OPTIMAL L_p-DISCREPANCY BOUNDS FOR SECOND ORDER DIGITAL SEQUENCES

BY

JOSEF DICK*

School of Mathematics and Statistics, The University of New South Wales,
Sydney NSW 2052, Australia
e-mail: josef.dick@unsw.edu.au

AND

AICKE HINRICHS

Institut für Funktionalanalysis, Johannes Kepler Universität Linz,
Altenbergerstraße 69, 4040 Linz, Austria
e-mail: aicke.hinrichs@jku.at

AND

LEV MARKHASIN

Institut für Stochastik and Anwendungen, Universität Stuttgart,
Pfaffenwaldring 57, 70569 Stuttgart, Germany
e-mail: lev.markhasin@mathematik.uni-stuttgart.de

AND

FRIEDRICH PILLICHSHAMMER**

Institut für Finanzmathematik und angewandte Zahlentheorie,
Johannes Kepler Universität Linz, Altenbergerstraße 69, 4040 Linz, Austria
e-mail: friedrich.pillichshammer@jku.at

ABSTRACT

The L_p-discrepancy is a quantitative measure for the irregularity of distribution modulo one of infinite sequences. In 1986, Proinov proved for all $p > 1$ a lower bound for the L_p-discrepancy of general infinite sequences in the d-dimensional unit cube, but it remained an open question whether this lower bound is best possible in the order of magnitude until recently. In 2014, Dick and Pillichshammer gave a first construction of an infinite sequence whose order of L_2-discrepancy matches the lower bound of Proinov. Here we give a complete solution to this problem for all finite $p > 1$. We consider so-called order 2 digital (t, d)-sequences over the finite field with two elements and show that such sequences achieve the optimal order of L_p-discrepancy simultaneously for all $p \in (1, \infty)$.

1. Introduction

Let $d, N \in \mathbb{N}$ (where $\mathbb{N} = \{1, 2, 3, \dots\}$) and let $\mathcal{P}_{N,d}$ be an N-element point set in the unit cube $[0, 1)^d$. The discrepancy function of $\mathcal{P}_{N,d}$ is defined as

$$(1) \qquad D_{\mathcal{P}_{N,d}}(\boldsymbol{x}) = \frac{1}{N} \sum_{\boldsymbol{z} \in \mathcal{P}_{N,d}} \chi_{[\boldsymbol{0}, \boldsymbol{x})}(\boldsymbol{z}) - x_1 \cdots x_d$$

where $\boldsymbol{x} = (x_1, \dots, x_d) \in [0, 1]^d$ and $[\boldsymbol{0}, \boldsymbol{x}) = [0, x_1) \times \cdots \times [0, x_d)$. By χ_A we mean the characteristic function of a set $A \in \mathbb{R}^d$, i.e., $\chi_A(\boldsymbol{x}) = 1$ if $\boldsymbol{x} \in A$ and 0 if $\boldsymbol{x} \notin A$. The term $\sum_{\boldsymbol{z}} \chi_{[\boldsymbol{0}, \boldsymbol{x})}(\boldsymbol{z})$ in (1) is equal to the number of points of $\mathcal{P}_{N,d}$ in $[\boldsymbol{0}, \boldsymbol{x})$. Hence, $D_{\mathcal{P}_{N,d}}$ is a normalized measure for the deviation of the proportion of the number of points of $\mathcal{P}_{N,d}$ in $[\boldsymbol{0}, \boldsymbol{x})$ from the 'fair' or 'expected' proportion of the number of points $\lambda_d([\boldsymbol{0}, \boldsymbol{x})) = x_1 \cdots x_d$ in this interval under the assumption of a perfect uniform distribution. Here λ_d denotes the d-dimensional Lebesgue measure.

The L_p-discrepancy of $\mathcal{P}_{N,d}$ is defined as the L_p-norm of the discrepancy function, i.e.,

$$L_{p,N}(\mathcal{P}_{N,d}) = \|D_{\mathcal{P}_{N,d}}|L_p([0,1]^d)\| \qquad \text{for } p \in [1,\infty].$$

For an infinite sequence $\mathcal{S}_d$ in $[0,1)^d$ and $N \in \mathbb{N}$ the discrepancy function is defined as $D_{\mathcal{S}_d}^N(\boldsymbol{x}) = D_{\mathcal{P}_{N,d}}(\boldsymbol{x})$, where the point set $\mathcal{P}_{N,d}$ consists of the first N terms of $\mathcal{S}_d$, and the L_p-discrepancy of $\mathcal{S}_d$ is defined as

$$L_{p,N}(\mathcal{S}_d) = \|D_{\mathcal{S}_d}^N|L_p([0,1]^d)\| \qquad \text{for } p \in [1,\infty].$$

The L_p-discrepancy is a quantitative measure for the irregularity of distribution of finite point sets and of infinite sequences. We refer to [21, 30] for extensive introductions to this topic. It is well known that a sequence $\mathcal{S}_d$ is uniformly distributed modulo one in the sense of Weyl [54] if and only if $L_{p,N}(\mathcal{S}_d)$ tends to zero for $N \to \infty$. The L_p-discrepancy is also closely related to the worst-case integration error in certain function spaces using quasi-Monte Carlo algorithms via variants of the Koksma–Hlawka inequality. This follows immediately from Hlawka's identity (which is also sometimes attributed to Zaremba); see [27, 55] or also [17, 32, 41].

The conceptual difference between the discrepancy of finite point sets and infinite sequences can be explained in the following way (cf. [37]): while for finite point sets we are interested in the behavior of the whole set $\{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N-1}\}$ with a fixed number of elements N, for infinite sequences we are interested in the discrepancy of all initial segments $\{\boldsymbol{x}_0\}$, $\{\boldsymbol{x}_0, \boldsymbol{x}_1\}$, $\{\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2\}$, $\ldots$, $\{\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{N-1}\}$, where $N = 2, 3, 4, \ldots$. In this sense the discrepancy of finite point sets can be viewed as a static setting and the discrepancy of infinite sequences as a dynamic setting. Very often the dynamic setting in dimension d is related to the static setting in dimension $d + 1$ (see, for example, [30, Chapter 2.2, Theorem 2.2, Example 2.2]). This will also be confirmed by our results.

It is well known that for every $p \in (1,\infty]$ and every $d \in \mathbb{N}$ there exists a positive constant $c_{p,d}$ with the following property: for every finite N-element point set $\mathcal{P}_{N,d}$ in $[0,1)^d$ with $N \geq 2$ we have

$$(2) \qquad L_{p,N}(\mathcal{P}_{N,d}) \geq c_{p,d} \frac{(\log N)^{\frac{d-1}{2}}}{N}.$$

This has been first shown in a celebrated paper by Roth [43] for $p \geq 2$ and by Schmidt [46] for $p \in (1, 2)$. As shown by Halász [24] the estimate is also true

for $p = 1$ and $d = 2$, i.e., there exists a positive constant $c_{1,2}$ with the following property: for every finite N-element point set $\mathcal{P}_{N,2}$ in $[0,1)^2$ with $N \geq 2$ we have

$$(3) \qquad L_{1,N}(\mathcal{P}_{N,2}) \geq c_{1,2} \frac{(\log N)^{\frac{1}{2}}}{N}.$$

Later Proinov [42] (see also [19] for a proof) extended these results to infinite sequences: for every $p \in (1,\infty]$ and every $d \in \mathbb{N}$, $d \geq 2$, there exists a positive constant $c_{p,d}$ with the following property: for every infinite sequence $\mathcal{S}_d$ in $[0,1)^d$ we have

$$(4) \qquad L_{p,N}(\mathcal{S}_d) \geq c_{p,d} \frac{(\log N)^{\frac{d}{2}}}{N} \qquad \text{for infinitely many } N \in \mathbb{N}.$$

For $d = 1$ this estimate is also valid for $p = 1$, i.e., there exists a positive constant $c_{1,1}$ with the following property: for every infinite sequence $\mathcal{S}_1$ in $[0,1)$ we have

$$(5) \qquad L_{1,N}(\mathcal{S}_1) \geq c_{1,1} \frac{(\log N)^{\frac{1}{2}}}{N}.$$

This can be shown by combining Proinov's method [42] (see also [19]) with the result of Halász (3).

