BMO AND EXPONENTIAL ORLICZ SPACE ESTIMATES OF THE DISCREPANCY FUNCTION IN ARBITRARY DIMENSION

By

DMITRIY BILYK[1] AND LEV MARKHASIN

Abstract. In the current paper, we obtain discrepancy estimates in exponential Orlicz and BMO spaces in arbitrary dimension $d \geq 3$. In particular, we use dyadic harmonic analysis to prove that the dyadic product BMO and $\exp(L^{2/(d-1)})$ norms of the discrepancy function of so-called digital nets of order two are bounded above by $(\log N)^{(d-1)/2}$. The latter bound has been recently conjectured in several papers and is consistent with the best known low-discrepancy constructions. Such estimates play an important role as an intermediate step between the well-understood L_p bounds and the notorious open problem of finding the precise L_∞ asymptotics of the discrepancy function in higher dimensions, which is still elusive.

1 Introduction and results

1.1 Definitions. The main object of the present paper is the discrepancy function. For a positive integer N, let $\mathcal{P}_N$ be a point set in the unit interval $[0, 1)^d$ with N points. The **discrepancy function** is defined as

$$D_{\mathcal{P}_N}(x) = \sum_{z \in \mathcal{P}_N} \chi_{[0,x)}(z) - N x_1 \cdots x_d,$$

where $x = (x_1, \ldots, x_d) \in [0, 1)^d$ and $[0, x) = [0, x_1) \times \ldots \times [0, x_d)$. By χ_A we denote the characteristic function of a set $A \in \mathbb{R}^d$, so the term $C_{\mathcal{P}_N}(x) = \sum_z \chi_{[0,x)}(z)$ is equal to the number of points of $\mathcal{P}_N$ in the interval $[0, x)$. Hence, $D_{\mathcal{P}_N}$ measures the deviation of the number of points of $\mathcal{P}_N$ in $[0, x)$ from the fair number of points $L_N(x) = N|[0, x)| = N x_1 \cdots x_d$, which would be achieved by a (practically impossible) perfectly uniform distribution of points, thus quantifying the extent of equidistribution of the point set $\mathcal{P}_N$ and its quality for numerical integration (quasi-Monte Carlo methods; see, e.g., [16]).

Asymptotic behavior of the discrepancy function in $L_p([0, 1)^d)$-spaces for $1 < p < \infty$ is well understood. The classical lower bound proved by Roth [35] for

[1]The first author was supported by the NSF grant DMS-1260516.

JOURNAL D'ANALYSE MATHÉMATIQUE, Vol. 135 (2018)
DOI 10.1007/s11854-018-0035-x

$p = 2$ and by Schmidt [38] for arbitrary $1 < p < \infty$ states that there exists a constant $c = c(p, d) > 0$ such that for every positive integer N and all point sets $\mathcal{P}_N$ in $[0, 1)^d$ with N points,

$$\left\| D_{\mathcal{P}_N} | L_p([0, 1)^d) \right\| \geq c \, (\log N)^{(d-1)/2} . \tag{1}$$

The best known value for c in L_2 can be found in [24]. Furthermore, these estimates are known to be sharp, i.e., there exists a constant $C = C(p, d) > 0$ such that for every positive integer N, there is a point set $\mathcal{P}_N$ in $[0, 1)^d$ with N points such that

$$\left\| D_{\mathcal{P}_N} | L_p([0, 1)^d) \right\| \leq C \, (\log N)^{(d-1)/2} . \tag{2}$$

This was proved by Davenport [12] for $p = 2, d = 2$, by Roth [36] for $p = 2$ and arbitrary d, and finally by Chen [9] in the general case. The best known value for C in L_2 can be found in [16] and [19].

The precise asymptotics of the $L_\infty([0, 1)^d)$-norm of the discrepancy function (star-discrepancy) is known as *the great open problem in discrepancy theory* [2]. The best currently known lower bound in dimensions $d \geq 3$ was obtained quite recently [6]. There exists a constant $c = c(d) > 0$ such that for every positive integer N and all point sets $\mathcal{P}_N$ in $[0, 1)^d$ with N points,

$$\left\| D_{\mathcal{P}_N} | L_\infty([0, 1)^d) \right\| \geq c \, (\log N)^{(d-1)/2+\eta_d} ,$$

where $0 < \eta_d < 1/2$. At the same time, the bound in the plane is well known ([37]):

$$\left\| D_{\mathcal{P}_N} | L_\infty([0, 1)^2) \right\| \geq c \log N. \tag{3}$$

Furthermore (e.g., [21]), there exists a constant $C > 0$ such that for every positive integer N, there is a point set $\mathcal{P}_N$ in $[0, 1)^d$ with N points such that

$$\left\| D_{\mathcal{P}_N} | L_\infty([0, 1)^d) \right\| \leq C \, (\log N)^{d-1} . \tag{4}$$

One can observe a gap between the known upper and lower bounds for the star discrepancy in dimensions $d \geq 3$. There is no agreement among the experts as to what should be the correct asymptotics in higher dimension, the two main conjectures being $(\log N)^{d-1}$ and $(\log N)^{d/2}$. We refer the reader to, e.g., [4] for a more detailed discussion.

1.2 Main results. Since the precise behavior of discrepancy in L_p-spaces ($1 < p < \infty$) is known while the L_∞ estimates remain elusive, it is natural and instructive to investigate what happens in intermediate spaces "close" to L_∞. Standard examples of such spaces are the exponential Orlicz spaces and (various versions of) **BMO**, which stands for bounded mean oscillation. In harmonic analysis, these spaces often play a role of a natural substitute for L_∞ as an endpoint of the L_p scale. We refer the reader to the next section for precise definitions and references.

This approach was initiated in [5] in the case of dimension $d = 2$. Examples used to prove upper bounds in the two-dimensional case were constructed as modifications of the celebrated Van der Corput set. In higher dimensions, we resort to the higher-order digital nets—a concept introduced by Dick [13], [14] and studied from the relevant point of view in [17], [15], and [30]. In particular, we rely strongly on the estimates of the Haar coefficients of the discrepancy function for such nets (see Lemma 3.1) recently obtained by the second author [30].

The first main result of this work is the upper bound in dyadic product BMO.

Theorem 1.1. *For each dimension $d \geq 3$, there exists a constant $C = C(d) > 0$ such that for every positive integer N, there is a point set $\mathcal{P}_N$ in $[0, 1)^d$ with N points such that*

$$\left\| D_{\mathcal{P}_N} \,|\, \mathrm{BMO}^d \right\| \leq C (\log N)^{(d-1)/2} . \tag{5}$$

This result is known in the plane (see [5, Theorem 1.7]); moreover, it is sharp. A simple modification of the proof of (1) yields the corresponding lower bound.

Theorem 1.2. *For each dimension $d \geq 3$, there exists a constant $c = c(d) > 0$ such that for every positive integer N and all point sets $\mathcal{P}_N$ in $[0, 1)^d$ with N points,*

$$\left\| D_{\mathcal{P}_N} \,|\, \mathrm{BMO}^d \right\| \geq c (\log N)^{(d-1)/2} . \tag{6}$$

These results say that in the case of discrepancy function, the BMO norm behaves more like L_p than like L_∞.

Furthermore, we extend the result of [5, Theorem 1.4] on exponential Orlicz spaces to the case of arbitrary dimension. The main theorem we prove in this direction is the following.

Theorem 1.3. *For each dimension $d \geq 3$, there exists a constant $C = C(d) > 0$ such that for every positive integer N, there is a point set $\mathcal{P}_N$ in $[0, 1)^d$ with N points for which*

$$\left\| D_{\mathcal{P}_N} \,\middle|\, \exp\left(L^{\frac{2}{d-1}}\right) \right\| \leq C (\log N)^{\frac{d-1}{2}} . \tag{7}$$

Some remarks are in order for this theorem. This bound has been recently conjectured in several different sources. A similar (albeit weaker) estimate has been recently proved for the so-called Chen–Skriganov nets, independently in [40] and [1], for the smaller $\exp\left(L^{2/(d+1)}\right)$ norm. The authors of both papers conjectured that it should be improved to the $\exp\left(L^{2/(d-1)}\right)$ estimate stated above. In addition, the same conjecture has been made in the survey paper [18, Section 9]. However, until now this claim remained unproved.

