Mathematika 63 (2017) 863–894 © 2017 University College London
doi:10.1112/S0025579317000213

DISCREPANCY OF SECOND ORDER DIGITAL SEQUENCES IN FUNCTION SPACES WITH DOMINATING MIXED SMOOTHNESS

JOSEF DICK, AICKE HINRICHS, LEV MARKHASIN AND FRIEDRICH PILLICHSHAMMER

Gewidmet dem Andenken an Klaus Friedrich Roth 1925–2015

Abstract. The discrepancy function measures the deviation of the empirical distribution of a point set in $[0,1]^d$ from the uniform distribution. In this paper, we study the classical discrepancy function with respect to the bounded mean oscillation and exponential Orlicz norms, as well as Sobolev, Besov and Triebel–Lizorkin norms with dominating mixed smoothness. We give sharp bounds for the discrepancy function under such norms with respect to infinite sequences.

§1. *Background.* Let d, N be positive integers and let $\mathcal{P}_{N,d}$ be a point set in the d-dimensional unit cube $[0,1)^d$ with N elements. The discrepancy function of $\mathcal{P}_{N,d}$ is defined as

$$D_{\mathcal{P}_{N,d}}(\boldsymbol{x}) = \frac{1}{N} \sum_{\boldsymbol{z} \in \mathcal{P}_{N,d}} \chi_{[\boldsymbol{0},\boldsymbol{x})}(\boldsymbol{z}) - x_1 \cdots x_d. \tag{1.1}$$

Here χ_A denotes the characteristic function of a subset $A \in \mathbb{R}^d$. Thus $D_{\mathcal{P}_{N,d}}(\boldsymbol{x})$ denotes the difference between the actual and expected proportions of points that fall into $[\boldsymbol{0},\boldsymbol{x}) = [0,x_1) \times \cdots \times [0,x_d)$, where $\boldsymbol{x} = (x_1,\ldots,x_d) \in [0,1]^d$.

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

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

Let $\mathcal{S}_d$ be an infinite sequence in $[0,1)^d$ and let $N \in \mathbb{N}$. We define the discrepancy function of $\mathcal{S}_d$ 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 elements of $\mathcal{S}_d$, and we define the L_p-discrepancy of $\mathcal{S}_d$ as

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

Received 29 April 2016.
MSC (2010): 11K06, 11K38 (primary), 32A37, 42C10, 46E30, 46E35, 65C05 (secondary).
The research of J. Dick was supported under Australian Research Council's Discovery Projects funding scheme (project number DP150101770). F. Pillichshammer is supported by the Austrian Science Fund (FWF): Project F5509-N26, which is a part of the Special Research Program "Quasi-Monte Carlo Methods: Theory and Applications".

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

The L_p-discrepancy is a quantitative measure for the irregularity of distribution of finite point sets and of infinite sequences. The reader is referred to [1, 17, 20] for extensive introductions to this topic.

The conceptual difference in the concept between the discrepancy of finite point sets and infinite sequences is pointed out in [27]: for finite point sets one is interested in the behaviour of the whole set $\{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N-1}\}$ with a fixed number of elements N, whereas for infinite sequences one is 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. In many situations, the difference between these two settings in discrepancy estimates comes at a cost of "one dimension". In particular, this means that the dynamic setting in dimension d is related to the static setting in dimension $d + 1$, although this connection is not so straightforward, especially in proving upper bounds. This observation will also be confirmed by our results.

It is well known from results of Roth [32] and Schmidt [35] 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 $N \geqslant 2$ we have

$$L_{p,N}(\mathcal{P}_{N,d}) \geqslant c_{p,d}\, N^{-1} (\log N)^{(d-1)/2}. \tag{1.2}$$

These results were extended to infinite sequences by Proinov [31] (see also [16] for a proof). For every $p \in (1, \infty]$ and every $d \in \mathbb{N}$, $d \geqslant 2$, 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

$$L_{p,N}(\mathcal{S}_d) \geqslant c_{p,d}\, N^{-1} (\log N)^{d/2} \quad \text{for infinitely many } N \in \mathbb{N}. \tag{1.3}$$

The lower bound (1.2) for finite point sets is known to be best possible in the order of magnitude in N, i.e., for every d, $N \in \mathbb{N}$, $N \geqslant 2$, one can find (and explicitly construct) an N-element point set $\mathcal{P}_{N,d}$ in $[0, 1)^d$ with L_p-discrepancy of order

$$L_{p,N}(\mathcal{P}_{N,d}) \ll_{p,d} N^{-1} (\log N)^{(d-1)/2}. \tag{1.4}$$

(For functions $f, g : D \subseteq \mathbb{N} \to \mathbb{R}$ with $g \geqslant 0$ the notation $f(N) \ll g(N)$ means that there exists some $C > 0$ such that $f(N) \leqslant C g(N)$ for all $N \in D$. If we want to stress that C depends on some parameters, say a, b, then this is indicated by writing $f(N) \ll_{a,b} g(N)$. If we have $f(N) \ll g(N)$ and $g(N) \ll f(N)$, then we write $f(N) \asymp g(N)$.)

The first proof of (1.4) was given by Roth [33] for $p = 2$ and arbitrary d and finally by Chen [7] in the general case, albeit using a probabilistic approach. Explicit constructions are due to Chen and Skriganov [8] and Skriganov [36] (the first time that for each $1 \leqslant p < \infty$ a construction method of an optimal point set is given), Dick and Pillichshammer [15], Dick [11] (the first proof where the construction method of the point set does not depend on p), and Markhasin [26].

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

It is also known that the lower bound (1.3) for infinite sequences is best possible for all $p \in (1, \infty)$. This was shown by the authors of the present paper in the recent work [**13**] where explicit constructions $\mathcal{S}_d$ of infinite sequences in arbitrary dimensions d were provided whose L_p-discrepancy satisfies

$$L_{p,N}(\mathcal{S}_d) \ll_{p,d} N^{-1}(\log N)^{d/2}. \tag{1.5}$$

The sequences and methods developed in [**13**] are the fundamental basis for the results that will be presented in this paper (see, for example, Lemma 4.5).

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 §4, see [**14**, **28**, **29**]) such that

$$L_{\infty,N}(\mathcal{S}_d) \ll_d N^{-1}(\log N)^d. \tag{1.6}$$

Regarding lower bounds, it is known 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

$$L_{\infty,N}(\mathcal{S}_d) \geqslant c_d\, N^{-1}(\log N)^{d/2+\eta_d} \quad \text{for infinitely many } N \in \mathbb{N}.$$

This result follows from a corresponding result for finite point sets by Bilyk *et al* [**3**]. For growing d the exponent η_d in this estimate tends to zero.

The quest for the exact order of the L_∞-discrepancy in the multivariate case is a very demanding open question. Only for $d = 1$ it is known that the exact order of the L_∞-discrepancy of infinite sequences is $(\log N)/N$. This follows from (1.6) together with a celebrated result of Schmidt [**34**].

§2. *Intermediate norms.* It is now a natural and instructive question to ask what happens in intermediate spaces "close" to L_∞. One standard example of such spaces is the space BMO^d, which stands for *bounded mean oscillation*. For the definition of this space we need the concept of Haar functions.

We define $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and $\mathbb{N}_{-1} = \mathbb{N}_0 \cup \{-1\}$. Let $\mathbb{D}_j = \{0, 1, \dots, 2^j - 1\}$ for $j \in \mathbb{N}_0$ and $\mathbb{D}_{-1} = \{0\}$. For $\boldsymbol{j} = (j_1, \dots, j_d) \in \mathbb{N}_{-1}^d$ 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 $I_{j,m} = [2^{-j}m, 2^{-j}(m + 1))$ the mth dyadic interval in $[0, 1)$ on level j. We put $I_{-1,0} = [0, 1)$ and call it the 0th 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 $\boldsymbol{m} = (m_1, \dots, m_d) \in \mathbb{D}_{\boldsymbol{j}}$ 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 mth *dyadic Haar function on level* j.

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

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 *dyadic Haar basis on* $[0,1)^d$.

It is well known that the system

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

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.

For an integrable function $f : [0,1]^d \to \mathbb{R}$ we define

$$\|f|\mathrm{BMO}^d\|^2 = \sup_{U \subseteq [0,1)^d} \lambda_d(U)^{-1} \sum_{\boldsymbol{j} \in \mathbb{N}_0^d} 2^{|\boldsymbol{j}|} \sum_{\substack{\boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}} \\ I_{\boldsymbol{j},\boldsymbol{m}} \subseteq U}} |\langle f, h_{\boldsymbol{j},\boldsymbol{m}} \rangle|^2, \tag{2.1}$$

where $h_{\boldsymbol{j},\boldsymbol{m}}$ is the $\boldsymbol{m}$th dyadic Haar function on level $\boldsymbol{j}$, λ_d is the d-dimensional Lebesgue measure and the supremum is taken over all measurable sets $U \subseteq [0,1)^d$.

The space BMO^d contains all integrable functions f with finite norm $\|f|\mathrm{BMO}^d\|$. Note that strictly speaking $\|f|\mathrm{BMO}^d\|$ is only a seminorm, since it vanishes on linear combinations of functions which are constant in one or more coordinate directions. This means that formally we need to consider a factor space over such functions. For a more detailed study of the spaces BMO^d we refer to [5] and [6].

We consider the BMO-seminorm of the discrepancy function of infinite sequences and call this the BMO-discrepancy.

For finite point sets the following is shown in [4, Theorems 1.1 and 1.2]. Suppose that $d \in \mathbb{N}$. For every finite point set $\mathcal{P}_{N,d}$ in $[0,1)^d$,

$$\|D_{\mathcal{P}_{N,d}}|\mathrm{BMO}^d\| \gg N^{-1}(\log N)^{(d-1)/2} \quad \text{for all } N \in \mathbb{N} \setminus \{1\}.$$

Furthermore, there exists a finite point set $\mathcal{P}_{N,d}$ in $[0,1)^d$ such that

$$\|D_{\mathcal{P}_{N,d}}|\mathrm{BMO}^d\| \ll_d N^{-1}(\log N)^{(d-1)/2} \quad \text{for all } N \in \mathbb{N} \setminus \{1\}.$$

The first main result of this paper is a corresponding result for infinite sequences.

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

THEOREM 2.1. *Suppose that* $d \in \mathbb{N}$. *For every infinite sequence* $\mathcal{S}_d$ *in* $[0,1)^d$,

$$\|D_{\mathcal{S}_d}^N|\mathrm{BMO}^d\| \gg N^{-1}(\log N)^{d/2} \quad \text{for infinitely many } N \in \mathbb{N}. \tag{2.2}$$

Furthermore, there exists an infinite sequence $\mathcal{S}_d$ *in* $[0,1)^d$ *such that*

$$\|D_{\mathcal{S}_d}^N|\mathrm{BMO}^d\| \ll_d N^{-1}(\log N)^{d/2} \quad \text{for all } N \in \mathbb{N} \setminus \{1\}. \tag{2.3}$$

The proof of this result will be presented in §5.1.

As another example we study the exponential Orlicz norm, which is closer to the L_∞-norm than any L_p-norm with $1 \leqslant p < \infty$. Thus bounds on the exponential Orlicz norm of the discrepancy function shed more light on the behaviour of the L_∞-norm. For a probability space (Ω, P), let $\mathbb{E}$ denote the expectation over (Ω, P). To define the Orlicz norm, let $\psi : [0,\infty) \to [0,\infty)$ be a convex function with $\psi(x) = 0$ if and only if $x = 0$. The Orlicz norm L^ψ of a (Ω, P)-measurable function f is now given by

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

where we set $\inf \emptyset = \infty$. The Orlicz space L^ψ consists of all functions f with finite Orlicz norm $\|f|L^\psi\|$.

For $\alpha > 0$, one obtains the exponential Orlicz norm $\exp(L^\alpha) = L^{\psi_\alpha}$ by choosing ψ_α to be a convex function which equals $\exp(x^\alpha) - 1$ for sufficiently large x, depending on α (for $\alpha \geqslant 1$ we can use $\psi_\alpha(x) = \exp(x^\alpha) - 1$ for all $x \geqslant 0$). More information can be found in [22].

For finite point sets the following is shown in [4, Corollary 1.4] (see [2, Theorem 1.4] for the case $d = 2$). There exists a finite point set $\mathcal{P}_{N,d}$ in $[0,1)^d$ such that

$$\|D_{\mathcal{P}_{N,d}}|\exp(L^\beta)\| \ll_d N^{-1}(\log N)^{d-1-1/\beta} \quad \text{for all } N \in \mathbb{N} \setminus \{1\} \text{ and for all } \frac{2}{d-1} \leqslant \beta < \infty.$$

We show a corresponding result for infinite sequences.

THEOREM 2.2. *There exists an infinite sequence* $\mathcal{S}_d$ *in* $[0,1)^d$ *such that*

$$\|D_{\mathcal{S}_d}^N|\exp(L^\beta)\| \ll_d N^{-1}(\log N)^{d-1/\beta} \quad \textit{for all } N \in \mathbb{N} \setminus \{1\} \textit{ and for all } \frac{2}{d-1} \leqslant \beta < \infty.$$

From Proposition 2.3 below in conjunction with (1.3) it follows that for every infinite sequence $\mathcal{S}_d$ in $[0,1)^d$ the discrepancy $\|D_{\mathcal{S}_d}^N|\exp(L^\beta)\|$ is at least of order $(\log N)^{d/2}/N$ for infinitely many $N \in \mathbb{N}$. However, a matching lower bound for $\|D_{\mathcal{S}_d}^N|\exp(L^\beta)\|$ seems to be presently beyond reach, even for finite point sets as pointed out in [4] (see [4, Remark after Theorem 1.3] for a short discussion

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

about this problem). There is only one singular result for *finite* point sets $\mathcal{P}_{N,2}$ in dimension $d = 2$, $\|D_{\mathcal{P}_{N,2}}|\exp(L^\beta)\| \geqslant c(\log N)^{1-1/\beta}/N$ for $2 \leqslant \beta < \infty$ and some $c > 0$, according to [2], for which the upper and lower bound match.

We close this section with some further technical results concerning the exponential Orlicz norm. The reader familiar with this topic may also skip straight to the beginning of the next section.

An equivalent way of stating the exponential Orlicz norm is using the following result, which can be shown using the Taylor series expansion of $\exp(x)$ and Stirling's formula.

PROPOSITION 2.3. *For any $\beta > 0$, the following equivalence holds*

$$\|f|\exp(L^\beta)\| \asymp_d \sup_{p>1} p^{-1/\beta}\|f|L_p([0,1]^d)\|.$$

Our proof will make use of the Littlewood–Paley inequality, which we state in the following as a lemma.