The lower bound (2) for finite point sets is known to be best possible in the order of magnitude in N, i.e., for every $d, N \in \mathbb{N}$, $N \geq 2$, one can find an N-element point set $\mathcal{P}_{N,d}$ in $[0,1)^d$ with L_p-discrepancy of order

$$(6) \qquad L_{p,N}(\mathcal{P}_{N,d}) \ll_{p,d} \frac{(\log N)^{\frac{d-1}{2}}}{N}.$$

For functions $f, g : D \subseteq \mathbb{N} \to \mathbb{R}$ with $g \geq 0$ we write $f(N) \ll g(N)$ if there exists some $C > 0$ such that $f(N) \leq Cg(N)$ for all $N \in D$. If we want to stress that C depends on some parameters, say a, b, then we indicate this by writing $f(N) \ll_{a,b} g(N)$. If we have $f(N) \ll_{a,b} g(N)$ and $g(N) \ll_{a,b} f(N)$ then we write $f(N) \asymp_{a,b} g(N)$.

The result in (6) was proved by Davenport [9] for $p = 2$, $d = 2$, by Roth [44] for $p = 2$ and arbitrary d and finally by Chen [6] in the general case. Other proofs were found by Frolov [23], Chen [7], Dobrovol'skiĭ [20], Skriganov [47, 48], Hickernell and Yue [25], and Dick and Pillichshammer [16]. For more details on the history of the subject see the monograph [1]. Apart from Davenport, who gave an explicit construction in dimension $d = 2$, these results are pure existence results and explicit constructions of point sets were not known until

the beginning of this millennium. First explicit constructions of point sets with optimal order of L_2-discrepancy have been provided in 2002 by Chen and Skriganov [8] for $p = 2$ and in 2006 by Skriganov [49] for general p. Other explicit constructions are due to Dick and Pillichshammer [18] for $p = 2$, and Dick [13] and Markhasin [36] for general p.

It is also known that the lower bound (4) for infinite sequences is best possible for the particular case $p \in (1, 2]$ (and also for $p = 1$ when $d = 1$). This was first shown by Dick and Pillichshammer [18] who gave an explicit construction of sequences whose L_2-discrepancy achieves the order of magnitude $(\log N)^{d/2}/N$ for all $N \geq 2$. For dimension $d = 1$ there exists a simple construction of a sequence with optimal order of L_p-discrepancy for all $p \in [1, \infty)$. Let $\mathcal{V} = (y_n)_{n \geq 0}$ be the van der Corput sequence (in base 2), i.e., $y_n = \sum_{j \geq 0} \frac{n_j}{2^{j+1}}$ whenever $n \in \mathbb{N}_0$ has binary expansion $n = \sum_{j \geq 0} n_j 2^j$ with digits $n_j \in \{0, 1\}$ (which is of course finite). Then let $\mathcal{V}^{\mathrm{sym}} = (z_n)_{n \geq 0}$ be the so-called symmetrized van der Corput sequence given by $z_{2n} = y_n$ and $z_{2n+1} = 1 - y_n$ for $n \in \mathbb{N}_0$. Then it has been shown in [29] that for all $p \in [1, \infty)$ we have

$$L_{p,N}(\mathcal{V}^{\mathrm{sym}}) \ll_p \frac{(\log N)^{\frac{1}{2}}}{N} \qquad \text{for all } N \geq 2.$$

A generalization of this result to van der Corput sequences in arbitrary base $b \geq 2$ has been shown quite recently by Kritzinger [28]. See also the recent survey article [22] and the references therein for more information about symmetrized van der Corput sequences.

In this paper we provide explicit constructions of infinite sequences in arbitrary dimensions d whose L_p-discrepancy is of order of magnitude $(\log N)^{d/2}/N$ for all $p \in [1, \infty)$. Thereby we prove that the lower estimate (4) is best possible for all $p \in (1, \infty)$.

The following is the main result of this work.

THEOREM 1.1: *There exists an explicit construction of an infinite sequence $\mathcal{S}_d$ in $[0, 1)^d$ with the property that*

$$L_{p,N}(\mathcal{S}_d) \ll_{p,d} \frac{(\log N)^{\frac{d}{2}}}{N} \qquad \text{for all } N \geq 2 \text{ and all } 1 \leq p < \infty.$$

A more concrete version of the main result will be stated in Section 2.3 as Theorem 2.2 and proved in Section 4. For $p = 2$ this result has been shown in

[18] (but with a more complex construction — see the discussion after Theorem 2.2) and a weaker result can be found in [13].

Proinov's lower bound (4) and Theorem 1.1 give the precise behavior of the L_p-discrepancy for $1 < p < \infty$. On the other hand, the L_∞-discrepancy remains elusive. We have constructions of infinite sequences $\mathcal{S}_d$ in $[0,1)^d$ (for example, order 1 digital (t,d)-sequences as presented in Section 2, see [17, 38, 39]) such that

$$L_{\infty,N}(\mathcal{S}_d) \ll_d \frac{(\log N)^d}{N}.$$

As to lower bounds, we know that there exists some $c_d > 0$ and $\eta_d \in (0, \frac{1}{2})$ such that for every sequence $\mathcal{S}_d$ in $[0,1)^d$ we have

$$L_{\infty,N}(\mathcal{S}_d) \geq c_d \frac{(\log N)^{\frac{d}{2}+\eta_d}}{N} \quad \text{for infinitely many } N \in \mathbb{N}.$$

This result follows from a corresponding result for finite point sets by Bilyk, Lacey and Vagharshakyan [4]. For growing d the exponent η_d in this estimate tends to zero.

In dimension $d = 1$ we even know that for every sequence $\mathcal{S}_1 \in [0,1)$ we have

$$L_{\infty,N}(\mathcal{S}_1) \geq c \frac{\log N}{N} \quad \text{for infinitely many } N \in \mathbb{N}$$

for some positive c. This is a famous result of Schmidt [45] (see also [2, 31]). Since the L_∞-discrepancy of the van der Corput sequence is of order $(\log N)/N$ the exact order of the L_∞-discrepancy of infinite sequences in dimension $d = 1$ is known. However, the quest for the exact order of the L_∞-discrepancy in the multivariate case is a very demanding open question.

It is a natural question to ask what happens in intermediate spaces "close" to L_∞. Standard examples of such spaces are BMO-spaces and exponential Orlicz spaces. For point sets, the norm of the discrepancy function in these spaces was studied in [3, 5]. The methods of this paper can also be used to give sharp bounds for sequences. This will be the subject of a follow-up paper [15].

Moreover, it is well-known that norms of the discrepancy function are intimately connected to integration errors of the corresponding quasi-Monte Carlo rules. That is the reason for recent work on the discrepancy function in function spaces like Sobolev spaces, Besov spaces and Triebel–Lizorkin spaces of dominating mixed smoothness; see [26, 33, 34, 35, 36, 52, 53]. Again the methods of our paper can be used to give sharp bounds for sequences. This will also be treated in [15].

The explicit construction in Theorem 1.1 is based on linear algebra over the finite field $\mathbb{F}_2$. In the subsequent section we provide a detailed introduction to the infinite sequences which lead to the optimal discrepancy bounds.

2. Digital nets and sequences

2.1. THE DIGITAL CONSTRUCTION SCHEME ACCORDING TO NIEDERREITER. The concepts of digital nets and sequences were introduced by Niederreiter [38] in 1987. These constructions are based on linear algebra over $\mathbb{F}_b$, the finite field of prime-power order b. A detailed overview of this topic is given in the books [17, 39] (see also [32, Chapter 5]). Here we restrict ourselves to the case $b = 2$. Let $\mathbb{F}_2$ be the finite field of order 2. We identify $\mathbb{F}_2$ with $\{0, 1\}$ equipped with arithmetic operations modulo 2.

First we recall the definition of digital nets according to Niederreiter, which we present here in a slightly more general form. For $n, q, d \in \mathbb{N}$ with $q \geq n$ let $C_1, \ldots, C_d \in \mathbb{F}_2^{q \times n}$ be $q \times n$ matrices over $\mathbb{F}_2 \cong \{0, 1\}$ (originally one used $n \times n$ matrices). For $k \in \{0, \ldots, 2^n - 1\}$ with dyadic expansion $k = k_0 + k_1 2 + \cdots + k_{n-1} 2^{n-1}$, where $k_j \in \{0, 1\}$, we define the dyadic digit vector $\vec{k}$ as $\vec{k} = (k_0, k_1, \ldots, k_{n-1})^\top \in \mathbb{F}_2^n$ (the symbol $\top$ means the transpose of a vector or a matrix; hence $\vec{k}$ is a column-vector). Then compute

$$(7) \qquad C_j \vec{k} =: (x_{j,k,1}, x_{j,k,2}, \ldots, x_{j,k,q})^\top \quad \text{for } j = 1, \ldots, d,$$

where the matrix vector product is evaluated over $\mathbb{F}_2$, and put

$$x_{j,k} = x_{j,k,1} 2^{-1} + x_{j,k,2} 2^{-2} + \cdots + x_{j,k,q} 2^{-q} \in [0, 1).$$

The k-th point $\boldsymbol{x}_k$ of the net $\mathcal{P}_{2^n, d}$ is given by $\boldsymbol{x}_k = (x_{1,k}, \ldots, x_{d,k})$. A net $\mathcal{P}_{2^n, d}$ constructed this way is called a *digital net (over $\mathbb{F}_2$) with generating matrices* $C_1, \ldots, C_d$. Note that a digital net consists of 2^n elements in $[0, 1)^d$.