The exponential integrability exponent $2/(d-1)$ is quite natural for a variety of reasons. First, it is consistent with the general ideology that the problem effectively has $d-1$ "free parameters" (see [4] for a detailed discussion), and therefore the Littlewood–Paley inequalities should be applied $d-1$ times: see Section 2.2, in particular, estimate (12) of Lemma 2.1. Furthermore, this estimate is consistent with the L_∞-discrepancy bound (4) of the order $(\log N)^{d-1}$, valid for digital nets; see Section 3.5.

For this very reason, the complementary lower bound is presently beyond reach. There are indications that it should be almost as difficult as one of the main open problems in the subject—the lower bound of the L_∞-discrepancy. The proof of the corresponding lower bound in dimension $d = 2$ ([5, Theorem 1.4]) uses techniques similar to the proof of the two-dimensional L_∞ bound (3), which are not available in higher dimensions. Besides, if one believes that the correct L_∞ bound is $(\log N)^{d/2}$, then estimate (7) is probably not sharp—in this case, the norm on the left-hand side should be the subgaussian $\exp(L^2)$; see Subsection 3.5 for details.

During the final stages of preparation of the present manuscript, we learned about a recent preprint of Skriganov [41] written almost simultaneously, where inequality (7) is proved for *random* digit shifts of an arbitrary digital (t, n, d)-net (although the author does not state the result in exponential form, but instead writes down L_p estimates with explicit dependence on p). The techniques of Skriganov's work exploit randomness in a crucial way. In contrast, our proof is deterministic and is applicable to any higher order digital net (in fact, it suffices to take order $\sigma = 2$). Concrete construction of such nets are given, e.g., in [15].

Interpolating the estimate of Theorem 1.3 with the well-known L_∞ bound (4), we obtain the following result, which is a direct analog of [5, Theorem 1.4].

Corollary 1.4. *For each β satisfying $2/(d-1) \leq \beta < \infty$, there exists a constant $C_\beta > 0$ such that for every positive integer N, there is a point set $\mathcal{P}_N$ in $[0,1)^d$ with N points such that*

$$\left\| D_{\mathcal{P}_N} \,|\, \exp(L^\beta) \right\| \leq C_\beta \, (\log N)^{(d-1)-\frac{1}{\beta}} . \tag{8}$$

Since this result is even more closely tied to the L_∞ estimates, no corresponding lower bounds are available.

Our strategy resonates with that of [5], but we also rely strongly on very recent results and constructions: digital nets of higher order [13], [14] and their explicit constructions [15], [18], Haar expansions of the discrepancy function of such nets used in the study of discrepancy in Besov spaces with dominating mixed smoothness and in L_2; see [17], [18], [15], [30]. For further results on this topic, see [10], [11], [22], [27], [28], [39], [43]. As general references for studies of the discrepancy function, we refer to the monographs [2], [16], [25], [31], [33] and surveys [4], [23], [29].

We write $A \preceq B$ if there exists an absolute constant $c > 0$ such that $A \leq cB$. We write $A \simeq B$ if $A \preceq B$ and $B \preceq A$. The implicit constants in this paper do not depend on the number of points N (but may depend on some other parameters, such as dimension, integrability index etc).

2 Preliminary facts

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

For $j \in \mathbb{N}_0$ and $m \in \mathbb{D}_j$, we call the interval $I_{j,m} = \big[2^{-j}m, 2^{-j}(m+1)\big)$ 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 $j \in \mathbb{N}_{-1}^d$ and $m = (m_1, \ldots, m_d) \in \mathbb{D}_j$, we call $I_{j,m} = I_{j_1,m_1} \times \ldots \times I_{j_d,m_d}$ the **m-th dyadic interval in** $[0,1)^d$ **at level** j. We call the number $|j|$ the **order** of the dyadic interval $I_{j,m}$. Its volume is then $|I_{j,m}| = 2^{-|j|}$.

An important combinatorial fact is that $\#\{j \in \mathbb{N}_0^d : |j| = n\} \simeq n^{d-1}$, where $\#S$ stands for the cardinality of the set S.

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 constant values 1 on $I_{j,m}^+$ and -1 on $I_{j,m}^-$. We put $h_{-1,0} = \chi_{I_{-1,0}}$ on $[0,1)$. Notice that we normalize the Haar functions in L_∞, rather than L_2.

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

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

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

It is well known that the system

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

is an orthonormal basis in $L_2([0,1)^d)$, an unconditional basis in $L_p([0,1)^d)$ for $1 < p < \infty$, and a conditional basis in $L_1([0,1)^d)$.

2.2 Littlewood–Paley inequalities. For a function $f \in L_2([0,1)^d)$, we have Parseval's identity

$$\left\| f|L_2([0,1)^d)\right\|^2 = \sum_{j\in\mathbb{N}_{-1}^d} 2^{|j|} \sum_{m\in\mathbb{D}_j} |\langle f, h_{j,m}\rangle|^2. \tag{9}$$

Littlewood–Paley inequalities are a generalization of this statement to L_p-spaces. For a function $f : [0,1]^d \to \mathbb{R}$, the (dyadic) **Littlewood–Paley square function** is defined as

$$Sf(x) = \left(\sum_{j\in\mathbb{N}_{-1}^d} 2^{2|j|} \sum_{m\in\mathbb{D}_j} |\langle f, h_{j,m}\rangle|^2 \chi_{I_{j,m}} \right)^{1/2}.$$

It is a classical fact and a natural extension of (9) that in dimension $d = 1$, the L_p-norm of f can be characterized using the square function; i.e., for each $1 < p < \infty$, there exist constants $A_p, B_p > 0$ such that

$$A_p \|Sf|L_p([0,1))\| \le \|f|L_p([0,1))\| \le B_p \|Sf|L_p([0,1))\|. \tag{10}$$

Two remarks are important. First, it is well known that $B_p \simeq \sqrt{p}$. Second, estimates (10) continue to hold for Hilbert space-valued functions f. This allows one to extend the inequalities to the case of multivariate functions $f : [0,1]^d \to \mathbb{R}$ by iterating the one-dimensional estimates d times, thus picking up constants A_p^d and $B_p^d \simeq p^{d/2}$.

However, if the function f is represented by a **hyperbolic sum** of Haar wavelets, i.e., a sum of Haar functions supported by intervals of fixed order, $f \in \operatorname{span}\{h_{j,m} : |j| = n\}$; in other words, when the number of "free parameters" is $d-1$, then the one-dimensional Littlewood–Paley inequalities (10) need only be applied $d-1$ times, yielding constants A_p^{d-1} and $B_p^{d-1} \simeq p^{(d-1)/2}$. We summarize the estimates useful for our purposes in the following lemma.

Lemma 2.1. *Let* $1 < p < \infty$.

(i) **Multiparameter Littlewood–Paley inequality**: *For each function* $f : [0,1]^d \to \mathbb{R}$,

$$\left\| f \,|\, L_p([0,1)^d)\right\| \preceq p^{d/2} \left\| Sf|L_p([0,1)^d)\right\|. \tag{11}$$

(ii) **Hyperbolic Littlewood–Paley inequality**: *Assume that the function* $f : [0,1]^d \to \mathbb{R}$ *is a hyperbolic sum of Haar functions, i.e.,*

$$f \in \operatorname{span}\{h_{j,m} : |j| = n\}$$

for some $n \in \mathbb{N}$. *Then*

$$\left\| f \,|\, L_p([0,1)^d) \right\| \preceq p^{\frac{d-1}{2}} \left\| Sf \,|\, L_p([0,1)^d) \right\|. \tag{12}$$

A more detailed discussion of the Littlewood–Paley inequalities and their applications in discrepancy theory can be found in [4].

