for some $a > 1$. On the other hand, for special choices of m and s the t-value of a digital (t, m, s)-net may be considerably lower than the t-value of the best (t, s)-sequence. This will be investigated in the next subsection.

4.2. *On the* $\mathcal{L}_2$ *discrepancy of shifted Niederreiter–Xing nets.* In this subsection we derive an upper bound on the classical $\mathcal{L}_2$ discrepancy of shifted Niederreiter–Xing nets (see [20]; see also [31] for a recent survey article). This enables us to show that (16) is essentially best possible.

Niederreiter and Xing [20, Corollary 3] showed that for every integer $d \geq 2$ there exists a sequence of digital $(t_k, t_k + d, s_k)$-nets over $\mathbb{Z}_2$ with $s_k \to \infty$ as $k \to \infty$ such that

$$\lim_{k\to\infty} \frac{t_k}{\log_2 s_k} = \left\lfloor \frac{d}{2} \right\rfloor, \tag{19}$$

and that this is best possible. (We remark that the sequence of digital nets from the result of Niederreiter and Xing can be constructed explicitly.) Therefore for any $d \geq 1$ there exists a sequence of digital $(t_k, t_k + d, s_k)$-nets over $\mathbb{Z}_2$ and a k_d such that

$$\left\lceil \frac{t_k}{\log_2 s_k} \right\rceil = d \quad \text{and} \quad s_k \geq 2d+2 \qquad \text{for all } k \geq k_d. \tag{20}$$

(Note that if a digital $(t_k, t_k + d, s_k)$-net exists, then there also exists a digital $(t_k + 1, t_k + d + 1, s_k)$-net. Further, for $d = 1$ there exists a digital $(t, t+1, s)$-net for all $t, s \geq 1$.) For a set P of 2^m points in $[0,1)^s$ let

$$\mathcal{D}_{m,s}(P) := \frac{2^m \mathcal{L}_{2,2^m,\gamma_c}(P)}{\sqrt{\binom{m+s+1}{s-1}}}. \tag{21}$$

The bound in Theorem 3 was obtained by averaging over all shifts. Hence for any digital (t, m, s)-net there is always a shift $\boldsymbol{\sigma}^*$ which yields an $\mathcal{L}_2$ discrepancy smaller than or equal to this bound. Let $P_k(d)$ denote a shifted digital $(t_k, t_k + d, s_k)$-net over $\mathbb{Z}_2$ satisfying (20), which is shifted by such a shift $\boldsymbol{\sigma}^*$. We prove an upper bound on $\mathcal{D}_{t_k+d,s_k}(P_k(d))$ for fixed d.

In the following let $k \geq k_d$. Let $\mathfrak{v} \subseteq \{1, \ldots, s_k\}$ and $l := |\mathfrak{v}|$. First we consider the case where $l \geq d+2$. Note that $m - t = d$ for the nets considered here. Then (12) and (14) yield

$$\mathcal{B}(\mathfrak{v}) \leq \frac{1}{3^l} \frac{1}{4^d} \frac{1}{2} \binom{d+l}{d+1} + \frac{2^l}{9^l}.$$

For $0 < l \leq d+1$ we deduce from (12) and (13) that

$$\mathcal{B}(\mathfrak{v}) \leq \frac{1}{3^l} \frac{1}{4^d} \frac{1}{2} \binom{d+l}{d+1} + \frac{8^l}{9^l} \frac{1}{4^d} \frac{1}{4} \binom{d}{l-1}.$$

Therefore

$$\begin{aligned} \frac{1}{3^{s_k}} \sum_{\substack{\mathfrak{v}\subseteq D\\ \mathfrak{v}\neq\emptyset}} \left(\frac{3}{2}\right)^{|\mathfrak{v}|} \mathcal{B}(\mathfrak{v}) \leq & \frac{1}{3^{s_k}} \sum_{l=1}^{d+1} \left(\frac{3}{2}\right)^l \binom{s_k}{l} \left(\frac{1}{3^l} \frac{1}{4^d} \frac{1}{2} \binom{d+l}{d+1} + \frac{8^l}{9^l} \frac{1}{4^d} \frac{1}{4} \binom{d}{l-1} \right) \\ & + \frac{1}{3^{s_k}} \sum_{l=d+2}^{s_k} \left(\frac{3}{2}\right)^l \binom{s_k}{l} \left(\frac{1}{3^l} \frac{1}{4^d} \frac{1}{2} \binom{d+l}{d+1} + \frac{2^l}{9^l} \right). \end{aligned}$$