LEMMA 2.4. *Let $1 < p < \infty$. Then for a function $f : [0,1)^d \to \mathbb{R}$ we have*

$$\|f|L_p([0,1)^d)\| \ll_d \; p^{d/2}\|Sf|L_p([0,1)^d)\|, \tag{2.4}$$

where the square function Sf is given by

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

If furthermore f is a hyperbolic sum of multiparameter Haar functions, i.e., $f \in \operatorname{span}\{h_{\boldsymbol{j},\boldsymbol{m}} : |\boldsymbol{j}| = n\}$ for some $n \in \mathbb{N}$, then the exponent of p in (2.4) can be reduced to $(d-1)/2$.

An essential tool in the proof of the bound on the exponential Orlicz norm of the discrepancy function is the following hyperbolic Chang–Wilson–Wolff inequality. It relates the exponential Orlicz norm of a function to the L_∞-norm of the corresponding square function. This inequality was also recently used in [4].

PROPOSITION 2.5 (Hyperbolic Chang–Wilson–Wolff inequality). *Assume that f is a hyperbolic sum of multiparameter Haar functions, i.e., $f \in \operatorname{span}\{h_{\boldsymbol{j},\boldsymbol{m}} : |\boldsymbol{j}| = n\}$ for some $n \in \mathbb{N}$. Then*

$$\|f|\exp(L^{2/(d-1)})\| \ll \|Sf|L_\infty([0,1)^d)\|.$$

The proof follows easily from the characterization in Proposition 2.3, Lemma 2.4 and the monotonicity of the L_p-norm.

Below we use the Hyperbolic Chang–Wilson–Wolff inequality to obtain bounds on the $\exp(L^{2/(d-1)})$-norm of the discrepancy function. To obtain the general result of Theorem 2.2 we need to interpolate between this result and the L_∞-norm of the discrepancy function. Concretely, we use the following result [4, Proposition 2.4], where also a short proof can be found.

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

PROPOSITION 2.6. *Let $0 < \alpha < \beta < \infty$. Consider a function $f \in L_\infty([0,1)^d)$. If $f \in \exp(L^\alpha)$, then also $f \in \exp(L^\beta)$ and we have*

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

§3. *Besov and Triebel–Lizorkin spaces with dominating mixed smoothness.* We are also interested in Besov spaces and Triebel–Lizorkin spaces with dominating mixed smoothness and their quasi-norm of the discrepancy function. The discrepancy of finite point sets in Besov spaces and Triebel–Lizorkin spaces and the corresponding integration orders where considered in [18, 23–26, 40, 41]. These spaces are particularly interesting, since they form natural scales of function spaces including the L_p-spaces and Sobolev spaces of, dominating mixed smoothness. To guide the reader not familiar with these function spaces we remark that the parameter s in the spaces $S^s_{p,q}B$ and $S^s_{p,q}F$ introduced below is the (dominating mixed) smoothness, the parameter p denotes integrability and q is a fine index, see [38]. It is remarkable that in the case $s > 0$ the results do not show the usual difference between the static setting of finite point sets and the dynamic setting of sequences considered here. Instead, the behaviour of the discrepancy is the same, compare (ii) in Theorems 3.1 and 3.2 with the corresponding results for point sets in the above mentioned references.

We follow [40]. Let $\mathfrak{S}(\mathbb{R}^d)$ denote the Schwartz space and $\mathfrak{S}'(\mathbb{R}^d)$ the space of tempered distributions on $\mathbb{R}^d$. For $\varphi \in \mathfrak{S}(\mathbb{R}^d)$ we denote by $\mathcal{F}\varphi$ the Fourier transform of φ and extend it to $\mathfrak{S}'(\mathbb{R}^d)$ in the usual way. For $f \in \mathfrak{S}'(\mathbb{R}^d)$ the Fourier transform is given as $\mathcal{F}f(\varphi) = f(\mathcal{F}\varphi), \varphi \in \mathfrak{S}(\mathbb{R}^d)$. Analogously we proceed with the inverse Fourier transform $\mathcal{F}^{-1}$.

Let $\varphi_0 \in \mathfrak{S}(\mathbb{R})$ satisfy $\varphi_0(x) = 1$ for $|x| \leqslant 1$ and $\varphi_0(x) = 0$ for $|x| > \frac{3}{2}$. Let $\varphi_k(x) = \varphi_0(2^{-k}x) - \varphi_0(2^{-k+1}x)$ where $x \in \mathbb{R}, k \in \mathbb{N}$ and $\varphi_{\boldsymbol{k}}(\boldsymbol{x}) = \varphi_{k_1}(x_1)\cdots\varphi_{k_d}(x_d)$ where $\boldsymbol{k} = (k_1,\dots,k_d) \in \mathbb{N}_0^d$ and $\boldsymbol{x} = (x_1,\dots,x_d) \in \mathbb{R}^d$. The set of functions $\{\varphi_{\boldsymbol{k}}\}$ are a dyadic resolution of unity since

$$\sum_{\boldsymbol{k}\in\mathbb{N}_0^d} \varphi_{\boldsymbol{k}}(\boldsymbol{x}) = 1$$

for all $\boldsymbol{x} \in \mathbb{R}^d$. The functions $\mathcal{F}^{-1}(\varphi_{\boldsymbol{k}}\mathcal{F}f)$ are entire analytic functions for every $f \in \mathfrak{S}'(\mathbb{R}^d)$.

Let $0 < p,q \leqslant \infty$ and $s \in \mathbb{R}$. The Besov space with dominating mixed smoothness $S^s_{p,q}B(\mathbb{R}^d)$ consists of all $f \in \mathfrak{S}'(\mathbb{R}^d)$ with finite quasi-norm

$$\|f|S^s_{p,q}B(\mathbb{R}^d)\| = \left(\sum_{\boldsymbol{k}\in\mathbb{N}_0^d} 2^{s|\boldsymbol{k}|q}\|\mathcal{F}^{-1}(\varphi_{\boldsymbol{k}}\mathcal{F}f)|L_p(\mathbb{R}^d)\|^q\right)^{1/q} \tag{3.1}$$

with the usual modification if $q = \infty$.

Let $\mathfrak{D}([0,1)^d)$ consist of all complex-valued infinitely differentiable functions on $\mathbb{R}^d$ with compact support in the interior of $[0,1)^d$ and let $\mathfrak{D}'([0,1)^d)$ be its dual space of all distributions in $[0,1)^d$. The Besov space with

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

dominating mixed smoothness $S^s_{p,q}B([0,1)^d)$ consists of all $f \in \mathfrak{D}'([0,1)^d)$ with finite quasi-norm

$$\|f|S^s_{p,q}B([0,1)^d)\| = \inf\{\|g|S^s_{p,q}B(\mathbb{R}^d)\| : g \in S^s_{p,q}B(\mathbb{R}^d),\, g|_{[0,1)^d} = f\}. \tag{3.2}$$

THEOREM 3.1. *Suppose that* $1 \leqslant p, q < \infty$ *and* $d \in \mathbb{N}$.

(i) *For every infinite sequence* $\mathcal{S}_d$ *in* $[0,1)^d$,

$$\|D^N_{\mathcal{S}_d}|S^0_{p,q}B([0,1)^d)\| \gg_{p,q,d} N^{-1}(\log N)^{d/q} \quad \text{for infinitely many } N \in \mathbb{N}. \tag{3.3}$$

Furthermore, there exists an infinite sequence $\mathcal{S}_d$ *in* $[0,1)^d$ *such that*

$$\|D^N_{\mathcal{S}_d}|S^0_{p,q}B([0,1)^d)\| \ll_{p,q,d} N^{-1}(\log N)^{d/q} \quad \text{for all } N \in \mathbb{N} \setminus \{1\}. \tag{3.4}$$

(ii) *Suppose further that* $0 < s < 1/p$. *For every infinite sequence* $\mathcal{S}_d$ *in* $[0,1)^d$,

$$\|D^N_{\mathcal{S}_d}|S^s_{p,q}B([0,1)^d)\| \gg_{p,q,s,d} N^{s-1}(\log N)^{(d-1)/q} \quad \text{for infinitely many } N \in \mathbb{N}. \tag{3.5}$$

Furthermore, there exists an infinite sequence $\mathcal{S}_d$ *in* $[0,1)^d$ *such that*

$$\|D^N_{\mathcal{S}_d}|S^s_{p,q}B([0,1)^d)\| \ll_{p,q,s,d} N^{s-1}(\log N)^{(d-1)/q} \quad \text{for all } N \in \mathbb{N} \setminus \{1\}. \tag{3.6}$$

In the case $d = 1$ this result has been recently shown in [**19**, Theorem 1] which is based on the symmetrized van der Corput sequence.

Let $0 < p < \infty$, $0 < q \leqslant \infty$ and $s \in \mathbb{R}$. The Triebel–Lizorkin space with dominating mixed smoothness $S^s_{p,q}F(\mathbb{R}^d)$ consists of all $f \in \mathfrak{S}'(\mathbb{R}^d)$ with finite quasi-norm

$$\|f|S^s_{p,q}F(\mathbb{R}^d)\| = \left\| \left(\sum_{\boldsymbol{k} \in \mathbb{N}_0^d} 2^{s|\boldsymbol{k}|q} |\mathcal{F}^{-1}(\varphi_{\boldsymbol{k}} \mathcal{F} f)(\cdot)|^q \right)^{1/q} \middle| L_p(\mathbb{R}^d) \right\| \tag{3.7}$$

with the usual modification if $q = \infty$.

The Triebel–Lizorkin space with dominating mixed smoothness $S^s_{p,q}F([0,1)^d)$ consists of all $f \in \mathfrak{D}'([0,1)^d)$ with finite quasi-norm

$$\|f|S^s_{p,q}F([0,1)^d)\| = \inf\{\|g|S^s_{p,q}F(\mathbb{R}^d)\| : g \in S^s_{p,q}F(\mathbb{R}^d),\, g|_{[0,1)^d} = f\}. \tag{3.8}$$

THEOREM 3.2. *Suppose that* $1 \leqslant p, q < \infty$ *and* $d \in \mathbb{N}$.

(i) *For every infinite sequence* $\mathcal{S}_d$ *in* $[0,1)^d$,

$$\|D^N_{\mathcal{S}_d}|S^0_{p,q}F([0,1)^d)\| \gg_{p,q,d} N^{-1}(\log N)^{d/q} \quad \text{for infinitely many } N \in \mathbb{N}. \tag{3.9}$$

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

Furthermore, there exists an infinite sequence $\mathcal{S}_d$ *in* $[0,1)^d$ *such that*

$$\|D_{\mathcal{S}_d}^N | S_{p,q}^0 F([0,1)^d)\| \ll_{p,q,d} N^{-1}(\log N)^{d/q} \quad \text{for all } N \in \mathbb{N}\setminus\{1\}. \tag{3.10}$$

(ii) *Suppose further that* $0 < s < 1/\min(p,q)$. *For every infinite sequence* $\mathcal{S}_d$ *in* $[0,1)^d$,

$$\|D_{\mathcal{S}_d}^N | S_{p,q}^s F([0,1)^d)\| \gg_{p,q,s,d} N^{s-1}(\log N)^{(d-1)/q} \quad \text{for infinitely many } N \in \mathbb{N}. \tag{3.11}$$

Furthermore, there exists an infinite sequence $\mathcal{S}_d$ *in* $[0,1)^d$ *such that*

$$\|D_{\mathcal{S}_d}^N | S_{p,q}^s F([0,1)^d)\| \geqslant c_{p,q,s,d}\, N^{s-1}(\log N)^{(d-1)/q} \quad \text{for all } N \in \mathbb{N}\setminus\{1\}. \tag{3.12}$$

For $0 < p < \infty$ the spaces $S_p^s W([0,1)^d) = S_{p,2}^s F([0,1)^d)$ are called Sobolev spaces with dominating mixed smoothness. Thus in the case $q = 2$ the above theorem implies results for Sobolev norms with dominating mixed smoothness.

We close this section with some further technical results about Besov and Triebel–Lizorkin spaces. The spaces $S_{p,q}^s B(\mathbb{R}^d)$, $S_{p,q}^s F(\mathbb{R}^d)$, $S_{p,q}^s B([0,1)^d)$ and $S_{p,q}^s F([0,1)^d)$ are quasi-Banach spaces. The following results will be helpful. The first one is an embedding result (see [**40**] or [**24**, Corollary 1.13]), and the second one gives a characterization of the Besov spaces in terms of Haar bases (see [**24**, Theorem 2.11]).

PROPOSITION 3.3. *Let* $p, q \in (0,\infty)$ *and* $s \in \mathbb{R}$. *Then we have*

$$S_{\max(p,q),q}^s B([0,1)^d) \hookrightarrow S_{p,q}^s F([0,1)^d) \hookrightarrow S_{\min(p,q),q}^s B([0,1)^d).$$

PROPOSITION 3.4. *Let* $p, q \in (0,\infty)$, $1/p - 1 < s < \min(1/p, 1)$ *and* $f \in S_{p,q}^s B([0,1)^d)$. *Then*

$$\|f | S_{p,q}^s B([0,1)^d)\|^q \asymp_{p,q,s,d} \sum_{\boldsymbol{j}\in\mathbb{N}_{-1}^d} 2^{|\boldsymbol{j}|(s-1/p+1)q} \left(\sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}} |\langle f, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^p \right)^{q/p}.$$

§4. *Upper bounds.* The explicit constructions in Theorems 2.1, 2.2, 3.1 and 3.2 are all the same and are 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.

4.1. *Digital nets and sequences.* Niederreiter introduced the concepts of digital nets and sequences [**28**] in 1987. A detailed overview of this topic can be found in the books [**14**, **21**, **29**]. Here we restrict ourselves to the dyadic case. Let $\mathbb{F}_2$ be the finite field of order 2 identified with the set $\{0,1\}$ equipped with arithmetic operations modulo 2.

The digital construction scheme. We begin with the definition of digital nets according to Niederreiter, which we present here in a slightly more general form.

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

For $n, q, d \in \mathbb{N}$ with $q \geqslant n$ let $C_1, \ldots, C_d \in \mathbb{F}_2^{q \times n}$ be $q \times n$ matrices over $\mathbb{F}_2$. 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

$$C_j \vec{k} =: (x_{j,k,1}, x_{j,k,2}, \ldots, x_{j,k,q})^\top \quad \text{for } j = 1, \ldots, d, \tag{4.1}$$

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 kth 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* $C_1, \ldots, C_d$. A digital net consists of 2^n elements in $[0, 1)^d$.