2.3 Bounded mean oscillation and exponential Orlicz spaces. There are different definitions of the space of functions of bounded mean oscillation in the multivariate case. The appropriate version in our setting is the so-called dyadic product BMO^d introduced in [3]. For an integrable function $f : [0,1]^d \to \mathbb{R}$, we define

$$\left\| f \,|\, \mathrm{BMO}^d \right\| = \sup_{U \subset [0,1)^d} \left(|U|^{-1} \sum_{j \in \mathbb{N}_0^d} 2^{|j|} \sum_{\substack{m \in \mathbb{D}_j \\ I_{j,m} \subset U}} |\langle f, h_{j,m} \rangle|^2 \right)^{1/2}, \tag{13}$$

where the supremum is taken over all measurable sets $U \subset [0,1)^d$. The space BMO^d contains all integrable functions f with finite norm $\| f \,|\, \mathrm{BMO}^d \|$. Notice that, technically, $\| f \,|\, \mathrm{BMO}^d \|$ is only a seminorm, since it vanishes on linear combinations of functions that are constant in some of the coordinate directions. Therefore, formally, we need to take a factor space over such functions.

To give some intuition behind this definition, we notice that when $d = 1$ and U is a dyadic interval, we have by Parseval's identity

$$|U|^{-1} \sum_{j \in \mathbb{N}_0} 2^{|j|} \sum_{\substack{m \in \mathbb{D}_j \\ I_{j,m} \subset U}} |\langle f, h_{j,m} \rangle|^2 = |U|^{-1} \int_U \left| f - \langle f \rangle_U \right|^2 dx,$$

where $\langle f \rangle_U$ is the mean of f over U—this is precisely the expression which arises in the definition of the one-dimensional dyadic BMO space. The precise technical definition of the norm (13) turns out to be the correct multiparameter dyadic extension which preserves the most natural properties of BMO, in particular, the celebrated $H_1 - \mathrm{BMO}$ duality: **dyadic product BMO** is the dual of the **dyadic Hardy space** H_1—the space of functions $f \in L_1$ with integrable Littlewood–Paley square function, i.e. such that $Sf \in L_1$; see [3].

We remark that a non-dyadic version of this space, the Chang–Fefferman product BMO, was introduced and studied in [7]. This space also admits a characterization similar to (13), but with smoother functions in place of Haar wavelets. For the relation between these spaces see e.g. [34].

In order to introduce the definition of the exponential Orlicz spaces, we start by briefly discussing general Orlicz spaces. We refer to [26] for more information. Let (Ω, P) be a probability space, and denote by $\mathbb{E}$ the expectation over (Ω, P). Let $\psi : [0,\infty) \to [0,\infty)$ be a convex function such that $\psi(x) = 0$ if and only if $x = 0$. For a (Ω, P)-measurable real valued function f, we define

$$\left\| f | L^{\psi} \right\| = \inf\{K > 0 : \mathbb{E}\psi(|f|/K) \le 1\},$$

where $\inf \varnothing = \infty$. The **Orlicz space (associated with ψ)** L^{ψ} consists of all functions f with finite norm $\|f|L^{\psi}\|$.

Let $\alpha > 0$, and let ψ_α be a convex function which equals $\mathrm{e}^{x^\alpha} - 1$ for x sufficiently large, depending upon α (for $\alpha \ge 1$, this function may be used for all $x \ge 0$). We write $\exp(L^\alpha) = L^{\psi_\alpha}$.

The following proposition yields a standard way to compute the $\exp(L^\alpha)$ norms. Its proof is a simple application of Taylor's series for e^x and Stirling's formula.

Proposition 2.2. *For every $\alpha > 0$,*

$$\left\| f | \exp(L^\alpha) \right\| \simeq \sup_{p>1} p^{-1/\alpha} \cdot \left\| f | L_p([0,1)^d) \right\| . \tag{14}$$

The next proposition is a variant of the famous Chang–Wilson–Wolff inequality [8], which states that boundedness of the square function implies certain exponential integrability of the original function. The hyperbolic version presented here can be deduced easily from the Littlewood–Paley inequality with sharp constants (12) and the previous proposition.

Proposition 2.3 (Hyperbolic Chang–Wilson–Wolff inequality)**.** *Assume that f is a hyperbolic sum of multiparameter Haar functions, i.e.,*

$$f \in \operatorname{span}\{h_{j,m} : |j| = n\}$$

for some $n \in \mathbb{N}$. Then

$$\left\| f | \exp\left(L^{2/(d-1)}\right) \right\| \preceq \left\| S(f) | L_\infty([0,1)^d) \right\| . \tag{15}$$

Proof. According to (12), we have

$$\|f|L_p([0,1)^d)\| \preceq p^{(d-1)/2} \|Sf|L_p([0,1)^d)\| \le p^{(d-1)/2} \|Sf|L_\infty([0,1)^d)\|.$$

Estimate (15) now follows from (14). □

We note that it is important here that the function f be a linear combination of Haar functions supported by rectangles of fixed volume: without this assumption, the correct norm in the left-hand side would be $\exp\left(L^{2/d}\right)$, which can be deduced from (11).

For all $1 \leq p < \infty$, we have $L_\infty \subset \exp(L^\alpha) \subset L_p$. Furthermore, it is obvious that $\left\|f \,|\exp(L^\alpha)\right\| \preceq \left\|f\,|\exp(L^\beta)\right\|$, i.e., $\exp(L^\beta) \subset \exp(L^\alpha)$, for $\alpha < \beta$. The next lemma shows that under the assumption that $f \in L_\infty$, the relation may be reversed. The argument is a simple interpolation between exponential Orlicz spaces and L_∞.

Proposition 2.4. *Let* $0 < \alpha < \beta < \infty$. *If a function* $f \in L_\infty([0,1)^d)$ *also satisifies* $f \in \exp(L^\alpha)$, *then* $f \in \exp(L^\beta)$ *and*

$$\left\|f\,|\exp(L^\beta)\right\| \preceq \left\|f\,|\exp(L^\alpha)\right\|^{\alpha/\beta} \cdot \left\|f|L_\infty([0,1)^d)\right\|^{1-\alpha/\beta}.$$

Proof. Set $q = \frac{\alpha}{\beta}p$. Then

$$\|f|L_p([0,1)^d)\| \leq \|f|L_q([0,1)^d)\|^{\alpha/\beta} \cdot \|f|L_\infty([0,1)^d)\|^{1-\alpha/\beta}$$

and

$$\begin{aligned}\left\|f\,|\exp(L^\beta)\right\| &\simeq \sup_{p>1} p^{-1/\beta} \cdot \|f|L_p([0,1)^d)\| \\ &\leq \sup_{p>1} p^{-1/\beta} \cdot \|f|L_q([0,1)^d)\|^{\alpha/\beta} \cdot \|f|L_\infty([0,1)^d)\|^{1-\alpha/\beta} \\ &\preceq \sup_{q>1} \left(q^{-1/\alpha} \cdot \|f|L_q([0,1)^d)\|\right)^{\alpha/\beta} \cdot \|f|L_\infty([0,1)^d)\|^{1-\alpha/\beta},\end{aligned}$$

which finishes the proof. □

2.4 Digital nets. Our next step is to define digital (t,n,d)-nets of order $\sigma \geq 1$. The original definition of digital nets goes back to Niederreiter [32], and the first constructions were given even earlier by Sobol' [42]. The concept of higher-order digital nets was introduced in [13], [14]. We quote the definitions from [13] and [14, Definitions 4.1, 4.3]. In the case of order $\sigma = 1$, we recover the original definition of digital nets.

For $\nu \in \{0,1,\ldots,2^n-1\}$ with the binary expansion $\nu = \nu_0+\nu_1 2+\cdots+\nu_{n-1}2^{n-1}$ with digits $\nu_0,\nu_1,\ldots,\nu_{n-1} \in \{0,1\}$, the binary digit vector $\overline{\nu}$ is given as $\overline{\nu} = (\nu_0,\nu_1,\ldots,\nu_{n-1})^\top \in \mathbb{F}_2^n$. Then, for $\sigma \in \mathbb{N}$, let $C_1,\ldots,C_d$ be $\sigma n \times n$ matrices over $\mathbb{F}_2$. We compute $C_i\overline{\nu} = (x_{i,\nu,1}, x_{i,\nu,2},\ldots,x_{i,\nu,\sigma n})^\top \in \mathbb{F}_2^{\sigma n}$ for $1 \leq i \leq d$. Finally, we define $x_{i,\nu} = x_{i,\nu,1}2^{-1}+x_{i,\nu,2}2^{-2}+\cdots+x_{i,\nu,\sigma n}2^{-\sigma n} \in [0,1)$ and $x_\nu = (x_{1,\nu},\ldots,x_{d,\nu})$. We call the point set $\mathcal{P}_{2^n} = \{x_0,x_1,\ldots,x_{2^n-1}\}$ a **digital net** (over $\mathbb{F}_2$).