Now we have

$$\frac{1}{3^{s_k}}\sum_{l=1}^{s_k}\left(\frac{3}{2}\right)^l\binom{s_k}{l}\frac{1}{3^l}\frac{1}{4^d}\frac{1}{2}\binom{d+l}{d+1} \le \frac{1}{2}\frac{1}{3^{s_k}}\frac{1}{4^d}\binom{d+s_k}{d+1}\sum_{l=0}^{s_k}\frac{1}{2^l}\binom{s_k}{l}$$
$$= \frac{1}{2}\frac{1}{2^{s_k}}\frac{1}{4^d}\binom{d+s_k}{s_k-1}$$

and

$$\frac{1}{3^{s_k}}\sum_{l=1}^{d+1}\left(\frac{3}{2}\right)^l\binom{s_k}{l}\frac{8^l}{9^l}\frac{1}{4^d}\frac{1}{4}\binom{d}{l-1}+\frac{1}{3^{s_k}}\sum_{l=d+2}^{s_k}\left(\frac{3}{2}\right)^l\binom{s_k}{l}\frac{2^l}{9^l}$$
$$= \frac{1}{3^{s_k}}\frac{1}{4^{d+1}}\sum_{l=1}^{d+1}\binom{s_k}{l}\left(\frac{4}{3}\right)^l\binom{d}{l-1}+\frac{1}{3^{s_k}}\sum_{l=d+2}^{s_k}\binom{s_k}{l}\frac{1}{3^l}$$
$$\le \frac{1}{3}\left(\frac{7}{12}\right)^d\frac{1}{3^{s_k}}\binom{s_k}{d+1}+\left(\frac{4}{9}\right)^{s_k},$$

as $\max_{l=1,\dots,d+1}\binom{s_k}{l}=\binom{s_k}{d+1}$ for $s_k \ge 2d+2$. Thus we obtain

$$\frac{1}{3^{s_k}}\sum_{\substack{\mathfrak{v}\subseteq D\\ \mathfrak{v}\neq\emptyset}}\left(\frac{3}{2}\right)^{|\mathfrak{v}|}\mathcal{B}(\mathfrak{v}) \le \frac{1}{2}\frac{1}{2^{s_k}}\frac{1}{4^d}\binom{d+s_k}{s_k-1}+\frac{1}{3}\left(\frac{7}{12}\right)^d\frac{1}{3^{s_k}}\binom{s_k}{d+1}+\left(\frac{4}{9}\right)^{s_k}.$$

Further we have

$$1-\left(1-\frac{1}{3\cdot 2^m}\right)^s \le \frac{s}{3\cdot 2^m}.$$

Hence it follows from the definition of $P_k(d)$ and (18) that

$$\mathcal{L}^2_{2,2^{t_k+d},\gamma_c}(P_k(d)) \le \frac{1}{2}\frac{1}{2^{s_k}}\frac{1}{4^d}\binom{d+s_k}{s_k-1}+\frac{1}{3}\left(\frac{7}{12}\right)^d\frac{1}{3^{s_k}}\binom{s_k}{d+1}+\left(\frac{4}{9}\right)^{s_k}+\frac{s_k}{3\cdot 2^{2(t_k+d)+s_k}}. \tag{22}$$

In order to get a bound on $\mathcal{D}^2_{t_k+d,s_k}(P_k(d))$ we need to multiply the inequality above with $4^{t_k+d}\left[\binom{t_k+d+s_k+1}{s_k-1}\right]^{-1}$. For the first term in the bound of (22) we get

$$\frac{1}{2}\frac{1}{2^{s_k}}\frac{1}{4^d}\binom{d+s_k}{s_k-1}4^{t_k+d}\left[\binom{t_k+d+s_k+1}{s_k-1}\right]^{-1}$$
$$= \frac{1}{2}\frac{1}{2^{s_k}}4^{t_k}\frac{(d+s_k)\cdots(d+2)}{(t_k+d+s_k+1)\cdots(t_k+d+3)}.$$

