EXPONENTIAL SQUARED INTEGRABILITY OF THE DISCREPANCY FUNCTION IN TWO DIMENSIONS

DMITRIY BILYK, MICHAEL T. LACEY, IOANNIS PARISSIS, AND ARMEN VAGHARSHAKYAN

Abstract. Let $\mathcal{A}_N$ be an N-point set in the unit square and consider the Discrepancy function

$$D_N(\vec{x}) := \sharp\big(\mathcal{A}_N \cap [\vec{0}, \vec{x})\big) - N|[\vec{0}, \vec{x})|,$$

where $\vec{x} = (x_1, x_2) \in [0,1]^2$, $[0, \vec{x}) = \prod_{t=1}^{2}[0, x_t)$, and $|[\vec{0}, \vec{x})|$ denotes the Lebesgue measure of the rectangle. We give various refinements of a well-known result of (Schmidt, 1972) on the L^∞ norm of D_N. We show that necessarily

$$\|D_N\|_{\exp(L^\alpha)} \gtrsim (\log N)^{1-1/\alpha}\,, \qquad 2 \le \alpha < \infty\,.$$

The case of $\alpha = \infty$ is the Theorem of Schmidt. This estimate is sharp. For the digit-scrambled van der Corput sequence, we have

$$\|D_N\|_{\exp(L^\alpha)} \lesssim (\log N)^{1-1/\alpha}\,, \qquad 2 \le \alpha < \infty\,,$$

whenever $N = 2^n$ for some positive integer n. This estimate depends upon variants of the Chang-Wilson-Wolff inequality (Chang et al., 1985). We also provide similar estimates for the BMO norm of D_N.

1. Main Theorems

The common theme of the subject of irregularities of distribution is to show that, no matter how N points are selected, their distribution must be far from uniform. In the present article, we are primarily interested in the precise behavior of such estimates near the L^∞ endpoint, phrased in terms of exponential Orlicz classes. We restrict our attention to the two-dimensional case.

Let $\mathcal{A}_N \subset [0,1]^2$ be a set of N points in the unit square. For $\vec{x} = (x_1, x_2) \in [0,1]^2$, we define the *Discrepancy function* associated to $\mathcal{A}_N$ as follows:

$$D_N(\vec{x}) := \sharp\big(\mathcal{A}_N \cap [0, \vec{x})\big) - N|[0, \vec{x})|\,,$$

where $[0, \vec{x})$ is the axis-parallel rectangle in the unit square with one vertex at the origin and the other at $\vec{x} = (x_1, x_2)$, and $|[0, \vec{x})| = x_1 \cdot x_2$ denotes the Lebesgue measure of the rectangle. This is the difference between the actual number of points in the rectangle $[0, \vec{x})$ and the expected number of points in this rectangle. The relative size of this function, in

All authors are grateful to the Fields Institute for hospitality and support, and to the National Science Foundation for support.

various senses, must necessarily increase with N. The principal result in this direction is due to Roth (Roth, 1954):

K. Roth's Theorem. *In all dimensions $d \geq 2$, we have the following estimate*

$$\|D_N\|_2 \gtrsim (\log N)^{(d-1)/2} \tag{1.1}$$

where the implied constant is only a function of dimension d.

The same bound holds for the L^p norm, for $1 < p < \infty$, (Schmidt, 1977b), and is known to be sharp as to the order of magnitude, see (Chen, 1980) and (Beck and Chen, 1987) for a history of this subject (for the case $d = 2$, see Corollary 1.3 below). The endpoint cases of $p = 1$ and $p = \infty$ are much harder.

We concentrate on the case of $p = \infty$ in this note, just in dimension $d = 2$, and refer the reader to (Beck, 1989; Bilyk et al., 2008; Bilyk and Lacey, 2008; Halász, 1981) for more information about the case of $d \geq 3$. For information about the case of $p = 1$, see (Halász, 1981; Lacey, 2006). As it has been shown in the fundamental theorem of W. Schmidt (Schmidt, 1972), in dimension $d = 2$, the lower bound on the L^∞ norm of the Discrepancy function is substantially greater than the L^p estimate (1):

W. Schmidt's Theorem. *For any set $\mathcal{A}_N \subset [0,1]^2$ we have*

$$\|D_N\|_\infty \gtrsim \log N\,. \tag{1.2}$$

This theorem is also sharp: one particular example is the famous van der Corput set (van der Corput, 1935) – a detailed discussion is contained in §3. In this paper, we give an interpolant between the results of Roth and Schmidt, which is measured in the scale of exponential Orlicz classes.

1.3. **Theorem.** *For any N-point set $\mathcal{A}_N \subset [0,1]^2$ we have*

$$\|D_N\|_{\exp(L^\alpha)} \gtrsim (\log N)^{1-1/\alpha}\,, \qquad 2 \leq \alpha < \infty\,.$$

Of course the lower bound of $(\log N)^{1/2}$, the case of $\alpha = 2$ above, is a consequence of Roth's bound. The other estimates require proof, which is a variant of Halász's argument (Halász, 1981). We give details below and also remark that this estimate in the context of the Small Ball Inequality (Talagrand, 1994; Temlyakov, 1995) is known (Dunker et al., 1998). In addition, we demonstrate that the previous theorem is sharp.

1.4. **Theorem.** *For all N, there is a choice of $\mathcal{A}_N$, specifically the digit-scrambled van der Corput set (see Definition 3.5), for which we have*

$$\|D_N\|_{\exp(L^\alpha)} \lesssim (\log N)^{1-1/\alpha}\,, \qquad 2 \leq \alpha < \infty\,.$$

In view of Proposition 2.2, taking $\alpha = 2$, the theorem above immediately yields the sharpness of the L^p lower bounds in $d = 2$ with explicit dependence of constants on p.

1.5. **Corollary.** *For every $1 \leq p < \infty$, the set $\mathcal{A}_N$ from Theorem 1.2 satisfies*

$$\|D_N\|_p \lesssim p^{1/2}(\log N)^{1/2},$$

where the implied constant is independent of p.

There is another variant of the Roth lower bound, which we state here.

1.6. **Theorem.** *We have the estimate*

$$\|D_N\|_{\mathrm{BMO}_{1,2}} \gtrsim (\log N)^{1/2} \,,$$

where the norm is the dyadic Chang-Fefferman product BMO *norm (see Definition 2.10), introduced in* (Chang and Fefferman, 1980).

Indeed, this Theorem is just a corollary to a standard proof of Roth's Theorem, and its main interest lies in the fact that the estimate above is sharp. It is useful to recall the simple observation that the BMO norm is insensitive to functions that are constant in either the vertical or horizontal direction. That is, we have $\|D_N\|_{\mathrm{BMO}_{1,2}} = \|\widetilde{D}_N\|_{\mathrm{BMO}_{1,2}}$, where

$$\begin{aligned}\widetilde{D}_N(x_1, x_2) = D_N(x_1, x_2) - \int_0^1 D_N(x_1, x_2)\, dx_1 \\ - \int_0^1 D_N(x_1, x_2)\, dx_2 + \int_0^1 \int_0^1 D_N(x_1, x_2)\, dx_1\, dx_2 \,.\end{aligned}$$

1.7. **Theorem.** *For $N = 2^n$, there is a choice of $\mathcal{A}_N$, specifically the digit-scrambled van der Corput set, for which we have*

$$\|D_N\|_{\mathrm{BMO}_{1,2}} \lesssim (\log N)^{1/2} \,.$$

The main point of these results is that they unify the theorems of Roth and Schmidt in a sharp fashion. This line of research is also of interest in higher dimensions, but the relevant conjectures do not seem to be as readily apparent. As such, we think that this is an interesting theme for further investigation.

In the next section we collect a variety of results needed to prove the main Theorems. These results are drawn from the theory of Irregularities of Distribution, Harmonic Analysis, Probability Theory and other subjects. In §3 we discuss the structure of the digit-scrambled van der Corput set. Section 4 is dedicated to the analysis of the Haar decomposition of the Discrepancy function for the van der Corput set. The proofs of the main theorems above are then taken up in the §5 and §6.

The results of this paper concern refinements of the L^∞-endpoint estimates for the Discrepancy Function. In three dimensions, even the correct form of Schmidt's Theorem is not yet known, making the discussion of these results in three dimensions entirely premature, though speculation about such results could inform the analysis of the more difficult three dimensional case. See (Bilyk and Lacey, 2008; Bilyk et al., 2008) for recent information about the higher dimensional versions of Schmidt's Theorem.

The authors thank the referee for an expert reading, and suggestions to improve the paper.

2. Preliminary Facts

We suppress many constants which do not affect the arguments in essential ways. $A \lesssim B$ means that there is an absolute constant $K > 0$ such that $A \le KB$. Thus $A \lesssim 1$ means that A is bounded by an absolute constant. And if $A \lesssim B \lesssim A$, we write $A \simeq B$.

Inequalities. We recall the square function inequalities for martingales, in a form convenient for us.

In one dimension, the class of dyadic intervals in the unit interval are $\mathcal{D} := \{[j2^{-k}, (j+1)2^{-k}) \mid j, k \in \mathbb{N}, 0 \le j < 2^k\}$. Let $\mathcal{D}_n$ denote the dyadic intervals of length 2^{-n}, and by abuse of notation, also the sigma field generated by these intervals. For an integrable function f on $[0,1]$, the conditional expectation is

$$f_n = \mathbb{E}(f \mid \mathcal{D}_n) := \sum_{I \in \mathcal{D}_n} \mathbf{1}_I \cdot |I|^{-1} \int_I f(y)\, dy\,.$$

The sequence of functions $\{f_n \mid n \ge 0\}$ is a *martingale*. The *martingale difference sequence* is $d_0 = f_0$, and $d_n = f_n - f_{n-1}$ for $n \ge 1$. The sequence of functions $\{d_n \mid n \ge 0\}$ are pairwise orthogonal. The *square function* is

$$\operatorname{S}(f) := \Big[\sum_{n=0}^{\infty} |d_n|^2\Big]^{1/2}.$$

We have the following extension of the Khintchine inequalities.

2.1. **Theorem.** *The inequalities below hold, for some absolute choice of constant $C > 0$.*

$$\|f\|_p \le C\sqrt{p}\|\operatorname{S}(f)\|_p\,, \qquad 2 \le p < \infty\,.$$

In addition, this inequality holds for Hilbert space valued functions f.

For real-valued martingales, this was observed by (Chang et al., 1985). The extension to Hilbert space valued martingales is useful for us and is proved in (Fefferman and Pipher, 1997). The best constants in these inequalities are known for $p \ge 3$ (Wang, 1991).

