The small ball inequality with restricted coefficients

by

Dimitrios Karslidis

A thesis submitted in partial fulfilment of the requirements for
the degree of

Doctor of Philosophy

in

The Faculty of Graduate and Postdoctoral Studies
(Mathematics)

The University of British Columbia
(Vancouver)

April 2016

Abstract

The main focus of this document is the small ball inequality. The small ball inequality is a functional inequality concerning the lower bound of the supremum norm of a linear combination of Haar functions supported on dyadic rectangles of a fixed volume. The sharp lower bound in this inequality, as yet unproven, is of considerable interest due to the inequality's numerous applications. We prove the optimal lower bound in this inequality under mild assumptions on the coefficients of a linear combination of Haar functions, and further investigate the lower bounds under more general assumptions on the coefficients. We also obtain lower bounds of such linear combinations of Haar functions in alternative function spaces such as exponential Orlicz spaces.

Preface

Much of the following document is adapted from two of the author's research papers: [29] and [30]. In particular, Proposition 5.6 and Section 6.1, Section 6.2 form the main content of [29], *On the signed small ball inequality with restricted coefficients*, while Section 6.3, Section 6.4 and all of Chapter 7 are adapted from [30], *Orlicz spaces bounds for special classes of hyperbolic sums.* The first of these two manuscripts has been accepted for publication and will appear in Indiana University Mathematics Journal and the second paper has been submitted.

Table of Contents

List of Figures

Acknowledgements

Primarily, I would like to thank my supervisor Malabika Pramanik. No part of this thesis would have been possible without her unwavering commitment and constant encouragement. Malabika has generously assisted me in all my scientific activities. She is always willing to give me advice and help me to navigate my way through the academic world.

I want to thank the members of my supervisory committee, Philip Loewen, Richard Froese and Akos Magyar for their valuable advice, their support, and their time. I would especially like to express my deep gratitude to Philip Loewen for supporting me in all my academic pursuits during my time at UBC. I would like to thank Fok-Shuen Leung for his constant support and his valuable advice. I want to express my gratitude to Dmitriy Bilyk for sharing his rich knowledge with me. We also have had many discussions on the small ball inequality from which I have benefited a lot.

I would also like to acknowledge the abundant support I received from my girlfriend, Pam Sargent, during the process of typing the thesis. She has helped me more than anybody else in drafting this document and has never ceased to believe in my abilities as a researcher. *Σε ευχαρίστω πολυ*!

Finally, I wish to thank my family, all of my friends and Idryma Paideias kai Europaikoy Politismou for their ample support all these years.

Chapter 1

Introduction

In this thesis, we focus on the small ball inequality, an inequality that conjectures a lower bound for the supremum norm of any linear combination of Haar functions. The precise definition of Haar functions will be given in Section 1.1, roughly speaking these are functions that take the value 1 on half of its support and -1 on the other half. This inequality, which remains a conjecture in dimensions $d \geq 3$, has important applications in several fields of mathematics. Here, we show that the conjectured inequality is valid under some mild additional assumptions on the linear combinations of Haar functions.

To get an idea of what this inequality claims, first consider dyadic intervals in $[0,1]$. These are intervals whose endpoints are consecutive dyadic rationals, *i.e.*, intervals of the form $\left[\frac{m}{2^n}, \frac{m+1}{2^n}\right)$ where $n \geq 0$ and $0 \leq m < 2^n$. For example, $\left[0, \frac{1}{4}\right)$ and $\left[\frac{1}{4}, \frac{1}{2}\right)$ are dyadic intervals of $[0,1]$ of length $1/4$. Now, for a dyadic interval I, consider the function h_I supported on I and taking the value -1 on the left half of I, 1 on the right half of I, and zero outside of the interval. Finally, form a linear combination of such functions h_I, where the dyadic interval I is of a fixed length, *i.e.*, $\sum_{|I|=2^{-n}} \alpha_I h_I$. For reasons described later on, one is often interested in determining how small the largest possible magnitude of such linear combinations can be. In other words, one seeks lower bounds of the supremum norm of functions of the form $\sum_{|I|=2^{-n}} \alpha_I h_I$. We will explain in the next paragraph why this is reasonably simple when $\sum_{|I|=2^{-n}} \alpha_I h_I$ is considered. The present work focuses on lower bounds of the supremum norm of the higher dimensional analogues of such functions. The small ball inequality is an inequality which conjectures a sharp lower bound for such quantities, as $\sum_{|R|=2^{-n}} \alpha_R h_R$, where h_R is a higher dimensional analogue of a function h_I and its precise definition is given in (1.3).

The key step in finding the largest possible magnitude of the function $\sum_{|I|=2^{-n}} \alpha_I h_I$ is to notice that dyadic intervals of a fixed length are disjoint. Therefore, the largest magnitude is the largest of numbers $|\alpha_I|$. However, this is not the case in higher dimensions, and this creates a barrier in solving the higher dimensional problem.

To further clarify the problem, consider a rectangle R, say with area $|R| = 2^{-n}$, in the unit square. Divide this rectangle into four subrectangles, R_i, $i = 1, \ldots, 4$, where R_1 corresponds to the upper left subrectangle, R_3 corresponds to the lower right subrectangle, and R_2 and R_4 correspond to the upper right and lower left subrectangles respectively. The length of each side of these subrectangles will be half of the length of the corresponding side of the original rectangle. Consider a function which takes the value $+1$ on R_1 and R_3, -1 on R_2 and R_4, and zero outside of the rectangle R. Denote such functions by h_R. A linear combination of such functions corresponding to planar rectangles of area 2^{-n} are functions of the form $H = \sum_{|R|=2^{-n}} \alpha_R h_R$. Letting the coefficients α_R vary yields different corresponding functions H. In this setting, the small ball inequality conjectures a lower bound for the largest value of H over all possible choices of coefficients in terms of n, for large values of n.

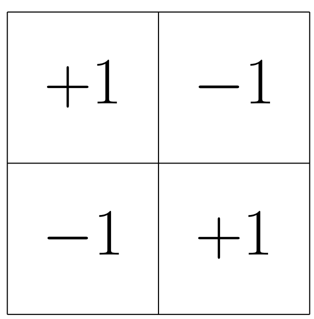


Figure 1.1: A Haar function, h_R, supported on a rectangle R.

The small ball inequality would be almost obvious if we only considered linear combinations of functions which correspond to non-overlapping rectangles. When the rectangles do overlap, the analysis is not straightforward. If we replace the planar rectangles by parallelepipeds, for example, the situation becomes much more difficult and delicate. In this thesis, we focus on obtaining lower bounds on the largest value of H in terms of n, in any dimension.

The formulation of the small ball inequality is given in Chapter 1. This chapter is introductory in nature and discusses various versions of the small ball inequality, and we give all of the preliminary facts needed to state the small ball inequality precisely. We also include the main results in Chapter 1 to better illustrate the connection of these results to the small ball inequality.

In Chapter 2, we outline the history of this problem and progress to date.

The main purpose of Chapter 3 is to highlight the connections between the small ball inequality and other areas of mathematics. Surprisingly, the validity of the small ball inequality is also a key to solving open problems in fields such as probability theory, approximation theory and the theory of irregularities of distributions. These connections have been dealt with precisely in Chapter 3, but we provide some motivation here in non-technical terms.

Let the unit square $T := [0,1]^2$ represent the city of Vancouver. To each point $\vec{t} = (t_1, t_2)$ of T, we associate the price, $X_{\vec{t}}$, of the property at that point. Since property prices in Vancouver are often high, it is natural to be interested in determining the likelihood that the property prices in the entire city will fall below a prescribed value, say $\varepsilon > 0$. In other words, we are concerned with estimating $\mathbb{P}(\sup_{\vec{t} \in T} X_{\vec{t}} < \varepsilon)$ for sufficiently small $\varepsilon > 0$, where $\mathbb{P}$ is the probability measure of the underlying space. Under certain normality on the probability distribution of $\{X_{\vec{t}}\}$, the small ball inequality will determine this probability. We give all the details of this connection in Section 3.1 of Chapter 3.

Now, suppose we have a collection of messages represented as binary strings of 0's and 1's. Due to physical constraints, we would like to determine a subcollection of the messages with the property that, given any message in the collection, there is a member of that subcollection which recovers the given message with prescribed accuracy. We obviously want to find the smallest subcollection that will do the job. This is related to the notion of entropy, and the small ball inequality plays a crucial role in solving such problems. This is further explained in Section 3.2 of Chapter 3.

In practice, there are many functions whose integrals cannot be written in terms of elementary functions. However, we may still be interested in estimating the area under such curves. One can, of course, approximate it by a Riemann sum, and it is natural to want to obtain a bound on the error in the approximation. It turns out that how well-distributed the tags of the partition of the Riemann sum are affects the error bound. The discrepancy function defined in Section 3.3 of Chapter 3 provides a way of determining how far the distribution of the tags are from a perfectly randomized scenario. In a quantifiable sense that has been made precise in Section 3.3, minimizing the error bound in the numerical integration problem corresponds to minimizing the discrepancy function, and the small ball inequality plays a vital role in bounding the discrepancy function. This connection is presented in Section 3.3 of Chapter 3.

In Chapter 4, we present a detailed proof of the two-dimensional small ball inequality. We also highlight the similarities between the arguments of this proof and the proof, given in Chapter 3, of the two-dimensional lower bound of the discrepancy function's L^∞-norm.

In Chapter 5, we provide the background needed to understand the main results of this thesis, presented in Chapter 6 and Chapter 7. In particular, we recall Khintchine's inequalities, a special case of the Littlewood-Paley inequalities for martingales, since Khintchine's inequalities are used in the proof of Theorem 1.4. We also introduce the Littlewood-Paley inequalities for martingales and connect them to the Littlewood-Paley inequalities for Haar functions, since such inequalities for Haar functions are used in the proofs of Theorem 1.3 and Theorem 1.6.

Chapters 6 and 7 are primarily devoted to proving the main results. In particular, Theorem 1.3 and Theorem 1.4 are proved in Chapter 6. In Chapter 7, we first give the necessary background information on exponential Orlicz spaces before moving on to proving Theorem 1.6. The reader will find this brief introduction to Orlicz spaces helpful in understanding the proof of Theorem 1.6. We conclude Chapter 7 with further investigation of hyperbolic sums in exponential Orlicz spaces.

Chapter 8, the final chapter of the thesis, provides a summary of the results of this thesis and outlines some related future projects.

1.1 Preliminary facts

Let $\mathbb{1}_I(x)$ be the characteristic function of the interval I, *i.e.*,

$$\mathbb{1}_I(x) = \begin{cases} 1, & x \in I \\ 0, & \text{otherwise} \end{cases}.$$

Consider the collection of the dyadic intervals of $[0,1]$:

$$\begin{aligned} &\mathcal{D} = \left\{ I = \left[\frac{m}{2^n}, \frac{m+1}{2^n}\right) : m, n \in \mathbb{Z}, n \geq 0, 0 \leq m < 2^n \right\} \quad \text{with} \\ &\mathcal{D}_* = \mathcal{D} \cup \{[-1,1]\}. \end{aligned} \tag{1.1}$$

For every interval $I \in \mathcal{D}$, its left and right dyadic halves are denoted by I_l and I_r respectively. We define the L^∞-normalized Haar function, h_I, corresponding to an interval I as:

$$h_I(x) = -\mathbb{1}_{I_l} + \mathbb{1}_{I_r}. \tag{1.2}$$

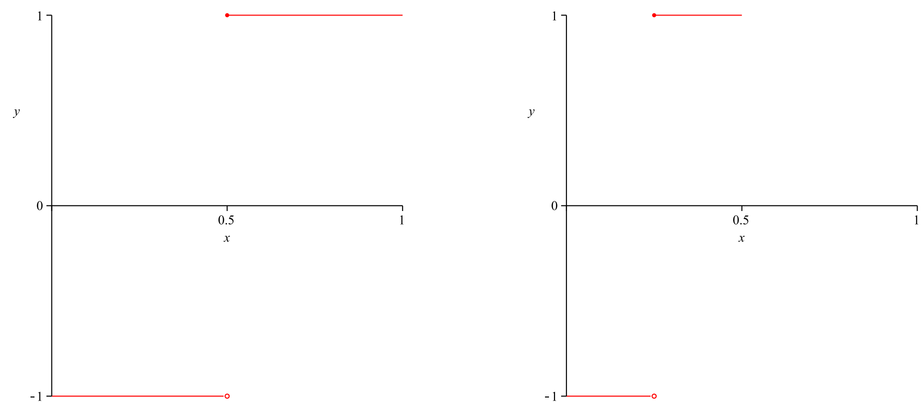
Note that $h_{[-1,1]}(x) = \mathbb{1}_{[0,1)}(x)$ for all x in $[0,1]$.

It is easy to extend Haar functions to higher dimensions. We consider the family of dyadic rectangles in dimension $d \geq 2$:

$$\mathcal{D}^d = \{R = R_1 \times \cdots \times R_d : R_j \in \mathcal{D}\},$$

i.e., every $R \in \mathcal{D}^d$ is a Cartesian product of dyadic intervals. The Haar functions supported on R are defined as a coordinate-wise product of one-dimensional Haar functions:

$$h_R(x_1, \ldots, x_d) = h_{R_1}(x_1) \cdots h_{R_d}(x_d), \text{ where } R = R_1 \times \cdots \times R_d, \; R_j \in \mathcal{D}. \tag{1.3}$$



(a) A Haar function supported on the interval $[0, 1)$.

(b) A Haar function supported on the interval $[0, 0.5)$.

Figure 1.2: Examples of Haar functions.

If we consider two distinct dyadic intervals, then either one will be strictly contained in the other, or they will be disjoint. Haar functions enjoy the following properties:

$$h_R^2(x) = \mathbb{1}_R(x), \tag{1.4}$$

$$\int_{[0,1]^d} h_R(x) h_{R'}(x) dx = 0, \text{ whenever } R \neq R', \tag{1.5}$$

$$\int_{[0,1]^d} h_R(x) dx = 0, \quad \forall\, R \in \mathcal{D}^d. \tag{1.6}$$

Let

$$\mathbb{H}_n^d = \{\vec{r} \in \mathbb{Z}_+^d : |\vec{r}| := r_1 + r_2 + \cdots + r_d = n\}$$

with

$$\mathbb{Z}_+^d = \{\vec{r} = (r_1, \ldots, r_d) : r_j \geq 0 \text{ and } r_j \in \mathbb{Z}, j = 1, \ldots, d\},$$

and

$$\mathcal{R}_{\vec{r}} = \{R \in \mathcal{D}^d : R_j \in \mathcal{D} \text{ and } |R_j| = 2^{-r_j}, j = 1, \dots, d\}.$$

In other words, the set $\mathcal{R}_{\vec{r}}$ is a collection of axis-parallel dyadic rectangles whose side length along the j^{th} axis is 2^{-r_j}. The rectangles in $\mathcal{R}_{\vec{r}}$ are disjoint and partition the d-dimensional unit cube, $[0,1)^d$. Also, it will be useful to observe that the cardinality of the set $\mathbb{H}_n^d$ denoted by $\#\mathbb{H}_n^d$ is of order n^{d-1}. This can be stated as $\#\mathbb{H}_n^d \simeq n^{d-1}$, which means that there exist positive constants $C_1(d)$ and $C_2(d)$ depending on the dimension d such that $C_1(d)n^{d-1} \le \#\mathbb{H}_n^d \le C_2(d)n^{d-1}$. This estimate agrees with our intuition, since each of the first $d-1$ coordinates of the vector $\vec{r} \in \mathbb{H}_n^d$ can be chosen in n different ways. The last coordinate is determined immediately once the first $d-1$ coordinates have been selected, since such $\vec{r}$ has prescribed l^1-norm, $|\vec{r}| = n$.

A function defined on $[0,1]^d$ of the form

$$f_{\vec{r}} = \sum_{R \in \mathcal{R}_{\vec{r}}} \alpha_R h_R \quad \text{with } \alpha_R = \pm 1$$

is called an $\vec{r}$-function with parameter $\vec{r} \in \mathbb{H}_n^d$. These ± 1-valued functions are also known in the literature as generalized Rademacher functions. It can easily be verified that $\vec{r}$-functions are orthonormal with respect to the L^2-norm.

In all dimensions d, $|\cdot|$ stands for the d-dimensional Lebesgue measure. The signum function of a real number x is defined by

$$\operatorname{sgn}(x) = \begin{cases} 1, & \text{if } x \ge 0 \\ -1, & \text{if } x < 0 \end{cases}. \tag{1.7}$$

1.2 Statement of the small ball inequality

As mentioned earlier, the small ball inequality is a functional inequality concerning the lower bound of the supremum norm of a sum of Haar functions supported on dyadic rectangles of a fixed volume. Such sums of Haar functions are also referred to as *hyperbolic sums.*

For brevity, let $\mathcal{A}_n^d = \{R \in \mathcal{D}^d : |R| = 2^{-n}\}$, *i.e.*, the set of all dyadic rectangles whose d-dimensional volume is equal to 2^{-n}.

Conjecture 1.1. (The small ball conjecture)*:*

In all dimensions $d \geq 2$, for any choice of the coefficients α_R, and all integers $n \geq 1$, one has the following inequality

$$n^{\frac{d-2}{2}} \left\| \sum_{R \in \mathcal{A}_n^d} \alpha_R h_R \right\|_\infty \geq C_d 2^{-n} \sum_{R \in \mathcal{A}_n^d} |\alpha_R|, \tag{1.8}$$

where C_d is a constant depending only on the dimension d, not on n or the choice of $\{\alpha_R\}$.

The critical feature of inequality (1.8) is the precise exponent of n occurring on the left side. If we replace $n^{\frac{d-2}{2}}$ with $n^{\frac{d-1}{2}}$, then we get the so-called "trivial bound":

$$n^{\frac{d-1}{2}} \left\| \sum_{R \in \mathcal{A}_n^d} \alpha_R h_R \right\|_{L^2} \geq C_d 2^{-n} \sum_{R \in \mathcal{A}_n^d} |\alpha_R|. \tag{1.9}$$

If we compare the conjectured small ball inequality with inequality (1.9), noting that the $L^\infty([0,1]^d)$-norm is at least as large as the $L^2([0,1]^d)$-norm, then we can see that the conjectured small ball inequality is better than (1.9) by a factor of $\sqrt{n}$.

Proof of (1.9). We calculate the L^2-norm of $H_n = \sum_{|R|=2^{-n}} \alpha_R h_R$. Using the orthogonality of Haar functions and the fact that $\|h_R\|_2^2 = 2^{-n}$, Parseval's identity gives us

$$\begin{aligned} \|H_n\|_2^2 = \sum_{|R|=2^{-n}} |\alpha_R|^2 \|h_R\|_2^2 &= 2^{-n} \sum_{|R|=2^{-n}} |\alpha_R|^2 \\ &\geq C_d^2 2^{-n} \frac{\left(\sum_{|R|=2^{-n}} |\alpha_R| \right)^2}{2^n n^{d-1}}, \end{aligned}$$

where in the last line, we used the Cauchy-Schwarz inequality and the fact that $\#\{R \in \mathcal{D}^d : |R| = 2^{-n}\} = 2^n \# \mathbb{H}_n^d \simeq 2^n n^{d-1}$. Now, taking the square root of both sides, we get

$$\|H_n\|_2 \geq C_d \Big(\sum_{|R|=2^{-n}} |\alpha_R| \Big) 2^{-n} n^{-\frac{d-1}{2}},$$

which is exactly (1.9). $\square$

There is a simplified form of the small ball inequality which is of interest as well. If we consider coefficients $\{\alpha_R\}$ from the set $\{-1, 1\}$ in the small ball conjecture, then we obtain the signed small ball conjecture, *i.e.*,

Conjecture 1.2. (The signed small ball conjecture) *If $\alpha_R = \pm 1$ for every $R \in \mathcal{A}_n^d$, then we have*

$$\Big\| \sum_{R \in \mathcal{A}_n^d} \alpha_R h_R \Big\|_\infty \geq C_d n^{\frac{d}{2}}, \tag{1.10}$$

where C_d is a constant depending only on the dimension d, not on n or the choice of $\{\alpha_R\}$.

Note that when $\{\alpha_R\} \subset \{-1, 1\}$, then the trivial bound in (1.9) is given by

$$\Big\| \sum_{R \in \mathcal{A}_n^d} \alpha_R h_R \Big\|_{L^2} \geq C_d n^{\frac{d-1}{2}}. \tag{1.11}$$

In Section 4.2 we will prove that the conjectured signed small ball inequality, inequality (1.10), is sharp. In other words, we will prove that there exists a positive constant $C_d > 0$ and a choice of the coefficients $\{\alpha_R\} \subset \{-1, 1\}$ such that

$$\Big\| \sum_{R \in \mathcal{A}_n^d} \alpha_R h_R \Big\|_\infty \leq C_d n^{\frac{d}{2}}.$$

In particular, this fact will imply that if inequality (1.8) holds, then it is sharp as well.

1.3 Main results

According to inequality (1.10), if one wants to attempt to solve the signed small ball conjecture, then one has to obtain a gain over the trivial bound given in (1.11), $n^{\frac{d-1}{2}}$, by a factor of $\sqrt{n}$ so that $\sqrt{n} \cdot n^{\frac{d-1}{2}}$ is the optimal lower bound in the signed small ball inequality. The best known gain over the trivial bound for general coefficients $\{\alpha_R\} \subset \{-1, 1\}$ for all $d \geq 3$ was obtained by D. Bilyk, M. Lacey, and A. Vagharshakyan in [5]. We will state their result precisely in Chapter 2. In this work, we follow a different approach to that taken in [5]. Here, we impose a restriction on the collection of coefficients and aim to get the optimal bound appearing in the signed small ball conjecture for this restricted collection of coefficients.

To better understand the restriction we impose on the coefficients, let us observe that, for given $n \geq 1$ and $d \geq 3$, any choice of coefficients $\{\alpha_R\}_{R \in \mathcal{A}_n^d} \subset \{-1, 1\}$ can be generated by a map

$$\alpha^* : \mathcal{A}_n^d \to \{-1, 1\}. \tag{1.12}$$

such that $\alpha^*(R) = \alpha_R$ for all $R \in \mathcal{A}_n^d$. Since

$$\mathcal{A}_n^d \subset \bigcup_{r_1=0}^{n} \mathcal{A}_{r_1}^1 \times \bigcup_{r_1=0}^{n} \mathcal{A}_{n-r_1}^{d-1},$$

the map α^* in (1.12) can be viewed as a restriction of a, not necessarily unique, map

$$\alpha : \bigcup_{r_1=0}^{n} \mathcal{A}_{r_1}^1 \times \bigcup_{r_1=0}^{n} \mathcal{A}_{n-r_1}^{d-1} \to \{-1, 1\}.$$

The signed small ball conjecture states that inequality (1.10) must hold for the collection of coefficients $\{\alpha_R := \alpha(R) = \alpha^*(R), R \in \mathcal{A}_n^d\}$. We will be interested in collections of coefficients generated by maps of the form:

$$\alpha : \bigcup_{r_1=0}^{n} \mathcal{A}_{r_1}^1 \times \bigcup_{r_1=0}^{n} \mathcal{A}_{n-r_1}^{d-1} \to \{-1, 1\}$$

with

$$\alpha(R_1 \times R_2 \times \cdots \times R_d) = \alpha_1(R_1) \cdot \alpha_2(R_2 \times \cdots \times R_d), \tag{1.13}$$

where

$$\alpha_1 : \bigcup_{r_1=0}^{n} \mathcal{A}^1_{r_1} \to \{-1, 1\} \quad \text{and} \quad \alpha_2 : \bigcup_{r_1=0}^{n} \mathcal{A}^{d-1}_{n-r_1} \to \{-1, 1\}.$$

With a slight abuse of notation, we denote the values of a map in (1.13) by $\alpha_R = \alpha_{R_1}\alpha_{R_2\times\cdots\times R_d}$, and we let

$$\mathcal{A}_{\text{split}} = \Big\{\alpha_R : R \in \mathcal{A}^d_n, \alpha_R = \alpha_{R_1} \cdot \alpha_{R_2\times\cdots\times R_d} \text{ with } \alpha_{R_1}, \alpha_{R_2\times\cdots\times R_d} \in \{\pm 1\}\Big\}.$$

We will show that for any integer $n \geq 1$, $d \geq 3$ and any choice of coefficients in $\mathcal{A}_{\text{split}}$, Conjecture 1.2 holds, *i.e.*,

Theorem 1.3. *For all integers $n \geq 1$, $d \geq 3$ and all choices of coefficients $\{\alpha_R\}_{R\in\mathcal{A}^d_n} \subset \mathcal{A}_{split}$ we have that*

$$\Big\| \sum_{R\in\mathcal{A}^d_n} \alpha_R h_R \Big\|_\infty \geq C_d n^{\frac{d}{2}}, \tag{1.14}$$

where C_d is a constant depending only on the dimension d, not on n or the choice of $\{\alpha_R\}$.