Let $r \ge 1$ be an integer to be chosen later. From (19) it follows that for large enough k we have $rt_k < s_k$. Further, $t_k > 0$. We get

$$\frac{(t_k+d+s_k+1)\cdots(t_k+d+3)}{(d+s_k)\cdots(d+2)} = \left(1+\frac{t_k+1}{d+s_k}\right)\cdots\left(1+\frac{t_k+1}{d+2}\right) \geq \prod_{j=1}^{r}\left(1+\frac{t_k+1}{jt_k+d+1}\right)^{t_k}.$$

Now we have

$$\prod_{j=1}^{r}\left(1+\frac{t_k+1}{jt_k+d+1}\right) \to \prod_{j=1}^{r}\left(1+\frac{1}{j}\right) = r+1 \quad \text{as } t_k \to \infty.$$

Therefore, for large enough k, we obtain

$$r^{t_k} \leq \frac{(t_k+d+s_k+1)\cdots(t_k+d+3)}{(d+s_k)\cdots(d+2)}$$

and

$$\frac{1}{2}\,\frac{1}{2^{s_k}}\,\frac{1}{4^d}\binom{d+s_k}{s_k-1}4^{t_k+d}\left[\binom{t_k+d+s_k+1}{s_k-1}\right]^{-1} \leq \frac{1}{2}\,\frac{1}{2^{s_k}}\left(\frac{4}{r}\right)^{t_k}$$

for all $k \geq K_1(r,d)$, for some well chosen $K_1(r,d)$. Further one can show that the other terms on the right hand side of (22) decay faster than $2^{-s_k}(4/r)^{t_k}$. From (20) it follows that $t_k \geq (d-1)\log_2 s_k$ for all $k \geq K_2(d)$. Let $r=8$; then we have $(4/r)^{t_k} \leq s_k^{1-d}$. Therefore there exists a K_d such that for all $k \geq K_d$ we have

$$\mathcal{D}^2_{t_k+d,s_k}(P_k(d)) \leq \frac{1}{2^{s_k}}\,\frac{1}{s_k^{d-1}}.$$

We summarize the result in the following theorem.

Theorem 5. *For any $d \geq 1$ there exists an integer $K_d > 0$ and a sequence of shifted digital (t_k, t_k+d, s_k)-nets $(P_k(d))_{k\geq 1}$ over $\mathbb{Z}_2$ with $s_k \to \infty$ as $k \to \infty$ and*

$$\left\lceil \frac{t_k}{\log_2 s_k} \right\rceil = d \quad \textit{for all } k \geq K_d,$$

such that for all $k \geq K_d$ we have

$$\mathcal{L}_{2,2^{t_k+d},\gamma_c}(P_k(d)) \leq \frac{1}{2^{t_k+d}}\,\frac{1}{2^{s_k/2}}\,\frac{1}{s_k^{(d-1)/2}}\sqrt{\binom{t_k+d+s_k+1}{s_k-1}}.$$

We use (21) again. Then by using (16) and the result above we conclude that for any $d > 0$ and for all $k \geq K_d$ we have

$$\frac{3}{16}\,\frac{1}{2^{2s_k}} \leq \mathcal{D}_{t_k+d,s_k}(P_k(d)) \leq \frac{1}{2^{s_k/2}}\,\frac{1}{s_k^{(d-1)/2}}. \tag{23}$$

This shows that Roth's lower bound is also in s of the best possible form. The small remaining gap in the constant is not surprising as the result in Theorem 5 was obtained by averaging over well distributed point sets. Some attempts have been made to improve the lower bound of Roth, but no considerable progress has been achieved (see [17]). For small point sets there exist other lower bounds which yield numerically better results than the bound of Roth, but do not show a higher convergence rate (see [17]).

We note that the results in this section are, apart from the digital shift, constructive as they are based on Niederreiter–Xing constructions of digital nets and sequences. It would also be desirable to have fully deterministic point sets with a small $\mathcal{L}_2$ discrepancy (like the constructions in [2]). In our context this amounts to finding an appropriate digital shift for a given digital (t, m, s)-net. This work remains to be done.

References

[1] J. Beck and W. W. L. Chen, *Irregularities of Distribution*, Cambridge Univ. Press, Cambridge, 1987.

[2] W. W. L. Chen and M. M. Skriganov, *Explicit constructions in the classical mean squares problem in irregularity of point distribution*, J. Reine Angew. Math. 545 (2002), 67–95.