Orlicz Spaces. For background on Orlicz Spaces, we refer the reader to (Lindenstrauss and Tzafriri, 1977). Consider a symmetric convex function ψ, which is zero at the origin, and is otherwise non-zero. Let (Ω, P) be a probability space, on which our functions are defined, and let $\mathbb{E}$ denote expectation over the probability space. We can define

$$\|f\|_{L^\psi} = \inf\{K > 0 \mid \mathbb{E}\psi(f \cdot K^{-1}) \le 1\}\,,$$

where we define the infimum over the empty set to be ∞. The set of functions $L^\psi = \{f \mid \|f\|_{L^\Psi} < \infty\}$ is a normed linear space, called the Orlicz space associated with ψ.

We are interested in, for instance, $\psi(x) = \mathrm{e}^{x^2} - 1$, in which case we denote the Orlicz space by $\exp(L^2)$. More generally, for $\alpha > 0$, we let $\psi_\alpha(x)$ be a symmetric convex function which equals $e^{|x|^\alpha} - 1$ for $|x|$ sufficiently large, depending upon α.[1] And we write $L^{\psi_\alpha} = \exp(L^\alpha)$. These are the spaces used in the statements of our main Theorems 1.1 and 1.2. It is obvious that, for all $1 \le p < \infty$ and $\alpha > 0$, we have $L^p \supset \exp(L^\alpha) \supset L^\infty$, hence Theorem 1.1 can be indeed viewed as interpolation between the estimates of Roth (1) and Schmidt (1). The following useful proposition is well-known and follows from elementary methods.

2.2. **Proposition.** *We have the following equivalence of norms valid for all $\alpha > 0$:*

$$\|f\|_{\exp(L^\alpha)} \simeq \sup_{p>1} p^{-1/\alpha} \|f\|_p \,.$$

We shall also make use of the duality relations for the exponential Orlicz classes. For $\alpha > 0$, let $\varphi_\alpha(x)$ be a symmetric convex function which equals $|x|(\log(3 + |x|))^\alpha$ for $|x|$ sufficiently large, depending upon α.[2] The Orlicz space L^{φ_α} is denoted as $L^{\varphi_\alpha} = L(\log L)^\alpha$. The propositions below are standard.

2.3. **Proposition.** *For $0 < \alpha < \infty$, the two Orlicz spaces $\exp(L^\alpha)$ and $L(\log L)^{1/\alpha}$ are Banach spaces which are dual to one another.*

2.4. **Proposition.** *Let E be a measurable subset of a probability set. We have*

$$\|\mathbf{1}_E\|_{L(\log L)^{1/\alpha}} \simeq \mathbb{P}(E) \cdot (1 - \log \mathbb{P}(E))^{1/\alpha} \,.$$

Chang-Wilson-Wolff Inequality. Each dyadic interval has a left and right half, $I_{\text{left}}, I_{\text{right}}$ respectively, which are also dyadic. Define the Haar function associated with I by

$$h_I := -\mathbf{1}_{I_{\text{left}}} + \mathbf{1}_{I_{\text{right}}}$$

Note that here the Haar functions are normalized in L^∞. In particular, the square function with this normalization has the form

$$\mathrm{S}(f)^2 = \sum_{I \in \mathcal{D}} \frac{\langle f, h_I \rangle^2}{|I|^2} \mathbf{1}_I \,, \qquad \text{for} \qquad f(x) = \sum_I \frac{\langle f, h_I \rangle}{|I|} h_I(x).$$

We can now deduce the Chang-Wilson-Wolff inequality.

2.5. **Chang-Wilson-Wolff Inequality .** *For all Hilbert space valued martingales, we have*

$$\|f\|_{\exp(L^2)} \lesssim \|\mathrm{S}(f)\|_\infty \,.$$

Indeed, we have

$$\|f\|_p \lesssim \sqrt{p} \cdot \|\mathrm{S}(f)\|_p \lesssim \sqrt{p} \cdot \|\mathrm{S}(f)\|_\infty \,.$$

Taking $p \to \infty$, and using Proposition 2.2, we deduce the inequality above.

---

[1] We are only interested in measuring the behavior of functions for large values of f, so this requirement is sufficient. For $\alpha > 1$, we can insist upon this equality for all x.

[2] For $\alpha \ge 1$, we can take this as the definition for all $|x| \ge 0$.

In dimension 2, a *dyadic rectangle* is a product of dyadic intervals, thus an element of $\mathcal{D}^2$. A Haar function associated to R is the product of the Haar functions associated with each side of R, namely for $R_1 \times R_2$,

$$h_{R_1\times R_2}(x_1, x_2) := \prod_{t=1}^{2} h_{R_t}(x_t)\,.$$

See Figure 1. Below, we will expand the definition of Haar functions, so that we can describe a basis for $L^2([0,1]^2)$.

We will concentrate on rectangles of a fixed volume, contained in $[0,1]^2$. The notion of the square function is also useful in the two dimensional context. It has the form

$$\text{(2.6)} \qquad \operatorname{S}(f)^2 = \sum_{R\in\mathcal{D}^2} \frac{\langle f, h_R\rangle^2}{|R|^2}\mathbf{1}_R\,, \qquad \text{for} \qquad f(x) = \sum_{R\in\mathcal{D}^2} \frac{\langle f, h_R\rangle}{|R|} h_R(x)\,.$$

Jill Pipher (Pipher, 1986) observed the following extension of the Chang-Wilson-Wolff inequality.

2.7. **Two Parameter Chang-Wilson-Wolff Inequality .** *For functions f in the plane as in* (2) *we have*

$$\|f\|_{\exp(L)} \lesssim \|\operatorname{S}(f)\|_\infty\,.$$

Namely, in the case of two-parameters, the exponential integrability has been reduced by a factor of two. This follows from a two-fold application of the Littlewood-Paley inequalities, with best constants, for Hilbert space valued functions. Details can be found in (Pipher, 1986; Fefferman and Pipher, 1997; Bilyk and Lacey, 2008). In fact, we will need the following variant.

2.8. **Theorem.** *Let $n \geq 1$ be an integer. Suppose that f on the plane has the expansion*

$$f = \sum_{\substack{R\in\mathcal{D}^2\\ |R|=2^{-n}}} \frac{\langle f, h_R\rangle}{|R|} h_R\,.$$

That is, f is in the linear span of Haar functions with a fixed volume. Then, we have the estimate

$$\|f\|_{\exp(L^2)} \lesssim \|S(f)\|_\infty\,.$$

Thus, if f is in the linear span of a 'one-parameter' family of rectangles, we regain the exponential-squared integrability. The proof is straightforward. As the volumes of the rectangles are fixed, one need only apply the one-parameter Chang-Wilson-Wolff inequality in, say, the x_1 variable, holding the x_2 variable fixed.

The following simple proposition reduces the proof of Theorem 1.2 to the case $\alpha = 2$.

2.9. **Proposition.** *Suppose that for $A \geq 1$, we have*

$$\|f\|_{\exp(L^2)} \leq \sqrt{A}\,, \qquad \|f\|_\infty \leq A\,.$$

It follows that

$$\|f\|_{\exp(L^\alpha)} \le A^{1-1/\alpha}\,, \qquad 2 \le \alpha < \infty\,.$$

Bounded Mean Oscillation. We recall facts about dyadic *BMO* spaces, see (Chang and Fefferman, 1985; 1980).

We need to subtract some terms from D_N, as it is not necessarily in the span of the Haar functions as we have defined them. The deficiency is that standard Haar functions on the unit square have zero means in both directions. Hence, for a dyadic interval $I \in \mathcal{D}$, we also need to consider

$$h_I^1 = \mathbf{1}_I = |h_I|\,.$$

And set $h_I^0 = h_I$, where '0' stands for 'zero integral' and '1' for 'non-zero integral.' In the plane, for $\epsilon_1\,,\,\epsilon_2 \in \{0,1\}$ set

$$h_{R_1\times R_2}^{\epsilon_1,\epsilon_2}(x_1,x_2) = \prod_{j=1}^{2} h_{R_j}^{\epsilon_j}(x_j)\,.$$

We will sometimes write $h_R = h_R^{0,0}$ in order to simplify our notation. With these definitions we have the following *orthogonal* basis for $L^2([0,1]^2)$.

$$\{h_{[0,1]^2}^{1,1}\} \cup \{h_{[0,1]\times I}^{1,0}\,,\, h_{I\times[0,1]}^{1,0} \mid I \in \mathcal{D}\} \cup \{h_R^{0,0} \mid R \in \mathcal{D}^2\}\,.$$

There are couple of different BMO spaces that are relevant here. Let us begin with the variants of the more familiar C. Fefferman, one-parameter, dyadic BMO spaces.

2.10. **Definition.** Define the space BMO_1 to be those square integrable functions f in the span of $\{h_{I\times[0,1]}^{0,1} \mid I \in \mathcal{D}\}$ which satisfy

$$\|f\|_{\mathrm{BMO}_1} \coloneqq \sup_{J\in\mathcal{D}} \Bigl[|J|^{-1} \sum_{\substack{I\in\mathcal{D}\\ I\subset J}} \frac{\langle f, h_{I\times[0,1]}^{0,1}\rangle^2}{|I|}\Bigr]^{1/2} < \infty\,.$$

Define BMO_2 similarly, with the roles of the first and second coordinate reversed.

2.11. **Definition.** Dyadic Chang-Fefferman $\mathrm{BMO}_{1,2}$ is defined to be those square integrable functions f in the linear span of $\{h_R \mid R \in \mathcal{D}^2\}$, for which we have

$$\|f\|_{\mathrm{BMO}_{1,2}} \coloneqq \sup_{U\subset[0,1]^2} \Bigl[|U|^{-1} \sum_{\substack{R\in\mathcal{D}^2\\ R\subset U}} \frac{\langle f, h_R\rangle^2}{|R|}\Bigr]^{1/2} < \infty\,.$$

We stress that the supremum is over *all* measurable subsets $U \subset [0,1]^2$, not just rectangles.

It is well-known that these 'uniform square integrability' conditions imply that the corresponding functions enjoy higher moments. This is usually phrased as the John-Nirenberg inequalities, which we state here in their sharp exponential form.

The John-Nirenberg Estimates. *We have the following estimate for* $f \in \mathrm{BMO}_1$*, and* $\varphi \in \mathrm{BMO}_{1,2}$.

$$\|f\|_{\exp(L)} \lesssim \|f\|_{\mathrm{BMO}_1}$$

$$\|\varphi\|_{\exp(\sqrt{L})} \lesssim \|\varphi\|_{\mathrm{BMO}_{1,2}} \tag{2.12}$$

Note that in the second inequality, (2), the number of parameters has doubled, hence the exponential integrability has decreased by a factor of two. Of course, if the square function of f is bounded, one sees immediately that the functions are necessarily in *BMO*. And in this circumstance the Chang-Wilson-Wolff inequalities give an essential strengthening of the John-Nirenberg estimates.

