ESTIMATES FOR DISCREPANCY AND CALDERON-ZYGMUND OPERATORS

A Thesis
Presented to
The Academic Faculty

by

Armen Vagharshakyan

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Mathematics

Georgia Institute of Technology
August 2010

ESTIMATES FOR DISCREPANCY AND CALDERON-ZYGMUND OPERATORS

Approved by:

Vladimir Koltchinskii, Committee Chair
School of Mathematics
Georgia Institute of Technology

Michael Lacey, Advisor
School of Mathematics
Georgia Institute of Technology

Jeff Geronimo
School of Mathematics
Georgia Institute of Technology

Christopher E. Heil
School of Mathematics
Georgia Institute of Technology

Dmitriy Bilyk
School of Mathematics
University of South Carolina

Date Approved: 26 April 2010

To my family.

ACKNOWLEDGEMENTS

THANK YOU !

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The thesis consists of two independent parts.

The first part is devoted to certain results in the discrepancy theory and related problems. Take $A \subset [0,1]^d$ to be an N point set in the d dimensional unit cube and consider the discrepancy function associated to it:

$$D_A(\vec{x}) := \sharp\{A \cap [\vec{0}, \vec{x}]\} - N \left|[\vec{0}, \vec{x}]\right|, \qquad \vec{x} \in [0,1]^d.$$

(here the $\sharp$ sign counts the number of elements in the set, and $[\vec{0}, \vec{x}]$ stands for the rectangle with antipodal corners $\vec{0}$ and $\vec{x}$). The function D_A measures how much the distribution of the finite set A deviates from the corresponding uniform distribution.

In a joint work with D. Bilyk and M. Lacey we extended the previous result of D. Bilyk and M. Lacey (see [4]) to dimensions $d > 3$, by improving the lower bound for the discrepancy function. Namely, we showed that there exists a positive $\eta(d) > 0$ for which we have:

$$||D_A||_\infty \gtrsim (\ln N)^{(d-1)/2+\eta(d)} \quad \textit{for any set} \quad \sharp A = N.$$

This result makes a partial step towards resolving the Discrepancy Conjecture. Being a theorem in the theory of irregularities of distributions, it also relates to corresponding results in approximation theory (namely, the Kolmogorov entropy of spaces of functions with bounded mixed derivatives) and in probability theory (namely, Small Ball Inequality - small deviation inequality for the Brownian sheet).

In another joint work with D. Bilyk and M. Lacey we treat a particular case of the Small Ball Inequality - the Signed Small Ball Inequality. We show that in this case our estimates can be further improved.

Yet another joint work with D. Bilyk, M. Lacey and I. Parissis provides sharp bounds

for the exponential Orlicz norm and the BMO norm of the discrepancy function in two dimensions.

The second part of the thesis deals with Calderón-Zygmund operators in weighted spaces. We prove that any sufficiently smooth one-dimensional Calderón-Zygmund convolution operator can be recovered through averaging of certain Haar shift operators (i.e. dyadic operators which can be efficiently expressed in terms of the Haar basis). This generalizes the estimates, which had been previously known (see [23]) for Haar shift operators, to Calderón-Zygmund operators. As a result, the A_2 conjecture is settled for this particular type of Calederón-Zygmund operators.

CHAPTER I

INTRODUCTION

1.1 Main References.

The thesis is based on three published and one accepted papers:

Chapter 2 ("Discrepancy and Small Ball Inequality") is based on a paper by D. Bilyk, M. Lacey and A. Vagharshakyan (see [6]),

Chapter 3 ("Signed Small Ball Inequality") is based on a paper by D. Bilyk, M. Lacey and A. Vagharshakyan (see [3]),

Chapter 4 ("Orlicz and BMO Norms of Discrepancy in Two Dimensions") is based on a paper by D. Bilyk, M. Lacey, I. Parissis and A. Vagharshakyan (see [5]),

Chapter 5 ("Recovering Singular Integrals from Haar Shifts") is based on a paper by A. Vagharshakyan (see [48]).

1.2 Reading Order.

For chapters 3 and 4 a review of sections 2.1 and 2.2.1 is necessary. Other than that, the chapters essentially don't depend on each other.

CHAPTER II

DISCREPANCY AND SMALL BALL INEQUALITY

2.1 Small Ball Conjecture

In this chapter we will prove results in dimension four and higher in three separate areas, Number Theory, Approximation Theory, and Probability Theory: (a) the theory of Irregularities of Distribution, (b) the Kolmogorov Entropy of spaces of functions with bounded mixed derivative, and (c) Small Deviation Inequalities for the Brownian Sheet. As far as the authors are aware, these are the first results on these questions which provide more information than that given by an average case analysis. Underlying these three results is a central inequality, the *Small Ball Inequality* for the Haar functions, which we state here. The related areas are addressed in the next section.

In one dimension, the class of dyadic intervals is $D = [j2^k, (j+1)2^k] \quad : \quad j, k \in \mathbb{Z}$. Each dyadic interval has a left and right half, indicated below, which are also dyadic. Define the Haar functions

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

Note that this is an L^∞ normalization of these functions, which we will keep throughout chapters 2,3 and 4. In contrast, chapter 5 will use a different normalization.

In dimension d, a *dyadic rectangle* is a product of dyadic intervals, thus an element of D^d. We define a Haar function associated to R to be the product of the Haar functions associated with each side of R, namely

$$h_{R_1 \times \cdots \times R_d}(x_1, \ldots, x_d) = \prod_{j=1}^d h_{R_j}(x_j).$$

This is the usual 'tensor' definition.

We will concentrate on rectangles with fixed volume and consider a local problem. This is the 'hyperbolic' assumption, that pervades the subject. Our concern is the following Theorem and Conjecture concerning a *lower bound* on the L^∞ norm of sums of hyperbolic Haar functions:

Small Ball Conjecture 2.1.1. *For dimension $d \geq 3$ we have the inequality*

$$2^{-n} \sum_{|R|=2^{-n}} |\alpha(R)| \lesssim n^{\frac{1}{2}(d-2)} \Big\| \sum_{|R| \geq 2^{-n}} \alpha(R)h_R \Big\|_\infty.$$

Average case analysis — that is passing through L^2 — shows that we always have

$$2^{-n} \sum_{|R|=2^{-n}} |\alpha(R)| \lesssim n^{\frac{1}{2}(d-1)} \Big\| \sum_{|R| \geq 2^{-n}} \alpha(R)h_R \Big\|_\infty.$$

Namely, the constant on the right is bigger than in the conjecture by a factor of $\sqrt{n}$. We refer to this as the 'average case estimate,' and refer to improvements over this as a 'gain over the average case estimate.' Random choices of coefficients $\alpha(R)$ show that the Small Ball Conjecture is sharp.

In dimension $d = 2$, the Conjecture was resolved by [44].[1]

Talagrand's Theorem 2.1.2. *For dimensions $d \geq 2$, we have*

$$2^{-n} \sum_{|R|=2^{-n}} |\alpha(R)| \lesssim \Big\| \sum_{|R|=2^{-n}} \alpha(R)h_R \Big\|_\infty.$$

Here, the sum on the right is taken over all rectangles with area at least 2^{-n}.

The main result of this chapter is the next Theorem, which shows that there is a gain over the trivial bound in the Small Ball Conjecture in dimensions $d \geq 3$. In dimension $d = 3$, this result was proved in [4]. The three-dimensional result and its present extension build upon the method devised by [1]. As far as we are aware, this is the first 'gain over the average case bound' known in dimensions four and higher.

Theorem 2.1.3. *(Bilyk, Lacey, Vagharshakyan, [6]) In dimension $d \geq 3$, there exists a number $\eta(d) > 0$ such that for all choices of coefficients $\alpha(R)$, we have the inequality*

$$n^{\frac{d-1}{2}-\eta(d)} \Big\| \sum_{|R| \geq 2^{-n}} \alpha(R)h_R \Big\|_\infty \gtrsim 2^{-n} \sum_{|R|=2^{-n}} |\alpha(R)|.$$

---

[1]This result should be compared to [41], as well as [17],[46].

We take this Theorem as basic to our study, and use its proof to derive results on the three other questions mentioned at the beginning of this section.

The principal difficulty in three and higher dimensions is that two dyadic rectangles of the same volume can share a common side length. Beck [1] found a specific estimate in this case, an estimate that is extended in [4]. In this chapter the main technical device is the extension of this estimate, in the simplest instance, to arbitrary dimensions, see Lemma 2.5.2. This Lemma, and its extension to longer products Theorem 2.8.2, is the main technical innovation of this chapter. The value of η that we can get out of this line of reasoning appears to be of the order d^{-2}, imputing additional interest to the methods of proof used to improve this estimate.

2.2 Related Results

2.2.1 Discrepancy Conjecture

In d dimensions, take A_N to be N points in the unit cube, and consider the Discrepancy Function

$$D_N(x) := \sharp A_N \cap [\vec{0}, \vec{x}) - N|[\vec{0}, \vec{x})| \,. \tag{2.2.1}$$

Here, $[\vec{0}, \vec{x}) = \prod_{j=1}^d [0, x_j)$, that is a rectangle with antipodal corners being $\vec{0}$ and $\vec{x}$. 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. The canonical result of this type is the following estimate proved in [40].

K. Roth's Theorem 2.2.2. *We have the universal estimate*

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

with the implied constant only depending upon dimension.

For all $1 < p < \infty$, $\|D_N\|_p$ admits the same lower bound, a result in [42]. The endpoint estimates of $p = 1, \infty$ are however much harder, with definitive information known only in two dimensions. The method of proof of this Theorem, and the L^p variants can be summarized as follows: Fix $2N \leq 2^n < N$, and just project the Discrepancy Function onto the (hyperbolic) Haar functions $\{h_R : |R| = 2^{-n}\}$. By the Bessel inequality, this provides a

lower bound on the L^2 norm of D_N. This same method of proof, with the Littlewood-Paley inequalities replacing the Bessel inequality, can be used to prove the L^p lower bound, for $1 < p < \infty$. See [2].

At L^∞, guided by the sharpness of the Small Ball Conjecture, we pose the Conjecture below, which represents a $\sqrt{\log N}$ gain over the lower bound proved by Roth.

The L^∞ Norm of Discrepancy Function Conjecture 2.2.3. *In dimension $d \geq 3$, we have the lower estimate valid for all point sets A_N.*

$$\|D_N\|_\infty \gtrsim (\log N)^{d/2}.$$

In dimension $d = 2$, this is the Theorem of [41]. In dimension $d = 3$, [1],[4] give partial information about this conjecture. In this chapter, we can prove the following result, which appears to be new in dimensions $d \geq 4$.

Theorem 2.2.4. *(Bilyk, Lacey, Vagharshakyan, [6]) In dimension $d \geq 3$ there is a positive $\eta = \eta(d) > 0$ for which we have the uniform estimate*

$$\|D_N\|_\infty \gtrsim (\log N)^{(d-1)/2+\eta}.$$

The proof of this result follows easily from the method of proof of Theorem 2.1.3, and will be presented below.

2.2.2 Metric Entropy of Mixed Derivative Sobolev Spaces

While the special structure of the Haar functions can be exploited to prove the Small Ball Conjecture, one would *not* anticipate that this special structure is in fact essential to the Conjecture. Thus, we formulate a smooth variant of the Small Ball Conjecture.

Fix a continuous non-constant function φ, supported on $[-1/2, 1/2]$, and of mean zero. For a dyadic interval I, let

$$\varphi_I(x) = \varphi\left(\tfrac{x-c(I)}{|I|}\right),$$

be a translation and rescaling of φ so that it is supported on I. Then, for a dyadic rectangle $R = R_1 \times \cdots \times R_d$, set

$$\varphi_R(x_1, \ldots, x_d) = \prod_{j=1}^{d} \varphi_{R_j}(x_j).$$

Smooth Small Ball Conjecture 2.2.5. *For dimension $d \geq 3$ we have the inequality*

$$2^{-n} \sum_{|R|=2^{-n}} |\alpha(R)| \lesssim n^{\frac{1}{2}(d-2)} \left\| \sum_{|R| \geq 2^{-n}} \alpha(R)\varphi_R \right\|_\infty.$$

The implied constant depends upon dimension d and φ only.

In this direction, we will prove a result in the same spirit as our Main Theorem.

Theorem 2.2.6. *(Bilyk, Lacey, Vagharshakyan, [6]) Suppose φ is continuous, supported on $[-1/2, 1/2]$, of mean zero, and such that $\langle \varphi, h_{[-1/2,1/2]} \rangle \neq 0$. For dimension $d \geq 3$, there is a positive $\eta = \eta(d)$ so that we have the inequality below*

$$2^{-n} \sum_{|R|=2^{-n}} |\alpha(R)| \lesssim n^{\frac{1}{2}(d-1)-\eta} \left\| \sum_{|R| \geq 2^{-n}} \alpha(R)\varphi_R \right\|_\infty. \tag{2.2.7}$$

The implied constant depends upon φ.

With this Theorem, we can establish new results on the metric entropy of certain Sobolev spaces of functions with mixed derivative in certain L^p spaces. In d dimensions, consider the map

$$\mathrm{Int}_d \, f(x_1, \cdots, x_d) = \int_0^{x_1} \cdots \int_0^{x_d} f(y_1, \cdots, y_d) \, dy_1 \cdots dy_d.$$

We consider this as a map from $L^p([0,1]^d)$ into $C([0,1]^d)$. Clearly, the image of Int_d consists of functions with L^p integrable mixed partial derivatives. Let us set

$$\mathrm{Ball}(MW^p([0,1]^d)) = \mathrm{Int}_d(\{f \in L^p([0,1]^d) : \|f\|_p \lesssim 1\}).$$

That is, this is the image of the unit ball of L^p. This is the unit ball of the space of functions with mixed derivative in L^p.

These sets are compact in in $C([0,1]^d)$, and it is of relevance to quantify the compactness, through the device of *covering numbers*. For $0 < \epsilon < 1$, set $N(\epsilon, p, d)$ to be the least number N of points $x_1, \cdots, x_N \in C([0,1]^d)$ so that

$$\mathrm{Ball}(MW^p([0,1]^d)) \subset \bigcup_{n=1}^N (x_n + \epsilon B_\infty).$$

Here, B_∞ is the unit ball of $C([0,1]^d)$. The task at hand is to uncover the correct order of growth of these numbers as $\epsilon \downarrow 0$. The case of $d = 2$ below follows from Talagrand [44], and the upper bound is known in full generality [14],[46].

6

Conjecture 2.2.8. *For $d \geq 2$ one has the estimate*

$$\log N(\epsilon, p, d) \simeq \epsilon^{-1}(\log 1/\epsilon)^{d-1/2}, \qquad \epsilon \downarrow 0.$$

It is well known [45] that results such as Theorem 2.2.6 can be used to give new lower bounds on these covering numbers.

Theorem 2.2.9. *(Bilyk, Lacey, Vagharshakyan, [6]) For $1 \leq p < \infty$, and $d \geq 3$, there is a $\eta > 0$ for which we have*

$$\log N(\epsilon, p, d) \gtrsim \epsilon^{-1}(\log 1/\epsilon)^{d-1+\eta}.$$

We have concentrated on the case of one mixed derivative, but various results on fractional derivatives are also interesting. See for instance [22], and [13].

2.2.3 Small Ball Inequality for the Brownian Sheet

Perhaps, it is worthwhile to explain the nomenclature 'Small Ball' at this point. The name comes from the probability theory. Assume that $X_t : T \to \mathbb{R}$ is a canonical Gaussian process indexed by a set T. The *Small Ball Problem* is concerned with estimates of $\mathbb{P}(\sup_{t \in T} |X_t| < \varepsilon)$ as ε goes to zero, i.e the probability that the random process takes values in an L^∞ ball of small radius. The reader is advised to consult a paper by Li and Shao [28] for a survey of this type of questions. A particular question of interest to us deals with the Brownian Sheet, that is, a centered Gaussian process indexed by the points in the unit cube $[0,1]^d$ and characterized by the covariance relation $\mathbb{E}X_s \cdot X_t = \prod_{j=1}^d \min(s_j, t_j)$.

Kuelbs and Li [21] have discovered a tight connection between the Small Ball probabilities and the properties of the reproducing kernel Hilbert space corresponding to the process, which in the case of the Brownian Sheet is $WM^2([0,1]^d)$, the space described in the previous subsection. Their result, applied to the setting of the Brownian sheet in [14], states that

Theorem 2.2.10. *In dimension $d \geq 2$, as $\varepsilon \downarrow 0$ we have*

$$-\log \mathbb{P}(\|B\|_{C([0,1]^d)} < \varepsilon) \simeq \varepsilon^{-2}(\log 1/\varepsilon)^\beta \quad \text{iff} \quad \log N(\varepsilon) \simeq \varepsilon^{-1}(\log 1/\varepsilon)^{\beta/2}.$$

Thus, in agreement with Conjecture 2.2.8, the conjectured form of the aforementioned probability in this case is the following:

The Small Ball Conjecture for the Brownian Sheet 2.2.11. *In dimensions $d \geq 2$, for the Brownian Sheet B we have*

$$- \log \mathbb{P}(\|B\|_{C([0,1]^d)} < \varepsilon) \simeq \varepsilon^{-2} (\log 1/\varepsilon)^{2d-1}, \quad \varepsilon \downarrow 0.$$

In dimension $d = 2$, this conjecture has been resolved by Talagrand in the already cited paper [44], in which he actually proved Conjecture 2.2.5 for a specific function φ and used it to deduce the lower bound in the inequality above.[2] In higher dimensions, the upper bounds are established, see [14], and the previously known lower bounds miss the conjecture by a single power of the logarithm.

Theorem 2.2.9 can be translated into the following result on the Small Ball Probability for the Brownian Sheet:

Theorem 2.2.12. *(Bilyk, Lacey, Vagharshakyan, [6]) In dimensions $d \geq 3$, there exists $\eta > 0$ such that for the Brownian Sheet B we have*

$$- \log \mathbb{P}(\|B\|_{C([0,1]^d)} < \varepsilon) \gtrsim \varepsilon^{-2} (\log 1/\varepsilon)^{2d-2+\eta}, \quad \varepsilon \downarrow 0.$$

2.3 Notations and Littlewood-Paley Inequality

Let $\vec{r} \in \mathbb{N}^d$ be a partition of n, thus $\vec{r} = (r_1, \ldots, r_d)$, where the r_j are nonnegative integers and $|\vec{r}| = \sum_t r_t = n$, which we refer to as the *length of the vector $\vec{r}$*. Denote all such vectors as $\mathbb{H}_n$. ('$\mathbb{H}$' for 'hyperbolic.') For vector $\vec{r}$ let $R_{\vec{r}}$ be all dyadic rectangles R such that for each coordinate k, $|R_k| = 2^{-r_k}$.

Definition 2.3.1. We call a function f an r *function with parameter $\vec{r}$* if

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

A fact used without further comment is that $f_{\vec{r}}^2 \equiv 1$.

---

[2]The work of Talagrand bears strong similarities to the prior work of [41] and [17]. The argument of Talagrand was subsequently clarified by [46], and [13].

As it has been already pointed out, the principal difficulty in three and higher dimensions is that the product of Haar functions is not necessarily a Haar function. On this point, we have the following

Proposition 2.3.2. *Suppose that $R_1, \ldots, R_k$ are rectangles such that there is no choice of $1 \leq j < j' \leq k$ and no choice of coordinate $1 \leq t \leq d$ for which we have $R_{j,t} = R_{j',t}$. Then, for a choice of sign $\varepsilon \in \{\pm 1\}$ we have*

$$\prod_{j=1}^{k} h_R = \varepsilon h_S, \qquad S = \bigcap_{j=1}^{k} R_k.$$

Proof. Expand the product as

$$\prod_{m=1}^{\ell} h_{R_m}(x_1, \ldots, x_d) = \prod_{m=1}^{\ell} \prod_{t=1}^{d} h_{R_{m,t}}(x_t).$$

Here $\varepsilon_m \in \{\pm 1\}$. Our assumption is that for each t, there is exactly one choice of $1 \leq m_0 \leq \ell$ such that $R_{m_0,t} = S_t$. And moreover, since the minimum value of $|R_{m,t}|$ is obtained exactly once, for $m \neq m_0$, we have that $h_{R_{m,t}}$ is constant on S_t. Thus, in the t coordinate, the product is

$$h_{S_t}(x_t) \prod_{1 \leq m \neq m_0 \leq \ell} h_{R_{m,t}}(S_t).$$

This proves our Lemma. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \square$

Remark 2.3.3. It is also a useful observation, that the product of Haar functions will have mean zero if the minimum value of $|R_{m,t}|$ is unique for at least one coordinate t.

Definition 2.3.4. For vectors $\vec{r}_j \in \mathbb{N}^d$, say that $\vec{r}_1, \ldots, \vec{r}_J$ are *strongly distinct* iff for coordinates $1 \leq t \leq d$ the integers $\{r_{j,t} : 1 \leq j \leq J\}$ are distinct. The product of strongly distinct r functions is also an r function, which follows from 'the product rule' (2.3.2).

The r functions we are interested in are:

$$f_{\vec{r}} = \sum_{R \in R_{\vec{r}}} \text{sgn}(\alpha(R)) \, h_R. \tag{2.3.5}$$

We recall some Littlewood-Paley inequalities, which are standard, and so we omit proofs.

Littlewood-Paley Inequalities 2.3.6. *In one dimension, we have the inequalities*

$$\Big\| \sum_{I \subset \mathbb{R}} a_I h_I(\cdot) \Big\|_p \lesssim \sqrt{p} \Big\| \Big[\sum_{I \subset \mathbb{R}} a_I^2 \mathbf{1}_I(\cdot) \Big]^{1/2} \Big\|_p, \qquad 2 < p < \infty. \qquad (2.3.7)$$

Moreover, these inequalities continue to hold in the case where the coefficients a_I take values in a Hilbert space H.

The growth of the constant is essential for us, in particular the factor $\sqrt{p}$ is, up to a constant, the best possible in this inequality. See [15],[50]. That these inequalities hold for Hilbert space valued sums is imperative for applications to higher dimensional sums of Haar functions. The relevant inequality is as follows.