Next we recall the definition of digital sequences, 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 $C_j = (c_{j,k,\ell})_{k,\ell \in \mathbb{N}}$ it is assumed 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

$$C_j \vec{k} =: (x_{j,k,1}, x_{j,k,2}, \ldots)^\top \quad \text{for } j = 1, \ldots, d, \tag{4.2}$$

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 kth point $\boldsymbol{x}_k$ of the sequence $\mathcal{S}_d$ is given by $\boldsymbol{x}_k = (x_{1,k}, \ldots, x_{d,k})$. We call a sequence $\mathcal{S}_d$ constructed this way a *digital sequence* (*over* $\mathbb{F}_2$) *with generating matrices* $C_1, \ldots, C_d$. 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.

Higher-order nets and sequences. The choice of the respective generating matrices is crucial for the distribution quality of digital nets and sequences. The following definitions 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 4.1. Let $n, q, \alpha \in \mathbb{N}$ with $q \geqslant \alpha n$ and let t be an integer such that $0 \leqslant t \leqslant \alpha n$. Let $C_1, \ldots, C_d \in \mathbb{F}_2^{q \times n}$. Denote the ith row vector of the matrix C_j by $\vec{c}_{j,i} \in \mathbb{F}_2^n$. If for all $1 \leqslant i_{j,\nu_j} < \cdots < i_{j,1} \leqslant q$ with

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

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* $\mathbb{F}_2$.

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

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

Definition 4.2. Let $\alpha \in \mathbb{N}$ and let $t \geqslant 0$ 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 $\mathbb{F}_2$*.

Definition 4.1 shows that if $\mathcal{P}_{2^n,d}$ is an order-α digital (t, n, d)-net, then for any $t \leqslant t' \leqslant \alpha n$, $\mathcal{P}_{2^n,d}$ is also an order-α digital (t', n, d)-net. An analogous result also applies to higher-order digital sequences.

We point out 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 $\alpha' \neq \alpha$. The quality parameter t may depend on α (i.e., $t = t(\alpha)$). The same is true for digital sequences. In particular, [**10**, Theorem 4.10] implies that an order-α digital (t, n, d)-net is an order-α' digital (t', n, d)-net for all $1 \leqslant \alpha' \leqslant \alpha$ with

$$t' = \lceil t\alpha'/\alpha \rceil \leqslant t. \tag{4.3}$$

The same result applies to order-α digital (t, d)-sequences which are also order-α' digital (t', d)-sequences for $1 \leqslant \alpha' \leqslant \alpha$ and t' as above. In other words, $t(\alpha') = \lceil t(\alpha)\alpha'/\alpha \rceil$ for all $1 \leqslant \alpha' \leqslant \alpha$. For more details consult [**9**, **10**] or [**14**, Ch. 15].

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 [**14**, **29**] for the classical case $\alpha = 1$ and [**12**] for general α. See also Lemma 4.4 below.

Explicit constructions of order-2 digital sequences. Explicit constructions of order-α digital nets and sequences have been provided by Dick [**9**, **10**]. For our purposes it suffices to consider only $\alpha = 2$.

We start with generating matrices $C_1, \ldots, C_{2d}$ of a digital net or sequence. Let $\vec{c}_{j,k}$ denote the kth row of C_j. We now define matrices $E_1, \ldots, E_d$, where the kth 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 $d = 1$. In this case we have

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

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

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

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 [9, Theorems 4.11 and 4.12] we obtain the following result.

PROPOSITION 4.3. *If $C_1, \dots, C_{2d} \in \mathbb{F}_2^{\mathbb{N}\times\mathbb{N}}$ generate an order-1 digital $(t', 2d)$-sequence over $\mathbb{F}_2$, then $E_1, \dots, E_d \in \mathbb{F}_2^{\mathbb{N}\times\mathbb{N}}$ generate an order-2 digital (t, d)-sequence over $\mathbb{F}_2$ with*

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

Examples of explicit constructions of suitable generating matrices over $\mathbb{F}_2$ were obtained by Sobol' [37], Niederreiter [28, 29], Niederreiter and Xing [30] and others. An overview is presented in [14, Ch. 8]. Any of these constructions is sufficient for our purpose, however, for completeness, we briefly describe a special case of Tezuka's construction [39], which is a generalization of Sobol's construction [37] and Niederreiter's construction [28] of the generating matrices.

We explain how to construct the entries $c_{j,k,\ell} \in \mathbb{F}_2$ of the generator matrices $C_j = (c_{j,k,\ell})_{k,\ell\geqslant 1}$ for $j = 1, 2, \dots, d'$ (for our purpose $d' = 2d$). To this end choose the polynomials $p_j \in \mathbb{F}_2[x]$ for $j = 1, \dots, d'$ to be the jth 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, $e_1 \leqslant e_2 \leqslant \cdots \leqslant e_{d'}$ (the ordering of polynomials with the same degree is irrelevant).

Let $j \in \{1, \dots, d'\}$ 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 $0 \leqslant z < e_j$. 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

$$c_{j,k,\ell} = a_\ell(i, j, z). \tag{4.4}$$

Every digital sequence with generating matrices $C_j = (c_{j,k,\ell})_{k,\ell\geqslant 1}$ for $j = 1, 2, \dots, d'$ 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 [39, (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 $t = \sum_{j=1}^{d'}(e_j - 1)$, see [29, Ch. 4.5] for the case of Niederreiter sequences.

Remark. Let $C_1, \dots, C_{2d}$ be $\mathbb{N} \times \mathbb{N}$ matrices which are constructed according to Tezuka's method as described above. Let $E_1, \dots, E_d$ 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.

We point out that the explicit construction of the order-2 digital (t, d)-sequences $\mathcal{S}_d$ shown above, with generating matrices $E_i = (e_{i,k,\ell})_{k,\ell\geqslant 1}$ for

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

which $e_{i,k,\ell} = 0$ for all $k > 2\ell$ and for all $i \in \{1, 2, \dots, d\}$, does not depend on the parameters p, q, s in Theorems 2.2, 3.1 and 3.2. Our explicit construction, which is 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 $\mathcal{S}_{d+1}$. This means that 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 (cf. [**11**]).

4.2. *Some auxiliary results.* A connection between higher-order digital nets over $\mathbb{F}_2$ and dyadic intervals is given by the following result (see [**13**, Lemma 3.1]).

LEMMA 4.4. *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 already mentioned, 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 [**14**, **29**]). □

The following lemma, implicitly shown in [**13**], is utmost important for our analysis. Throughout the paper let ld denote the logarithm in base 2, i.e., $\operatorname{ld} x = (\log x)/\log 2$.

LEMMA 4.5. *Let $\mathcal{S}_d$ be an order-2 digital (t, d)-sequence over $\mathbb{F}_2$ with generating matrices $E_i = (e_{i,k,\ell})_{k,\ell \geqslant 1}$ for which $e_{i,k,\ell} = 0$ for all $k > 2\ell$ and for all $i \in \{1, 2, \dots, d\}$ and let $\mathcal{P}_{N,d}$ be the point set consisting of the first N elements of $\mathcal{S}_d$.*

(i) *If $|\boldsymbol{j}| + t/2 \geqslant \operatorname{ld} N$ we have: if $I_{\boldsymbol{j},\boldsymbol{m}}$ contains points of $\mathcal{P}_{N,d}$, then*

$$|\langle D^N_{\mathcal{S}_d}, h_{\boldsymbol{j},\boldsymbol{m}} \rangle| \ll 2^{t/2}\, N^{-1} 2^{-|\boldsymbol{j}|}.$$

At least $2^{|\boldsymbol{j}|} - N$ such intervals contain no points of $\mathcal{P}_{N,d}$ and in such cases we have

$$|\langle D^N_{\mathcal{S}_d}, h_{\boldsymbol{j},\boldsymbol{m}} \rangle| \ll 2^{-2|\boldsymbol{j}|}.$$

(ii) *If $|\boldsymbol{j}| + t/2 < \operatorname{ld} N$, more precisely $n_\mu \leqslant |\boldsymbol{j}| + t/2 < n_{\mu+1}$, where $N = 2^{n_r} + \cdots + 2^{n_1}$ with $n_r > \cdots > n_1 \geqslant 0$, then*

$$|\langle D^N_{\mathcal{S}_d}, h_{\boldsymbol{j},\boldsymbol{m}} \rangle| \ll 2^t\, N^{-1} (2^{-|\boldsymbol{j}|} + (2n_{\mu+1} - t - 2|\boldsymbol{j}|)^{d-1} 2^{-n_{\mu+1}}).$$

Proof. This is [**13**, equations (13), (14) and (16)]. □

LEMMA 4.6. *Let $r \in \mathbb{N}_0$ and $s \in \mathbb{N}$. Then*

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

For a proof of this result we refer to [**14**, Proof of Lemma 16.26].

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

LEMMA 4.7. *Let $K \in \mathbb{N}$, $A > 1$ and $q, s \geqslant 0$. 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 [26, 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 [25, Lemma 5.1].

LEMMA 4.8. *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 [25, Lemma 5.2].

LEMMA 4.9. *Let $\boldsymbol{z} = (z_1, \ldots, z_d) \in [0,1)^d$ and $g(\boldsymbol{x}) = \chi_{[\boldsymbol{0},\boldsymbol{x})}(\boldsymbol{z})$ 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 $|\langle g, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \ll 2^{-|\boldsymbol{j}|}$.*

In the case of finite digital nets over $\mathbb{F}_2$ the number of elements is a fixed power of 2, e.g., $N = 2^m$. For infinite digital sequences one needs discrepancy bounds for general $N \in \mathbb{N}$. The general plan is to split the initial N-element segments of the sequence into parts such that each part has a digital net structure (this method was also used in the proof of Lemma 4.5 in [13]). These parts are handled with methods developed for finite point sets. A similar approach was also used in [11, 15], however, the major challenge in these papers is the handling of the interactions between digital nets of different size, which arises from squaring the discrepancy function. This approach leads to a complicated proof for the L_2 discrepancy in [15] and even non-optimal results for the L_p discrepancy for general N in [11]. We overcome this problem by using a Haar series expansion of the discrepancy function which leads to a simpler bound on the various norms of the discrepancy function. See also [4, 13].

4.3. *The proof of the upper bound in Theorem 2.1.*

Proof of equation (2.3) in Theorem 2.1. The proof uses ideas from [4]. Let $N = 2^{n_r} + \cdots + 2^{n_1}$ with $n_r > \cdots > n_1 \geqslant 0$ and let $\mathcal{P}_{N,d}$ be the point set which consists of the first N terms of $\mathcal{S}_d$. We need to prove that

$$\lambda_d(U)^{-1} \sum_{\boldsymbol{j} \in \mathbb{N}_0^d} 2^{|\boldsymbol{j}|} \sum_{\substack{\boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}} \\ I_{\boldsymbol{j},\boldsymbol{m}} \subseteq U}} |\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^2 \ll_d 2^{2t} N^{-2} (\log N)^d.$$

It is clear that for each $U \subseteq [0,1)^d$ and $\boldsymbol{j} \in \mathbb{N}_0^d$ there are at most $2^{|\boldsymbol{j}|}\lambda_d(U)$ values of $\boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}}$ such that $I_{\boldsymbol{j},\boldsymbol{m}} \subseteq U$.

We split the above sum over $\boldsymbol{j} \in \mathbb{N}_0^d$ into three parts according to the size of $|\boldsymbol{j}|$.

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

First we consider those $\boldsymbol{j}$ for which $\operatorname{ld} N - t/2 \leqslant |\boldsymbol{j}| < \operatorname{ld} N$. From Lemma 4.5 we know that

$$|\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \ll 2^{t/2}\, N^{-1} 2^{-|\boldsymbol{j}|}$$

if $I_{\boldsymbol{j},\boldsymbol{m}}$ contains points of $\mathcal{P}_{N,d}$ and

$$|\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \ll 2^{-2|\boldsymbol{j}|}$$

otherwise. With these estimates we obtain

$$\begin{aligned}
&\lambda_d(U)^{-1} \sum_{\substack{\boldsymbol{j}\in\mathbb{N}_0^d \\ \operatorname{ld} N - t/2 \leqslant |\boldsymbol{j}| < \operatorname{ld} N}} 2^{|\boldsymbol{j}|} \sum_{\substack{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}} \\ I_{\boldsymbol{j},\boldsymbol{m}} \subseteq U}} |\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^2 \\
&\quad \ll \lambda_d(U)^{-1} \sum_{\substack{\boldsymbol{j}\in\mathbb{N}_0^d \\ \operatorname{ld} N - t/2 \leqslant |\boldsymbol{j}| < \operatorname{ld} N}} \\
&\qquad \times 2^{|\boldsymbol{j}|} \Bigg(\sum_{\substack{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}} \\ I_{\boldsymbol{j},\boldsymbol{m}} \subseteq U \\ I_{\boldsymbol{j},\boldsymbol{m}} \cap \mathcal{P}_{N,d} \neq \emptyset}} 2^t\, N^{-2} 2^{-2|\boldsymbol{j}|} + \sum_{\substack{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}} \\ I_{\boldsymbol{j},\boldsymbol{m}} \subseteq U \\ I_{\boldsymbol{j},\boldsymbol{m}} \cap \mathcal{P}_{N,d} = \emptyset}} 2^{-4|\boldsymbol{j}|} \Bigg) \\
&\quad \leqslant \lambda_d(U)^{-1} \sum_{\substack{\boldsymbol{j}\in\mathbb{N}_0^d \\ \operatorname{ld} N - t/2 \leqslant |\boldsymbol{j}| < \operatorname{ld} N}} 2^{|\boldsymbol{j}|} \max(2^t\, N^{-2} 2^{-2|\boldsymbol{j}|}, 2^{-4|\boldsymbol{j}|}) 2^{|\boldsymbol{j}|} \lambda_d(U) \\
&\quad \ll 2^t\, N^{-2} \sum_{\substack{\boldsymbol{j}\in\mathbb{N}_0^d \\ \operatorname{ld} N - t/2 \leqslant |\boldsymbol{j}| < \operatorname{ld} N}} 1 \ll 2^t\, N^{-2} \sum_{k < \operatorname{ld} N} k^{d-1} \ll 2^t\, N^{-2} (\log N)^d.
\end{aligned}$$

Now we consider those $\boldsymbol{j} \in \mathbb{N}_0^d$ for which $|\boldsymbol{j}| < \operatorname{ld} N - t/2$. More precisely assume that we have