Now let $0 \le t \le \sigma n$ be an integer. For every $1 \le i \le d$, we write $C_i = (c_{i,1}, \ldots, c_{i,\sigma n})^\top$, where $c_{i,1}, \ldots, c_{i,\sigma n} \in \mathbb{F}_2^n$ are the row vectors of C_i. If for all $1 \le \lambda_{i,1} < \ldots < \lambda_{i,\eta_i} \le \sigma n$, $1 \le i \le d$ with

$$\lambda_{1,1} + \cdots + \lambda_{1,\min(\eta_1,\sigma)} + \cdots + \lambda_{d,1} + \cdots + \lambda_{d,\min(\eta_d,\sigma)} \le \sigma n - t$$

the vectors $c_{1,\lambda_{1,1}}, \ldots, c_{1,\lambda_{1,\eta_1}}, \ldots, c_{d,\lambda_{d,1}}, \ldots, c_{d,\lambda_{d,\eta_d}}$ are linearly independent over $\mathbb{F}_2$, then $\mathcal{P}_{2^n}$ is called an **order σ digital (t, n, d)-net** (over $\mathbb{F}_2$).

The smaller the quality parameter t and the greater the order σ, the better structure the point set has. In particular, every point set $\mathcal{P}_{2^n}$ constructed with the digital method is at least an order σ digital $(\sigma n, n, d)$-net. Every order σ_2 digital (t, n, d)-net is an order σ_1 digital $(\lceil t\sigma_1/\sigma_2 \rceil, n, d)$-net if $1 \le \sigma_1 \le \sigma_2$; see [13]. It is well known that digital (t, n, d)-nets are perfectly distributed with respect to dyadic intervals (in the standard terminology (see, e.g., [16]), order 1 digital (t, n, d)-nets are (t, n, d)-nets): *every dyadic interval of order $n - t$ contains exactly 2^t points of the (t, n, d)-net.* A version of this property continues to hold for higher-order nets.

Lemma 2.5. *Let $\mathcal{P}_{2^n}$ be an order σ digital (t, n, d)-net. Then every dyadic interval of order n contains at most $2^{\lceil t/\sigma \rceil}$ points of $\mathcal{P}_{2^n}$.*

It is a classical fact that such sets satisfy the best known star discrepancy estimate (4); see [16, Theorem 5.10].

Lemma 2.6. *Let $\mathcal{P}_{2^n}$ be an order σ digital (t, n, d)-net. Then*

$$\left\| D_{\mathcal{P}_{2^n}} | L_\infty([0,1)^d) \right\| \preceq n^{d-1}. \tag{16}$$

Constructions of order σ digital (t_2, n, d)-nets can be obtained via so-called digit interlacing of order 1 digital $(t_1, n, \sigma d)$-nets, and several constructions of order 1 digital nets are known. For details, examples, and further literature, we refer to [18] and [15]. We only point out here that there are constructions with a good quality parameter t which, in particular, does not depend on n.

3 Proofs of the theorems

We prove the main theorems in the case when the number of points is a power of 2, i.e., $N = 2^n$. The reduction to the general case is standard; see, e.g., Subsection 6.3 in [5]. Our examples are the higher-order digital nets described in the previous section with the minimal non-trivial value of the order $\sigma = 2$.

We rely on the recent estimates of the Haar coefficients of the discrepancy function of order 2 digital nets obtained by the second author.

Lemma 3.1 ([30, Lemma 5.9]). *Let $\mathcal{P}_{2^n}$ be an order 2 digital (t, n, d)-net. Let $j \in \mathbb{N}_{-1}^d$ and $m \in \mathbb{D}_j$.*

(i) *If $|j| \geq n - \lceil t/2 \rceil$, then $|\langle D_{\mathcal{P}_{2^n}}, h_{j,m}\rangle| \preceq 2^{-|j|}$ and $|\langle D_{\mathcal{P}_{2^n}}, h_{j,m}\rangle| \preceq 2^{-2|j|+n}$ for all but 2^n values of m.*

(ii) *If $|j| < n - \lceil t/2 \rceil$, then $|\langle D_{\mathcal{P}_{2^n}}, h_{j,m}\rangle| \preceq 2^{-n}(2n - t - 2|j|)^{d-1}$.*

In fact, we mostly need the second part of this lemma, i.e., the Haar coefficients for small values of $|j|$ (in other words, for large intervals). Comparing this estimate to the corresponding two-dimensional bound for the Van der Corput set obtained in [5, Lemma 4.1], which states that $|\langle D_{\mathcal{P}_{2^n}}, h_{j,m}\rangle| \preceq 2^{-n}$, we see that in our case we have an additional logarithmic factor. However, this factor is completely harmless, as one can see from the following elementary computation.

Lemma 3.2. *Let K be a positive integer, $A > 1$, and $q, r > 0$. Then*

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

where the implicit constant is independent of K.

Proof. We have

$$\sum_{k=0}^{K-1} A^k (K-k)^q k^r \leq A^K K^r \sum_{k=0}^{K-1} A^{k-K} (K-k)^q = A^K K^r \sum_{k=1}^{K} A^{-k} k^q \preceq A^K K^r. \qquad \square$$

We now turn to the proofs of the main theorems, which are similar in spirit to the arguments in [5].

3.1 Proof of Theorem 1.1.

Let $\mathcal{P}_{2^n}$ be an order 2 digital (t, n, d)-net with the quality parameter t depending only on the dimension d. We recall that

$$\#\{ j \in \mathbb{N}_0^d : |j| = n\} \simeq n^{d-1} \quad \text{and} \quad \#\mathbb{D}_j = 2^{|j|}.$$

We fix an arbitrary measurable set $U \subset [0,1)^d$. We need to prove

$$|U|^{-1} \sum_{j \in \mathbb{N}_0^d} 2^{|j|} \sum_{\substack{m \in \mathbb{D}_j \\ I_{j,m} \subset U}} |\langle D_{\mathcal{P}_{2^n}}, h_{j,m}\rangle|^2 \preceq n^{d-1}.$$

We split the sum above into three cases: large, intermediate, and small intervals, according to the cases in Lemma 3.1. We observe that, in each case, there are at most $2^{|j|}|U|$ values of $m \in \mathbb{D}_j$ such that $I_{j,m} \subset U$.

Starting with large intervals, we apply (ii) of Lemma 3.1 and Lemma 3.2 to obtain

$$\begin{aligned}|U|^{-1} \sum_{\substack{j\in\mathbb{N}_0^d \\ |j|<n-\lceil t/2\rceil}} 2^{|j|} \sum_{\substack{m\in\mathbb{D}_j \\ I_{j,m}\subset U}} |\langle D_{\mathcal{P}_{2^n}}, h_{j,m}\rangle|^2 \\ &\preceq |U|^{-1} \sum_{\substack{j\in\mathbb{N}_0^d \\ |j|<n-\lceil t/2\rceil}} 2^{|j|}\, 2^{|j|} |U|\, 2^{-2n}\, (2n-t-2|j|)^{2(d-1)} \\ &\preceq 2^{-2n} \sum_{k=0}^{n-\lceil t/2\rceil-1} 2^{2k}\, (2n-t-2k)^{2(d-1)}\, (k+1)^{d-1} \\ &\preceq 2^{-2n}\, 2^{2n}\, (n-t/2)^{d-1} \preceq n^{d-1}.\end{aligned}$$

Next, we consider intermediate intervals and apply (i) of Lemma 3.1 to obtain

$$\begin{aligned}|U|^{-1} \sum_{\substack{j\in\mathbb{N}_0^d \\ n-\lceil t/2\rceil\le|j|<n}} 2^{|j|} \sum_{\substack{m\in\mathbb{D}_j \\ I_{j,m}\subset U}} |\langle D_{\mathcal{P}_{2^n}}, h_{j,m}\rangle|^2 &\preceq |U|^{-1} \sum_{\substack{j\in\mathbb{N}_0^d \\ n-\lceil t/2\rceil\le|j|<n}} 2^{|j|}\, 2^{|j|} |U|\, 2^{-2|j|} \\ &\le \sum_{k=n-\lceil t/2\rceil}^{n-1} (k+1)^{d-1} \preceq n^{d-1}.\end{aligned}$$

We now turn to the case of small intervals, where $|j| \ge n$. These boxes are too small to capture any cancellation, hence we treat the linear and counting parts of the discrepancy function separately.