Theorem 2.3.8. *We have the inequalities below for hyperbolic sums of r functions in dimension $d \geq 3$.*

$$\Big\| \sum_{|\vec{r}|=n} f_{\vec{r}} \Big\|_p \lesssim (pn)^{(d-1)/2}, \qquad 2 < p < \infty.$$

We recall a vector valued Harmonic Analysis inequality.

Proposition 2.3.9. *Let F_j be a sigma field generated by dyadic rectangles in dimension 2. We then have*

$$\Big\| \Big[\sum_j \mathbb{E}(\varphi_j \,|\, F_j)^2 \Big]^{1/2} \Big\|_p \lesssim p \Big\| \Big[\sum_j \varphi_j^2 \Big]^{1/2} \Big\|_p, \qquad 2 < p < \infty. \qquad (2.3.10)$$

Proof. This is one of many examples of a vector valued inequality in the Harmonic Analysis literature. This particular inequality admits a simple proof by duality, recalled here for convenience.

Since $p > 2$, we can appeal to a duality argument. We can choose $g \in L^{(p/2)'}$ of norm

one so that

$$\left\|\sum_j \mathbb{E}(\varphi_j \,|\, F_j)^2\right\|_{p/2} = \sum_j \langle \mathbb{E}(\varphi_j \,|\, F_j)^2, g\rangle$$

$$\leq \sum_j \langle \mathbb{E}(\varphi_j^2 \,|\, F_j), g\rangle$$

$$= \sum_j \langle \varphi_j^2, \mathbb{E}(g \,|\, F_j)\rangle$$

$$\leq \sum_j \langle \varphi_j^2, \mathrm{M}\, g\rangle$$

$$\leq \left\|\sum_j \varphi_j^2\right\|_{p/2} \|\mathrm{M}\, g\|_{(p/2)'}$$

$$\lesssim ((p/2)' - 1)^{-2} \left\|\sum_j \varphi_j^2\right\|_{p/2}.$$

Here we have used Jensen's inequality and the self-duality of the conditional expectation operators. The operator $\mathrm{M}\, g$ is the (strong) maximal function on the plane, namely

$$\mathrm{M}\, g(x) = \sup_R \frac{\mathbf{1}_R}{|R|} \int_R |g(y)|\, dy,$$

where the supremum is over all dyadic rectangles R. This maps L^q into L^q for all $1 < q < \infty$, an inequality appealed to in the last line of the display above. Moreover, it is well known that the norm of the operator behaves as

$$\|\mathrm{M}\|_{q \to q} \lesssim (q - 1)^{-2}, \qquad 1 < q < 2.$$

$\square$

2.4   Proof of a Small Ball Inequality

The proof of the Theorem is by duality, namely we construct a function Ψ of L^1 norm about one, which is used to provide a lower bound on the L^∞ norm of the sum of Haar functions. The details of this argument are similar to those of [4].

The function Ψ will take the form of a Riesz product, but in order to construct it, we need some definitions. Fix $0 < \varepsilon < 1$ to be a small number, ultimately of order $1/d^2$. Define

relevant parameters by

$$q = \lfloor an^\varepsilon \rfloor, \qquad b = \tfrac{1}{4}, \tag{2.4.1}$$

$$\widetilde{\rho} = aq^b n^{-(d-1)/2}, \qquad \rho = \sqrt{q}\, n^{-(d-1)/2}. \tag{2.4.2}$$

Here a is a small positive constant, we use the notation $b = \tfrac{1}{4}$ throughout, so as not to obscure those aspects of the argument that dictate this choice. $\widetilde{\rho}$ is a 'false' L^2 normalization for the sums we consider, while the larger term ρ is the 'true' L^2 normalization. Our 'gain over the average case estimate' in the Small Ball Conjecture is $q^b \simeq n^{\varepsilon/4}$.

Divide the integers $\{1, 2, \ldots, n\}$ into q disjoint intervals of equal length $I_1, \ldots, I_q$, ordered from smallest to largest. Let $\mathbb{A}_t = \{\vec{r} \in \mathbb{H}_n : r_1 \in I_t\}$. Let

$$F_t = \sum_{\vec{r} \in \mathbb{A}_t} f_{\vec{r}}, \qquad H_n = \sum_{|R|=2^{-n}} \alpha(R) h_R. \tag{2.4.3}$$

Here, the $f_{\vec{r}}$ are as in (2.3.5). The Riesz product is a 'short product':

$$\Psi = \prod_{t=1}^{q} (1 + \widetilde{\rho} F_t),$$

One can view the $\widetilde{\rho} F_t$ as a 'poor man's $\mathrm{sgn}(F_t)$', in that the Riesz product above tends to weight the region where the functions F_t align. Note the subtle way in which the false L^2 normalization enters into the product. It means that the product is, with high probability, positive. And of course, for a positive function F, we have $\mathbb{E}F = \|F\|_1$, with expectations being typically easier to estimate. This heuristic is made precise below.

Proposition 2.3.2 suggests that we should decompose the product Ψ into

$$\Psi = 1 + \Psi^{\mathrm{sd}} + \Psi^{\neg},$$

where the two pieces are the 'strongly distinct' and 'not strongly distinct' pieces. To be specific, for integers $1 \le u \le q$, let

$$\Psi_k^{\mathrm{sd}} = \widetilde{\rho}^k \sum_{1 \le v_1 < \cdots < v_k \le q} \sum_{\vec{r}_t \in \mathbb{A}_{v_t}}^{\mathrm{sd}} \prod_{t=1}^{u} f_{\vec{r}_t},$$

where $\sum^{\mathrm{sd}}$ is taken to be over all $\vec{r}_t \in \mathbb{A}_{v_t}$ $1 \le m \le k$ such that:

the vectors $\{\vec{r}_t : 1 \le m \le k\}$ are strongly distinct. \tag{2.4.4}

12

Then define

$$\Psi^{\mathrm{sd}} = \sum_{k=1}^{q} \Psi_k^{\mathrm{sd}}.$$

With this definition, it is clear that we have

$$\langle H_n, \Psi^{\mathrm{sd}} \rangle = \langle H_n, \Psi_1^{\mathrm{sd}} \rangle \gtrsim q^b \cdot n^{-\frac{d-1}{2}} \cdot 2^{-n} \sum_{|R|=2^{-n}} |\alpha_R|,$$

so that q^b is our 'gain over the trivial estimate', once we prove that $\|\Psi^{\mathrm{sd}}\|_1 \lesssim 1$ (estimate (2.4.11) below). Proving this inequality is the main goal of the technical estimates of the following Lemma:

Lemma 2.4.5. *We have these estimates:*

$$\mathbb{P}(\Psi < 0) \lesssim \exp(-A q^{1-2b})\,; \tag{2.4.6}$$

$$\|\Psi\|_2 \lesssim \exp(a' q^{2b})\,; \tag{2.4.7}$$

$$\mathbb{E}\Psi = 1\,; \tag{2.4.8}$$

$$\|\Psi\|_1 \lesssim 1\,; \tag{2.4.9}$$

$$\|\Psi^-\|_1 \lesssim 1\,; \tag{2.4.10}$$

$$\|\Psi^{\mathrm{sd}}\|_1 \lesssim 1\,. \tag{2.4.11}$$

Here, $0 < a' < 1$, in (2.4.7), is a small constant, decreasing to zero as a in (2.4.1) goes to zero; and $A > 1$, in (2.4.6) is a large constant, tending to infinity as a in (2.4.1) goes to zero.

Proof. We give the proof of the Lemma, assuming our main inequalities proved in the subsequent sections.

Proof of (2.4.6). Using the distributional estimate (2.6.3) of Theorem 2.6.1 proved in Section 5, and the definition of Ψ we estimate

13

$$\mathbb{P}(\Psi < 0) \leq \sum_{t=1}^{q} \mathbb{P}(\widetilde{\rho} \, F_t < -1)$$

$$= \sum_{t=1}^{q} \mathbb{P}(\rho F_t < -a^{-1}q^{1/2-b})$$

$$\lesssim \exp(-ca^{-2}q^{1-2b}).$$

Proof of (2.4.7). The proof of this is detailed enough and uses the results of subsequent sections, so we postpone it to Section 6, Lemma 2.6.4 below.

Proof of (2.4.8). Expand the product in the definition of Ψ. The leading term is one. Every other term is a product

$$\prod_{k \in V} \widetilde{\rho} \, F_k \,,$$

where V is a non-empty subset of $\{1, \ldots, q\}$. This product is in turn a linear combination of products of r functions. Among each such product, the maximum in the first coordinate is unique. This fact tells us that the expectation of these products of r functions is zero. So the expectation of the product above is zero. The proof is complete.

Proof of (2.4.9). We use the first two estimates of our Lemma. Observe that

$$\|\Psi\|_1 = \mathbb{E}\Psi - 2\mathbb{E}\Psi \mathbf{1}_{\Psi<0}$$

$$\leq 1 + 2\mathbb{P}(\Psi < 0)^{1/2}\|\Psi\|_2$$

$$\lesssim 1 + \exp(-Aq^{1-2b}/2 + a'q^{2b}).$$

We have taken $b = 1/4$ so that $1 - 2b = 2b$. For sufficiently small a in (2.4.1), we will have $A \gtrsim a'$. We see that (2.4.9) holds.

Indeed, Lemma 2.6.4 proves a uniform estimate, namely

$$\sup_{V \subset \{1,\ldots,q\}} \mathbb{E} \prod_{v \in V} (1 + \widetilde{\rho}F_t)^2 \lesssim \exp(a'q^{2b}).$$

Hence, the argument above proves

$$\sup_{V \subset \{1,\ldots,q\}} \left\| \prod_{t \in V} (1 + \widetilde{\rho}F_t) \right\|_1 \lesssim 1. \tag{2.4.12}$$

14

Proof of (2.4.10). The primary facts are (2.4.12) and Theorem 2.8.2; we use the notation devised for that Theorem.

We use the triangle inequality, estimate (2.4.7) of Lemma 2.4.5, Hölder's inequality, with indices q^{2b} and $(q^{2b})' = q^{2b}/(q^{2b} - 1)$, the inclusion-exclusion identity (2.8.1) and estimate (2.8.3) of Theorem 2.8.2 in the calculation below. Notice that we have

$$\sup_{V \subset \{1,\ldots,q\}} \left\| \prod_{t \in V} (1 + \widetilde{\rho} F_t) \right\|_{(q^{2b})'} \leq \sup_{V \subset \{1,\ldots,q\}} \left\| \prod_{t \in V} (1 + \widetilde{\rho} F_t) \right\|_1^{(q^{2b}-1)/q^{2b}} \times \left\| \prod_{t \in V} (1 + \widetilde{\rho} F_t) \right\|_2^{2q^{-2b}}$$

$$\lesssim 1.$$

We now estimate

$$\|\Psi^{\neg}\|_1 \leq \sum_{G \text{ admissible}} \left\| \widetilde{\rho}^{|V(G)|} \operatorname{SumProd}(X(G)) \cdot \prod_{t \in \{1,\ldots,q\}-V(G)} (1 + \widetilde{\rho} F_t) \right\|_1 \quad (2.4.13)$$

$$\leq \sum_{G \text{ admissible}} \left\| \widetilde{\rho}^{|V(G)|} \operatorname{SumProd}(X(G)) \right\|_{q^{2b}} \cdot \left\| \prod_{t \in \{1,\ldots,q\}-V(G)} (1 + \widetilde{\rho} F_t) \right\|_{(q^{2b})'}$$

$$\lesssim \sum_{G \text{ admissible}} \left\| \widetilde{\rho}^{|V(G)|} \operatorname{SumProd}(X(G)) \right\|_{q^{2b}}$$

$$= \sum_{v=2}^{q} \sum_{G:\,|V(G)|=v} \left\| \widetilde{\rho}^{|V(G)|} \operatorname{SumProd}(X(G)) \right\|_{q^{2b}}$$

$$\lesssim \sum_{v=2}^{q} \binom{q}{v} v^{2dv} [q^{C'} n^{-\eta}]^v$$

$$\lesssim q^{C''} n^{-\eta} \lesssim n^{-\varepsilon'} \lesssim 1 .$$

Proof of (2.4.11). This follows from (2.4.10) and (2.4.9), and the identity $\Psi = 1 + \Psi^{\text{sd}} + \Psi^{\neg}$ together with the triangle inequality.

$\square$

2.5 The Analysis of the Coincidence

Following the language of J. Beck [1], a *coincidence* occurs if we have two vectors $\vec{r} \neq \vec{s}$ with e.g. $r_2 = s_2$. He observed that sums over products of r functions in which there are coincidences obey favorable L^2 estimates. We refer to (extensions of) this observation as the *Beck Gain.* We introduce relevant notation for this situation. For $1 \leq k \leq d$ and

$1 \leq t_1, t_2 \leq q$, set

$$\Phi_{t_1,t_2,k} = \sum_{\substack{\vec{r} \in \mathbb{A}_{t_1}; \vec{s} \in \mathbb{A}_{t_2} \\ \vec{r} \neq \vec{s} \\ r_k = s_k}} f_{\vec{r}} \cdot f_{\vec{s}}. \tag{2.5.1}$$

Notice that due to our construction of the Riesz Product, there are no coincidences in the first coordinate in the decomposition of Ψ, although the case $k = 1$ is important for the proof of the L^2 estimate (2.4.7) . In the sum above, there are $2d - 3$ free parameters among the vectors $\vec{r}$ and $\vec{s}$. That is, the pair of vectors $(\vec{r}, \vec{s})$ are completely specified by their values in $2d - 3$ coordinates. The following lemma suggests that these parameters behave as if they were orthogonal.

The Simplest Instance of the Beck Gain 2.5.2. *We have the estimates below, valid for an absolute implied constant that is only a function of dimension $d \geq 3$.*

$$\sup \|\Phi_{t_1,t_2,k}\|_p \lesssim p^{d-1/2} \cdot n^{d-3/2}, \qquad 2 \leq p < \infty, \tag{2.5.3}$$

where the supremum is taken over all $1 \leq k \leq d$ and $1 \leq t_1, t_2 \leq q$.

This estimate is smaller by $1/2$ power of n than what one might naively expect, and so we say that we have an average gain of $1/4$ power of n in the products above. (Here, the average is in reference to the two functions we form the product of.) This Lemma, in dimension $d = 3$ appears in [4]. We will give an inductive proof of this estimate, that requires that we revisit the three dimensional case. In the next section, we also derive other estimates from the one above.

The estimate above may admit an improvement, in that the power of p is perhaps too large by a single power, due to our use of Proposition 2.3.9. (There should also be a dependence upon q, but on this point, and in many others, the arguments are suboptimal, and so we do not pursue this point here.)

Conjecture 2.5.4. *We have the estimates below, valid for an absolute implied constant that is only a function of dimension $d \geq 3$.*

$$\sup \|\Phi_{t_1,t_2,k}\|_p \lesssim (pn)^{d-3/2}, \qquad 2 \leq p < \infty.$$

Proof of Lemma 2.5.2

The proof is inductive on dimension. We shall suppress dependence on t_1, t_2. In fact, we shall prove the Theorem for the quantity

$$\Phi_1 = \sum_{\substack{\vec{r} \neq \vec{s} \in \mathbb{H}_n \\ r_1 = s_1}} f_{\vec{r}} \cdot f_{\vec{s}},$$

and the claimed statement will follow with only minor adjustments. To set up the induction, we need some definitions.

Definition 2.5.5. Given a set of r functions $\{f_{\vec{r}}\}$ and subset $\mathbb{C} \subset \mathbb{H}_{n_1} \times \cdots \times \mathbb{H}_{n_t}$, set

$$\mathrm{SumProd}(\mathbb{C}) = \sum_{(\vec{r}_1, \ldots, \vec{r}_t) \in \mathbb{C}} \prod_{s=1}^{t} f_{\vec{r}_s} .$$

Below, we will be interested in pairs and four-tuples of r functions. It is an important element of the argument, allowing us to run the induction, that we consider products of r functions where the vectors are in hyperbolic collections $\mathbb{H}_n$, for different values of n.

The main quantity we induct on is then

$$B(d, n, p) = \sup_{\mathbb{B}} \|\mathrm{SumProd}(\mathbb{B})\|_p , \qquad d, n, p \geq 3 . \tag{2.5.6}$$

Here, the supremum is formed over all $\mathbb{B} \subset \mathbb{H}_{n_1} \times \mathbb{H}_{n_2}$ and all r functions subject to these conditions:

- There is a coincidence in the first coordinate: For all $(\vec{r}, \vec{s}) \in \mathbb{B}$, we have $\vec{r} \neq \vec{s}$ and $r_1 = s_1$.

- $n_1, n_2 \leq n$. That is the lengths of the vectors $\vec{r}$ and $\vec{s}$ are permitted to be different.

- No other restriction is placed upon the pairs of vectors in $\mathbb{B}$.

Our main estimate on these quantities is as follows.

Lemma 2.5.7. *We have the inequality below valid for all dimensions $d \geq 3$.*

$$B(d, n, p) \lesssim p^{d-1/2} n^{d-3/2} , \qquad p, n \geq 3 .$$

The inductive argument for Lemma 2.5.7 has the underlying strategy of reducing dimension by application of the Littlewood-Paley inequalities. But, this causes the collections of vectors to lose some of their symmetry. Regaining the symmetry causes us to introduce additional types of collections of vectors. Two of these collections are as follows.

$$C(d,n,p) = \sup_{\mathbb{C}} \|\mathrm{SumProd}(\mathbb{C})\|_p, \qquad d,n,p \geq 3.$$

Here, the supremum is formed over all $\mathbb{C} \subset \mathbb{H}_{n_1} \times \mathbb{H}_{n_2}$ and all r functions subject to these conditions

- There is a coincidence in the first coordinate: For all $(\vec{r}, \vec{s}) \in \mathbb{C}$, we have $\vec{r} \neq \vec{s}$ and $r_1 = s_1$.

- For all $(\vec{r}, \vec{s}) \in \mathbb{C}$, we have $r_2 > s_2$ and $r_3 < s_3$.

- $n_1, n_2 \leq n$.

- There is no other restriction on the pairs of vectors in $\mathbb{C}$.

The only difference between the present collections and the collections in $B(d,n,p)$ is that in the present collections we assume locations of maximums in the second and third coordinates, thereby permitting application of the Littlewood-Paley inequalities in those two coordinates.

The second collection is less sophisticated. We simply assume that the maximum always occurs in say, the first coordinate. Define

$$D(d,n,p) = \sup_{\mathbb{D}} \|\mathrm{SumProd}(\mathbb{D})\|_p, \qquad d,n,p \geq 3. \qquad (2.5.8)$$

Here, the supremum is formed over all $\mathbb{D} \subset \mathbb{H}_{n_1} \times \mathbb{H}_{n_2}$ and all r functions subject to these conditions

- There is a coincidence in the first coordinate: For all $(\vec{r}, \vec{s}) \in \mathbb{D}$, we have $\vec{r} \neq \vec{s}$ and $r_1 = s_1$.

- For all $(\vec{r}, \vec{s}) \in \mathbb{D}$, and all $2 \leq j \leq d$, we have $r_j \geq s_j$.

18

- $n_2 < n_1 \leq n$.

That is, we require that in each coordinate where there is a maximum, the maximum occurs in the vector $\vec{r}$.

Lemma 2.5.9. *We have the inequality below valid for all dimensions $d \geq 3$.*

$$C(d,n,p)\,,\ D(d,n,p) \lesssim p^{d-1/2} \cdot n^{d-3/2}\,, \qquad p,n \geq 3\,.$$

We turn to the proofs of the Lemma 2.5.7 and Lemma 2.5.9, and begin by explaining the logic of our induction. Let $B(d)$ stand for the inequalities in Lemma 2.5.7 in dimension d, and likewise for $C(d)$ and $D(d)$. We prove:

- The inequalities $D(d)$ for all dimensions d.

- The inequalities $B(3)$ and $C(3)$. At the same time, assuming $B(d-1)$, $d \geq 4$, we prove $C(d)$.

- Assuming $C(d)$ and $D(d)$, we prove $B(d)$.

These clearly combine to prove the two Lemmas, and so complete the proof of Lemma 2.5.2.

The Inequalities $D(d)$.

The definition of $D(d)$ permits the possibility of equality for a large number of coordinates of the two vectors. Let us exclude that case in this definition. Define

$$D_{\neq}(d,n,p) = \sup_{\mathbb{D}} \|\mathrm{SumProd}(\mathbb{D})\|_p\,, \qquad d,n,p \geq 3\,, \tag{2.5.10}$$

where $\mathbb{D}$ is as in (2.5.8), but with the additional condition that for $2 \leq j \leq d$ we have $r_j > s_j$. Then, we are free to apply the Littlewood-Paley inequality in each of the coordinates from 2 to d.

Fix a collection of vectors $\mathbb{D}$, and a collection of r functions which achieves the supremum in (2.5.10). For this collection, and a choice of vector $\vec{\rho} \in \mathbb{N}^{d-1}$, let

$$\mathbb{D}_{\vec{\rho}} = \{(\vec{r},\vec{s}) \in \mathbb{D}\ :\ r_{j+1} = \rho_j\,,\ 1 \leq j \leq d-1\}\,.$$

19

Of course there are at most $\lesssim n^{d-1}$ values of $\vec{\rho}$ for which the collection above is non-empty. Then,

$$D_{\neq}(d, n, p) \lesssim p^{(d-1)/2} \left\| \left[\sum_{\vec{\rho}} \mathrm{SumProd}(\mathbb{D}_{\vec{\rho}})^2 \right]^{1/2} \right\|_p$$

$$\lesssim p^{(d-1)/2} n^{(d-1)/2} \sup_{\vec{\rho}} \left\| \sum_{\vec{\rho}} \mathrm{SumProd}(\mathbb{D}_{\vec{\rho}}) \right\|_p .$$

But, the coordinate r_1 is completely specified in $\mathbb{D}_{\vec{\rho}}$, and therefore does not contribute to the last norm. And so the first coordinate of $\vec{s}$ is specified. Therefore, there are at most $d - 2$ free choices of parameters in the vector s. By application of the Littlewood–Paley inequalities, we have

$$D_{\neq}(d, n, p) \lesssim (pn)^{d-3/2} .$$

This is better than the claimed inequality.