A variant of digital nets are so-called *digitally shifted digital nets*. Here one chooses $(\vec{\sigma}_1, \ldots, \vec{\sigma}_d) \in (\mathbb{F}_2^\mathbb{N})^d$ with $\vec{\sigma}_j = (\sigma_{j,1}, \sigma_{j,2}, \ldots)^\top \in \mathbb{F}_2^\mathbb{N}$ with all but finitely many components different from zero and replaces (7) by

$$C_j \vec{k} + \vec{\sigma}_j =: (x_{j,k,1}, x_{j,k,2}, x_{j,k,3}, \ldots,)^\top \in \mathbb{F}_2^\mathbb{N} \quad \text{for } j = 1, \ldots, d,$$

and puts

$$x_{j,k} = x_{j,k,1} 2^{-1} + x_{j,k,2} 2^{-2} + x_{j,k,3} 2^{-3} + \cdots \in [0, 1).$$

We also recall the definition of digital sequences according to Niederreiter, which are infinite versions of digital nets. Let $C_1, \ldots, C_d \in \mathbb{F}_2^{\mathbb{N} \times \mathbb{N}}$ be $\mathbb{N} \times \mathbb{N}$ matrices over $\mathbb{F}_2$. For $C_j = (c_{j,k,\ell})_{k,\ell \in \mathbb{N}}$ we assume that for each $\ell \in \mathbb{N}$ there exists a $K(\ell) \in \mathbb{N}$ such that $c_{j,k,\ell} = 0$ for all $k > K(\ell)$. For $k \in \mathbb{N}_0$ with dyadic expansion $k = k_0 + k_1 2 + \cdots + k_{m-1} 2^{m-1} \in \mathbb{N}_0$, define the infinite dyadic digit vector of k by $\vec{k} = (k_0, k_1, \ldots, k_{m-1}, 0, 0, \ldots)^\top \in \mathbb{F}_2^{\mathbb{N}}$. Then compute

$$(8) \qquad C_j \vec{k} =: (x_{j,k,1}, x_{j,k,2}, \ldots)^\top \quad \text{for } j = 1, \ldots, d,$$

where the matrix vector product is evaluated over $\mathbb{F}_2$, and put

$$x_{j,k} = x_{j,k,1} 2^{-1} + x_{j,k,2} 2^{-2} + \cdots \in [0,1).$$

The k-th point $\boldsymbol{x}_k$ of the sequence $\mathcal{S}_d$ is given by $\boldsymbol{x}_k = (x_{1,k}, \ldots, x_{d,k})$. A sequence $\mathcal{S}_d$ constructed this way is called a *digital sequence (over $\mathbb{F}_2$) with generating matrices $C_1, \ldots, C_d$*. Note that since $c_{j,k,\ell} = 0$ for all k large enough, the numbers $x_{j,k}$ are always dyadic rationals, i.e., have a finite dyadic expansion.

The variant of *digitally shifted digital sequences* is defined in the same way as was done for digitally shifted digital nets.

2.2. HIGHER ORDER NETS AND SEQUENCES. Our approach is based on higher order digital nets and sequences constructed explicitly in [10, 11]. We state here simplified versions of these definitions which are sufficient for our purpose.

The distribution quality of digital nets and sequences depends on the choice of the respective generating matrices. In the following definitions we put some restrictions on $C_1, \ldots, C_d$ with the aim to quantify the quality of equidistribution of the digital net or sequence.

Definition 2.1: Let $n, q, \alpha \in \mathbb{N}$ with $q \geq \alpha n$ and let t be an integer such that $0 \leq t \leq \alpha n$. Let $C_1, \ldots, C_d \in \mathbb{F}_2^{q \times n}$. Denote the i-th row vector of the matrix C_j by $\vec{c}_{j,i} \in \mathbb{F}_2^n$. If for all $1 \leq i_{j,\nu_j} < \cdots < i_{j,1} \leq q$ with

$$\sum_{j=1}^{d} \sum_{l=1}^{\min(\nu_j, \alpha)} i_{j,l} \leq \alpha n - t$$

the vectors

$$\vec{c}_{1,i_{1,\nu_1}}, \ldots, \vec{c}_{1,i_{1,1}}, \ldots, \vec{c}_{d,i_{d,\nu_d}}, \ldots, \vec{c}_{d,i_{d,1}}$$

are linearly independent over $\mathbb{F}_2$, then the digital net with generating matrices $C_1, \ldots, C_d$ is called an **order α digital (t, n, d)-net over $\mathbb{F}_2$**.

The case $\alpha = 1$ corresponds to the classical case of (t, n, d)-nets according to Niederreiter's definition in [38].

Next we consider digital sequences for which the initial segments are order α digital (t, n, d)-nets over $\mathbb{F}_2$.

Definition 2.2: Let $\alpha \in \mathbb{N}$ and let $t \geq 0$ be an integer. Let $C_1, \ldots, C_d \in \mathbb{F}_2^{\mathbb{N} \times \mathbb{N}}$ and let $C_{j, \alpha n \times n}$ denote the left upper $\alpha n \times n$ submatrix of C_j. If for all $n > t/\alpha$ the matrices $C_{1, \alpha n \times n}, \ldots, C_{d, \alpha n \times n}$ generate an order α digital (t, n, d)-net over $\mathbb{F}_2$, then the digital sequence with generating matrices $C_1, \ldots, C_d$ is called an **order α digital (t, d)-sequence over $\mathbb{F}_2$**.

Again, the case $\alpha = 1$ corresponds to the classical case of (t, d)-sequences according to Niederreiter's definition in [38].

From Definition 2.1 it is clear that if $\mathcal{P}_{2^n, d}$ is an order α digital (t, n, d)-net, then for any $t \leq t' \leq \alpha n$, $\mathcal{P}_{2^n, d}$ is also an order α digital (t', n, d)-net. An analogous result also applies to higher order digital sequences.

Note that a digital net can be an order α digital (t, n, d)-net over $\mathbb{F}_2$ and at the same time an order α' digital (t', n, d)-net over $\mathbb{F}_2$ for $\alpha' \neq \alpha$. This means that the quality parameter t may depend on α (i.e., $t = t(\alpha)$). The same holds for digital sequences. In particular [11, Theorem 4.10] implies that an order α digital (t, n, d)-net is an order α' digital (t', n, d)-net for all $1 \leq \alpha' \leq \alpha$ with

$$(9) \qquad\qquad t' = \lceil t\alpha'/\alpha \rceil \leq t.$$

The same result applies to order α digital (t, d)-sequences which are also order α' digital (t', d)-sequences with $1 \leq \alpha' \leq \alpha$ and t' as above. In other words, $t(\alpha') = \lceil t(\alpha)\alpha'/\alpha \rceil$ for all $1 \leq \alpha' \leq \alpha$. More information can be found in [17, Chapter 15].

In [18] it has been shown that every order α digital (t, d)-sequence over $\mathbb{F}_2$ with $\alpha \geq 5$ has optimal order of the L_2-discrepancy. In this paper we show that even order 2 digital (t, d)-sequences over $\mathbb{F}_2$ achieve the optimal order of magnitude in N of the L_p-discrepancy for all $p \in (1, \infty)$.

Higher order digital nets and sequences have also a geometrical interpretation. Roughly speaking, the definitions imply that special intervals or unions of intervals of prescribed volume contain the correct share of points with respect to a perfect uniform distribution. See [39, 17] for the classical case $\alpha = 1$ and [14] or [17, 18] for general α. See also Lemma 3.2 below.