Discrepancy. Below, we will refer to the two parts of the Discrepancy function as the 'linear' and the 'counting' part. Specifically, they are

$$L_N(\vec{x}) = N x_1 \cdot x_2 \,,$$

$$C_{\mathcal{P}}(\vec{x}) = \sum_{\vec{p} \in \mathcal{P}} \mathbf{1}_{[\vec{p},\vec{1})}(\vec{x}) \,.$$

Here, $\mathcal{P}$ is the subset of the unit square of cardinality N. In proving upper bounds on the Discrepancy function, one of course needs to capture a cancellation between these two, that is large enough to nearly completely cancel the nominal normalization by N.

We recall some definitions and facts about Discrepancy which are well represented in the literature, and apply to general selection of point sets, see (Roth, 1954; Schmidt, 1977; Beck and Chen, 1987).

We call a function f an r *function with parameter* $\vec{r} = (r_1, r_2)$ if $\vec{r} \in \mathbb{N}^2$, and

$$f = \sum_{R \in \mathcal{R}_{\vec{r}}} \varepsilon_R \, h_R \,, \qquad \varepsilon_R \in \{\pm 1\} \,,$$

where we set $\mathcal{R}_{\vec{r}} := \{R = R_1 \times R_2 \mid R \in \mathcal{D}^2 \,, R \subset [0,1]^2 \,,\ |R_t| = 2^{-r_t} \,,\ t = 1,2\}$. We will use $f_{\vec{r}}$ to denote a generic r function. A fact used without further comment is that $f_{\vec{r}}^2 \equiv 1$.

Let $|\vec{r}| = \sum_{t=1}^2 r_t = n$, which we refer to as the index of the r function. And let $\mathbb{H}_n^2 := \{\vec{r} \in \{0,1,\ldots,n\}^2 \mid |\vec{r}| = n\}$, i.e., the set of all $\vec{r}$'s such that rectangles in $\mathcal{R}_{\vec{r}}$ have area 2^{-n}. It is fundamental to the subject that $\sharp \mathbb{H}_n^2 = n+1$. We refer to $\{f_{\vec{r}} \mid r \in \mathbb{H}_n^2\}$ as hyperbolic r functions. The next four Propositions are standard.

2.13. **Proposition.** *For any selection* $\mathcal{A}_N$ *of* N *points in the unit cube the following holds. Fix* n *with* $2N < 2^n \leq 4N$*. For each* $\vec{r} \in \mathbb{H}_n^2$*, there is an* r *function* $f_{\vec{r}}$ *with*

$$\langle D_N, f_{\vec{r}} \rangle \gtrsim 1 \,.$$

Proof. There is a very elementary one dimensional fact: for all dyadic intervals I,

$$\int_0^1 x \cdot h_I(x)\, dx = \tfrac{1}{4} |I|^2 \,.$$

This immediately implies that

$$\langle x_1 \cdot x_2\,,\, h_R^{0,0}(x_1, x_2)\rangle = 4^{-2}|R|^2\,. \tag{2.14}$$

Thus, the inner product with the linear part of the Discrepancy function is completely straightforward. We have $\langle L, h_R^{0,0}\rangle \geq 4^{-2}N|R|^2 \geq 4|R|$ for $R \in \mathcal{R}_{\vec{r}}$ with $\vec{r} \in \mathbb{H}_n^2$.

Call a rectangle $R \in \mathcal{R}_{\vec{r}}$ *good* if R does **not** intersect $\mathcal{A}_N$, otherwise call it *bad*. Set

$$f_{\vec{r}} := \sum_{R \in \mathcal{R}_{\vec{r}}} \operatorname{sgn}(\langle D_N, h_R\rangle) h_R\,.$$

Each bad rectangle contains at least one point in $\mathcal{A}_N$, and $2^n \geq 2N$, so there are at least N good rectangles. Moreover, one should observe that the counting function $\sharp(\mathcal{A}_N \cap [0, \vec{x}))$ is orthogonal to h_R for each good rectangle R. That is,

$$\langle C_{\mathcal{A}_N}, h_R^{0,0}\rangle = 0\,, \qquad \text{whenever} \quad R \cap \mathcal{A}_N = \emptyset\,.$$

Critical to this property is the fact that Haar functions have mean zero on each line parallel to the coordinate axes.

Thus, by (2), for a *good* rectangle $R \in \mathcal{R}_{\vec{r}}$ we have

$$\langle D_N, h_R\rangle = -\langle L_N, h_R\rangle = -N\langle |[0, \vec{x})|, h_R(\vec{x})\rangle = -N2^{-2n-4} \lesssim -2^{-n}\,.$$

Hence, to complete the proof, we can estimate

$$\langle D_N, f_{\vec{r}}\rangle \geq \sum_{\substack{R \in \mathcal{R}_{\vec{r}} \\ R \text{ is good}}} |\langle D_N, h_R\rangle| \gtrsim 2^{-n}\sharp\{R \in \mathcal{R}_{\vec{r}} \mid R \text{ is good}\} \gtrsim 1\,.$$

□

2.15. **Proposition.** *Let $f_{\vec{s}}$ be any* r *function with $|\vec{s}| > n$. We have*

$$|\langle D_N, f_{\vec{s}}\rangle| \lesssim N2^{-|\vec{s}|}\,.$$

Proof. This is a brute force proof. Consider the linear part of the Discrepancy function. By (2), we have

$$|\langle L_N, f_{\vec{s}}\rangle| \lesssim N2^{-|\vec{s}|}\,,$$

as claimed.

Consider the part of the Discrepancy function that arises from the point set. Observe that for any point $\vec{x}_0$ in the point set, we have

$$|\langle \mathbf{1}_{[\vec{0}, \vec{x}_0)}, f_{\vec{s}}\rangle| \lesssim 2^{-|\vec{s}|}\,.$$

Indeed, of the different Haar functions that contribute to $f_{\vec{s}}$, there is at most one with non zero inner product with the function $\mathbf{1}_{[\vec{0}, \vec{x})}(\vec{x}_0)$ as a function of $\vec{x}$. It is the one rectangle which contains x_0 in its interior. Thus the inequality above follows. Summing it over the N points in the point set completes the proof of the Proposition. □

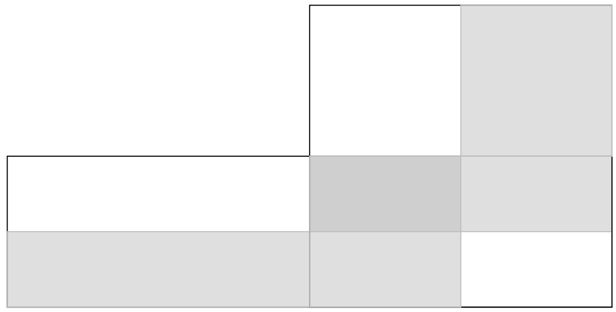

FIGURE 1. Two Haar functions.

2.16. **Proposition.** *In dimension $d = 2$ the following holds. Fix a collection of* r *functions* $\{f_{\vec r} \mid \vec r \in \mathbb{H}_n^2\}$. *Fix an integer* $2 \le v \le n$ *and* $\vec s$ *with* $0 \le s_1, s_2 \le n$ *and* $|\vec s| \ge n + v - 1$. *Let* $\operatorname{Count}(\vec s; v)$ *be the number of ways to choose distinct* $\vec r_1, \dots, \vec r_v \in \mathbb{H}_n^2$ *so that* $\prod_{w=1}^{v} f_{\vec r_w}$ *is an* $\vec s$ *function. We have*

$$\operatorname{Count}(\vec s; v) = \binom{|\vec s| - n - 1}{v - 2}.$$

Proof. Fix a vector $\vec s$ with $|\vec s| > n$, and suppose that

$$\prod_{w=1}^{v} f_{\vec r_w}$$

is an $\vec s$ function. Then, the maximum of the first coordinates of the $\vec r_w$ must be s_1, and similarly for the second coordinate. Thus, the vector s completely specifies two of the $\vec r_w$.

The remaining $v - 2$ vectors must be distinct, and take values in the first coordinate that are greater than $n - s_2$ and less than s_1. Hence there are at most $|\vec s| - n - 1$ possible choices for these vectors. This completes the proof. □

In two dimensions, the decisive product rule holds. If $R, R' \in \mathcal{D}^2$ are distinct, have the same area and non-empty intersection, then we have

$$h_R \cdot h_{R'} = \pm h_{R \cap R'}.$$

This rule is illustrated in Figure 1 and can be generalized as follows.

2.17. **Proposition.** *In dimension $d = 2$ the following holds. Let* $\vec r_1, \dots, \vec r_k$ *be elements of* $\mathbb{H}_n^2$ *where one of the vectors occurs an odd number of times. Then, the product* $\prod_{j=1}^{k} f_{\vec r}$ *is also an* r *function. If the* $\vec r_j$ *are distinct and* $k \ge 2$, *the product has index larger than* n.

3. The Digit-Scrambled van der Corput Set

In this section we introduce the digit-scrambled van der Corput set, that is, a variation of the classical van der Corput set described, e.g., in (Matoušek, 1999, Section 2.1), and prove some auxiliary lemmas that will help us exploit its properties. This set will be our main construction for the upper bounds in Theorems 1.2 and 1.5, although strictly speaking, Theorem 1.5 is satisfied by the standard van der Corput point distribution. The reasons we need this modified version of the van der Corput set will become clear by the end of this section.

First, we introduce some additional definitions and notations.

3.1. **Definition.** For $x \in [0,1)$ define $\mathrm{d}_i(x)$ to be the i'th digit in the binary expansion of x, that is

$$\mathrm{d}_i(x) = \lfloor 2^i x \rfloor \bmod 2.$$

3.2. **Definition.** For $x \in [0,1)$ we define the *digit reversal* function by means of the expression

$$\mathrm{d}_i\left(\mathrm{rev}_n(x)\right) = \begin{cases} \mathrm{d}_{n+1-i}(x), & i = 1, 2 \cdots n, \\ 0, & \text{otherwise,} \end{cases}$$

in other words, setting $\mathrm{d}_i(x) = x_i$, we have $\mathrm{rev}_n(0.x_1x_2...x_n) = 0.x_n...x_2x_1$.

3.3. **Definition.** Let $x, \sigma \in [0,1)$ where σ has n binary digits. We define the number $x \oplus \sigma$ as

$$\mathrm{d}_i(x \oplus \sigma) = \mathrm{d}_i(x) + \mathrm{d}_i(\sigma) \bmod 2,$$

i.e. the i^{th} digit of x changes if $\mathrm{d}_i(\sigma) = 1$ and stays the same if $\mathrm{d}_i(\sigma) = 0$. In the literature this operation is called *digit scrambling* or *digital shift*.

3.4. *Remark.* We stress at this point that when we define a digit scrambling we only use the first n binary digits of the number $\sigma \in [0,1)$. As a result, for each given positive integer n there are exactly 2^n such digital shifts, that is, the number of digital shifts is finite. The choice of a real number $\sigma \in [0,1)$ to represent this operation is just a matter of notational convenience.

We are now ready to define the digit-scrambled van der Corput set.