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

for some $\mu \in \{0, 1, \ldots, r\}$, where we set $n_0 = 0$ and $n_{r+1} = \operatorname{ld} N$. From Lemma 4.5 we know that

$$|\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \ll 2^t\, N^{-1} (2^{-|\boldsymbol{j}|} + (2n_{\mu+1} - t - 2|\boldsymbol{j}|)^{d-1} 2^{-n_{\mu+1}}).$$

Then we have

$$\begin{aligned}
&\lambda_d(U)^{-1} \sum_{\substack{\boldsymbol{j}\in\mathbb{N}_0^d \\ |\boldsymbol{j}| < \operatorname{ld} N - t/2}} 2^{|\boldsymbol{j}|} \sum_{\substack{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}} \\ I_{\boldsymbol{j},\boldsymbol{m}} \subseteq U}} |\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^2 \\
&\quad \ll \sum_{\mu=0}^{r} \sum_{\substack{\boldsymbol{j}\in\mathbb{N}_0^d \\ n_\mu \leqslant |\boldsymbol{j}| + t/2 < n_{\mu+1}}} 2^{2|\boldsymbol{j}|}\, 2^{2t}\, N^{-2} (2^{-|\boldsymbol{j}|} + (2n_{\mu+1} - t - 2|\boldsymbol{j}|)^{d-1} 2^{-n_{\mu+1}})^2.
\end{aligned}$$

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

From here we can proceed in the same way as in [13, Proof of Theorem 2.2]. This way we obtain

$$\lambda_d(U)^{-1} \sum_{\substack{\boldsymbol{j}\in\mathbb{N}_0^d \\ |\boldsymbol{j}|<\operatorname{ld} N - t/2}} 2^{|\boldsymbol{j}|} \sum_{\substack{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}} \\ I_{\boldsymbol{j},\boldsymbol{m}}\subseteq U}} |\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^2 \ll 2^{2t}\, N^{-2} (\log N)^d.$$

Finally we need to study the case $|\boldsymbol{j}| \geqslant \operatorname{ld} N$ which is the most involved one among the three cases considered. As in [4] we treat the counting part and the volume part of the discrepancy function separately. For the volume part $L(\boldsymbol{x}) = x_1 \cdots x_d$ it is easy to show that $|\langle L, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \asymp 2^{-2|\boldsymbol{j}|}$. Thus we have

$$\begin{aligned}\lambda_d(U)^{-1} \sum_{\substack{\boldsymbol{j}\in\mathbb{N}_0^d \\ |\boldsymbol{j}|\geqslant \operatorname{ld} N}} 2^{|\boldsymbol{j}|} \sum_{\substack{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}} \\ I_{\boldsymbol{j},\boldsymbol{m}}\subseteq U}} |\langle L, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^2 &\ll \sum_{\substack{\boldsymbol{j}\in\mathbb{N}_0^d \\ |\boldsymbol{j}|\geqslant \operatorname{ld} N}} 2^{-2|\boldsymbol{j}|} \\ &\ll \sum_{k\geqslant \operatorname{ld} N} k^{d-1}\, 2^{-2k} \ll N^{-2} (\log N)^{d-1}.\end{aligned}$$

Let $C(\boldsymbol{x}) = N^{-1} \sum_{\boldsymbol{z}\in\mathcal{P}_{N,d}} \chi_{[\boldsymbol{0},\boldsymbol{x})}(\boldsymbol{z})$ be the counting part of the discrepancy function.

Let $\mathcal{J}$ be the family of all dyadic intervals $I_{\boldsymbol{j},\boldsymbol{m}} \subseteq U$ with $|\boldsymbol{j}| \geqslant \operatorname{ld} N$ such that $\langle C, h_{\boldsymbol{j},\boldsymbol{m}}\rangle \neq 0$. Consider the subfamily $\widetilde{\mathcal{J}} \subseteq \mathcal{J}$ of all maximal (with respect to inclusion) dyadic intervals in $\mathcal{J}$. We show that we have

$$\sum_{I_{\boldsymbol{j},\boldsymbol{m}}\in\widetilde{\mathcal{J}}} \lambda_d(I_{\boldsymbol{j},\boldsymbol{m}}) \ll (\log N)^{d-1} \lambda_d(U). \tag{4.5}$$

To see this consider an interval $I_{\boldsymbol{j},\boldsymbol{m}} \in \mathcal{J}$. The condition $\langle C, h_{\boldsymbol{j},\boldsymbol{m}}\rangle \neq 0$ implies that the interior of $I_{\boldsymbol{j},\boldsymbol{m}}$ contains at least one point of $\mathcal{P}_{N,d}$. Now we need the assumption that the generating matrices $E_i = (e_{i,k,\ell})_{k,\ell\geqslant 1}$ are chosen such that $e_{i,k,\ell} = 0$ for all $k > 2\ell$ and all $i \in \{1, 2, \ldots, d\}$. This implies that the points of $\mathcal{P}_{N,d}$ have binary coordinates of length $2n$ (where $n = \lfloor \operatorname{ld} N \rfloor + 1$). Since the interior of $I_{\boldsymbol{j},\boldsymbol{m}}$ contains at least one point of $\mathcal{P}_{N,d}$ we obtain that each side of $I_{\boldsymbol{j},\boldsymbol{m}}$ has length at least 2^{-2n}, i.e., $j_i \in \{0, 1, \ldots, 2n\}$ for all $i \in \{1, 2, \ldots, d\}$.

Now fix $r_1, r_2, \ldots, r_{d-1} \in \{0, 1, \ldots, 2n\}$ and consider the family $\widetilde{\mathcal{J}}_{r_1,\ldots,r_{d-1}} \subseteq \widetilde{\mathcal{J}}$ of those intervals $I_{\boldsymbol{j},\boldsymbol{m}} \in \mathcal{J}$ for which $j_i = r_i$ for $i \in \{1, 2, \ldots, d-1\}$ (i.e., with fixed lengths of their first $d-1$ sides). If two of these intervals have a non-empty intersection, then their first $d-1$ sides have to coincide and hence one would have to be contained in the other. This however contradicts the maximality of the intervals in $\widetilde{\mathcal{J}}$. Thus we have shown that the intervals in $\widetilde{\mathcal{J}}_{r_1,\ldots,r_{d-1}}$ are pairwise disjoint. From this observation we find that

$$\begin{aligned}\sum_{I_{\boldsymbol{j},\boldsymbol{m}}\in\widetilde{\mathcal{J}}} \lambda_d(I_{\boldsymbol{j},\boldsymbol{m}}) &= \sum_{r_1,\ldots,r_{d-1}=0}^{2n} \sum_{I_{\boldsymbol{j},\boldsymbol{m}}\in\widetilde{\mathcal{J}}_{r_1,\ldots,r_{d-1}}} \lambda_d(I_{\boldsymbol{j},\boldsymbol{m}}) \\ &\leqslant \sum_{r_1,\ldots,r_{d-1}=0}^{2n} \lambda_d(U) \ll (2n)^{d-1} \lambda_d(U),\end{aligned}$$

which implies (4.5).

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

For dyadic intervals J, we define

$$C^J(\boldsymbol{x}) = \sum_{\boldsymbol{z}\in\mathcal{P}_{N,d}\cap J} \chi_{[\boldsymbol{0},\boldsymbol{x})}(\boldsymbol{z}).$$

It is easy to see that $\langle C, h_{j,\boldsymbol{m}}\rangle = \langle C^J, h_{j,\boldsymbol{m}}\rangle$ whenever $I_{j,\boldsymbol{m}} \subseteq J$. From $n = \lfloor \mathrm{ld}\, N\rfloor + 1$ we know that $N < 2^n$. 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 4.4, in an interval $I_{j,\boldsymbol{m}}$ 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 $I_{j,\boldsymbol{m}}$. Therefore, for any $J \in \widetilde{\mathcal{J}}$ we have $0 \leqslant C^J(\boldsymbol{x}) \leqslant 2^{\lceil t/2\rceil} N^{-1}$, which implies

$$\|C^J\|_{L_2(J)} \leqslant \|2^{\lceil t/2\rceil} N^{-1}\|_{L_2(J)} \leqslant N^{-1} 2^{\lceil t/2\rceil} \lambda_d(J)^{1/2}. \tag{4.6}$$

Using the orthogonality of Haar functions, Bessel's inequality, (4.5), and (4.6), we obtain

$$\begin{aligned}
\lambda_d(U)^{-1} \sum_{I_{j,\boldsymbol{m}}\in\mathcal{J}} 2^{|j|} |\langle C, h_{j,\boldsymbol{m}}\rangle|^2 &\leqslant \lambda_d(U)^{-1} \sum_{J\in\widetilde{\mathcal{J}}} \sum_{I_{j,\boldsymbol{m}}\subseteq J} 2^{|j|} |\langle C^J, h_{j,\boldsymbol{m}}\rangle|^2 \\
&\leqslant \lambda_d(U)^{-1} \sum_{J\in\widetilde{\mathcal{J}}} \|C^J\|_{L_2(J)}^2 \\
&\leqslant \lambda_d(U)^{-1}\, 2^{t+1}\, N^{-2} \sum_{J\in\widetilde{\mathcal{J}}} \lambda_d(J) \\
&\ll 2^{t+1}\, N^{-2} (\log N)^{d-1}.
\end{aligned}$$

The result now follows by putting the three parts together and taking square roots. □

4.4. *The proof of Theorem 2.2.*

Proof of Theorem 2.2. We prove now the bound on the exponential Orlicz norm of the discrepancy function. Our proof uses ideas from [4]. Let again $N = 2^{n_r} + \cdots + 2^{n_1}$ with $n_r > \cdots > n_1 \geqslant 0$ and let $\mathcal{P}_{N,d}$ be the point set which consists of the first N terms of $\mathcal{S}_d$. We first restrict ourselves to the $\exp(L^{2/(d-1)})$-norm. The general result will follow directly by interpolating between the $\exp(L^{2/(d-1)})$-norm and the L_∞-norm using Proposition 2.6 and the fact that for digital sequences we have

$$\|D_{\mathcal{S}_d}^N | L_\infty([0,1)^d)\| \ll_d 2^t\, N^{-1} (\log N)^d,$$

which follows from the fact that $\mathcal{S}_d$ is a digital (t',d)-sequence (i.e., order-1 digital sequence) and [29, Theorem 4.13].

We first use the Haar series expansion for the discrepancy function, the triangle inequality and the Chang–Wilson–Wolff inequality (Lemma 2.5)

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

to obtain

$$\begin{aligned}
&\|D_{\mathcal{S}_d}^N|\exp(L^{2/(d-1)})\| \\
&\quad= \left\| \sum_{k=0}^{\infty} \sum_{\substack{\boldsymbol{j}\in\mathbb{N}_{-1}^d \\ |\boldsymbol{j}|=k}} 2^{|\boldsymbol{j}|} \sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}} \langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle h_{\boldsymbol{j},\boldsymbol{m}} \middle| \exp(L^{2/(d-1)}) \right\| \\
&\quad\leqslant \sum_{k=0}^{\infty} \left\| \sum_{\substack{\boldsymbol{j}\in\mathbb{N}_{-1}^d \\ |\boldsymbol{j}|=k}} 2^{|\boldsymbol{j}|} \sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}} \langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle h_{\boldsymbol{j},\boldsymbol{m}} \middle| \exp(L^{2/(d-1)}) \right\| \\
&\quad\leqslant \sum_{k=0}^{\infty} \left\| \left(\sum_{\substack{\boldsymbol{j}\in\mathbb{N}_{-1}^d \\ |\boldsymbol{j}|=k}} 2^{2|\boldsymbol{j}|} \sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}} |\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^2 \chi_{I_{\boldsymbol{j},\boldsymbol{m}}} \right)^{1/2} \middle| L_\infty([0,1)^d) \right\|. \qquad (4.7)
\end{aligned}$$

To prove a bound on (4.7), we divide the sum over k into three parts, i.e., we consider the L_∞-norm of the square function where $|\boldsymbol{j}|$ lies in a certain range.

For $\operatorname{ld} N - t/2 \leqslant |\boldsymbol{j}| < \operatorname{ld} N$, we have from Lemma 4.5 that

$$|\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \ll 2^{t/2}\, N^{-1} 2^{-|\boldsymbol{j}|}$$

if $I_{\boldsymbol{j},\boldsymbol{m}}$ contains points of $\mathcal{P}_{N,d}$ and

$$|\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \ll 2^{-2|\boldsymbol{j}|}$$

otherwise. Hence, for $\operatorname{ld} N - t/2 \leqslant k < \operatorname{ld} N$ we have

$$\begin{aligned}
&\sup_{\boldsymbol{x}\in[0,1)^d} \left(\sum_{\substack{\boldsymbol{j}\in\mathbb{N}_{-1}^d \\ |\boldsymbol{j}|=k}} 2^{2|\boldsymbol{j}|} \sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}} |\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^2 \chi_{I_{\boldsymbol{j},\boldsymbol{m}}} \right)^{1/2} \\
&\qquad\ll \left(\sum_{\substack{\boldsymbol{j}\in\mathbb{N}_{-1}^d \\ |\boldsymbol{j}|=k}} \max\{2^t\, N^{-2}, 2^{-2|\boldsymbol{j}|}\} \right)^{1/2} \\
&\qquad\leqslant \left(2^t\, N^{-2} \sum_{\substack{\boldsymbol{j}\in\mathbb{N}_{-1}^d \\ |\boldsymbol{j}|=k}} 1 \right)^{1/2} \ll 2^{t/2}\, N^{-1} (\log N)^{(d-1)/2}.
\end{aligned}$$

Therefore

$$\begin{aligned}
&\sum_{\substack{k\in\mathbb{N}_0 \\ \operatorname{ld} N - t/2 \leqslant k < \operatorname{ld} N}} \left\| \left(\sum_{\substack{\boldsymbol{j}\in\mathbb{N}_{-1}^d \\ |\boldsymbol{j}|=k}} 2^{2|\boldsymbol{j}|} \sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}} |\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^2 \chi_{I_{\boldsymbol{j},\boldsymbol{m}}} \right)^{1/2} \middle| L_\infty([0,1)^d) \right\| \\
&\qquad\ll 2^t\, N^{-1} (\log N)^{(d-1)/2}.
\end{aligned}$$

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

Next we consider the case where $0 \leqslant |\boldsymbol{j}| < \operatorname{ld} N - t/2$. Assume that we have