If there are a set $J \subset \{2, \dots, d\}$ of coordinates for which $r_j = s_j$ for all $j \in J$, then after arbitrarily specifying these values, we have will be in position to apply the inequality $D_{\neq}(d - |J|, n, p)$. This will clearly give a smaller estimate. As the number of possible choices for J is only a function of dimension, this completes the proof.

The Bounds $B(3)$ and $C(3)$. Assuming $B(d-1)$, $d \geq 4$, we prove $C(d)$.

In this section, we will prove the estimates for $C(3)$. As well, we present the inductive proof of $C(d)$ assuming $B(d-1)$, for $d \geq 4$.

For the proof of $C(3)$ there is an ancillary collection that we will have recourse to. Let

$$M(n, p) = \sup_{\mathbb{M}} \|\mathrm{SumProd}(\mathbb{M})\|_p$$

where the supremum is formed over all choices of $\mathbb{M} \subset \mathbb{H}_{n_1} \times \mathbb{H}_{n_2}$ and all r functions subject to these conditions.

- $\vec{r}, \vec{s}$ are three dimensional vectors.

- There is a coincidence in the first coordinate: For all $(\vec{r}, \vec{s}) \in \mathbb{C}$, we have $\vec{r} \neq \vec{s}$ and $r_1 = s_1$.
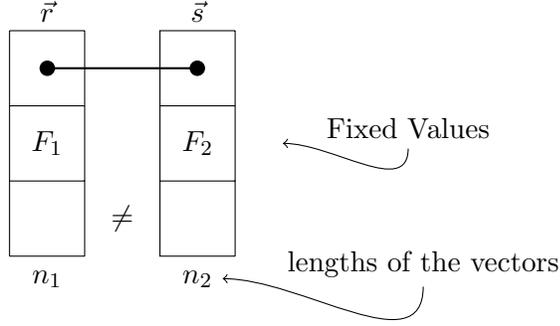
Figure 1: The collections $\mathbb{M}$, with a coincidence in the top row, the second row taking fixed values, and no coincidence in the bottom row.

- The second coordinates are fixed: There are integers F_1, F_2 so that for all $(\vec{r}, \vec{s}) \in \mathbb{M}$ we have $r_2 = F_1$ and $s_2 = F_2$.

- There is no coincidence in the third coordinate: For all $(\vec{r}, \vec{s}) \in \mathbb{M}$ we have $r_3 \neq s_3$.

- $n_1, n_2 \leq n$.

See Figure 1 for an illustration of this collection. We remark that in the case $n_1 \neq n_2$, a coincidence can occur in the third coordinate, a case that will come up below.

Lemma 2.5.11. *We have the inequalities*

$$M(n, p) \lesssim \sqrt{p} \cdot \sqrt{n}.$$

Proof. Notice that the value of the maximum in the third coordinate completely specifies the pair of vectors $(\vec{r}, \vec{s})$. Therefore, one application of the Littlewood-Paley inequalities completes the proof. For any collection $\mathbb{M}$ as above, let $\mathbb{M}_a$ be the $(\vec{r}, \vec{s}) \in \mathbb{M}$ where the maximum in the third coordinate is a, $\max\{r_3, s_3\} = a$. Note that this can only consist, at most, of two pairs of vectors.

$$\|\mathrm{SumProd}(\mathbb{M})\|_p \lesssim \sqrt{p} \Big\| \sum_a \mathrm{SumProd}(\mathbb{M}_a)^2 \Big\|_{p/2}^{1/2} \lesssim \sqrt{p} \cdot \sqrt{n}.$$

$\square$

Fix a dimension $d \geq 3$. Let $\mathbb{B}$ be the collection which satisfies the conditions associated with (2.5.6) that contains $\mathbb{C}$. We introduce a conditional expectation into the argument, to gain some additional symmetry. Let $F_{a,b}$ be the dyadic sigma field in the second and third coordinates generated by dyadic rectangles of side lengths 2^{-a-1} and 2^{-b-1} respectively.

We have this equality.

$$\sum_{\substack{(\vec{r},\vec{s})\in\mathbb{C} \\ r_2=a\,,\,s_3=b}} f_{\vec{r}}\cdot f_{\vec{s}} = \mathbb{E}\Big(\sum_{\substack{(\vec{r},\vec{s})\in\mathbb{B} \\ r_2=a\,,\,s_3=b}} f_{\vec{r}}\cdot f_{\vec{s}}\,|\,F_{a,b}\Big) - \mathrm{SumProd}(\mathbb{D}_{a,b}),$$

where $\mathbb{D}_{a,b}$ consists of pairs of vectors $(\vec{r},\vec{s}) \in \mathbb{B}$ such that $r_1 = s_1$, $a = r_2 = s_2$ and $b = r_3 = s_3$. In three dimensions, the set $\mathbb{D}_{a,b}$ is empty, since the requirements for a pair of vectors being in the set $D_{a,b}$ forces $\vec{r} = \vec{s}$, a contradiction.

Assuming that $d > 3$, using the assumption of $B(d-2)$ (in the case of $d = 4$ we just apply the Littlewood-Paley inequality in the last coordinate), we see that

$$\|\mathrm{SumProd}(\mathbb{D}_{a,b})\|_{p/2} \lesssim p^{d-5/2}\cdot n^{d-7/2}\,.$$

Here, we have 'lost two dimensions' due to the roles of a,b. Therefore, using a trivial estimate in the parameters a,b,

$$p\left\|\left[\sum_{a,b}\mathrm{SumProd}(\mathbb{D}_{a,b})^2\right]^{1/2}\right\|_p \lesssim p^{d-3/2}n^{d-5/2}\,.$$

This estimate is smaller than what the other terms will give us.

Therefore, using (2.3.10) we can estimate

$$\|\mathrm{SumProd}(\mathbb{C})\|_p \lesssim p^{d-3/2}n^{d-5/2} + p^2\left\|\sum_{a,b}\left|\sum_{\substack{(\vec{r},\vec{s})\in\mathbb{B} \\ r_2=a\,,\,s_3=b}} f_{\vec{r}}\cdot f_{\vec{s}}\right|^2\right\|_{p/2}^{1/2}.$$

We concentrate on the latter term, and in particular expand the square.

$$\sum_{a,b}\left|\sum_{\substack{(\vec{r},\vec{s})\in\mathbb{B} \\ r_2=a\,,\,s_3=b}} f_{\vec{r}}\cdot f_{\vec{s}}\right|^2 \lesssim n^{2d-3}$$

$$+ \mathrm{SumProd}(\mathbb{B}'_1) + \mathrm{SumProd}(\mathbb{B}'_2)$$

$$+ \mathrm{SumProd}(\mathbb{B}'') \qquad (2.5.12)$$

where these terms arise as follows. In forming the square on the left in (2.5), we have two pairs $(\vec{r}, \vec{s}), (\underline{\vec{r}}, \underline{\vec{s}}) \in C$ with $r_2 = r_2'$ and $s_3 = s_3'$. We form the product

$$f_{\vec{r}} \cdot f_{\vec{s}} \cdot f_{\underline{\vec{r}}} \cdot f_{\underline{\vec{s}}}$$

- If the two pairs are equal, the product in (2.5) is one. There are $\lesssim n^{2d-3}$ ways to select such pairs. This is the right hand side of (2.5).

- The collection $\mathbb{B}_1'$ consists of vectors such that $\vec{r} = \underline{\vec{r}}$ but $\vec{s} \neq \underline{\vec{s}}$, the product in (2.5) is equal to $f_{\vec{s}} \cdot f_{\underline{\vec{s}}}$ ($\mathbb{B}_2'$ is defined symmetrically). Notice that necessarily we have $s_1 = s_1'$, which is equal to r_1, and $s_3 = s_3'$. Let us set

$$\mathbb{B}_c' = \{(\vec{s}, \underline{\vec{s}}) \ : \ s_1 = \underline{s}_1 = c; \ s_3 = \underline{s}_3\}.$$

We have 'lost' one parameter in $\mathbb{B}_c'$ and have one more coincidence, therefore, we can apply the induction hypothesis $B(d-1)$ to see that

$$\|\mathrm{SumProd}(\mathbb{B}_c')\|_p \lesssim p^{d-3/2} n^{d-5/2}.$$

It is easy to see that

$$\mathrm{SumProd}(\mathbb{B}_1') = \sum_{\vec{r} \in \mathbb{H}_{n_1}} \mathrm{SumProd}(\mathbb{B}_{r_1}').$$

Thus we have

$$\left\|\mathrm{SumProd}(\mathbb{B}_1')\right\|_p \leq \sum_{\vec{r} \in \mathbb{H}_{n_1}} \left\|\mathrm{SumProd}(\mathbb{B}_{r_1}')\right\|_p \leq n^{d-1} \cdot p^{d-3/2} n^{d-5/2} = p^{d-3/2} n^{d-7/2}.$$

This controls the term in (2.5).

- The last term arises from two pairs of vectors $(\vec{r}, \vec{s}), (\underline{\vec{r}}, \underline{\vec{s}}) \in C$ that consist of four distinct vectors. Let us set

$$\mathbb{B}'' = \{(\vec{r}, \vec{s}, \underline{\vec{s}}, \underline{\vec{r}}) \ : \ (\vec{r}, \vec{s}), (\underline{\vec{r}}, \underline{\vec{s}}) \in C, \ \vec{r} \neq \underline{\vec{r}}, \ \vec{s} \neq \underline{\vec{s}}\}$$

Here, for the sake of cleaner graphics, we have deliberately written $\vec{s}, \underline{\vec{s}}$ as the middle two vectors in the four-tuples in $\mathbb{B}''$.
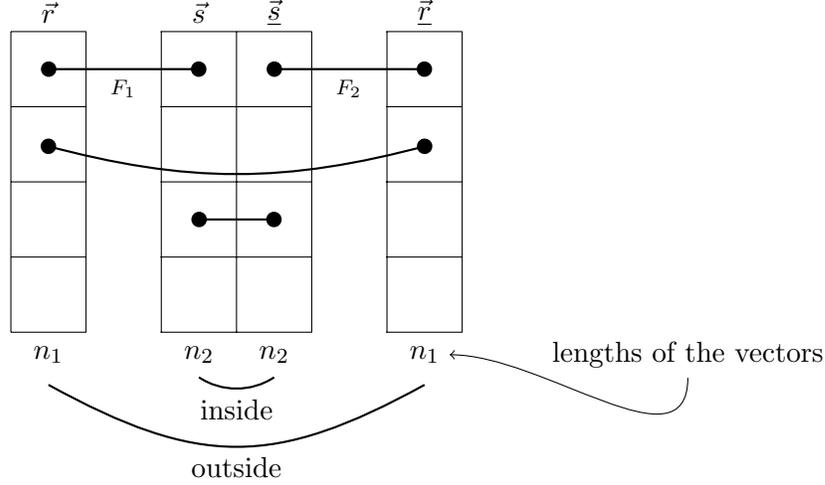
Figure 2: The Decomposition of $\mathbb{B}''_{F_1,F_2}$, in the four dimensional case (the coincidences are indicated by the connected black circles).

It remains to bound the term in (2.5.12). We reduce this four-fold product back to a product of two-fold products. For integers F_1, F_2, let $\mathbb{B}''_{F_1,F_2}$ be those $(\vec{r}, \vec{s}, \underline{\vec{s}}, \underline{\vec{r}}) \in \mathbb{B}''$ with $r_1 = s_1 = F_1$ and $\underline{r}_1 = \underline{s}_1 = F_2$. Let $\mathbb{B}''_{\text{outside},F_1,F_2}$ be the projection of four-tuples in $\mathbb{B}''_{F_1,F_2}$ onto the first and fourth coordinates, and $\mathbb{B}''_{\text{inside},F_1,F_2}$ the projection onto the second and third coordinates. See Figure 2.

For any pair $(\vec{r}, \underline{\vec{r}}) \in \mathbb{B}''_{\text{outside},F_1,F_2}$, and any two pairs

$$(\vec{s}, \underline{\vec{s}}), \ (\vec{\sigma}, \underline{\vec{\sigma}}) \in \mathbb{B}''_{\text{inside},F_1,F_2},$$

we have

$$(\vec{r}, \vec{s}, \underline{\vec{s}}, \underline{\vec{r}}), \ (\vec{r}, \vec{\sigma}, \underline{\vec{\sigma}}, \underline{\vec{r}}) \in \mathbb{B}''_{F_1,F_2}.$$

Therefore, we have the product formula

$$\text{SumProd}(\mathbb{B}''_{F_1,F_2}) = \text{SumProd}(\mathbb{B}''_{\text{outside},F_1,F_2}) \times \text{SumProd}(\mathbb{B}''_{\text{inside},F_1,F_2}).$$

Notice that the pairs of vectors in $\mathbb{B}''_{\text{outside},F_1,F_2}$ have their first coordinates fixed, and have a coincidence in the second coordinate. The fixed first coordinates need not be the same, so that the lengths of the remaining coordinates are, in general, distinct. Still, we may conclude that

$$\|\text{SumProd}(\mathbb{B}''_{\text{outside},F_1,F_2})\|_p \lesssim p^{d-3/2} n^{d-5/2}.$$

This estimate is uniform in F_1, F_2. In the case of dimension $d = 3$, this follows from Lemma 2.5.11, while for $d > 3$ it follows from the induction hypothesis. A similar inequality holds for $\mathbb{B}''_{\text{inside},F_1,F_2}$.

Therefore, we can estimate the term in (2.5.12) as follows:

$$\left\| \text{SumProd}(\mathbb{B}'') \right\|_{p/2}^{1/2} \lesssim pn \sup_{F_1,F_2} \| \text{SumProd}(\mathbb{B}''_{\text{outside},F_1,F_2}) \times \text{SumProd}(\mathbb{B}''_{\text{inside},F_1,F_2}) \|_{p/2}^{1/2}$$

$$\lesssim pn \sup_{F_1,F_2} \| \text{SumProd}(\mathbb{B}''_{\text{outside},F_1,F_2}) \|_p^{1/2} \times \| \text{SumProd}(\mathbb{B}''_{\text{inside},F_1,F_2}) \|_p^{1/2}$$

$$\lesssim (pn)^{d-3/2} .$$

Our proof is complete. Assuming $B(d-1)$, $d \geq 4$, we have proved $C(d)$. We have also proved $C(3)$. The fact that $B(3)$ holds follows from the argument below.

Assuming $C(d)$ and $D(d)$, we prove $B(d)$.

Fix $p, n \geq 3$, a collection of vectors $\mathbb{B}$ and r functions which achieve the supremum in (2.5.6). Write this collection as

$$\mathbb{B} = \mathbb{D} \cup \bigcup_{2 \leq i \neq j \leq d} \mathbb{C}_{i,j}$$

where $\mathbb{C}_{i,j}$ consists of those pairs $(\vec{r}, \vec{s}) \in \mathbb{B}$ such that i is the first coordinate for which $r_i > s_i$ and j is the first coordinate for which $r_j < s_j$. Then, the collections $\mathbb{C}_{i,j}$ are pairwise disjoint, and the collection $\mathbb{D}$ consists of all pairs not in some $\mathbb{C}_{i,j}$. Thus,

$$\text{SumProd}(\mathbb{B}) = \text{SumProd}(\mathbb{D}) + \sum_{2 \leq i \neq j \leq d} \text{SumProd}(\mathbb{C}_{i,j}) .$$

After a harmless permutation of indices, the inequalities $C(d)$ apply to the collections $\mathbb{C}_{i,j}$. The (unconditional) inequalities D apply to the collection $\mathbb{D}$. The proof is complete.

2.6 Corollaries of the Beck Gain

Theorem 2.3.8 implies an exponential estimate of order $\exp(L^{2/(d-1)})$ for sums of $\vec{r}$ functions. In fact, we can derive a *subgaussian* estimate for such sums, for moderate deviations, and moreover, in order to have a gain of order n^{c/d^2} in our Main Theorem, we need to use this estimate.

Theorem 2.6.1. *Using the notation of (2.4.2) and (2.4.3), we have this estimate, valid for all $1 \leq t \leq q$.*

$$\|\rho F_t\|_p \lesssim \sqrt{p}, \qquad 1 \leq p \leq cn^{\frac{1-2\varepsilon}{2d-1}}. \tag{2.6.2}$$

As a consequence, we have the distributional estimate

$$\mathbb{P}(|\rho F_t| > x) \lesssim \exp(-cx^2), \qquad x < cn^{\frac{1-2\varepsilon}{4d-2}}. \tag{2.6.3}$$

Here $0 < c < 1$ is an absolute constant.

To use (2.6.3), we need $q^b = a^b n^{\varepsilon \cdot b} < cn^{\frac{1}{4d-6}}$, and so $\varepsilon \simeq 1/d$ is the optimal value for ε that this proof will give.

Proof. Recall that

$$F_t = \sum_{\vec{r} \in \mathbb{A}_t} f_{\vec{r}}.$$

where $\mathbb{A}_t = \{\vec{r} \in \mathbb{H}_n : r_1 \in I_t\}$, and I_t in an interval of integers of length n/q, so that $\sharp \mathbb{A}_t \simeq n^{d-1}/q \simeq \rho^{-2}$.

Apply the Littlewood-Paley inequality in the first coordinate. This results in the estimate

$$\|\rho F_t\|_p \lesssim \sqrt{p} \left\| \left[\sum_{s \in I_j} \left| \rho \sum_{\vec{r}:r_1=s} f_{\vec{r}} \right|^2 \right]^{1/2} \right\|_p$$

$$\lesssim \sqrt{p} \|1 + \rho^2 \Phi_{t,t,1}\|_{p/2}^{1/2}$$

$$\lesssim \sqrt{p} \left\{ 1 + \|\rho^2 \Phi_{t,t,1}\|_{p/2}^{1/2} \right\},$$

where $\Phi_{t,t,1}$ is defined in (2.5.1). Here it is important to use the constants in the Littlewood-Paley inequalities that give the correct order of growth of $\sqrt{p}$. Of course the terms $\Phi_{t,t,1}$ are controlled by the estimate in (2.5.3). In particular, we have

$$\|\rho^2 \Phi_{t,t,1}\|_p \lesssim \frac{q}{n^{d-1}} p^{d-1/2} n^{d-3/2} \lesssim q \, p^{d-1/2} n^{-1/2} \lesssim 1.$$

Hence (2.6.2) follows.

The second distributional inequality is a well known consequence of the norm inequality. Namely, one has the inequality below, valid for all x:

$$\mathbb{P}(\rho F_t > x) \leq C^p p^{p/2} x^{-p}, \qquad 1 \leq p \leq c n^{\frac{1-2\varepsilon}{2d-1}}.$$

If x is as in (2.6.3), we can take $p \simeq x^2$ to prove the claimed exponential squared bound. $\quad\square$

We shall now use the Beck Gain to prove the crucial L^2 estimate (2.4.7) of Lemma 2.4.5. We actually need a slightly more general inequality:

Lemma 2.6.4. *We have the following estimate:*

$$\sup_{V \subset \{1,\ldots,q\}} \mathbb{E} \prod_{v \in V} (1 + \widetilde{\rho} F_t)^2 \lesssim \exp(a' q^{2b}).$$

The supremum over V will be an immediate consequence of the proof below, and so we don't address it specifically.

Proof of (2.4.7). Let us give the essential initial observation. We expand

$$\mathbb{E} \prod_{j=1}^q (1 + \widetilde{\rho} F_j)^2 = \mathbb{E} \prod_{j=1}^q (1 + 2\widetilde{\rho} F_j + (\widetilde{\rho} F_j)^2).$$

Hold the last $d-1$ coordinates, $x_2, \ldots, x_d$, fixed and let F be the sigma field generated by $F_1, \ldots, F_{q-1}$. We have

$$\mathbb{E}(1 + 2\widetilde{\rho} F_q + (\widetilde{\rho} F_q)^2 \mid F) = 1 + \mathbb{E}((\widetilde{\rho} F_q)^2 \mid F)$$

$$= 1 + a^2 q^{2b-1} + \widetilde{\rho}^2 \mathbb{E}(\Phi_{q,q,1} \mid F),$$

where $\Phi_{q,q,1}$ is defined in (2.5.1). Then, we see that

$$\mathbb{E} \prod_{v=1}^q (1 + 2\widetilde{\rho} F_t + (\widetilde{\rho} F_t)^2) = \mathbb{E}\Big\{ \prod_{v=1}^{q-1} (1 + 2\widetilde{\rho} F_t + (\widetilde{\rho} F_t)^2) \times \mathbb{E}(1 + 2\widetilde{\rho} F_t + (\widetilde{\rho} F_t)^2 \mid F) \Big\}$$

$$\leq (1 + a^2 q^{2b-1}) \mathbb{E} \prod_{v=1}^{q-1} (1 + 2\widetilde{\rho} F_t + (\widetilde{\rho} F_t)^2) \qquad (2.6.5)$$

$$+ \mathbb{E}|\widetilde{\rho}^2 \Phi_{q,q,1}| \cdot \prod_{v=1}^{q-1} (1 + 2\widetilde{\rho} F_t + (\widetilde{\rho} F_t)^2) \qquad (2.6.6)$$

This is the main observation: one should induct on (2.6.5), while treating the term in (2.6.6) as an error, as the Beck Gain estimate (2.5.3) applies to it.