Moreover, the above inequality is sharp.

Theorem 1.4. *For any integer $n \geq 1$ and $d \geq 3$, there exists a collection of coefficients $\{\alpha_R\}_{R\in\mathcal{A}^d_n} \subset \mathcal{A}_{split}$ such that*

$$\Big\| \sum_{R\in\mathcal{A}^d_n} \alpha_R h_R \Big\|_\infty \leq C_d n^{\frac{d}{2}},$$

where C_d is a constant depending only on the dimension d.

We have already seen that the L^2-norm of the signed hyperbolic sum, $H_n = \sum_{|R|=2^{-n}} \alpha_R h_R$ with $\{\alpha_R\} \subset \{-1, 1\}$, can be calculated explicitly up to a multiplicative constant by using the orthogonality of Haar functions, *i.e.*, $\|H_n\|_2 \simeq n^{\frac{d-1}{2}}$. It turns out that, for any $p < \infty$, the L^p-norm of the signed hyperbolic sum has the same behaviour as its L^2-norm, *i.e.*, $\|H_n\|_p \simeq n^{\frac{d-1}{2}}$, where the implicit constants depend on p and on the dimension d. We will see that this fact is a consequence of the Littlewood-Paley inequalities. In order to gain insight and to better understand the signed small ball conjecture, we will also look for bounds of signed hyperbolic sums in function spaces which lie between L^p and L^∞. Exponential Orlicz spaces, denoted by $\exp(L^a)$, are one type of such function spaces. These will be defined and further discussed in Chapter 7. Exponential Orlicz spaces have already appeared in the study of problems related to the small ball inequality. Based on the work of Bilyk *et al.* in [6], in which they prove an analogous result in dimension 2 for the discrepancy function instead of hyperbolic sums, we conjecture the following:

Conjecture 1.5. *If $\alpha_R = \pm 1$ for every $R \in \mathcal{A}_n^d$, then for $d \geq 2$ and $n > 1$ we have*

$$\left\| \sum_{|R|=2^{-n}} \alpha_R h_R \right\|_{\exp(L^a)} \geq C(d,a) n^{\frac{d-1}{2}+\frac{1}{2}-\frac{1}{a}}, \ 2 \leq a < \infty, \tag{1.15}$$

where $C(d,a)$ is a constant depending only on the dimension d and the scale of integrability a, not n or the choice of $\{\alpha_R\}$.

The validity of this conjecture remains unknown in $d \geq 3$. However, adjusting the arguments given by V. Temlyakov in [49] shows that the conjecture is true in dimension $d = 2$. We will also verify a weaker version of Conjecture 1.5 in all dimensions $d \geq 3$ in Chapter 7. In particular, we will prove the following theorem:

Theorem 1.6. *For any integer $n > 1$ and any choice of coefficients $\alpha_R \in \mathcal{A}_{split}$, we have*

$$\left\| \sum_{|R|=2^{-n}} \alpha_R h_R \right\|_{\exp(L^a)} \geq C(d,a)\ n^{\frac{d-1}{2}+\frac{1}{2}-\frac{1}{a}}, \ 2 \leq a < \infty \tag{1.16}$$

for $d \geq 3$.

Theorem 1.3 and Theorem 1.4 showcase the usage of the collection $\mathcal{A}_{\text{split}}$ in proving optimal bounds for the conjectured signed small ball inequality. Therefore, it is natural to ask whether the conjectured inequalities (1.10) and (1.15) hold under structural assumptions on the coefficients. Such questions have been discussed in Chapter 8.

Chapter 2

Recent history of the small ball inequality

In this chapter, we focus on outlining known results pertaining to the resolution of the small ball conjecture.

In dimension $d = 2$, the small ball conjecture was first proved by M. Talagrand [47] in 1994, and Temlyakov [49] gave a second proof in 1995. However, in more general dimensions $d \geq 3$, the small ball conjecture remains unsolved.

Remark 2.1. In what follows, the expression $A \gtrsim B$ means that there is a constant K such that $A \geq KB$. In our setting, K does not depend on n or $\{\alpha_R\}$.

The first improvement over the trivial bound in higher dimensions was given in dimension $d = 3$ by J. Beck [2]. In particular, the lower bound he obtained was better than the trivial bound by a factor proportional to $(\log n)^{-1}$. In 2008, a body of work authored by Bilyk, Lacey, and Vagharshakyan [3, 4] made significant progress toward understanding the small ball inequality. They proved the following theorem:

Theorem 2.2. *For $d \geq 3$, there exists $0 < \eta(d) < \frac{1}{2}$ such that, for all choices of coefficients α_R and all non-negative integers $n \geq 1$,*

$$n^{\frac{d-1}{2}-\eta(d)} \left\| \sum_{R \in \mathcal{A}_n^d} \alpha_R h_R \right\|_\infty \gtrsim 2^{-n} \sum_{R \in \mathcal{A}_n^d} |\alpha_R|. \tag{2.1}$$

In the proof of Theorem 2.2, no specific values of $\eta(d)$ are given, though it is clear from the method of proof that $\eta(d)$ must be small. The small ball conjecture says that (2.1) should hold with $\eta = \frac{1}{2}$.

The proofs of the two-dimensional small ball inequality given by Talagrand and Temlyakov were fundamentally different. Temlyakov's proof used a duality argument whereas Talagrand's used a combinatorial argument. Here we give a rough outline of Temlyakov's duality argument as similar techniques have been employed in proving Theorem 1.3 and Theorem 1.6.

To begin the argument, Temlyakov constructed a function, called a test function, $\Psi \in L^1([0,1]^2)$ such that

$$\|\Psi\|_1 \leq C, \tag{2.2}$$

where C is an absolute constant which does not depend on the choice of the coefficients $\{\alpha_R\}$ or the integer $n \geq 1$. He further required that

$$\left\langle \Psi, \sum_{|R|=2^{-n}} \alpha_R h_R \right\rangle := \int_{[0,1]^d} \Psi \cdot \sum_{|R|=2^{-n}} \alpha_R h_R dx \geq C 2^{-n} \sum_{|R|=2^{-n}} |\alpha_R| \tag{2.3}$$

for all coefficients $\{\alpha_R\}$ and every $n \geq 1$. He then deduced the two-dimensional small ball inequality from a simple application of Hölder's inequality to (2.3) which exploited (2.2).

The structure of the test function that Temlyakov constructed restricted the proof's validity to the two-dimensional small ball inequality. In particular, the test function that Temlyakov constructed was in the form of a Riesz product, *i.e.*,

$$\Psi = \prod_{k=0}^{n} \left(1 + f_k \right), \tag{2.4}$$

where f_k is the two-dimensional $\vec{r}$-function defined by

$$f_k := f_{(k,n-k)} = \sum_{R \in \mathcal{R}_{\vec{r}}} \operatorname{sgn}(\alpha_R) h_R, \quad \vec{r} = (k, n-k).$$

In dimension $d = 2$, the product of Haar functions supported on dyadic rectangles of a fixed volume is again a Haar function. This property, known as the Product Rule, allows one to write $\Psi - 1$ as a sum of Haar functions, and verify that Ψ satisfies (2.2) and (2.3). The precise formulation of the Product Rule is given in Lemma 3.23, stated in Section 3.3. Unfortunately, this property does not hold in dimensions $d \geq 3$, and so Temlyakov's proof only applies in the two-dimensional case. Temlyakov's proof will be presented in more detail in Chapter 4.

The absence of the Product Rule in higher dimensions makes the general small ball conjecture a very challenging problem. However, Beck [2] and Bilyk, Lacey and Vagharshakyan [3, 4] successfully modified Temlyakov's approach to yield partial results in higher dimensions. To adapt Temlyakov's duality approach, they instead considered the function

$$\Psi = \prod_{t=1}^{q} (1 + \tilde{\rho} F_t), \tag{2.5}$$

where $q > 0$ is a constant depending on n, $\tilde{\rho} > 0$ is a constant depending on q, and

$$F_t = \sum_{\vec{r} \in \mathbb{A}_t} f_{\vec{r}}. \tag{2.6}$$

Here, $\mathbb{A}_t$ is the collection of vectors whose first-coordinate depends on t in such a way that $\{\mathbb{A}_t\}_{t=1}^{q}$ partitions $\mathbb{H}_n^d$, *i.e.*,

$$\bigsqcup_{t=1}^{q} \mathbb{A}_t = \mathbb{H}_n^d.$$

Expanding the product in (2.5) allowed them to write

$$\Psi = 1 + \Psi^{sd} + \Psi^{\neg}, \tag{2.7}$$

where Ψ^{sd} is a sum of a product of $\vec{r}$-functions for which the Product Rule can be applied, and $\Psi^{\neg}$ is a sum of a product of $\vec{r}$-functions for which the Product Rule fails. They then employed the Littlewood-Paley inequalities to obtain estimates on the L^p-norm of a product of $\vec{r}$-functions for which the Product Rule fails. Combining these estimates with combinatorial arguments, Bilyk, Lacey and Vagharshakyan proved that

$$\|\Psi^{sd}\|_1 \leq C$$

for $q = n^\eta$. As a result, they obtained

$$\begin{aligned}\langle H_n, \Psi^{sd}\rangle &= \int_{[0,1]^d} H_n(x)\Psi^{sd}(x)dx \\ &\gtrsim q n^{-\frac{d-1}{2}} \cdot 2^{-n} \sum_{R\in\mathcal{A}_n^d} |\alpha_R|.\end{aligned} \tag{2.8}$$

Finally, they used Hölder's inequality to finish the proof of Theorem 2.2. This theorem gives an improvement over the trivial bound by a factor of q. Before their work, Beck had already used probabilistic and combinatorial methods to prove the L^1-boundedness of Ψ^{sd} in dimension $d = 3$ for a value of q which was logarithmic in n and smaller than n^η. Beck's work heavily influenced [3, 4].

As mentioned in Chapter 1, the complete resolution of the small ball conjecture seems to be a difficult problem, but the signed small ball conjecture, stated as Conjecture 1.2 on page 9, is thought to be more manageable, due to the additional simplifying assumptions on the coefficients $\{\alpha_R\}$. In the case of the signed small ball inequality, the trivial bound reduces to

$$\left\| \sum_{\vec{r}\in\mathbb{H}_n^d} f_{\vec{r}} \right\|_{L^2} \gtrsim n^{\frac{d-1}{2}}, \tag{2.9}$$

where

$$H_n = \sum_{R\in\mathcal{A}_n^d} \alpha_R h_R = \sum_{\vec{r}\in\mathbb{H}_n^d} f_{\vec{r}}.$$

The trivial bound can be obtained by using the orthogonality of generalized Rademacher functions and the fact that $\#\mathbb{H}_n^d$, the cardinality of the set $\mathbb{H}_n^d$, is of order n^{d-1}. The exponent of the integer n appears naturally in the trivial bound (2.9). This comes from the volume constraint on dyadic rectangles, $|R| = 2^{-n}$, or equivalently $|\vec{r}| = n$, which reduces the number of "free" parameters in the vector

$\vec{r} \in \mathbb{H}_n^d$ to $d-1$. Therefore, the total number of terms in the sum is of order n^{d-1}. In [5], Bilyk, Lacey, and Vagharshakyan explicitly quantified the improvement over the trivial bound, *i.e.*, they proved the signed version of the Theorem 2.2, explicitly providing the range of values for η. In particular, they showed that, for every $\epsilon > 0$,

$$\left\| \sum_{R \in \mathcal{A}_n^d} \alpha_R h_R \right\|_\infty \geq C_d \cdot n^{\frac{d-1}{2} + \frac{1}{8d} - \epsilon}, \tag{2.10}$$

where $\alpha_R \in \{\pm 1\}$. Under an additional assumption on the length of the first side of the dyadic rectangle R, Bilyk, Lacey, I. Parissis and Vagharshakyan [7] made an improvement over (2.10) when $d = 3$. Specifically, they showed that

$$\left\| \sum_{\substack{|R| = 2^{-n} \\ |R_1| \geq 2^{-n/2}}} \alpha_R h_R \right\|_\infty \gtrsim n^{\frac{9}{8}}, \tag{2.11}$$

where $\alpha_R \in \{\pm 1\}$, and the sum is taken over all dyadic rectangles $R = R_1 \times \cdots \times R_d \in \mathcal{A}_n^d$ with the length of the first side R_1 of R to be bounded below by $2^{-n/2}$.

Chapter 3

Connections and applications

As indicated in the introduction, the small ball inequality lies at the interface of many areas of mathematics. This chapter, and the next, is devoted to a detailed analysis of these connections. The material in these chapters has been adapted from the excellent survey written by Bilyk [10].

3.1 Connection to probability theory

For convenience, we start by recalling some definitions from probability theory.

Definition 3.1. (a) If T is an arbitrary index set, a *centred Gaussian process*, $g = \{g_t\,;\, t \in T\}$, is a stochastic process with the property that, for any positive integer m and for every choice of $t_1, \dots, t_m \in T$, $(g_{t_1}, \dots, g_{t_m})$ has a multivariate normal distribution. That is, the joint density function of $(g_{t_1}, \dots, g_{t_m})$ is given by

$$f(\vec{x}) = \{(2\pi)^m \det(\mathbf{V})\}^{-1/2} \exp\left\{ -\frac{1}{2}\vec{x}\,\mathbf{V}^{-1}\vec{x}' \right\}, \quad \vec{x} = (x_1, \dots, x_m) \in \mathbb{R}^m,$$

where $\mathbf{V} = (v_{ij})$, $v_{ij} = \mathbb{E}(g_{t_i} g_{t_j})$, is the $m \times m$ covariance matrix depending on $t_1, \dots, t_m$, and $\vec{x}'$ is the transpose of the vector $\vec{x}$.

(b) A Brownian sheet, $\mathbb{B}$, is a centred multi-parameter Gaussian process which satisfies the covariance relation

$$\mathbb{E}\left(\mathbb{B}(\vec{s})\mathbb{B}(\vec{t})\right) = \prod_{j=1}^{d} \min(s_j, t_j),$$

for $\vec{s}, \vec{t} \in T := [0,1]^d$, with $\vec{s} = (s_1, \ldots, s_d)$ and $\vec{t} = (t_1, \ldots, t_d)$.

The name "small ball inequality" comes from probability theory, where one is interested in determining the exact asymptotic behaviour of the small deviation probability $\phi(\varepsilon) := -\log \mathbb{P}(\|\mathbb{B}\|_{L^\infty([0,1]^d)} < \varepsilon)$ as $\varepsilon \to 0^+$. In general, research in small deviation problems aims to uncover the asymptotic behaviour of the probability that a random process is accumulated in a small ball around the mean of the process in some norm.

It is a known fact that in dimension $d = 1$ (when $\mathbb{B}$ is Brownian motion), $\lim_{\varepsilon\to 0^+} \frac{\phi(\varepsilon)}{\varepsilon^{-2}} = \frac{\pi^2}{8}$, see [19]. In higher dimensions, the situation is more difficult and subtle, though upper bounds of the asymptotic behaviour of $\phi(\varepsilon)$ as $\varepsilon \to 0^+$ are completely understood. In particular, for all $d \geq 2$ and $\varepsilon > 0$ small,

$$\phi(\varepsilon) \leq \frac{C}{\varepsilon^2}\left(\log\frac{1}{\varepsilon}\right)^{2d-1}. \tag{3.1}$$

This inequality was proved by T. Dunker, T. Kühn, M. Lifshits and W. Linde [18].

Lower bounds of $\phi(\varepsilon)$ for small $\varepsilon > 0$ are not as well-understood as the upper bounds, though it is believed that they should have the same form.

Conjecture 3.2. *If $d \geq 2$ and $\phi(\varepsilon)$ is the small deviation probability associated to the Brownian sheet $\mathbb{B}$, then for small $\varepsilon > 0$,*

$$\frac{C'}{\varepsilon^2}\left(\log\frac{1}{\varepsilon}\right)^{2d-1} \leq \phi(\varepsilon) \leq \frac{C}{\varepsilon^2}\left(\log\frac{1}{\varepsilon}\right)^{2d-1}, \tag{3.2}$$

where C and C' are absolute constants.

Note that the right side of inequality (3.2) is equivalent to (3.1). Conjecture 3.2 was shown to be true by Talagrand [47] in the two-dimensional case but remains open in higher dimensions $d \geq 3$.

Remark 3.3. The two-sided inequality in Conjecture 3.2 is usually expressed more compactly by writing $\phi(\varepsilon) \simeq \frac{1}{\varepsilon^2}\Big(\log\frac{1}{\varepsilon}\Big)^{2d-1}$ as $\varepsilon \to 0^+$.

Though the precise lower bound appearing in Conjecture 3.2 has not been proven, there is a well-known lower bound on the L^2-norm in dimensions $d \geq 2$. More specifically, E. Csáki [17] has shown that

$$\frac{C}{\varepsilon^2}\left(\log\frac{1}{\varepsilon}\right)^{2d-2} \leq -\log\mathbb{P}\left(\|B\|_{L^2([0,1]^d)} < \varepsilon\right). \tag{3.3}$$

In fact, since $\|\mathbb{B}\|_{L^2} \leq \|\mathbb{B}\|_{L^\infty}$, we have

$$\phi(\varepsilon) \geq \frac{C}{\varepsilon^2}\left(\log\frac{1}{\varepsilon}\right)^{2d-2}, \tag{3.4}$$

where C is an absolute constant. Note that the lower bound of $\phi(\varepsilon)$ in (3.4) is smaller than the lower bound stated in Conjecture 3.2 by a factor of $\log\frac{1}{\varepsilon}$ for sufficiently small $\varepsilon > 0$.

There is a continuous analogue of the conjectured small ball inequality, which we state below, and we will see that the validity of this continuous analogue implies the truth of Conjecture 3.2. Recall that $\mathcal{D}_* = \mathcal{D} \cup \{[-1,1]\}$.

Conjecture 3.4. *In dimensions $d \geq 2$, there exists an orthonormal basis of $L^2([0,1]^d)$, $\{u_R\}$, enumerated using the index set $\mathcal{D}_*^d$, such that, for any choice of coefficients α_R,*

$$n^{\frac{d-2}{2}}\left\|\sum_{|R|=2^{-n}}\alpha_R\eta_R\right\|_\infty \gtrsim 2^{-\frac{3n}{2}}\sum_{|R|=2^{-n}}|\alpha_R|, \tag{3.5}$$

where

$$\eta_R(x_1,\cdots,x_d) := \int_0^{x_1}\cdots\int_0^{x_d} u_R(y_1,\ldots,y_d)dy_1\cdots dy_d,\ (x_1,\cdots,x_d) \in [0,1]^d.$$

Conjecture 3.4 remains open in dimensions $d \geq 3$, but Talagrand's [47] proof of Conjecture 3.2 in dimension $d = 2$ came from proving Conjecture 3.4 in $d = 2$. In his work, he constructed an orthonormal basis $\{u_R\}$ of $L^2([0,1]^2)$, and corresponding functions η_R, from a specific orthonormal basis $\{u_I\}$ of $L^2([0,1])$, and the corresponding functions η_I. More precisely, in dimension 1, he divided a dyadic interval $I = [m \cdot 2^{-k}, (m+1) \cdot 2^{-k})$ into 4 successive dyadic subintervals of I of length $\frac{1}{4}|I|$, *i.e.*,

$$I = \bigcup_{j=1}^{4} I_j, \quad I_j = [(4m+j-1) \cdot 2^{-k-2}, (4m+j) \cdot 2^{-k-2}), \quad j = 1, \dots, 4. \tag{3.6}$$

Note that the centre of the dyadic interval I, denoted by $c(I)$, is equal to $(4m+2) \cdot 2^{-k-2}$. He then considered the function

$$u_I(x) = \frac{1}{|I|^{\frac{1}{2}}}(-\mathbb{1}_{I_1}(x) + \mathbb{1}_{I_2}(x) + \mathbb{1}_{I_3}(x) - \mathbb{1}_{I_4}(x)), \tag{3.7}$$

and after a simple integration over a dyadic interval I, he showed that

$$\eta_I(x) = \begin{cases} -(x - (4m+2)2^{-k-2} + 2^{-k-1})|I|^{-\frac{1}{2}}, & x \in I_1 \\ (x - (4m+2)2^{-k-2})|I|^{-\frac{1}{2}}, & x \in I_2 \cup I_3 \\ (2^{-k-1} - (x - (4m+2)2^{-k-2}))|I|^{-\frac{1}{2}}, & x \in I_4 \\ 0, & \text{otherwise.} \end{cases} \tag{3.8}$$

The function η can be rewritten in the equivalent form $\eta_I(x) = |I|^{\frac{1}{2}} w\left(\frac{x-c(I)}{|I|}\right)$, where $c(I)$ is the centre of the dyadic interval I. Here, w is a continuous non-constant function supported on the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$ with mean zero. The explicit formula of w is given by

$$w(x) = \begin{cases} -(x + \frac{1}{2}), & -\frac{1}{2} \le x \le -\frac{1}{4} \\ x, & -\frac{1}{4} \le x \le \frac{1}{4} \\ -(x - \frac{1}{2}), & \frac{1}{4} \le x \le \frac{1}{2} \\ 0, & \text{otherwise.} \end{cases}$$

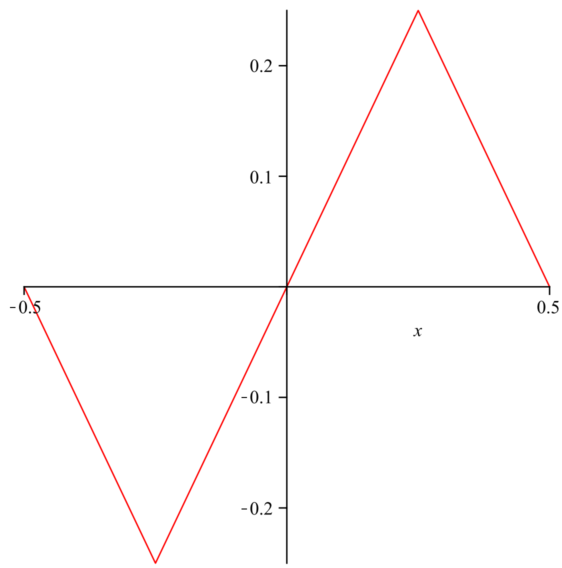


Figure 3.1: The function w used in the proof of the two-dimensional small ball probability conjecture by Talagrand.

In dimensions $d \geq 2$, to a dyadic rectangle $R = R_1 \times \cdots \times R_d \in \mathcal{D}_*^d$, he associated the function

$$u_R(y_1, \ldots, y_d) = u_{R_1}(y_1) \cdots u_{R_d}(y_d) \tag{3.9}$$

so that $\{u_R\}$ formed an orthonormal basis for $L^2([0,1]^d)$. For this special choice of $\{u_R\}$, $\eta_R = \prod_{j=1}^d \eta_{R_j}(x_j)$ where $\eta_{R_j}(x_j) = |R_j|^{\frac{1}{2}} w\left(\frac{x_j - c(R_j)}{|R_j|}\right)$ and $c(R_j)$ is the centre of the dyadic interval R_j, $j = 1, \ldots, d$. Note that $\eta_R = 2^{-n/2} \prod_{j=1}^d w\left(\frac{x_j - c(R_j)}{|R_j|}\right)$ when $R \in \mathcal{A}_n^d$.