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

for some $\mu \in \{0, 1, \dots, r\}$, where we set $n_{r+1} = \operatorname{ld} N$ and $n_0 = 0$. From Lemma 4.5 we have that

$$|\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \ll 2^t\, N^{-1}(2^{-|\boldsymbol{j}|} + (2n_{\mu+1} - t - 2|\boldsymbol{j}|)^{d-1}2^{-n_{\mu+1}}).$$

Hence, for $n_\mu \leqslant k + t/2 < n_{\mu+1}$, we have

$$\begin{aligned}
&\sup_{\boldsymbol{x}\in[0,1)^d}\Bigg(\sum_{\substack{\boldsymbol{j}\in\mathbb{N}_{-1}^d\\|\boldsymbol{j}|=k}} 2^{2|\boldsymbol{j}|}\sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}} |\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^2\chi_{I_{\boldsymbol{j},\boldsymbol{m}}}\Bigg)^{1/2}\\
&\qquad \ll 2^t\, N^{-1}\Bigg(\sum_{\substack{\boldsymbol{j}\in\mathbb{N}_{-1}^d\\|\boldsymbol{j}|=k}} 2^{2|\boldsymbol{j}|}(2^{-|\boldsymbol{j}|} + (2n_{\mu+1} - t - 2|\boldsymbol{j}|)^{d-1}2^{-n_{\mu+1}})^2\Bigg)^{1/2},
\end{aligned}$$

and therefore we need to estimate

$$\begin{aligned}
&\sum_{\substack{k\in\mathbb{N}_0\\k\leqslant \operatorname{ld} N - t/2}} \Bigg\|\Bigg(\sum_{\substack{\boldsymbol{j}\in\mathbb{N}_{-1}^d\\|\boldsymbol{j}|=k}} 2^{2|\boldsymbol{j}|}\sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}} |\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^2\chi_{I_{\boldsymbol{j},\boldsymbol{m}}}\Bigg)^{1/2}\Bigg| L_\infty([0,1)^d)\Bigg\|\\
&\quad \ll 2^t\, N^{-1}\sum_{\mu=0}^{r}\sum_{\substack{k\in\mathbb{N}_0\\ n_\mu\leqslant k+t/2<n_{\mu+1}}}\\
&\qquad \times\Bigg(\sum_{\substack{\boldsymbol{j}\in\mathbb{N}_{-1}^d\\|\boldsymbol{j}|=k}} 2^{2|\boldsymbol{j}|}(2^{-|\boldsymbol{j}|} + (2n_{\mu+1} - t - 2|\boldsymbol{j}|)^{d-1}2^{-n_{\mu+1}})^2\Bigg)^{1/2}\\
&\quad \ll 2^t\, N^{-1}\sum_{\mu=0}^{r}\\
&\qquad \times\sum_{\substack{k\in\mathbb{N}_0\\ n_\mu\leqslant k+t/2<n_{\mu+1}}} (k^{d-1}2^{2k}(2^{-k} + (2n_{\mu+1} - t - 2k)^{d-1}2^{-n_{\mu+1}})^2)^{1/2}\\
&\quad = 2^t\, N^{-1}\sum_{\mu=0}^{r}\sum_{\substack{k\in\mathbb{N}_0\\ n_\mu\leqslant k+t/2<n_{\mu+1}}} k^{(d-1)/2}(1 + (2n_{\mu+1} - t - 2k)^{d-1}2^{-n_{\mu+1}-k})\\
&\quad \ll 2^t\, N^{-1}\sum_{k=0}^{\lceil \operatorname{ld} N\rceil} k^{(d-1)/2} + 2^t\, N^{-1} r\sum_{\ell=0}^{\lceil \operatorname{ld} N\rceil}(2\ell - t)^{d-1}2^{-\ell}\\
&\quad \ll 2^t\, N^{-1}(\log N)^{(d+1)/2}.
\end{aligned}$$

It remains to estimate the terms where $k \geqslant \operatorname{ld} N$. As in [4] we treat the volume part of the discrepancy function separately from the counting part.

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

For the volume part $L(\boldsymbol{x}) = x_1 \cdots x_d$ we have $|\langle L, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \asymp 2^{-2|\boldsymbol{j}|}$. Hence

$$\begin{aligned}
&\left\| \sum_{\substack{k=0\\ k\geqslant \mathrm{ld}\, N}}^{\infty} \sum_{\substack{\boldsymbol{j}\in\mathbb{N}_{-1}^d\\ |\boldsymbol{j}|=k}} 2^{|\boldsymbol{j}|} \sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}} \langle L, h_{\boldsymbol{j},\boldsymbol{m}}\rangle h_{\boldsymbol{j},\boldsymbol{m}} \,\middle|\, \exp(L^{2/(d-1)}) \right\| \\
&\quad\leqslant \sum_{\substack{k=0\\ k\geqslant \mathrm{ld}\, N}}^{\infty} \left\| \sum_{\substack{\boldsymbol{j}\in\mathbb{N}_{-1}^d\\ |\boldsymbol{j}|=k}} 2^{|\boldsymbol{j}|} \sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}} \langle L, h_{\boldsymbol{j},\boldsymbol{m}}\rangle h_{\boldsymbol{j},\boldsymbol{m}} \,\middle|\, \exp(L^{2/(d-1)}) \right\| \\
&\quad\leqslant \sum_{\substack{k=0\\ k\geqslant \mathrm{ld}\, N}}^{\infty} \left\| \left(\sum_{\substack{\boldsymbol{j}\in\mathbb{N}_{-1}^d\\ |\boldsymbol{j}|=k}} 2^{2|\boldsymbol{j}|} \sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}} |\langle L, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^2 \chi_{I_{\boldsymbol{j},\boldsymbol{m}}} \right)^{1/2} \,\middle|\, L_\infty([0,1)^d) \right\| \\
&\quad\ll \sum_{\substack{k=0\\ k\geqslant \mathrm{ld}\, N}}^{\infty} \left(\sum_{\substack{\boldsymbol{j}\in\mathbb{N}_{-1}^d\\ |\boldsymbol{j}|=k}} 2^{-2|\boldsymbol{j}|} \right)^{1/2} \ll \sum_{\substack{k=0\\ k\geqslant \mathrm{ld}\, N}}^{\infty} k^{(d-1)/2} 2^{-k} \ll N^{-1} (\log N)^{(d-1)/2}.
\end{aligned}$$

It remains to estimate the counting part $C(\boldsymbol{x}) = N^{-1} \sum_{\boldsymbol{z}\in\mathcal{P}_{N,d}} \chi_{[\boldsymbol{0},\boldsymbol{x})}(\boldsymbol{z})$. Let $n \in \mathbb{N}$ be such that $2^{n-1} < N \leqslant 2^n$. Let again $\mathcal{J}$ be the set of all dyadic intervals $I_{\boldsymbol{j},\boldsymbol{m}}$ with $|\boldsymbol{j}| \geqslant n$, i.e., $|I_{\boldsymbol{j},\boldsymbol{m}}| \leqslant 2^{-n}$, such that $\langle C, h_{\boldsymbol{j},\boldsymbol{m}}\rangle \neq 0$. As above, this implies that for $I_{\boldsymbol{j},\boldsymbol{m}} \in \mathcal{J}$ we have that at least one point of $\mathcal{P}_{N,d}$ lies in the interior of $I_{\boldsymbol{j},\boldsymbol{m}}$. Hence $j_k \leqslant 2n$, or in other words $|I_{j_k,m_k}| \geqslant 2^{-2n}$, for each coordinate $k = 1, 2, \ldots, d$.

We also define the unique parent of each $I_{\boldsymbol{j},\boldsymbol{m}} \in \mathcal{J}$, denoted by $\widetilde{I}_{\boldsymbol{j},\boldsymbol{m}}$, which satisfies: (i) $I_{\boldsymbol{j},\boldsymbol{m}} \subseteq \widetilde{I}_{\boldsymbol{j}',\boldsymbol{m}'}$; (ii) $|\boldsymbol{j}'| = n$, i.e., $|\widetilde{I}_{\boldsymbol{j}',\boldsymbol{m}'}| = 2^{-n}$; and (iii) $j_k = j_k'$ (which implies that $I_{j_k,m_k} = \widetilde{I}_{j_k',m_k'}$) for all $k = 1, 2, \ldots, d$. In other words, to find the parent, we expand the dth side of $I_{\boldsymbol{j},\boldsymbol{m}}$ to a dyadic interval such that the volume of the resulting interval has volume 2^{-n}. By reordering the sum over the intervals with respect to their parents we obtain

$$\sum_{\substack{\boldsymbol{j}\in\mathbb{N}_{-1}^d\\ |\boldsymbol{j}|\geqslant n}} 2^{|\boldsymbol{j}|} \sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}} \langle C, h_{\boldsymbol{j},\boldsymbol{m}}\rangle h_{\boldsymbol{j},\boldsymbol{m}} = \sum_{\substack{\widetilde{I}_{\boldsymbol{j}',\boldsymbol{m}'}:|\boldsymbol{j}'|=n\\ j_k'\leqslant 2n: k=1,\ldots,d}} \sum_{\substack{I_{\boldsymbol{j},\boldsymbol{m}}\subseteq\widetilde{I}_{\boldsymbol{j}',\boldsymbol{m}'}\\ j_k=j_k': k=1,\ldots,d-1}} 2^{|\boldsymbol{j}|} \langle C, h_{\boldsymbol{j},\boldsymbol{m}}\rangle h_{\boldsymbol{j},\boldsymbol{m}}. \tag{4.8}$$

For an arbitrary, but fixed, parent interval $\widetilde{I}_{\boldsymbol{j}',\boldsymbol{m}'}$ we consider the innermost sum above

$$\begin{aligned}
&\sum_{\substack{I_{\boldsymbol{j},\boldsymbol{m}}\subseteq\widetilde{I}_{\boldsymbol{j}',\boldsymbol{m}'}\\ j_k=j_k': k=1,\ldots,d-1}} 2^{|\boldsymbol{j}|} \langle C, h_{\boldsymbol{j},\boldsymbol{m}}\rangle h_{\boldsymbol{j},\boldsymbol{m}} \\
&\quad= N^{-1} \sum_{\boldsymbol{z}\in\mathcal{P}_{N,d}\cap\widetilde{I}_{\boldsymbol{j}',\boldsymbol{m}'}} \sum_{\substack{I_{\boldsymbol{j},\boldsymbol{m}}\subseteq\widetilde{I}_{\boldsymbol{j}',\boldsymbol{m}'}\\ j_k=j_k': k=1,\ldots,d-1}} 2^{|\boldsymbol{j}|} \langle \chi_{[\boldsymbol{z},\boldsymbol{1})}, h_{\boldsymbol{j},\boldsymbol{m}}\rangle h_{\boldsymbol{j},\boldsymbol{m}}.
\end{aligned} \tag{4.9}$$

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

The summands in the last sum split into products of one-dimensional factors

$$\begin{aligned}
&2^{|j|}\langle \chi_{[z,\mathbf{1})}, h_{\boldsymbol{j},\boldsymbol{m}}\rangle h_{\boldsymbol{j},\boldsymbol{m}}(\boldsymbol{x}) \\
&\quad = \prod_{k=1}^{d} 2^{j_k}\langle \chi_{[z_k,1)}, h_{j_k,m_k}\rangle h_{j_k,m_k}(x_k) \\
&\quad = \left(\prod_{k=1}^{d-1} 2^{j'_k}\langle \chi_{[z_k,1)}, h_{j'_k,m'_k}\rangle h_{j'_k,m'_k}(x_k)\right) 2^{j_d}\langle \chi_{[z_d,1)}, h_{j_d,m_d}\rangle h_{j_d,m_d}(x_d) \\
&\quad = 2^{|j'_*|}\langle \chi_{[z_*,\mathbf{1})}, h_{\boldsymbol{j}'_*,\boldsymbol{m}'_*}\rangle 2^{j_d}\langle \chi_{[z_d,1)}, h_{j_d,m_d}\rangle h_{j_d,m_d}(x_d),
\end{aligned}$$

where by $*$ we denote the projection of a d-dimensional vector to its first $d-1$ coordinates (i.e., for instance, $z_* = (z_1, \dots, z_{d-1})$ for $z = (z_1, \dots, z_{d-1}, z_d)$). Thus we can write the innermost sum on the right-hand side of (4.9) as

$$\begin{aligned}
&\sum_{\substack{I_{j,m}\subseteq \widetilde{I}_{j',m'} \\ j_k = j'_k : k=1,\dots,d-1}} 2^{|j|}\langle \chi_{[z,\mathbf{1})}, h_{\boldsymbol{j},\boldsymbol{m}}\rangle h_{\boldsymbol{j},\boldsymbol{m}}(\boldsymbol{x}) \\
&\quad = 2^{|j'_*|}\langle \chi_{[z_*,\mathbf{1})}, h_{\boldsymbol{j}'_*,\boldsymbol{m}'_*}\rangle h_{\boldsymbol{j}'_*,\boldsymbol{m}'_*}(x_*) \sum_{I_{j_d,m_d}\subseteq \widetilde{I}_{j'_d,m'_d}} 2^{j_d}\langle \chi_{[z_d,1)}, h_{j_d,m_d}\rangle h_{j_d,m_d}(x_d).
\end{aligned}$$

The last sum is now the Haar series expansion of the indicator function $\chi_{[z_d,1)}$ restricted to the interval $\widetilde{I}_{j'_d,m'_d}$ without the constant term, i.e., we have

$$\begin{aligned}
&\sum_{I_{j_d,m_d}\subseteq \widetilde{I}_{j'_d,m'_d}} 2^{j_d}\langle \chi_{[z_d,1)}, h_{j_d,m_d}\rangle h_{j_d,m_d}(x_d) \\
&\qquad = \chi_{\widetilde{I}_{j'_d,m'_d}}(x_d)(\chi_{[z_d,1)}(x_d) - 2^{j'_d}|[z_d,1)\cap \widetilde{I}_{j'_d,m'_d}|). \qquad (4.10)
\end{aligned}$$

Using the triangle inequality it follows that the last expression is bounded pointwise by 2. Further we have $|2^{|j'_*|}\langle \chi_{[z_*,\mathbf{1})}, h_{\boldsymbol{j}'_*,\boldsymbol{m}'_*}\rangle| \leqslant 1$. Since there are at most $2^{\lceil t/2\rceil}$ points of $\mathcal{P}_{N,d}$ in $\widetilde{I}_{\boldsymbol{j}',\boldsymbol{m}'}$, we obtain that

$$\sum_{\substack{I_{j,m}\subseteq \widetilde{I}_{j',m'} \\ j_k = j'_k : k=1,\dots,d-1}} 2^{|j|}\langle C, h_{\boldsymbol{j},\boldsymbol{m}}\rangle h_{\boldsymbol{j},\boldsymbol{m}}(\boldsymbol{x}) = N^{-1}\alpha_{j'_d}(x_d) h_{\boldsymbol{j}'_*,\boldsymbol{m}'_*}(\boldsymbol{x}_*),$$