Let us set up notation to implement this line of approach. Set

$$N(V;r) = \left\| \prod_{t=1}^{V} (1 + \widetilde{\rho} F_t) \right\|_r , \qquad V = 1, \ldots, q .$$

We will obtain a very crude estimate for these numbers for $r = 4$. Fortunately, this is relatively easy for us to obtain. Namely, q is small enough that we can use the inequalities (2.6.2) to see that

$$N(V;4) \leq \prod_{v=1}^{V} \| 1 + \widetilde{\rho} F_t \|_{4V}$$

$$\leq (1 + Cq^{1/2+b})^V$$

$$\leq (Cq)^q .$$

We have the estimate below from Hölder's inequality

$$N(V;2(1 - 1/q)^{-1}) \leq N(V;2)^{1-1/q} \cdot N(V;4)^{1/q} . \tag{2.6.7}$$

We see that (2.6.5), (2.6.6) and (2.6.7) give us the inequality

$$N(V+1;2)^2 \leq (1 + a^2 q^{2b-1}) N(V;2)^2 + C \cdot N(V;2(1-1/q)^{-1})^2 \cdot \| \widetilde{\rho}^2 \Phi_{V+1,V+1,1} \|_q$$

$$\leq (1 + a^2 q^{2b-1}) N(V;2)^2 + C N(V;2)^{2-2/q} \cdot N(V;4)^{2/q} \| \widetilde{\rho}^2 \Phi_{V+1,V+1,1} \|_q$$

$$\leq (1 + a^2 q^{2b-1}) N(V;2)^2 + C q^{d+2} n^{-1/2} N(V;2)^{2-2/q} .$$

In the last line we have used the inequality (2.5.3). Of course we only apply this as long as $N(V;2) \geq 1$. Assuming this is true for all $V \geq 1$, we see that

$$N(V+1;2)^2 \leq (1 + a^2 q^{2b-1} + C q^{d+2} n^{-1/2}) N(V;2)^2 .$$

And so, by induction,

$$N(q;2) \lesssim (1 + a^2 q^{2b-1} + C q^{d+2} n^{-1/2})^{q/2} \lesssim e^{2a q^{2b}} .$$

Here, the last inequality will be true for large n, provided that ε in the definition of q (2.4.1) is small. Indeed, we need

$$a^2 q^{2b-1} \geq C q^{d+2} n^{-1/2}$$

28

Or equivalently,

$$a^2 n^{1/2} \gtrsim q^{d+5/2} \,.$$

Comparing to the definition of q in (2.4.1), we see that the proof is finished. □

One should notice that the results of this section suggest that our methods give a gain of the order $\frac{1}{d}$.

2.7 The Beck Gain with Fixed Parameters.

We will need to analyze longer products of r functions. These longer products will be reduced to the case of a a slightly more general version of the Beck Gain Lemma 2.5.2. Namely, we will consider sums of products of two r fucntions, but impose the additional restriction for some coordinates in the pair of vectors to have fixed values. Let $\vec{a} \in \mathbb{N}^{F_1}$ and $\vec{b} \in \mathbb{N}^{F_2}$ be integer vectors with lengths $|\vec{a}|, |\vec{b}| < n$. We will be estimating the quantity:

$$B(F_1, F_2) = \sup_{\vec{a}, \vec{b}, j_1 < j_2} \sup_{\mathbb{B}} \|\text{SumProd}(\mathbb{B})\|_p \,, \qquad d, n, p \geq 3 \,.$$

The inner supremum is formed over all $\mathbb{B} \subset \mathbb{H}_n \times \mathbb{H}_n$ and all r functions subject to these conditions:

- $\vec{r} \in \mathbb{A}_{j_1}$, $\vec{s} \in \mathbb{A}_{j_2}$, where $j_1 < j_2$ (i.e. s_1 is the maximum in the first coordinate.)

- There is a coincidence in the second coordinate: For all $(\vec{r}, \vec{s}) \in \mathbb{B}$, we have $\vec{r} \neq \vec{s}$ and $r_2 = s_2$.

- For $k = 1, \ldots, F_1$, we have $r_{k+2} = a_k$. (F_1 coordinates of $\vec{r}$ are fixed.)

- For $k = 1, \ldots, F_2$, we have $s_{F_1+k+2} = b_k$. (F_2 coordinates of $\vec{s}$ are fixed, and these coordinates are distinct from the other vector.)

We have the following estimate, which gives an average Beck Gain of $n^{1/8}$ for each of the two functions in the product.

Lemma 2.7.1. *We have the inequality below valid for all dimensions $d \geq 3$.*

$$B(F_1, F_2) \lesssim p^{d-1-\frac{F_1+F_2}{2}-\frac{1}{4}} n^{d-1-\frac{F_1+F_2}{2}-\frac{1}{4}} \,, \qquad p, n \geq 3 \,.$$

Proof. We will reduce this situation to the Beck Gain proven before. Let $\mathbb{B}$ be as above. First of all, we shall apply the Littlewood-Paley inequality in the first coordinate. Notice that the maximum in this coordinate is automatically s_1.

$$\|\mathrm{SumProd}(\mathbb{B})\|_p \lesssim \sqrt{p} \left\| \sum_{c \in I_{j_2}} \left| \sum_{\substack{(\vec{r},\vec{s}) \in \mathbb{B} \\ s_1 = c}} f_{\vec{r}} \cdot f_{\vec{s}} \right|^2 \right\|_{p/2}^{1/2}$$

We concentrate on the latter term, and in particular expand the square.

$$\sqrt{p} \left\| \sum_{c \in \mathbb{I}_{j_2}} \left| \sum_{\substack{(\vec{r},\vec{s}) \in \mathbb{B} \\ s_1 = c}} f_{\vec{r}} \cdot f_{\vec{s}} \right|^2 \right\|_{p/2}^{1/2} = \sqrt{p} \left\| \sum_{\substack{(\vec{r},\vec{s},\underline{\vec{r}},\underline{\vec{s}}) \in \mathbb{B} \times \mathbb{B} \\ s_1 = \underline{s}_1}} f_{\vec{r}} \cdot f_{\vec{s}} \cdot f_{\underline{\vec{r}}} \cdot f_{\underline{\vec{s}}} \right\|_{p/2}^{1/2}$$

$$\leq \sqrt{p}n \max_{c \neq \underline{c}} \left\| \sum_{\substack{(\vec{r},\vec{s},\underline{\vec{r}},\underline{\vec{s}}) \in \mathbb{B} \times \mathbb{B} \\ s_1 = \underline{s}_1; \, r_2 = s_2 = c; \, \underline{r}_2 = \underline{s}_2 = \underline{c}}} f_{\vec{r}} \cdot f_{\vec{s}} \cdot f_{\underline{\vec{r}}} \cdot f_{\underline{\vec{s}}} \right\|_{p/2}^{1/2} \quad (2.7.2)$$

$$+ \sqrt{p}\sqrt{n} \max_{c} \left\| \sum_{\substack{(\vec{r},\vec{s},\underline{\vec{r}},\underline{\vec{s}}) \in \mathbb{B} \times \mathbb{B} \\ s_1 = \underline{s}_1; \, r_2 = s_2 = \underline{r}_2 = \underline{s}_2 = c}} f_{\vec{r}} \cdot f_{\vec{s}} \cdot f_{\underline{\vec{r}}} \cdot f_{\underline{\vec{s}}} \right\|_{p/2}^{1/2} \quad (2.7.3)$$

We start with the estimates for the first term above (2.7.2):

$$\sqrt{p}n \max_{c \neq \underline{c}} \left\| \sum_{\substack{(\vec{r},\vec{s},\underline{\vec{r}},\underline{\vec{s}}) \in \mathbb{B} \times \mathbb{B} \\ s_1 = \underline{s}_1; \, r_2 = s_2 = c; \, \underline{r}_2 = \underline{s}_2 = \underline{c}}} f_{\vec{r}} \cdot f_{\vec{s}} \cdot f_{\underline{\vec{r}}} \cdot f_{\underline{\vec{s}}} \right\|_{p/2}^{1/2}$$

$$= \sqrt{p}n \max_{c \neq \underline{c}} \left\| \left(\sum_{(\vec{r},\underline{\vec{r}}) \in \mathbb{B}_1} f_{\vec{r}} \cdot f_{\underline{\vec{r}}} \right) \times \left(\sum_{(\vec{s},\underline{\vec{s}}) \in \mathbb{B}_2} f_{\vec{s}} \cdot f_{\underline{\vec{s}}} \right) \right\|_{p/2}^{1/2}$$

$$\leq \sqrt{p}n \max_{c \neq \underline{c}} \left\| \sum_{(\vec{r},\underline{\vec{r}}) \in \mathbb{B}_1} f_{\vec{r}} \cdot f_{\underline{\vec{r}}} \right\|_{p}^{1/2} \left\| \sum_{(\vec{s},\underline{\vec{s}}) \in \mathbb{B}_2} f_{\vec{s}} \cdot f_{\underline{\vec{s}}} \right\|_{p}^{1/2}$$

Here $\mathbb{B}_1$ is defined to consist of pairs $(\vec{r}, \underline{\vec{r}}) \in \mathbb{A}_{j_1}^2$ which satisfy the following:

- For $k = 1, \ldots, F_1$, we have $r_{k+2} = \underline{r}_{k+2} = a_k$.

- $r_2 = c$, $\underline{r}_2 = \underline{c}$.

And similarly $\mathbb{B}_2$ consists of pairs $(\vec{s}, \underline{\vec{s}}) \in \mathbb{A}_{j_2}^2$ with the properties:

- For $k = 1, \ldots, F_2$, we have $s_{k+F_1+2} = \underline{s}_{k+F_1+2} = b_k$.

- $s_2 = c$, $\underline{s}_2 = \underline{c}$.

30

- Moreover, we have $s_1 = \underline{s}_1$.

Notice that because of the last condition and the fact that $c \neq \underline{c}$ (i.e., $\vec{s} \neq \underline{\vec{s}}$), the Beck Gain (Lemma 2.5.2) applies to this family of pairs, giving a gain of $n^{1/2}$, while $\mathbb{B}_1$ will be estimated by simple parameter counting, supplying no gain. We have

$$\left\| \mathrm{SumProd}(\mathbb{B}_1) \right\|_p \lesssim (pn)^{d-2-F_1},$$

$$\left\| \mathrm{SumProd}(\mathbb{B}_2) \right\|_p \lesssim p^{d-3/2-F_2} n^{d-2-F_2-1/2}.$$

And thus we can estimate the term (2.7.2) by

$$\sqrt{pn} \left\| \sum_{\substack{(\vec{r}, \vec{s}, \underline{\vec{r}}, \underline{\vec{s}}) \in \mathbb{B} \times \mathbb{B} \\ s_1 = \underline{s}_1; \, r_2 = s_2 = c; \, \underline{r}_2 = \underline{s}_2 = \underline{c}}} f_{\vec{r}} \cdot f_{\vec{s}} \cdot f_{\underline{\vec{r}}} \cdot f_{\underline{\vec{s}}} \right\|_{p/2}^{1/2} \lesssim$$

$$\lesssim \sqrt{pn} \left((pn)^{d-2-F_1} \right)^{1/2} \left(p^{d-3/2-F_2} n^{d-2-F_2-1/2} \right)^{1/2} = (pn)^{d-1-\frac{F_1+F_2}{2}-\frac{1}{4}}.$$

The second term (2.7.3) satisfies the same bound in n. This can be shown by simple parameter counting, the gain comes from the loss of one parameter since $c = \underline{c}$.

We remark that in this version of the Beck gain 'error terms' do not arise, since we apply Littlewood-Paley inequality only in the first coordinate, where we already have a natural order. Thus we do not need to use the conditional expectation argument as in the proof of Lemma 2.5.2.

$\square$

2.8 The Beck Gain for Longer Coincidences

In the present section we treat longer coincidences. This requires a careful analysis of the variety of ways that a product can fail to be strongly distinct. That is, we need to understand the variety of ways that coincidences can arise, and how coincidences can contribute to a smaller norm. Following Beck, we will use the language of Graph Theory to describe these general patterns of coincidences.

2.8.1 Graph Theory Nomenclature

We adopt familiar nomenclature from Graph Theory, although there is no graph theoretical fact that we need, rather the use of this language is just a convenient way to do some

bookkeeping. The class of graphs that we are interested in satisfies particular properties. A $d-1$ *colored graph* G is the tuple $(V(G), E_2, E_3, \ldots, E_d)$, of the *vertex set* $V(G) \subset \{1, \ldots, q\}$, and *edge sets E_2, E_3, ... E_d, of colors $2, 3, \ldots, d$ respectively.* Edge sets are are subsets of

$$E_j \subset V(G) \times V(G) - \{(k, k) \mid k \in V(G)\}.$$

Edges are symmetric, thus if $(v, v') \in E_j$ then necessarily $(v', v) \in E_j$.

A *clique of color j* is a maximal subset $Q \subset V(G)$ such that for all $v \neq v' \in Q$ we have $(v, v') \in E_j$. By *maximality*, we mean that no strictly larger set of vertices $Q' \supset Q$ satisfies this condition.

Call a graph G *admissible* iff

- The edges sets, in all $d - 1$ colors, decompose into a union of cliques.

- If Q_k's are cliques of color k $(k = 2, \ldots, d)$, then $\bigcap_{k=2}^d Q_k$ contains at most one vertex.

- Every vertex is in at least one clique.

A graph G is *connected* iff for any two vertices in the graph, there is a path that connects them. A *path* in the graph G is a sequence of vertices $v_1, \ldots, v_k$ with an edge of *any color,* spanning adjacent vertices , that is $(v_j, v_{j+1}) \in \cup_{k=2}^d E_k$.

2.8.2 Reduction to Admissible Graphs

It is clear that admissible graphs as defined above are naturally associated to sums of products of r functions. Given admissible graph G on vertices V, we set $X(G)$ to be those tuples of r vectors

$$\vec{r}_v \in \prod_{v \in V} \mathbb{A}_v$$

so that if (v, v') is an edge of color j in G, then $r_{v,j} = r_{v',j}$.

We shall introduce the following counting parameter: for an admissible graph G, its index, $ind(G)$, is defined as

$$ind(G) = \sum_{Q \text{ is a clique}} (\sharp Q - 1).$$

Effectively, the index of G is the least number of equalities, needed to define $X(G)$, in other words, the number of coincidences. In particular, for the graphs, corresponding to the simplest case of the Beck Gain, the index is one.

With these definitions at hand, it is not hard to obtain the Inclusion-Exclusion formula, relating admissible graphs and the 'not strongly distinct' part of the Riesz product:

$$\Psi^{\neg} = \sum_{G \text{ admissible}} (-1)^{ind(G)+1} \, \widetilde{\rho}^{\,|V(G)|} \, \mathrm{SumProd}(X(G)) \cdot \prod_{t \notin V(G)} (1 + \widetilde{\rho} F_t). \qquad (2.8.1)$$

We will prove the following Theorem:

Theorem 2.8.2. Beck Gain for Graphs *(Bilyk, Lacey, Vagharshakyan, [6]) For an admissible graph G on vertices V we have the estimate below for positive, finite constants C_0, C_1, C_2, C_3:*

$$\rho^{|V|} \|\mathrm{SumProd}(X(G))\|_p \le [C_0 |V|^{C_1} p^{C_2} q^{C_3} n^{-\eta}]^{|V|}, \qquad 2 < p < \infty. \qquad (2.8.3)$$

The most significant term on the right is $n^{-\eta}$. It shows that as the number of coincidences goes up, the corresponding 'Beck Gain' improves. Notice that for the other terms on the right, C_0 is a constant; $|V| \le q \le n^{\epsilon}$, where we can choose $0 < \epsilon$ as a function of η; and while the inequality above holds for all $2 \le p < \infty$, we will only need to apply it for $p \lesssim q^{2b} \le n^{\epsilon/2}$. That is, the $n^{-\eta}$ is the dominant term on the right. This Theorem, together with the fact that there are at most $|V|^{2d|V|}$ admissible graphs on the vertex set V, yields the boundedness of the sum in (2.4.13).

2.8.3 Norm Estimates for Admissible Graphs

We begin the proof of Theorem 2.8.2 with a further reduction to connected admissible graphs. Let us write $G \in \mathrm{BG}(C_0, C_1, C_2, C_3, \eta)$ if the estimates (2.8.3) holds. ('BG' for 'Beck Gain.') We need to see that all admissible graphs are in $\mathrm{BG}(C_0, C_1, C_2, C_3, \eta)$ for non-negative, finite choices of the relevant constants.

Lemma 2.8.4. *Let C_0, C_1, C_2, C_3, η be non-negative constants. Suppose that G is an admissible graph, and that it can be written as a union of subgraphs $G_1, \ldots, G_k$ on disjoint*

vertex sets, where all $G_j \in \mathrm{BG}(C_0, C_1, C_2, C_3, \eta)$. Then,

$$G \in \mathrm{BG}(C_0, C_1, C_2, C_2 + C_3, \eta).$$

With this Lemma, we will identify a small class of graphs for which we can verify the property (2.8.3) directly, and then appeal to this Lemma to deduce Lemma 2.8.2. Accordingly, we modify our notation. If G is a class of graphs, we write $G \subset \mathrm{BG}(\eta)$ if there are constants C_0, C_1, C_2, C_3 such that $G \subset \mathrm{BG}(C_0, C_1, C_2, C_3, \eta)$.

Proof. We then have by Proposition 2.8.5

$$\mathrm{SumProd}(X(G)) = \prod_{j=1}^{k} \mathrm{SumProd}(X(G_j)).$$

Using Hölder's inequality, we can estimate

$$\rho^{|V|} \|\mathrm{SumProd}(X(G))\|_p \leq \prod_{j=1}^{k} \rho^{|V_j|} \|\mathrm{SumProd}(X(G_j))\|_{kp}$$

$$\leq \prod_{j=1}^{k} [C_0 (kp)^{C_1} q^{C_2} n^{-\eta}]^{|V_j|}$$

$$\leq [C_0 p^{C_1} q^{C_2 + C_1} n^{-\eta}]^{|V|}.$$

Here, we use the fact that since the graphs are non-empty, we necessarily have $k \leq q$.

$\square$

Proposition 2.8.5. *Let $G_1, \ldots, G_p$ be admissible graphs on pairwise disjoint vertex sets $V_1, \ldots, V_p$. Extend these graphs in the natural way to a graph G on the vertex set $V = \bigcup V_t$. Then, we have*

$$\mathrm{SumProd}(X(G)) = \prod_{t=1}^{p} \mathrm{SumProd}(X(G_t)).$$

2.8.4 Connected Graphs Have the Beck Gain.

We single out for special consideration the connected admissible graphs G. Let $G_{\mathrm{connected}}$ be the collection of of all admissible connected graphs on $V \subset \{1, \ldots, q\}$.

Lemma 2.8.6. *We have $G_{\mathrm{connected}} \subset \mathrm{BG}(\eta)$ for some $\eta > 0$.*

The point of this proof is that we will reduce this question to a much simpler key fact, namely Lemma 2.7.1, which we restate here in our current notation.[3]

Let $G_{\text{fixed}}(2)$ be the set of graphs—and sets of r functions associated with the graphs—with these properties:

- G is a connected graph on two vertices $\{v, v'\}$. That is, there is at least one edge that connects these to vertices. Denote by $C \subset \{2, \ldots, d\}$ the set of coordinates corresponding to the edges.

- There are a set of coordinates $F_v, F_{v'} \subset \{2, \ldots, d\}$ that are disjoint from the set of edges, and two vectors $\vec{a} \in \mathbb{N}^{F_v}$ and $\vec{a}' \in \mathbb{N}^{F_{v'}}$, so that we define

$$Y(G) := \{(\vec{r}_v, \vec{r}_{v'}) \in \mathbb{H}_n \; : \; r_{v,j} = r_{v',j} \; \forall j \in C \, ; \; r_{v,k} = a_k \; \forall k \in F_v \, ; \; r_{v',k} = a_k \; \forall k \in F_{v'}\}$$

These are in essence the assumptions of Lemma 2.7.1. This Lemma proves that

$$\|\text{SumProd}(Y(G))\|_p \lesssim p^d n^\sigma \, , \qquad \sigma = d - 1 - \frac{F_v + F_{v'}}{2} - \frac{1}{4} \, .$$

By abuse of notation, let us summarize this inequality by the inclusion $G_{\text{fixed}}(2) \subset \text{BG}(C_0, C_1, d/2, 0, 1/8)$. Or, even more briefly, as $G_{\text{fixed}}(2) \subset \text{BG}(1/8)$. That is, there is a gain of $\frac{1}{8}$ for each vertex. It follows from the proof of Lemma 2.8.4, that if G is any graph whose connected components are each elements of $G_{\text{fixed}}(2)$, then $G \in \text{BG}(1/8)$.

Our line of attack on this Lemma is to take a general connected graph G, use the triangle inequality to assign fixed values to a number of edges, making the connected components of the new graph to be elements of $G_{\text{fixed}}(2)$. The proportion of vertices that will be in one of these graphs will be at least $1/2d$ of all vertices. And therefore connected graphs will be in $\text{BG}(1/16d)$.

Remark 2.8.7. A heuristic guides this argument. The normalization $\rho^{|V|}$ in (2.8.3) assigns a weight $n^{-1/2}$ to each free parameter of $X(G)$, ignoring losses of parameters from the edges of G. If (v, v') is an edge in the graph, and we assign the edge one of n possible values, the

---

full power of n is exactly compensated by the collective weight of the two parameters in the edge. Therefore, we are free to fix a fixed proportion of edges in the graph, obtaining a Beck Gain on the remaining proportion. In this argument, if the edge is in a clique of size at least $k \geq 3$, specifying a single value on this clique actually leads to a positive gain of $n^{-k/2+1}$. In other words, graphs, all of whose cliques are of size two, are extremal with respect to this analysis (see Lemma 2.8.8). This heuristic is made precise in the proof below.

By 'deleting a clique' we shall mean fixing a value of the coincidence which corresponds to that clique. Let $G \in G_{\text{connected}}$. Following the heuristic above, in the first step of the algorithm we delete all cliques of size at least 3 in G.

After this step G breaks down into connected components, which are admissible graphs with cliques only of size 2 (and, possibly, some singletons). Next, we want to obtain an estimate for such graphs.

Lemma 2.8.8. *Suppose $\widetilde{G} \in G_{\text{connected}}$ has cliques of size at most 2. Then $\widetilde{G} \in BG(\frac{1}{16d})$.*

To prove this statement we shall use the following property of $\widetilde{G}$:

- The degree of each vertex in $\widetilde{G}$ is at most $d - 1$ (since the degree in each color is at most one).

Let $\widetilde{V}$ be the set of vertices of $\widetilde{G}$, and $\widetilde{E}$ be the set of all its edges. The point is to select a maximal subset $\widetilde{E}_{\text{indpndt}}$ of *independent* edges. That is, no two edges in $\widetilde{E}_{\text{indpndt}}$, regardless of color, have a common vertex. It is an elementary fact that we can take

$$|\widetilde{E}_{\text{indpndt}}| \geq \tfrac{1}{2d-3}|\widetilde{E}|\,.$$

Indeed, each edge in $\widetilde{G}$ shares a vertex with at most $2d-4$ distinct edges, which observation directly implies the inequality above.