Proposition 3.5. *Conjecture 3.4 implies Conjecture 3.2.*

The proof of this proposition has been adapted from [10].

Proof of Proposition 3.5. The proof relies on Levy's [35] representation of a Brownian sheet, $\mathbb{B}$, in terms of a linear combination of independent $\mathcal{N}(0,1)$ random variables, *i.e.*,

$$\mathbb{B}(\vec{s}) = \sum_{k\in\mathbb{N}} \gamma_k \eta_k(\vec{s}),\ \ \vec{s} \in [0,1]^d, \quad \text{in distribution,} \tag{3.10}$$

where $\{\gamma_k\}$ are the independent $\mathcal{N}(0,1)$ random variables,

$$\eta_k(\vec{s}) = \int_0^{s_1} \cdots \int_0^{s_d} u_k(y_1,\ldots,y_d)dy_1\cdots dy_d,\ \ \vec{s} = (s_1,\cdots,s_d) \in [0,1]^d,$$

and $\{u_k\}$ is any orthonormal basis of $L^2([0,1]^d)$. Since there are countably many dyadic rectangles in $[0,1]^d$, there is a bijection between $\mathbb{N}$ and $\mathcal{D}^d_*$. By a slight abuse of notation, we will suppress the bijection and now write γ_R and η_R instead of γ_k and η_k, respectively. We decompose the Brownian sheet $\mathbb{B}$ into two pieces:

$$\mathbb{B} = X_{\vec{t}} + Y_{\vec{t}},$$

where

$$X_{\vec{t}} = \sum_{|R|=2^{-n}} \gamma_R(\omega)\eta_R(\vec{t}), \quad Y_{\vec{t}} = \sum_{|R|\neq 2^{-n}} \gamma_R(\omega)\eta_R(\vec{t}),$$

with $\vec{t} \in [0,1]^d$. Note that $X_{\vec{t}}$ and $Y_{\vec{t}}$ are independent for each $\vec{t} \in [0,1]^d$. We will choose the precise value of the integer n later. By Anderson's lemma, stated in Lemma 3.6 below after the proof of this proposition, we get that

$$\begin{aligned}
\mathbb{P}(\|\mathbb{B}\|_{L^\infty([0,1]^d)} < \varepsilon) \leq & \mathbb{P}\left(\left\| \sum_{|R|=2^{-n}} \gamma_R\eta_R \right\|_\infty < \varepsilon\right) \\
\leq & \mathbb{P}\left(C n^{-\frac{d-2}{2}} 2^{-\frac{3n}{2}} \sum_{|R|=2^{-n}} |\gamma_R| < \varepsilon \right) \\
\leq & \mathbb{P}\Big(\sum_{|R|=2^{-n}} |\gamma_R| < A \Big),
\end{aligned}$$

where $A = \frac{\varepsilon}{C} n^{\frac{d-2}{2}} 2^{\frac{3n}{2}}$ and C is the implicit constant appearing in (3.5) which depends on the dimension d. Here, the conjectured small ball inequality (3.5) was used to get the second inequality. Now, by applying Chebyshev's inequality, we can estimate the distribution function of the sum of the absolute values of iid Gaussians $\{\gamma_R\}$ as

$$\mathbb{P}\Big(\sum_{|R|=2^{-n}} |\gamma_R| < A\Big) \le e^A \mathbb{E} e^{-\sum_{|R|=2^{-n}} |\gamma_R|}. \tag{3.11}$$

To calculate the quantity on the right side of (3.11), we first recall that the formula for the moment generating function of the random variable $|\mathcal{N}(0,1)|$ is given by $M(t) = \mathbb{E}e^{|\mathcal{N}(0,1)|t}$. A straightforward but tedious calculation then shows that $M(t) = 2e^{\frac{t^2}{2}} \int_{-\infty}^{t} \frac{e^{(-1/2)x^2}}{\sqrt{2\pi}} dx$. Also, note that $M(-1) \le 2e^{\frac{1}{2}} \cdot 0.16 \le e^{\frac{1}{2}} e^{-C_0}$, where C_0 is an absolute constant larger than $1/2$. Note that such a constant can be found, since $0.32 < e^{-1/2}$ and e^{-x} is continuous at $x = 1/2$. This, together with the fact that $\{\gamma_R\}$ is a collection of K iid random variables, where $K := \#\{R \in \mathcal{D}^d : |R| = 2^{-n}\} = 2^n \cdot \#\mathbb{H}_n^d \ge C_d 2^n n^{d-1}$, gives us that

$$e^A \mathbb{E} e^{-\sum_{|R|=2^{-n}} |\gamma_R|} = e^A (M(-1))^K \le \exp\left(\frac{1}{C} \varepsilon n^{\frac{d-2}{2}} 2^{\frac{3n}{2}} - C_1 2^n n^{d-1}\right), \tag{3.12}$$

where $C_1 = \big(C_0 - \frac{1}{2}\big) C_d$ is a positive constant which depends only on the dimension d. Now, for small $\varepsilon > 0$, if we let n be a positive integer satisfying

$$\frac{1}{2} C C_1 2^{-\frac{n+1}{2}} (n+1)^{\frac{d}{2}} \le \varepsilon \le \frac{1}{2} C C_1 2^{-\frac{n}{2}} n^{\frac{d}{2}}, \tag{3.13}$$

then the right side of inequality (3.13) implies that

$$\frac{1}{C} \varepsilon n^{\frac{d-2}{2}} 2^{\frac{3n}{2}} \le \frac{1}{2} C_1 2^n n^{d-1}. \tag{3.14}$$

This means that

$$\exp\left(\frac{1}{C} \varepsilon n^{\frac{d-2}{2}} 2^{\frac{3n}{2}} - C_1 2^n n^{d-1}\right) \le \exp\Big(-\frac{1}{2} C_1 2^n n^{d-1}\Big). \tag{3.15}$$

Note that the left side of (3.13) implies that $\log \frac{1}{\varepsilon} \lesssim n$. Hence,

$$\begin{aligned}\mathbb{P}\Big(\sum_{|R|=2^{-n}} |\gamma_R| < A\Big) &\le \exp\Big(-\frac{1}{2}C_1 2^n n^{d-1}\Big)\\ &\le \exp\Big(-\frac{1}{2}C_1 \frac{(2^{\frac{n}{2}})^2}{n^d} n^{2d-1}\Big)\\ &\le \exp\bigg(-\frac{C'}{\varepsilon^2}\Big(\log\frac{1}{\varepsilon}\Big)^{2d-1}\bigg),\end{aligned}$$

where C' is a positive constant. Therefore,

$$\mathbb{P}(\|\mathbb{B}\|_{L^\infty([0,1]^d)} < \varepsilon) \le \exp\bigg(-\frac{C'}{\varepsilon^2}\Big(\log\frac{1}{\varepsilon}\Big)^{2d-1}\bigg).$$

That is,

$$\phi(\varepsilon) \gtrsim \frac{1}{\varepsilon^2}\Big(\log\frac{1}{\varepsilon}\Big)^{2d-1}.$$

□

The following lemma, which was used in the proof of Proposition 3.5, is a standard fact in probability theory. We state it here without proof.

Lemma 3.6. (Anderson's Lemma [1]). Let $X_t,\ Y_t,\ t \in T$ be independent centered Gaussian random processes. Then

$$\mathbb{P}\Big(\sup_{t\in T} |X_t + Y_t| < c\Big) \le \mathbb{P}\Big(\sup_{t\in T} |X_t| < c\Big).$$

3.2 Connection to approximation theory

We begin this section by recalling some notions from approximation theory which will be used throughout.

Definition 3.7. (a) Let $\hat{\mathcal{K}}$ be a subset of a metric space $(\mathcal{K}, \rho)$. The set $\hat{\mathcal{K}}$ is called an ε-net for $\mathcal{K}$, if for every $x \in \mathcal{K}$ there exists $y \in \hat{\mathcal{K}}$ such that $\rho(x, y) \le \varepsilon$.

(b) Points $y_1, \dots, y_m \in \mathcal{K}$ are called ε-separated if $\rho(y_i, y_j) > \varepsilon$ for all $i \ne j$.

Note that if $(\mathcal{K}, \rho)$ is a compact metric space, then $\mathcal{K}$ contains a finite ε-net for every $\varepsilon > 0$.

Definition 3.8. Let $\tilde{\mathcal{K}}$ be a compact subset of a metric space $(\mathcal{K}, \rho)$. Then the *covering number* of the set $\tilde{\mathcal{K}}$, $N(\varepsilon, \tilde{\mathcal{K}})$, is the minimum of the cardinalities of ε-nets for $\tilde{\mathcal{K}}$, *i.e.*,

$$N(\varepsilon, \tilde{\mathcal{K}}) := \min \left\{ N : \exists \{x_k\}_{k=1}^N \subset \tilde{\mathcal{K}},\ \tilde{\mathcal{K}} \subset \bigcup_{k=1}^N B(x_k, \varepsilon) \right\},$$

where $B(x_k, \varepsilon) = \{x \in \tilde{\mathcal{K}} : \rho(x, x_k) \leq \varepsilon\}$ is the ball centred at x_k of radius ε. The ε-*entropy* of the set $\tilde{\mathcal{K}}$ is then defined by

$$H(\varepsilon, \tilde{\mathcal{K}}) := \log(N(\varepsilon, \tilde{\mathcal{K}})).$$

Definition 3.9. Let $(\mathcal{K}, \rho)$ be a compact metric space. Given $\varepsilon > 0$, we define the packing number, $\mathcal{M}(\varepsilon, \mathcal{K})$, of the set $\mathcal{K}$ to be

$$\mathcal{M}(\varepsilon, \mathcal{K}) = \max\{m \in \mathbb{N} : \exists\ \varepsilon\text{-separated points } y_1, \dots, y_m \in \mathcal{K}\}.$$

The study of covering numbers and entropies goes back A. Kolmogorov. Around 1950, Kolmogorov developed a strong interest in problems in information theory. Inspired by Shannon's ideas, he introduced the notion of a covering number to measure the size of a set, A, in a metric space. From the point of view of information theory, the entropy of A has the following usage. Consider the elements of A as "messages" encoded in binary strings of 0's and 1's. Suppose that $\varepsilon > 0$ is a number small enough so that any message x in A can be recovered with reasonable accuracy from some x' in A with $\rho(x, x') \leq \varepsilon$. One should think of x as the original and x' as the transmitted message. The quantity $N(\varepsilon, A)$ then captures the economic yet effective collection of transmittable messages.

Kolmogorov used the concept of covering numbers to give a simpler proof of Vitushkin's theorem on superposition of functions. To do this, he found an upper bound for the ε-entropy of $W_\infty^r(I^n)$, the class of functions of n variables defined on the unit cube I^n in $\mathbb{R}^n$ and having uniformly bounded partial derivatives up to order r. In particular, he showed that

$$\log(N(\varepsilon, W_\infty^r(I^n))) \leq C\left(\frac{1}{\varepsilon}\right)^{n/r}.$$

This was the starting point of investigating the ε-entropy of compact sets in function spaces. We refer the interested reader to [32] for a thorough description of the development and history of covering numbers.

The ε-entropy of a set A also has applications in approximation theory, specifically, linear approximations. To present these concepts in an abstract setting, we must first introduce the notion of the *n-width* of a set, A, in a normed space, another tool invented by Kolmogorov to measure the size of A. Let X be a normed space equipped with a norm $\|\cdot\|$, and let A be a subset of X. The *n-width* of the set A, $d_n(A)$, is defined to be

$$d_n(A) := \inf_{L_n} \sup_{x\in A} \inf_{\xi\in L_n} \|x-\xi\|,$$

where the outer infimum is taken over all n-dimensional subspaces $L_n \subset X$. We will refer to the quantity $\inf_{\xi\in L_n}\|x-\xi\|$ as the deviation of the approximation of x by L_n, $\sup_{x\in A}\inf_{\xi\in L_n}\|x-\xi\|$ as the deviation of the approximation of the set A by L_n, and $d_n(A)$ as the optimal deviation of the approximation of A. For small $\varepsilon > 0$, in order to ensure that $d_n(A) < \varepsilon$, n must be at least as large as some simple lower bound depending on ε, and is, roughly, $H(\varepsilon, A)$. More precisely, G. Lorentz [36] showed that

$$n \geq \frac{H(e^2\varepsilon, A) - C_0}{2 + \log\frac{1}{\varepsilon} + \log[H(e^2\varepsilon, A) - C_0]},$$

where C_0 is an absolute constant and A is a compact set in a separable Banach space X. From the above inequality, we see that, in order to see how n depends on ε for small $\varepsilon > 0$, one needs to understand how H depends on ε.

Here, we survey some results concerning the asymptotic behaviour of $\sigma_p(\varepsilon) := H(\varepsilon, B(MW^p))$, the ε-entropy of a specific class of functions with bounded mixed derivatives. As we will see, this asymptotic behaviour can be determined explicitly if one can prove the continuous version of the small ball conjecture. This topic gained a lot of attention after the connection between the entropy of this class of functions with bounded mixed derivatives and small deviation probabilities was discovered (see Theorem 3.12).

Next, we define a specific integration operator that will be used to give the precise definition of our space of functions with bounded mixed derivatives.

Definition 3.10. The integration operator $\mathcal{T}_d$, acting on functions defined on the unit cube $[0,1]^d$, is defined by

$$(\mathcal{T}_d f)(x_1, \ldots, x_d) := \int_0^{x_1} \cdots \int_0^{x_d} f(y_1, \ldots, y_d) dy_1 \cdots dy_d. \tag{3.16}$$

Let B_∞ and $B(L^p)$ denote the unit ball in $L^\infty([0,1]^d)$ and $L^p([0,1]^d)$ respectively. Then, $MW^p([0,1]^d) := \mathcal{T}_d(L^p([0,1]^d))$ is the space of functions on $[0,1]^d$ with mixed derivatives $\frac{\partial^d f}{\partial x_1 \partial x_2 \cdots \partial x_d}$ in L^p, which, throughout, we will equip with the sup-norm, $\|\cdot\|_\infty$. We then define the subset $B(MW^p) \subset MW^p$ by $B(MW^p) := \mathcal{T}_d(B(L^p))$. To define the covering number of the set $B(MW^p)$, we must first show that $B(MW^p)$ is a precompact set with respect to the L^∞-norm.

Proposition 3.11. *The closure of the set* $B(MW^p)$, $\overline{B(MW^p)}^{\|\cdot\|_\infty}$, *is a compact set with respect to the* L^∞*-metric for* $1 < p < \infty$ *and for all* $d \geq 1$.

Proof. To conclude that $\overline{B(MW^p)}^{\|\cdot\|_\infty}$ is a compact set, we will use the Arzelà-Ascoli theorem. To do this, we need to show that

$$\mathcal{T}_d(B(L^p)) \subset B_\infty(0, c) \tag{3.17}$$

for some positive constant c, where $B_\infty(0,c)$ is the ball centred at zero of radius c measured in L^∞-norm, and that the family of functions $\{\mathcal{T}_d(f)\}_{f \in B(L^p)}$ is uniformly equicontinuous, *i.e.*,

$$\forall \varepsilon > 0,\ \exists \delta > 0 \text{ such that } |\mathcal{T}_d(f)(\vec{s}) - \mathcal{T}_d(f)(\vec{t})| < \varepsilon, \forall \vec{s}, \vec{t} \in [0,1]^d \text{ with} \tag{3.18}$$

$$\|\vec{s}-\vec{t}\|_2 := \Big(\sum_{i=1}^{d} |s_i - t_i|^2\Big)^{1/2} < \delta \text{ and } \forall f \in B(L^p).$$

First, for any $f \in B(L^p)$, we have that

$$|\mathcal{T}_d(f)(\vec{t})| \leq \int_0^{t_1} \cdots \int_0^{t_d} |f(y)| dy \leq \|f\|_p \leq 1,$$

for every $\vec{t} \in [0,1]^d$. This shows that (3.17) holds.

Now we will prove condition (3.18). To do this, we will show that

$$|\mathcal{T}_d(f)(\vec{s}) - \mathcal{T}_d(f)(\vec{t})| \leq \sum_{i=1}^{d} |t_i - s_i|^{1/q} \quad \forall \vec{t}, \vec{s} \in [0,1]^d \text{ and } \forall f \in B(L^p), \tag{3.19}$$

where $q = \frac{p}{p-1}$ is the dual exponent of p, as (3.18) follows from (3.19) immediately. Suppose $d \geq 1$ and $\vec{t}, \vec{s} \in [0,1]^d$ with $t_i \neq s_i$ for $i = 1, \ldots d$. For any such $\vec{s}$ and $\vec{t}$, we define the following sets:

$$I_1 := \{i \in \{1, \ldots, d\} : s_i > t_i\},$$
$$I_2 := \{i \in \{1, \ldots, d\} : s_i < t_i\}.$$

Note that these sets depend on the choice of $\vec{s}, \vec{t} \in [0,1]^d$, and that $I_1 \cup I_2 = \{1, \ldots, d\}$. Let $[\vec{0}, \vec{s}] := \prod_{i=1}^{d} [0, s_i]$ and $[\vec{0}, \vec{t}] := \prod_{i=1}^{d} [0, t_i]$. Note that the corner that is furthest from the origin of each of these d-dimensional boxes is given by $\vec{s}$ and $\vec{t}$ respectively, and that the intersection of these boxes is non-empty. In fact,

$$[\vec{0}, \vec{s}] \bigcap [\vec{0}, \vec{t}] = \prod_{i=1}^{d} \Big[0, \min(t_i, s_i)\Big],$$

where

$$\min(t_i, s_i) = s_i, \ \forall i \in I_2 \quad \text{and} \quad \min(t_i, s_i) = t_i \ \forall i \in I_1.$$

We have that

$$[\vec{0},\vec{s}] = \left([\vec{0},\vec{s}]\bigcap[\vec{0},\vec{t}]\right)\bigcup\left(\bigcup_{i\in I_1} A_i\right), \tag{3.20}$$

where $\bigcup_{i\in I_1} A_i$ is the complement of the set $[\vec{0},\vec{s}]\bigcap[\vec{0},\vec{t}]$ in $[\vec{0},\vec{s}]$, and each A_i depends on the choice of $\vec{s}$ and $\vec{t}$. In fact, for $i \in I_1$, $A_i = B_1\times\cdots\times B_d$ with $B_j = [0, s_j] \subset [0,1]$, $j \neq i$, and $B_i = (t_i, s_i]$. Moreover,

$$[\vec{0},\vec{t}] = \left([\vec{0},\vec{s}]\bigcap[\vec{0},\vec{t}]\right)\bigcup\left(\bigcup_{i\in I_2} \tilde{A}_i\right), \tag{3.21}$$

where, for $i \in I_2$, $\tilde{A}_i = \tilde{B}_1\times\cdots\times\tilde{B}_d$ with $\tilde{B}_j = [0, t_j] \subset [0,1]$, $j \neq i$, and $\tilde{B}_i = (s_i, t_i]$. Note that the complement of the set $[\vec{0},\vec{s}]\bigcap[\vec{0},\vec{t}]$ in $[\vec{0},\vec{t}]$ is $\bigcup_{i\in I_2}\tilde{A}_i$. Therefore, using relations (3.20), (3.21), and Hölder's inequality, we obtain

$$\begin{aligned} |\mathcal{T}_d(f)(\vec{s}) - \mathcal{T}_d(f)(\vec{t})| &= \left|\int_{[\vec{0},\vec{s}]} f(y)dy - \int_{[\vec{0},\vec{t}]} f(y)dy\right| \\ &\leq \sum_{i\in I_1}\int_{A_i}|f(y)|dy + \sum_{i\in I_2}\int_{\tilde{A}_i}|f(y)|dy \\ &\leq \sum_{i\in I_1}\|f\|_p|t_i - s_i|^{1/q} + \sum_{i\in I_2}\|f\|_p|t_i - s_i|^{1/q} \quad (*) \\ &\leq \sum_{i=1}^{d}|t_i - s_i|^{1/q}, \end{aligned} \tag{3.22}$$

for any $f \in B(L^p)$, for all $\vec{s},\vec{t} \in [0,1]^d$ with $s_i \neq t_i,\ i = 1,\ldots,d$.

From (3.22), we see that (3.19) holds for points $\vec{s},\vec{t} \in [0,1]^d$ with $s_i \neq t_i,\ i = 1,\ldots,d$. Now, we will show that (3.19) holds for general points $\vec{s},\vec{t} \in [0,1]^d$. Fix any two points $\vec{s}$ and $\vec{t}$ in $[0,1]^d$ and consider the set

$$I_0 = \{i \in \{1,\ldots,d\} : s_i = t_i\},$$

i.e., I_0 is the subset of indices in $\{1,\ldots,d\}$ for which the corresponding coordinates of the vectors $\vec{s}$ and $\vec{t}$ coincide. Let I_0^c denote the complement of the set I_0 in $\{1,\ldots,d\}$.

Now, let

$$S_{I_0} := \prod_{i \in I_0} [0, s_i], \; S_{I_0^c} := \prod_{i \in I_0^c} [0, s_i], \text{ and } T_{I_0^c} = \prod_{i \in I_0^c} [0, t_i]$$

be boxes corresponding to I_0 and I_0^c. We will let $y_{S_{I_0}}, y_{S_{I_0^c}}$, and $y_{T_{I_0^c}}$ denote points belonging to S_{I_0}, $S_{I_0^c}$, and $T_{I_0^c}$ respectively. Using Fubini's theorem, we get that

$$\begin{aligned} |\mathcal{T}_d(f)(\vec{s}) - \mathcal{T}_d(f)(\vec{t})| &= \left| \int_{S_{I_0}} \left[\int_{S_{I_0^c}} f(y) dy_{S_{I_0^c}} - \int_{T_{I_0^c}} f(y) dy_{T_{I_0^c}} \right] dy_{S_{I_0}} \right| \\ &\leq \int_{S_{I_0}} \left| \int_{S_{I_0^c}} f(y) dy_{S_{I_0^c}} - \int_{T_{I_0^c}} f(y) dy_{T_{I_0^c}} \right| dy_{S_{I_0}}. \end{aligned}$$

In order to estimate the absolute value of the difference of the two integrals in the above inequality, we appeal to the case we considered previously. Now, for $y_{S_{I_0}} \in S_{I_0}$, with the aid of $(*)$ in (3.22), we get that

$$\begin{aligned} \left| \int_{S_{I_0^c}} f(y) dy_{S_{I_0^c}} - \int_{T_{I_0^c}} f(y) dy_{T_{I_0^c}} \right| \leq & \sum_{i \in I_1} \left(\int_{S_{I_0^c}} |f(y)|^p dy_{S_{I_0^c}} \right)^{1/p} |t_i - s_i|^{1/q} + \\ & \sum_{i \in I_2} \left(\int_{T_{I_0^c}} |f(y)|^p dy_{T_{I_0^c}} \right)^{1/p} |t_i - s_i|^{1/q}. \end{aligned}$$

Here, I_1 and I_2 are subsets of I_0^c for which $I_0^c = I_1 \cup I_2$. Finally, integrating the above inequality with respect to $y_{S_{I_0}}$ and applying Jensen's inequality gives

$$\begin{aligned} |\mathcal{T}_d(f)(\vec{s}) - \mathcal{T}_d(f)(\vec{t})| &\leq \sum_{i \in I_1} \|f\|_p |t_i - s_i|^{1/q} + \sum_{i \in I_2} \|f\|_p |t_i - s_i|^{1/q} \\ &\leq \|f\|_p \sum_{i \in I_0^c} |t_i - s_i|^{1/q} \leq \sum_{i=1}^{d} |t_i - s_i|^{1/q} \end{aligned}$$

for any $f \in B(L^p)$ and any $\vec{t}, \vec{s} \in [0, 1]^d$. This completes the proof of the proposition.