2.3. Explicit constructions of order 2 digital sequences. Explicit constructions of order α digital nets and sequences have been provided by Dick [10, 11]. For our purposes it suffices to consider only $\alpha = 2$.

Let $C_1, \ldots, C_{2d}$ be generating matrices of a digital net or sequence and let $\vec{c}_{j,k}$ denote the k-th row of C_j. Define matrices $E_1, \ldots, E_d$, where the k-th row of E_j is given by $\vec{e}_{j,k}$, in the following way. For all $j \in \{1, 2, \ldots, d\}$, $u \in \mathbb{N}_0$ and $v \in \{1, 2\}$ let

$$\vec{e}_{j,2u+v} = \vec{c}_{2(j-1)+v, u+1}.$$

We illustrate the construction for $d = 1$. Then

$$C_1 = \begin{pmatrix} \vec{c}_{1,1} \\ \vec{c}_{1,2} \\ \vdots \end{pmatrix}, \ C_2 = \begin{pmatrix} \vec{c}_{2,1} \\ \vec{c}_{2,2} \\ \vdots \end{pmatrix} \Rightarrow E_1 = \begin{pmatrix} \vec{c}_{1,1} \\ \vec{c}_{2,1} \\ \vec{c}_{1,2} \\ \vec{c}_{2,2} \\ \vdots \end{pmatrix}.$$

This procedure is called *interlacing* (in this case the so-called interlacing factor is 2).

Recall that above we assumed that $c_{j,k,\ell} = 0$ for all $k > K(\ell)$. Let $E_j = (e_{j,k,\ell})_{k,\ell \in \mathbb{N}}$. Then the construction yields that $e_{j,k,\ell} = 0$ for all $k > 2K(\ell)$.

From [10, Theorem 4.11 and Theorem 4.12] we obtain the following result.

PROPOSITION 2.1: *If $C_1, \ldots, C_{2d} \in \mathbb{F}_2^{\mathbb{N} \times \mathbb{N}}$ generate an order 1 digital $(t', 2d)$-sequence over $\mathbb{F}_2$, then $E_1, \ldots, E_d \in \mathbb{F}_2^{\mathbb{N} \times \mathbb{N}}$ generate an order 2 digital (t, d)-sequence over $\mathbb{F}_2$ with*

$$t = 2t' + d.$$

Explicit constructions of suitable generating matrices $C_1, \ldots, C_s$ over $\mathbb{F}_2$ were obtained by Sobol' [50], Niederreiter [38, 39], Niederreiter–Xing [40] and others (see [17, Chapter 8] for an overview). Any of these constructions is sufficient for our purpose; however, for completeness, we briefly describe a special case of Tezuka's construction [51], which is a generalization of Sobol's construction [50] and Niederreiter's construction [38] of the generating matrices.

We explain how to construct the entries $c_{j,k,\ell} \in \mathbb{F}_2$ of the generating matrices $C_j = (c_{j,k,\ell})_{k,\ell \geq 1}$ for $j = 1, 2, \ldots, s$ (for our purpose $s = 2d$). To this end choose the polynomials $p_1 = x$ and $p_j \in \mathbb{F}_2[x]$ for $j = 2, \ldots, s$ to be the $(j-1)$-th irreducible polynomial in a list of irreducible polynomials over $\mathbb{F}_2$ that is sorted in increasing order according to their degree $e_j = \deg(p_j)$, that

is, $e_2 \leq e_3 \leq \cdots \leq e_{s-1}$ (the ordering of polynomials with the same degree is irrelevant). We also put $e_1 = \deg(x) = 1$.

Let $j \in \{1, \ldots, s\}$ and $k \in \mathbb{N}$. Take $i - 1$ and z to be respectively the main term and remainder when we divide $k - 1$ by e_j, so that $k - 1 = (i - 1)e_j + z$, with $0 \leq z < e_j$. Now consider the Laurent series expansion

$$\frac{x^{e_j - z - 1}}{p_j(x)^i} = \sum_{\ell=1}^{\infty} a_\ell(i, j, z) x^{-\ell} \in \mathbb{F}_2((x^{-1})).$$

For $\ell \in \mathbb{N}$ we set

(10) $$c_{j,k,\ell} = a_\ell(i, j, z).$$

Every digital sequence with generating matrices $C_j = (c_{j,k,\ell})_{k,\ell \geq 1}$ for $j = 1, 2, \ldots, s$ found in this way is a special instance of a Sobol' sequence which in turn is a special instance of so-called generalized Niederreiter sequences (see [51, (3)]). Note that in the construction above we always have $c_{j,k,\ell} = 0$ for all $k > \ell$. The t-value for these sequences is known to be $t = \sum_{j=1}^{s}(e_j - 1)$; see [39, Chapter 4.5] for the case of Niederreiter sequences.

REMARK: Let $C_1, \ldots, C_{2d}$ be $\mathbb{N} \times \mathbb{N}$ matrices which are constructed according to Tezuka's method as described above. Let $E_1, \ldots, E_d$ be the generator matrices of the corresponding order 2 digital sequence. Then we always have $e_{j,k,\ell} = 0$ for all $k > 2\ell$, where $e_{j,k,\ell}$ is the entry in row k and column ℓ of the matrix E_j.

Now we can state a more concrete version of our main result.

THEOREM 2.2: *For every order 2 digital (t, d)-sequence $\mathcal{S}_d$ over $\mathbb{F}_2$, with generating matrices $E_i = (e_{i,k,\ell})_{k,\ell \geq 1}$ for which $e_{i,k,\ell} = 0$ for all $k > 2\ell$ and for all $i \in \{1, 2, \ldots, d\}$, we have*

$$L_{p,N}(\mathcal{S}_d) \ll_{p,d} 2^t \frac{(\log N)^{\frac{d}{2}}}{N} \qquad \text{for all } N \geq 2 \text{ and all } 1 \leq p < \infty.$$

We remark that this result is not only a generalization of the main result in [18] from L_2- to L_p-discrepancy for general finite p, but also a considerable improvement in the following sense. In [18] the explicit construction is based on higher order sequences of order $\alpha = 5$. Here, on the other hand, we show that even $\alpha = 2$ suffices in order to achieve the optimal discrepancy bound with respect to the lower bound (4). This means that for the explicit construction

of a sequence in dimension d one can begin with a classical digital sequence in dimension $s = 2d$ rather than $s = 5d$.

Note that the explicit construction of the order 2 digital (t, d)-sequences $\mathcal{S}_d$ shown above, with generating matrices $E_i = (e_{i,k,\ell})_{k,\ell \geq 1}$ for which $e_{i,k,\ell} = 0$ for all $k > 2\ell$ and for all $i \in \{1, 2, \ldots, d\}$, does not depend on the parameter p in Theorem 2.2. Our explicit construction (based on Tezuka's construction and the interlacing of the generating matrices) is also extensible in the dimension, i.e., if we have constructed the sequence $\mathcal{S}_d$, we can add one more coordinate to obtain the sequence $\mathcal{S}_{d+1}$. In other words, we can define a sequence $\mathcal{S}_\infty$ of points in $[0, 1)^{\mathbb{N}}$ and obtain the sequence $\mathcal{S}_d$, $d \in \mathbb{N}$, by projecting $\mathcal{S}_\infty$ to the first d coordinates.

3. Haar bases

The proof of Theorem 2.2 is based on Haar functions. This is in contrast to the proof of the result in [18] (the L_2-discrepancy of order 5 digital sequences is of optimal order) which is based on Walsh functions.

We define $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and $\mathbb{N}_{-1} = \mathbb{N}_0 \cup \{-1\}$. Let $\mathbb{D}_j = \{0, 1, \ldots, 2^j - 1\}$ for $j \in \mathbb{N}_0$ and $\mathbb{D}_{-1} = \{0\}$. For $\boldsymbol{j} = (j_1, \ldots, j_d) \in \mathbb{N}_{-1}^d$ let $\mathbb{D}_{\boldsymbol{j}} = \mathbb{D}_{j_1} \times \cdots \times \mathbb{D}_{j_d}$. For $\boldsymbol{j} \in \mathbb{N}_{-1}^d$ we write $|\boldsymbol{j}| = \max(j_1, 0) + \cdots + \max(j_d, 0)$.