The case of the linear part $L_{\mathcal{P}_{2^n}}(x) = 2^n x_1 \cdot \ldots \cdot x_d$ is simple. It is easy to verify that $|\langle L_{\mathcal{P}_{2^n}}, h_{j,m}\rangle| \simeq 2^{-2|j|+n}$; thus we obtain

$$\begin{aligned}|U|^{-1} \sum_{\substack{j\in\mathbb{N}_0^d \\ |j|\ge n}} 2^{|j|} \sum_{\substack{m\in\mathbb{D}_j \\ I_{j,m}\subset U}} |\langle L_{\mathcal{P}_{2^n}}, h_{j,m}\rangle|^2 &\preceq |U|^{-1} \sum_{\substack{j\in\mathbb{N}_0^d \\ |j|\ge n}} 2^{|j|}\, 2^{|j|} |U|\, 2^{-4|j|+2n} \\ &\preceq 2^{2n} \sum_{k=n}^{\infty} 2^{-2k}\, (k+1)^{d-1} \preceq n^{d-1}.\end{aligned}$$

Estimating the counting part $C_{\mathcal{P}_{2^n}}$ is a bit more involved. Denote by $\mathcal{J}$ the family of dyadic intervals $I_{j,m} \subset U$ with $|j| \ge n$ such that $\langle C_{\mathcal{P}_{2^n}}, h_{j,m}\rangle \ne 0$. Consider the subfamily $\widetilde{\mathcal{J}} \subset \mathcal{J}$, which consists of *maximal* (with respect to inclusion) dyadic intervals in $\mathcal{J}$. We first demonstrate the following fact, which provides control of the total size of the intervals in this family:

$$\sum_{I_{j,m}\in\widetilde{\mathcal{J}}} |I_{j,m}| = \sum_{I_{j,m}\in\widetilde{\mathcal{J}}} 2^{-|j|} \preceq n^{d-1}|U|. \tag{17}$$

Indeed, consider an interval $I_{j,m} \in \mathfrak{J}$. Since $\langle C_{\mathcal{P}_{2^n}}, h_{I_{j,m}}\rangle \neq 0$, at least one point $z \in \mathcal{P}_{2^n}$ must be contained in the interior of $I_{j,m}$, which in turn means that each side of $I_{j,m}$ has length at least 2^{-2n} (since $\mathcal{P}_{2^n}$ is an order 2 digital net, whose points have binary coordinates of length $2n$), i.e., $0 \le j_k \le 2n$ for $k = 1, \dots, d$.

Fix integer parameters $r_1, \dots, r_{d-1}$ between 0 and $2n$. Consider the family $\widetilde{\mathfrak{J}}_{r_1,\dots,r_{d-1}} \subset \widetilde{\mathfrak{J}}$ consisting of those intervals $I_{j,m} \in \widetilde{\mathfrak{J}}$ for which $j_k = r_k$ for $k = 1, \dots, d-1$, i.e., the lengths of their first $d-1$ sides are fixed. All intervals in this family are disjoint: if two of them intersected, then their first $d-1$ sides would have to coincide, and hence one would have to be contained in the other, which contradicts maximality. Therefore,

$$\sum_{I_{j,m} \in \widetilde{\mathfrak{J}}} 2^{-|j|} = \sum_{r_1,\dots,r_{d-1}=0}^{2n} \sum_{I_{j,m} \in \widetilde{\mathfrak{J}}_{r_1,\dots,r_{d-1}}} |I_{j,m}| \le \sum_{r_1,\dots,r_{d-1}=0}^{2n} |U| \preceq n^{d-1}|U|,$$

which proves (17).

For a dyadic interval J, we define $C^J_{\mathcal{P}_{2^n}}(x) = \sum_{z \in \mathcal{P}_{2^n} \cap J} \chi_{[0,x)}(z)$, i.e. $C^J_{\mathcal{P}_{2^n}}(x)$ is the part of the counting function that counts only the points from J. It is clear that $\langle C_{\mathcal{P}_{2^n}}, h_{j,m}\rangle = \langle C^J_{\mathcal{P}_{2^n}}, h_{j,m}\rangle$ whenever $I_{j,m} \subset J$.

We recall Lemma 2.5, which implies that every dyadic interval of volume at most 2^{-n} contains no more than $2^{\lceil t/2 \rceil}$ points. Therefore, for each interval $J \in \widetilde{\mathfrak{J}}$,

$$\left\| C^J_{\mathcal{P}_{2^n}} \right\|_{L_2(J)} \le \sum_{p \in \mathcal{P}_{2^n} \cap J} \left\| \chi_{[0,\cdot)}(z) \right\|_{L_2(J)} \le 2^{\lceil t/2 \rceil} |J|^{1/2}. \tag{18}$$

Using the orthogonality of Haar functions, Bessel inequality, (18), and (17), we find that

$$\begin{aligned}
|U|^{-1} \sum_{I_{j,m} \in \mathfrak{J}} 2^{|j|} |\langle C_{\mathcal{P}_{2^n}}, h_{j,m}\rangle|^2 &\le |U|^{-1} \sum_{J \in \widetilde{\mathfrak{J}}} \sum_{I_{j,m} \subset J} 2^{|j|} |\langle C^J_{\mathcal{P}_{2^n}}, h_{j,m}\rangle|^2 \\
&\le |U|^{-1} \sum_{J \in \widetilde{\mathfrak{J}}} \left\| C^J_{\mathcal{P}_{2^n}} \right\|^2_{L_2(J)} \\
&\le |U|^{-1} 2^{t+1} \sum_{J \in \widetilde{\mathfrak{J}}} |J| \preceq n^{d-1},
\end{aligned}$$

which concludes the proof for small intervals and therefore proves Theorem 1.1.

3.2 Proof of Theorem 1.2. We now turn to the proof of the matching lower bound for the space BMO^d. The proof is a simple adaptation of the ideas of the original proof [35] of the lower bound for the L_2-discrepancy (1).

Fix an arbitrary point set $\mathcal{P}_N \subset [0,1)^d$ with N points. Choose the scale $n \in \mathbb{N}$ so that $2N \le 2^n < 4N$. This choice guarantees that for each $j \in \mathbb{N}_0^d$ with $|j| = n$, there are at least 2^{n-1} values of $m \in \mathbb{D}_j$ such that $I_{j,m} \cap \mathcal{P}_N = \varnothing$, i.e., at least half of all intervals do not contain any points of $\mathcal{P}_N$. As discussed before, for such empty intervals,

$$|\langle D_{\mathcal{P}_N}, h_{j,m}\rangle| = |\langle L_N, h_{j,m}\rangle| \simeq N 2^{-2|j|} \simeq 2^{-n}.$$

We use the definition of the BMO^d norm (13) and choose the measurable set $U = [0,1)^d$ to obtain

$$\begin{aligned}\left\| D_{\mathcal{P}_N} \,|\, \mathrm{BMO}^d \right\|^2 &\geq \sum_{j \in \mathbb{N}_0^d} 2^{|j|} \sum_{m \in \mathbb{D}_j} |\langle D_{\mathcal{P}_N}, h_{j,m}\rangle|^2 \geq \sum_{\substack{j \in \mathbb{N}_0^d \\ |j|=n}} 2^{|j|} \sum_{\substack{m \in \mathbb{D}_j \\ I_{j,m} \cap \mathcal{P}_N = \emptyset}} |\langle L_N, h_{j,m}\rangle|^2 \\ &\succeq \sum_{\substack{j \in \mathbb{N}_0^d \\ |j|=n}} 2^n \cdot 2^{n-1} \cdot 2^{-2n} \simeq n^{d-1},\end{aligned}$$

which finishes the proof, since $n \simeq \log N$.