□

One branch of research in approximation theory is concerned with determining the asymptotic behaviour of $\sigma_p(\varepsilon) := \log(N(\varepsilon, p, d))$ as $\varepsilon \to 0^+$, where

$$N(\varepsilon, p, d) := \min \left\{ N : \exists \{x_k\}_{k=1}^N \subset B(MW^p),\ B(MW^p) \subset \bigcup_{k=1}^N \left(x_k + \varepsilon B_\infty\right) \right\}$$

is the covering number of the set $B(MW^p)$. Recall that this is the smallest number of balls of radius ε needed to cover the set $B(MW^p)$, or, equivalently, the size of the smallest ε-net of $B(MW^p)$. There is an interesting relation between the covering number $N(\varepsilon, p, d)$ and the packing number $\mathcal{M}(\varepsilon, B(MW^p))$. Here, for convenience, we will denote the packing number of $B(MW^p)$ by $\mathcal{M}(\varepsilon, p, d)$ rather than $\mathcal{M}(\varepsilon, B(MW^p))$. In particular, it is easy to show that

$$\mathcal{M}(2\varepsilon, p, d) \le N(\varepsilon, p, d) \le \mathcal{M}(\varepsilon, p, d). \tag{3.23}$$

J. Kuelbs and V. Li [33] have established a deep connection between the asymptotic behaviours of $\phi(\varepsilon)$ and $\sigma_p(\varepsilon)$. In particular, from their results, one can conclude the following:

Theorem 3.12. *Suppose $d \ge 2$, $b > 0$, and $\varepsilon > 0$ is small. Then*

$$\phi(\varepsilon) \simeq \varepsilon^{-2} \left(\log \frac{1}{\varepsilon}\right)^b \quad \text{if and only if} \quad \sigma_2(\varepsilon) \simeq \varepsilon^{-1} \left(\log \frac{1}{\varepsilon}\right)^{\frac{b}{2}}.$$

Here, the symbol " $\simeq$ " is as in Remark 3.3. As mentioned in the previous section, the upper bound for $\phi(\varepsilon)$ has been proved for the special case when $b = 2d - 1$. So, according to Theorem 3.12,

$$\sigma_2(\varepsilon) \le C\varepsilon^{-1} \left(\log \frac{1}{\varepsilon}\right)^{\frac{2d-1}{2}}$$

for small $\varepsilon > 0$, where C is an absolute constant. Thus, in the setting of approximation theory, the small ball conjecture is equivalent to:

Conjecture 3.13. *For all $d \geq 2$ and $\varepsilon > 0$ small, we have that*

$$\sigma_2(\varepsilon) \geq C'\varepsilon^{-1}\Big(\log\frac{1}{\varepsilon}\Big)^{\frac{2d-1}{2}},$$

where C' is an absolute constant.

Conjecture 3.13 was solved in dimension $d = 2$ by Kuelbs and Li [33]. They were able to prove the conjecture in $d = 2$ by using Talagrand's result on the asymptotic behaviour of $\phi(\varepsilon)$ in $d = 2$ and Theorem 3.12. Also, Temlyakov [50] gave a second proof of Conjecture 3.13 in $d = 2$ by using the analogue of the small ball inequality for trigonometric polynomials. In particular, he showed that, in $d = 2$, for all $1 < p < \infty$, and for $\varepsilon > 0$ small,

$$\sigma_p(\varepsilon) \geq C_p\varepsilon^{-1}\Big(\log\frac{1}{\varepsilon}\Big)^{\frac{3}{2}}.$$

Conjecture 3.13 remains open in $d \geq 3$, though, using Theorem 3.12, it's not hard to see that Conjecture 3.13 would follow from a proof of Conjecture 3.4 in $d \geq 3$. However, there is a way to deduce Conjecture 3.13 from Conjecture 3.4 without invoking Theorem 3.12, which is what we do here.

Proposition 3.14. *If Conjecture 3.4 is true, then so is Conjecture 3.13.*

In order to prove Proposition 3.14, we will need a preliminary result. Consider a bi-valued function $\beta : \mathcal{A}_n^d \to \{-1, 1\}$, $\mathcal{A}_n^d = \{R \in \mathcal{D}^d, |R| = 2^{-n}\}$. Let $K = \#\mathcal{A}_n^d$. Then, by defining $\beta_R := \beta(R)$ for $R \in \mathcal{A}_n^d$, we can associated β with a vector $\vec{\beta} \in \{-1, 1\}^K$ whose "Rth" coordinate is β_R . It has already been noted that $K = 2^n\#\mathbb{H}_n^d \simeq 2^n n^{d-1}$.

Lemma 3.15. *There is an integer $n_0 \geq 1$ such that, if $n \geq n_0$ and $K = 2^n\#\mathbb{H}_n^d$, then there exists a set $X \subset \{-1, 1\}^K$ satisfying:*

1. $\sum_{|R|=2^{-n}} |\beta'_R - \beta_R| \gtrsim K$ *for any* $\vec{\beta} \neq \vec{\beta}' \in X$.

2. $\log\#X \gtrsim K$.

The proof of Lemma 3.15 will follow the proof of Proposition 3.14.

Proof of Proposition 3.14. According to inequality (3.23), for a fixed small $\varepsilon > 0$, we need to produce a 2ε-separated collection of functions $\tilde{\mathcal{F}} = \{\tilde{F}_{\vec{\beta}}\} \subset B(MW^2)$ for which $\log \#\tilde{\mathcal{F}} \gtrsim \frac{1}{\varepsilon}\Big(\log \frac{1}{\varepsilon}\Big)^{\frac{2d-1}{2}}$.

Now, to each such function β or equivalently vector $\vec{\beta} \in \{-1,1\}^K$, we can further associate the function $\tilde{F}_{\vec{\beta}}$ given by

$$\tilde{F}_{\vec{\beta}} = \frac{c}{2^{\frac{n}{2}} n^{\frac{d-1}{2}}} \sum_{|R|=2^{-n}} \beta_R \eta_R, \tag{3.24}$$

where $c > 0$ is a small constant and $\eta_R := \mathcal{T}_d u_R$, as given in the statement of Conjecture 3.4. Moreover, $\tilde{F}_{\vec{\beta}} \in B(MW^2)$ for every $\vec{\beta} \in \{-1,1\}^K$. Indeed, since $\eta_R = \mathcal{T}_d u_R$ and $\{u_R\}_{R\in\mathcal{A}_n^d}$ is an orthonormal set of functions, we obtain

$$\tilde{F}_{\vec{\beta}} = \mathcal{T}_d\Bigg(\frac{c}{2^{\frac{n}{2}} n^{\frac{d-1}{2}}} \sum_{|R|=2^{-n}} \beta_R u_R \Bigg),$$

and

$$\Bigg\| \frac{c}{2^{\frac{n}{2}} n^{\frac{d-1}{2}}} \sum_{|R|=2^{-n}} \beta_R u_R \Bigg\|_2^2 = \frac{c^2}{2^n n^{d-1}} 2^n \cdot \#\mathbb{H}_n^d \leq 1$$

for c sufficiently small.

Let $X \subset \{-1,1\}^K$ be a set satisfying the conditions of Lemma 3.15. Then, $\tilde{\mathcal{F}} = \{\tilde{F}_{\vec{\beta}}\}_{\vec{\beta}\in X}$ is a 2ε-separated collection of functions in $B(MW^2)$ for which

$$\log \#\tilde{\mathcal{F}} \simeq \frac{1}{\varepsilon}\Big(\log \frac{1}{\varepsilon}\Big)^{\frac{2d-1}{2}} \quad \text{as } \varepsilon \to 0^+,$$

for a suitable choice of $n \geq n_0$. Indeed, let $\varepsilon > 0$ be any small number strictly less than $\frac{1}{2} n_0^{1/2} 2^{-n_0}$, where n_0 is the integer which appears in Lemma 3.15, and let $n \geq n_0$ be an integer for which $n^{1/2}2^{-n} \geq 2\varepsilon \geq (n+1)^{1/2}2^{-(n+1)}$. First, we show that $\tilde{\mathcal{F}}$ is a 2ε-separated family with respect to the L^∞-norm. From Conjecture 3.4, if $\vec{\beta}, \vec{\beta}' \in X$

are such that $\vec{\beta} \neq \vec{\beta}'$, then we get that

$$\begin{aligned}
\|\tilde{F}_{\vec{\beta}} - \tilde{F}_{\vec{\beta}'}\|_\infty =& \left\| \frac{c}{2^{\frac{n}{2}} n^{\frac{d-1}{2}}} \sum_{|R|=2^{-n}} (\beta_R - \beta'_R)\eta_R \right\|_\infty \\
\gtrsim& 2^{-\frac{n}{2}} n^{-\frac{d-1}{2}} \cdot n^{-\frac{d-2}{2}} 2^{-\frac{3n}{2}} \sum_{|R|=2^{-n}} |\beta_R - \beta'_R| \\
\gtrsim& n^{-\frac{2d-3}{2}} 2^{-2n} \cdot 2^n \cdot n^{d-1} = n^{1/2} 2^{-n}.
\end{aligned} \tag{3.25}$$

Hence, $\tilde{\mathcal{F}}$ is a 2ε-separated family with respect to the L^∞-norm. Since $2\varepsilon \geq (n+1)^{1/2} 2^{-(n+1)}$, we therefore have that

$$\begin{aligned}
\log \#\tilde{\mathcal{F}} = \log \# X \gtrsim K &\gtrsim 2^n n^{d-1} \\
&\gtrsim 2^n n^{-\frac{1}{2}} n^{d-\frac{1}{2}} \\
&\gtrsim \frac{1}{\varepsilon}\left(\log \frac{1}{\varepsilon} \right)^{d-\frac{1}{2}}.
\end{aligned}$$

□

Now, to prove Lemma 3.15, it will be useful to first introduce some notions from coding theory.

In coding theory, a subset $X \subset \{-1, 1\}^K$ is called a binary code of length K, and we can endow such a binary code with the Hamming distance. This is a distance function, d_H, defined on $X \times X$ by

$$d_H(x, y) = \#\{j = 1, \ldots, K : x_j \neq y_j\},$$

i.e., the Hamming distance between x and y is the number of components in which x and y do not coincide. It is easy to verify that (X, d_H) is, in fact, a metric space. The minimum Hamming distance of a binary code X is then defined to be the smallest Hamming distance between its elements, $\min_{x,y \in X, x \neq y} d_H(x, y)$. So, in the language of coding theory, to prove Lemma 3.15, we must find a binary code X of length

$K = 2^n \# \mathbb{H}_n^d$, such that

$$\min_{\vec{\beta}, \vec{\beta}' \in X, \vec{\beta} \neq \vec{\beta}'} d_H(\vec{\beta}, \vec{\beta}') \geq \frac{K}{4}, \tag{3.26}$$

and

$$\log \# X \gtrsim K. \tag{3.27}$$

Since (3.26) tells us that $\sum_{|R|=2^{-n}} |\beta'_R - \beta_R| \geq 2\frac{K}{4} \gtrsim K$ for any $\vec{\beta} \neq \vec{\beta}' \in X$, it's clear that (3.26) implies the first condition on X in Lemma 3.15.

It is easy to show that there exists a binary code X satisfying (3.26). If this binary code X is equal to $\{-1, 1\}^K$, then $\# X = 2^K$, and (3.27) follows immediately. In general, X will be a strict subset of $\{-1, 1\}^K$, and so $\# X$ will be some fraction of 2^K. Gilbert-Varshamov's theorem, stated below, provides us with a lower bound of this fraction, and this lower bound will be used in the proof of (3.27).

Theorem 3.16. (Gilbert-Varshamov's theorem) Let $A(m, k)$ denote the maximal size of a binary code of length m and minimum Hamming distance at least k. Then

$$A(m, k) \geq \frac{2^m}{\sum_{j=0}^{k-1} \binom{m}{j}}.$$

We will postpone the proof of Gilbert-Varshamov's theorem until we have completed the proof of Lemma 3.15.

Proof of Lemma 3.15. We will apply Gilbert-Varshamov's theorem with $m = K$ and $k = \frac{m}{4}$ in order to show that $\log \# X \gtrsim K$. First, note that there is a maximal code X of length m and minimum Hamming distance $\frac{m}{4}$, so Gilbert-Varshamov's theorem is applicable. Using Stirling's formula, $m! \simeq \frac{1}{\sqrt{2\pi m}} \left(\frac{m}{e}\right)^m$, we obtain

$$\binom{m}{m/4} \simeq \frac{1}{\sqrt{m}} \left(\frac{1}{4}\right)^{-\frac{m}{4}} \left(\frac{3}{4}\right)^{-3\frac{m}{4}}.$$

Applying Gilbert-Varshamov's theorem and recalling the fact that $\binom{m}{j}$ is an increasing function of j when $0 \leq j < k$, we get that

$$\begin{aligned}
\#X &\geq \frac{2^m}{\sum_{j=0}^{k-1} \binom{m}{j}} \geq \frac{2^m}{k\binom{m}{k}} = \frac{2^m}{m/4\binom{m}{m/4}} \\
&\simeq \frac{1}{\sqrt{m}} 2^m \left(\frac{1}{4}\right)^{m/4} \left(\frac{3}{4}\right)^{3m/4} \gtrsim \frac{1}{\sqrt{m}} C^m \gtrsim \frac{b^m}{\sqrt{m}} \frac{C^m}{b^m} \\
&\gtrsim \tilde{C}^m,
\end{aligned}$$

for a sufficiently large value of m, or, equivalently, for a big enough integer n. Here, $C = 2(1/4)^{1/4}(3/4)^{3/4} > 1$, b is any number satisfying $1 < b < C$, and $\tilde{C} = \frac{C}{b}$. Hence, $\log \#X \gtrsim m$. □

Finally, to conclude the section, we prove Gilbert-Varshamov's theorem.

Proof of Gilbert-Varshamov's Theorem. First, note that, given any $x \in \{-1,1\}^m$, $\#\{y \in X : d_H(x,y) = j\} = \binom{m}{j}$ for $j = 1, \ldots, m$. If we let $B_H(x,k)$ denote the Hamming ball centred at x of the radius k, *i.e.*, $B_H(x,k) = \{y \mid d_H(x,y) < k\}$, then $\#B_H(x,k) = \sum_{j=0}^{k-1} \binom{m}{j}$. Let X be a maximal code of length m and minimum Hamming distance at least k. Then $\bigcup_{x\in X} B_H(x,k) = \{-1,1\}^m$. As a result, we obtain

$$2^m = \# \bigcup_{x\in X} B_H(x,k) \leq \sum_{x\in X} \#B_H(x,k) = \#X \sum_{j=0}^{k-1} \binom{m}{j},$$

which completes the proof. □

3.3 Connection to discrepancy theory

3.3.1 Conjectures in discrepancy theory

Discrepancy theory studies the deviation of the actual data from the ideal data. One of the most challenging problems in geometric discrepancy theory is, given a set $\mathcal{P}_N = \{p_1, p_2, \ldots, p_N\} \subset [0,1]^d$ with cardinality $\#\mathcal{P}_N = N$, to obtain a lower bound on the discrepancy function,

$$D_N(x_1, \dots, x_d) = \#\{\mathcal{P}_N \cap [0, x_1) \times \cdots \times [0, x_d)\} - N x_1 \cdots x_d, \tag{3.28}$$

with respect to some norm. Relation (3.28) represents the difference between the actual and expected number of points in the box $[0, x_1) \times \cdots \times [0, x_d)$. When the points $p_1, \dots, p_N$ are uniformly distributed, these quantities are the same and so $D_N(x_1, \dots, x_d) = 0$. Lower bounds of the discrepancy function with respect to the L^∞-norm indicate the best case scenario for the distribution of points, *i.e.*, how close to uniformly distributed the distribution of the points can be. Motivation for studying the behaviour of the discrepancy function comes from numerical integration. In numerical integration, one attempts to approximate integrals of functions $f : [0,1]^d \to \mathbb{R}$ that cannot be computed analytically. Moreover, while doing this, one wants to minimize the error in their approximation. The Koksma-Hlawka inequality connects the error, which appears in the process of the numerical integration, with the discrepancy function. This inequality says that, for an N-point set $P = \{p_1, \dots, p_N\} \subset [0,1]^d$, and for all $p \in [1, \infty]$,

$$\left| \int_{[0,1]^d} f(x)dx - \frac{1}{N} \sum_{i=1}^{N} f(p_i) \right| \leq V_q(f) \frac{\|D_N\|_p}{N}. \tag{3.29}$$

Here, q is the dual exponent of p, and $V_q(f)$ is a quantity depending on f and p. In the one-dimensional case, $d = 1$, $V_q(f) = \|f'\|_q$ when f is a smooth function. The interested reader can consult [31, 27] for the precise definition of $V_q(f)$ in dimensions $d \geq 2$. From inequality (3.29), we see that an N-point set with smallest possible L^p-norm gives rise to the smallest error. Hence, we want to find an N-point set $\mathcal{P}_N \subset [0,1]^d$ which minimizes the L^p-norm of D_N. This motivates the study of uniform lower bounds of the discrepancy function D_N. The first quantitative result in this direction, stated below, was obtained by K. F. Roth [39]. Note that we will prove Roth's theorem in the subsequent subsection.

Theorem 3.17. (Roth, 1954) In all dimensions $d \geq 2$, for any N-point set $\mathcal{P}_N \subset [0,1]^d$, one has

$$\|D_N\|_2 \geq C_d (\log N)^{\frac{d-1}{2}}, \tag{3.30}$$

where C_d is an absolute constant that depends only on the dimension d.

Furthermore, inequality (3.30) in Theorem 3.17 is sharp in the sense that there exists an N-point set $\mathcal{P}_N$ for which $\|D_N\|_2 \leq C_d(\log N)^{\frac{d-1}{2}}$. This fact was shown by Roth [40, 41], Chen [14] and Frolov [20]. The best known constant C_d in (3.30) was obtained by A. Hinrichs and L. Markhasin [26]. In general, the L^p-behaviour of the discrepancy function is well understood for $1 < p < \infty$. In fact, the L^p-behaviour of the discrepancy function is similar to the L^2-behaviour, *i.e.*, $\|D_N\|_p \geq C(d,p)(\log N)^{\frac{d-1}{2}}$, for $1 < p < \infty$. This was first shown by W. M. Schmidt [42, 43]. He originally considered the values of p for which the dual exponent q is an even integer using Roth's approach. One can also obtain such lower bounds on the L^p-norm of D_N for the full range of $1 < p < \infty$ by using the Littlewood-Paley inequalities. This will be explained in more detail after proving Roth's theorem and after introducing the Littlewood-Paley inequalities. In addition, examples of N-point sets with $\|D_N\|_p \leq C(d,p)(\log N)^{\frac{d-1}{2}}$ were constructed by Chen and Skriganov [15, 16] for $p = 2$ and by Skriganov [46] for $p > 1$. Determining the L^p-behaviour of the discrepancy functions for $p = 1, \infty$ is one of the most challenging open problems in the field of geometric discrepancy.

Conjecture 3.18. (L^∞-Conjecture) For all $d \geq 2$ and all $\mathcal{P}_N \subset [0,1]^d$ with $\#\mathcal{P}_N = N$,

$$\|D_N\|_\infty \geq C_d (\log N)^{\frac{d}{2}}, \tag{3.31}$$

where C_d is an absolute constant depending only on the dimension d.

Since $L^\infty([0,1]^d)$ is continuously embedded in $L^2([0,1]^d)$, Roth's theorem immediately implies that

$$\|D_N\|_\infty \geq C_d (\log N)^{\frac{d-1}{2}}. \tag{3.32}$$

Comparing (3.31) and (3.32), we see that there is a gap of order $\sqrt{\log N}$ between the conjectured lower bound of the discrepancy function and the lower bound provided by

Roth's theorem. This gap was eliminated by Schmidt [42] in $d = 2$. More precisely, he showed that $\|D_N\|_\infty \geq C \log N$. One decade later, Halász [25] gave another proof of the same result by combining the technique of using Riesz products with Roth's orthogonal function method. This is where the small ball inequality enters the area of discrepancy theory. More precisely, Halász's approach can be transferred to prove the small ball inequality in $d = 2$. After we prove Roth's theorem, we will analyze Halász's proof and, in Chapter 4, show how his technique can be implemented to prove the small ball inequality in $d = 2$. The connection between problems in discrepancy theory and the small ball inequality is based purely on the methods of proof used. Currently, there is no indication that the results in discrepancy theory will yield results concerning the small ball inequality, or vice versa.

In 1989, by employing techniques using Riesz products, Beck [2] was able to make some progress toward improving (3.32) in higher dimensions. Specifically, in dimension $d = 3$, he showed that

$$\|D_N\|_\infty \geq C \log N \cdot (\log \log N)^{\frac{1}{8}-\varepsilon}, \quad \text{for small } \varepsilon > 0. \tag{3.33}$$

In other words, he improved the lower bound, $\log N$, by a factor of order $(\log \log N)^{\frac{1}{8}-\varepsilon}$, for some small $\varepsilon > 0$. In dimensions $d \geq 3$, almost two decades later, Bilyk, Lacey, Vagharshakyan [3, 4] improved the bound, $(\log N)^{\frac{d-1}{2}}$, by a factor of order $(\log N)^\theta$ by showing

$$\|D_N\|_\infty \geq C_d (\log N)^{\frac{d-1}{2}+\theta}, \tag{3.34}$$

where θ is a small positive constant depending on the dimension d. Their work was based on Beck's work and improved Beck's result. It is still unknown whether Conjecture 3.18 is sharp, though examples of sets have been constructed by J. H. Halton and J. M. Hammersley [23, 24] which obey the upper bound: $\|D_N\|_\infty \leq C_d (\log N)^{d-1}$. This leads to another L^∞-conjecture:

Conjecture 3.19. *For all $d \geq 2$ and all point sets $\mathcal{P}_N$ in $[0,1]^d$ with $\#\mathcal{P}_N = N$ we have*

$$\|D_N\|_\infty \geq C_d (\log N)^{d-1}, \tag{3.35}$$

where C_d is an absolute constant.

The L^1-behaviour of the discrepancy function is another related open problem.

Conjecture 3.20. (L^1-Conjecture) For all $d \geq 2$ and all N-point sets in $[0,1]^d$,

$$\|D_N\|_1 \geq C_d(\log N)^{\frac{d-1}{2}}. \tag{3.36}$$

In dimension $d = 2$, Conjecture 3.20 was solved by Halász [25]. Vagharshakyan [51] improved the constant C_d in Halász's proof. It has also been conjectured that, for $0 < p < 1$, the L^p-norm of D_N satisfies the lower bound analogous to that appearing in (3.36).

3.3.2 Proof of Theorem 3.17

Here, we prove that, for any N-point set in $[0,1]^d$ and for all $d \geq 2$, the L^2-norm of D_N is of order $(\log N)^{\frac{d-1}{2}}$. Having expanded the discrepancy function with respect to the orthogonal basis of Haar functions:

$$D_N = \sum_{R \in \mathcal{D}^d_*} \frac{\langle D_N, h_R \rangle}{\|h_R\|_2^2} h_R,$$

Roth's approach was to identify the terms of the expansion which do not contribute significantly to $\|D_N\|_2$. The following proposition quantifies what this means more precisely.

Proposition 3.21. *Let $\mathcal{P}_N \subset [0,1]^d$ be a Npoint-set with $\#\mathcal{P}_N = N$, and let $n \in \mathbb{N}$ be such that $2^{n-2} \leq N < 2^{n-1}$. Then, for every $\vec{r} \in \mathbb{H}^d_n$, there exists an $\vec{r}$-function, $f_{\vec{r}}$, such that*

$$\langle D_N, f_{\vec{r}} \rangle := \int_{[0,1]^d} D_N \cdot f_{\vec{r}} dx \geq c_d > 0, \tag{3.37}$$

where the constant c_d depends only on the dimension d.

We will first assume that Proposition 3.21 holds to prove Theorem 3.17, and then we will prove Proposition 3.21 immediately afterward.