We delete all other edges of $\widetilde{G}$ (i.e. we fix some choice of parameters for the corresponding coincidences) and thus $\widetilde{G}$ breaks down into a number of components each of which is either a singleton or a graph with two vertices and one edge. The latter components correspond exactly to the situation in which the Beck gain of the previous section is applicable. Let

36

us denote these pairs by $G_k \in G_{\text{fixed}}(2)$, $k = 1, \ldots, N = |\widetilde{E}_{\text{indpndt}}|$; the singletons – by v_j,

$j = 1, \ldots, |\widetilde{V}| - 2N$. Let also $E' = \widetilde{E} - \widetilde{E}_{\text{indpndnt}}$ denote the set of all deleted edges in $\widetilde{G}$.

Denote also by F_k the number of fixed parameters in $X(G_k)$ and F'_j will be the number of

fixed parameters in $\vec{r}_{v_j}$. We have the following relations:

$$2|E'| = 2|\widetilde{E} - \widetilde{E}_{\text{indpndnt}}| = \sum_{k=1}^{N} F_k + \sum_{j=1}^{|\widetilde{V}|-2N} F'_j, \tag{2.8.9}$$

and, since $\widetilde{G}$ is connected, it has at least $|V(G)| - 1$ edges, thus

$$N \geq \frac{|\widetilde{E}|}{2d-3} \geq \frac{|\widetilde{V}| - 1}{2d-3} \geq \frac{|\widetilde{V}|}{2(2d-3)} \geq \frac{|\widetilde{V}|}{4d}. \tag{2.8.10}$$

Besides, by Proposition 2.8.5, we obtain the following equality (the sum below is taken over

all choices of parameters on the 'deleted' edges):

$$\text{SumProd}(X(\widetilde{G})) = \sum \prod_{k=1}^{N} \text{SumProd}(X(G_k)) \cdot \prod_{j=1}^{|\widetilde{V}|-2N} \text{SumProd}(X(v_j)).$$

Now we apply the triangle inequality, Hölder's inequality, the relations (2.8.9) and

(2.8.10), and the Beck gain in the form of Lemma 2.7.1 to estimate ($\kappa = |\widetilde{V}| - N < q$):

$$\rho^{|\widetilde{V}|} \|\text{SumProd}(X(\widetilde{G}))\|_p \leq n^{|E'|} \cdot \prod_{k=1}^{N} \rho^2 \|\text{SumProd}(X(G_k))\|_{\kappa p} \cdot \prod_{j=1}^{|\widetilde{V}|-2N} \rho \|f_{\vec{r}_{v_j}}\|_{\kappa p}$$

$$\lesssim n^{|E'|} \cdot \prod_{k=1}^{N} \left[\rho^2 (\kappa p\, n)^{d-1-\frac{F_k}{2}-\frac{1}{4}} \right] \cdot \prod_{j=1}^{|\widetilde{V}|-2N} \left[\rho(\kappa p\, n)^{\frac{d-1}{2}-\frac{F'_j}{2}} \right]$$

$$\lesssim \left[C p^{\frac{d-1}{2}} q^{\frac{d}{2}} \right]^{|\widetilde{V}|} \cdot n^{-\frac{N}{4}} \lesssim \left[p^{\frac{d-1}{2}} q^{\frac{d}{2}} n^{-\frac{1}{16d}} \right]^{|\widetilde{V}|}.$$

This proves Lemma 2.8.8. The point of passing to the collection of independent edges

is that $\text{SumProd}(X(\widetilde{G}))$ splits into a product of terms associated with graphs in $G_{\text{fixed}}(2)$.

Each of these graphs leads to a gain of at least $\frac{1}{8}$ for each vertex. But by (2.8.9), there are

at least $\frac{1}{2d}|V(G)|$ vertices for which we will get this gain. This shows that $G \in BG(1/16d)$.

We can now proceed to prove Lemma 2.8.6 – the proof will be in the same spirit.

After we delete "large" (of size at least 3) cliques of G, this graph decomposed into some

singletons and components as in Lemma 2.8.8 (but with some parameters fixed). Denote these components by $\widetilde{G}_k$, $k = 1, \ldots, n_1$ and the singletons by u_j, $j = 1, \ldots, n_2$. Let f_k be the number of fixed parameters in $X(\widetilde{G}_k)$ and and f'_j – the number of fixed parameters in $\vec{r}_{u_j}$. Notice that the proof of Lemma 2.8.8 can be trivially adapted to the case when some parameters are fixed to obtain the estimate:

$$\rho^{|\widetilde{V}_k|} \|\mathrm{SumProd}(X(\widetilde{G}_k))\|_p \leq \left[C p^{\frac{d-1}{2}} q^{\frac{d}{2}} n^{-\frac{1}{16d}} \right]^{|\widetilde{V}_k|} n^{-\frac{f_k}{2}}.$$

Also, if we denote by K the total number of fixed cliques, one can see that, since all the cliques had size at least 3, we have the inequality:

$$3K \leq \sum_{k=1}^{n_1} f_k + \sum_{j=1}^{n_2} f'_j.$$

Let us write the set of vertices of G as $V = V_1 \cup V_2$, where V_1 are the vertices involved in at least one of the deleted cliques and V_2 are all the other vertices. It is easy to see that $V_2 \subset \cup_{k=1}^{n_1} V(\widetilde{G}_k)$. Indeed, all the vertices that became singletons had to be a part of one of the deleted cliques. Thus,

$$|V_2| \leq \sum_{k=1}^{n_1} |V(\widetilde{G}_k)|.$$

Besides, it is easy to see that

$$|V_1| \leq \sum_{k=1}^{n_1} f_k + \sum_{j=1}^{n_2} f'_j,$$

because at least one parameter is fixed in each vertex from a deleted clique. Using these relations, similarly to the proof of Lemma 2.8.8, taking $\kappa = n_1 + n_2 < q$, we can write:

$$\rho^{|V|} \|\mathrm{Prod}(X(G))\|_p \leq n^K \cdot \prod_{k=1}^{n_1} \rho^{|V(\widetilde{G}_k)|} \|\mathrm{Prod}(X(\widetilde{G}_k))\|_{\kappa p} \cdot \prod_{j=1}^{n_2} \rho \|f_{\vec{r}_{u_j}}\|_{\kappa p}$$

$$\lesssim n^K \cdot \prod_{k=1}^{n_1} \left[C p^{\frac{d-1}{2}} q^d n^{-\frac{1}{16d}} \right]^{|V(\widetilde{G}_k)|} n^{-\frac{f_k}{2}} \cdot \prod_{j=1}^{n_2} \left[p^{\frac{d-1}{2}} q^d n^{-\frac{f'_j}{2}} \right]$$

$$\lesssim \left[C p^{\frac{d-1}{2}} q^d \right]^{|V|} \cdot n^{K - \frac{1}{2}(\sum f_k + \sum f'_j) - \frac{1}{16d} \sum |V(\widetilde{G}_k)|}$$

$$\lesssim \left[C p^{\frac{d-1}{2}} q^d \right]^{|V|} n^{-\frac{1}{6}|V_1| - \frac{1}{16d}|V_2|} \lesssim \left[C p^{\frac{d-1}{2}} q^d n^{-\frac{1}{16d}} \right]^{|V|}.$$

2.9 The Lower Bound on the Discrepancy Function

We give the proof of Theorem 2.2.4, which is essentially a corollary to the proof of our Main Theorem, Theorem 2.1.3. As such, we will give a somewhat abbreviated proof. Indeed, the analogy between the lower bound on Discrepancy Functions and the Small Ball Inequality is well known to experts.

The proof is by duality. Fix N, and take $2N \leq 2^n < 4N$. It is a familiar fact [2] that for each $|\vec{r}| = n$ we can construct a r function $f_{\vec{r}}$ such that

$$\langle D_N, f_{\vec{r}} \rangle > c > 0 \,,$$

where c depends only on dimension. We use these functions in the construction of the test function, following § 2.4, with this one change. Before, see (2.4.3), we took $I_1, \ldots, I_q$ to be a partition of $\{1, 2, \ldots, n\}$ into q disjoint intervals of equal length. Instead, we take

$$I_t := \{j \in \mathbb{N} \ : \ |j - tn/q| < q/4\} \,. \tag{2.9.1}$$

This is the only change we make in the construction of Ψ^{sd}. It follows that $\|\Psi^{\mathrm{sd}}\|_1 \lesssim 1$.

Recall that $\Psi^{\mathrm{sd}} = \sum_{k=1}^{q} \Psi_k^{\mathrm{sd}}$, see (2.4.4). By construction, we have

$$\langle D_N, \Psi_1^{\mathrm{sd}} \rangle = \sum_{t=1}^{q} \widetilde{\rho} \sum_{\vec{r} \in \mathbb{A}_t} \langle D_N, f_{\vec{r}} \rangle$$

$$\gtrsim q^b n^{(d-1)/2} \simeq n^{\epsilon/4 + (d-1)/2} \,.$$

This is a 'gain over the average case estimate' as one can see by comparison to Theorem 2.2.2. It remains to see that the higher order terms Ψ_k^{sd} contribute smaller terms than the one above.

By construction, Ψ_k^{sd} is itself a sum of r functions $f_{\vec{s}}$ with $|\vec{s}| > n$. Indeed, it follows from the separation in (2.9.1) that we necessarily have

$$n + k\tfrac{n}{2q} \leq |\vec{s}| \leq nd \,.$$

Second, it is a well known fact that $|\langle D_N, f_{\vec{s}} \rangle| < N2^{-|\vec{s}|}$. Third, we fix $\vec{s}$ as above, and set Count$(\vec{s})$ to be the number of distinct ways can we select $\vec{r}_1, \ldots, \vec{r}_k$, all of length n, so that

the product $f_{\vec{r}_1} \cdots f_{\vec{r}_k}$ is an r function of parameter $\vec{s}$. A very crude bound here is sufficient,

$$\mathrm{Count}(\vec{s}) \le |\vec{s}|^{(d-1)k} \, .$$

Thus, we can estimate

$$\langle D_N, \Psi_k^{\mathrm{sd}} \rangle \le \sum_{j \ge n + k\frac{n}{2q}} \left(\sum_{\vec{s} \,:\, |\vec{s}| = j} \mathrm{Count}(\vec{s}) |\langle D_N, f_{\vec{s}} \rangle| \right)$$

$$\lesssim n^{d(k+3)} 2^{-kn/2q} \, .$$

As $q = n^\epsilon$, this is clearly summable in $k \ge 1$ to at most a constant. This completes the proof.

2.10   The Proof of the Smooth Small Ball Inequality

We prove Theorem 2.2.6. There is no loss of generality in assuming that $|\alpha(R)| \le 1$ for all R of volume at least 2^{-n}, since both sides of (2.2.7) are homogeneous and sums have finitely many terms. With φ as in the theorem, set

$$\varphi_{\vec{r}} = \sum_{R \,:\, |R_j| = 2^{-r_j}} \alpha(R) \varphi_R \, .$$

And let $\Phi = \sum_{|\vec{r}|=n} \varphi_{\vec{r}}$. Define the r functions as in (2.3.5). It is the assumption that $c_\varphi = \langle \varphi, h_{[-1/2,1/2]} \rangle \ne 0$, and in fact we will assume that this inner product is positive. Thus,

$$\langle \varphi_{\vec{r}}, f_{\vec{r}} \rangle = c_\varphi 2^{-n} \sum_{R \,:\, |R_j| = 2^{r_j}} |\alpha(R)| \, . \tag{2.10.1}$$

As $\varphi \in C[-1/2, 1/2]$, we have

$$|\langle \varphi, h_J \rangle| \le C_\varphi |J| \tag{2.10.2}$$

for all dyadic intervals J.

It is important to note that

$$|\langle \varphi_{\vec{r}}, f_{\vec{s}} \rangle| \lesssim \begin{cases} 0 & \exists j \,:\, s_j < r_j \\ C_\varphi 2^{-|\vec{r}-\vec{s}|} & \text{otherwise} \end{cases} \tag{2.10.3}$$

40

The first line follows from the fact that φ is supported on $[-1/2, 1/2]$, so that if e. g. $s_1 < r_1$, the fact that φ has mean zero proves this estimate. The second estimate follows from (2.10.2) and the assumption that the coefficients $\alpha(R)$ are at most one in absolute value.

Let us take the intervals I_t in (2.9.1), and let us assume that

$$\sum_{|R|=2^n} |\alpha(R)| \le 4 \sum_{t=1}^{q} \sum_{\vec{r} \in \mathbb{A}_t} \sum_{R : |R_j| = 2^{-r_j}} |\alpha(R)|. \qquad (2.10.4)$$

If this inequality fails, it is an easy matter to redefine the I_t so that the inequality above is true, and adjacent intervals I_t, I_{t+1} are seperated by n/q.

We then follow § 2.4 as before to define our test function Ψ^{sd}. It follows that $\|\Psi^{\mathrm{sd}}\|_1 \lesssim 1$. Using (2.10.4), (2.10.1) and (2.10.3), we have

$$\langle \Phi, \Psi_1^{\mathrm{test}} \rangle \ge c 2^{-n} \widetilde{\rho} \sum_{t=1}^{q} \sum_{\vec{r} \in \mathbb{A}_t} \sum_{R : |R_j| = 2^{r_j}} |\alpha(R)|$$

$$\gtrsim 2^{-n} n^{-(d-1)/2 + \epsilon/4} \sum_{|R| = 2^{-n}} |\alpha(R)|.$$

This is the main term.

It remains to see that the inner products $|\langle \Phi, \Psi_k^{\mathrm{sd}} \rangle|$ are small $k \ge 1$. The details of this calculation are very similar to the corresponding calcluations in the previous section, hence they are omitted.

CHAPTER III

SIGNED SMALL BALL INEQUALITY

3.1 Signed Small Ball Conjecture

In this chapter we discuss a more restrictive formulation of the Small Ball Conjecture that we considered in chapter 2 (see (2.1.1)). This new conjecture still appears to be of interest.

Signed Small Ball Conjecture 3.1.1. *We have the inequality (2.1.1), in the case where the coefficients $\alpha(R) \in \{\pm 1\}$, for $|R| = 2^{-n}$. Namely, under these assumptions on the coefficients $\alpha(R)$ we have the inequality*

$$n^{d/2} \lesssim \Big\| \sum_{|R|=2^{-n}} \alpha(R) h_R \Big\|_\infty .$$

The main result of this chapter is the next Theorem, in which we give an explicit gain over the trivial bound in the Signed Small Ball Conjecture in dimensions $d \geq 3$.

Theorem 3.1.2. *(Bilyk, Lacey, Vagharshakyan, [3]) In dimension $d \geq 3$, for choices of coefficients $\alpha(R) \in \{\pm 1\}$, we have the inequality*

$$n^{\eta(d)} \lesssim \Big\| \sum_{|R|=2^{-n}} \alpha(R) h_R \Big\|_\infty , \quad \text{for all} \quad \eta(d) < \frac{d-1}{2} + \frac{1}{8d} .$$

The main simplification in the current chapter, in comparison to the previous one, lies in the equalities (3.3.2), which allow us to avoid analyzing longer coincidences. The value of η appears to be the optimal one we can get out of this line of reasoning, imputing additional interest to the methods of proof used to improve this estimate.

3.2 Notations and Littlewood-Paley Inequality

Recall the definition of $\vec{r}$-functions, introduced in section 2.3 (see (2.3.1)).

The r-functions that we are interested in this chapter are:

$$f_{\vec{r}} = \sum_{R \in \mathcal{R}_{\vec{r}}} \alpha(R) \, h_R$$

In the following sections we will use the Littlewood-Paley inequalities (2.3.7) and (2.3.8), mentioned in the previous chapter.

3.3   Proof of a Signed Small Ball Inequality

The proof of the main theorem of this chapter - Theorem 3.1.2, is by duality, namely we construct a function Ψ of L^1 norm about one, which is used to provide a lower bound on the L^∞ norm of the sum of Haar functions.

The function Ψ will take the form of a Riesz product, but in order to construct it, we need some definitions first. Fix $0 < \kappa < 1$, with the interesting choices of κ being close to zero. Define relevant parameters by

$$q = \lfloor an^\varepsilon \rfloor, \qquad \varepsilon = \frac{1}{2d} - \kappa, \qquad b = \tfrac{1}{4},$$
$$\widetilde{\rho} = aq^b n^{-(d-1)/2}, \qquad \rho = \sqrt{q}\, n^{-(d-1)/2}.$$

Here a is a small positive constant, we use the notation $b = \frac{1}{4}$ throughout, so as not to obscure those aspects of the argument that that dictate these choices of parameters. $\widetilde{\rho}$ is a 'false' L^2 normalization for the sums we consider, while the larger term ρ is the 'true' L^2 normalization. Our 'gain over the average case estimate' in the Small Ball Conjecture is $q^b \simeq n^{\varepsilon/4} = n^{1/8d - \kappa/4} = n^{\eta(d) - (d-1)/2}$.

Divide the integers $\{1, 2, \ldots, n\}$ into q disjoint increasing intervals of equal length $I_1, \ldots, I_q$, and let $\mathbb{A}_t = \{\vec{r} \in \mathbb{H}_n : r_1 \in I_t\}$. Let

$$F_t = \sum_{\vec{r} \in \mathbb{A}_t} f_{\vec{r}}, \qquad H = \sum_{\vec{r} \in \mathbb{H}_n} f_r = \sum_{t=1}^{q} F_t.$$

The Riesz product is a 'short product.'

$$\Psi = \prod_{t=1}^{q} (1 + \widetilde{\rho} F_t), \qquad \Psi_{\neq j} = \prod_{\substack{t=1 \\ t \neq j}}^{q} (1 + \widetilde{\rho} F_t), \quad 1 \leq j \leq q.$$

43

Note the subtle way that the false L^2 normalization enters into the product. It means that the product is, with high probability, positive. And of course, for a positive function F, we have $\mathbb{E}F = \|F\|_1$, with expectations being typically easier to estimate. This heuristic is made precise below. Notice also that $\mathbb{E}\Psi = 1$.

We need a final bit of notation. Set

$$\Phi_t = \sum_{\substack{\vec{r} \neq \vec{s} \in \mathbb{A}_t \\ r_1 = s_1}} f_{\vec{r}} \cdot f_{\vec{s}}.$$

Note that in this sum, there are $2d - 3$ free parameters among the vectors $\vec{r}$ and $\vec{s}$. That is, the pair of vectors $(\vec{r}, \vec{s})$ are completely specified by there values in $2d - 3$ coordinates.

Our main Lemma is below. Note that in (3.3.1), the assertion is that the $2d - 3$ parameters in the definition of Φ_t behave, with respect to L^p norms, as if they are independent.

Lemma 3.3.1. *We have these estimates.*

$$\|\Psi\|_1 \lesssim 1\,,$$

$$\|\Psi\|_2\,,\ \max_{1 \leq j \leq n} \|\Psi_{\neq j}\|_2 \lesssim e^{a' q^{2b}}\,,$$

$$\|\Phi_t\|_p \lesssim p^{d-1/2} n^{d-3/2} q^{-1/2}\,, \qquad 2 < p < \infty\,.$$

In (3.3.1), the value of a' is a decreasing function of $0 < a < 1$.

The proof of this Lemma is taken up next section. Assuming the Lemma, we proceed as follows. An important simplification in the Signed Small Ball Inequality comes from the equalities

$$\begin{aligned}
\langle F_j, \Psi \rangle &= \langle \sum_{\vec{r} \in \mathbb{A}_j} f_r, \Psi \rangle \\
&= \sum_{\vec{r} \in \mathbb{A}_j} \langle f_r, (1 + \widetilde{\rho} F_j)\Psi_{\neq j} \rangle \\
&= \widetilde{\rho} \sum_{\vec{r} \in \mathbb{A}_j} \langle f_r^2, \Psi_{\neq j} \rangle + \widetilde{\rho} \langle \Phi_j, \Psi_{\neq j} \rangle \\
&= \widetilde{\rho} \sharp \mathbb{A}_j + \widetilde{\rho} \langle \Phi_j, \Psi_{\neq j} \rangle\,.
\end{aligned} \tag{3.3.2}$$

We have used the fact that there has to be a coincidence in the first coordinate in order for the product of r functions to have non-zero integral. The first term in the third line is the

44

'diagonal' term, while the second term arises from different vectors which coincide in the first coordinate. Therefore, we can estimate

$$
\begin{aligned}
\|H\|_\infty &\gtrsim \langle H, \Psi \rangle \\
&= \sum_{j=1}^{q} \widetilde{\langle F_j, \Psi \rangle} \\
&= \widetilde{\rho} \, \sharp \mathbb{H}_n + \sum_{j=1}^{q} \widetilde{\rho} \langle \Phi_j, \Psi_{\neq j} \rangle
\end{aligned}
$$

It is clear that

$$
\widetilde{\rho} \, \sharp \mathbb{H}_n \simeq a^{5/4} n^{\frac{d-1}{2} + \frac{\varepsilon}{4}} = a^{5/4} n^{\eta(d)} \,,
$$

which is our principal estimate. The other term we treat as an error term. Using Hölder's inequality, and (3.3.1) and (3.3.1) we see that

$$
\|\Psi_{\neq j}\|_{(q^{2b})'} \leq \|\Psi_{\neq j}\|_{1}^{(q^{2b}-2)/q^{2b}} \|\Psi_{\neq j}\|_{2}^{2q^{-2b}} \lesssim 1 \,.
$$

Therefore, we can estimate as below, where we use the estimate above and (3.3.1).

$$
\begin{aligned}
\left| \sum_{j=1}^{q} \widetilde{\rho} \langle \Phi_j, \Psi_{\neq j} \rangle \right| &\lesssim \sum_{j=1}^{q} \widetilde{\rho} \|\Phi_j\|_{q^{2b}} \|\Psi_{\neq j}\|_{(q^{2b})'} \\
&\lesssim q \cdot \frac{aq^b}{n^{(d-1)/2}} \cdot q^{2b(d-1/2)} n^{d-3/2} \cdot q^{-1/2} \\
&\simeq aq^{2bd+1/2} n^{(d-2)/2} \ll n^{\eta(d)} \,.
\end{aligned}
$$

This term will be smaller than the term in (3.3). The proof of our main result is complete, modulo the proof of Lemma 3.3.1.