3.5. **Definition.** For an integer $n \geq 1$ and a number $\sigma \in [0,1)$ we define the σ-digit scrambled *van der Corput* set $\mathcal{V}_{n,\sigma}$ as

$$\mathcal{V}_{n,\sigma} = \{v_{n,\sigma}(\tau) : \ \tau = 0, 1, \ldots, 2^n - 1\},$$

where

$$v_{n,\sigma}(\tau) = \left(\frac{\tau}{2^n}, \mathrm{rev_n}\left(\frac{\tau}{2^n} \oplus \sigma\right)\right) + (2^{-n-1}, 2^{-n-1}).$$

It is clear that the digit-scrambled van der Corput set has cardinality $|\mathcal{V}_{n,\sigma}| = 2^n$. We should notice that the roles of x and y coordinates are symmetric, since we can write $\mathcal{V}_{n,\sigma} = \{(\mathrm{rev}_n(\tau/2^n \oplus \sigma'), \tau/2^n) + (2^{-n-1}, 2^{-n-1}) : \ \tau = 0, 1, \ldots, 2^n - 1\}$ with $\sigma' = \mathrm{rev}_n(\sigma)$.

With the notation introduced above, the standard van der Corput set

$$\mathcal{V}_n = \{(0.x_1x_2...x_n1, 0.x_n...x_2x_1 1) : x_i = 0, 1\}$$

is just $\mathcal{V}_n = \mathcal{V}_{n,0}$. Note that our definition differs from the classical by the shift $(2^{-n-1}, 2^{-n-1})$. This shift 'pads' the binary expansion of the elements by a final 1 in the $(n+1)^{\text{st}}$ place, and ensures that the average value of each coordinate is $\frac{1}{2}$:

$$2^{-n} \sum_{(x,y)\in\mathcal{V}_{n,\sigma}} x = 2^{-n} \sum_{(x,y)\in\mathcal{V}_{n,\sigma}} y = \frac{1}{2}. \tag{3.6}$$

This is just a technical modification that will simplify our formulas and calculations.

The following proposition describes which points of the van der Corput set $\mathcal{V}_{n,\sigma}$ fall into any given dyadic rectangle.

3.7. **Proposition.** *Let $k, l \in \mathbb{N}$ and $i \in \{0, 1, \ldots, 2^k - 1\}$, $j \in \{0, 1, \ldots, 2^l - 1\}$. Consider a dyadic rectangle*

$$R = \left[\frac{i}{2^k}, \frac{i+1}{2^k}\right) \times \left[\frac{j}{2^l}, \frac{j+1}{2^l}\right).$$

Then the set $\mathcal{V}_{n,\sigma} \cap R$ consists of the points $v_{n,\sigma}(\tau)$ where

$$\mathrm{d}_m\Big(\frac{\tau}{2^n}\Big) = \begin{cases} \mathrm{d}_m(\frac{i}{2^k}), & m = 1, 2 \cdots, k, \\ \mathrm{d}_{n+1-m}(\frac{j}{2^k}) + \mathrm{d}_m(\sigma) \mod 2, & m = n+1-l, \cdots, n. \end{cases}$$

Proof. Let (x, y) be any point $[0, 1)^2$. It is easy to see that $(x, y) \in R$ if and only if

$$\mathrm{d}_q(x) = \mathrm{d}_q\Big(\frac{i}{2^k}\Big) \quad \text{for all} \quad q = 1, 2, \ldots, k, \quad \text{and}$$

$$\mathrm{d}_r(y) = \mathrm{d}_r\Big(\frac{j}{2^l}\Big) \quad \text{for all} \quad r = 1, 2, \ldots, l.$$

The proposition is now a simple consequence of the structure of the van der Corput set. □

Some remarks are in order:

3.8. *Remarks.*
When $k+l < n$ there are exactly $2^{n-(k+l)}$ points of the van der Corput set inside the canonical rectangle R. Indeed, the conditions of Proposition 3.6 only specify the first k and last l binary digits of the $x-$coordinates of the points $v_{n,\sigma}(\tau)$.
When $k+l > n$ it might happen that the set of conditions in proposition 3.6 is void (observe that the system is overdetermined in this case).
Finally, when $k + l = n$, that is when the rectangle R has volume $|R| = 2^{-n}$, the system of equations in 3.6 gives a unique point of the van der Corput set inside R. So, for fixed n, the van der Corput set $\mathcal{V}_{n,\sigma}$ is a *net*: every dyadic rectangle of volume $N^{-1} = 2^{-n}$ contains exactly one point. This has the well-known consequence, see (Matoušek, 1999), that

$$\|D_N(\mathcal{V}_{n,\sigma})\|_\infty \lesssim \log N. \tag{3.9}$$

This fact is independent of the digit scrambling σ and holds in particular for the standard van der Corput set $\mathcal{V}_n$ ((van der Corput, 1935), (Roth, 1954)). In view of Schmidt's Theorem (1) this means that the van der Corput set is *extremal* in terms of measuring the Discrepancy function in L^∞. However, the same is not true if one is interested in meeting the lower bound in Roth's Theorem, that is, the standard van der Corput set $\mathcal{V}_n$ is not extremal in terms of measuring the Discrepancy function in L^2. The lemma below explains this fact. In particular it shows that the L^2 discrepancy of $\mathcal{V}_n$ is big because of a single 'zero-order' Haar coefficient, i. e. the mean $\int D_N$. The lemma also shows that digit scrambling provides a remedy for this shortcoming. This fact has been observed by Chen in (Chen, 1983) where the author uses digit scrambling in order to obtain the best possible L^p upper bounds for a general class of 'one point in a box' sets in general dimension (see the case $k + l = n$ in the remarks above). We also note that similar calculations, albeit slightly less general, have been carried out in (Halton and Zaremba, 1969). We include a proof of this Lemma for the sake of completeness.

3.10. **Lemma.** *We have*

$$\int_0^1\int_0^1 D_N(\mathcal{V}_{n,\sigma})\,dxdy = \frac{1}{4}\left(\frac{n}{2} - \sum_{k=1}^n \mathrm{d}_k(\sigma)\right).$$

In particular

$$\int_0^1\int_0^1 D_N(\mathcal{V}_n)\,dxdy = \frac{n}{8}.$$

On the other hand, if $\sum_{k=1}^n \mathrm{d}_k(\sigma) = n/2$, i.e. half of the digits are scrambled, then

$$\int_0^1\int_0^1 D_N(\mathcal{V}_{n,\sigma})\,dxdy = 0.$$

Proof. As usually, we write $N = 2^n$. We have

$$I := \int_0^1\int_0^1 D_N(\mathcal{V}_{n,\sigma})(x,y)\,dxdy = -N/4 + \sum_{\tau=0}^{N-1}\int_0^1\int_0^1 \mathbf{1}_{[0,x]\times[0,y]}\left(v_{n,\sigma}(\tau/N)\right)dxdy$$
$$= -N/4 + \sum_{\tau=0}^{N-1}\left(1 - \frac{\tau}{N} - \frac{1}{2N}\right)\left(1 - \mathrm{rev}_n\left(\frac{\tau}{N}\oplus\sigma\right) - \frac{1}{2N}\right).$$

Using (3) we get

$$I = -\frac{N}{4} + \frac{1}{2} - \frac{1}{4N} + \sum_{\tau=0}^{N-1}\frac{\tau}{N}\cdot\mathrm{rev}_n\left(\frac{\tau}{N}\oplus\sigma\right). \tag{3.11}$$

Now expand the sum above using the binary representation of the summands as follows:

$$\sum_{\tau=0}^{N-1}\frac{\tau}{N}\cdot\mathrm{rev}_n\left(\frac{\tau}{N}\oplus\sigma\right) = \sum_{\tau=0}^{N-1}\sum_{k=1}^{n}\sum_{l=1}^{n}\frac{\mathrm{d}_k\left(\frac{\tau}{N}\right)\mathrm{d}_l\left(\mathrm{rev}_n\left(\frac{\tau}{N}\oplus\sigma\right)\right)}{2^{k+l}}$$

$$= \sum_{\tau=0}^{N-1} \sum_{k=1}^{n} \sum_{l=1}^{n} \frac{\mathrm{d}_k\left(\frac{\tau}{N}\right) \mathrm{d}_{n+1-l}\left(\frac{\tau}{N} \oplus \sigma\right)}{2^{k+l}} \tag{3.12}$$

$$= \sum_{k=1}^{n} \sum_{l=1}^{n} \frac{1}{2^{k+l}} \sum_{\tau=0}^{N-1} \mathrm{d}_k\left(\frac{\tau}{N}\right) \mathrm{d}_{n+1-l}\left(\frac{\tau}{N} \oplus \sigma\right).$$

Finally observe that if $s, t \in \{1, 2, \ldots, n\}$ then

$$\sum_{\tau=0}^{N-1} \mathrm{d}_s\left(\frac{\tau}{N}\right) \mathrm{d}_t\left(\frac{\tau}{N} \oplus \sigma\right) = \begin{cases} \frac{N}{2}\left(1 - \mathrm{d}_s(\sigma)\right), & s = t \\ \frac{N}{4}, & s \neq t. \end{cases} \tag{3.13}$$

Indeed, when $s = t$, the terms in the sum above are non-zero exactly when $\mathrm{d}_s(\frac{\tau}{N}) = 1$ and $\mathrm{d}_s(\sigma) = 0$, and hence the first equality. The case $s \neq t$ is similar.

Using (3) and (3) we get

$$\sum_{\tau=0}^{N-1} \frac{\tau}{N} \operatorname{rev}_n\left(\frac{\tau}{N} \oplus \sigma\right) = \frac{n}{8} - \frac{1}{4} \sum_{k=1}^{n} \mathrm{d}_k(\sigma) + \frac{N}{4} - \frac{1}{2} + \frac{1}{4N},$$

which, combined with (3), completes the proof. □

Remark. We should point out that in (Kritzer and Pillichshammer, 2006) it has been shown that the L^2 norm of the Discrepancy of the digit-scrambled van der Corput set depends only on the number of 1's in σ, and not their distribution.

4. Haar Coefficients for the Digit-Scrambled van der Corput Set

In this section we will work with the digit-scrambled van der Corput set $\mathcal{V}_{n,\sigma}$ as defined in Section 3, where $\sigma \in [0, 1)$ is arbitrary and $N = 2^n$. We will just write D_N for the discrepancy function of $\mathcal{V}_{n,\sigma}$. The following Lemma records the main estimate for the Haar coefficients of D_N and is the core of the proof for the upper bounds in Theorems 1.2 and 1.5.