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

[4] H. Davenport, *Note on irregularities of distribution*, Mathematika 3 (1956), 131–135.

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

[6] J. Dick, I. H. Sloan, X. Wang and H. Woźniakowski, *Good lattice rules in weighted Korobov spaces with general weights*, submitted.

[7] M. Drmota and R. F. Tichy, *Sequences*, *Discrepancies and Applications*, Lecture Notes in Math. 1651, Springer, Berlin, 1997.

[8] H. Faure, *Discrépance de suites associées à un système de numération* (*en dimension s*), Acta Arith. 41 (1982), 337–351.

[9] F. J. Hickernell, *The mean square discrepancy of randomized nets*, ACM Trans. Model. Comput. Simul. 6 (1996), 274–296.

[10] —, *A generalized discrepancy and quadrature error bound*, Math. Comp. 67 (1998), 299–322.

[11] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, New York, 1974.

[12] G. Larcher, *Digital point sets*: *analysis and application*, in: Random and Quasi-Random Point Sets, P. Hellekalek and G. Larcher (eds.), Lecture Notes in Statist. 138, Springer, 1998, 167–222.

[13] G. Larcher, H. Niederreiter and W. Ch. Schmid, *Digital nets and sequences constructed over finite rings and their application to quasi-Monte Carlo integration*, Monatsh. Math. 121 (1996), 231–253.

[14] G. Larcher and F. Pillichshammer, *Sums of distances to the nearest integer and the discrepancy of digital nets*, Acta Arith. 106 (2003), 379–408.

[15] G. Leobacher and F. Pillichshammer, *Bounds for the weighted L^p discrepancy and tractability of integration*, J. Complexity 19 (2003), 529–547.

[16] J. Matoušek, *Geometric Discrepancy*, Algorithms Combinatorics 18, Springer, Berlin, 1999.

[17] —, *On the $\mathcal{L}_2$ discrepancy for anchored boxes*, J. Complexity 14 (1998), 527–556.

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

[19] —, *Random Number Generation and Quasi-Monte Carlo Methods*, CBMS-NSF Regional Conf. Ser. in Appl. Math. 63, SIAM, Philadelphia, 1992.

[20] H. Niederreiter and C. P. Xing, *Constructions of digital nets*, Acta Arith. 102 (2002), 189–197.

[21] —, —, *Rational Points on Curves over Finite Fields*, London Math. Soc. Lecture Note Ser. 285, Cambridge Univ. Press, Cambridge, 2001.

[22] A. B. Owen, *Randomly permuted* (t, m, s)*-nets and* (t, s)*-sequences*, in: Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing (Las Vegas, NV, 1994), Lecture Notes in Statist. 106, Springer, New York, 1995, 299–317.

[23] F. Pillichshammer, *Improved upper bounds for the star discrepancy of digital nets in dimension* 3, Acta Arith. 108 (2003), 167–189.

[24] G. Pirsic, *Schnell konvergierende Walshreihen über Gruppen*, master's thesis, University of Salzburg, 1995; http://www.ricam.oeaw.ac.at/people/page/pirsic/.

[25] T. J. Rivlin and E. B. Saff, *Joseph L. Walsh Selected Papers*, Springer, New York, 2000.

[26] K. F. Roth, *On irregularities of distribution IV*, Acta Arith. 37 (1980), 67–75.

[27] —, *On irregularities of distribution*, Mathematika 1 (1954), 73–79.

[28] W. Ch. Schmid, *Shift-nets*: *a new class of binary digital* (t, m, s)*-nets*, in: Monte Carlo and Quasi-Monte Carlo Methods 1996, H. Niederreiter *et al.* (eds.), Lecture Notes in Statist. 127, Springer, 1997, 369–381.

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

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

[31] C. P. Xing and H. Niederreiter, *Digital nets*, *duality and algebraic curves*, in: Monte Carlo and Quasi-Monte Carlo Methods 2002, H. Niederreiter (ed.), Springer, Berlin, 2004, 155–166.

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

School of Mathematics
University of New South Wales
Sydney 2052, Australia
E-mail: josi@maths.unsw.edu.au

Institut für Finanzmathematik
Universität Linz
Altenbergstraße 69
A-4040 Linz, Austria
E-mail: friedrich.pillichshammer@jku.at

Received on 13.11.2003
and in revised form on 14.12.2004 (4666)