Proof of Theorem 3.17. Here, we use a duality argument and construct a test function, F, such that $\|F\|_2 \simeq \log^{\frac{d-1}{2}} N$ and $\langle D_N, F\rangle \gtrsim (\log N)^{d-1}$. Using these two lower bounds, the Cauchy-Schwarz inequality will imply:

$$\|D_N\|_2 \geq \frac{\langle D_N, F\rangle}{\|F\|_2} \gtrsim (\log N)^{\frac{d-1}{2}},$$

which is the estimate stated in Roth's theorem.

Define F by

$$F = \sum_{\vec{r}\in\mathbb{H}_n^d} f_{\vec{r}},$$

where $f_{\vec{r}}$ is the $\vec{r}$-function associated to $\vec{r} \in \mathbb{H}_n^d$ which comes from Proposition 3.21. Using the orthogonality of $\vec{r}$-functions and the fact that $\log N \simeq n$, we get that

$$\|F\|_2 = \left(\sum_{\vec{r}\in\mathbb{H}_n^d} \|f_{\vec{r}}\|_2^2\right)^{1/2} = (\#\mathbb{H}_n^d)^{1/2} \simeq n^{\frac{d-1}{2}} \simeq (\log N)^{\frac{d-1}{2}}.$$

On the other hand, Proposition 3.21 ensures that

$$\langle D_N, F\rangle \geq c_d \#\mathbb{H}_n^d \gtrsim n^{d-1} \gtrsim (\log N)^{d-1},$$

which completes the proof. □

Remark 3.22. The same approach can, essentially, be used to prove the L^p lower bounds of the discrepancy function. This will require us to estimate the L^p- norm of the function F, but, using the Littlewood-Paley inequalities, one can show that this has the same behaviour as the L^2-norm of F, *i.e.*, $\|F\|_p \simeq (\log N)^{\frac{d-1}{2}}$. Now, together with Hölder's inequality and Proposition 3.21, this will give us the estimate proved by Schmidt,

$$\|D_N\|_p \geq \frac{\langle D_N, F\rangle}{\|F\|_q} \gtrsim (\log N)^{\frac{d-1}{2}}, \tag{3.38}$$

where q is the dual exponent of $p \in (1, \infty)$. The Littlewood-Paley inequalities are needed to replace the orthogonality argument used in calculating the L^2-norm of F.

We refer the reader to Remark 5.8 for more details regarding the L^p-estimates of F.

Proof of Proposition 3.21. First, we observe that the discrepancy function associated to a set $\mathcal{P}_N \subset [0,1)^d$ with $\#\mathcal{P}_N = N$ can be written in an equivalent form:

$$D_N(\vec{x}) = \sum_{\vec{p}\in\mathcal{P}_N} \mathbb{1}_{[\vec{p},\vec{1})}(\vec{x}) - Nx_1\cdots x_d, \tag{3.39}$$

where $\vec{p} = (p_1,\ldots,p_d) \in \mathcal{P}_N \subset [0,1)^d$ and $[\vec{p},\vec{1}) = [p_1,1)\times\cdots\times[p_d,1)$. Fix a dyadic rectangle $R = I_1\times\cdots\times I_d \in \mathcal{D}^d$, and assume that the point $\vec{p}\in\mathcal{P}_N$ is not contained in R. Therefore, there exists $j\in\{1,\ldots,d\}$ such that p_j is not contained in I_j. As a result, $\int_{[0,1]} \mathbb{1}_{[p_j,1)}(x_j)h_{I_j}(x_j)dx_j = 0$, since either p_j is to the right of I_j, in which case $I_j\cap[p_j,1) = \varnothing$ and so $\int_{[0,1]} \mathbb{1}_{[p_j,1)}(x_j)h_{I_j}(x_j)dx_j = 0$, or p_j is to the left of I_j and $\int_{[0,1]} \mathbb{1}_{[p_j,1)}(x_j)h_{I_j}(x_j)dx_j = \int_{[0,1]} h_{I_j}(x_j)dx_j = 0$. Therefore, $\int_{[0,1]^d} \mathbb{1}_{[\vec{p},\vec{1})}\, h_R = 0$, and we get that

$$\left\langle \sum_{\vec{p}\in\mathcal{P}_N\cap R=\varnothing} \mathbb{1}_{[\vec{p},\vec{1})}, h_R \right\rangle = 0. \tag{3.40}$$

Also, since $\int_{I_j} x_j h_{I_j} dx_j = \frac{|I_j|^2}{4}$, $j = 1,\ldots,d$, we get

$$\langle x_1\cdots x_d, h_R\rangle = 4^{-d}|R|^2. \tag{3.41}$$

Using (3.40) and (3.41), we get a crucial relation which will be essential for the rest of the proof,

$$\langle D_N, h_R\rangle = -N4^{-d}|R|^2, \tag{3.42}$$

for all dyadic rectangles R which do not contain points from $\mathcal{P}_N$.

Recall that the set

$$\mathcal{R}_{\vec{r}} = \{R\in\mathcal{D}^d : R_j\in\mathcal{D} \text{ and } |R_j| = 2^{-r_j}, j = 1,\ldots,d\}$$

is a collection of disjoint axis-parallel dyadic rectangles whose sides are of prescribed lengths. Now, for each $\vec{r}\in\mathbb{H}_n^d$, we are ready to define an $\vec{r}$-function consisting of two parts, where the first part is a sum over the dyadic rectangles in $\mathcal{R}_{\vec{r}}$ which do not contain points from $\mathcal{P}_N$, and the second part is a sum over the dyadic rectangles in $\mathcal{R}_{\vec{r}}$ which contain points from $\mathcal{P}_N$. That is, we decompose $f_{\vec{r}}$ as

$$f_{\vec{r}} = \sum_{R\in\mathcal{R}_{\vec{r}}:R\cap\mathcal{P}_N=\varnothing} -1 h_R + \sum_{R\in\mathcal{R}_{\vec{r}}:R\cap\mathcal{P}_N\neq\varnothing} \operatorname{sgn}\langle D_N, h_R\rangle h_R, \tag{3.43}$$

where

$$\operatorname{sgn}(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ -1, & \text{otherwise} \end{cases}.$$

The inner product of the discrepancy function, D_N, with the second term of the decomposition in (3.43) is a positive number. Taking into account relation (3.42), this gives us

$$\langle D_N, f_{\vec{r}}\rangle \geq \sum_{R\in\mathcal{R}_{\vec{r}}:R\cap\mathcal{P}_N=\varnothing} N4^{-d}|R|^2. \tag{3.44}$$

Since $2^{n-2} \leq N < 2^{n-1}$ and $\#\mathcal{R}_{\vec{r}} = 2^n$, the number of rectangles in $\mathcal{R}_{\vec{r}}$ which intersect the N-point set $\mathcal{P}_N$ is at most 2^{n-1}. Thus, we have

$$\#\{R \in \mathcal{R}_{\vec{r}} : R \cap \mathcal{P}_N = \varnothing\} \geq 2^n\Big(1 - \frac{1}{2}\Big) \geq 2^{n-1}.$$

Therefore, recalling that $|R| = 2^{-n}$, we get the desired estimate:

$$\langle D_N, f_{\vec{r}}\rangle \geq \sum_{R\in\mathcal{R}_{\vec{r}}:R\cap\mathcal{P}_N=\varnothing} N4^{-d}|R|^2 \geq 2^{n-1}2^{n-2} \cdot \frac{2^{-2n}}{4^d} = c_d.$$

□

3.3.3 Halász's proof of Conjecture 3.18 in $d = 2$

Here, we give all the essential details of Halász's proof of Schmidt's result for the discrepancy function in $d = 2$,

$$\|D_N\|_\infty \geq C \log N. \tag{3.45}$$

Halász's approach was to use a duality argument in combination with Riesz product techniques and Roth's principle, Proposition 3.21. Using Riesz products turned out to be very fruitful in proving the two-dimensional small ball inequality, as we will see in the next chapter.

To present Halász's proof, we need two auxiliary lemmas. The first lemma shows how Haar functions interact when multiplied.

Lemma 3.23. (Product Rule) If $R \neq R' \in \mathcal{D}^2$ are not disjoint dyadic rectangles and $|R| = |R'|$, then

$$h_R \cdot h_{R'} = \pm 1 h_{R \cap R'},$$

i.e., the product of two Haar functions is again a Haar function.

We let the reader verify the result of this simple lemma. This lemma breaks down in dimensions $d \geq 3$, and the absence of the Product Rule in $d \geq 3$ is a major obstacle in solving the L^∞-discrepancy conjecture.

Lemma 3.24. *Let $f_{\vec{r}}$ be any $\vec{r}$-function with parameter s,* i.e., $|\vec{r}| := r_1 + \cdots + r_d = s$. *Then, for some constant $b_d > 0$,*

$$\langle D_N, f_{\vec{r}} \rangle \leq b_d N 2^{-s}.$$

The interested reader can find the proof of the above lemma in [10].

Halász's proof. First, we construct a test function Φ which satisfies two key estimates:

$$\|\Phi\|_1 \lesssim 1, \tag{3.46}$$

and

$$|\langle D_N, \Phi \rangle| \gtrsim \log N. \tag{3.47}$$

Then, once we have shown that (3.46) and (3.47) hold, applying Hölder's inequality in (3.47) will yield the conjectured estimate (3.45).

The function Φ will be a Riesz product, where the building blocks of the product will be $\vec{r}$-functions coming from Proposition 3.21. To be more precise, we fix n of order $\log N$, *i.e.*, $n \simeq \log N$, and we define the test function Φ by

$$\Phi := \prod_{k=0}^{n}(1+\gamma f_k) - 1, \tag{3.48}$$

where $\gamma \in (0,1)$ is a small constant which will be determined later, and

$$f_k := f_{(k,n-k)} = \sum_{R\in\mathcal{R}_{\vec{r}}} \alpha_R h_R$$

is the two-dimensional $\vec{r} = (k, n-k)$-function constructed in Proposition 3.21.

We proceed to establish properties (3.46) and (3.47) for our test function Φ. Expanding the product in (3.48), we get

$$\tilde{\Phi} := \prod_{k=0}^{n}(1+\gamma f_k) = 1 + \Phi_1 + \sum_{k=2}^{n} \Phi_k, \tag{3.49}$$

where

$$\Phi_1 = \gamma \sum_{k=0}^{n} f_k, \tag{3.50}$$

and

$$\Phi_k = \gamma^k \sum_{0\le j_1<\cdots<j_k\le n} f_{j_1}\cdot f_{j_2}\cdots f_{j_k}, \quad k \geq 2. \tag{3.51}$$

Note that the functions f_k are ± 1-valued. Since $\gamma < 1$, $1+\gamma f_k > 0$ for all k, and so $\tilde{\Phi}$ is a positive function. In addition, with the aid of Lemma 3.23, we have that $\int_{[0,1]^d} \Phi_k dx = 0$ for all $k \geq 1$, since a finite product of $\vec{r}$-functions is a linear combination of Haar functions. Therefore,

$$\|\tilde{\Phi}\|_1 = \int_{[0,1]^d} \tilde{\Phi} dx = 1. \tag{3.52}$$

Moreover, noticing that $\Phi = \tilde{\Phi} - 1$, from the triangle inequality we obtain

$$\|\Phi\|_1 \leq \|\tilde{\Phi}\|_1 + 1 = 2, \tag{3.53}$$

which is (3.46).

Now, we concentrate on showing (3.47). Recall that, from Proposition 3.21, $\langle D_N, f_k\rangle \geq c_d$ for $k = 0, \ldots, n$, for some constant $c_d < 1$. From this it is easy to see that

$$\langle D_N, \Phi_1\rangle \geq c_d \gamma \log N. \tag{3.54}$$

Hence, to complete the proof of (3.47), it suffices to show that, for some absolute constant c, we have

$$\langle D_N, \sum_{k=2}^{n} \Phi_k\rangle \leq 2c\gamma^2 \log N, \tag{3.55}$$

for a small value of $\gamma < 1$ such that $\gamma < \min(\frac{1}{2}, \frac{c_d}{2c})$. Indeed, since $\Phi = \Phi_1 + \sum_{k=2}^{n} \Phi_k$, by applying the reverse triangle inequality, (3.54) and (3.55), we get

$$|\langle D_N, \Phi\rangle| \geq \langle D_N, \Phi_1\rangle - \langle D_N, \sum_{k=2}^{n} \Phi_k\rangle \geq C(d, \gamma) \log N, \tag{3.56}$$

where $C(d, \gamma) = c_d\gamma - 2c\gamma^2$. Note that $C(d, \gamma) > 0$ since $\gamma < \frac{c_d}{2c}$. From the definition of Φ_k, $k = 2, \ldots, n$, and Lemma 3.23, we can see that product of $\vec{r}$-functions, $f_{j_1} \cdot f_{j_2} \cdots f_{j_k}$, is again an $\vec{r}$-function with $\vec{r} = (j_k, n - j_1)$. We rewrite the function Φ_k with respect to the parameter s as

$$\Phi_k = \gamma^k \sum_{s=n+1}^{2n} \sum_{\substack{0 \leq j_1 < j_2 < \cdots < j_k \leq n: \\ n - j_1 + j_k = s}} f_{\vec{r}}^{\vec{j}}, \tag{3.57}$$

where $f_{\vec{r}}^{\vec{j}}$ is an $\vec{r}$-function depending on $j_1 < j_2 < \cdots < j_k$ with $\vec{r} = (j_k, n - j_1)$. Given a value of parameter the s and values of j_1 and j_k such that $s = j_k + n - j_1$, the number of ways to select distinct values of $j_2 < \cdots < j_{k-1}$ between j_1 and j_k is equal to $\binom{s-n-1}{k-2}$. Furthermore, the number of ways to select the indices j_1 and j_k such that $j_k + n - j_1 = s$ for a given value of s is $2n - s + 1$. Using this, (3.57) and

Lemma 3.24, we get the following estimates:

$$\begin{aligned}
\langle D_N, \sum_{k=2}^{n} \Phi_k \rangle &\leq b_2 N \sum_{k=2}^{n} \gamma^k \sum_{s=n+1}^{2n} 2^{-s}(2n-s+1)\binom{s-n-1}{k-2} \\
&\leq b_2 N\gamma^2 \sum_{s=n+1}^{2n} 2^{-s}(2n-s+1) \sum_{k=2}^{s-n+1} \binom{s-n-1}{k-2}\gamma^{k-2} \\
&\leq b_2 N\gamma^2 \sum_{s=n+1}^{2n} 2^{-s}(2n-s+1)(1+\gamma)^{s-n-1} \\
&\leq b_2\gamma^2 N 2^{-n-1} \sum_{s=n+1}^{2n} \left(\frac{1+\gamma}{2}\right)^{s-n-1} (2n-s+1) \\
&\leq 2nb_2\gamma^2 \sum_{s=n+1}^{2n} \left(\frac{1+\gamma}{2}\right)^{s-n-1} \\
&\leq 2nb_2 \frac{2\gamma^2}{1-\gamma} \leq c\frac{\gamma^2}{1-\gamma} \log N.
\end{aligned}$$

Here, we have also used the fact that $N2^{-n-1} < 1$, since $2^{n-2} \leq N < 2^{n-1}$ and $n \simeq \log N$. Now, choosing $\gamma < \min(\frac{1}{2}, \frac{c_d}{2c})$, we have that

$$c\frac{\gamma^2}{1-\gamma} \log N \leq 2c\gamma^2 \log N,$$

where $c = 4b_2$, which completes the proof of (3.55). □

As we have already mentioned, the L^1-conjectured lower bound of the discrepancy function was solved by Halász [25] in $d = 2$, where, once more, he used Riesz product techniques. For the proof of L^1-conjectured lower bound of the discrepancy function, the test function Φ was defined by

$$\Phi := \prod_{k=0}^{n} \left(1 + \frac{i\gamma}{\sqrt{\log N}} f_k\right) - 1.$$

In [25], by introducing complex numbers, he showed that $\|\Phi\|_\infty \lesssim 1$. As in the previous calculations, he also proved that $|\langle D_N, \Phi\rangle| \gtrsim \sqrt{\log N}$ for an appropriate value of γ. These two estimates immediately imply the L^1-conjectured lower bound of the discrepancy function in $d = 2$.

Chapter 4

The two-dimensional small ball inequality

The two-dimensional small ball inequality states that, for any collection of coefficients $\{\alpha_R\}$ and any integer $n \geq 1$, one must have the following inequality

$$\Big\| \sum_{|R|=2^{-n}} \alpha_R h_R \Big\|_\infty \geq C_2 2^{-n} \sum_{|R|=2^{-n}} |\alpha_R|, \tag{4.1}$$

where C_2 is an absolute constant. Here, we will prove the two-dimensional small ball inequality, and will show that it is sharp. In fact, we will show in Section 4.2 that for all $d \geq 2$, the d-dimensional small ball inequality is sharp.

4.1 Temlyakov's proof of the two-dimensional small ball inequality

Although there are multiple proofs of the two-dimensional small ball inequality [47, 49, 11], here, we will only present the one given by Temlyakov [49]. The main reason that we focus on the details of Temlyakov's proof is to highlight the connection between the small ball inequality and discrepancy theory. As we already mentioned, this connection lies in the method of proofs. In particular, we will demonstrate the similarities between Temlyakov's proof of the small ball conjecture in $d = 2$ and Halász's proof of the conjectured L^∞-lower bound of the discrepancy function in $d = 2$. As in Halász's proof, the Product Rule, stated as Lemme 3.23 on page 46, was the key component in Temlyakov's proof. We can extend this product rule to a finite product of Haar functions, *i.e.*, if $\{R^j\}_{j=1}^k \subset \mathcal{D}^2$ is a finite sequence of non-disjoint dyadic rectangles, all of the same volume, then $\prod_{j=1}^k h_{R^j} = \pm 1 h_{\cap_{j=1}^k R^j}$. As a result, the product of distinct $\vec{r}$-functions, $f_{\vec{r}_j}$, $\vec{r}_j \in \mathbb{H}_n^2$, is again an $\vec{r}$-function, *i.e.*, $\prod_{j=1}^k f_{\vec{r}_j} = f_{\vec{r}}$. However, here $|\vec{r}| = r_1 + r_2 > n$.

It is easy to see that this rule does not hold in higher dimensions. For example, in $d = 3$, suppose R and R' are distinct non-disjoint rectangles of the same volume and whose first sides coincide, *i.e.*, $R_1 = R_1'$. Then the product of the Haar functions supported on these dyadic rectangles is not a Haar function, since $h_{R_1} \cdot h_{R_1'} = h_{R_1}^2 = \mathbb{1}_{R_1}$. The failure of this rule in higher dimensions makes the small ball inequality a hard problem in $d \geq 3$.

Proof of the small ball inequality for $d = 2$. Here, we solve the small ball conjecture for $d = 2$ via a duality approach. To do this, we need to construct a test function, Ψ, which satisfies the following two relations:

$$\|\Psi\|_1 \leq C, \tag{4.2}$$

where C is an absolute constant which does not depend on the choice of the coefficients $\{\alpha_R\}$ or the integer $n \geq 1$, and

$$\left\langle \Psi, \sum_{|R|=2^{-n}} \alpha_R h_R \right\rangle := \int_{[0,1]^d} \Psi \cdot \sum_{|R|=2^{-n}} \alpha_R h_R dx \geq C 2^{-n} \sum_{|R|=2^{-n}} |\alpha_R|. \tag{4.3}$$

Once we construct such a test function, the small ball conjecture in $d = 2$ follows from a simple application of Hölder's inequality to (4.3) for $p = 1$ and $q = \infty$.

As in Halász's proof, the test function Ψ is a Riesz product. Here, the building blocks of Ψ are again $\vec{r}$-functions, but this time they are defined by

$$f_k := f_{(k,n-k)} = \sum_{|R|=2^{-n}:|R_1|=2^{-k}} \operatorname{sgn}(\alpha_R) h_R,$$

and the test function Ψ is defined by

$$\Psi = \prod_{k=0}^{n} \Big(1 + \gamma f_k\Big). \tag{4.4}$$

Here, $\gamma \in (0,1]$ is any constant. For simplicity, we choose the constant γ to be equal to 1, but we will see that the proof works for any positive constant $\gamma \leq 1$.

Now, we are ready to show that Ψ satisfies (4.2) and (4.3). First, since $(1 + f_k)$ is nonnegative for all $k \in \{0, \dots, n\}$, we observe that the test function Ψ is nonnegative. Expanding the product in (4.4) gives

$$\Psi = 1 + \Psi_1 + \sum_{k=2}^{n} \Psi_k, \tag{4.5}$$

where

$$\Psi_1 = \sum_{k=0}^{n} f_k, \quad \Psi_k = \sum_{0 \leq j_1 < \cdots < j_k \leq n} f_{j_1} \cdots f_{j_k}, \quad k = 2, \dots, n. \tag{4.6}$$

As we remarked after restating the Product Rule, for all $k \in \{2, \dots, n\}$, Ψ_k is a sum of $\vec{r}$-functions. Thus, since $\int_{[0,1]^d} f_{\vec{r}} = 0$ for any $\vec{r}$-function, we get that $\int_{[0,1]^d} \Psi_k = 0$ for all $k \in \{1, \dots, n\}$. Therefore, $\|\Psi\|_1 = \int_{[0,1]^d} \Psi = 1 + 0 + 0 = 1$, which shows that Ψ satisfies (4.2).

Now, set $H_n = \sum_{|R|=2^{-n}} \alpha_R h_R$ and note that $\langle \Psi, H_n \rangle = \int_{[0,1]^2} H_n dx + \langle \Psi_1, H_n \rangle + \sum_{k=2}^{n} \langle \Psi_k, H_n \rangle$. Since $\int_{[0,1]^2} H_n dx = 0$, in order to finish the proof of (4.3), it suffices to show

$$\langle \Psi_1, H_n \rangle \geq C 2^{-n} \sum_{|R|=2^{-n}} |\alpha_R|, \tag{4.7}$$

and

$$\langle \Psi_k, H_n \rangle = 0, \quad k = 2, \dots, n. \tag{4.8}$$

It is easy to check that (4.8) holds. Indeed, as we have already mentioned, for $k = 2, \dots, n$, Ψ_k is a sum of $\vec{r}$-functions with $|\vec{r}| > n$, *i.e.*, Ψ_k is a linear combination of Haar functions supported on dyadic rectangles of volume strictly less than 2^{-n}. As a result, for every $k \in \{2, \dots, n\}$, when we multiply Ψ_k with H_n we get a linear combination of functions of the form $h_R h_{R'}$, where $R = R_1 \times R_2$, $R' = R_1' \times R_2'$ are two-dimensional dyadic rectangles with $|R| = 2^{-n}$ and $|R'| < 2^{-n}$ respectively. Since $|R| \neq |R'|$, $R_i \neq R_i'$ for either $i = 1$ or $i = 2$. Hence, $h_{R_i} h_{R_i'}$ is a Haar function. Due to the fact that the 1-dimensional integral of Haar functions is equal to zero, when we integrate the product $\Psi_k \cdot H_n$ over the unit square $[0,1]^2$, we therefore get (4.8).