For $j \in \mathbb{N}_0$ and $m \in \mathbb{D}_j$ we call the interval $I_{j,m} = \left[2^{-j}m, 2^{-j}(m + 1)\right)$ the m-th dyadic interval in $[0, 1)$ on level j. We put $I_{-1,0} = [0, 1)$ and call it the 0-th dyadic interval in $[0, 1)$ on level -1. Let $I_{j,m}^+ = I_{j+1,2m}$ and $I_{j,m}^- = I_{j+1,2m+1}$ be the left and right half of $I_{j,m}$, respectively. For $\boldsymbol{j} \in \mathbb{N}_{-1}^d$ and $\boldsymbol{m} = (m_1, \ldots, m_d) \in \mathbb{D}_{\boldsymbol{j}}$ we call $I_{\boldsymbol{j},\boldsymbol{m}} = I_{j_1,m_1} \times \cdots \times I_{j_d,m_d}$ the $\boldsymbol{m}$-th dyadic interval in $[0, 1)^d$ on level $\boldsymbol{j}$. We call the number $|\boldsymbol{j}|$ the order of the dyadic interval $I_{\boldsymbol{j},\boldsymbol{m}}$. Its volume is $2^{-|\boldsymbol{j}|}$.

Let $j \in \mathbb{N}_0$ and $m \in \mathbb{D}_j$. Let $h_{j,m}$ be the function on $[0, 1)$ with support in $I_{j,m}$ and the constant values 1 on $I_{j,m}^+$ and -1 on $I_{j,m}^-$. We put $h_{-1,0} = \chi_{I_{-1,0}}$ on $[0, 1)$. The function $h_{j,m}$ is called the m-th dyadic Haar function on level j.

Let $\boldsymbol{j} \in \mathbb{N}_{-1}^d$ and $\boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}}$. The function $h_{\boldsymbol{j},\boldsymbol{m}}$, given as the tensor product

$$h_{\boldsymbol{j},\boldsymbol{m}}(\boldsymbol{x}) = h_{j_1,m_1}(x_1) \cdots h_{j_d,m_d}(x_d)$$

for $\boldsymbol{x} = (x_1, \ldots, x_d) \in [0, 1)^d$, is called a *dyadic Haar function* on $[0, 1)^d$. The system of dyadic Haar functions $h_{\boldsymbol{j},\boldsymbol{m}}$ for $\boldsymbol{j} \in \mathbb{N}_{-1}^d$, $\boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}}$ is called the *dyadic Haar basis on* $[0, 1)^d$.

It is well known that the system

$$\left\{ 2^{\frac{|\boldsymbol{j}|}{2}} h_{\boldsymbol{j},\boldsymbol{m}} \; : \; \boldsymbol{j} \in \mathbb{N}_{-1}^d, \; \boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}} \right\}$$

is an orthonormal basis of $L_2([0,1)^d)$, an unconditional basis of $L_p([0,1)^d)$ for $1 < p < \infty$ and a conditional basis of $L_1([0,1)^d)$. For any function $f \in L_2([0,1)^d)$ we have Parseval's identity

$$\|f|L_2([0,1)^d)\|^2 = \sum_{\boldsymbol{j} \in \mathbb{N}_{-1}^d} 2^{|\boldsymbol{j}|} \sum_{\boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}}} |\langle f, h_{\boldsymbol{j},\boldsymbol{m}} \rangle|^2,$$

where $\langle \cdot, \cdot \rangle$ denotes the usual L_2-inner product, i.e., $\langle f, g \rangle = \int_{[0,1]^d} f(\boldsymbol{x}) g(\boldsymbol{x}) \, d\boldsymbol{x}$. The terms $\langle f, h_{\boldsymbol{j},\boldsymbol{m}} \rangle$ are called the *Haar coefficients* of the function f.

The following Littlewood–Paley type estimate for the L_p-norm for $p \in (1, \infty)$ is a special case of [34, Theorem 2.11, Corollary 1.13].

PROPOSITION 3.1: *Let $p \in (1, \infty)$ and $f \in L_p([0,1)^d)$. Then*

$$\|f|L_p([0,1)^d)\|^2 \ll_{p,d} \sum_{\boldsymbol{j} \in \mathbb{N}_{-1}^d} 2^{2|\boldsymbol{j}|(1-1/\bar{p})} \left(\sum_{\boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}}} |\langle f, h_{\boldsymbol{j},\boldsymbol{m}} \rangle|^{\bar{p}} \right)^{2/\bar{p}},$$

where $\bar{p} = \max(p, 2)$.

We present a connection between higher order digital nets over $\mathbb{F}_2$ and dyadic intervals.

LEMMA 3.2: *Let $\mathcal{P}_{2^n,d}$ be an order α digital (t,n,d)-net over $\mathbb{F}_2$. Then every dyadic interval of order $n - \lceil t/\alpha \rceil$ contains at most $2^{\lceil t/\alpha \rceil}$ points of $\mathcal{P}_{2^n,d}$.*

Proof. As mentioned in Section 2.2, every order α digital (t,n,d)-net over $\mathbb{F}_2$ is an order 1 digital $(\lceil t/\alpha \rceil, n, d)$-net over $\mathbb{F}_2$. Then every dyadic interval of order $n - \lceil t/\alpha \rceil$ contains exactly $2^{\lceil t/\alpha \rceil}$ points of $\mathcal{P}_{2^n,d}$ (see [17, 39]). ■

The following lemma is a slight generalization of [36, Lemma 5.9]. The result was originally proved for order 2 digital (t,n,d)-nets. The extension to digitally shifted order 2 digital (t,n,d)-nets follows with almost exactly the same arguments as the proof of [36, Lemma 5.9] (not repeated here). We restrict ourselves to the finite field $\mathbb{F}_2$.

LEMMA 3.3: *Let $\mathcal{P}_{2^n,d}$ be a digitally shifted order 2 digital (t,n,d)-net over $\mathbb{F}_2$. Let $\boldsymbol{j} \in \mathbb{N}_{-1}^d$ with $|\boldsymbol{j}| + t/2 \leq n$ and $\boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}}$. Then*

$$|\langle D_{\mathcal{P}_{2^n,d}}, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \ll 2^{-2n+t}(2n-t-2|\boldsymbol{j}|)^{d-1}.$$

4. The proof of Theorem 2.2

For the proof of the main result we need some auxiliary lemmas. The first two lemmas are elementary:

LEMMA 4.1: *Let $r \in \mathbb{N}_0$ and $s \in \mathbb{N}$. Then*

$$\#\{(a_1,\ldots,a_s) \in \mathbb{N}_0^s : a_1 + \cdots + a_s = r\} \leq (r+1)^{s-1}.$$

For a proof of this result we refer to [17, Proof of Lemma 16.26].

LEMMA 4.2: *Let $K \in \mathbb{N}$, $A > 1$ and $q, s \geq 0$. Then we have*

$$\sum_{r=0}^{K-1} A^r (K-r)^q r^s \ll A^K K^s,$$

where the implicit constant is independent of K.

For a proof we refer to [36, Lemma 5.2].

The subsequent two lemmas are required in order to estimate the Haar coefficients of the discrepancy function. The first one is a special case of [34, Lemma 5.1].

LEMMA 4.3: *Let $f(\boldsymbol{x}) = x_1 \cdots x_d$ for $\boldsymbol{x} = (x_1,\ldots,x_d) \in [0,1)^d$. Let $\boldsymbol{j} \in \mathbb{N}_{-1}^d$ and $\boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}}$. Then $|\langle f, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \asymp 2^{-2|\boldsymbol{j}|}$.*

The next lemma is a special case of [34, Lemma 5.2].

LEMMA 4.4: *Let $\boldsymbol{z} = (z_1,\ldots,z_d) \in [0,1)^d$ and $g(\boldsymbol{x}) = \chi_{[0,\boldsymbol{x})}(\boldsymbol{z})$ for $\boldsymbol{x} = (x_1,\ldots,x_d) \in [0,1)^d$. Let $\boldsymbol{j} \in \mathbb{N}_{-1}^d$ and $\boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}}$. Then $\langle g, h_{\boldsymbol{j},\boldsymbol{m}}\rangle = 0$ if $\boldsymbol{z}$ is not contained in the interior of the dyadic interval $I_{\boldsymbol{j},\boldsymbol{m}}$. If $\boldsymbol{z}$ is contained in the interior of $I_{\boldsymbol{j},\boldsymbol{m}}$ then $|\langle g, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \ll 2^{-|\boldsymbol{j}|}$.*

Now we have collected all auxiliary results in order to provide the proof of Theorem 2.2.