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

Let now x_d be fixed. Equation (4.10) implies that for a given $(d-1)$-dimensional interval $\widetilde{I}_{\boldsymbol{j}'_*,\boldsymbol{m}'_*}$ there exists only one d-dimensional interval of the form $\widetilde{I}_{\boldsymbol{j}',\boldsymbol{m}'} = \widetilde{I}_{\boldsymbol{j}'_*,\boldsymbol{m}'_*} \times \widetilde{I}_{j'_d,m'_d}$ with $|\widetilde{I}_{\boldsymbol{j}',\boldsymbol{m}'}| = 2^{-n}$ such that $\alpha_{j'_d}(x_d) \neq 0$. We can now apply (4.8) and take the L_p-norm in the first $d-1$ variables to obtain

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

$$\begin{aligned}
&\left\| \sum_{\substack{\boldsymbol{j}\in\mathbb{N}_{-1}^d\\ |\boldsymbol{j}|\geqslant n}} 2^{|\boldsymbol{j}|} \sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}} \langle C, h_{\boldsymbol{j},\boldsymbol{m}}\rangle h_{\boldsymbol{j},\boldsymbol{m}} \right| L_p\left([0,1)^{d-1}\right) \Bigg\| \\
&\quad = \left\| \sum_{\substack{\widetilde{I}_{\boldsymbol{j}',\boldsymbol{m}'}:|\boldsymbol{j}'|=n\\ j_k'\leqslant 2n:k=1,\dots,d}} \sum_{\substack{I_{\boldsymbol{j},\boldsymbol{m}}\subseteq\widetilde{I}_{\boldsymbol{j}',\boldsymbol{m}'}\\ j_k=j_k':k=1,\dots,d-1}} 2^{|\boldsymbol{j}|}\langle C, h_{\boldsymbol{j},\boldsymbol{m}}\rangle h_{\boldsymbol{j},\boldsymbol{m}} \right| L_p([0,1)^{d-1}) \Bigg\| \\
&\quad = N^{-1} \left\| \sum_{\substack{\widetilde{I}_{\boldsymbol{j}'_*,\boldsymbol{m}'_*}:|\boldsymbol{j}'|=n\\ j_k'\leqslant 2n:k=1,\dots,d-1}} \alpha_{j_d'}(x_d) h_{\boldsymbol{j}'_*,\boldsymbol{m}'_*} \right| L_p\left([0,1)^{d-1}\right) \Bigg\| \\
&\quad \ll p^{(d-1)/2} N^{-1} \left\| \left(\sum_{\substack{\widetilde{I}_{\boldsymbol{j}'_*,\boldsymbol{m}'_*}:|\boldsymbol{j}'|=n\\ j_k'\leqslant 2n:k=1,\dots,d-1}} |\alpha_{j_d'}(x_d)|^2 \chi_{\widetilde{I}_{\boldsymbol{j}'_*,\boldsymbol{m}'_*}} \right)^{1/2} \right| L_p([0,1)^{d-1}) \Bigg\| \\
&\quad \ll p^{(d-1)/2} n^{(d-1)/2} N^{-1},
\end{aligned}$$

where in the penultimate step we have employed the $(d-1)$-dimensional Littlewood–Paley inequality (see Lemma 2.4) and in the last step the fact that the number of choices of $\boldsymbol{j}'_*$ is of order n^{d-1} at most.

By integrating this bound with respect to the last variable x_d and using Proposition 2.3 we obtain

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

from which the result now follows. □

4.5. *The proof of the upper bounds in Theorem 3.1.*

Proof of equations (3.4) and (3.6) in Theorem 3.1. According to Proposition 3.4 we have

$$\|D_{\mathcal{S}_d}^N | S_{p,q}^s B\|^q \ll_{p,q,s,d} \sum_{\boldsymbol{j}\in\mathbb{N}_{-1}^d} 2^{|\boldsymbol{j}|(s-1/p+1)q} \left(\sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}} |\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^p \right)^{q/p}. \quad (4.11)$$

Let $|\boldsymbol{j}| + t/2 \geqslant \operatorname{ld} N$, then according to Lemma 4.5 we have

$$|\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \ll 2^{t/2} N^{-1} 2^{-|\boldsymbol{j}|}$$

and for a fixed $\boldsymbol{j}$, at least $2^{|\boldsymbol{j}|} - N$ intervals $I_{\boldsymbol{j},\boldsymbol{m}}$ contain no points of $\mathcal{P}_{N,d}$ (the point set consisting of the first N elements of $\mathcal{S}_d$) and in such cases we have according to Lemma 4.5

$$|\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \ll 2^{-2|\boldsymbol{j}|}.$$

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

Now we estimate (4.11) by applying Minkowski's inequality to obtain

$$\begin{aligned}
&\sum_{|\boldsymbol{j}|+t/2\geqslant \operatorname{ld} N} 2^{|\boldsymbol{j}|(s-1/p+1)q}\left(\sum_{\boldsymbol{m}\in\mathbb{D}_j}|\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^p\right)^{q/p}\\
&\quad\ll \sum_{|\boldsymbol{j}|+t/2\geqslant \operatorname{ld} N} 2^{|\boldsymbol{j}|(s-1/p+1)q}(N2^{pt/2}N^{-p}2^{-p|\boldsymbol{j}|})^{q/p}\\
&\qquad+\sum_{|\boldsymbol{j}|+t/2\geqslant \operatorname{ld} N} 2^{|\boldsymbol{j}|(s-1/p+1)q}(\max\{2^{|\boldsymbol{j}|}-N,0\}2^{-2p|\boldsymbol{j}|})^{q/p}\\
&\quad\leqslant N^{q/p-q}2^{qt/2}\sum_{|\boldsymbol{j}|+t/2\geqslant \operatorname{ld} N} 2^{|\boldsymbol{j}|(s-1/p)q}+\sum_{|\boldsymbol{j}|+t/2\geqslant \operatorname{ld} N} 2^{|\boldsymbol{j}|(s-1)q}\\
&\quad\ll N^{q/p-q}2^{qt/2}N^{(s-1/p)q}2^{-t/2(s-1/p)q}(\log N)^{d-1}\\
&\qquad+N^{(s-1)q}2^{-t/2(s-1)q}(\log N)^{d-1}\\
&\quad\leqslant 2^{(1-s+1/p)tq/2}N^{(s-1)q}(\log N)^{d-1}. \qquad (4.12)
\end{aligned}$$

Now let $n_r > \cdots > n_{\mu+1} \geqslant |\boldsymbol{j}| + t/2 > n_\mu > \cdots > n_1$. Then according to Lemma 4.5 we have

$$|\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \ll 2^t N^{-1}(2^{-|\boldsymbol{j}|}+(2n_{\mu+1}-t-2|\boldsymbol{j}|)^{d-1}2^{-n_{\mu+1}}).$$

We use this bound to estimate the terms of (4.11) for which $|\boldsymbol{j}| + t/2 < \operatorname{ld} N$ (again we set $n_0 = 0$, $n_{r+1} = \operatorname{ld} N$). We use Minkowski's inequality

$$\begin{aligned}
&\sum_{|\boldsymbol{j}|+t/2<\operatorname{ld} N} 2^{|\boldsymbol{j}|(s-1/p+1)q}\left(\sum_{\boldsymbol{m}\in\mathbb{D}_j}|\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^p\right)^{q/p}\\
&\quad=\sum_{\mu=0}^{r}\sum_{n_\mu\leqslant|\boldsymbol{j}|+t/2<n_{\mu+1}} 2^{|\boldsymbol{j}|(s-1/p+1)q}\left(\sum_{\boldsymbol{m}\in\mathbb{D}_j}|\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^p\right)^{q/p}\\
&\quad\ll\sum_{\mu=0}^{r}\sum_{n_\mu\leqslant|\boldsymbol{j}|+t/2<n_{\mu+1}} 2^{|\boldsymbol{j}|(s-1/p+1)q}2^{q|\boldsymbol{j}|/p}\,2^{qt}\,N^{-q}\\
&\qquad\times(2^{-|\boldsymbol{j}|}+(2n_{\mu+1}-t-2|j|)^{d-1}2^{-n_{\mu+1}})^{\lceil q\rceil}\\
&\quad=2^{qt}\,N^{-q}\sum_{\mu=0}^{r}\sum_{n_\mu\leqslant|\boldsymbol{j}|+t/2<n_{\mu+1}} 2^{|\boldsymbol{j}|(s+1)q}\\
&\qquad\times\sum_{i=0}^{\lceil q\rceil}2^{-|\boldsymbol{j}|i}2^{-n_{\mu+1}(\lceil q\rceil-i)}(2n_{\mu+1}-t-2|\boldsymbol{j}|)^{(d-1)(\lceil q\rceil-i)}\\
&\quad\leqslant 2^{qt}\,N^{-q}\sum_{\mu=0}^{r}\sum_{i=0}^{\lceil q\rceil}2^{-n_{\mu+1}(\lceil q\rceil-i)}\\
&\qquad\times\sum_{n_\mu\leqslant|\boldsymbol{j}|+t/2<n_{\mu+1}} 2^{|\boldsymbol{j}|(sq+\lceil q\rceil-i)}(2n_{\mu+1}-t-2|\boldsymbol{j}|)^{(d-1)(\lceil q\rceil-i)}
\end{aligned}$$

(4.13)

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

and obtain from Lemmas 4.6 and 4.7 for $s > 0$

$$\begin{aligned}
&\sum_{|\boldsymbol{j}|+t/2<\mathrm{ld}\, N} 2^{|\boldsymbol{j}|(s-1/p+1)q} \left(\sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}} |\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^p \right)^{q/p} \\
&\quad \ll 2^{qt}\, N^{-q} \sum_{\mu=0}^{r} \sum_{i=0}^{\lceil q\rceil} 2^{-n_{\mu+1}(\lceil q\rceil - i)} 2^{(n_{\mu+1}-t/2)(sq+\lceil q\rceil - i)} (\log N)^{d-1} \\
&\quad \leqslant 2^{qt}\, N^{-q} 2^{-tsq/2 - t\lceil q\rceil/2} \sum_{\mu=0}^{r} 2^{n_{\mu+1}sq} \sum_{i=0}^{\lceil q\rceil} 2^{ti/2} (\log N)^{d-1} \\
&\quad \ll 2^{qt-tsq/2}\, N^{-q} 2^{n_{r+1}sq} (\log N)^{d-1} = 2^{tq(1-s/2)}\, N^{(s-1)q} (\log N)^{d-1}.
\end{aligned} \tag{4.14}$$

The case $s = 0$ needs to be handled slightly differently. We continue from (4.13) and obtain

$$\begin{aligned}
&\sum_{|\boldsymbol{j}|+t/2<\mathrm{ld}\, N} 2^{|\boldsymbol{j}|(-1/p+1)q} \left(\sum_{\boldsymbol{m}\in\mathbb{D}_{\boldsymbol{j}}} |\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^p \right)^{q/p} \\
&\quad \ll 2^{qt}\, N^{-q} \sum_{\mu=0}^{r} \sum_{i=0}^{\lceil q\rceil - 1} 2^{-n_{\mu+1}(\lceil q\rceil - i)} \\
&\qquad \times \sum_{n_\mu \leqslant |\boldsymbol{j}|+t/2<n_{\mu+1}} 2^{|\boldsymbol{j}|(\lceil q\rceil - i)} (2n_{\mu+1} - t - 2|\boldsymbol{j}|)^{(d-1)(\lceil q\rceil - i)} \\
&\qquad + 2^{qt}\, N^{-q} \sum_{\mu=0}^{r} \sum_{n_\mu \leqslant |\boldsymbol{j}|+t/2<n_{\mu+1}} (2n_{\mu+1} - t - 2|\boldsymbol{j}|) \\
&\quad \ll 2^{qt}\, N^{-q} (\log N)^d + 2^{qt} N^{-q} \\
&\qquad \times \sum_{\mu=0}^{r} (n_{\mu+1} - n_\mu)(\log N)^{d-1} \ll 2^{qt}\, N^{-q} (\log N)^d.
\end{aligned} \tag{4.15}$$

Combining (4.12), (4.14) and (4.15) we obtain

$$\| D_{\mathcal{S}_d}^N | S_{p,q}^s B \|^q \ll_{p,q,s,d} \begin{cases} 2^{tq} N^{-q}\, (\log N)^d & \text{if } s = 0, \\ 2^{tq} N^{(s-1)q}\, (\log N)^{d-1} & \text{if } 0 < s < 1/p, \end{cases}$$

and the result follows. □

4.6. *The proof of the upper bounds in Theorem 3.2.*

Proof of equations (3.10) and (3.12) in Theorem 3.2. From Proposition 3.3 we know that

$$S_{\max(p,q),q}^s B \hookrightarrow S_{p,q}^s F.$$

Thus we have

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

$$\|D_{\mathcal{S}_d}^N|S_{p,q}^s F\| \ll \|D_{\mathcal{S}_d}^N|S_{\max(p,q),q}^s B\| \ll \begin{cases} 2^t N^{-1} (\log N)^{d/q} & \text{if } s = 0, \\ 2^t N^{(s-1)} (\log N)^{(d-1)/q} & \text{if } 0 < s < 1/\max(p,q). \end{cases} \quad \square$$

§5. *Lower bounds.* In this section we prove the lower bounds on the respective norms of the discrepancy function of arbitrary sequences in the unit cube.