4.1. **Lemma.** *For any dyadic rectangle $R \in \mathcal{D}^2$ we have*

$$|\langle D_N, h_R \rangle| \lesssim \frac{1}{N}.$$

We need to consider dyadic rectangles of the form $R = \left[\frac{i}{2^k}; \frac{i+1}{2^k}\right) \times \left[\frac{j}{2^l}; \frac{j+1}{2^l}\right)$, where $k, l \in \mathbb{N}$ and $i \in \{0, 1, \ldots, 2^k - 1\}$, $j \in \{0, 1, \ldots, 2^l - 1\}$. The proof will be divided in two cases, depending on whether the volume of R is 'big' or 'small'.

We will use an auxiliary function to help us write down formulas for the inner product of the counting part with the Haar function corresponding to the rectangle R. In particular, $\phi : \mathbb{R} \to \mathbb{R}$ is the periodic function

$$\phi(x) = \begin{cases} \{x\}, & 0 < \{x\} < \frac{1}{2} \\ 1 - \{x\}, & \frac{1}{2} < \{x\} < 1, \end{cases}$$

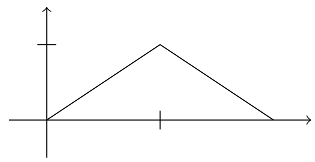

FIGURE 2. The graph of the function ϕ.

where $\{x\}$ is the fractional part of x. Observe that the function ϕ is the periodic extension of the anti-derivative of the Haar function on [0,1]. See Figure 2.

Let $p = (p_x, p_y) \in [0, 1)^2$. A moment's reflection allows us to write

$$\langle \mathbf{1}_{[\vec{p},\vec{1})}, h_R\rangle = \begin{cases} |R|\phi(2^k p_x)\phi(2^l p_y), & p \in R, \\ 0, & \text{otherwise.} \end{cases} \tag{4.2}$$

We also record two simple properties of the function ϕ that will be useful in what follows. First, for $x \in \mathbb{R}$,

$$\phi(x) + \phi\Big(x \oplus \frac{1}{2}\Big) = \frac{1}{2}. \tag{4.3}$$

Second, ϕ is a 'Lipschitz' function with constant 1. For $x, y \in \mathbb{R}$,

$$\big|\phi(y) - \phi(x)\big| \le |\{y\} - \{x\}|. \tag{4.4}$$

Proof of Lemma 4.1 when $|R| < \frac{4}{N}$. We fix a dyadic rectangle R with $|R| < \frac{4}{N}$. We treat the linear part and the counting part separately.

For the linear part we have that

$$\langle L_N, h_R\rangle = \frac{N|R|^2}{4^2} \lesssim \frac{1}{N}.$$

Now notice that since $k + l > n - 2$, there are at most 2 points in $\mathcal{V}_{n,\sigma} \cap R$. Since ϕ is obviously bounded by 1, formula (4) implies

$$|\langle C_{\mathcal{V}_{n,\sigma}}, h_R\rangle| \le |R| \sum_{p \in \mathcal{V}_{n,\sigma} \cap R} \phi(2^k p_x)\phi(2^l p_y) \le 4|R| \lesssim \frac{1}{N}.$$

Summing up the estimates for the linear and the counting part completes the proof. □

Proof of Lemma 4.1 when $|R| \ge \frac{4}{N}$. The proof of the case $|R| \ge \frac{4}{N}$ is much more involved as this is the typical case where the rectangle contains 'many' points of the point set $\mathcal{V}_{n,\sigma}$. Before going into the details of the proof we will discuss the structure of the set $R \cap \mathcal{V}_{n,\sigma}$ in order to organize and simplify the calculations that follow.

First, notice that the condition $|R| \ge \frac{4}{N}$ implies that $n - (k + l) \ge 2$. In other words, there are at least 4 points in the set $R \cap \mathcal{V}_{n,\sigma}$ according to Proposition 3.6 and Remark 3.7. To be more precise, let us look at a point $p = (x, y) \in \mathcal{V}_{n,\sigma}$. The x-coordinate can be written in the

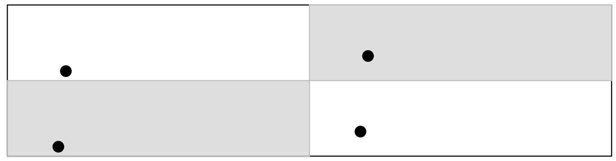

FIGURE 3. The quadruple Q.

form $x = 0.x_1x_2\dots x_n1$, where $x_i = \mathrm{d}_i(x)$, for $i = 1, 2, \dots, n$. The first k and the last l binary digits of x are determined by the fact that $x \in R$ (Proposition 3.6). That leaves us with at least 2 'free' digits for x

$$x = 0.x_1 \dots x_k, *, \dots, *, x_{n-l+1} \dots x_n 1.$$

We group all points in $\mathcal{V}_{n,\sigma} \cap R$ in quadruples according to the choices for the first and last 'free' digits x_{k+1} and x_{n-l}. In particular, we consider *quadruples* (Q) of points in $\mathcal{V}_{n,\sigma} \cap R$ with x-coordinates of the form:

$$\text{(Q)} \qquad \begin{array}{l} 0.x_1 \dots x_k\ 0\ x_{k+2} \dots, x_{n-l-1}\ 0\ x_{n-l+1} \dots x_n 1, \\ 0.x_1 \dots x_k\ 0\ x_{k+2} \dots, x_{n-l-1}\ 1\ x_{n-l+1} \dots x_n 1, \\ 0.x_1 \dots x_k\ 1\ x_{k+2} \dots, x_{n-l-1}\ 0\ x_{n-l+1} \dots x_n 1, \\ 0.x_1 \dots x_k\ 1\ x_{k+2} \dots, x_{n-l-1}\ 1\ x_{n-l+1} \dots x_n 1. \end{array}$$

There are exactly $2^{n-(k+l)-2} = \frac{N|R|}{4}$ such quadruples. Let's index the quadruples Q arbitrarily as Q_r, $r = 1, 2, \dots, \frac{N|R|}{4}$. Observe that we can write

$$\langle D_N, h_R \rangle = \sum_{p \in \mathcal{V}_{n,\sigma} \cap R} \langle \mathbf{1}_{[\vec{p},\vec{1})}, h_R \rangle - \frac{N|R|^2}{16} = \sum_{r=1}^{\frac{N|R|}{4}} \Big(\sum_{p \in Q_r} \langle \mathbf{1}_{[\vec{p},\vec{1})}, h_R \rangle - \frac{|R|}{4} \Big). \tag{4.5}$$

The following Proposition exploits large cancellation within these quadruples.

4.6. **Proposition.**

$$\Big| \sum_{p \in Q_r} \langle \mathbf{1}_{[\vec{p},\vec{1})}, h_R \rangle - \frac{|R|}{4} \Big| \lesssim \frac{1}{N^2 |R|}.$$

Let assume Proposition 4.2 for a moment in order to complete the proof of Lemma 4.1. Indeed, Proposition 4.2 together with equation (4) immediately yield

$$\langle D_N, h_R \rangle \lesssim \sum_{r=1}^{\frac{N|R|}{4}} \frac{1}{N^2|R|} \lesssim \frac{1}{N}.$$

This completes the proof modulo Proposition 4.2. □

Proof of Proposition 4.2. For the proof of the proposition we will fix a $Q = Q_r$ and suppress the index r since it does not play any role. Suppose $p = (u, v)$ is any of the points with x-coordinate as in (Q) and y-coordinate v such that $p \in \mathcal{V}_{n,\sigma}$. Then it is easy to see that the quadruple (Q) consists of the four points which can be written in the form:

$$\text{(Q)} \qquad \begin{cases} (u, v), \\ (u \oplus 2^{-k-1}, v \oplus 2^{-n+k}), \\ (u \oplus 2^{-n+l}, v \oplus 2^{-l-1}), \\ (u \oplus 2^{-n+l} \oplus 2^{-k-1}, v \oplus 2^{-n+k} \oplus 2^{-l-1}). \end{cases}$$

See also Figure 3.

We invoke equation (4) to write

$$\sum_{p \in Q} \langle \mathbf{1}_{[\vec{p}, \vec{1})}, h_R \rangle - \frac{|R|}{4} = |R| \Big(\sum_{p \in Q} \phi(2^k p_x) \phi(2^l p_y) - \frac{1}{4} \Big) =: |R| \Big(\Sigma - \frac{1}{4} \Big). \tag{4.7}$$

We have

$$\begin{aligned} \Sigma &= \phi(2^k u)\phi(2^l v) \\ &\quad + \phi(2^k u \oplus \frac{1}{2})\phi(2^l(v \oplus 2^{-n+k})) \\ &\quad + \phi(2^k(u \oplus 2^{-n+l}))\phi(2^l v \oplus \frac{1}{2}) \\ &\quad + \phi(2^k u \oplus 2^k \cdot 2^{-n+l} \oplus \frac{1}{2})\phi(2^l v \oplus 2^l \cdot 2^{-n+k} \oplus \frac{1}{2}). \end{aligned}$$

Using equation (4) we get

$$\Sigma = \frac{1}{4} + \Big[\phi(2^k u) - \phi(2^k(u \oplus 2^{-n+l})) \Big] \Big[\phi(2^l v) - \phi(2^l(v \oplus 2^{-n+k})) \Big].$$

Finally, using the fact the the function ϕ is Lipschitz (4) we have

$$\Big| \Sigma - \frac{1}{4} \Big| \leq (2^{-n+l+k})^2 = \frac{1}{N^2 |R|^2}.$$

This estimate together with equation (4) completes the proof. □

Lemma 4.1 has an analogue in the case of Haar functions $h^{1,0}_{[0,1] \times I}$ and $h^{0,1}_{I \times [0,1]}$, where $I \in \mathcal{D}$. Observe also that the inner product that corresponds to $h^{1,1}_{[0,1]^2}$ is the content of Lemma 3.8 of the previous section.

4.8. **Lemma.** *For $I \in \mathcal{D}$ we have the estimates*

$$|\langle D_N, h^{0,1}_{I \times [0,1]} \rangle| \lesssim |I|,$$

$$|\langle D_N, h^{1,0}_{[0,1] \times I} \rangle| \lesssim |I|.$$

Proof. It suffices to prove just the first estimate in the statement of the Lemma. The proof proceeds in a more or less analogous fashion as the proof of Lemma 4.1. We fix a dyadic interval $I = \left[\frac{i}{2^k}, \frac{i+1}{2^k}\right)$ and write $h_I = h^{0,1}_{I\times[0,1]}$. We need an analogue of formula (4) which in this case becomes

$$\langle \mathbf{1}_{[\vec{p},\vec{1})}, h_I\rangle = \begin{cases} |I|\phi(2^k p_x)(1-p_y), & p_x \in I, \\ 0, & \text{otherwise.} \end{cases} \tag{4.9}$$

As in the proof of Lemma 4.1, we need to consider separately the case of small volume and large volume rectangles. The small volume case here is $|I| \le \frac{2}{N}$. Note that in this case there are at most $2^{n-k} \le 2$ points of the van der Corput set whose x coordinate lies in I. Using equation (4) we trivially get the desired estimate as in the proof of the corresponding case of Lemma 4.1.