3.3 Proof of Theorem 1.3. We now turn our attention to the proof of the upper bound in the Orlicz space $\exp\left(L^{2/(d-1)}\right)$. Once again we consider three different cases, namely, large intervals, intermediate intervals, and small intervals.

We start with the large intervals. Applying the triangle inequality, Chang–Wilson–Wolff inequality (Proposition 2.3), and (ii) of Lemma 3.1, we obtain

$$\begin{aligned}&\left\| \sum_{\substack{j \in \mathbb{N}_{-1}^d \\ |j| < n - \lceil t/2 \rceil}} 2^{|j|} \sum_{m \in \mathbb{D}_j} \langle D_{\mathcal{P}_{2^n}}, h_{j,m}\rangle\, h_{j,m} \,\Big|\, \exp\left(L^{2/(d-1)}\right) \right\| \\ &\quad\leq \sum_{k=0}^{n-\lceil t/2 \rceil} \left\| \sum_{\substack{j \in \mathbb{N}_{-1}^d \\ |j|=k}} 2^{|j|} \sum_{m \in \mathbb{D}_j} \langle D_{\mathcal{P}_{2^n}}, h_{j,m}\rangle\, h_{j,m} \,\Big|\, \exp\left(L^{2/(d-1)}\right) \right\| \\ &\quad\preceq \sum_{k=0}^{n-\lceil t/2 \rceil} \left\| \left(\sum_{\substack{j \in \mathbb{N}_{-1}^d \\ |j|=k}} 2^{2|j|} \sum_{m \in \mathbb{D}_j} |\langle D_{\mathcal{P}_{2^n}}, h_{j,m}\rangle|^2 \chi_{I_{j,m}} \right)^{1/2} \Big|\, L_\infty \right\| \\ &\quad\preceq \sum_{k=0}^{n-\lceil t/2 \rceil} \left\| \left(\sum_{\substack{j \in \mathbb{N}_{-1}^d \\ |j|=k}} 2^{2k}\, 2^{-2n} (2n - t - 2|j|)^{2(d-1)} \sum_{m \in \mathbb{D}_j} \chi_{I_{j,m}} \right)^{1/2} \Big|\, L_\infty \right\| \\ &\quad\preceq 2^{-n} \sum_{k=0}^{n-\lceil t/2 \rceil} \left(2^{2k} (2n - t - 2k)^{2(d-1)} (k+1)^{d-1} \right)^{1/2} \\ &\quad\preceq 2^{-n}\, 2^n n^{(d-1)/2} = n^{(d-1)/2}.\end{aligned}$$

Now we consider the medium sized intervals, applying (i) of Lemma 3.1, obtaining

$$\begin{aligned}
&\Bigg\| \sum_{\substack{j\in\mathbb{N}^d_{-1}\\ n-\lceil t/2\rceil\le|j|<n}} 2^{|j|} \sum_{m\in\mathbb{D}_j} \langle D_{\mathcal{P}_{2^n}}, h_{j,m}\rangle\, h_{j,m} \Big| \exp\left(L^{2/(d-1)}\right)\Bigg\| \\
&\qquad \le \sum_{k=n-\lceil t/2\rceil}^{n-1} \Bigg\| \sum_{j\in\mathbb{N}^d_{-1}:\,|j|=k} 2^{|j|} \sum_{m\in\mathbb{D}_j} \langle D_{\mathcal{P}_{2^n}}, h_{j,m}\rangle\, h_{j,m} \,\Big|\exp\left(L^{2/(d-1)}\right)\Bigg\| \\
&\qquad \le \sum_{k=n-\lceil t/2\rceil}^{n-1} \Bigg\| \Bigg(\sum_{j\in\mathbb{N}^d_{-1}:\,|j|=k} 2^{2k} \sum_{m\in\mathbb{D}_j} |\langle D_{\mathcal{P}_{2^n}}, h_{j,m}\rangle|^2\, \chi_{I_{j,m}} \Bigg)^{1/2} \Big| L_\infty \Bigg\| \\
&\qquad \preceq \sum_{k=n-\lceil t/2\rceil}^{n-1} \left((k+1)^{d-1}\right)^{1/2} \preceq n^{(d-1)/2}.
\end{aligned}$$

In the case of small intervals, we again treat the linear and the counting parts separately. Since $|\langle L_{\mathcal{P}_{2^n}}, h_{j,m}\rangle| \preceq 2^{-2|j|+n}$, we obtain

$$\begin{aligned}
&\Bigg\| \sum_{\substack{j\in\mathbb{N}^d_{-1}\\ |j|\ge n}} 2^{|j|} \sum_{m\in\mathbb{D}_j} \langle L_{\mathcal{P}_{2^n}}, h_{j,m}\rangle\, h_{j,m} \Big| \exp\left(L^{2/(d-1)}\right)\Bigg\| \\
&\qquad \le \sum_{k=n}^{\infty} \Bigg\| \sum_{j\in\mathbb{N}^d_{-1}:|j|=k} 2^{|j|} \sum_{m\in\mathbb{D}_j} \langle L_{\mathcal{P}_{2^n}}, h_{j,m}\rangle\, h_{j,m} \Big|\exp\left(L^{2/(d-1)}\right)\Bigg\| \\
&\qquad \le \sum_{k=n}^{\infty} \Bigg\| \Bigg(\sum_{j\in\mathbb{N}^d_{-1}:\,|j|=k} 2^{2k} \sum_{m\in\mathbb{D}_j} |\langle L_{\mathcal{P}_{2^n}}, h_{j,m}\rangle|^2 \chi_{I_{j,m}} \Bigg)^{1/2} \Big| L_\infty \Bigg\| \\
&\qquad \preceq 2^n \sum_{k=n}^{\infty} \left(2^{-2k}\,(k+1)^{d-1}\right)^{1/2} \preceq n^{(d-1)/2}.
\end{aligned}$$

The estimate of the counting part is somewhat harder. Recall that $\mathcal{J}$ denotes the family of all dyadic intervals $I_{j,m}\subset U$ with $|j|\ge n$, i.e., $|I_{j,m}|\le 2^{-n}$, such that $\langle C_{\mathcal{P}_{2^n}}, h_{j,m}\rangle \ne 0$. As noticed earlier, if $I_{j,m}\in\mathcal{J}$, this implies that $I_{j,m}$ contains at least one point of $\mathcal{P}_{2^n}$ in its interior; therefore, $j_k\le 2n$ (i.e., $|I_{j_k,m_k}|\ge 2^{-2n}$) for each $k=1,\dots,d$.

In addition, for each $I_{j,m}\in\mathcal{J}$, we can find its unique **parent**, i.e., the interval $\widetilde{I}_{j',m'}$ that satisfies

(i) $I_{j,m}\subset \widetilde{I}_{j',m'}$;

(ii) $|j'| = 2^{-n}$, i.e., $|\widetilde{I}_{j',m'}| = 2^{-n}$; and

(iii) $j_k = j'_k$ (which implies that $I_{j_k,m_k} = \widetilde{I}_{j'_k,m'_k}$) for all $k=1,\dots,d-1$.

In other words, to find the parent, we expand the d-th side of $I_{j,m}$ so that the resulting interval has volume 2^{-n}.