5.1. *The proof of the lower bound in Theorem 2.1.* To prove equation (2.2) we actually prove a lower bound for a norm close to the L_2-norm of f which is dominated by the BMO-norm. For simplicity of notation, let us consider here the L_2-normalized version of the Haar system consisting of the functions $g_{\boldsymbol{j},\boldsymbol{m}} = 2^{|\boldsymbol{j}|/2} h_{\boldsymbol{j},\boldsymbol{m}}$. Moreover, for a subset $\mathcal{E}$ of indices $(\boldsymbol{j},\boldsymbol{m})$ with $\boldsymbol{j} \in \mathbb{N}_{-1}^d$, $\boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}}$ let $P_{\mathcal{E}} f$ denote the orthogonal projection of $f \in L_2([0,1]^d)$ onto the span of $\{g_{\boldsymbol{j},\boldsymbol{m}} : (\boldsymbol{j},\boldsymbol{m}) \in \mathcal{E}\}$ and let $\|f|\mathcal{E}\| = \|P_{\mathcal{E}} f|L_2\|$. Then Parseval's equality yields

$$\|f|\mathcal{E}\|^2 = \sum_{(\boldsymbol{j},\boldsymbol{m})\in\mathcal{E}} \langle f, g_{\boldsymbol{j},\boldsymbol{m}}\rangle^2.$$

We also fix the index sets

$$\mathcal{D}^d = \{(\boldsymbol{j},\boldsymbol{m}) : \boldsymbol{j} \in \mathbb{N}_{-1}^d, \boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}}\} \quad \text{and} \quad \mathcal{D}_0^d = \{(\boldsymbol{j},\boldsymbol{m}) : \boldsymbol{j} \in \mathbb{N}_0^d, \boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}}\}.$$

Considering $U = [0,1)^d$ in the definition (2.1) of the BMO-norm, we obviously have

$$\|f|\mathrm{BMO}^d\| \geqslant \|f|\mathcal{D}_0^d\|.$$

Now the lower bound (2.2) in Theorem 2.1 follows immediately from the next lemma.

LEMMA 5.1. *For every $d \in \mathbb{N}$ there exists a positive constant $c_d > 0$ with the following property: for every infinite sequence $\mathcal{S}_d$ in $[0,1)^d$ we have*

$$\|D_{\mathcal{S}_d}^N|\mathcal{D}_0^d\| \geqslant c_d\, N^{-1} (\log N)^{d/2} \quad \textit{for infinitely many } N \in \mathbb{N}.$$

For the proof we combine the following lemma with the method of Proinov.

LEMMA 5.2. *Let $f : [0,1]^{d+1} \to \mathbb{R}$ be in $L_2([0,1]^d)$. Then we have*

$$\|f|\mathcal{D}_0^{d+1}\| \leqslant \sup_{t_{d+1}\in[0,1]} \|f(\cdot, t_{d+1})|\mathcal{D}_0^d\|.$$

Proof. We are going to show the identity

$$\int_0^1 \|f(\cdot, t_{d+1})|\mathcal{D}_0^d\|^2\, dt_{d+1} = \|f|\mathcal{E}^{d+1}\|^2, \tag{5.1}$$

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

where

$$\mathcal{E}^{d+1} = \{(\boldsymbol{j}, \boldsymbol{m}) : \boldsymbol{j} \in \mathbb{N}_0^d \times \mathbb{N}_{-1}, \boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}}\}.$$

Then the claimed inequality is a consequence of (5.1) together with the obvious inequalities

$$\|f|\mathcal{D}_0^{d+1}\| \leqslant \|f|\mathcal{E}^{d+1}\|$$

and

$$\int_0^1 \|f(\cdot, t_{d+1})|\mathcal{D}_0^d\|^2 \, dt_{d+1} \leqslant \sup_{t_{d+1}\in[0,1]} \|f(\cdot, t_{d+1})|\mathcal{D}_0^d\|^2.$$

For the proof of (5.1), let

$$f = \sum_{(\boldsymbol{j},\boldsymbol{m})\in\mathcal{D}^{d+1}} a_{\boldsymbol{j},\boldsymbol{m}} g_{\boldsymbol{j},\boldsymbol{m}}$$

be the Haar decomposition of f which converges in L_2 and, therefore, almost everywhere. For each $(\boldsymbol{j}, \boldsymbol{m}) \in \mathcal{D}^{d+1}$ with $\boldsymbol{j} = (j_1, \ldots, j_d, j_{d+1})$ and $\boldsymbol{m} = (m_1, \ldots, m_d, m_{d+1})$, we let $\boldsymbol{j}_* = (j_1, \ldots, j_d)$ and $\boldsymbol{m}_* = (m_1, \ldots, m_d)$. Then

$$f(\cdot, t_{d+1}) = \sum_{(\boldsymbol{j}_*,\boldsymbol{m}_*)\in\mathcal{D}^d} b_{\boldsymbol{j}_*,\boldsymbol{m}_*}(t_{d+1}) g_{\boldsymbol{j}_*,\boldsymbol{m}_*}$$

in $L_2([0, 1]^d)$ for almost every $t_{d+1} \in [0, 1]$, where

$$b_{\boldsymbol{j}_*,\boldsymbol{m}_*}(t_{d+1}) = \sum_{(j_{d+1},m_{d+1})\in\mathcal{D}^1} a_{\boldsymbol{j},\boldsymbol{m}} g_{j_{d+1},m_{d+1}}(t_{d+1}).$$

It follows that

$$\|f(\cdot, t_{d+1})|\mathcal{D}_0^d\|^2 = \sum_{(\boldsymbol{j}_*,\boldsymbol{m}_*)\in\mathcal{D}_0^d} b_{\boldsymbol{j}_*,\boldsymbol{m}_*}(t_{d+1})^2.$$

By integrating over $t_{d+1} \in [0, 1]$ and using the orthogonality of the one-dimensional Haar system we finally conclude that

$$\begin{aligned}\int_0^1 \|f(\cdot, t_{d+1})|\mathcal{D}_0^d\|^2 \, dt_{d+1} &= \sum_{(\boldsymbol{j}_*,\boldsymbol{m}_*)\in\mathcal{D}_0^d} \int_0^1 b_{\boldsymbol{j}_*,\boldsymbol{m}_*}(t_{d+1})^2 \, dt_{d+1} \\ &= \sum_{(\boldsymbol{j}_*,\boldsymbol{m}_*)\in\mathcal{D}_0^d} \sum_{(j_{d+1},m_{d+1})\in\mathcal{D}^1} a_{\boldsymbol{j},\boldsymbol{m}}^2 = \|f|\mathcal{E}^{d+1}\|^2. \quad \square\end{aligned}$$

Additionally, we need the following lower bound for finite point sets.

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

PROPOSITION 5.3. *For every $d \in \mathbb{N}$ there exists a positive constant $c_d > 0$ with the following property: for every integer $N \geqslant 2$ and every N-element point set $\mathcal{P}_{N,d}$ in $[0,1)^d$ we have*

$$\|D_{\mathcal{P}_{N,d}}|\mathcal{D}_0^d\| \geqslant c_d\, N^{-1}(\log N)^{(d-1)/2}.$$

This has been shown in [2] for $d = 2$ and [4] for $d \geqslant 3$ with the BMO^d-norm instead of the $\mathcal{D}_0^d$-norm. But the proofs actually work literally for the $\mathcal{D}_0^d$-norm. The proof of the following result relies on Proinov's idea.

LEMMA 5.4. *For $d \in \mathbb{N}$ let $\mathcal{S}_d = (\boldsymbol{y}_k)_{k\geqslant 0}$, where $\boldsymbol{y}_k = (y_{1,k}, \ldots, y_{d,k})$ for $k \in \mathbb{N}_0$, be an arbitrary sequence in $[0,1)^d$. Then for every $N \in \mathbb{N}$ there exists an $n \in \{1, 2, \ldots, N\}$ such that*

$$n\|D_{\mathcal{S}_d}^n|\mathcal{D}_0^d\| + 1 \geqslant N\|D_{\mathcal{P}_{N,d+1}}|\mathcal{D}_0^{d+1}\|$$

where $\mathcal{P}_{N,d+1}$ is the finite point set in $[0,1)^{d+1}$ consisting of the points

$$\boldsymbol{x}_k = (y_{1,k}, \ldots, y_{d,k}, k/N) \quad \text{for } k = 0, 1, \ldots, N-1.$$

Proof. Choose $n \in \{1, 2, \ldots, N\}$ such that

$$\|nD_{\mathcal{S}_d}^n|\mathcal{D}_0^d\| = \max_{k=1,2,\ldots,N} \|kD_{\mathcal{S}_d}^k|\mathcal{D}_0^d\|. \tag{5.2}$$

Consider a subinterval of the $(d+1)$-dimensional unit cube of the form $E = \prod_{i=1}^{d+1}[0, t_i)$ with $\boldsymbol{t} = (t_1, t_2, \ldots, t_{d+1}) \in [0,1)^{d+1}$ and put

$$m = m(t_{d+1}) := \lceil Nt_{d+1} \rceil \in \{1, 2, \ldots, N\}.$$

Then a point $\boldsymbol{x}_k$, $k = 0, 1, \ldots, N-1$, belongs to E, if and only if $\boldsymbol{y}_k \in \prod_{i=1}^{d}[0, t_i)$ and $0 \leqslant k < Nt_{d+1}$. Writing $E_* = \prod_{i=1}^{d}[0, t_i)$ we have

$$\sum_{\boldsymbol{z}\in\mathcal{P}_{N,d+1}} \chi_E(\boldsymbol{z}) = \sum_{\boldsymbol{y}\in\mathcal{P}_{m,d}} \chi_{E_*}(\boldsymbol{y}),$$

where $\mathcal{P}_{m,d}$ is the point set which consists of the first m terms of the sequence $\mathcal{S}_d$. Therefore we have

$$\begin{aligned} N D_{\mathcal{P}_{N,d+1}}(\boldsymbol{t}) &= \sum_{\boldsymbol{z}\in\mathcal{P}_{N,d+1}} \chi_E(\boldsymbol{z}) - Nt_1t_2\cdots t_{d+1} \\ &= \sum_{\boldsymbol{y}\in\mathcal{P}_{m,d}} \chi_{E_*}(\boldsymbol{y}) - mt_1t_2\cdots t_d + mt_1t_2\cdots t_d - Nt_1t_2\cdots t_{d+1} \\ &= m D_{\mathcal{P}_{m,d}}(\boldsymbol{t}_*) + t_1t_2\cdots t_d(m - Nt_{d+1}), \end{aligned}$$

where $\boldsymbol{t}_* = (t_1, \ldots, t_d)$. With

$$\varphi(\boldsymbol{t}) = t_1t_2\cdots t_d(m - Nt_{d+1}) = t_1t_2\cdots t_d(m(t_{d+1}) - Nt_{d+1})$$

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

we obtain

$$N D_{\mathcal{P}_{N,d+1}}(\boldsymbol{t}) = m D_{\mathcal{P}_{m,d}}(\boldsymbol{t}_*) + \varphi(\boldsymbol{t}). \tag{5.3}$$

From the definition of m it is clear that $|m - N t_{d+1}| \leqslant 1$ and, therefore, also $|\varphi(\boldsymbol{t})| \leqslant 1$. It follows that

$$\begin{aligned} N \| D_{\mathcal{P}_{N,d+1}} | \mathcal{D}_0^{d+1} \| &\leqslant \| m D_{\mathcal{S}_d}^m | \mathcal{D}_0^{d+1} \| + \| \varphi | \mathcal{D}_0^{d+1} \| \\ &\leqslant \| m D_{\mathcal{S}_d}^m | \mathcal{D}_0^{d+1} \| + \| \varphi | L_2 \| \leqslant \| m D_{\mathcal{S}_d}^m | \mathcal{D}_0^{d+1} \| + 1. \end{aligned}$$

Using Lemma 5.2 we obtain

$$\| m(\cdot) D_{\mathcal{S}_d}^{m(\cdot)} | \mathcal{D}_0^{d+1} \| + 1 \leqslant \sup_{t_{d+1} \in [0,1]} \| m(t_{d+1}) D_{\mathcal{S}_d}^{m(t_{d+1})} | \mathcal{D}_0^d \| + 1 \leqslant \| n D_{\mathcal{S}_d}^n | \mathcal{D}_0^d \| + 1.$$

This implies the result. □

Now we can give the proof of Lemma 5.1 which also finishes the proof of equation (2.2) in Theorem 2.1.

Proof of Lemma 5.1. We use the notation from Lemma 5.4. For the finite point set $\mathcal{P}_{N,d+1}$ in $[0, 1)^{d+1}$ we obtain from Proposition 5.3 that

$$N \| D_{\mathcal{P}_{N,d+1}} | \mathcal{D}_0^{d+1} \| \geqslant c_{d+1} (\log N)^{d/2}$$

for some real $c_{d+1} > 0$ which is independent of N. According to Lemma 5.4 there exists an $n \in \{1, 2, \ldots, N\}$ such that

$$n \| D_{\mathcal{S}_d}^n | \mathcal{D}_0^d \| + 1 \geqslant N \| D_{\mathcal{P}_{N,d+1}} | \mathcal{D}_0^{d+1} \| \geqslant c_{d+1} (\log N)^{d/2}. \tag{5.4}$$

Thus we have shown that for every $N \in \mathbb{N}$ there exists an $n \in \{1, 2, \ldots, N\}$ such that

$$n \| D_{\mathcal{S}_d}^n | \mathcal{D}_0^d \| + 1 \geqslant c_{d+1} (\log n)^{d/2}. \tag{5.5}$$

It remains to show that (5.5) holds for infinitely many $n \in \mathbb{N}$. Assume to the contrary that (5.5) holds for finitely many $n \in \mathbb{N}$ only and let m be the largest integer with this property. Then choose $N \in \mathbb{N}$ large enough such that

$$c_{d+1} (\log N)^{d/2} > \max_{k=1,2,\ldots,m} k \| D_{\mathcal{S}_d}^k | \mathcal{D}_0^d \| + 1.$$

For this N we can find an $n \in \{1, 2, \ldots, N\}$ for which (5.4) and (5.5) hold true. However, (5.4) implies that $n > m$ which leads to a contradiction since m is the largest integer such that (5.5) is true. Thus we have shown that (5.5) holds for infinitely many $n \in \mathbb{N}$ and this completes the proof. □

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

5.2. *The proof of the lower bounds in Theorems 3.1 and 3.2.* Before we give the proof we collect some auxiliary results.