3.4 Analysis of Coincidences

Following the language of J. Beck [1], a *coincidence* occurs if we have two vectors $\vec{r} \neq \vec{s}$ with e.g. $r_1 = s_1$, precisely the condition that we imposed in the definition of Φ_t, (3.3). He observed that sums over products of r functions in which there are coincidences obey favorable L^2 estimates. We refer to (extensions of) this observation as the *Beck Gain.*

The Simplest Instance of the Beck Gain 3.4.1. *We have the estimates below, valid for an absolute implied constant that is only a function of dimension $d \geq 3$.*

$$\sup_{1 \leq j \leq n} \|\Phi_j\|_p \lesssim p^{d-1/2} \cdot n^{d-3/2} q^{-1/2}, \qquad 1 \leq p < \infty.$$

This Lemma, in dimension $d = 3$ appears in [4]. The proof in higher dimensions, which was given in the previous chapter, is inductive. We omit the proof here as it is rather lengthy and refer the reader to the previous chapter for details. Strictly speaking, the estimate of the previuos chapter does not contain Lemma 3.4.1, as it does not include $q^{-1/2}$. However, this can be easily fixed in the proof due to the fact that the value of the first coordinate can be chosen in n/q ways rather than n. We also emphasize that the estimate above may admit an improvement, in that the power of p is perhaps too large by a single power.

Conjecture 3.4.2. *We have the estimates below, valid for an absolute implied constant that is only a function of dimension $d \geq 3$.*

$$\sup_{1 \leq j \leq n} \|\Phi_j\|_p \lesssim (pn)^{d-3/2} q^{-1/2}, \qquad 1 \leq p < \infty.$$

With this conjecture we could prove our main theorem for all $\eta(d) < \frac{d-1}{2} + \frac{1}{8d-8}$.

The Beck Gain Lemma 3.4.1 has several important consequences. Theorem 2.3.8 implies an exponential estimate for sums of r functions. However, with the Beck Gain at hand, we can derive a subgaussian estimate for such sums, for moderate deviations.

Theorem 3.4.3. *Using the notation of (3.3) and (3.3), we have this estimate, valid for all $1 \leq t \leq q$.*

$$\|\rho F_t\|_p \lesssim \sqrt{p}, \qquad 1 \leq p \leq c \left(\frac{n}{q}\right)^{\frac{1}{2d-1}}.$$

As a consequence, we have the distributional estimate

$$\mathbb{P}(|\rho F_t| > x) \lesssim \exp(-cx^2), \qquad x < c \left(\frac{n}{q}\right)^{\frac{1}{4d-2}}.$$

Here $0 < c < 1$ is an absolute constant.

Proof. Recall that

$$F_t = \sum_{\vec{r} \in \mathbb{A}_t} f_{\vec{r}} \,.$$

where $\mathbb{A}_t = \{ \vec{r} \in \mathbb{H}_n \ : \ r_1 \in I_t \}$, and I_t in an interval of integers of length n/q, so that $\sharp \mathbb{A}_t \simeq n^2/q \simeq \rho^{-2}$.

Apply the Littlewood-Paley inequality in the first coordinate. This results in the estimate

$$\| \rho F_t \|_p \lesssim \sqrt{p} \left\| \left[\sum_{s \in I_j} \Big| \rho \sum_{\vec{r} : r_1 = s} f_{\vec{r}} \Big|^2 \right]^{1/2} \right\|_p$$

$$\lesssim \sqrt{p} \| 1 + \rho^2 \Phi_t \|_{p/2}^{1/2}$$

$$\lesssim \sqrt{p} \Big\{ 1 + \| \rho^2 \Phi_t \|_{p/2}^{1/2} \Big\} \,.$$

Here, it is important to use the constants in the Littlewood-Paley inequalities that give the correct order of growth of $\sqrt{p}$. Of course the terms Φ_t are controlled by the estimate in (3.4.1). In particular, we have

$$\| \rho^2 \Phi_t \|_p \lesssim \frac{q}{n^{d-1}} p^{d-1/2} n^{d-3/2} q^{-1/2} \lesssim p^{d-1/2} n^{-1/2} q^{1/2} \,.$$

Hence (3.4.3) follows.

The second distributional inequality is a well known consequence of the norm inequality. Namely, one has the inequality below, valid for all x:

$$\mathbb{P}(\rho F_t > x) \leq C^p p^{p/2} x^{-p} \,, \qquad 1 \leq p \leq c \left(\frac{n}{q} \right)^{\frac{1}{2d-1}} \,.$$

If x is as in (3.4.3), we can take $p \simeq x^2$ to prove the claimed exponential squared bound. $\qquad \square$

Proof of (3.3.1). Observe that

$$\mathbb{P}(\Psi < 0) \lesssim q \exp(ca^{-2} q^{1-2b}) \,.$$

Indeed, using (3.4.3), we have

$$\mathbb{P}(\Psi < 0) \le \sum_{t=1}^{q} \mathbb{P}(\widetilde{\rho}\, F_t < -1)$$

$$= \sum_{t=1}^{q} \mathbb{P}(\rho F_t < -a^{-1} q^{1/2-b})$$

$$\lesssim q \exp(ca^{-2} q^{1-2b}).$$

Note that to be able to use (3.4.3) we need to have $a^{-1} q^{1/2-b} \le c \left(\frac{n}{q}\right)^{\frac{1}{2d-1}}$, which leads to $\varepsilon < \frac{1}{2d}$. Then, assuming (3.3.1), we have

$$\|\Psi\|_1 = \mathbb{E}\Psi - 2\mathbb{E}\Psi \mathbf{1}_{\Psi<0}$$

$$\le 1 + 2\mathbb{P}(\Psi < 0)^{1/2} \|\Psi\|_2$$

$$\lesssim 1 + \exp(-a^{-2} q^{1-2b}/2 + aq^{2b}).$$

For sufficiently small $0 < a < 1$, the proof is finished. Note that this last step forces $b = 1/4$

on us.

$\square$

Proof of (3.3.1). The supremum over j will be an immediate consequence of the proof

below, and so we don't address it specifically.

Let us give the initial, essential observation. We expand

$$\mathbb{E} \prod_{j=1}^{q}(1 + \widetilde{\rho}F_j)^2 = \mathbb{E} \prod_{j=1}^{q}(1 + 2\widetilde{\rho}F_j + (\widetilde{\rho}F_j)^2).$$

Hold the x_2 and x_3 coordinates fixed, and let $\mathcal{F}$ be the sigma field generated by $F_1, \ldots, F_{q-1}$.

We have

$$\mathbb{E}\big(1 + 2\widetilde{\rho}F_q + (\widetilde{\rho}F_q)^2 \,\big|\, \mathcal{F}\big) = 1 + \mathbb{E}\big((\widetilde{\rho}F_q)^2 \,\big|\, \mathcal{F}\big)$$

$$= 1 + a^2 q^{2b-1} + \widetilde{\rho}^2 \Phi_q,$$

where Φ_q is defined in (3.3). Then, we see that

$$\mathbb{E}\prod_{v=1}^{q}(1+2\widetilde{\rho}F_t+(\widetilde{\rho}F_t)^2) = \mathbb{E}\Big\{\prod_{v=1}^{q-1}(1+2\widetilde{\rho}F_t+(\widetilde{\rho}F_t)^2)\ \times\ \mathbb{E}\big(1+2\widetilde{\rho}F_t+(\widetilde{\rho}F_t)^2\,\big|\,\mathcal{F}\big)\Big\}$$

$$\leq (1+a^2q^{2b-1})\mathbb{E}\prod_{v=1}^{q-1}(1+2\widetilde{\rho}F_t+(\widetilde{\rho}F_t)^2)$$

$$+\ \mathbb{E}|\widetilde{\rho}^2\Phi_q|\cdot\prod_{v=1}^{q-1}(1+2\widetilde{\rho}F_t+(\widetilde{\rho}F_t)^2)$$

This is the main observation: one should induct on (3.4), while treating the term in (3.4) as an error, as the 'Beck Gain' estimate (3.4.1) applies to it.

Let us set up notation to implement this line of approach. Set

$$N(V;r) = \Big\|\prod_{v=1}^{V}(1+\widetilde{\rho}F_v)\Big\|_r, \qquad V = 1,\dots,q.$$

We will obtain a very crude estimate for these numbers for $r = 4$. Fortunately, this is relatively easy for us to obtain. Namely, q is small enough that we can use the inequalities (3.4.3) to see that

$$N(V;4) \leq \prod_{v=1}^{V}\|1+\widetilde{\rho}F_v\|_{4V}$$

$$\leq (1+Cq^b)^V$$

$$\leq (Cq)^q.$$

For a large choice of $\tau > 1$, which is a function of the choice of $\kappa > 0$ in (3.3), we have the estimate below from Hölder's inequality

$$N(V;2(1-1/\tau q)^{-1}) \leq N(V;2)^{1-2/\tau q}\cdot N(V;4)^{2/\tau q}.$$

We see that (3.4), (3.4) and (3.4) give us the inequality

$$N(V+1;2)^2 \leq (1+a^2q^{2b-1})N(V;2)^2 + C\cdot N(V;2(1-1/\tau q)^{-1})^2\cdot\|\widetilde{\rho}^2\Phi_V\|_{\tau q}$$

$$\leq (1+a^2q^{2b-1})N(V;2)^2 + CN(V;2)^{2-4/\tau q}\cdot N(V;4)^{4/\tau q}\|\widetilde{\rho}^2\Phi_V\|_{\tau q}$$

$$\leq (1+a^2q^{2b-1})N(V;2)^2 + C_\tau q^{d-1/2+4/\tau}n^{-1/2}N(V;2)^{2-2/\tau q}.$$

In the last line we have used the the inequality (3.4.1) and the constant C_τ is only a function of $\tau > 1$, which is fixed.

Of course we only apply this as long as $N(V; 2) \geq 1$. Assuming this is true for all $V \geq 1$, we see that

$$N(V+1; 2)^2 \leq (1 + a^2 q^{2b-1} + C_\tau q^{d-1/2+4/\tau} n^{-1/2}) N(V; 2)^2 .$$

And so, by induction,

$$N(q; 2) \lesssim (1 + a^2 q^{2b-1} + C_\tau q^{d-1/2+4/\tau} n^{-1/2})^{q/2} \lesssim e^{2aq^{2b}} .$$

Here, the last inequality will be true for large n, provided τ is much bigger than $1/\kappa$. Indeed, we need

$$a^2 q^{2b-1} \geq C_\tau q^{d-1/2+4/\tau} n^{-1/2}$$

Or equivalently,

$$cn^{1/2} \geq q^{d+4/\tau} .$$

Comparing to the definition of q in (3.3), we see that the proof is finished.

$\square$

CHAPTER IV

ORLICZ AND BMO NORMS OF DISCREPANCY IN TWO DIMENSIONS

4.1 Formulation of the Result

Recall the definition of Discrepancy Function (2.2.1) from chapter 2. In the present chapter, we are primarily interested in the precise behavior of estimates for the Discrepancy Function near the L^∞ endpoint, phrased in terms of exponential Orlicz classes. We restrict our attention to the two-dimensional case. According to (2.2.1), the definition of Discrepancy Function in two-dimensions would go as follows:

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 various senses, must necessarily increase with N.

As noted in chapter 2, the principal result in this direction is due to Roth [40]. Let's cite it again:

K. Roth's Theorem 4.1.1. *In all dimensions $d \geq 2$, we have the following estimate*

$$\|D_N\|_2 \gtrsim (\log N)^{(d-1)/2} \tag{4.1.2}$$

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

The same bound holds for the L^p norm, for $1 < p < \infty$, [42], and is known to be sharp as to the order of magnitude, see [10] and [2] for a history of this subject (for the case $d = 2$, see Corollary 4.1.6 below). The endpoint cases of $p = 1$ and $p = \infty$ are much harder.

We concentrate on the case of $p = \infty$ in this chapter, just in dimension $d = 2$. The case $d \geq 3$ was discussed in chapter 2. As for information about the case of $p = 1$, see [17],[24]. As it has been shown in the fundamental theorem of W. Schmidt [41], in dimension $d = 2$, the lower bound on the L^∞ norm of the Discrepancy function is substantially greater than the L^p estimate (4.1.2):

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

$$\|D_N\|_\infty \gtrsim \log N \,. \tag{4.1.3}$$

This theorem is also sharp: one particular example is the famous van der Corput set [49] – a detailed discussion is contained in section 4.3. In this chapter, we give an interpolant between the results of Roth and Schmidt, which is measured in the scale of exponential Orlicz classes.

Theorem 4.1.4. *(Bilyk, Lacey, Parissis, Vagharshakyan, [5]) 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 [17]. We give details below and also remark that this estimate in the context of the Small Ball Inequality [44],[46] is known [14]. In addition, we demonstrate that the previous theorem is sharp.

Theorem 4.1.5. *(Bilyk, Lacey, Parissis, Vagharshakyan, [5]) For all N, there is a choice of $\mathcal{A}_N$, specifically the digit-scrambled van der Corput set (see Definition 4.3.5), for which we have*

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

In view of Proposition 4.2.2, 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.

Corollary 4.1.6. *For every $1 \leq p < \infty$, the set $\mathcal{A}_N$ from Theorem 4.1.5 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.

Theorem 4.1.7. *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 4.2.11), introduced in [8].*

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

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

Theorem 4.1.8. *(Bilyk, Lacey, Parissis, Vagharshakyan, [5]) 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 of this chapter. These results are drawn from the theory of Irregularities of Distribution,

Harmonic Analysis, Probability Theory and other subjects. In section 4.3 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 sections 4.5 and 4.6.

The results of this chapter of thesis 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. Higher dimensional versions of Schmidt's theorem were discussed in chapter 2.

4.2   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 $A \leq KB$. Thus $A \lesssim 1$ means that A is bounded by an absolute constant. And if $A \lesssim B \lesssim A$, we write $A \simeq B$.

4.2.1   Martingale 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 $\mathcal{D} = \{[j2^{-k}, (j+1)2^{-k}) : j, k \in \mathbb{N}, 0 \leq j < 2^k\}$. 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

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

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

$$\mathrm{S}(f) = \left[\sum_{n=0}^{\infty} |d_n|^2\right]^{1/2} \,.$$

We have the following extension of the Khintchine inequalities.

Theorem 4.2.1. *The inequalities below hold, for some absolute choice of constant $C > 0$.*

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

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

For real-valued martingales, this was observed by [7]. The extension to Hilbert space valued martingales is useful for us and is proved in [15]. The best constants in these inequalities are known for $p \ge 3$ [50].

4.2.2 Orlicz Spaces

For background on Orlicz Spaces, we refer the reader to [29]. 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

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

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

We are interested in, for instance, $\psi(x) = e^{x^2} - 1$, 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 the main Theorems of this chapter: 4.1.4 and 4.1.5. 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 4.1.4 can be indeed viewed as interpolation between the estimates of Roth (4.1.2) and Schmidt (4.1.3). The following useful proposition is well-known and follows from elementary methods.

---

[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.

Proposition 4.2.2. *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.

Proposition 4.2.3. *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.*

Proposition 4.2.4. *Let E be a measurable subset of a probability set. We have*

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

4.2.3   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

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

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

Chang-Wilson-Wolff Inequality   4.2.5. *For all Hilbert space valued martingales, we have*

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

Indeed, we have

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

---

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

Taking $p \to \infty$, and using Proposition 4.2.2, we deduce the inequality above.

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 3. 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

$$\mathrm{S}(f)^2 = \sum_{R \in \mathcal{D}^2} \frac{\langle f, h_R \rangle^2}{|R|^2} \mathbf{1}_R, \qquad \text{for} \quad f(x) = \sum_{R \in \mathcal{D}^2} \frac{\langle f, h_R \rangle}{|R|} h_R(x). \qquad (4.2.6)$$

Jill Pipher [38] observed the following extension of the Chang-Wilson-Wolff inequality.

Two Parameter Chang-Wilson-Wolff Inequality 4.2.7. *For functions f in the plane as in* (4.2.6) *we have*

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

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 [38],[15],[4]. In fact, we will need the following variant.

Theorem 4.2.8. *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 4.1.5 to the case $\alpha = 2$.

Proposition 4.2.9. *Suppose that for $A \geq 1$, we have*

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

It follows that

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

4.2.4   Bounded Mean Oscillation

We recall facts about dyadic BMO spaces, see [9],[8].

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 ϵ_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 $L^2([0,1]^2)$.

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

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.

Definition 4.2.10. Define the space BMO_1 to be those square integrable functions f in the span of $\{h^{0,1}_{I \times [0,1]} : I \in \mathcal{D}\}$ which satisfy

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

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

Definition 4.2.11. Dyadic Chang-Fefferman $\mathrm{BMO}_{1,2}$ is defined to be those square integrable functions f in the linear span of $\{h_R : R \in \mathcal{D}^2\}$, for which we have

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

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 $\varphi \in \mathrm{BMO}_{1,2}$.*

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

$$\|\varphi\|_{\exp(\sqrt{L})} \lesssim \|\varphi\|_{\mathrm{BMO}_{1,2}} \tag{4.2.12}$$

Note that in the second inequality, (4.2.12), 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.

4.2.5 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 [40],[43],[2].

In consistency with the definition of $\vec{r}$-function in any dimension (see (2.3.1)), 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 $\mathcal{R}_{\vec{r}} = \{R = R_1 \times R_2 \,:\, R \in \mathcal{D}^2 \,, R \subset [0,1]^2 \,, |R_t| = 2^{-r_t} \,, t = 1, 2\}$. 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 $\mathbb{H}_n^2 = \{\vec{r} \in \{0, 1, \ldots, n\}^2 \,:\, |\vec{r}| = n\}$, 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 $\{f_{\vec{r}} \,:\, r \in \mathbb{H}_n^2\}$ as hyperbolic r functions. The next four Propositions are standard.

Proposition 4.2.13. *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

$$\langle x_1 \cdot x_2 \,, \, h_R^{0,0}(x_1, x_2) \rangle = 4^{-2} |R|^2 \,. \tag{4.2.14}$$

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

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

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

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

Hence, to complete the proof, we can estimate

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

$\square$

Proposition 4.2.15. *Let $f_{\vec{s}}$ be any r function with $|\vec{s}| > n$. We have*

$$|\langle D_N, f_{\vec{s}} \rangle| \lesssim N 2^{-|\vec{s}|} \,.$$

Proof. This is a brute force proof. Consider the linear part of the Discrepancy function. By (4.2.5), we have

$$|\langle L_N, f_{\vec{s}} \rangle| \lesssim N 2^{-|\vec{s}|} \,,$$

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. $\qquad\square$

Proposition 4.2.16. *In dimension $d = 2$ the following holds. Fix a collection of r functions $\{f_{\vec{r}} \;:\; \vec{r} \in \mathbb{H}_n^2\}$. 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 $Count(\vec{s}; v)$ 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*

$$Count(\vec{s}; v) = \begin{array}{c} |\vec{s}| - n - 1 \\ v - 2 \end{array}\,.$$

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. $\qquad\square$

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

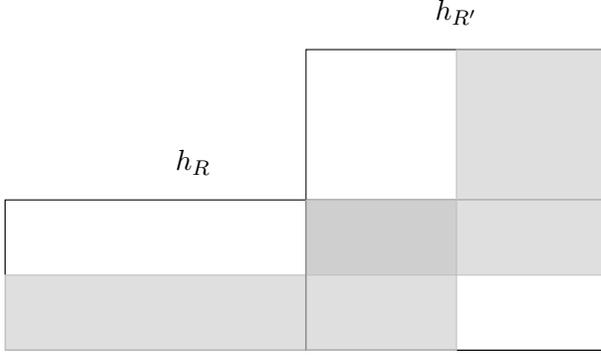$$h_R \cdot h_{R'} = \pm h_{R \cap R'}.$$

Figure 3: Two Haar functions.

This rule is illustrated in Figure 3 and can be generalized as follows.

Proposition 4.2.17. *In dimension $d = 2$ the following holds. Let $\vec{r}_1, \ldots, \vec{r}_k$ 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 $k \geq 2$, the product has index larger than n.*

4.3   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 [30]*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 4.1.5 and 4.1.8, although strictly speaking, Theorem 4.1.8 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.

Definition 4.3.1. For $x \in [0, 1)$ define $\mathrm{d}_i(x)$ to be the i'th digit in the binary expansion of x, that is

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

Definition 4.3.2. 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 $d_i(x) = x_i$, we have $\mathrm{rev}_n(0.x_1x_2...x_n) = 0.x_n...x_2x_1$.

Definition 4.3.3. Let $x, \sigma \in [0, 1)$ where σ has n binary digits. We define the number $x \oplus \sigma$ as

$$d_i(x \oplus \sigma) = d_i(x) + d_i(\sigma) \mod 2,$$

i.e. the ith digit of x changes if $d_i(\sigma) = 1$ and stays the same if $d_i(\sigma) = 0$. In the literature this operation is called *digit scrambling* or *digital shift*.

Remark 4.3.4. 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.

Definition 4.3.5. 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

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

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 $\mathcal{V}_{n,\sigma} = \{(\mathrm{rev}_n(\tau/2^n \oplus \sigma'), \tau/2^n) + (2^{-n-1}, 2^{-n-1}) : \quad \tau = 0, 1, \ldots, 2^n - 1\}$ with $\sigma' = \mathrm{rev}_n(\sigma)$.

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

$$\mathcal{V}_n = \{(0.x_1x_2...x_n1, 0.x_n...x_2x_11) : x_i = 0, 1\}$$

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 $(n+1)^{\mathrm{st}}$ place, and ensures that the average value of each coordinate is $\frac{1}{2}$:

$$2^{-n} \sum_{(x,y) \in \mathcal{V}_{n,\sigma}} x = 2^{-n} \sum_{(x,y) \in \mathcal{V}_{n,\sigma}} y = \frac{1}{2}. \tag{4.3.6}$$

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.