Now, we focus on the proof of (4.7). First, note that the function H_n can be written as

$$H_n = \sum_{|R|=2^{-n}} \alpha_R h_R = \sum_{\vec{r} \in \mathbb{H}_n^2} \sum_{R \in \mathcal{R}_{\vec{r}}} \alpha_R h_R, \tag{4.9}$$

since $\cup_{\vec{r} \in \mathbb{H}_n^2} \mathcal{R}_{\vec{r}} = \{R \in \mathcal{D}^2 : |R| = 2^{-n}\}$. Using the above expression for H_n and the orthogonality of Haar functions, we have

$$\int_{[0,1]^2} \Psi_1 H_n dx = \sum_{\vec{r} \in \mathbb{H}_n^2} \sum_{R \in \mathcal{R}_{\vec{r}}} \int_{[0,1]^2} |\alpha_R| h_R^2 dx = 2^{-n} \sum_{|R|=2^{-n}} |\alpha_R|, \tag{4.10}$$

which shows that (4.7) holds. □

From the proof of the two-dimensional small ball inequality we see that $C_2 = 1$. If we had worked with the test function $\Psi = \prod_{k=0}^n \Big(1 + \gamma f_k\Big)$, $\gamma < 1$, then we would have gotten that $C_2 = \gamma < 1$. In Halász's proof, the constant γ had a specific value. In particular, he used the test function $\Phi = \prod_{k=0}^n (1 + \gamma f_k) - 1 = \Phi_1 + \sum_{k=2}^n \Phi_k$, where Φ_1 and Φ_k, $k \in \{2, \dots, n\}$ were defined in (3.50) and (3.51) respectively. The value of γ was chosen in such a way that the lower bound of $\langle D_N, \sum_{k=2}^n \Phi_k \rangle$ was smaller than the lower bound of $\langle D_N, \Phi_1 \rangle$. Also, Ψ and Φ differed in that Halász's $\vec{r}$-functions $f_k := f_{(k,n-k)}$ were chosen so that they had the property that $\langle D_N, f_k \rangle \geq c_d$ for all $k \in \{0, \dots, n\}$, whereas Temlyakov's didn't have any specific properties. The '-1' term was included in the definition of Φ for a technical reason. Specifically, including that term allowed Halász to avoid calculating the integral $\int_{[0,1]^2} D_N(x) dx$

when estimating the lower bound of $\langle D_N, \Phi\rangle := \int_{[0,1]^2} D_N(x)\Phi(x)dx$.

In the proof of the two-dimensional small ball inequality we estimated the L^∞-norm of H_n from below by estimating the integral

$$\langle \Psi, H_n\rangle := \int_{[0,1]^2} \Psi H_n dx.$$

The test function Ψ was defined as it was because it identifies small sets where H_n is very large. As a result, the integral $\langle \Psi, H_n\rangle$ is a good estimate for the L^∞-norm of H_n. Indeed, consider for simplicity that $\{\alpha_R\} \subset \{-1, 1\}$, *i.e.*, $\operatorname{sgn}(\alpha_R) = \alpha_R$. Then, in this case, $H_n = \sum_{\vec{r}\in\mathbb{H}_n^2} f_{\vec{r}} = \sum_{k=0}^n f_k$ and $\Psi = \prod_{k=0}^n \Big(1 + f_k\Big) = 2^{n+1}\, \mathbb{1}_E$, where

$$E = \{\vec{x} \in [0,1]^2 : f_k(\vec{x}) = 1 \text{ for every } \ k \in \{0, \ldots, n\}\}.$$

In other words, the function Ψ is supported on the set E where H_n is maximized. As a consequence of this fact, we have that $\int_{[0,1]^2} \Psi H_n dx = \int_E \Psi H_n dx = \|H_n\|_\infty \int_E \Psi = \|H_n\|_\infty$, since $\|\Psi\|_1 = \int_E \Psi dx = 1$.

Remark 4.1. Temlyakov actually proved a stronger result. His proof carries through if one replaces H_n with $\sum_{|R|\geq 2^{-n}} \alpha_R h_R$ to give

$$\Big\| \sum_{|R|\geq 2^{-n}} \alpha_R h_R \Big\|_\infty \gtrsim 2^{-n} \sum_{|R|=2^{-n}} |\alpha_R|.$$

Riesz products had already been used by Sidon [44, 45] to study lacunary Fourier series. Recall that an increasing sequence $\{\lambda_j\}_{j=1}^\infty \subset \mathbb{N}$ is called lacunary if there exists $q > 1$ so that $\lambda_{j+1}/\lambda_j > q$. Sidon used Riesz products to prove that if f has a lacunary series, *i.e.*, there exists a lacunary sequence Λ such that the Fourier coefficients of f,

$$\hat{f}(k) = \int_{[0,1]} f(x) e^{-2\pi i k x} dx,$$

are supported on the sequence Λ, then

$$\|f\|_\infty \gtrsim \sum_{k=1}^{\infty} |\hat{f}(k)|,$$

and

$$\|f\|_1 \gtrsim \|f\|_2,$$

where f is a bounded 1-periodic function. For the first inequality, he used the Riesz product

$$P_N(x) = \prod_{k=1}^{N} (1 + \cos(2\pi\lambda_k x + \delta_k)),$$

where δ_k was chosen so that $e^{i\delta_k} = \hat{f}(k)/|\hat{f}(k)|$. He showed that $\|P_N\|_1 = 1$ for every $N \in \mathbb{N}$ and $\|f\|_\infty \geq \left| \int_{[0,1]} f(x)\overline{P_N}(x) \right| = \frac{1}{2}\sum_{k=1}^{N} |f(\lambda_k)|$. He completed the proof by letting N go to infinity. The same ideas were implemented for the proof of the second inequality.

4.2 The sharpness of the small ball inequality in all dimensions

In this section, we show that the small ball inequality is sharp, *i.e.*, there exists a constant $C_d > 0$ and a collection of $\{\alpha_R\}$, in fact $\{\alpha_R\} \subset \{-1, 1\}$, such that

$$\left\| \sum_{R \in \mathcal{A}_n^d} \alpha_R h_R \right\|_\infty \leq C_d n^{-\frac{d-2}{2}} \cdot 2^{-n} \sum_{R \in \mathcal{A}_n^d} |\alpha_R| = C_d n^{\frac{d}{2}}.$$

It suffices to show that the above inequality holds on average, *i.e.*,

$$\mathbb{E} \left\| \sum_{R \in \mathcal{A}_n^d} \alpha_R h_R \right\|_\infty \leq C_d n^{\frac{d}{2}},$$

where $\{\alpha_R(\omega)\}_{R\in\mathcal{A}_n^d}$ are iid ±1-valued random variables on a probability space $(\Omega,\mathbb{P})$ with $\mathbb{P}(\omega\in\Omega:\alpha_R(\omega)=1)=\mathbb{P}(\omega\in\Omega:\alpha_R(\omega)=-1)=\frac{1}{2}$. Indeed, recall Markov's inequality: Let (X,Σ,μ) be a measure space. If $f\in L^1$ we have that

$$\mu\big(\{x:|f(x)|>a\}\big)\leq\frac{\|f\|_{L^1}}{a},$$

and for $a=2\|f\|_{L^1}$ we have that

$$\mu\big(\{x:|f(x)|>a\}\big)\leq\frac{1}{2}.$$

So, when $\mu=\mathbb{P}$ ($\mathbb{P}$ a probability measure) and $f(\omega)=\left\|\sum_{R\in\mathcal{A}_n^d}\alpha_R(\omega)h_R\right\|_\infty$ we have

$$\mathbb{P}\big(\{\omega:|f(\omega)|>2\mathbb{E}(f)\}\big)\leq\frac{1}{2}$$

which implies that

$$\mathbb{P}\big(\{\omega:|f(\omega)|\leq2\mathbb{E}(f)\}\big)>\frac{1}{2}.$$

This implies that there exists $\omega_o\in\Omega$ such that

$$\begin{aligned}|f(\omega_0)|=\left\|\sum_{R\in\mathcal{A}_n^d}\alpha_R(\omega_0)h_R\right\|_\infty&\leq2\mathbb{E}\left(\left\|\sum_{R\in\mathcal{A}_n^d}\alpha_R(\omega)h_R\right\|_\infty\right)\\&\lesssim n^{\frac{d}{2}}.\end{aligned}$$

To present the proof of the sharpness of the d-dimensional small ball inequality, we need a preliminary lemma. This lemma is a standard fact in probability theory and we state it here without a proof.

Lemma 4.2. (Bernstein's inequality [28]) If Z_i, $i=1,\ldots,N$, are independent and identically distributed ±1-valued random variables with $\mathbb{P}(Z_i=1)=\mathbb{P}(Z_i=-1)=$

0.5 and $(c_1, \ldots, c_N) \in \mathbb{R}^N$, then

$$\mathbb{P}\left(\left|\sum_{l=1}^{N} c_i Z_i\right| > \lambda\right) \le 2\exp\left(-\frac{\lambda^2}{4\sigma^2}\right),$$

where $\sigma^2 = \sum_{i=1}^N \sigma_i^2$ with $\sigma_i^2 = \mathbb{E}((c_i Z_i)^2)$ and $\mathbb{E}(Z_i) = 0$, for $i = 1, \ldots, N$.

Theorem 4.3. *The small ball inequality is sharp on average,* i.e.,

$$\mathbb{E}\left\|\sum_{R\in\mathcal{A}_n^d} \alpha_R h_R\right\|_\infty \le C_d n^{\frac{d}{2}},$$

where $\{\alpha_R(\omega)\}_{R\in\mathcal{A}_n^d}$ *are iid* ± 1 *valued random variables on a probability space* $(\Omega, \mathbb{P})$ *with* $\mathbb{P}(\alpha_R(\omega) = 1) = \mathbb{P}(\alpha_R(\omega) = -1) = \frac{1}{2}$.

Proof. First, we observe that the function $\sum_{R\in\mathcal{A}_n^d} \alpha_R(\omega) h_R(\cdot)$ is constant on dyadic cubes of side length $2^{-(n+1)}$ for each $\omega \in \Omega$. This can be seen by noting that each rectangle of volume 2^{-n} can be covered by the dyadic cubes of side length $2^{-(n+1)}$, and so each cube will lie in that part of the rectangle where the Haar function is constant.

As a consequence, the total number of such cubes is $M = 2^{(n+1)d}$. If we denote each such cube by Q_k, $k = 1, \ldots M$, then we can create M random variables

$$X_k(\omega) = \sum_{R\in\mathcal{A}_n^d} \alpha_R(\omega) h_R(x_k),$$

where $x_k \in Q_k$ for $k = 1, \ldots, M$. Applying Bernstein's inequality to X_k, for $k = 1, \ldots, M$, we get

$$\mathbb{P}\big(|X_k| > t\big) \le 2\exp\left(-\frac{t^2}{4\#\mathbb{H}_n^d}\right). \tag{4.11}$$

Let $\psi(t) = \exp(t^2)$ and define the following random variables:

$$Y_k(\omega) = \frac{X_k(\omega)}{Cn^{\frac{d-1}{2}}},$$

where the constant C will be chosen later. Taking into account the fact that

$$\{\omega : \psi(|Y_k(\omega)|) > t\} = \{\omega \in \Omega : |Y_k(\omega)| > \psi^{-1}(t)\},$$

we have that

$$\begin{aligned}
\mathbb{E}\big(\psi(|Y_k|)\big) &= \int_0^\infty \mathbb{P}\big(\psi(|Y_k(\omega)|) > t\big)dt \\
&= \int_0^\infty \mathbb{P}\big(|Y_k(\omega)| > \psi^{-1}(t)\big)dt \\
&= \int_0^\infty \mathbb{P}\left(|X_k(\omega)| > \frac{Cn^{\frac{d-1}{2}}\sqrt{\log t}}{4\#\mathbb{H}_n^d}\right) \\
&= \int_0^\infty \min\left\{1, 2\exp\left(-\frac{C^2 n^{d-1}\log t}{4\#\mathbb{H}_n^d}\right)\right\} dt \qquad \text{(by (4.11))} \\
&\leq \int_0^\infty \min\left\{1, 2t^{-K}\right\} dt,
\end{aligned}$$

where $\#\mathbb{H}_n^d \leq C_d n^{d-1}$ and $K = \frac{C^2}{4C_d}$. Now we choose K sufficiently large, or equivalently we choose C sufficiently large, so that

$$\int_0^\infty \min\left\{1, 2t^{-K}\right\} dt \leq C',$$

where C' is an absolute constant, so that

$$\mathbb{E}\big(\psi(|Y_k|)\big) \leq C'. \tag{4.12}$$

Since $\psi(t) = \exp(t^2)$ is a convex, increasing function on $[0, \infty)$ and

$$\psi\left(\sup_{k=1,\ldots,M}\{|Y_k(\omega)|\}\right) = \sup_{k=1,\ldots,M}\left\{\psi\big(|Y_k(\omega)|\big)\right\},$$

we get the following by applying Jensen's inequality:

$$\begin{aligned}
\psi\left(\mathbb{E}\left(\sup_{k=1,\ldots,M}\left\{|Y_k(\omega)|\right\}\right)\right) &\leq \mathbb{E}\left(\psi\left(\sup_{k=1,\ldots,M}\left\{|Y_k(\omega)|\right\}\right)\right)\\
&\leq \mathbb{E}\left(\sup_{k=1,\ldots,M}\left\{\psi(|Y_k(\omega)|)\right\}\right)\\
&\leq \mathbb{E}\left(\sum_{k=1}^{M}\psi\big(|Y_k(\omega)|\big)\right)\\
&\leq C'M = C'2^{(n+1)d},
\end{aligned}$$

where C' is an absolute constant, and the last line was obtained by using (4.12). Since $\psi^{-1}(t) = \sqrt{\log t}$, we have that

$$\mathbb{E}\left(\sup_{k=1,\ldots,M}\left\{|Y_k(\omega)|\right\}\right) \leq (n+1)^{\frac{1}{2}} d^{\frac{1}{2}} \sqrt{C' \log 2},$$

which implies that

$$\mathbb{E}\left(\sup_{k=1,\ldots,M}\left\{|X_k(\omega)|\right\}\right) \leq C'' n^{\frac{d-1}{2}} n^{\frac{1}{2}} d^{\frac{1}{2}} = C'' n^{\frac{d}{2}} d^{\frac{1}{2}},$$

where C'' is an absolute constant. Hence, we have shown that

$$\mathbb{E}\left\|\sum_{R\in\mathcal{A}_n^d}\alpha_R(\omega)h_R\right\|_\infty \lesssim n^{\frac{d}{2}}, \tag{4.13}$$

since $\sup_{k=1,\ldots,M}\left\{|X_k(\omega)|\right\} = \left\|\sum_{R\in\mathcal{A}_n^d}\alpha_R(\omega)h_R\right\|_\infty, \ \forall\omega\in\Omega.$ □

Chapter 5

Elements of dyadic harmonic analysis

5.1 Khintchine's inequalities

Khintchine's inequalities allow us to, up to a constant, calculate the p-norms of a linear combination of iid random variables. We will use Khintchine's inequalities to prove Theorem 1.4.

Let $\{X_i, 1 \leq i \leq N\}$ be a collection of ± 1-valued iid random variables on a probability space $(\Omega, \mathbb{P})$ such that $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$. Then, Khintchine's inequalities state that there exist constants $A_p > 0$ and $B_p > 0$ such that

$$A_p \left[\sum_{i=1}^{N} |a_i|^2 \right]^{\frac{1}{2}} \leq \left\| \sum_{i=1}^{N} a_i X_i \right\|_p \leq B_p \left[\sum_{i=1}^{N} |a_i|^2 \right]^{\frac{1}{2}}, \ 1 < p < \infty,$$

for any finite sequence of real numbers $(a_i)_{i=1}^N$.

The proofs of these inequalities rely on two key facts about the distribution function of the random variable $S_N = \sum_{i=1}^N a_i X_i$. In particular, the distribution function of S_N, $d_{S_N}(\lambda) = \mathbb{P}(\{\omega \in \Omega : |S_N| > \lambda\})$, satisfies the sub-Gaussian bound

$$d_{S_N}\left(\Big[\sum_{i=1}^{N} a_i^2\Big]^{\frac{1}{2}}\lambda\right) \leq 2e^{-\frac{1}{2}\lambda^2}, \text{ for all } \lambda > 0,$$

and the L^p-norm of S_N can be computed in terms of d_{S_N}:

$$\|S_N\|_p^p = \int_0^\infty p\lambda^{p-1} d_{S_N}(\lambda)d\lambda, \ 0 < p < \infty.$$

Khintchine's inequalities can be extended to the case of complex-valued square summable sequences $\{a_j\}$. When $p = 2$, the orthogonality of the random variables $X_i, 1 \leq i \leq N$, reduces Khintchine's inequalities to Parseval's identity:

$$\|S_N\|_2^2 = \sum_{i=1}^{N} |a_i|^2.$$

Moreover, $B_p = 1$ when $0 < p \leq 2$ and $A_p = 1$ for $p \geq 2$, as the trivial estimates $\|S_N\|_p \leq \|S_N\|_2, \ 0 < p \leq 2$ and $\|S_N\|_2 \leq \|S_N\|_p, \ p \geq 2$ show respectively. For a complete proof of these inequalities, we refer the interested reader to [21]. For the proof of the optimal values of A_p and B_p, we refer the reader to [22, 53].

A direct application of Khintchine's inequalities shows that the p-norms of a linear combination of Rademacher functions are comparable, *i.e.*, there exist positive constants $C_i(p,q), i = 1, 2$ such that

$$C_1(p,q)\Big\|\sum_{k=0}^{N} a_k r_k\Big\|_p \leq \Big\|\sum_{k=0}^{N} a_k r_k\Big\|_q \leq C_2(p,q)\Big\|\sum_{k=0}^{N} a_k r_k\Big\|_p, \ q > p > 0. \tag{5.1}$$

Here, $r_k = \sum_{I \in \mathcal{D}:|I|=2^{-k}} h_I$ are Rademacher functions which are iid random variables on the probability space $([0,1), |\cdot|)$.

We would like to extend the above inequalities to more general functions. More specifically, we would like to extend them to $\sum_{k=0}^{N}\sum_{|I|=2^{-k}} \alpha_I h_I$ or, in the multivariable case, to $\sum_{R \in \mathcal{D}^d:|R|=2^{-N}} \alpha_R h_R$, where the coefficients $\{\alpha_I\}$ or $\{\alpha_R\}$ depend not only on the scale of the dyadic rectangles, but also on the rectangles themselves. For this reason, the Littlewood-Paley inequalities will be useful.

5.2 The Littlewood-Paley inequalities

In this section, we will present the Littlewood-Paley inequalities for martingales. These inequalities can be considered as a generalization of Khintchine's inequalities. The special case of the Littlewood-Paley inequalities for Haar functions will be used for the proof of the higher dimensional analogue of inequalities (5.1). To introduce the Littlewood-Paley inequalities for martingales, we need some preliminary definitions from probability theory.

Suppose $X \in L^1(\Omega, \mathcal{B}, \mathbb{P})$ is an integrable real-valued random variable and let $\mathcal{G} \subset \mathcal{B}$ a be sub-σ-algebra. We define the conditional expectation of X given $\mathcal{G}$ to be a random variable, denoted by $\mathbb{E}(X|\mathcal{G})$, which satisfies

(i) $\mathbb{E}(X|\mathcal{G})$ is $\mathcal{G}$-measurable and integrable.

(ii) For all $G \in \mathcal{G}$, we have

$$\int_G X d\mathbb{P} = \int_G \mathbb{E}(X|\mathcal{G}) d\mathbb{P}.$$

We mention only those properties conditional expectations possess that we will use throughout this chapter.

(1) Linearity: If $X, Y \in L^1$, and $\alpha, \beta \in \mathbb{R}$, we have

$$\mathbb{E}(\alpha X + \beta Y|\mathcal{G}) = \alpha \mathbb{E}(X|\mathcal{G}) + \beta \mathbb{E}(Y|\mathcal{G}).$$

(2) $\mathbb{E}(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}(X)$.

(3) Product Rule: Let X, Y be random variables satisfying $X, Y \cdot X \in L^1$. If $Y \in \mathcal{G}$,

i.e., Y is measurable with respect to $\mathcal{G}$, then

$$\mathbb{E}(XY|\mathcal{G}) = Y\mathbb{E}(X|\mathcal{G}).$$

(4) Smoothing: If $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{B}$, then

$$\mathbb{E}(\mathbb{E}(X|\mathcal{G}_2)|\mathcal{G}_1) = \mathbb{E}(X|\mathcal{G}_1).$$

(5) If $X \in L^1$ is independent of the σ-algebra $\mathcal{G}$, then

$$\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X).$$

The proof of these properties can be found in [38].

Definition 5.1. Suppose we are given integrable random variables $\{X_n\}_{n\geq 0}$ and σ-algebras $\{\mathcal{G}_n\}_{n\geq 0}$ which are sub-σ-algebras of $\mathcal{B}$. Then $\{(X_n, \mathcal{G}_n)\}_{n\geq 0}$ is a *martingale* if:

(i) $\{\mathcal{G}_n\}_{n\geq 0}$ is an increasing sequence of σ-algebras, *i.e.*,

$$\mathcal{G}_0 \subset \mathcal{G}_1 \subset \mathcal{G}_2 \subset \cdots \subset \mathcal{B}.$$

(ii) X_n is adapted, *i.e.*, $X_n \in \mathcal{G}_n$, for each $n \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$.

(iii) For $0 \leq m < n$,

$$\mathbb{E}(X_n|\mathcal{G}_m) = X_m.$$

By the smoothing property, condition (iii) in the definition of a martingale is equivalent to $\mathbb{E}(X_{n+1}|\mathcal{G}_n) = X_n$. In many applications, $\mathcal{G}_n$ is the smallest σ-algebra generated by $\{X_i : 0 \leq i \leq n\}$, *i.e.* $\mathcal{G}_n = \sigma(X_0, \ldots, X_n)$. Note that every martingale $\{(X_n, \mathcal{G}_n)\}_{n\geq 0}$ can be written as $X_n = \sum_{j=0}^{n} d_j$, where $\{(d_j, \mathcal{G}_j)\}_{j\geq 0}$ is a *martingale difference*, *i.e.*, $d_j \in \mathcal{G}_j$ with $d_j \in L^1$ and $\mathbb{E}(d_{j+1}|\mathcal{G}_j) = 0$ for $j \geq 0$. Indeed, simply define $d_0 = X_0$ and $d_j = X_j - X_{j-1},\ j \geq 1$. We can see that a martingale is a natural generalization of a sum of independent random variables. Indeed, suppose $\{Z_n\}_{n\geq 0}$ is an independent sequence of integrable random variables such that $\mathbb{E}(Z_n) = 0,\ n \geq 0$. Set $Y_n = \sum_{i=0}^{n} Z_i$ and $\mathcal{G}_n = \sigma(Z_0, \ldots, Z_n)$. Then $\{(Y_n, \mathcal{G}_n)\}_{n\geq 0}$ is a martingale since $\mathbb{E}(Y_{n+1}|\mathcal{G}_n) = \mathbb{E}(Y_n|\mathcal{G}_n) + \mathbb{E}(Z_{n+1}|\mathcal{G}_n) = Y_n + \mathbb{E}(Z_{n+1}) = Y_n$. Note that, in this case, $\{(Z_i, \mathcal{G}_i)\}_{i\geq 0}$ is a martingale difference. Moreover, we can create a martingale from a random variable $X \in L^1$ via the smoothing property. Specifically, if we take $X_n = \mathbb{E}(X|\mathcal{G}_n),\ n \geq 0$, then $\{(X_n, \mathcal{G}_n)\}_{n\geq 0}$ is a martingale.

Now, if $\{(X_n, \mathcal{G}_n)\}_{n\geq 0}$ is a martingale such that $\mathbb{E}(X_n^2) < \infty$, then $\{d_j\}$ is an orthogonal martingale difference, *i.e.*,

$$\mathbb{E}(d_i d_j) = 0, \quad i \neq j.$$

Indeed, using properties (2) and (3) of conditional expectation, if $j > i$, then we get

$$\begin{aligned}\mathbb{E}(d_i d_j) &= \mathbb{E}\left(\mathbb{E}(d_i d_j|\mathcal{G}_i)\right)\\ &= \mathbb{E}(d_i \mathbb{E}(d_j|\mathcal{G}_i)) = 0.\end{aligned}$$

As an immediate consequence, we obtain

$$\mathbb{E}(X_n^2) = \mathbb{E}\left(\sum_{j=0}^{n} d_j^2\right), \quad n \geq 0,$$

which is equivalent to

$$\|X_n\|_2 = \left\| \left(\sum_{j=0}^{n} d_j^2\right)^{\frac{1}{2}} \right\|_2. \tag{5.2}$$

The Littlewood-Paley inequalities are an L^p-version of (5.2), but they can also be viewed as an extension of Khintchine's inequalities to more general martingales.