We now turn to the main part of the proof, namely the estimate

$$|\langle D_N, h^{1,0}_{I\times[0,1]}\rangle| \lesssim |I|,$$

when $|I| > \frac{2}{N}$. Instead of the quadruples (Q), we now group the points of the van der Corput set with x-coordinate in I, into *pairs* (P) of the form:

$$\begin{aligned} &0.x_1 \dots x_k\ 0\ x_{k+2} \dots x_n 1, \\ &0.x_1 \dots x_k\ 1\ x_{k+2} \dots x_n 1. \end{aligned} \tag{P}$$

If (u, v) is one of the two points in (P), we also have the description:

$$\begin{cases} (u, v), \\ (u \oplus 2^{-k-1}, v \oplus 2^{-n+k}). \end{cases} \tag{P}$$

There are 2^{n-k-1} such pairs and let's index them arbitrarily as P_r, $r = 1, 2, \dots, 2^{n-k-1}$. We write

$$\langle D_N, h_I\rangle = \sum_{p\in\mathcal{V}_{n,\sigma}\cap I\times[0,1]} \langle \mathbf{1}_{[\vec{p},\vec{1})}, h_I\rangle - \frac{N|I|^2}{8} = \sum_{r=1}^{2^{n-k-1}} \sum_{p\in P_r} \langle \mathbf{1}_{[\vec{p},\vec{1})}, h_I\rangle - \frac{N|I|^2}{8}.$$

Now for any pair (P) we use (4) to write

$$\begin{aligned} \sum_{p\in P} \langle \mathbf{1}_{[\vec{p},\vec{1})}, h_I\rangle &= |I|\,\phi(2^k u)(1-v) + |I|\,\phi(2^k(u \oplus 2^{-k-1}))(1 - v \oplus 2^{-n+k}) \\ &= |I|\left[\phi(2^k u) + \phi(2^k u \oplus 2^{-1})\right](1-v) \\ &\quad + |I|\,\phi(2^k u \oplus 2^{-1})\,(v - v \oplus 2^{-n+k}) \\ &= \frac{1}{2}|I|(1-v) + |I|\,\phi(2^k u \oplus 2^{-1})\,(v - v \oplus 2^{-n+k}). \end{aligned}$$

where in the last equality we have used (4). Using the fact that $|v - v \oplus 2^{-n+k}| = 2^{-n+k}$ and assuming $\mathrm{d}_{n-k}(v) = 0$, it is routine to check that

$$\langle D_N, h_I \rangle = |I| \left\{ \frac{1}{2} \sum_{r=1}^{2^{n-k-1}} (1 - v_r) - 2^{n-k-3} + O(1) \right\}, \tag{4.10}$$

where v_r are y-coordinates of the form

$$v_r = 0.Y_1 \dots Y_{n-k-1} 0 y_{n-k+1} \dots y_n 1$$

The digits y_{n-k+1} up to y_n are fixed because of the digit reversal structure of the van der Corput set. We can then estimate the sum in the previous expression as follows:

$$\sum_{r=1}^{2^{n-k-1}} (1 - v_r) = 2^{n-k-1} - \frac{1}{2} 2^{n-k-1} \Big(1 - 2^{-n+k+1} \Big) + O(1) = 2^{n-k-2} + O(1).$$

Substituting in (4) we get

$$\langle D_N, h_I \rangle = |I| \left\{ \frac{1}{2} \Big(2^{n-k-2} + O(1) \Big) - 2^{n-k-3} + O(1) \right\} \lesssim |I|,$$

which completes the proof. □

5. BMO Estimates for the Discrepancy Function

This section is devoted to the proofs of Theorems 1.4 and 1.5. We recall that the Dyadic Chang-Fefferman $\mathrm{BMO}_{1,2}$ is defined to consist of those square integrable functions f in the linear span of $\{h_R \mid R \in \mathcal{D}^2\}$, for which we have

$$\|f\|_{\mathrm{BMO}_{1,2}} := \sup_{U \subset [0,1]^2} \left[|U|^{-1} \sum_{\substack{R \in \mathcal{D}^2 \\ R \subset U}} \frac{\langle f, h_R \rangle^2}{|R|} \right]^{1/2} < \infty \,.$$

We begin with the proof of Theorem 1.4 which is essentially just a repetition of the argument used in Proposition 2.11.

Proof of Theorem 1.4. We fix a distribution $\mathcal{A}_N$ of N points in the unit square and take n such that $2N < 2^n \le 4N$. For the special choice of $U = [0,1]^2$ we have

$$\|D_N\|^2_{\mathrm{BMO}_{1,2}} \ge \sum_{\vec{r} \in \mathbb{H}_n} \sum_{\substack{R \in \mathcal{R}_{\vec{r}} \\ R \cap \mathcal{A}_N = \emptyset}} \frac{\langle D_N, h_R \rangle^2}{|R|}.$$

Consider a rectangle $R \in \mathcal{R}_{\vec{r}}$ which does not contain any points of $\mathcal{A}_N$. Then

$$\langle D_N, h_R \rangle = -\langle L_N, h_R \rangle = -\frac{|R|^2}{4^2}.$$

As a result,

$$\|D_N\|^2_{\mathrm{BMO}_{1,2}} \gtrsim \sum_{\vec{r}\in\mathbb{H}_n} \sum_{\substack{R\in\mathcal{R}_{\vec{r}} \\ R\cap\mathcal{A}_N=\emptyset}} N^2|R|^3 \gtrsim \frac{1}{N}\sum_{\vec{r}\in\mathbb{H}_n} \sharp\{R\in\mathcal{R}_{\vec{r}},\ R\cap\mathcal{A}_N=\emptyset\}.$$

For fixed $\vec{r}\in\mathbb{H}_n$ we have $\sharp\{R\in\mathcal{R}_{\vec{r}}, R\cap\mathcal{A}_N=\emptyset\}\geq N$, arguing as in the proof of Proposition 2.11. Thus we get

$$\|D_N\|^2_{\mathrm{BMO}_{1,2}} \gtrsim \sum_{\vec{r}\in\mathbb{H}_n} 1 \gtrsim n.$$

This completes the proof since $n\simeq\log N$. □

We proceed with the proof of the upper bound in Theorem 1.5. Our extremal set of cardinality $N=2^n$ will be $\mathcal{V}_{n,\sigma}$ for arbitrary $\sigma\in[0,1)$, as defined in Definition 3.5. We will just write D_N for the Discrepancy function of the digit-scrambled van der Corput set.

Proof of Theorem 1.5. We fix a measurable set $U\subset[0,1]^2$ and consider only rectangles R in the family $\{R\in\mathcal{D}^2, R\subset U\}$. We will sometimes suppress the fact that our rectangles are contained in U to simplify the notation.

The are two estimates that are relevant here, one for large rectangles and one for small volume rectangles. For the large volume case, $|R|\geq 2^{-n}$, we have

$$\begin{aligned}|U|^{-1}\sum_{|R|\geq 2^{-n}}\frac{\langle D_N,h_R\rangle^2}{|R|} &= |U|^{-1}\sum_{k=0}^{n}\sum_{\vec{r}\in\mathbb{H}_k}\sum_{R\in\mathcal{R}_{\vec{r}}}\frac{\langle D_N,h_R\rangle^2}{|R|}\\ &\lesssim N^{-2}|U|^{-1}\sum_{k=0}^{n}2^k\sum_{\vec{r}\in\mathbb{H}_k}\sum_{R\in\mathcal{R}_{\vec{r}}}1,\end{aligned}$$

where we have used the estimate $\langle D_N,h_R\rangle\lesssim\frac{1}{N}$ of Proposition 4.1. Now observe that for fixed k and $\vec{r}\in\mathbb{H}_k$ there are at most $2^k|U|$ rectangles $R\in\mathcal{R}_{\vec{r}}$ contained in U. Furthermore, there are k choices for the 'geometry' $\vec{r}\in\mathbb{H}_k$. We thus get

$$|U|^{-1}\sum_{|R|\geq 2^{-n}}\frac{\langle D_N,h_R\rangle^2}{|R|}\lesssim N^{-2}\sum_{k=0}^{n}k(2^k)^2\lesssim\frac{n(2^n)^2}{N^2}=n.$$

In the small volume term we treat the linear and the counting parts separately.

For the linear part we use (2) to get $\langle L_N,h_R\rangle=4^{-2}N|R|^2$. So we have

$$\begin{aligned}|U|^{-1}\sum_{|R|<2^{-n}}\frac{\langle L_N,h_R\rangle^2}{|R|} &= |U|^{-1}\sum_{k=n+1}^{\infty}\sum_{\vec{r}\in\mathbb{H}_k}\sum_{R\in\mathcal{R}_{\vec{r}}}\frac{\langle L_N,h_R\rangle^2}{|R|}\\ &\simeq N^2|U|^{-1}\sum_{k=n+1}^{\infty}\sum_{\vec{r}\in\mathbb{H}_k}(2^{-k})^3\sum_{R\in\mathcal{R}_{\vec{r}}}1.\end{aligned}$$

Now arguing as in the large volume case we have $\sum_{R\in\mathcal{R}_{\vec{p}}} 1 \lesssim 2^k|U|$, and thus

$$|U|^{-1} \sum_{|R|<2^{-n}} \frac{\langle L_N, h_R\rangle^2}{|R|} \lesssim N^2 \sum_{k=n+1}^{\infty} k(2^{-k})^2 \lesssim n.$$

It remains to bound the counting part that corresponds to small volume rectangles, i.e.

$$|U|^{-1} \sum_{|R|<2^{-n}} \frac{\langle C_{\mathcal{V}_{n,\sigma}}, h_R\rangle^2}{|R|}.$$

Let $\mathcal{R}$ be the maximal dyadic rectangles R of area at most 2^{-n}, contained inside U, and such that h_R has non-zero inner product with the counting part. It is essential to note that

$$\text{(5.1)} \qquad \sum_{R\in\mathcal{R}} |R| \lesssim n|U|\,.$$

Indeed, for each rectangle $R \in \mathcal{R}$, the function h_R is, as we have observed, orthogonal to each $\mathbf{1}_{[\vec{p},\vec{1})}$ with $\vec{p}$ not in the interior of R. Thus, R must contain one element of the van der Corput set in its interior. On the other hand $\mathcal{V}_{n,\sigma}$ is a net so R contains exactly one point. Now look at all the rectangles in $R \in \mathcal{R}$, $R = R_x \times R_y$, with a fixed side length $|R_x|$. The length of this side must be at least 2^{-n} in order for the rectangle to contain a point of the van der Corput set in its interior, so there are at most n choices for $|R_x|$. On the other hand, the rectangles in $\mathcal{R}$ with the same side length must be disjoint since they are maximal and dyadic. Since they are all contained in U, their union has volume at most U. Summing over all possible side lengths $|R_x|$ proves (5).