Proof of Theorem 2.2. According to the monotonicity of the L_p-norm it suffices to prove the result for $p > 1$. Let $\mathcal{S}_d$ be an order 2 digital (t,d)-sequence over $\mathbb{F}_2$

with generating matrices $C_1, \ldots, C_d$, with $C_j = (c_{j,k,\ell})_{k,\ell \geq 1}$ for which $c_{i,k,\ell} = 0$ for all $k > 2\ell$. Let $N \in \mathbb{N}$ with dyadic expansion $N = 2^{n_r} + \cdots + 2^{n_1}$ with $n_r > \cdots > n_1 \geq 0$.

We first prove the following claim: For $\mu = 1, \ldots, r$ let

$$\mathcal{Q}_{2^{n_\mu}, d} = \{ \boldsymbol{x}_{2^{n_1} + \cdots + 2^{n_{\mu-1}}}, \boldsymbol{x}_{2^{n_1} + \cdots + 2^{n_{\mu-1}} + 1}, \ldots, \boldsymbol{x}_{-1 + 2^{n_1} + \cdots + 2^{n_\mu}} \},$$

where for $\mu = 1$ we set $2^{n_1} + \cdots + 2^{n_{\mu-1}} = 0$. Then the point set $\mathcal{P}_{N,d}$ consisting of the first N elements of the sequence $\mathcal{S}_d$ is a union of $\mathcal{Q}_{2^{n_\mu}, d}$ for $\mu = 1, \ldots, r$ and $\mathcal{Q}_{2^{n_\mu}, d}$ is a digitally shifted order 2 digital (t, n_μ, d)-net over $\mathbb{F}_2$ with generating matrices $C_{1, 2n_\mu \times n_\mu}, \ldots, C_{d, 2n_\mu \times n_\mu}$, i.e., the left upper $2n_\mu \times n_\mu$ submatrices of $C_1, \ldots, C_d$.

For the proof of this claim let $C_{j, \mathbb{N} \times n_\mu}$ denote the matrix which consists of the first n_μ columns of C_j. Only the first $2n_\mu$ rows of $C_{j, \mathbb{N} \times n_\mu}$ can be nonzero since $c_{j,k,\ell} = 0$ for all $k > 2\ell$ and hence C_j is of the form

$$C_j = \begin{pmatrix} C_{j, 2n_\mu \times n_\mu} & D_{j, 2n_\mu \times \mathbb{N}} \\ \hline 0_{\mathbb{N} \times n_\mu} & F_{j, \mathbb{N} \times \mathbb{N}} \end{pmatrix} \in \mathbb{F}_2^{\mathbb{N} \times \mathbb{N}},$$

where $0_{\mathbb{N} \times n_\mu}$ denotes the $\mathbb{N} \times n_\mu$ zero matrix. Note that the entries of each column of the matrix $F_{j, \mathbb{N} \times \mathbb{N}}$ become eventually zero. Any

$$k \in \{ 2^{n_1} + \cdots + 2^{n_{\mu-1}}, 2^{n_1} + \cdots + 2^{n_{\mu-1}} + 1, \ldots, -1 + 2^{n_1} + \cdots + 2^{n_\mu} \}$$

can be written in the form

$$k = 2^{n_1} + \cdots + 2^{n_{\mu-1}} + a = 2^{n_{\mu-1}} \ell + a$$

with $a \in \{ 0, 1, \ldots, 2^{n_\mu} - 1 \}$ and $\ell = 1 + 2^{n_\mu - n_{\mu-1}} + \cdots + 2^{n_1 - n_{\mu-1}}$ if $\mu > 1$ and $\ell = 0$ if $\mu = 1$. Hence the dyadic digit vector of k is given by

$$\vec{k} = (a_0, a_1, \ldots, a_{n_\mu - 1}, l_0, l_1, l_2, \ldots)^\top =: \begin{pmatrix} \vec{a} \\ \vec{\ell} \end{pmatrix},$$

where $a_0, \ldots, a_{n_\mu - 1}$ are the dyadic digits of a and $l_0, l_1, l_2, \ldots$ are the dyadic digits of ℓ. With this notation we have

$$
C_j \vec{k} = \begin{pmatrix} C_{j,2n_\mu \times n_\mu} \vec{a} \\ 0 \\ 0 \\ \vdots \end{pmatrix} + \left(\begin{array}{c} D_{j,2n_\mu \times \mathbb{N}} \\ \hline \\ F_{j,\mathbb{N} \times \mathbb{N}} \end{array} \right) \vec{\ell}.
$$

For the point set $\mathcal{Q}_{2^{n_\mu}, d}$ under consideration, the vector

$$
(11) \qquad \vec{\sigma}_{\mu,j} := \left(\begin{array}{c} D_{j,\alpha n_\mu \times \mathbb{N}} \\ \hline \\ F_{j,\mathbb{N} \times \mathbb{N}} \end{array} \right) \vec{\ell}
$$

is constant and its components become eventually zero (i.e., only a finite number of components is nonzero). Furthermore, $C_{j,2n_\mu \times n_\mu} \vec{a}$ for $a = 0, 1, \ldots, 2^{n_\mu} - 1$ and $j = 1, \ldots, s$ generate an order 2 digital (t, n_μ, s)-net over $\mathbb{F}_2$ (which is also an order 1 digital (t, n_μ, d)-net over $\mathbb{F}_2$, which follows from [12, Proposition 1]).

This means that the point set $\mathcal{Q}_{2^{n_\mu}, d}$ is a digitally shifted order 2 digital (t, n_μ, d)-net over $\mathbb{F}_2$ with generating matrices $C_{1,2n_\mu \times n_\mu}, \ldots, C_{d,2n_\mu \times n_\mu}$ and hence the claim is proven.

According to Proposition 3.1 with $f = D_{\mathcal{S}_d}^N$ we have

$$
(12) \qquad (L_{p,N}(\mathcal{S}_d))^2 \ll_{p,d} \sum_{\boldsymbol{j} \in \mathbb{N}_{-1}^d} 2^{2|\boldsymbol{j}|(1 - 1/\bar{p})} \left(\sum_{\boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}}} |\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}} \rangle|^{\bar{p}} \right)^{2/\bar{p}},
$$

where $\bar{p} = \max(p, 2)$.

Now we split up the sum in (12) according to the size $|\boldsymbol{j}|$ of the involved dyadic intervals. Let us consider "small" dyadic intervals first, so let $|\boldsymbol{j}| + t/2 \geq \operatorname{ld} N$. We calculate $|\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}} \rangle|$. Choose $n \in \mathbb{N}$ such that $2^{n-1} < N \leq 2^n$. Then the point set $\mathcal{P}_{N,d}$ of the first N elements of $\mathcal{S}_d$ is a subset of $\mathcal{P}_{2^n, d}$ consisting of the first 2^n elements of $\mathcal{S}_d$. From the construction of $\mathcal{S}_d$ it follows that $\mathcal{P}_{2^n, d}$ is an order 2 digital (t, n, d)-net over $\mathbb{F}_2$. Therefore, according to Lemma 3.2, in an interval $I_{\boldsymbol{j},\boldsymbol{m}}$ there are at most $2^{\lceil t/2 \rceil}$ points of $\mathcal{P}_{2^n, d}$ and hence there are at

most $2^{\lceil t/2 \rceil}$ points of $\mathcal{P}_{N,d}$ in $I_{\boldsymbol{j},\boldsymbol{m}}$. Hence we get from Lemmas 4.3 and 4.4

$$
(13) \qquad |\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \leq \frac{1}{N}\frac{2^{\lceil t/2 \rceil}}{2^{|\boldsymbol{j}|}} + \frac{1}{2^{2|\boldsymbol{j}|}} \ll \frac{1}{N}\frac{2^{t/2}}{2^{|\boldsymbol{j}|}}.
$$

The estimate (13) will be applied to dyadic intervals on level $\boldsymbol{j}$ which contain points from $\mathcal{P}_{N,d}$. The cardinality of such intervals is at most N. At least $2^{|\boldsymbol{j}|} - N$ contain no points of $\mathcal{P}_{N,d}$, hence in such cases we get from Lemmas 4.3 and 4.4

$$
(14) \qquad |\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \ll \frac{1}{2^{2|\boldsymbol{j}|}}.
$$