LEMMA 5.5. *Let $p, q \in (0, \infty)$ and $0 \leqslant s < 1/p$. For every $d \in \mathbb{N}$ there exists a positive constant $c_{p,q,s,d} > 0$ with the following property: for every infinite sequence $\mathcal{S}_d$ in $[0, 1)^d$ and every $N \in \mathbb{N}$ we have*

$$\|D_{\mathcal{S}_d}^N | S_{p,q}^s B([0,1)^d)\| \geqslant c_{p,q,s,d}\, N^{s-1} (\log N)^{(d-1)/q}.$$

Proof. We apply Proposition 3.4 and the now classical idea of Roth's proof of the bound on the L_2-discrepancy [**32**]. Fix an arbitrary infinite sequence $\mathcal{S}_d$. Then for every $N \in \mathbb{N}$ we can choose $n \in \mathbb{N}$ such that $2N \leqslant 2^n < 4N$. Let $\boldsymbol{j} \in \mathbb{N}_0^d$ with $|\boldsymbol{j}| = n$. Then there are at least $2^{n-1} = 2^{|\boldsymbol{j}|-1}$ values of $\boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}}$ for which $I_{\boldsymbol{j},\boldsymbol{m}} \cap \mathcal{P}_{N,d} = \emptyset$, where $\mathcal{P}_{N,d}$ is again the set consisting of the first N elements of $\mathcal{S}_d$. For such $I_{\boldsymbol{j},\boldsymbol{m}}$ which do not contain any points of $\mathcal{P}_{N,d}$ we give the Haar coefficients as $|\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle| \asymp 2^{-2|\boldsymbol{j}|}$, relying on Lemma 4.8 and 4.9. We apply Proposition 3.4 to obtain (with some constant $c > 0$)

$$\begin{aligned}
\|D_{\mathcal{S}_d}^N | S_{p,q}^s B([0,1)^d)\|^q &\asymp \sum_{\boldsymbol{j} \in \mathbb{N}_{-1}^d} 2^{|\boldsymbol{j}|(s-1/p+1)q} \left(\sum_{\boldsymbol{m} \in \mathbb{D}_{\boldsymbol{j}}} |\langle D_{\mathcal{S}_d}^N, h_{\boldsymbol{j},\boldsymbol{m}}\rangle|^p \right)^{q/p} \\
&\geqslant c \sum_{|\boldsymbol{j}|=n} 2^{|\boldsymbol{j}|(s-1/p+1)q}\, 2^{-2|\boldsymbol{j}|q}\, 2^{(|\boldsymbol{j}|-1)q/p} \\
&\asymp \sum_{|\boldsymbol{j}|=n} 2^{|\boldsymbol{j}|(s-1)q} \asymp n^{d-1} 2^{n(s-1)q}.
\end{aligned}$$

□

LEMMA 5.6. *Let $p, q \in (0, \infty)$. Let $f : [0,1]^{d+1} \to \mathbb{R}$ be in $S_{p,q}^0 B([0,1]^d)$. Then we have*

$$\|f | S_{p,q}^0 B([0,1]^{d+1})\| \leqslant \sup_{t_{d+1} \in [0,1]} \|f(\cdot, t_{d+1}) | S_{p,q}^0 B([0,1]^d)\|.$$

Proof. We write $\langle f, g\rangle_d$ to keep track of the dimension. We have

$$\begin{aligned}
|\langle f, h_{\boldsymbol{j},\boldsymbol{m}}\rangle_{d+1}| &= |\langle \langle f, h_{\boldsymbol{j}',\boldsymbol{m}'}\rangle_d, h_{j_{d+1},\boldsymbol{m}_{d+1}}\rangle_1| \\
&\leqslant \sup_{t_{d+1} \in [0,1]} |\langle f(\cdot, t_{d+1}), h_{\boldsymbol{j}',\boldsymbol{m}'}\rangle_d| 2^{-\max(j_{d+1},0)}.
\end{aligned}$$

The result then follows from the fact that $\mathbb{D}_{j_{d+1}}$ contains $2^{j_{d+1}}$ elements and Proposition 3.4. □

We need the following lower bound for finite point sets from [**40**].

PROPOSITION 5.7. *For every $d \in \mathbb{N}$ and all $1 \leqslant p, q < \infty$ and $1/p - 1 < s < 1/p$ there exists a positive constant $c_{p,q,s,d} > 0$ with the following property: for every integer $N \geqslant 2$ and every N-element point set $\mathcal{P}_{N,d}$ in $[0,1)^d$ we have*

$$\|D_{\mathcal{P}_{N,d}} | S_{p,q}^s B([0,1)^d)\| \geqslant c_{p,q,s,d}\, N^{s-1} (\log N)^{(d-1)/q}.$$

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

LEMMA 5.8. *Let $1 \leqslant p, q < \infty$. For $d \in \mathbb{N}$ let $\mathcal{S}_d = (\boldsymbol{y}_k)_{k \geqslant 0}$, where $\boldsymbol{y}_k = (y_{1,k}, \ldots, y_{d,k})$ for $k \in \mathbb{N}_0$, be an arbitrary sequence in $[0,1)^d$. Then for every $N \in \mathbb{N}$ there exists an $n \in \{1, 2, \ldots, N\}$ such that*

$$n \| D^n_{\mathcal{S}_d} | S^0_{p,q} B([0,1)^d) \| + 1 \geqslant N \| D_{\mathcal{P}_{N,d+1}} | S^0_{p,q} B([0,1)^{d+1}) \|,$$

where $\mathcal{P}_{N,d+1}$ is the finite point set in $[0,1)^{d+1}$ consisting of the points

$$\boldsymbol{x}_k = (y_{1,k}, \ldots, y_{d,k}, k/N) \quad \text{for } k = 0, 1, \ldots, N-1.$$

Proof. The proof is analogous to the proof of Lemma 5.4. We choose $n \in \{1, 2, \ldots, N\}$ such that

$$\| n D^n_{\mathcal{S}_d} | S^0_{p,q} B([0,1)^d) \| = \max_{k=1,\ldots,N} \| k D^k_{\mathcal{S}_d} | S^0_{p,q} B([0,1)^d) \|.$$

For $\boldsymbol{t} \in [0,1)^{d+1}$ we have (5.3)

$$N D_{\mathcal{P}_{N,d+1}}(\boldsymbol{t}) = m D_{\mathcal{P}_{m,d}}(\boldsymbol{t}_*) + \varphi(\boldsymbol{t})$$

(see the proof of Lemma 5.4 for the notation). Then analogously to the proof of Lemma 5.4 we show that

$$N \| D_{\mathcal{P}_{N,d+1}} | S^0_{p,q} B([0,1)^{d+1}) \| \leqslant \| m D^m_{\mathcal{S}_d} | S^0_{p,q} B([0,1)^{d+1}) \| + 1.$$

Finally we use Lemma 5.6 to obtain

$$\begin{aligned} \| m(\cdot) D^{m(\cdot)}_{\mathcal{S}_d} | S^0_{p,q} B([0,1)^{d+1}) \| &\leqslant \sup_{t_{d+1} \in [0,1]} \| m(t_{d+1}) D^{m(t_{d+1})}_{\mathcal{S}_d} | S^0_{p,q} B([0,1)^d) \| \\ &\leqslant \| n D^n_{\mathcal{S}_d} | S^0_{p,q} B([0,1)^d) \|. \end{aligned}$$

□

Proof of equations (3.3) and (3.5) in Theorem 3.1. The case where $s > 0$ is covered by Lemma 5.5, hence we only have to consider the case $s = 0$. We use the notation from Lemma 5.4. For the finite point set $\mathcal{P}_{N,d+1}$ in $[0,1)^{d+1}$ we obtain from Proposition 5.7 that

$$N \| D_{\mathcal{P}_{N,d+1}} | S^0_{p,q} B([0,1)^{d+1}) \| \geqslant c_{p,q,0,d+1} (\log N)^{d/q}$$

for some constant $c_{p,q,0,d+1} > 0$ which is independent of N. Now we apply Lemma 5.8 and obtain an $n \in \{1, 2, \ldots, N\}$ such that

$$\begin{aligned} n \| D^n_{\mathcal{S}_d} | S^0_{p,q} B([0,1)^d) \| + 1 &\geqslant N \| D_{\mathcal{P}_{N,d+1}} | S^0_{p,q} B([0,1)^{d+1}) \| \\ &\geqslant c_{p,q,0,d+1} (\log N)^{d/q}. \end{aligned} \tag{5.6}$$

The proof that there are infinitely many $n \in \mathbb{N}$ for which the (5.6) holds works analogously to the proof of Lemma 5.1. □

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

Proof of equations (3.9) and (3.11) in Theorem 3.2. We apply Proposition 3.3, namely $S^s_{p,q}F \hookrightarrow S^s_{\min(p,q),q}B$ and we have for a constant $c > 0$ from Theorem 3.1

$$\begin{aligned}
&\|D^N_{\mathcal{S}_d}|S^s_{p,q}F\| \\
&\quad \geqslant c\,\|D^N_{\mathcal{S}_d}|S^s_{\min(p,q),q}B\| \\
&\quad \geqslant c_{p,q,s,d} \begin{cases} N^{-1}\,(\log N)^{d/q} & \text{if } s = 0, \\ N^{(s-1)}\,(\log N)^{(d-1)/q} & \text{if } 0 < s < 1/\min(p,q). \end{cases}
\end{aligned}$$

□

Acknowledgements. The authors would like to thank William Chen for his very helpful remarks and suggestions.

References

1. J. Beck and W. W. L. Chen, *Irregularities of Distribution*, Cambridge University Press (Cambridge, 1987).
2. 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.
3. D. Bilyk, M. T. Lacey and A. Vagharshakyan, On the small ball inequality in all dimensions. *J. Funct. Anal.* **254** (2008), 2470–2502.
4. D. Bilyk and L. Markhasin, BMO and exponential Orlicz space estimates of the discrepancy function in arbitrary dimension. *J. Anal. Math.* (to appear).
5. S.-Y. A. Chang and R. Fefferman, A continuous version of duality of H^1 with BMO on the bidisc. *Ann. of Math.* (2) **112** (1980), 179–201.
6. S.-Y. A. Chang, J. M. Wilson and T. H. Wolff, Some weighted norm inequalities concerning the Schrödinger operators. *Comment. Math. Helv.* **60** (1985), 217–246.
7. W. W. L. Chen, On irregularities of distribution. *Mathematika* **27** (1981), 153–170.
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. 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.
10. J. Dick, Walsh spaces containing smooth functions and quasi-Monte Carlo rules of arbitrary high order. *SIAM J. Numer. Anal.* **46** (2008), 1519–1553.
11. J. Dick, Discrepancy bounds for infinite-dimensional order two digital sequences over $\mathbb{F}_2$. *J. Number Theory* **136** (2014), 204–232.
12. J. Dick and J. Baldeaux, Equidistribution properties of generalized nets and sequences. In *Monte Carlo and Quasi-Monte Carlo Methods 2008* (eds P. L'Ecuyer and A. Owen), Springer (Berlin–Heidelberg, 2009), 305–323.
13. J. Dick, A. Hinrichs, L. Markhasin and F. Pillichshammer, Optimal L_p-discrepancy bounds for second order digital sequences. *Israel J. Math.* **221** (2017), 489–510.
14. J. Dick and F. Pillichshammer, *Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration*, Cambridge University Press (Cambridge, 2010).
15. 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.
16. 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* (*Radon Series of Computational and Applied Mathematics*) (eds P. Kritzer, H. Niederreiter, F. Pillichshammer and A. Winterhof), De Gruyter (Berlin/Boston, 2014), 63–86.
17. M. Drmota and R. F. Tichy, *Sequences, Discrepancies and Applications* (*Lecture Notes in Mathematics* **1651**), Springer (Berlin, 1997).
18. A. Hinrichs, Discrepancy of Hammersley points in Besov spaces of dominating mixed smoothness. *Math. Nachr.* **283** (2010), 478–488.
19. R. Kritzinger, L_p- and $S^r_{p,q}B$-discrepancy of the symmetrized van der Corput sequence and modified Hammersley point sets in arbitrary bases. *J. Complexity* **33** (2016), 145–168.
20. L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, John Wiley (New York, NY, 1974) , reprinted by Dover (Mineola, NY, 2006).

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press

21. G. Leobacher and F. Pillichshammer, *Introduction to Quasi-Monte Carlo Integration and Applications* (*Compact Textbooks in Mathematics*), Birkhäuser/Springer (Cham, 2014).
22. J. Lindenstrauss and L. Tzafriri, *Classical Banach Spaces, I*, Springer (Berlin, 1977).
23. L. Markhasin, Discrepancy of generalized Hammersley type point sets in Besov spaces with dominating mixed smoothness. *Unif. Distrib. Theory* **8** (2013), 135–164.
24. L. Markhasin, Discrepancy and integration in function spaces with dominating mixed smoothness. *Dissertationes Math. (Rozprawy Mat.)* **494** (2013), 1–81.
25. L. Markhasin, Quasi-Monte Carlo methods for integration of functions with dominating mixed smoothness in arbitrary dimension. *J. Complexity* **29** (2013), 370–388.
26. L. Markhasin, L_p- and $S^r_{p,q}B$-discrepancy of (order 2) digital nets. *Acta Arith.* **168** (2015), 139–159.
27. J. Matoušek, *Geometric Discrepancy. An Illustrated Guide* (*Algorithms and Combinatorics* **18**), Springer (Berlin, 1999).
28. H. Niederreiter, Point sets and sequences with small discrepancy. *Monatsh. Math.* **104** (1987), 273–337.
29. H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods* (*CBMS-NSF Regional Conference Series in Applied Mathematics* **63**), SIAM (Philadelphia, 1992).
30. H. Niederreiter and C. P. Xing, Low-discrepancy sequences and global function fields with many rational places. *Finite Fields Appl.* **2** (1996), 241–273.
31. P. D. Proinov, On irregularities of distribution. *C. R. Acad. Bulgare Sci.* **39** (1986), 31–34.
32. K. F. Roth, On irregularities of distribution. *Mathematika* **1** (1954), 73–79.
33. K. F. Roth, On irregularities of distribution. IV. *Acta Arith.* **37** (1980), 67–75.
34. W. M. Schmidt, Irregularities of distribution. VII. *Acta Arith.* **21** (1972), 45–50.
35. W. M. Schmidt, Irregularities of distribution X. In *Number Theory and Algebra*, Academic Press (New York, NY, 1977), 311–329.
36. M. M. Skriganov, Harmonic analysis on totally disconnected groups and irregularities of point distributions. *J. Reine Angew. Math.* **600** (2006), 25–49.
37. I. M. Sobol', The distribution of points in a cube and the approximate evaluation of integrals. *Zh. Vychisl. Mat. Mat. Fiz.* **7** (1967), 784–802.
38. T. Tao, A type diagram for function spaces. What's new – a blog by T. Tao, 11 March 2010, https://terrytao.wordpress.com/tag/besov-spaces.
39. S. Tezuka, Polynomial arithmetic analogue of Halton sequences. *ACM Trans. Model. Comput. Simul.* **3** (1993), 99–107.
40. H. Triebel, *Bases in Function Spaces, Sampling, Discrepancy, Numerical Integration*, European Mathematical Society (Zürich, 2010).
41. H. Triebel, Numerical integration and discrepancy. A new approach. *Math. Nachr.* **283** (2010), 139–159.

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

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

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

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

https://doi.org/10.1112/S0025579317000213 Published online by Cambridge University Press