Now, we can write

$$|U|^{-1} \sum_{|R|<2^{-n}} \frac{\langle C_{\mathcal{V}_{n,\sigma}}, h_R\rangle^2}{|R|} \leq |U|^{-1} \sum_{R\in\mathcal{R}} \sum_{R'\subseteq R} \frac{\langle C_{\mathcal{V}_{n,\sigma}}, h_{R'}\rangle^2}{|R'|}.$$

Note that we have inequality instead of equality, since a rectangle R can be contained in several maximal rectangles. However, this does not create any problem.

Let $R \in \mathcal{R}$ be fixed and let $\vec{p}_R$ be the unique point of $\mathcal{V}_{n,\sigma}$ contained in R. We can use Bessel's inequality to bound the inner sum:

$$\sum_{R'\subseteq R} \frac{\langle C_{\mathcal{V}_{n,\sigma}}, h_{R'}\rangle^2}{|R'|} \leq \|\mathbf{1}_{(\vec{p}_R,\vec{1}]}\|^2_{L^2(R)} \leq |R|\,.$$

Thus, by (5)

$$|U|^{-1} \sum_{|R|<2^{-n}} \frac{\langle C_{\mathcal{V}_{n,\sigma}}, h_R\rangle^2}{|R|} \lesssim |U|^{-1} \sum_{R\in\mathcal{R}} |R| \lesssim n.$$

The proof is finished, since we have shown that for any measurable set $U \subset [0,1]^2$

$$\Big(|U|^{-1} \sum_{\substack{R\in\mathcal{D}^2 \\ R\subset U}} \frac{\langle D_N, h_R\rangle^2}{|R|}\Big)^{\frac{1}{2}} \lesssim n^{\frac{1}{2}} \simeq \sqrt{\log N}.$$

□

6. The $\exp(L^\alpha)$ Estimates for the Discrepancy Function.

6.1. **Lower bound: The Proof of Theorem 1.1.** The proof is by way of duality and is very similar to Halász's proof (Halász, 1981) of Schmidt's Theorem, see (1). Fix the point distribution $\mathcal{A}_N \subset [0,1]^2$. Set $2N < 2^n \leq 4N$, so that $n \simeq \log N$. Proposition 2.11 provides us with r functions $f_{\vec r}$ for $\vec r \in \mathbb{H}_n^2$. Let $\mathbb{G}_N^2 \subset \mathbb{H}_N^2$ be those elements of $\mathbb{H}_N^2$ whose first coordinate is a multiple of a sufficiently large integer a. We construct the following functions:

$$\Psi := \prod_{\vec r\in\mathbb{G}_N^2} (1+f_{\vec r}), \qquad \widetilde{\Psi} := \Psi - 1.$$

The 'product rule' 2.14 easily implies that Ψ is a positive function of L^1 norm one. In fact, letting $g = \sharp\,\mathbb{G}_n^2$, it is clear that

$$\Psi = 2^g \mathbf{1}_E\,, \qquad \mathbb{P}(E) = 2^{-g}\,.$$

Therefore, by Proposition 2.4,

$$\|\widetilde{\Psi}\|_{L(\log L)^{1/\alpha}} \simeq g^{1/\alpha} \simeq n^{1/\alpha}\,.$$

The fact that $\langle D_N, \widetilde{\Psi}\rangle \gtrsim n$ is well-known (Halász, 1981), (Matoušek, 1999). In fact, if we expand

$$\widetilde{\Psi} = \sum_{k=1}^{g} \Psi_k\,,$$

$$\Psi_k = \sum_{\{\vec r_1,\dots,\vec r_k\}\subset\mathbb{G}_n^2} \prod_{\ell=1}^{k} f_{\vec r_\ell}\,,$$

then, using the 'product rule' 2.14, it is not hard to see that we have

$$\langle D_N, \Psi_1\rangle \gtrsim g \gtrsim \frac{n}{a}\,,$$

and the other, higher order terms can be summed up, using Propositions 2.12 and 2.13, to give a much smaller estimate for a sufficiently large.

Thus, we can estimate

$$n \lesssim \langle D_N, \widetilde{\Psi}\rangle \lesssim \|D_N\|_{\exp(L^\alpha)} \cdot n^{1/\alpha}\,,$$

and so Theorem 1.1 holds.

6.2. **Upper bound: The Proof of Theorem 1.2 in the case that $N = 2^n$.** In this section we shall obtain the upper bound of the $\exp(L^2)$ norm of the discrepancy of the digit-scrambled van der Corput set. We shall consider the case of $N = 2^n$, leaving the general case to later. Lemma 3.8 tells us that we should choose $\mathcal{V}_{n,\sigma}$ with half the digits 'scrambled', i.e. $\sum_{i=1}^{n} \mathrm{d}_i(\sigma) = \lfloor n/2 \rfloor$ – this will be the only restriction on σ and for simplicity we shall assume that n is even. We expand D_N in the Haar series and break the expansion into several parts (in view of our choice of σ, $h^{1,1}$ does not play a role in the expansion):

$$D_N = \sum_{R\in\mathcal{D}^2} \frac{\langle D_N, h_R\rangle}{|R|} h_R + \sum_{R=I\times[0,1]} \frac{\langle D_N, h_R^{0,1}\rangle}{|R|} h_R^{0,1} + \sum_{R=[0,1]\times I} \frac{\langle D_N, h_R^{1,0}\rangle}{|R|} h_R^{1,0}$$

$$(6.1) \qquad = \sum_{R:|R|>2^{-n}} \frac{\langle D_N, h_R\rangle}{|R|} h_R + \sum_{R:|R|\le 2^{-n}} \frac{\langle C_N, h_R\rangle}{|R|} h_R - \sum_{R:|R|\le 2^{-n}} \frac{\langle L_N, h_R\rangle}{|R|} h_R$$

$$(6.2) \qquad + \sum_{R=I\times[0,1]} \frac{\langle D_N, h_R^{0,1}\rangle}{|R|} h_R^{0,1} + \sum_{R=[0,1]\times I} \frac{\langle D_N, h_R^{1,0}\rangle}{|R|} h_R^{1,0}$$

For the first sum in the expansion (6.2) above we have:

$$\begin{aligned}
\Big\| \sum_{R:|R|>2^{-n}} \frac{\langle D_N, h_R\rangle}{|R|} h_R \Big\|_{\exp(L^2)} &\le \sum_{k=0}^{n-1} \Big\| \sum_{R:|R|=2^{-k}} \frac{\langle D_N, h_R\rangle}{|R|} h_R \Big\|_{\exp(L^2)} \\
&\lesssim \sum_{k=0}^{n-1} \Big\| \Big(\sum_{R:|R|=2^{-k}} \frac{\langle D_N, h_R\rangle^2}{|R|^2} \mathbf{1}_R \Big)^{\frac12} \Big\|_{\infty} \\
&\lesssim \sum_{k=0}^{n-1} \frac{1}{N} \cdot \sqrt{k+1} \cdot 2^k \approx \sqrt{n},
\end{aligned}$$

where we have used the hyperbolic version of the Chang-Wilson-Wolff inequality (Theorem 2.7), the estimate of the Haar coefficients of D_N (Lemma 4.1), and the fact that each point in $[0,1]^2$ lives in $k+1$ dyadic rectangles of volume 2^{-k}.

The last sum in (6.2) is easy to estimate. Since $\langle L_N, h_R\rangle = 4^{-d} N |R|^2$, we have:

$$\begin{aligned}
\Big\| \sum_{R:|R|\le 2^{-n}} \frac{\langle L_N, h_R\rangle}{|R|} h_R \Big\|_{\exp(L^2)} &\le 4^{-d} \sum_{k=n}^{\infty} \Big\| \sum_{R:|R|=2^{-k}} N 2^{-k} h_R \Big\|_{\exp(L^2)} \\
&\lesssim N \sum_{k=n}^{\infty} 2^{-k} \Big\| \Big(\sum_{R:|R|=2^{-k}} \mathbf{1}_R \Big)^{\frac12} \Big\|_{\infty} \\
&\lesssim N \sum_{k=n}^{\infty} \sqrt{k+1} \cdot 2^{-k} \approx \sqrt{n},
\end{aligned}$$

where we have once again applied Theorem 2.7.

The second sum in (6.2) is the hardest. We consider rectangles R of volume $|R| \le 2^{-n}$. Recall that, in order for $\langle C_N, h_R \rangle$ to be non-zero, R must contain points of $\mathcal{V}_{n,\sigma}$ in the interior. The structure of the van der Corput set then implies that we must at least have $|R_1|, |R_2| \ge 2^{-n}$. For each such rectangle R, one can find a unique 'parent': a dyadic rectangle $\widetilde{R} \subset [0,1]^2$ with $|\widetilde{R}| = 2^{-n}$, $\widetilde{R}_1 = R_1$, and $R \subset \widetilde{R}$. We can now write

$$\Big\| \sum_{R : |R| < 2^{-n}} \frac{\langle C_N, h_R \rangle}{|R|} h_R \Big\|_p = \Big\| \sum_{k=0}^{n} \sum_{\substack{\widetilde{R} : |\widetilde{R}| = 2^{-n} \\ |\widetilde{R}_1| = 2^{-k}}} \sum_{\substack{R \subset \widetilde{R} \\ R_1 = \widetilde{R}_1}} \frac{\langle C_N, h_R \rangle}{|R|} h_R \Big\|_p \tag{6.3}$$

A given rectangle $\widetilde{R}$ as above contains precisely one point (p_1, p_2) from the set $\mathcal{V}_{n,\sigma}$. Thus,

$$\sum_{\substack{R \subset \widetilde{R} \\ R_1 = \widetilde{R}_1}} \frac{\langle C_N, h_R \rangle}{|R|} h_R(x_1, x_2) = C_{\widetilde{R}}(x_2) \frac{\langle h_{\widetilde{R}_1}, \mathbf{1}_{[p_1,1]} \rangle}{|\widetilde{R}_1|} h_{\widetilde{R}_1}(x_1), \tag{6.4}$$

where

$$C_{\widetilde{R}}(x_2) = \begin{cases} \sum_{I \subset \widetilde{R}_2} \frac{\langle h_I, \mathbf{1}_{[p_2,1]} \rangle}{|I|} h_I(x_2) = \mathbf{1}_{[p_2,1]}(x_2) - \int_{\widetilde{R}_2} \mathbf{1}_{[p_2,1]}(x) dx / |\widetilde{R}_2|, \ x_2 \in \widetilde{R}_2, \\ 0, \ x_2 \notin \widetilde{R}_2. \end{cases}$$