Now we estimate the terms of (12) for which $|\boldsymbol{j}| + t/2 \geq \operatorname{ld} N$. Applying Minkowski's inequality and inserting (13) and (14) we obtain

$$
\sum_{|\boldsymbol{j}|+t/2\geq \operatorname{ld} N} 2^{2|\boldsymbol{j}|(1-1/\bar{p})} \left(\sum_{\boldsymbol{m}\in \mathbb{D}_{\boldsymbol{j}}} |\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^{\bar{p}} \right)^{2/\bar{p}}
$$

$$
\ll \sum_{|\boldsymbol{j}|+t/2\geq \operatorname{ld} N} 2^{2|\boldsymbol{j}|(1-1/\bar{p})} \left(N\frac{1}{N^{\bar{p}}}\frac{2^{\bar{p}t/2}}{2^{\bar{p}|\boldsymbol{j}|}} \right)^{2/\bar{p}}
$$

$$
(15) \qquad\qquad + \sum_{|\boldsymbol{j}|+t/2\geq \operatorname{ld} N} 2^{2|\boldsymbol{j}|(1-1/\bar{p})} \left(\frac{(2^{|\boldsymbol{j}|}-N)}{2^{2\bar{p}|\boldsymbol{j}|}} \right)^{2/\bar{p}}
$$

$$
\leq N^{2/\bar{p}-2}2^t \sum_{|\boldsymbol{j}|+t/2\geq \operatorname{ld} N} \frac{1}{2^{2|\boldsymbol{j}|/\bar{p}}} + \sum_{|\boldsymbol{j}|+t/2\geq \operatorname{ld} N} \frac{1}{2^{2|\boldsymbol{j}|}}
$$

$$
\ll N^{2/\bar{p}-2}2^t N^{-2/\bar{p}}2^{t/\bar{p}}(\operatorname{ld} N)^{d-1} + \frac{(\log N)^{d-1}}{N^2}2^t
$$

$$
\leq 2^{2t}\frac{(\log N)^{d-1}}{N^2}.
$$

We now turn to the more demanding case of "large" intervals where $|\boldsymbol{j}| + t/2 < \operatorname{ld} N$. More precisely, assume that we have

$$
n_\mu \leq |\boldsymbol{j}| + t/2 < n_{\mu+1}
$$

for some $\mu \in \{0, 1, \ldots, r\}$, where we set $n_0 = 0$ and $n_{r+1} = \operatorname{ld} N$. For $\kappa \in \{\mu+1, \ldots, r\}$ we use the estimation from Lemma 3.3 to get

$$
|\langle D_{\mathcal{Q}_{2^{n_\kappa},d}}, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \ll \frac{(2n_\kappa - t - 2|\boldsymbol{j}|)^{d-1}}{2^{2n_\kappa - t}},
$$

while for $M = 2^{n_\mu} + \cdots + 2^{n_1}$, with $\mu > 0$, we have according to (13) with $\widetilde{\mathcal{Q}}_{M,d} = \bigcup_{i=1}^{\mu} \mathcal{Q}_{2^{n_i},d}$ that

$$|\langle D_{\widetilde{\mathcal{Q}}_{M,d}}, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \ll \frac{1}{M}\frac{1}{2^{|\boldsymbol{j}|-t/2}},$$

since $|\boldsymbol{j}| + t/2 \geq \operatorname{ld} M$. If $\mu = 0$, this case does not occur.

Using the linearity of the discrepancy function and the triangle inequality leads to

$$|\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \leq \frac{M}{N}|\langle D_{\widetilde{\mathcal{Q}}_{M,d}}, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| + \frac{1}{N}\sum_{\kappa=\mu+1}^{r} 2^{n_\kappa}|\langle D_{\mathcal{Q}_{2^{n_\kappa},d}}, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|$$

$$\ll \frac{1}{N}\left(\frac{2^{t/2}}{2^{|\boldsymbol{j}|}} + \sum_{\kappa=\mu+1}^{r}\frac{(2n_\kappa - t - 2|\boldsymbol{j}|)^{d-1}}{2^{n_\kappa - t}}\right)$$

(16)

$$\ll \frac{2^t}{N}\left(\frac{1}{2^{|\boldsymbol{j}|}} + \sum_{k=0}^{\infty}\frac{(2n_{\mu+1} + 2k - t - 2|\boldsymbol{j}|)^{d-1}}{2^{n_{\mu+1}+k}}\right)$$

$$\ll \frac{2^t}{N}\left(\frac{1}{2^{|\boldsymbol{j}|}} + \frac{(2n_{\mu+1} - t - 2|\boldsymbol{j}|)^{d-1}}{2^{n_{\mu+1}}}\right),$$

where we used [17, Lemma 13.24] in the last step.

Now we use this bound to estimate the terms of (12) for which $|\boldsymbol{j}|+t/2 < \operatorname{ld} N$. We obtain from Minkowski's inequality, Lemma 4.1 and Lemma 4.2

$$\sum_{|\boldsymbol{j}|+t/2<\operatorname{ld} N} 2^{2|\boldsymbol{j}|(1-1/\bar{p})}\left(\sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}}|\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^{\bar{p}}\right)^{2/\bar{p}}$$

$$= \sum_{\mu=0}^{r}\sum_{n_\mu \leq |\boldsymbol{j}|+t/2<n_{\mu+1}} 2^{2|\boldsymbol{j}|(1-1/\bar{p})}\left(\sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}}|\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^{\bar{p}}\right)^{2/\bar{p}}$$

$$\ll \sum_{\mu=0}^{r}\sum_{n_\mu \leq |\boldsymbol{j}|+t/2<n_{\mu+1}} 2^{2|\boldsymbol{j}|(1-1/\bar{p})}2^{2|\boldsymbol{j}|/\bar{p}}\frac{2^{2t}}{N^2}\left(\frac{1}{2^{|\boldsymbol{j}|}} + \frac{(2n_{\mu+1} - t - 2|\boldsymbol{j}|)^{d-1}}{2^{n_{\mu+1}}}\right)^2$$

$$= \frac{2^{2t}}{N^2}\sum_{\mu=0}^{r}\sum_{n_\mu \leq |\boldsymbol{j}|+t/2<n_{\mu+1}} 2^{2|\boldsymbol{j}|}\left(\frac{1}{2^{2|\boldsymbol{j}|}} + 2\frac{(2n_{\mu+1} - t - 2|\boldsymbol{j}|)^{d-1}}{2^{|\boldsymbol{j}|+n_{\mu+1}}}\right.$$

$$\left. + \frac{(2n_{\mu+1} - t - 2|\boldsymbol{j}|)^{2d-2}}{2^{2n_{\mu+1}}}\right)$$

$$= \frac{2^{2t}}{N^2} \sum_{\mu=0}^{r} \left(\sum_{n_\mu \leq |\boldsymbol{j}|+t/2 < n_{\mu+1}} 1 + \frac{2}{2^{n_{\mu+1}}} \sum_{n_\mu < |\boldsymbol{j}|+t/2 \leq n_{\mu+1}} 2^{|\boldsymbol{j}|}(2n_{\mu+1}-t-2|\boldsymbol{j}|)^{d-1} \right.$$

$$\left. + \frac{1}{2^{2n_{\mu+1}}} \sum_{n_\mu \leq |\boldsymbol{j}|+t/2 < n_{\mu+1}} 2^{2|\boldsymbol{j}|}(2n_{\mu+1}-t-2|\boldsymbol{j}|)^{2d-2} \right)$$

$$\ll \frac{2^{2t}}{N^2} \sum_{\mu=0}^{r} \left((\log N)^{d-1}(n_{\mu+1}-n_\mu) + \frac{2^{n_{\mu+1}-t/2}}{2^{n_{\mu+1}}}(\log N)^{d-1} \right.$$

$$\left. + \frac{2^{2n_{\mu+1}-t}}{2^{2n_{\mu+1}}}(\log N)^{d-1} \right)$$

$$\leq \frac{2^{2t}}{N^2} \left((\log N)^d + 2(\log N)^{d-1}(r+1) \right)$$

(17)

$$\ll 2^{2t} \frac{(\log N)^d}{N^2}.$$

Combining (15) and (17) we obtain

$$(L_{p,N}(\mathcal{S}_d))^2 \ll_{p,d} 2^{2t} \frac{(\log N)^d}{N^2}.$$

Now the result follows by taking square roots. ∎

References

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

[2] R. Béjian, *Minoration de la discrépance d'une suite quelconque sur T*, Acta Arith. **41** (1982), 185–202.

[3] D. Bilyk, M. T. Lacey, I. Parissis and A. Vagharshakyan, *Exponential squared integrability of the discrepancy function in two dimensions*, Mathematika **55** (2009), 2470–2502.