We can now reorganize the sum according to the parents

$$\sum_{\substack{j\in\mathbb{N}^d_{-1}\\ |j|\geq n}} 2^{|j|} \sum_{m\in\mathbb{D}_j} \langle C_{\mathcal{P}_{2^n}}, h_{j,m}\rangle\, h_{j,m} = \sum_{\substack{\widetilde{I}_{j',m'}:\,|j'|=n\\ j'_k\leq 2n:\,k=1,\ldots,d}} \sum_{\substack{I_{j,m}\subset\widetilde{I}_{j',m'}\\ j_k=j'_k:k=1,\ldots,d-1}} 2^{|j|}\langle C_{\mathcal{P}_{2^n}}, h_{j,m}\rangle h_{j,m}. \tag{19}$$

Fix an arbitrary parent interval $\widetilde{I}_{j',m'}$ and consider the innermost sum above
(20)
$$\sum_{\substack{I_{j,m}\subset\widetilde{I}_{j',m'}\\ j_k=j'_k:\,k\leq d-1}} 2^{|j|}\langle C_{\mathcal{P}_{2^n}}, h_{j,m}\rangle\, h_{j,m} = \sum_{p\in\mathcal{P}_{2^n}\cap\widetilde{I}_{j',m'}} \sum_{\substack{I_{j,m}\subset\widetilde{I}_{j',m'}\\ j_k=j'_k:\,k\leq d-1}} 2^{|j|}\langle \chi_{[p,1)}, h_{j,m}\rangle\, h_{j,m}.$$

We notice that the expression inside the last sum splits into a product of one-dimensional factors:

$$\begin{aligned}
2^{|j|}\langle\chi_{[p,1)}, h_{j,m}\rangle h_{j,m}(x) &= \prod_{j=1}^{d} 2^{j_k}\langle\chi_{[p_k,1)}, h_{j_k,m_k}\rangle\, h_{j_k,m_k}(x_k)\\
&= \left(\prod_{j=1}^{d-1} 2^{j'_k}\langle\chi_{[p_k,1)}, h_{j'_k,m'_k}\rangle\, h_{j'_k,m'_k}(x_k)\right)\cdot 2^{j_d}\langle\chi_{[p_d,1)}, h_{j_d,m_d}\rangle\, h_{j_d,m_d}(x_d)\\
&= 2^{|j'_*|}\langle\chi_{[p_*,1)}, h_{j'_*,m'_*}\rangle\, h_{j'_*,m'_*}(x_1,\ldots,x_{d-1})\cdot 2^{j_d}\langle\chi_{[p_d,1)}, h_{j_d,m_d}\rangle\, h_{j_d,m_d}(x_d),
\end{aligned}$$

where $*$ denotes the projection of the d-dimensional vector to its first $d-1$ coordinates, e.g., if $j=(j_1,\ldots,j_d)$, then $j_*=(j_1,\ldots,j_{d-1})$. The expression in (20) can now be rewritten as

$$\begin{aligned}
&\sum_{\substack{I_{j,m}\subset\widetilde{I}_{j',m'}\\ j_k=j'_k:\,k\leq d-1}} 2^{|j|}\langle\chi_{[p,1)}, h_{j,m}\rangle\, h_{j,m}(x)\\
&\quad= 2^{|j'_*|}\langle\chi_{[p_*,1)}, h_{j'_*,m'_*}\rangle\, h_{j'_*,m'_*}(x_*)\cdot \sum_{I_{j_d,m_d}\subset\widetilde{I}_{j'_d,m'_d}} 2^{j_d}\langle\chi_{[p_d,1)}, h_{j_d,m_d}\rangle\, h_{j_d,m_d}(x_d),
\end{aligned}$$

leaving us with the task of examining this ultimately one-dimensional sum. However, one can easily see that this sum is precisely the Haar expansion of the function $\chi_{[p_d,1)}$ restricted to the interval $\widetilde{I}_{j'_d,m'_d}$, except for the constant term, i.e.,

$$\sum_{I_{j_d,m_d}\subset\widetilde{I}_{j'_d,m'_d}} 2^{j_d}\langle\chi_{[p_d,1)}, h_{j_d,m_d}\rangle h_{j_d,m_d}(x_d) = \chi_{\widetilde{I}_{j'_d,m'_d}}(x_d)\cdot\big(\chi_{[p_d,1)}(x_d)-2^{j_d}\big|[p_d,1)\cap\widetilde{I}_{j'_d,m'_d}\big|\big), \tag{21}$$

which, in particular, is bounded pointwise by 2. Obviously,

$$\big|2^{|j'_*|}\langle\chi_{[p_*,1)}, h_{j'_*,m'_*}\rangle\big| \leq 1.$$

Recall that, according to Lemma 2.5, there are at most $2^{\lceil t/2\rceil}$ points $p \in \mathcal{P}_{2^n} \cap \widetilde{I}_{j',m'}$. Therefore,

$$\sum_{\substack{I_{j,m}\subset \widetilde{I}_{j',m'}\\ j_k=j'_k:\, k=1,\ldots,d-1}} 2^{|j|}\langle C_{\mathcal{P}_{2^n}}, h_{j,m}\rangle\, h_{j,m}(x) = \alpha_{j'_d}(x_d) h_{j'_*,m'_*}(x_*),$$

where $|\alpha_{j'_d}(x_d)| \leq 2^{\lceil t/2\rceil+1} \preceq 1$.

Let x_d be fixed for the moment. Due to (21), for a given $(d-1)$-dimensional dyadic interval $\widetilde{I}_{j'_*,m'_*}$, there exists at most one d-dimensional dyadic interval $\widetilde{I}_{j',m'} = \widetilde{I}_{j'_*,m'_*} \times \widetilde{I}_{j'_d,m'_d}$ with $|\widetilde{I}_{j',m'}| = 2^{-n}$ such that $\alpha_{j'_d}(x_d) \neq 0$. Therefore, applying (19) and taking L_p-norms in the first $d-1$ variables, we obtain

$$\begin{aligned}
&\Bigg\| \sum_{\substack{j\in\mathbb{N}^d_{-1}\\ |j|\geq n}} 2^{|j|} \sum_{m\in\mathbb{D}_j} \langle C_{\mathcal{P}_{2^n}}, h_{j,m}\rangle\, h_{j,m} \Big| L_p(dx_*)\Bigg\| \\
&\quad = \Bigg\| \sum_{\substack{\widetilde{I}_{j',m'}:|j'|=n\\ j'_k\leq 2n:\, k=1,\ldots,d}} \sum_{\substack{I_{j,m}\subset\widetilde{I}_{j',m'}\\ j_k=j'_k:\, k=1,\ldots,d-1}} 2^{|j|}\langle C_{\mathcal{P}_{2^n}}, h_{j,m}\rangle\, h_{j,m} \,\Big|\, L_p(dx_*)\Bigg\| \\
&\quad = \Bigg\| \sum_{\substack{\widetilde{I}_{j'_*,m'_*}:|j'|=n\\ j'_k\leq 2n,\, k=1,\ldots,d-1}} \alpha_{j'_d}(x_d) h_{j'_*,m'_*} \,\Big|\, L_p(dx_*)\Bigg\| \\
&\quad \preceq p^{\frac{d-1}{2}} \Bigg\| \Bigg(\sum_{\substack{\widetilde{I}_{j'_*,m'_*}:\,|j'|=n\\ j'_k\leq 2n,\, k=1,\ldots,d-1}} |\alpha_{j'_d}(x_d)|^2 \chi_{\widetilde{I}_{j'_*,m'_*}} \Bigg)^{1/2} \Big|\, L_p(dx_*)\Bigg\| \preceq p^{\frac{d-1}{2}} n^{\frac{d-1}{2}},
\end{aligned}$$

where in the last line we have employed the $(d-1)$-dimensional Littlewood–Paley inequality (11) and the fact that there are of the order of n^{d-1} choices of j'_* in the sum. Integrating this bound with respect to x_d and applying Proposition 2.2, we arrive at

$$\Bigg\| \sum_{\substack{j\in\mathbb{N}^d_{-1}\\ |j|\geq n}} 2^{|j|} \sum_{m\in\mathbb{D}_j} \langle C_{\mathcal{P}_{2^n}}, h_{j,m}\rangle\, h_{j,m} \Big| \exp\left(L^{2/(d-1)}\right)\Bigg\| \preceq n^{(d-1)/2},$$

which finishes the proof of Theorem 1.3.