In any case, we have $|C_{\widetilde{R}}(x_2)| \le 2$. Now we fix $x_2 \in [0,1]$. For fixed x_2 and $\widetilde{R}_1$, there is a unique $\widetilde{R}$ such that the sum in (6.2) is non-zero. Thus, using (6.2)

$$\begin{aligned} \sum_{R : |R| \ge 2^{-n}} \frac{\langle C_N, h_R \rangle}{|R|} h_R(x_1, x_2) &= \sum_{k=0}^{n} \sum_{\widetilde{R}_1 : |\widetilde{R}_1| = 2^{-k}} \frac{C_{\widetilde{R}}(x_2) \langle h_{\widetilde{R}_1}, \mathbf{1}_{[p_1,1]} \rangle}{|\widetilde{R}_1|} h_{\widetilde{R}_1}(x_1) \\ &= \sum_{k=0}^{n} \sum_{\widetilde{R}_1 : |\widetilde{R}_1| = 2^{-k}} \frac{\alpha_{\widetilde{R}_1}(x_2)}{|\widetilde{R}_1|} h_{\widetilde{R}_1}(x_1), \end{aligned}$$

where the Haar coefficient $\alpha_{\widetilde{R}_1}(x_2)$ satisfies $|\alpha_{\widetilde{R}_1}(x_2)| \lesssim |\widetilde{R}_1|$. Next, we apply the one-dimensional Littlewood-Paley inequality in the variable x_1:

$$\Big\| \sum_{R : |R| \ge 2^{-n}} \frac{\langle C_N, h_R \rangle}{|R|} h_R \Big\|_{L^p(x_1)} \lesssim p^{\frac{1}{2}} \Big\| \Big(\sum_{\widetilde{R}_1 : |\widetilde{R}_1| \ge 2^{-n}} \frac{|\alpha_{\widetilde{R}_1}(x_2)|^2}{|\widetilde{R}_1|^2} \mathbf{1}_{\widetilde{R}_1} \Big)^{\frac{1}{2}} \Big\|_{L^p(x_1)} \le p^{\frac{1}{2}} n^{\frac{1}{2}}.$$

We now integrate this estimate in x_2 to obtain

$$\Big\| \sum_{R : |R| \ge 2^{-n}} \frac{\langle C_N, h_R \rangle}{|R|} h_R \Big\|_p \lesssim p^{\frac{1}{2}} n^{\frac{1}{2}},$$

and thus

$$\Big\| \sum_{R:|R|\geq 2^{-n}} \frac{\langle C_N, h_R\rangle}{|R|} h_R \Big\|_{\exp(L^2)} \lesssim n^{\frac{1}{2}},$$

in view of Proposition 2.2. Thus, we have estimated the $\exp(L^2)$ norms of all the terms in (6.2) by $n^{\frac{1}{2}}$. The estimates for $(0,1)$ and $(1,0)$ Haars in (6.2) can be easily incorporated, invoking similar one-dimensional arguments and Lemma 4.3. We skip these computations for the sake of brevity. We thus arrive to

$$\|D_N\|_{\exp(L^2)} \lesssim \sqrt{n} \approx \sqrt{\log N}.$$

Proposition 2.8 and inequality (3.7) finish the proof of Theorem 1.2 for all $\alpha \geq 2$.

6.3. **Upper bound: The Proof of Theorem 1.2 in the General Case.** We use a standard argument to generalize the previous proof to the case of arbitrary N. Fix $2^{n-1} < N < N' := 2^n$. Set $\frac{1}{2} < t = N2^{-n} + 2^{-n-1} < 1$. Consider the following function

$$\Delta_N(x_1, x_2) := D_{N'}(tx_1, x_2) - \tfrac{1}{2}x_1 \cdot x_2, \qquad (x_1, x_2) \in [0,1]^2.$$

Here, $D_{N'}$ is the Discrepancy Function of a shifted van der Corput set $\mathcal{V}_{n,\sigma}$. (The '$-\frac{1}{2}x_1 \cdot x_2$' above arises from the precise definition of the van der Corput set.)

The observation is that Δ_N is in fact the Discrepancy Function of the set of points $\{v_{n,\sigma}(\tau) \;:\; \tau = 0, 1, \ldots, N'\}$, where this notation is given in Definition 3.5. For the linear part of the Discrepancy Function, note that

$$N'(tx_1) \cdot x_2 - \tfrac{1}{2}x_1 \cdot x_2 = Nx_1 \cdot x_2.$$

And for the counting part, note that $\mathbf{1}_{[v_{n,\sigma}(\tau),1)}(tx_1, x_2)$ will be identically zero on $[0,1]^2$ iff $N < \tau \leq N'$. Thus, Δ_N is a Discrepancy Function.

So it suffices for us to estimate the $\exp(L^\alpha)$ norm of Δ_N. But this is straight forward.

$$\begin{aligned}\|\Delta_N\|_{\exp(L^\alpha)} &\leq 1 + \|D_{N'}(tx_1, x_2)\|_{\exp(L^\alpha)} \\ &\leq 1 + t^{-1}\|D_{N'}(x_1, x_2)\|_{\exp(L^\alpha)} \lesssim (\log N)^{1/\alpha}, \qquad 2 \leq \alpha < \infty.\end{aligned}$$

References

Beck, József. 1989. *A two-dimensional van Aardenne-Ehrenfest theorem in irregularities of distribution*, Compositio Math. **72**, no. 3, 269–339.MR1032337 (91f:11054) ↑2

Beck, József and William W. L. Chen. 1987. *Irregularities of distribution*, Cambridge Tracts in Mathematics, vol. 89, Cambridge University Press, Cambridge.MR903025 (88m:11061) ↑2, 8

Bilyk, Dmitriy and Michael T. Lacey. 2008. *On the small ball inequality in three dimensions*, Duke Math. J. **143**, no. 1, 81–115, available at `arXiv:math.CA/060981d`. MR 2414745 ↑2, 3, 6

Bilyk, Dmitriy, Michael T. Lacey, and Armen Vagharshakyan. 2008. *On the small ball inequality in all dimensions*, J. Funct. Anal. **254**, no. 9, 2470–2502, available at `arXiv:0705.4619`. MR 2409170 ↑2, 3

Chang, Sun-Yung A. and Robert Fefferman. 1980. *A continuous version of duality of H^1 with BMO on the bidisc*, Ann. of Math. (2) **112**, no. 1, 179–201. MR **584078 (82a:**32009) ↑3, 7

———. 1985. *Some recent developments in Fourier analysis and H^p-theory on product domains*, Bull. Amer. Math. Soc. (N.S.) **12**, no. 1, 1–43.MR 86g:42038 ↑7

Chang, S.-Y. A., J. M. Wilson, and T. H. Wolff. 1985. *Some weighted norm inequalities concerning the Schrödinger operators*, Comment. Math. Helv. **60**, no. 2, 217–246.MR800004 (87d:42027) ↑1, 4

Chen, W. W. L. 1980. *On irregularities of distribution*, Mathematika **27**, no. 2, 153–170 (1981). MR **610701 (82i:**10044) ↑2

______. 1983. *On irregularities of distribution. II*, Quart. J. Math. Oxford Ser. (2) **34**, no. 135, 257–279. MR **711520 (85c:**11065) ↑13

van der Corput, J. G. 1935. *Verteilungsfunktionen I*, Akad. Wetensch. Amdterdam, Proc. **38**, 813–821. ↑2, 13

Dunker, Thomas, Thomas Kühn, Mikhail Lifshits, and Werner Linde. 1998. *Metric entropy of the integration operator and small ball probabilities for the Brownian sheet*, C. R. Acad. Sci. Paris Sér. I Math. **326**, no. 3, 347–352 (English, with English and French summaries).MR2000b:60195 ↑2

Fefferman, R. and J. Pipher. 1997. *Multiparameter operators and sharp weighted inequalities*, Amer. J. Math. **119**, no. 2, 337–369.MR1439553 (98b:42027) ↑4, 6

Halász, G. 1981. *On Roth's method in the theory of irregularities of point distributions*, Recent progress in analytic number theory, Vol. 2 (Durham, 1979), Academic Press, London, pp. 79–94. MR **637361 (83e:**10072) ↑2, 22

Halton, J. H. and S. K. Zaremba. 1969. *The extreme and L^2 discrepancies of some plane sets*, Monatsh. Math. **73**, 316–328. ↑13

Kritzer, P. and F. Pillichshammer. 2006. *An exact formula for the L_2 discrepancy of the shifted Hammersley point set*, Uniform Distribution Theory **1**, no. 1, 1-13. ↑14

Lacey, Michael T. 2006. *On the Discrepancy Function in Arbitrary Dimension, Close to L^1*, to appear in Analysis Mathematica, available at arXiv:math.NT/060981d. ↑2

Lindenstrauss, Joram and Lior Tzafriri. 1977. *Classical Banach spaces. I*, Springer-Verlag, Berlin. Sequence spaces; Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. 92. MR 0500056 (58 #17766) ↑4

Matoušek, Jiří. 1999. *Geometric discrepancy*, Algorithms and Combinatorics, vol. 18, Springer-Verlag, Berlin. An illustrated guide. MR **1697825 (2001a:**11135) ↑11, 12, 22

Pipher, Jill. 1986. *Bounded double square functions*, Ann. Inst. Fourier (Grenoble) **36**, no. 2, 69–82 (English, with French summary).MR850744 (88h:42021) ↑6

Roth, K. F. 1954. *On irregularities of distribution*, Mathematika **1**, 73–79. MR 0066435 (16,575c) ↑2, 8, 13

Schmidt, Wolfgang M. 1972. *Irregularities of distribution. VII*, Acta Arith. **21**, 45–50. MR 0319933 (47 #8474) ↑1, 2

______. 1977a. *Lectures on irregularities of distribution*, Tata Institute of Fundamental Research Lectures on Mathematics and Physics, vol. 56, Tata Institute of Fundamental Research, Bombay. MR **554923 (81d:**10047) ↑8

______. 1977b. *Irregularities of distribution. X*, Number theory and algebra, Academic Press, New York, pp. 311–329. MR 0491574 (58 #10803) ↑2

Talagrand, Michel. 1994. *The small ball problem for the Brownian sheet*, Ann. Probab. **22**, no. 3, 1331–1354.MR 95k:60049 ↑2

Temlyakov, V. N. 1995. *An inequality for trigonometric polynomials and its application for estimating the entropy numbers*, J. Complexity **11**, no. 2, 293–307.MR 96c:41052 ↑2

Wang, Gang. 1991. *Sharp square-function inequalities for conditionally symmetric martingales*, Trans. Amer. Math. Soc. **328**, no. 1, 393–419. MR **1018577 (92c:**60067) ↑4

School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540, USA.
E-mail address: bilyk@math.ias.edu

School of Mathematics, Georgia Institute of Technology, Atlanta GA 30332, USA.
E-mail address: lacey@math.gatech.edu

Institutionen för Matematik, Kungliga Tekniska Högskolan, SE 100 44, Stockholm, SWEDEN.
E-mail address: ioannis.parissis@gmail.com

School of Mathematics, Georgia Institute of Technology, Atlanta GA 30332, USA.
E-mail address: armenv@math.gatech.edu