Proposition 4.3.7. *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

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

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

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

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

The proposition is now a simple consequence of the structure of the van der Corput set. $\qquad \square$

Some remarks are in order:

Remarks 4.3.8.

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 4.3.7 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 4.3.7 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 4.3.7 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 [30], that

$$\|D_N(\mathcal{V}_{n,\sigma})\|_\infty \lesssim \log N. \tag{4.3.9}$$

This fact is independent of the digit scrambling σ and holds in particular for the standard van der Corput set $\mathcal{V}_n$ ([49], [40]). In view of Schmidt's Theorem (4.1.3) 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 [11] 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 [18]. We include a proof of this Lemma for the sake of completeness.

Lemma 4.3.10. *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

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

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

Using (4.3.6) we get

$$I = -\frac{N}{4} + \frac{1}{2} - \frac{1}{4N} + \sum_{\tau=0}^{N-1} \frac{\tau}{N} \cdot \mathrm{rev}_n \left(\frac{\tau}{N} \oplus \sigma \right). \tag{4.3.11}$$

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

$$
\begin{aligned}
\sum_{\tau=0}^{N-1} \frac{\tau}{N} \cdot \operatorname{rev}_n\left(\frac{\tau}{N} \oplus \sigma\right) &= \sum_{\tau=0}^{N-1} \sum_{k=1}^{n} \sum_{l=1}^{n} \frac{\mathrm{d}_k\left(\frac{\tau}{N}\right) \mathrm{d}_l\left(\operatorname{rev}_n\left(\frac{\tau}{N} \oplus \sigma\right)\right)}{2^{k+l}} \\
&= \sum_{\tau=0}^{N-1} \sum_{k=1}^{n} \sum_{l=1}^{n} \frac{\mathrm{d}_k\left(\frac{\tau}{N}\right) \mathrm{d}_{n+1-l}\left(\frac{\tau}{N} \oplus \sigma\right)}{2^{k+l}} \\
&= \sum_{k=1}^{n} \sum_{l=1}^{n} \frac{1}{2^{k+l}} \sum_{\tau=0}^{N-1} \mathrm{d}_k\left(\frac{\tau}{N}\right) \mathrm{d}_{n+1-l}\left(\frac{\tau}{N} \oplus \sigma\right).
\end{aligned}
\tag{4.3.12}
$$

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

$$
\sum_{\tau=0}^{N-1} \mathrm{d}_s\left(\frac{\tau}{N}\right) \mathrm{d}_t\left(\frac{\tau}{N} \oplus \sigma\right) =
\begin{cases}
\frac{N}{2}\left(1 - \mathrm{d}_s(\sigma)\right), & s = t \\[2mm]
\frac{N}{4}, & s \neq t.
\end{cases}
\tag{4.3.13}
$$

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 (4.3.13) and (4.3.12) 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 (4.3.11), completes the proof. $\qquad \square$

Remark. We should point out that in [20] 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.

4.4 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 4.3, 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 4.1.5 and 4.1.8.

Lemma 4.4.1. *For any dyadic rectangle $R \in \mathcal{D}^2$ we have*

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

67

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

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

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 4.

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

$$\langle \mathbf{1}_{[\vec{p},\vec{1})}, h_R \rangle = \begin{cases} |R|\phi(2^k p_x)\phi(2^l p_y), & p \in R, \\ 0, & \text{otherwise.} \end{cases} \tag{4.4.2}$$

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

$$\phi(x) + \phi\left(x \oplus \frac{1}{2}\right) = \frac{1}{2}. \tag{4.4.3}$$

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

$$|\phi(y) - \phi(x)| \le |\{y\} - \{x\}|. \tag{4.4.4}$$

Proof of Lemma 4.4.1 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 (4.4.2) implies

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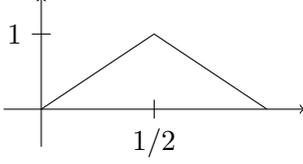
Figure 4: The graph of the function ϕ.

Summing up the estimates for the linear and the counting part completes the proof. $\qquad \square$

Proof of Lemma 4.4.1 when $|R| \geq \frac{4}{N}$. The proof of the case $|R| \geq \frac{4}{N}$ 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 $|R| \geq \frac{4}{N}$ implies that $n - (k+l) \geq 2$. In other words, there are at least 4 points in the set $R \cap \mathcal{V}_{n,\sigma}$ according to Proposition 4.3.7 and Remark 4.3.8. 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 form $x = 0.x_1 x_2 \ldots x_n 1$, where $x_i = \mathrm{d}_i(x)$, for $i = 1, 2, \ldots, n$. The first k and the last l binary digits of x are determined by the fact that $x \in R$ (Proposition 4.3.7). That leaves us with at least 2 'free' digits for x

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

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:

$$
\begin{aligned}
0.x_1 \ldots x_k \; 0 \; x_{k+2} \ldots, x_{n-l-1} \; 0 \; x_{n-l+1} \ldots x_n 1, \\
0.x_1 \ldots x_k \; 0 \; x_{k+2} \ldots, x_{n-l-1} \; 1 \; x_{n-l+1} \ldots x_n 1, \\
0.x_1 \ldots x_k \; 1 \; x_{k+2} \ldots, x_{n-l-1} \; 0 \; x_{n-l+1} \ldots x_n 1, \\
0.x_1 \ldots x_k \; 1 \; x_{k+2} \ldots, x_{n-l-1} \; 1 \; x_{n-l+1} \ldots x_n 1.
\end{aligned}
\tag{Q}
$$

There are exactly $2^{n-(k+l)-2} = \frac{N|R|}{4}$ such quadruples. Let's index the quadruples Q
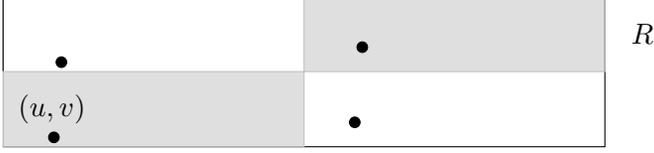
69

Figure 5: The quadruple Q.

arbitrarily as Q_r, $r = 1, 2, \ldots, \frac{N|R|}{4}$. Observe that we can write

$$\langle D_N, h_R \rangle = \sum_{p \in \mathcal{V}_{n,\sigma} \cap R} \langle \mathbf{1}_{[\vec{p},\vec{1})}, h_R \rangle - \frac{N|R|^2}{16} = \sum_{r=1}^{\frac{N|R|}{4}} \left(\sum_{p \in Q_r} \langle \mathbf{1}_{[\vec{p},\vec{1})}, h_R \rangle - \frac{|R|}{4} \right). \quad (4.4.5)$$

The following Proposition exploits large cancellation within these quadruples.

Proposition 4.4.6.

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

Let assume Proposition 4.4.6 for a moment in order to complete the proof of Lemma 4.4.1. Indeed, Proposition 4.4.6 together with equation (4.4.5) 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 4.4.6. $\qquad \square$

Proof of Proposition 4.4.6. 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:

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

See also Figure 5.

We invoke equation (4.4.2) to write

$$\sum_{p \in Q} \langle \mathbf{1}_{[\vec{p},\vec{1})}, h_R \rangle - \frac{|R|}{4} = |R| \left(\sum_{p \in Q} \phi(2^k p_x) \phi(2^l p_y) - \frac{1}{4} \right) =: |R|(\Sigma - \frac{1}{4}). \quad (4.4.7)$$

We have

$$\Sigma = \phi(2^k u)\phi(2^l v)$$

$$+ \phi(2^k u \oplus \frac{1}{2})\phi(2^l(v \oplus 2^{-n+k}))$$

$$+ \phi(2^k(u \oplus 2^{-n+l}))\phi(2^l v \oplus \frac{1}{2})$$

$$+ \phi(2^k u \oplus 2^k \cdot 2^{-n+l} \oplus \frac{1}{2})\phi(2^l v \oplus 2^l \cdot 2^{-n+k} \oplus \frac{1}{2}).$$

Using equation (4.4.3) we get

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

Finally, using the fact the the function ϕ is Lipschitz (4.4.4) we have

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

This estimate together with equation (4.4.7) completes the proof. $\qquad\square$

Lemma 4.4.1 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 4.3.10 of the previous section.

Lemma 4.4.8. *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 4.4.1. 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 (4.4.2) which in this case becomes

$$\langle \mathbf{1}_{[\vec{p},\vec{1})}, h_I\rangle = \begin{cases} |I|\phi(2^k p_x)(1 - p_y), & p_x \in I, \\ \\ 0, & \text{otherwise.} \end{cases} \tag{4.4.9}$$

71

As in the proof of Lemma 4.4.1, we need to consider separately the case of small volume and large volume rectangles. The small volume case here is $|I| \leq \frac{2}{N}$. Note that in this case there are at most $2^{n-k} \leq 2$ points of the van der Corput set whose x coordinate lies in I. Using equation (4.4.9) we trivially get the desired estimate as in the proof of the corresponding case of Lemma 4.4.1.

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:

$$0.x_1 \ldots x_k \ 0 \ x_{k+2} \ldots x_n 1,$$
$$0.x_1 \ldots x_k \ 1 \ x_{k+2} \ldots x_n 1. \tag{P}$$

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, $r = 1, 2, \ldots, 2^{n-k-1}$. 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 (4.4.9) to write

$$\sum_{p \in P} \langle \mathbf{1}_{[\vec{p}, \vec{1})}, h_I \rangle = |I| \ \phi(2^k u)(1 - v) + |I| \ \phi(2^k(u \oplus 2^{-k-1}))(1 - v \oplus 2^{-n+k})$$

$$= |I| \ \left[\phi(2^k u) + \phi(2^k u \oplus 2^{-1}) \right] (1 - v)$$

$$+ |I| \ \phi(2^k u \oplus 2^{-1}) \ (v - v \oplus 2^{-n+k})$$

$$= \frac{1}{2} |I|(1 - v) + |I| \ \phi(2^k u \oplus 2^{-1}) \ (v - v \oplus 2^{-n+k}).$$

where in the last equality we have used (4.4.3). 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

$$\langle D_N, h_I \rangle = |I| \left\{ \frac{1}{2} \sum_{r=1}^{2^{n-k-1}} (1 - v_r) - 2^{n-k-3} + \mathcal{O}(1) \right\}, \tag{4.4.10}$$

72

where v_r are y-coordinates of the form

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

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:

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

Substituting in (4.4.10) we get

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

which completes the proof. $\qquad\square$

4.5 BMO Estimates for the Discrepancy Function

This section is devoted to the proofs of Theorems 4.1.7 and 4.1.8. 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 $\{ h_R : R \in \mathcal{D}^2 \}$, 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 4.1.7 which is essentially just a repetition of the argument used in Proposition 4.2.13.

Proof of Theorem 4.1.7. We fix a distribution $\mathcal{A}_N$ of N points in the unit square and take n such that $2N < 2^n \leq 4N$. For the special choice of $U = [0,1]^2$ we have

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

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,

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

73

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 4.2.13. Thus we get

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

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

We proceed with the proof of the upper bound in Theorem 4.1.8. Our extremal set of cardinality $N = 2^n$ will be $\mathcal{V}_{n,\sigma}$ for arbitrary $\sigma \in [0, 1)$, as defined in Definition 4.3.5. We will just write D_N for the Discrepancy function of the digit-scrambled van der Corput set.

Proof of Theorem 4.1.8. 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

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

$$\lesssim N^{-2} |U|^{-1} \sum_{k=0}^{n} 2^k \sum_{\vec{r} \in \mathbb{H}_k} \sum_{R \in \mathcal{R}_{\vec{r}}} 1,$$

where we have used the estimate $\langle D_N, h_R \rangle \lesssim \frac{1}{N}$ of Proposition 4.4.1. 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 (4.2.14) to get $\langle L_N, h_R \rangle = 4^{-2} N |R|^2$. So we have

$$|U|^{-1} \sum_{|R| < 2^{-n}} \frac{\langle L_N, h_R \rangle^2}{|R|} = |U|^{-1} \sum_{k=n+1}^{\infty} \sum_{\vec{r} \in \mathbb{H}_k} \sum_{R \in \mathcal{R}_{\vec{r}}} \frac{\langle L_N, h_R \rangle^2}{|R|}$$

$$\simeq N^2 |U|^{-1} \sum_{k=n+1}^{\infty} \sum_{\vec{r} \in \mathbb{H}_k} (2^{-k})^3 \sum_{R \in \mathcal{R}_{\vec{r}}} 1.$$

74

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

$$\sum_{R \in \mathcal{R}} |R| \lesssim n|U|. \tag{4.5.1}$$

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 (4.5.1).

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:

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

75

Thus, by (4.5.1)

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

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

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

$\square$

4.6 Orlicz Norm Estimates for the Discrepancy Function.

This section is constructed according to the following table:

Table 1: Orlicz Norm Estimates

Subsection	Topic
4.6.1	Lower Bound
4.6.2	Upper Bound for $N = 2^n$
4.6.3	Upper Bound in the General Case

4.6.1 Lower Bound

The proof is by way of duality and is very similar to Halász's proof [17] of Schmidt's Theorem, see (4.1.3). Fix the point distribution $\mathcal{A}_N \subset [0,1]^2$. Set $2N < 2^n \le 4N$, so that $n \simeq \log N$. Proposition 4.2.13 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' 4.2.17 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 4.2.4,

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

The fact that $\langle D_N, \widetilde{\Psi} \rangle \gtrsim n$ is well-known [17], [30]. In fact, if we expand

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

$$\Psi_k = \sum_{\{\vec{r}_1, \ldots, \vec{r}_k\} \subset \mathbb{G}_n^2} \prod_{\ell=1}^{k} f_{\vec{r}_\ell} \, ,$$

then, using the 'product rule' 4.2.17, 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 4.2.15 and 4.2.16, 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 4.1.4 holds.

4.6.2 Upper Bound in the Case $N = 2^n$.

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 4.3.10 tells us that we should choose $\mathcal{V}_{n,\sigma}$ with half the digits 'scrambled', i.e. $\sum_{i=1}^{n} d_i(\sigma) = \lfloor n/2 \rfloor$ – 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):

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

$$= \sum_{R : |R| > 2^{-n}} \frac{\langle D_N, h_R \rangle}{|R|} h_R + \sum_{R : |R| \leq 2^{-n}} \frac{\langle C_N, h_R \rangle}{|R|} h_R - \sum_{R : |R| \leq 2^{-n}} \frac{\langle L_N, h_R \rangle}{|R|} h_R \quad (4.6.1)$$

$$+ \sum_{R = I \times [0,1]} \frac{\langle D_N, h_R^{0,1} \rangle}{|R|} h_R^{0,1} + \sum_{R = [0,1] \times I} \frac{\langle D_N, h_R^{1,0} \rangle}{|R|} h_R^{1,0} \quad (4.6.2)$$

For the first sum in the expansion (4.6.1) above we have:

$$\Big\| \sum_{R:|R|>2^{-n}} \frac{\langle D_N, h_R\rangle}{|R|} h_R \Big\|_{\exp(L^2)} \le \sum_{k=0}^{n-1} \Big\| \sum_{R:|R|=2^{-k}} \frac{\langle D_N, h_R\rangle}{|R|} h_R \Big\|_{\exp(L^2)}$$

$$\lesssim \sum_{k=0}^{n-1} \Big\| \Big(\sum_{R:|R|=2^{-k}} \frac{\langle D_N, h_R\rangle^2}{|R|^2} \mathbf{1}_R \Big)^{\frac{1}{2}} \Big\|_{\infty}$$

$$\lesssim \sum_{k=0}^{n-1} \frac{1}{N} \cdot \sqrt{k+1} \cdot 2^k \approx \sqrt{n},$$

where we have used the hyperbolic version of the Chang-Wilson-Wolff inequality (Theorem 4.2.8), the estimate of the Haar coefficients of D_N (Lemma 4.4.1), and the fact that each point in $[0,1]^2$ lives in $k+1$ dyadic rectangles of volume 2^{-k}.

The last sum in (4.6.1) is easy to estimate. Since $\langle L_N, h_R\rangle = 4^{-d} N|R|^2$, we have:

$$\Big\| \sum_{R:|R|\le2^{-n}} \frac{\langle L_N, h_R\rangle}{|R|} h_R \Big\|_{\exp(L^2)} \le 4^{-d} \sum_{k=n}^{\infty} \Big\| \sum_{R:|R|=2^{-k}} N2^{-k} h_R \Big\|_{\exp(L^2)}$$

$$\lesssim N \sum_{k=n}^{\infty} 2^{-k} \Big\| \Big(\sum_{R:|R|=2^{-k}} \mathbf{1}_R \Big)^{\frac{1}{2}} \Big\|_{\infty}$$

$$\lesssim N \sum_{k=n}^{\infty} \sqrt{k+1} \cdot 2^{-k} \approx \sqrt{n},$$

where we have once again applied Theorem 4.2.8.

The second sum in (4.6.1) 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

$$\Big\| \sum_{R:|R|<2^{-n}} \frac{\langle C_N, h_R\rangle}{|R|} h_R \Big\|_p = \Big\| \sum_{k=0}^{n} \sum_{\substack{\widetilde{R}:\,|\widetilde{R}|=2^{-n} \\ |\widetilde{R}_1|=2^{-k}}} \sum_{\substack{R\subset\widetilde{R} \\ R_1=\widetilde{R}_1}} \frac{\langle C_N, h_R\rangle}{|R|} h_R \Big\|_p \qquad (4.6.3)$$

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

$$\sum_{\substack{R\subset\widetilde{R} \\ R_1=\widetilde{R}_1}} \frac{\langle C_N, h_R\rangle}{|R|} h_R(x_1, x_2) = C_{\widetilde{R}}(x_2) \frac{\langle h_{\widetilde{R}_1}, \mathbf{1}_{[p_1,1]}\rangle}{|\widetilde{R}_1|} h_{\widetilde{R}_1}(x_1), \qquad (4.6.4)$$

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 $|C_{\widetilde{R}}(x_2)| \leq 2$. 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 (4.6.4) is non-zero. Thus, using (4.6.3)

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

$$= \sum_{k=0}^{n} \sum_{\widetilde{R}_1:|\widetilde{R}_1|=2^{-k}} \frac{\alpha_{\widetilde{R}_1}(x_2)}{|\widetilde{R}_1|} h_{\widetilde{R}_1}(x_1),$$

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:

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

We now integrate this estimate in x_2 to obtain

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

and thus

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

in view of Proposition 4.2.2. Thus, we have estimated the $\exp(L^2)$ norms of all the terms in (4.6.1) by $n^{\frac{1}{2}}$. The estimates for $(0,1)$ and $(1,0)$ Haars in (4.6.2) can be easily incorporated, invoking similar one-dimensional arguments and Lemma 4.4.8. 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 4.2.9 and inequality (4.3.9) finish the proof of Theorem 4.1.5 for all $\alpha \geq 2$.

4.6.3 Upper Bound 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 4.3.5. 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 $\mathbf{1}_{[v_{n,\sigma}(\tau),1)}(tx_1, x_2)$, restricted to $[0, 1]^2$ will be the indicator of a rectangle with one corner anchored at the upper right hand corner. Moreover, it will 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.

$$\|\Delta_N\|_{\exp(L^\alpha)} \leq 1 + \|D_{N'}(tx_1, x_2)\|_{\exp(L^\alpha)}$$

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

Remark 4.6.5. We make a final remark on the other upper bound of the dyadic BMO estimate of the digit-scrambled van der Corput set in Theorem 4.1.8. It is natural to guess that this estimate should hold for all N, and for BMO. A natural way to prove this is via the approach developed in [39],[47], but carrying out this argument is not completely straight forward.

RECOVERING SINGULAR INTEGRALS FROM HAAR SHIFTS

5.1   Introduction.

We will represent one-dimensional Calderón-Zygmund convolution operators with suffi-
ciently smooth kernels /see conditions (5.2.5), (5.2.6)/ by means of a properly chosen av-
eraging of certain Haar shift operators with bounded coefficients. By *Haar shift operators*
we mean linear operators that can be expressed in an efficient manner with the Haar basis
/see Remark (5.2.8)/.

The use of Haar shift operators to represent singular integral operators goes back to the
work of T.Figiel [16]. Later S. Petermichl derived a representation of the Hilbert transform
[36]. Similar representations were derived for Beurling [12], Riesz [35] transforms and the
truncated Hilbert transform (S. Petermichl, oral communication). The reason why these
representations are useful is that one can deduce deep facts about singular integral oper-
ators, based on the analysis of Haar shift operators. In Petermichl's original paper [36]
a deep property of Hankel operators associated to matrix symbols was deduced. These
representations also allowed to deduce the linear A_2 bound for the Hilbert [34] and Riesz
transforms [37]. The study of the Haar shift operators is interesting in itself [31],[33], and
has become an important model of the singular integral operators /see for instance their
use in [26],[25]/.

To illustrate this, as a corollary to the main result of this chapter, Theorem (5.2.4)
below, and the main result of [23], we see that we have a proved a sharp A_2 inequality for
the Calderon-Zygmund operators, a question of current interest:

Corollary 5.1.1. *Let*

$$T(f)(x) = P.V. \int_{\mathbb{R}} K(x-t)f(t)dt$$

be a one dimensional Calderon-Zygmund convolution operator whose kernel K is odd and

satisfies (5.2.5) and (5.2.6), then

$$\|Tf\|_{L_2(\omega)} \lesssim \|\omega\|_{A_2} \|f\|_{L_2(\omega)}.$$

By $\|\omega\|_{A_2}$ we mean the A_2 constant of the weight ω. (See [34],[27] for a definition.)

This generalizes the result of S. Petermichl, obtained for the Hilbert transform [34], and improves the estimates of A. Lerner, S. Ombrosi and C. Perez (equation 1.9 in [27]) for these particular type of Calderon-Zygmund operators.

5.2 Formulation of the Result.

In order to formulate the main theorem of this chapter we introduce some notations.

For any $\beta = \{\beta_l\} \in \{0,1\}^{\mathbb{Z}}$ and for any $r \in [1,2)$ define the dyadic grid $\mathbb{D}_{r,\beta}$ to be the collection of intervals

$$\mathbb{D}_{r,\beta} = \left\{ r2^n \left([0;1) + k + \sum_{i<n} 2^{i-n}\beta_i \right) \right\}_{n\in\mathbb{Z},\, k\in\mathbb{Z}}$$

This parametrization of dyadic grids appears explicitly in [19], and implicitly in section 9.1 of [32]. Note, that the dyadic grid we use is different from the one used in [36].