3.4 Proof of Corollary 1.4.

We set $\alpha = 2/(d-1)$ and use Proposition 2.4 to interpolate between the $\exp\left(L^{2/(d-1)}\right)$ estimate (7) of Theorem 1.3 and the L_∞

estimate (16) of Lemma 2.6, obtaining

$$\left\|D_{\mathcal{P}_{2^n}} \,|\, \exp(L^{\beta})\right\| \preceq \left\|D_{\mathcal{P}_{2^n}} \,|\, \exp(L^{2/(d-1)})\right\|^{\frac{2}{(d-1)\beta}} \cdot \left\|D_{\mathcal{P}_{2^n}} \,|L_\infty([0,1)^d)\right\|^{1-\frac{2}{(d-1)\beta}}$$
$$\preceq n^{\frac{d-1}{2}\cdot\frac{2}{(d-1)\beta}}\, n^{(d-1)\cdot\left(1-\frac{2}{(d-1)\beta}\right)} = n^{(d-1)-\frac{1}{\beta}}.$$

3.5 Orlicz space estimates and star-discrepancy. To conclude, we outline an argument which demonstrates how estimates in exponential Orlicz spaces may be related to the "great open problem" of the subject [2], i.e., sharp bounds on the L_∞-discrepancy. Let us assume that for a certain order 2 digital net $\mathcal{P}_{2^n}$ with $N = 2^n$ points for some $\alpha > 0$, the discrepancy function satisfies the exponential bound

$$\left\|D_{\mathcal{P}_{2^n}} \,|\, \exp(L^{\alpha})\right\| \preceq (\log N)^{(d-1)/2} \simeq n^{(d-1)/2}. \tag{22}$$

This trivially leads to the following distributional estimate: for each $\lambda > 0$,

$$\mu\left\{x \in [0,1]^d : \left|D_{\mathcal{P}_{2^n}}(x)\right| > \lambda\right\} \preceq \exp\left(-\left(\frac{\lambda}{n^{(d-1)/2}}\right)^{\alpha}\right),$$

where μ is Lebesgue measure. The fact that $\mathcal{P}_{2^n}$ is a binary digital net (i.e., all points have binary coordinates of length $2n$) implies that its discrepancy function does not change much on dyadic intervals of side length 2^{-2n}. Therefore, for those values of λ for which the set $\left\{\left|D_{\mathcal{P}_{2^n}}(x)\right| > \lambda\right\}$ is non-empty, we must have

$$\mu\left\{x \in [0,1]^d : \left|D_{\mathcal{P}_{2^n}}(x)\right| > \lambda\right\} \succeq 2^{-2nd}.$$

Comparing the last two estimates, we observe that they cannot simultaneously hold if $\lambda \succeq n^{\frac{d-1}{2}+\frac{1}{\alpha}}$; i.e., in this case, $\left\{\left|D_{\mathcal{P}_{2^n}}(x)\right| > \lambda\right\} = \varnothing$, in other words

$$\left\|D_{\mathcal{P}_{2^n}}\right\|_{L_\infty} \preceq n^{\frac{d-1}{2}+\frac{1}{\alpha}}. \tag{23}$$

Recall that two main conjectures about the correct asymptotics of the discrepancy function predict the sharp order of growth of either $(\log N)^{d-1}$ or $(\log N)^{d/2}$. Our Theorem 1.3 is consistent with the first hypothesis: in this case, (22) holds with $\alpha = 2/(d-1)$, and hence (23) becomes $\left\|D_{\mathcal{P}_{2^n}}\right\|_{L_\infty} \preceq n^{d-1}$ which matches the best known upper bound (4).

In striving to prove the second conjecture along these lines, estimate (22) should hold with $\alpha = 2$, i.e., one would need to construct a digital net whose discrepancy function is subgaussian.

We notice that Skriganov [41, Lemma 6.2] uses a somewhat different discretization approach which yields similar results and shows that an estimate in Orlicz space $\exp\left(L^{2/(d-1)}\right)$ yields the L_∞ upper bound for the discrepancy function of the order $(\log N)^{d-1}$.

3.6 Acknowledgements. Both authors express gratitude to the organizers of the workshop “Discrepancy, Numerical Integration and Hyperbolic Cross Approximation” (HCM, Bonn, Germany, 2013), the international conference MC-QMC 2014 (KU Leuven, Belgium), the Hausdorff trimester program “Harmonic Analysis and PDE” (HIM, Bonn, Germany, 2014), the semester program “High-dimensional approximation” (ICERM, Brown University, Providence, RI, USA, 2014), where they had an opportunity to meet and discuss their work.

References

[1] G. Amirkhanyan, D. Bilyk, and M. Lacey, *Estimates of the discrepancy function in exponential Orlicz spaces*, (2014), availiable at https://arxiv.org/abs/1306.1766.

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

[3] A. Bernard, *Espaces H^1 de martingales à deux indices. Dualité avec des martingales de type “BMO’*, Bull. Sci. Math. **103** (1979), 297–303.

[4] D. Bilyk, *On Roth’s orthogonal function method in discrepancy theory*, Unif. Distrib. Theory **6** (2011), 143–184.

[5] 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.

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

[7] S.-Y. A. Chang and R. Fefferman, *A continuous version of duality of H1 with BMO on the bidisc*, Ann. of Math. **112** (1980), 179–201.

[8] S.-Y. A. Chang, J. M. Wilson, and T. H. Wolff, *Some weighted norm inequalities concerning the Schrdinger operators*, Comment. Math. Helv. **60** (1985), 217–246.

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

[10] 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.

[11] W. W. L. Chen and M. M. Skriganov, *Orthogonality and digit shifts in the classical mean squares problem in irregularities of point distribution*, in *Diophantine Approximation*, Springer, Vienna, 2008, pp. 141–159.

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

[13] 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.

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

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

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

[17] 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.

[18] 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*, Walter DeGruyter GmbH, Berlin/Boston, 2014, pp. 63–86.

[19] H. Faure, F. Pillichshammer, G. Pirsic, and W. Ch. Schmid, *L_2 discrepancy of generalized two-dimensional Hammersley point sets scrambled with arbitrary permutations*, Acta Arith. **141** (2010), 395–418.

[20] 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.

[21] J. H. Halton, *On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals*, Numer. Math. **2** (1960), 84–90.

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

[23] A. Hinrichs, *Discrepancy, integration and tractability*, in *Monte Carlo and Quasi-Monte Carlo Methods 2012*, Springer, Berlin-Heidelberg, 2013, pp. 129–172.

[24] A. Hinrichs and L. Markhasin, *On lower bounds for the L_2-discrepancy*, J. Complexity **27** (2011), 127–132.

[25] L. Kuipers and H. Niederreiter, Uniform Distribution of Sequences, John Wiley & Sons, Ltd., New York, 1974.

[26] J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces. I.*, Springer, Berlin, 1977.

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

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

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

[30] L. Markhasin, *L_2- and $S^r_{p,q}B$-discrepancy of (order 2) digital nets*, Acta Arith. **168** (2015), 139–160.

[31] J. Matoušek, *Geometric discrepancy. An Illustrated Guide*, Springer-Verlag, Berlin, 1999.

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

[33] E. Novak and H. Woźniakowski, *Tractability of Multivariate Problems. Volume II: Standard Information for Functionals* European Mathematical Society Publishing House, Zürich, 2010.

[34] J. Pipher and L. Ward, *BMO from dyadic BMO on the bydisc*, J. London Math. Soc. **77** (2008), 524–544.

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

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

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

[38] W. M. Schmidt, *Irregularities of distribution X*, in *Number Theory and Algebra*, Academic Press, New York, 1977, pp. 311-329.

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

[40] M. M. Skriganov, *On mean values of the L_q-discrepancies of point distributions*, St. Petersburg Math. J. **24** (2013), 991–1012.

[41] M. M. Skriganov, *Dyadic shift randomizations in classical discrepancy theory*, Mathematikca **62** (2016), 183–209.

[42] 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.

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

Dmitriy Bilyk
School of Mathematics
University of Minnesota
206 Church St. SE
Minneapolis, MN, 55408, USA
email: dbilyk@math.umn.edu

Lev Markhasin
Institut für Stochastik und Anwendungen
Universität Stuttgart
Pfaffenwaldring 57
70569 Stuttgart, Germany
email: lev.markhasin@mathematik.uni-stuttgart.de

(Received December 10, 2014 and in revised form July 12, 2015)