Theorem 5.2. (Littlewood-Paley inequalities) *Let* $1 < p < \infty$*. There are positive* *real numbers M_p and N_p such that, if $\{(X_n, \mathcal{G}_n)\}_{n\geq 0}$ is a martingale, then*

$$M_p \left\| \left(\sum_{j=0}^{n} d_j^2 \right)^{\frac{1}{2}} \right\|_p \leq \|X_n\|_p \leq N_p \left\| \left(\sum_{j=0}^{n} d_j^2 \right)^{\frac{1}{2}} \right\|_p, \quad n \geq 0. \tag{5.3}$$

If $1 < p < \infty$ and $\sup_{n\geq 0} \|X_n\|_p < \infty$ then X_n converges to a random variable X both almost everywhere and with respect to the L^p-norm. Moreover, (5.3) becomes

$$M_p \left\| \left(\sum_{j=0}^{\infty} d_j^2 \right)^{\frac{1}{2}} \right\|_p \leq \|X\|_p \leq N_p \left\| \left(\sum_{j=0}^{\infty} d_j^2 \right)^{\frac{1}{2}} \right\|_p. \tag{5.4}$$

Here, the quantity

$$S(X) = \left(\sum_{j=0}^{\infty} d_j^2 \right)^{\frac{1}{2}}$$

is called the *square function* of X.

These inequalities were proven by D. Burkholder [12]. A major component of Burkholder's proof of the Littlewood-Paley inequalities the following lemma.

Lemma 5.3. *If $\{d_j\}$ is a martingale difference and $\{\varepsilon_j\} \subset \{-1, 1\}$, then*

$$\left\| \sum_{j=0}^{n} \varepsilon_j d_j \right\|_p \leq c_p \left\| \sum_{j=0}^{n} d_j \right\|_p, \quad 1 < p < \infty. \tag{5.5}$$

We will not prove inequality (5.5), but will show how to derive the Littlewood-Paley inequalities from Lemma 5.3.

Proof of the Littlewood-Paley inequalities. Consider the Rademacher functions $r_j = \sum_{I\in\mathcal{D}:|I|=2^{-j}} h_I,\ j = 0, \ldots, n$ on $[0, 1]$. Using inequality (5.5), we immediately get that

$$c_p^{-1} \|X_n\|_p \leq \left\| \sum_{j=0}^{n} r_j(t) d_j \right\|_p \leq c_p \|X_n\|_p, \quad \text{for every } t \in [0, 1]. \tag{5.6}$$

Raising both sides of the above inequality to the power p and integrating over the interval $[0,1]$, we obtain

$$(c_p^{-1})^p \|X_n\|_p^p \leq \int_{[0,1]} \left\| \sum_{j=0}^{n} r_j(t) d_j \right\|_p^p dt \leq (c_p)^p \|X_n\|_p^p. \tag{5.7}$$

Applying Fubini's Theorem and Khintchine's inequalities to the middle expression of the above inequality, we have

$$(c_p^{-1})^p \|X_n\|_p^p \leq B_p^p \left\| \left(\sum_{j=0}^{n} d_j^2 \right)^{\frac{1}{2}} \right\|_p^p. \tag{5.8}$$

Thus,

$$\|X_n\|_p \leq B_p c_p \left\| \left(\sum_{j=0}^{n} d_j^2 \right)^{\frac{1}{2}} \right\|_p, \tag{5.9}$$

which is the right side of (5.3) with $N_p = B_p c_p$. The left side of (5.3) is proven similarly. $\square$

Also, Burkholder [13] found the values of the optimal constants N_p and M_p occurring in the Littlewood-Paley inequalities:

$$N_p = \max\left\{p, \frac{p}{p-1}\right\} - 1,$$
$$M_p = \left(\max\left\{p, \frac{p}{p-1}\right\} - 1\right)^{-1}.$$

5.3 Connection of martingale differences with Haar functions

Recall from (1.1) that $\mathcal{D}$ is the set of dyadic intervals of $[0,1]$, *i.e.*,

$$\mathcal{D} = \bigcup_{k=0}^{\infty} \mathcal{A}_k^1, \tag{5.10}$$

where $\mathcal{A}_k^1$ is the collection of dyadic intervals of length 2^{-k} and $\mathcal{D}_* = \mathcal{D} \cup [-1, 1]$. We will be interested in the martingale:

$$\{(\mathbb{E}(f|\Delta_k), \Delta_k)\}_{k \geq 0}, \tag{5.11}$$

where $\Delta_k = \sigma(\mathcal{A}_k^1)$ and $f \in L^1([0,1]^d)$. Note that

$$\mathbb{E}(f|\Delta_k) = \sum_{I \in \mathcal{A}_k^1} \langle f \rangle_I \mathbb{1}_I(x),$$

where

$$\langle f \rangle_I = \frac{1}{|I|} \int_I f(x) dx$$

is the average of f over I. In the next proposition, we relate the martingale difference

$$d_k = \mathbb{E}(f|\Delta_k) - \mathbb{E}(f|\Delta_{k-1})$$

to Haar functions.

Proposition 5.4. *For every $f \in L^1([0,1])$ and for all $k \geq 1$, we have*

$$d_k(x) = \sum_{I \in \mathcal{A}_{k-1}^1} \frac{\langle f, h_I \rangle}{|I|} h_I(x), \quad x \in [0,1]. \tag{5.12}$$

Proof. First, we observe that, for every $I \in \mathcal{A}_{k-1}$, we have

$$2\langle f \rangle_I = \langle f \rangle_{I_l} + \langle f \rangle_{I_r}, \tag{5.13}$$

where I_l and I_r are the left and right halves of the dyadic interval I. Also, we can express $\mathbb{E}(f|\Delta_k)$ as

$$\mathbb{E}(f|\Delta_k) = \sum_{I \in \mathcal{A}_{k-1}^1} \langle f \rangle_{I_l} \mathbb{1}_{I_l} + \langle f \rangle_{I_r} \mathbb{1}_{I_r} \tag{5.14}$$

and $\mathbb{E}(f|\Delta_{k-1})$, using (5.13), as

$$\mathbb{E}(f|\Delta_{k-1}) = \sum_{I\in\mathcal{A}^1_{k-1}} \frac{1}{2}(\langle f\rangle_{I_l} + \langle f\rangle_{I_r})(\mathbb{1}_{I_l} + \mathbb{1}_{I_r}). \tag{5.15}$$

Subtracting (5.15) from (5.14), we get

$$d_k = \sum_{I\in\mathcal{A}^1_{k-1}} \frac{1}{2}(-\langle f\rangle_{I_l} + \langle f\rangle_{I_r})(-\mathbb{1}_{I_l} + \mathbb{1}_{I_r}) = \sum_{I\in\mathcal{A}^1_{k-1}} \frac{\langle f, h_I\rangle}{|I|} h_I, \tag{5.16}$$

which concludes the proof. □

It is well-known that the collection of Haar functions, $\{h_I\}_{I\in\mathcal{D}_*}$, is an unconditional Schauder basis for $L^p([0,1])$, $1 < p < \infty$. As a consequence, for $f \in L^p([0,1])$ and $1 < p < \infty$, we get that

$$f = \sum_{k=0}^{\infty} d_k \tag{5.17}$$

with respect to the L^p-norm. Furthermore, the square function of f is given by

$$S(f) = \left(\left| \int_0^1 f(x)dx \right|^2 + \sum_{I\in\mathcal{D}} \frac{|\langle f, h_I\rangle|^2}{|I|^2}.\, \mathbb{1}_I \right)^{\frac{1}{2}}. \tag{5.18}$$

If f has a Haar expansion,

$$f = \sum_{I\in\mathcal{D}_*} \alpha_I h_I, \tag{5.19}$$

then its square function is

$$S(f) = \left(\sum_{I\in\mathcal{D}_*} |\alpha_I|^2\, \mathbb{1}_I \right)^{\frac{1}{2}}, \tag{5.20}$$

and (5.4) becomes

$$M_p \left\| \left(\sum_{I\in\mathcal{D}_*} |\alpha_I|^2\, \mathbb{1}_I \right)^{\frac{1}{2}} \right\|_p \le \|f\|_p \le N_p \left\| \left(\sum_{I\in\mathcal{D}_*} |\alpha_I|^2\, \mathbb{1}_I \right)^{\frac{1}{2}} \right\|_p \tag{5.21}$$

for $1 < p < \infty$.

For functions of the form

$$f(\vec{x}) = \sum_{R\in\mathcal{D}_*^d} \alpha_R h_R(\vec{x}), \quad \vec{x} = (x_1, \ldots, x_d) \in [0,1]^d, \tag{5.22}$$

we define the *product dyadic square function*

$$S_d(f) = \left(\sum_{R\in\mathcal{D}_*^d} |\alpha_R|^2 \mathbb{1}_R \right)^{\frac{1}{2}}. \tag{5.23}$$

Inequalities (5.21) hold even when the coefficients α_I are elements of some Hilbert space (see [52]). Using this fact, one can prove the d-dimensional version of inequalities (5.21):

Theorem 5.5. (Product Littlewood-Paley Inequalities [8]) *For functions of the form* $f = \sum_{R\in\mathcal{D}_*^d} \alpha_R h_R$, *we have*

$$(M_p)^d \|S_d(f)\|_p \le \|f\|_p \le (N_p)^d \|S_d(f)\|_p, \quad 1 < p < \infty. \tag{5.24}$$

The next proposition compares different L^p-norms, $p \in (0,\infty)$, of a function. In particular, when f is a special linear combination of Haar functions, the next proposition shows that the norms $\|f\|_p$ are comparable for all p. The proof follows the strategy outlined in Corollary 1.4 of [48, page 5-6], and we extend that corollary to dimensions greater than one. Specifically, we establish the following:

Proposition 5.6. *Let f be a linear combination of Haar functions of the form*

$$f(\vec{x}) = \sum_{R\in\mathcal{A}_n^d} \alpha_R h_R(\vec{x}),$$

such that the square function, $S_d(f)$, is constant on $[0,1]^d$ (i.e., $S_d(f)(\vec{x}) = c(n,d)$ for every $\vec{x} \in [0,1]^d$, where the constant $c(n,d)$ depends on the integer $n \ge 1$ and on the dimension d). Then there exist positive constants $c_1(p,q,d)$ and $c_2(p,q,d)$ such *that*

$$c_2\|f\|_q \le \|f\|_p \le c_1\|f\|_q \quad \textit{for every} \quad 0<p<q<\infty. \tag{5.25}$$

Proof. First, we observe that the right side of inequality (5.25) is a simple consequence of Hölder's inequality ($c_1 = 1$). Therefore, our goal is to prove that

$$c_2\|f\|_q \le \|f\|_p \quad \text{for every} \quad 0<p<q<\infty.$$

We will show the above inequality by considering the following three cases:

Case 1: $1<p<q<\infty$.

Case 2: $0<p\le 1<q<\infty$.

Case 3: $0<p<q\le 1$.

In Case 1, using the Littlewood-Paley inequalities and the fact that $S_d(f)$ is independent of $\vec{x}$, we get that

$$\|f\|_p \ge (A_p)^d\|S_d(f)\|_p = (A_p)^d\|S_d(f)\|_q \ge \left(\frac{A_p}{B_q}\right)^d\|f\|_q. \tag{5.26}$$

Thus, $c_2\|f\|_q \le \|f\|_p$ with $c_2 = \left(\frac{A_p}{B_q}\right)^d$.

Now, we turn our attention to the proof of Case 2. First, choose $\alpha \in (0,1)$ such that $p > \alpha q$. Set $r = \frac{p}{\alpha q}$ and $s = \frac{r}{r-1}$ (*i.e.*, $\frac{1}{r}+\frac{1}{s}=1$). Also, set $b = 1-\alpha$ so that $\alpha + b = 1$. Note that $\alpha r = \frac{p}{q} < 1$, and $bs > 1$ since $(1-\alpha)\frac{r}{r-1} > 1$ if and only if $1 > \alpha r$. Now, applying Hölder's inequality, we get that

$$\|f\|_q^q = \int_{[0,1]^d} |f(x)|^{\alpha q}|f(x)|^{bq}dx \le \|f\|_{\alpha r q}^{\alpha q}\|f\|_{bsq}^{bq}. \tag{5.27}$$

But $bs > 1$ shows that $bsq > q > 1$, and Case 1 implies that there exists a positive constant c_2 such that

$$(c_2)^{bq}\|f\|_{bsq}^{bq} \le \|f\|_q^{bq}. \tag{5.28}$$

Combining inequalities (5.27) and (5.28), we get that

$$\|f\|_q^q \le (c_2)^{-bq}\|f\|_{\alpha r q}^{\alpha q}\|f\|_q^{bq}$$

which implies that

$$\|f\|_q^{q(1-b)} \le (c_2)^{-bq}\|f\|_{\alpha rq}^{\alpha q}.$$

As a result, $\|f\|_q \le (c_2)^{-\frac{b}{\alpha}}\|f\|_{\alpha rq}$. Thus, Case 2 follows since $\alpha rq = p$.

Finally, we prove Case 3. Define the numbers α, b, r and s in the same way as in Case 2 and note that inequality (5.27) still holds. Since $q \le 1$ implies $bsq \le bs$, using the right side of inequality (5.25), we get

$$\|f\|_{bsq}^{bq} \le \|f\|_{bs}^{bq}. \tag{5.29}$$

Now, putting together inequalities (5.27) and (5.29), we get that

$$\|f\|_q^q \le \|f\|_{\alpha rq}^{\alpha q}\|f\|_{bs}^{bq}. \tag{5.30}$$

Since $bs > 1 \ge q$, Case 2 gives us that $\|f\|_{bs}^{bq} \le (c_2)^{-bq}\|f\|_q^{bq}$. Combining this with inequality (5.30), we get that $\|f\|_q^q \le (c_2)^{-bq}\|f\|_{\alpha rq}^{\alpha q}\|f\|_q^{bq}$ which is equivalent to $\|f\|_q^{q(1-b)} \le (c_2)^{-bq}\|f\|_{arq}^{\alpha q}$. As a consequence, $\|f\|_q \le (c_2)^{-\frac{b}{a}}\|f\|_{arq}$. Since $\alpha rq = p$, this completes the proof of Case 3. □

Remark 5.7. To summarize, the Littlewood-Paley inequalities are used in the proof of (5.25) for the case $q > p > 1$. In particular, the assumption that the square function is constant on $[0,1]^d$ permits the p-norms of the square function to be comparable. Combined with the Littlewood-Paley inequalities, this fact proves inequality (5.25) for $q > p > 1$. To prove (5.25) for the full range of the values of p and q, one needs to combine the proof of the special case $q > p > 1$ and a subtle treatment of Hölder's inequality. If we assume that the coefficients $\{\alpha_R\}$ satisfy the inequalities

$$c_1 \le |\alpha_R| \le c_2, \quad \text{for every } R \in \mathcal{A}_n^d,$$

where c_1 and c_2 are positive constants independent of $R \in \mathcal{A}_n^d$ instead, then inequalities (5.25) still hold, as it is still true that the p-norms of the square function are comparable.

Remark 5.8. It is clear that if

$$f(\vec{x}) = \sum_{R\in\mathcal{A}_n^d} \alpha_R h_R(\vec{x}),$$

with $\{\alpha_R\}_{R\in\mathcal{A}_n^d} \subset \{-1,1\}$, then the square function, $(S_d f)$, is independent of $\vec{x} \in [0,1]^d$. More precisely,

$$S_d(f)(\vec{x}) = \left[\sum_{R\in\mathcal{A}_n^d} |\alpha_R|^2 \, \mathbb{1}_R(\vec{x})\right]^{\frac{1}{2}} = \left[\#\mathbb{H}_n^d\right]^{\frac{1}{2}},$$

where $\#\mathbb{H}_n^d$ denotes the cardinality of the set $\mathbb{H}_n^d$. An immediate consequence of this is that

$$\|F_{r_1}\|_{L^1([0,1]^{d-1})} \gtrsim \|F_{r_1}\|_{L^2([0,1]^{d-1})} \gtrsim (n-r_1)^{\frac{d-2}{2}}, \quad r_1 = 0,1,\ldots,n, \tag{5.31}$$

where

$$F_{r_1}(\vec{x}') = \sum_{R'\in\mathcal{A}_{n-r_1}^{d-1}} \alpha_{R'} h_{R'}(\vec{x}') \tag{5.32}$$

with $\vec{x}' = (x_2, x_3, \ldots, x_d) \in [0,1)^{d-1}$ and $\{\alpha_{R'}\} \subset \{-1,1\}$. Estimate (5.31) will be crucial in showing Lemma 6.4 and Theorem 1.3. Moreover, for $r_1 = 0$, F_{r_1} is the function used in the proof of the lower bound of the L^p-norm of the discrepancy function when $1 < p < \infty$, and, if $n \simeq \log N$, then $\|F_{r_1}\|_p \simeq (\log N)^{\frac{d-2}{2}}$. In addition, define

$$A_{r_1} = \sum_{\vec{r}'\in\mathbb{H}_{n-r_1}^{d-1}} \sum_{R\in\mathcal{R}_{\vec{r}}} \alpha_R h_R,$$

with $(r_1, \vec{r}') \in \mathbb{H}_n^d$, $r_1 = 0,1,\ldots,n$, and $\{\alpha_R\} \subset \{-1.1\}$. Using the same reasoning as for F_{r_1}, we get

$$\|A_{r_1}\|_1 \gtrsim \|A_{r_1}\|_2 \gtrsim (n-r_1)^{\frac{d-2}{2}}, \tag{5.33}$$

$r_1 = 0,1,\ldots,n$. These inequalities will be essential in proving Lemma 7.5.

Chapter 6

Two proofs of Theorem 1.3

The main objective of this chapter is to prove Theorem 1.3. In other words, we prove the signed small ball conjecture for all $d \geq 3$ under an additional constraint on the coefficients $\{\alpha_R\}_{R \in \mathcal{A}_n^d}$. We will give two proofs of Theorem 1.3. The first one will be based on a direct estimate of the L^∞-norm of the hyperbolic sum H_n, and the second one will be based on a duality argument which uses Riesz products. The second proof is particularly important since the techniques used there can also be used to obtain lower bounds of hyperbolic sums in other function spaces, such as exponential Orlicz spaces. We will discuss this more in Chapter 7. However, before proving Theorem 1.3, we first motivate the additional constraint on the coefficients that appears in Theorem 1.3, the splitting condition.

6.1 Motivation for the splitting condition

Recall that a coefficient $\alpha_R \in \{-1, 1\}$ satisfies the splitting condition if $\alpha_R = \alpha_{R_1} \cdot \alpha_{R'}$ with $\alpha_{R_1}, \alpha_{R'} \in \{-1, 1\}$ and $R = R_1 \times R' = R_1 \times R_2 \times \cdots \times R_d$. If one assumes that the coefficients of the signed hyperbolic sum satisfy the splitting condition, then the signed hyperbolic sum H_n takes the form

$$H_n(\vec{x}) = \sum_{r_1=0}^{n} B_{r_1}(x_1) F_{r_1}(\vec{x}'), \quad \vec{x} = (x_1, \vec{x}') \in [0,1]^d, \tag{6.1}$$

with

$$B_{r_1}(x_1) = \sum_{|R_1|=2^{-r_1}} \alpha_{R_1} h_{R_1}(x_1)$$

and

$$F_{r_1}(\vec{x}') = \sum_{|R'|=2^{-(n-r_1)}} \alpha_{R'} h_{R'}(\vec{x}'), \quad r_1 = 0, 1, \ldots, n.$$

The functions $\{B_{r_1}\}_{r_1=0}^n$ can be viewed as random variables defined on the probability space $(\Omega, \mathcal{B}, P) = ([0,1), \mathcal{B}([0,1)), |\cdot|)$, where $|\cdot|$ is the Lebesgue measure and $\mathcal{B}([0,1))$ is the Borel σ -algebra on the interval $[0,1)$, as we will explain later. The power of the splitting condition is that, when used to rewrite the hyperbolic sum as in (6.1), the random variables $\{B_{r_1}\}_{r_1=0}^n$ are independent, and the independence of these random variables will facilitate the computation of the supremum norm of H_n and lead to the gain of $n^{1/2}$ over the trivial bound. We prove the independence of $\{B_{r_1}\}_{r_1=0}^n$ after this preliminary discussion.

To prove the independence of the generalized Rademacher functions $\{B_{r_1}\}_{r_1=0}^n$, we will need to represent them in terms of iid binary random variables. To give all the details of this representation, we need to recall some facts from probability theory about the dyadic sequence $\{d_i(x)\}_{i\geq 1} \subset \{0,1\}$ associated to the dyadic expansion of a point $x \in [0,1)$,

$$x = \sum_{i=1}^{\infty} \frac{d_i(x)}{2^i}. \tag{6.2}$$

By convention, we use the terminating dyadic expansion for the dyadic rationals $\{x = \frac{k}{2^m} : k, m \in \mathbb{Z}, m \geq 0, 0 \leq k < 2^m\}$.

Fact 1: Each d_i is a random variable.

Fact 2: The sequence $\{d_i, i \geq 1\}$ is iid. In other words, each d_i is identically distributed with $P[d_i = 1] = P[d_i = 0] = \frac{1}{2}$, and the d_i are independent.
The proof of these facts can be found in [38, page 99-100].

To each such dyadic sequence we associate a sequence of the iid ± 1-valued random variables, $\{X_i\}_{i\geq 1}$, defined by

$$X_i = \begin{cases} -1, & \text{if } d_i = 0 \\ 1, & \text{if } d_i = 1 \end{cases},$$

where $i \in \mathbb{Z}_+ = \{1, 2, \ldots\}$. For every $i \in \mathbb{Z}_+$, we let

$$\vec{X}_i = (X_1, \ldots, X_i),$$

$$\vec{d}_i = (d_1, \ldots, d_i),$$

$$\vec{b}_i = (b_1, b_2, \ldots, b_i) \in \{-1, 1\}^i,$$

with the convention that $\{-1, 1\}^0 = \{0\}$, $\vec{d}_0 = 0$, $\vec{X}_0 = 0$ and $\vec{b}_0 = 0$. Now, fix $k \in \{0, 1, \ldots, n\}$ and a collection of numbers $\{\alpha_I\}_{I \in \mathcal{D}: |I| = 2^{-k}} \subset \{-1, 1\}$, and define the map:

$$a_k : \{-1, 1\}^k \to \{\alpha_I\}_{I \in \mathcal{D}: |I| = 2^{-k}}, \quad a_k(\vec{b}_k) = \alpha_I, \tag{6.3}$$

where I is the dyadic interval with $|I| = 2^{-k}$ given by

$$I = \left[\sum_{i=1}^{k} \frac{b_i + 1}{2^{i+1}}, \sum_{i=1}^{k} \frac{b_i + 1}{2^{i+1}} + 2^{-k} \right).$$

Observe that every $x \in I$ has the property that $\vec{X}_k(x) = \vec{b}_k$. Note that $a_0(\vec{b}_0) = \alpha_{[0,1)}$ and that this map is onto. Now, it is easy to verify that

$$B_{r_1}(x_1) = a_{r_1}(\vec{X}_{r_1}(x_1)) X_{r+1}(x_1), \quad r_1 = 0, 1, \ldots, n. \tag{6.4}$$

Hence,

$$H_n(\vec{x}) = \sum_{r_1=0}^{n} a_{r_1}(\vec{X}_{r_1}(x_1)) X_{r_1+1}(x_1) F_{r_1}(\vec{x}'), \quad \vec{x} = (x_1, \vec{x}') \in [0, 1]^d. \tag{6.5}$$

Proposition 6.1. *The random variables $\{B_k = a_k(\vec{X}_k) X_{k+1}\}_{k=0}^n$ are iid binary random variables on the probability space $([0, 1), \mathcal{B}([0, 1)), P = |\cdot|)$.*

Proof. First, for fixed values of $0 \leq k \leq n$ and $b_0 \in \{-1, 1\}$, we show that $P(X_{k+1}a_k(\vec{X}_k) = b_0) = \frac{1}{2}$. We will then use this to show that the random variables $\{a_k(\vec{X}_k)X_{k+1}\}_{k=0}^n$ are independent.