Place the usual uniform probability measure $\mathbb{P}$ on the space $\{0,1\}^{\mathbb{Z}}$, explicitly

$$\mathbb{P}(\beta : \beta_l = 0) = \mathbb{P}(\beta : \beta_l = 1) = \frac{1}{2}, \qquad \text{for all } l \in \mathbb{Z}.$$

We define two functions. Take h to be the function supported on $[0,1]$ defined by

$$h(x) = \begin{cases} 7, & 0 < x < 1/4, \\ -1, & 1/4 \le x \le 1/2, \\ 1, & 1/2 \le x \le 3/4, \\ -7, & 3/4 \le x \le 1. \end{cases} \tag{5.2.1}$$

and g to be the function supported on $[0,1]$ defined by

$$g(x) = \begin{cases} -1, & 0 \le x \le 1/4, \\ 1, & 1/4 \le x \le 1/2, \\ 1, & 1/2 \le x < 3/4, \\ -1, & 3/4 \le x \le 1. \end{cases} \tag{5.2.2}$$

Note that the function g appears in [36] paired with the usual Haar function. In contrast, our function h, defined by (5.2.1), differs a little bit from the Haar function. In some sense, our choice of the function h makes the convolution $h * g$ 'less smooth', and this property will be crucial for the proof. We'll make this statement precise in (5.3.12), which will permit us to invert a Fourier transform.

For any function f and any interval $I = [a; a + l]$ we define the function f_I to be the scaling of f to I which preserves the L_2-norm, namely

$$f_I(x) = f_{[a,a+l]} = \frac{1}{\sqrt{l}} f\left(\frac{x-a}{l}\right). \tag{5.2.3}$$

Now we are ready to state our main theorem:

Theorem 5.2.4. *Let $K : (-\infty, 0) \cup (0, \infty) \to \mathbb{R}$ be an odd, twice differentiable function (in the sense that K' is absolutely continuous) which satisfies*

$$\lim_{x \to \infty} K(x) = \lim_{x \to \infty} K'(x) = 0 \tag{5.2.5}$$

and

$$x^3 K''(x) \in L_\infty(\mathbb{R}). \tag{5.2.6}$$

Then there exists a coefficient-function $\gamma : (0, \infty) \to \mathbb{R}$, satisfying

$$\|\gamma\|_\infty \le C \|x^3 K''(x)\|_\infty$$

so that

$$K(x-y) = \int_{\{0,1\}^{\mathbb{Z}}} \int_1^2 \sum_{I \in \mathbb{D}_{r,\beta}} \gamma(|I|) \, h_I(x) g_I(y) \, \frac{dr}{r} \, d\mathbb{P}(\beta) \tag{5.2.7}$$

for all $x \ne y$. Here C is some absolute constant and the series on the right of (5.2.7) is a.e. absolutely convergent.

Remark 5.2.8. Note, that for for fixed r, β, and a function $\gamma \in L_\infty(R)$, the linear operator

$$f \mapsto \sum_{I \in \mathbb{D}_{r,\beta}} \gamma(|I|) \langle g_I, f \rangle h_I(x)$$

is an example of a Haar shift operator as defined in [23]. Note that this operator, expressed as a matrix in the Haar basis, has a bounded diagonals, but is even better than that: one

only needs to use Haar coefficients associated with dyadic intervals that intersect and have lengths that differ by at most a factor of 2.

5.3   Proof of the Recovering Theorem

5.3.1   Derivation of an Integral Equation.

The following lemma derives a concise formula for properly averaged Haar shift operators:

Lemma 5.3.1. *Suppose the functions h and g are defined by (5.2.1),(5.2.2) and suppose $\gamma \in L_\infty(R_+)$. Then for any $x \neq y$, we have*

$$\int_{\{0,1\}^{\mathbb{Z}}} \int_1^2 \sum_{I \in \mathbb{D}_{r,\beta}} \gamma(|I|) \, h_I(x) g_I(y) \, \frac{dr}{r} \, d\mathbb{P}(\beta) = \int_0^\infty \frac{\gamma(r)}{r^2} \, (h * g_1) \left(\frac{x - y}{r} \right) dr \qquad (5.3.2)$$

where $g_1(x) \equiv g(-x)$, and the functions h_I, g_I are defined by (5.2.3). Here the series on the left of (5.3.2) is a.e. absolutely convergent.

Remark 5.3.3. This lemma appears in [19] for the case $\gamma \equiv 1$.

Remark 5.3.4. The notation g_1 is introduced in the right hand side of (5.3.2) to emphasize the role of convolution.

Proof. The following calculation justifies the a.e. convergence of series,

$$\sum_{I \in \mathbb{D}_{r,\beta}} |\gamma(|I|) \, h_I(x) g_I(y)| \leq \|\gamma\|_\infty \sum_{\substack{I \in \mathbb{D}_{r,\beta} \\ x \in I, \, y \in I}} |h_I(x) g_I(y)| \leq$$

$$\leq \|\gamma\|_\infty \|h\|_\infty \|g\|_\infty \sum_{\substack{I \in \mathbb{D}_{r,\beta} \\ x \in I, \, |I| \geq |x-y|}} \frac{1}{|I|} \leq \frac{2\|\gamma\|_\infty \|h\|_\infty \|g\|_\infty}{|x - y|}.$$

84

Now recalling the definition of the dyadic grid $\mathbb{D}_{r,\beta}$ we get

$$\int_{\{0,1\}^{\mathbb{Z}}} \int_1^2 \sum_{I \in \mathbb{D}_{r,\beta}} \gamma(|I|) h_I(x) g_I(y) \frac{dr}{r} d\mathbb{P}(\beta)$$

$$= \int_{\{0,1\}^{\mathbb{Z}}} \int_1^2 \sum_{k \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \frac{\gamma(r2^n)}{r2^n} h\left(\frac{x}{r2^n} - k - \sum_{i<n} 2^{i-n}\beta_i\right) g\left(\frac{y}{r2^n} - k - \sum_{i<n} 2^{i-n}\beta_i\right) \frac{dr}{r} d\mathbb{P}(\beta)$$

$$= \int_1^2 \sum_{n \in \mathbb{Z}} \frac{\gamma(r2^n)}{r2^n} \int_{\mathbb{R}} h\left(\frac{x}{r2^n} - s\right) g\left(\frac{y}{r2^n} - s\right) ds \frac{dr}{r}$$

$$= \sum_{n \in \mathbb{Z}} \int_1^2 \frac{\gamma(r2^n)}{r2^n} (h * g_1)\left(\frac{x-y}{r2^n}\right) \frac{dr}{r}$$

$$= \sum_{n \in \mathbb{Z}} \int_{2^n}^{2^{n+1}} \frac{\gamma(r)}{r^2} (h * g_1)\left(\frac{x-y}{r}\right) dr$$

$$= \int_0^\infty \frac{\gamma(r)}{r^2} (h * g_1)\left(\frac{x-y}{r}\right) dr.$$

$\square$

Having this lemma at hand, the claim of theorem 5.2.4 is equivalent to the following: For h and g defined by (5.2.1) and (5.2.2), find a function $\gamma \in L_\infty(\mathbb{R})$ which would satisfy the following integral equation

$$K(x) = \int_0^\infty \frac{\gamma(r)}{r^2} (h * g_1)\left(\frac{x}{r}\right) dr, \tag{5.3.5}$$

for all $x > 0$. The case $x < 0$ would be satisfied automatically as both K and $h * g_1$ are odd.

5.3.2 Derivation of Recursive Equation

In this step we'll use the functional equation (5.3.5) to get a recursive equation (5.3.7) for the coefficient function γ.

Differentiating (5.3.5) twice, we get

$$K''(x) = \int_0^\infty \frac{\gamma(r)}{r^4} (h * g_1)''\left(\frac{x}{r}\right) dr, \qquad x > 0.$$

or equivalently

$$x^3 K''(x) = \int_0^\infty t^2 \gamma\left(\frac{x}{t}\right) (h * g_1)''(t) dt, \qquad x > 0. \tag{5.3.6}$$

Using the definitions (5.2.1),(5.2.2) for h and g, we see that $h * g_1$ is a continuous, piecewise linear, odd function. The graph of $h * g_1$ on the positive axis is illustrated in Figure 1.
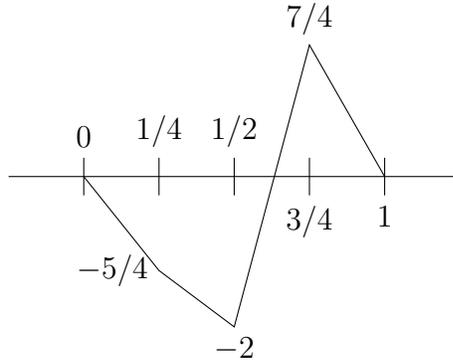
85

Figure 6: $h * g_1$ on the positive axis.

Thus, the function $(h * g_1)''$ is a linear combination of Dirac measures, which one can calculate from the graph above, in particular

$$(h * g_1)''(x) = 2\delta(x - 1/4) + 18\delta(x - 1/2) - 22\delta(x - 3/4) + 7\delta(x - 1), \qquad x > 0,$$

where δ is the usual Dirac delta function centered at the point 0.

With this, (5.3.6) becomes

$$x^3 K''(x) = 2\left(\frac{1}{4}\right)^2 \gamma(4x) + 18\left(\frac{1}{2}\right)^2 \gamma(2x) - 22\left(\frac{3}{4}\right)^2 \gamma\left(\frac{4}{3}x\right) + 7\gamma(x), \qquad x > 0. \quad (5.3.7)$$

Let's modify (5.3.7) to a form, which will be more convenient to us. Denote

$$m(x) = e^{3x} K''(e^x),$$

and

$$c(x) = \gamma(e^x). \tag{5.3.8}$$

In terms of these new notations the equation (5.3.7) becomes

$$m(x) = \tfrac{1}{8}c(x + \ln 4) + \tfrac{9}{2}c(x + \ln 2) - 22\left(\tfrac{3}{4}\right)^2 c\left(x + \ln\left(\tfrac{4}{3}\right)\right) + 7c(x), \qquad -\infty < x < \infty. \quad (5.3.9)$$

The condition (5.2.6) of theorem 5.2.4 provides that $m \in L_\infty(\mathbb{R})$. We want to find $c \in L_\infty(\mathbb{R})$ which would solve (5.3.9).

Remark 5.3.10. In the case of Hilbert transform we have $m(x) \equiv 2$, thus a constant function $c(x) \equiv C$ for a proper constant C would solve (5.3.9).

5.3.3 Fourier Transform

We'll use Fourier transform in order to solve the recursive functional equation (5.3.9). (Here we'll deal with Fourier transform of L_∞ functions, which is understood in a distributional sense.)

Apply Fourier transform to both sides of (5.3.9) to get

$$m^*(\omega) = a(\omega)c^*(\omega), \tag{5.3.11}$$

where

$$a(\omega) = \tfrac{1}{8}e^{i\omega \ln 4} + \tfrac{9}{2}e^{i\omega \ln 2} - 22\left(\frac{3}{4}\right)^2 e^{i\omega \ln\left(\frac{4}{3}\right)} + 7.$$

Now the function a is a Fourier transform of a finite Borel measure on $\mathbb{R}$. Also note the following important property of a: our choice of functions h and g provided that one of the terms of a dominates the rest

$$12\tfrac{3}{8} = 22\left(\frac{3}{4}\right)^2 > \frac{1}{8} + \frac{9}{2} + 7 = 11\tfrac{5}{8}. \tag{5.3.12}$$

In particular, we have $|a(\omega)| \geq \frac{3}{4}$ for all ω.

Recall that the space of Fourier transforms of finite Borel measures on $\mathbb{R}$, equipped with the L_∞ norms of these Fourier transforms, is a Banach algebra under pointwise multiplication. Therefore a is invertible, too. (The inverse of a can be written in terms of a Neumann series of exponents.) But this means that a^{-1} is a multiplier of the space $L_\infty(\mathbb{R})$. Hence, there exists a function $c \in L_\infty(\mathbb{R})$, which solves the equation (5.3.11) and $\|c\|_\infty < C\|m\|_\infty$ (for some absolute constant C). Using (5.3.8) we can further restore the coefficient-function γ. It would solve the integral equation (5.3.6) and would satisfy the same bound as the function c, i.e.

$$\|\gamma\|_\infty \leq C\|m\|_\infty.$$

This fact, along with the conditions (5.2.5) on kernel K is sufficient to make the integral in equation (5.3.5) convergent, and to justify the passage from (5.3.6) back to (5.3.5).

Remark 5.3.13. The conditions (5.2.5) and (5.2.6) are somewhat necessary. Indeed, if some functions $h, g : [0, 1] \to \mathbb{R}$ are constant on all dyadic intervals with sufficiently small length

and if the coefficient function γ is in $L_\infty(\mathbb{R})$ then the lemma (5.3.1) still holds. Thus, whatever kernel K is restored by the averaging of corresponding Haar shift operator, it must satisfy (5.3.5) and (5.3.6). If additionally h is odd and g is even with respect to the point $1/2$, then $h * g$ would be a piecewise linear function with bounded support, vanishing at 0. So, (5.3.5) and (5.3.6) would imply that K has to satisfy (5.2.5) and (5.2.6).

REFERENCES

[1] BECK, J., "A two-dimensional van Aardenne-Ehrenfest theorem in Irregularities of Distribution," *Compositio Math.*, vol. 72, no. 3, pp. 269–339.

[2] BECK, J. and CHEN, W. W. L., *Irregularities of Distribution*, vol. 89 of *Cambridge Tracts in Mathematics.* Cambridge University Press.

[3] BILYK, D., LACEY, M., and VAGHARSHAKYAN, A., "On the Signed Small Ball Inequality," *Online J. Anal. Comb.*, no. 3, pp. Art. 6, 7, 2008.

[4] BILYK, D. and LACEY, M. T., "On the Small Ball Inequality in three dimensions," *Duke Math. J.*, vol. 143, no. 1, pp. 81–115, 2008.

[5] BILYK, D., LACEY, M. T., PARISSIS, I., and VAGHARSHAKYAN, A., "Exponential squared integrability for the Discrepancy function in two dimensions," *Mathematika*, vol. 55.

[6] BILYK, D., LACEY, M. T., and VAGHARSHAKYAN, A., "On the Small Ball Inequality in all dimensions," *J. Funct. Anal.*, vol. 254, no. 9, pp. 2470–2502, 2008.

[7] CHANG, S.-Y. A., WILSON, J. M., and WOLFF, T. H., "Some weighted norm inequalities concerning the Schrödinger operators," *Comment. Math. Helv.*, vol. 60, no. 2, pp. 217–246.

[8] CHANG, S.-Y. A. and FEFFERMAN, R., "A continuous version of duality of H^1 with BMO on the bidisc," *Ann. of Math. (2)*, vol. 112, no. 1, pp. 179–201.

[9] CHANG, S.-Y. A. and FEFFERMAN, R., "Some recent developments in Fourier analysis and H^p-theory on product domains," *Bull. Amer. Math. Soc. (N.S.)*, vol. 12, no. 1, pp. 1–43.

[10] CHEN, W. W. L., "On Irregularities of Distribution," *Mathematika*, vol. 27, no. 2, pp. 153–170 (1981).

[11] CHEN, W. W. L., "On Irregularities of Distribution. ii," *Quart. J. Math. Oxford Ser. (2)*, vol. 34, no. 135, pp. 257–279.

[12] DRAGIČEVIĆ, O. and VOLBERG, A., "Sharp estimate of the Ahlfors-Beurling operator via averaging martingale transforms," *Michigan Math. J.*

[13] DUNKER, T., "Estimates for the Small Ball Probabilities of the fractional Brownian sheet," *J. Theoret. Probab.*, vol. 13, no. 2, pp. 357–382.

[14] DUNKER, T. and KÜHN, T., "Metric entropy of the integration operator and Small Ball Probabilities for the Brownian sheet," *C. R. Acad. Sci. Paris Sér. I Math.*, vol. 326, no. 3, pp. 347–352.

[15] FEFFERMAN, R. and PIPHER, J., "Multiparameter operators and sharp weighted inequalities," *Amer. J. Math.*, vol. 119, no. 2, pp. 337–369.

[16] FIGIEL, T., "Singular integral operators: a martingale approach," vol. 158, pp. 95–110, 1990.

[17] HALÁSZ, G., "On roth's method in the theory of Irregularities of Point Distributions," pp. 79–94.

[18] HALTON, J. H. and ZAREMBA, S. K., "The extreme and L^2 discrepancies of some plane sets," *Monatsh. Math.*, vol. 73, pp. 316–328, 1969.

[19] HYTÖNEN, T., "On Petermichl's dyadic shift and the Hilbert transform," *C. R. Math. Acad. Sci. Paris*, vol. 346, no. 21-22, pp. 1133–1136.

[20] KRITZER, P. and PILLICHSHAMMER, F., "An exact formula for the L_2 discrepancy of the shifted Hammersley point set," *Uniform Distribution Theory*, vol. 1, no. 1, 2006.

[21] KUELBS, J. and LI, W. V., "Metric entropy and the Small Ball Problem for Gaussian measures," *J. Funct. Anal.*, vol. 116, no. 1, pp. 133–157.

[22] KÜHN, T. and LINDE, W., "Optimal series representation of fractional Brownian Sheets," *Bernoulli*, vol. 8, no. 5, pp. 669–696.

[23] LACEY, M., PETERMICHL, S., and REGUERA, M. C., "Sharp A_2 inequality for Haar shift operators," *Math. Ann., to appear.*

[24] LACEY, M. T., "On the Discrepancy function in arbitrary dimension, close to L^1," *to appear in Analysis Mathematica.*

[25] LACEY, M. T., PETERMICHL, S., and PIPHER, J. C., "Multiparameter Riesz commutators," *Amer. J. Math.*, vol. 131, no. 3, pp. 731–769.

[26] LACEY, M. T., PIPHER, J. C., and PETERMICHL, S., "Iterated Riesz commutators: a simple proof of boundedness," *to appear in Proceeding of El Escorial, 2008.*

[27] LERNER, A. K., OMBROSI, S., and PÉREZ, C., "A_1 bounds for Calderón-Zygmund operators related to a problem of Muckenhoupt and Wheeden," *Math. Res. Lett.*

[28] LI, W. V. and SHAO, Q.-M., "Gaussian processes: inequalities, Small Ball Probabilities and applications," pp. 533–597.

[29] LINDENSTRAUSS, J. and TZAFRIRI, L., *Classical Banach spaces. I.* Springer-Verlag. Sequence spaces; Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. 92.

[30] MATOUŠEK, J., *Geometric Discrepancy*, vol. 18 of *Algorithms and Combinatorics.* Springer-Verlag. An illustrated guide.

[31] NAZAROV, F., TREIL, S., and VOLBERG, A., "The Bellman functions and two-weight inequalities for Haar multipliers," *J. Amer. Math. Soc.*, vol. 12, no. 4, pp. 909–928.

[32] NAZAROV, F., TREIL, S., and VOLBERG, A., "The Tb-theorem on non-homogeneous spaces," *Acta Math.*, vol. 190, no. 2, pp. 151–239.

[33] NAZAROV, F., TREIL, S., and VOLBERG, A., "Two weight inequalities for individual Haar multipliers and other well localized operators," *Math. Res. Lett.*, vol. 15, no. 3, pp. 583–597.

[34] PETERMICHL, S., "The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic," *Amer. J. Math.*, vol. 129, no. 5, pp. 1355–1375.

[35] PETERMICHL, S., TREIL, S., and VOLBERG, A., "Why the Riesz transforms are averages of the dyadic shifts?," *Publ. Mat.*, no. Vol. Extra, pp. 209–228.

[36] PETERMICHL, S., "Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol," *C. R. Acad. Sci. Paris Sér. I Math.*, vol. 330, no. 6, pp. 455–460.

[37] PETERMICHL, S., "The sharp weighted bound for the Riesz transforms," *Proc. Amer. Math. Soc.*, vol. 136, no. 4, pp. 1237–1249.

[38] PIPHER, J., "Bounded double square functions," *Ann. Inst. Fourier (Grenoble)*, vol. 36, no. 2, pp. 69–82.

[39] PIPHER, J. and WARD, L. A., "BMO from dyadic BMO on the bidisc," *J. Lond. Math. Soc. (2)*, vol. 77, no. 2, pp. 524–544.

[40] ROTH, K. F., "On Irregularities of Distribution," *Mathematika*, vol. 1, pp. 73–79.

[41] SCHMIDT, W. M., "Irregularities of Distribution. vii," *Acta Arith.*, vol. 21, pp. 45–50.

[42] SCHMIDT, W. M., "Irregularities of Distribution. x," pp. 311–329.

[43] SCHMIDT, W. M., *Lectures on Irregularities of Distribution*, vol. 56 of *Tata Institute of Fundamental Research Lectures on Mathematics and Physics*. Tata Institute of Fundamental Research.

[44] TALAGRAND, M., "The Small Ball Problem for the Brownian sheet," *Ann. Probab.*, vol. 22, no. 3, pp. 1331–1354.

[45] TEMLYAKOV, V. N., "Approximation of functions with bounded mixed derivative," *Proc. Steklov Inst. Math.*, no. 1(178), pp. vi+121.

[46] TEMLYAKOV, V. N., "An inequality for trigonometric polynomials and its application for estimating the entropy numbers," *J. Complexity*, vol. 11, no. 2, pp. 293–307.

[47] TREIL, S., "H^1 and dyadic H^1,"

[48] VAGHARSHAKYAN, A., "Recovering singular integrals from Haar shifts," *Proceedings of AMS, to appear.*

[49] VAN DER CORPUT, J. G., "Verteilungsfunktionen i," *Akad. Wetensch. Amdterdam, Proc.*, vol. 38, pp. 813–821.

[50] WANG, G., "Sharp dquare-function inequalities for conditionally symmetric martingales," *Trans. Amer. Math. Soc.*, vol. 328, no. 1, pp. 393–419.