[4] D. Bilyk, M. T. Lacey and A. Vagharshakyan, *On the small ball inequality in all dimensions*, J. Funct. Anal. **254** (2008), 2470–2502.

[5] D. Bilyk and L. Markhasin, *BMO and exponential Orlicz space estimates of the discrepancy function in arbitrary dimension*, J. d'Analyse Mathématique, to appear. arXiv:1411.5794.

[6] W. W. L. Chen, *On irregularities of distribution*, Mathematika **27** (1981), 153–170.

[7] W. W. L. Chen, *On irregularities of distribution II*, Quart. J. Math. Oxford **34** (1983), 257–279.

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

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

[10] J. Dick, *Explicit constructions of quasi-Monte Carlo rules for the numerical integration of high-dimensional periodic functions*, SIAM J. Numer. Anal. **45** (2007), 2141–2176.

[11] J. Dick, *Walsh spaces containing smooth functions and quasi-Monte Carlo rules of arbitrary high order*, SIAM J. Numer. Anal. **46** (2008), 1519–1553.

[12] J. Dick, *On Quasi-Monte Carlo rules achieving higher order convergence*, in *Monte Carlo and Quasi-Monte Carlo Methods 2008* (P. L'Ecuyer and A. Owen, eds.), Springer, Berlin, 2009, pp. 73–96.

[13] J. Dick, *Discrepancy bounds for infinite-dimensional order two digital sequences over* $\mathbb{F}_2$, J. Number Theory **136** (2014), 204–232.

[14] J. Dick and J. Baldeaux, *Equidistribution properties of generalized nets and sequences*, in *Monte Carlo and Quasi-Monte Carlo Methods 2008* (P. L'Ecuyer and A. Owen, eds.), Springer, Berlin, 2009, pp. 305–323.

[15] J. Dick, A. Hinrichs, L. Markhasin and F. Pillichshammer, *Discrepancy bounds for higher order digital sequences in function spaces with dominating mixed smoothness*, Mathematika, to appear. arXiv:1604.08713.

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

[17] J. Dick and F. Pillichshammer, *Digital Nets and Sequences. Discrepancy Theory and Quasi-Monte Carlo Integration*, Cambridge University Press, Cambridge, 2010.

[18] J. Dick and F. Pillichshammer, *Optimal $\mathcal{L}_2$ discrepancy bounds for higher order digital sequences over the finite field* $\mathbb{F}_2$, Acta Arith. **162** (2014), 65–99.

[19] J. Dick and F. Pillichshammer, *Explicit constructions of point sets and sequences with low discrepancy*, in *Uniform Distribution and Quasi-Monte Carlo Methods — Discrepancy, Integration and Applications* (P. Kritzer, H. Niederreiter, F. Pillichshammer and A. Winterhof, eds.), Radon Series of Computational and Applied Mathematics, De Gruyter, Berlin, 2014, pp. 63–86.

[20] N. M. Dobrovol'skiĭ, *An effective proof of Roth's theorem on quadratic dispersion* (Russian), Uspekhi Mat. Nauk **39** (1984), 155–156; English translation: Russian Math. Surv. **39** (1984), 117–118.

[21] M. Drmota and R.F. Tichy, *Sequences, Discrepancies and Applications*, Lecture Notes in Mathematics **1651**, Springer-Verlag, Berlin, 1997.

[22] H. Faure, P. Kritzer and F. Pillichshammer, *From van der Corput to modern constructions of sequences for quasi-Monte Carlo rules*, Indag. Math. (N.S.) **26** (2015), 760–822.

[23] K. K. Frolov, *Upper bound of the discrepancy in metric L_p, $2 \leq p < \infty$*, Dokl. Akad. Nauk SSSR **252** (1980), 805–807.

[24] G. Halász, *On Roth's method in the theory of irregularities of point distributions*, in *Recent Progress in Analytic Number Theory*, Vol. 2, Academic Press, London–New York, 1981, pp. 79–94.

[25] F. J. Hickernell and R.-X. Yue, *The mean square discrepancy of scrambled (t, s)-sequences*, SIAM J. Numer. Anal. **38** (2000), 1089–1112.

[26] A. Hinrichs, *Discrepancy of Hammersley points in Besov spaces of dominating mixed smoothness*, Math. Nachr. **283** (2010), 478–488.

[27] E. Hlawka, *Über die Diskrepanz mehrdimensionaler Folgen* mod 1, Math. Z. **77** (1961), 273–284.

[28] R. Kritzinger, L_p- and $S_{p,q}^r B$-discrepancy of the symmetrized van der Corput sequence and modified Hammersley point sets in arbitrary bases, J. Complexity **33** (2016), 145–168.

[29] R. Kritzinger and F. Pillichshammer, L_p-discrepancy of the symmetrized van der Corput sequence, Arch. Math. **104** (2015), 407–418.

[30] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, John Wiley, New York, 1974; Reprint, Dover Publications, Mineola, NY, 2006.

[31] G. Larcher, *On the star discrepancy of sequences in the unit interval*, J. Complexity **31** (2015), 474–485.

[32] G. Leobacher and F. Pillichshammer, *Introduction to Quasi-Monte Carlo Integration and Applications*, Compact Textbooks in Mathematics, Birkhäuser/Springer, Cham, 2014.

[33] L. Markhasin, *Discrepancy of generalized Hammersley type point sets in Besov spaces with dominating mixed smoothness*, Unif. Distrib. Theory **8** (2013), 135–164.

[34] L. Markhasin, *Quasi-Monte Carlo methods for integration of functions with dominating mixed smoothness in arbitrary dimension*, J. Complexity **29** (2013), 370–388.

[35] L. Markhasin, *Discrepancy and integration in function spaces with dominating mixed smoothness*, Dissertationes Math. **494** (2013), 1–81.

[36] L. Markhasin, L_p- and $S_{p,q}^r B$-discrepancy of (order 2) digital nets, Acta Arith. **168** (2015), 139–159.

[37] J. Matoušek, *Geometric Discrepancy. An Illustrated Guide*, Algorithms and Combinatorics, 18, Springer-Verlag, Berlin, 1999.

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

[39] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, No. 63 in CBMS-NSF Series in Applied Mathematics, SIAM, Philadelphia, 1992.

[40] H. Niederreiter and C.P. Xing, *Low-discrepancy sequences and global function fields with many rational places*, Finite Fields Appl. **2** (1996), 241–273.

[41] E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems. Volume I: Linear Information*, European Mathematical Society Publishing House, Zürich, 2008.

[42] P. D. Proinov, *On irregularities of distribution*, C. R. Acad. Bulgare Sci. **39** (1986), 31–34.

[43] K. F. Roth, *On irregularities of distribution*, Mathematika **1** (1954), 73–79.

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

[45] W. M. Schmidt, *Irregularities of distribution. VII*, Acta Arith. **21** (1972), 45–50.

[46] W. M. Schmidt, *Irregularities of Distribution. X*, Number Theory and Algebra, Academic Press, New York, 1977, pp. 311–329.

[47] M. M. Skriganov, *Lattices in algebraic number fields and uniform distribution mod* 1 (Russian), Algebra i Analiz **1** (1989), 207–228; English translation: Leningrad Math. J. **1** (1990), 535–558.

[48] M. M. Skriganov, *Constructions of uniform distributions in terms of geometry of numbers* (Russian), Algebra i Analiz **6** (1994), 200–230; English translation: St. Petersburg Math. J. **6** (1995), 635–664.

[49] M. M. Skriganov, *Harmonic analysis on totally disconnected groups and irregularities of point distributions*, J. Reine Angew. Math. **600** (2006), 25–49.

[50] I. M. Sobol', *The distribution of points in a cube and the approximate evaluation of integrals*, Zh. Vychisl. Mat. i Mat. Fiz. **7** (1967), 784–802.

[51] S. Tezuka, *Polynomial arithmetic analogue of Halton sequences*, ACM Trans. Model. Computer Simul. **3** (1993), 99–107.

[52] H. Triebel, *Bases in Function Spaces, Sampling, Discrepancy, Numerical Integration*, European Mathematical Society Publishing House, Zürich, 2010.

[53] H. Triebel, *Numerical integration and discrepancy. A new approach*, Math. Nachr. **283** (2010), 139–159.

[54] H. Weyl, *Über die Gleichverteilung mod. Eins*, Math. Ann. **77** (1916), 313–352.

[55] S. C. Zaremba, *Some applications of multidimensional integration by parts*, Ann. Polon. Math. **21** (1968), 85–96.