Without loss of generality, we can assume that $0 < k \leq n$. Fixing $b_0 \in \{-1, 1\}$, we get

$$\begin{aligned} P(X_{k+1}a_k(\vec{X}_k) = b_0) &= \sum_{\vec{b}_k \in \{-1,1\}^k} P\left(X_{k+1}a_k(\vec{X}_k) = b_0, \vec{X}_k = \vec{b}_k\right) \\ &= \sum_{\vec{b}_k \in \{-1,1\}^k} P\left(X_{k+1} = \frac{b_0}{a_k(\vec{b}_k)}, \vec{X}_k = \vec{b}_k\right) \\ &= 2^k \frac{1}{2}\left(\frac{1}{2}\right)^k \\ &= \frac{1}{2}. \end{aligned}$$

Now, using the independence of the random variables $\{X_i\}_{i=1}^{n+1}$, for any $\vec{b}_{n+1} \in \{-1, 1\}^{n+1}$, we get that

$$\begin{aligned} P(X_1 a_0(\vec{X}_0) = b_1, \ldots, X_{n+1}a_n(\vec{X}_n) = b_{n+1}) &= P(X_1 = \widetilde{b}_1, \ldots, X_{n+1} = \widetilde{b}_{n+1}) \\ &= \left(\frac{1}{2}\right)^{n+1} \\ &= \prod_{k=0}^{n} P(X_{k+1}a_k(\vec{X}_k) = b_{k+1}), \end{aligned}$$

where $\widetilde{b}_1, \ldots, \widetilde{b}_{n+1}$ are ± 1-valued deterministic constants depending on $b_1, \ldots, b_{n+1}$. This last equality gives the independence of the random variables $\{a_k(\vec{X}_k)X_{k+1}\}_{k=0}^n$. □

6.2 The first proof of Theorem 1.3

The next proposition gives us one of the main ingredients of the first proof of Theorem 1.3: an explicit formula for the L^∞-norm of a linear combination of iid ± 1-valued random variables.

Proposition 6.2. *Let* $(\Omega, \mathcal{F}, P)$ *be a probability space and* $\{U_i\}_{i=1}^n$ *be a finite collection of* ± 1*-valued independent random variables for which* $P(U_i = \pm 1) > 0$ *for every* $i = 1, \dots, n$. *If* $(c_i)_{i=1}^n$ *is any finite deterministic sequence of real numbers, then*

$$\left\| \sum_{i=1}^n c_i U_i \right\|_\infty = \sum_{i=1}^n |c_i|.$$

Proof. It is clear that $|\sum_{i=1}^n c_i U_i(\omega)| \leq \sum_{i=1}^n |c_i|$ for every $\omega \in \Omega$. So, to complete the proof, we need to find $\omega_0 \in \Omega$ such that $U_i(\omega_0) = \operatorname{sgn}(c_i),\ i = 1, \dots, n$, where sgn is the signum function defined in (1.7). Let $A = \bigcap_{i=1}^n \{\omega \in \Omega : U_i(\omega) = \operatorname{sgn}(c_i)\}$. Using the independence of the random variables $\{U_i\}_{i=1}^n$, we get that $P(A) = \prod_{i=1}^n P(U_i = \operatorname{sgn}(c_i)) > 0$, which shows that A is not a set of measure zero. Thus,

$$\left\| \sum_{i=1}^n c_i U_i \right\|_\infty = \sum_{i=1}^n |c_i|.$$

□

The first proof of Theorem 1.3. Using the definition of the supremum norm and (6.5), we get

$$\|H_n\|_\infty = \sup_{\vec{x}' \in [0,1]^{d-1}} \sup_{x_1 \in [0,1]} \left| \sum_{r_1=0}^n F_{r_1}(\vec{x}') a_{r_1} \left(\vec{X}_{r_1}(x_1) \right) X_{r_1+1}(x_1) \right|. \tag{6.6}$$

Now, if we can show that

$$\|H_n\|_\infty \geq \sum_{r_1=0}^n \|F_{r_1}\|_{L^1([0,1]^{d-1})}, \tag{6.7}$$

then we can use (5.31) to obtain the desired inequality. To obtain (6.7), we need to rewrite the inner supremum appearing on the right side in (6.6). Fix $\vec{x}' = (x_2, x_3, \cdots, x_d) \in [0,1]^{d-1}$. Applying both Proposition 6.1 and Proposition 6.2 to

$U_{r_1} = \alpha_{r_1}(\vec{X}_{r_1})X_{r_1+1}$ and $c_{r_1} = F_{r_1}(x_2, x_3, \cdots, x_d)$, $r_1 = 0, 1, \ldots, n$, we get that

$$\sup_{x_1 \in [0,1]} \left| \sum_{r_1=0}^{n} F_{r_1}(\vec{x}') a_{r_1}\Big(\vec{X}_{r_1}(x_1)\Big) X_{r_1+1}(x_1) \right| = \sum_{r_1=0}^{n} |F_{r_1}(\vec{x}')|. \tag{6.8}$$

Combining (6.6) and (6.8), we get

$$\|H_n\|_\infty = \sup_{\vec{x}' \in [0,1]^{d-1}} \sum_{r_1=0}^{n} |F_{r_1}(\vec{x}')|. \tag{6.9}$$

Inequality (6.7) easily follows since

$$\|H_n\|_\infty \geq \left\| \sum_{r_1=0}^{n} |F_{r_1}| \right\|_{L^1([0,1]^{d-1})} = \sum_{r_1=0}^{n} \|F_{r_1}\|_{L^1([0,1]^{d-1})}.$$

Now, using inequality (5.31), estimate (6.7) becomes

$$\|H_n\|_\infty \geq c_d \sum_{r_1=0}^{n} \|F_{r_1}\|_{L^2([0,1]^{d-1})}. \tag{6.10}$$

Using the orthogonality of Haar functions, we have that

$$\|F_{r_1}\|_{L^2([0,1]^{d-1})} = \|S_{d-1}(F_{r_1})\|_{L^2([0,1]^{d-1})} \gtrsim (n - r_1)^{\frac{d-2}{2}}. \tag{6.11}$$

Finally, putting inequalities (6.10) and (6.11) together, we obtain the desired result.

□

Remark 6.3. The two basic components of the first proof are the following:

$$\|H_n\|_\infty = \sup_{\vec{x}' \in [0,1]^{d-1}} \sum_{r_1=0}^{n} |F_{r_1}(\vec{x}')|, \tag{6.12}$$

$$\|F_{r_1}\|_{L^1([0,1]^{d-1})} \gtrsim \|F_{r_1}\|_{L^2([0,1]^{d-1})}, \tag{6.13}$$

for $r_1 = 0, 1, \ldots, n$. According to Remark 5.7, if we consider coefficients $\{\alpha_{R_1}\alpha_{R'}\}$

in H_n for which $\alpha_{R_1} \in \{-1, 1\}$ and $0 < c_1 \leq |\alpha_{R'}| \leq c_2$ for all $R_1 \times R' \in \mathcal{A}_n^d$, then inequality (6.13) still holds. Moreover, for the same collection of coefficients, the splitting condition ensures that equality (6.12) also remains true. Based on these two facts, the small ball conjecture, Conjecture 1.1, can be proved for this special collection of coefficients. Indeed, using (6.12), (6.13), and the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}
\|H_n\|_\infty &\gtrsim \sum_{r_1=0}^{n} \|F_{r_1}\|_{L^2([0,1]^{d-1})} \\
&\gtrsim \sum_{r_1=0}^{n} \Bigg(\sum_{\vec{r}' \in \mathbb{H}_{n-r_1}^{d-1}} \sum_{R' \in \mathcal{R}_{\vec{r}'}} |\alpha_{R'}|^2 2^{-(n-r_1)} \Bigg)^{1/2} \\
&\gtrsim 2^{-n} n^{-(d-2)/2} \sum_{r_1=0}^{n} 2^{r_1} \sum_{\vec{r}' \in \mathbb{H}_{n-r_1}^{d-1}} \sum_{R' \in \mathcal{R}_{\vec{r}'}} |\alpha_{R'}| \\
&\gtrsim 2^{-n} n^{-(d-2)/2} \sum_{r_1=0}^{n} \sum_{\vec{r}' \in \mathbb{H}_{n-r_1}^{d-1}} \sum_{|R_1|=2^{-r_1}} \sum_{R' \in \mathcal{R}_{\vec{r}'}} |\alpha_{R_1}||\alpha_{R'}| \\
&\gtrsim 2^{-n} n^{-(d-2)/2} \sum_{|R|=2^{-n}} |\alpha_R|,
\end{aligned}$$

which shows the validity of Conjecture 1.1

6.3 The second proof of Theorem 1.3

The second proof of Theorem 1.3 will use a duality argument. That is, for every collection of coefficients $\{\alpha_R\} \subset \mathcal{A}_{\text{split}}$, we will construct a test function $\psi \in L^1([0,1]^d)$ such that

$$\|\psi\|_{L^1} \leq C,$$

where C is a constant which does not depend on the choice of coefficients $\{\alpha_R\} \subset \mathcal{A}_{\text{split}}$ or the integer $n \geq 1$. In addition, the test function ψ will also satisfy

$$\langle \psi, H_n \rangle := \int_{[0,1]^d} \psi H_n dx \gtrsim n^{\frac{d}{2}},$$

where $H_n = \sum_{|R|=2^{-n}} \alpha_R h_R$ with $\{\alpha_R\} \subset \mathcal{A}_{\text{split}}$. Theorem 1.3 follows from the above two inequalities by a simple application of Hölder's inequality to $\langle \psi, H_n \rangle$. We will prove the two inequalities in Lemma 6.4. This lemma is also essential in proving Theorem 1.6. The construction of the test function ψ will be similar to that in Temlyakov's proof of the two-dimensional small ball inequality, given in Section 4.1 of Chapter 4.

Lemma 6.4. *For any $n \geq 1$, $d \geq 3$, and any choice of coefficients $\{\alpha_R\}_{R \in \mathcal{A}_n^d} \subset \mathcal{A}_{split}$, there exists a function $\psi \in L^1([0,1]^d)$ with $\|\psi\|_{L^1} = 1$ such that*

$$\langle \psi, H_n \rangle = \int_{[0,1]^d} \psi H_n dx \gtrsim n^{\frac{d}{2}}. \tag{6.14}$$

Here, the implicit constant depends only on the dimension d and not on the collection of coefficients $\{\alpha_R\}_{R \in \mathcal{A}_n^d} \subset \mathcal{A}_{split}$ or the integer $n \geq 1$.

Proof. Recall the expression of the function

$$H_n(\vec{x}) = \sum_{R \in \mathcal{A}_n^d} \alpha_R h_R(\vec{x}) \tag{6.15}$$

given in (6.1) and define the test function ψ as the Riesz product

$$\psi(\vec{x}) = \prod_{r=0}^{n} \Big(1 + B_r(x_1)\text{sgn}(F_r(\vec{x}'))\Big). \tag{6.16}$$

Here, we have replaced r_1 with r for our convenience. We claim that ψ has the property that

$$\|\psi\|_{L^1([0,1]^d)} = 1 \tag{6.17}$$

and satisfies (6.14). Indeed, we first observe that, for any $\vec{x} \in [0,1]^d$, $\psi(\vec{x}) \geq 0$, since each factor in the product forming ψ is nonnegative. Expanding the product in (6.16), we get

$$\psi = 1 + \psi_1 + \psi_2, \tag{6.18}$$

with

$$\psi_1(\vec{x}) = \sum_{r=0}^{n} B_r(x_1)\mathrm{sgn}(F_r(\vec{x}')), \tag{6.19}$$

and

$$\psi_2(\vec{x}) = \sum_{k=2}^{n} \sum_{0 \leq s_1 < \cdots < s_k \leq n} \prod_{j=1}^{k} B_{s_j}(x_1)\mathrm{sgn}(F_{s_j}(\vec{x}')). \tag{6.20}$$

Now, it is easy to see that

$$\|\psi\|_{L^1([0,1]^d)} = \int_{[0,1]^d} \psi(x)dx = 1 + \int_{[0,1]^d} \psi_1(x)dx + \int_{[0,1]^d} \psi_2(x)dx.$$

Hence, to complete the proof of (6.17), it suffices to show that

$$\int_{[0,1]^d} \psi_1(x)dx = 0 \tag{6.21}$$

and

$$\int_{[0,1]^d} \psi_2(x)dx = 0. \tag{6.22}$$

To show (6.21), we first recall from Proposition 6.1 that $\{B_r\}_{r=0}^n$ are independent random variables of mean zero. Now, for $r \in \{0, \ldots, n\}$, this implies that $\int_{[0,1]} B_r(x_1)dx_1 := \mathbb{E}(B_r) = 0$, where $\mathbb{E}$ is the expected value of B_r with respect to the one-dimensional Lebesgue measure, and this proves (6.21).

Now, fix $k \in \{2, \ldots, n\}$ and $\{s_j\}_{j=1}^k \subset \{0, \ldots, n\}$ with $s_1 < \cdots < s_k$. Again, using the fact that $\{B_r\}_{r=0}^n$ are independent random variables of mean zero, we get that $\int_{[0,1]} \prod_{j=1}^k B_{s_j}(x_1)dx_1 = \mathbb{E}(\prod_{j=1}^k B_{s_j}) = \prod_{j=1}^k \mathbb{E}(B_{s_j}) = \prod_{j=1}^k \int_{[0,1]} B_{s_j}(x_1)dx_1 = 0$, which proves (6.22).

To show that ψ satisfies (6.14), we first use (6.18) to get

$$\langle \psi, H_n \rangle = \int_{[0,1]^d} H_n(x)dx + \langle \psi_1, H_n \rangle + \langle \psi_2, H_n \rangle. \tag{6.23}$$

Now, since $\{B_r\}_{r=0}^n$ are of mean zero, we get that $\int_{[0,1]^d} H_n(x)dx = 0$. Therefore, to conclude (6.14), we show that

$$\langle \psi_1, H_n \rangle \gtrsim n^{\frac{d}{2}} \tag{6.24}$$

and

$$\langle \psi_2, H_n \rangle = 0. \tag{6.25}$$

Since $\{B_r\}_{r=0}^n$ are ± 1-valued independent random variables of mean zero, using (5.31), we get that

$$\begin{aligned} \langle \psi_1, H_n \rangle &= \int_{[0,1]^{d-1}} \sum_{r,r'=0}^{n} \operatorname{sgn}(F_r) F_{r'} \int_{[0,1]} \Big(B_r B_{r'} \Big) \\ &= \sum_{r=0}^{n} \|F_r\|_{L^1([0,1]^{d-1})} \\ &\gtrsim \sum_{r=0}^{n} \|F_r\|_{L^2([0,1]^{d-1})} \\ &\gtrsim \sum_{r=0}^{n} (n-r)^{\frac{d-2}{2}} \gtrsim n^{\frac{d}{2}}, \end{aligned} \tag{6.26}$$

which proves (6.24). Again, using the fact that $\{B_r\}_{r=0}^n$ are ± 1-valued independent random variables of mean zero, from the definitions of the functions ψ_2 and H_n we get that

$$\begin{aligned} \langle \psi_2, H_n \rangle = &\int_{[0,1]^{d-1}} \sum_{r=0}^{n} \sum_{k=2}^{n} \sum_{0 \le s_1 < \cdots < s_k \le n} F_r \prod_{j=1}^{k} \operatorname{sgn}(F_{s_j}) \int_{[0,1]} \prod_{j=1}^{k} \Big(B_{s_j} B_r \Big) \\ =&0, \end{aligned} \tag{6.27}$$

which completes the proof. $\square$

The second proof of Theorem 1.3. Lemma 6.4 ensures the existence of the function ψ with $\|\psi\|_{L^1([0,1]^d)} = 1$ for which

$$\int_{[0,1]^d} \psi H_n dx \gtrsim n^{\frac{d}{2}}. \tag{6.28}$$

Applying Hölder's inequality in (6.28), we get that $\|H_n\|_\infty \gtrsim n^{\frac{d}{2}}$. □

6.4 The sharpness of the small ball inequality with restricted coefficients

In this section, we prove Theorem 1.4: the sharpness of the small ball inequality when the coefficients satisfy the splitting condition. Precisely, this means we will prove that, for all $n \geq 1$ and $d \geq 3$, there exists a choice of coefficients $\{\alpha_R\}_{R \in \mathcal{A}_n^d} \subset \mathcal{A}_{\text{split}}$ for which

$$\left\| \sum_{R \in \mathcal{A}_n^d} \alpha_R h_R \right\|_\infty \lesssim n^{\frac{d}{2}}. \tag{6.29}$$

The proof of Theorem 1.4 relies on two lemmas: the first establishes the independence of a product of random variables, and the second estimates the expected value of the point-wise maximum of a collection of random variables.

Lemma 6.5. *Fix an integer $l \geq 1$. Let $\{X_i, Y_j^i : 1 \leq i \leq m, 1 \leq j \leq l\}$ be a collection of independent random variables taking the values ± 1 with probability $\frac{1}{2}$. Then the random variables $\{X_i Y_j^i : 1 \leq i \leq m, 1 \leq j \leq l\}$ are independent.*

Proof. Let $\overrightarrow{W}_i = (X_i Y_1^i, \ldots, X_i Y_l^i)$, $i = 1, \ldots, m$. First, we claim that, for $\vec{\epsilon_i} \in$

$\{-1,1\}^l$ and $i=1,\dots,m$, $\mathbb{P}\Big(\overrightarrow{W}_i=\vec{\epsilon_i}\Big)=\frac{1}{2^l}$. To see this, observe that

$$\begin{aligned}\mathbb{P}\Big(\overrightarrow{W}_i=\vec{\epsilon_i}\Big) =&\mathbb{P}(X_iY_1^i=\epsilon_1^i,\dots,X_iY_l^i=\epsilon_l^i)\\ =&\sum_{b_i\in\{-1,1\}}\mathbb{P}\Big(Y_1^i=\frac{\epsilon_1^i}{b_i},\dots,Y_l^i=\frac{\epsilon_l^i}{b_i},X_i=b_i\Big)\\ =&\frac{1}{2^l}\frac{2}{2}\\ =&\frac{1}{2^l}.\end{aligned}$$

Now, we prove the independence of $\{X_iY_j^i:1\le i\le m,1\le j\le l\}$. To prove this, we need to show that

$$\mathbb{P}\Big(\overrightarrow{W}_1=\vec{\epsilon}_1,\dots,\overrightarrow{W}_m=\vec{\epsilon}_m\Big)=\Big(\frac{1}{2^l}\Big)^m.$$

Let $\vec{b}=(b_1,\dots,b_m)\in\{-1,1\}^m$, $\vec{X}=(X_1,\dots,X_m)$ and $\vec{Y}_{l,i}=(Y_1^i,\dots,Y_l^i)$, $i=1,\dots,m$. The independence of the random variables $\{X_i,Y_j^i:1\le i\le m,1\le j\le l\}$ gives us that

$$\begin{aligned}\mathbb{P}\Big(\overrightarrow{W}_1=\vec{\epsilon}_1,\dots,\overrightarrow{W}_m=\vec{\epsilon}_m\Big)=&\sum_{\vec{b}\in\{-1,1\}^m}\mathbb{P}\Big(\vec{Y}_{l,1}=\frac{\vec{\epsilon}_1}{b_1},\dots,\vec{Y}_{l,m}=\frac{\vec{\epsilon}_m}{b_m},\vec{X}=\vec{b}\Big)\\ =&\Big(\frac{1}{2^l}\Big)^m\frac{2^m}{2^m}\\ =&\Big(\frac{1}{2^l}\Big)^m.\end{aligned}$$

□

Lemma 6.6. *Let $Y_1,\dots,Y_N$ be random variables on a probability space $(\Omega,\mathbb{P})$ and $\phi:[0,\infty)\to\mathbb{R}$ be a convex and strictly increasing function. If $\mathbb{E}\phi(|Y_k|)\le C,k=1,\dots,N$, then we have that*

$$\mathbb{E}\Big(\sup_{1\le k\le N}|Y_k|\Big)\le\phi^{-1}(N),$$

where C is an absolute constant and ϕ^{-1} is the inverse function of ϕ.

The proof of this lemma, which we omit here, can be found in Lacey's excellent survey [34] on the small inequality and discrepancy theory. The proof is a generalization of the proof of the special case when $\phi(t) = \exp(t^2)$, $t \geq 0$, which was presented in Chapter 4 in the proof of the sharpness of the d-dimensional small ball inequality.

Proof of Theorem 1.4. Suppose that the coefficients $\{\alpha_{R_1}\alpha_{R'}\}$ in $\mathcal{A}_{split}$ are chosen randomly and independently, *i.e.*, consider a probability space $(\Omega, \mathbb{P})$ where $\{\alpha_{R_1}\}$ and $\{\alpha_{R'}\}$ are two families of ± 1-valued iid random variables which are independent of each other. To prove (6.29), it suffices to show that

$$\mathbb{E}\Big\| \sum_{R \in \mathcal{A}_n^d} \alpha_R h_R \Big\|_\infty \lesssim n^{\frac{d}{2}}, \tag{6.30}$$

where $\mathbb{E}$ denotes the expected value with respect to the probability $\mathbb{P}$.

Let Q_k be a dyadic cube in $\mathcal{D}^d$ of side-length $2^{-(n+1)}$, $k = 1, \dots, 2^{(n+1)d}$. Note that the collection of these dyadic cubes partitions the unit cube $[0,1)^d$. Set

$$Z_k(\omega, \vec{x}) = \sum_{|R|=2^{-n}} \alpha_R h_R|_{Q_k} = \sum_{\vec{r} \in \mathbb{H}_n^d} \sum_{Q_k \subset R \in \mathcal{R}_{\vec{r}}} \alpha_{R_1}(\omega)\alpha_{R'}(\omega) h_R(\vec{x}), (\omega, \vec{x}) \in \Omega \times Q_k,$$

and fix any $\vec{x} \in Q_k$. Note that $\#\{R \in \mathcal{A}_n^d, Q_k \subset R\} = \#\mathbb{H}_n^d$. Using this fact and applying Lemma 6.5 to the random variables $\{\alpha_{R_1}\}$ and $\{\alpha_{R'}\}$, we get that Z_k is a linear combination of $\#\mathbb{H}_n^d$ many iid ± 1-valued random variables. Applying Khintchine's inequalities yields

$$C_d^{-1} n^{\frac{d-1}{2}} \leq \mathbb{E}(|Z_k|) \leq C_d n^{\frac{d-1}{2}}. \tag{6.31}$$

Set $Y_k = \frac{|Z_k|}{\mathbb{E}|Z_k| A}$, where A is a large constant to be determined. We claim that there exists A, depending on the dimension d, for which

$$\mathbb{E}\phi(Y_k) \lesssim 1, \tag{6.32}$$

where $\phi(t) = \exp(\frac{t^2}{2})$, $t > 0$, for every $k \in \{1, \ldots, 2^{(n+1)d}\}$. To proceed, we will first show how (6.30) follows from (6.32), and then we will prove the claim.

To see how (6.30) follows from the (6.32), let $N = 2^{(n+1)d}$. Lemma 6.6 implies that

$$\mathbb{E}\left(\sup_{1\le k\le N} |Y_k|\right) \lesssim \sqrt{n}. \tag{6.33}$$

Combining (6.31) and (6.33) and using the fact that Z_k is constant on the dyadic cube Q_k for a fixed $\omega \in \Omega$, we get that

$$\begin{aligned}
\mathbb{E}\Big\| \sum_{R\in\mathcal{A}_n^d} \alpha_R h_R \Big\|_\infty =& \mathbb{E}\left(\sup_{1\le k\le N} |Z_k|\right) \\
=& \mathbb{E}\left(\sup_{1\le k\le N} Y_k \mathbb{E}|Z_k| A\right) \\
\lesssim& \, n^{\frac{d-1}{2}} \mathbb{E}\left(\sup_{1\le k\le N} Y_k\right) \\
\lesssim& \, n^{\frac{d-1}{2}} \sqrt{n} \\
\lesssim& n^{\frac{d}{2}}.
\end{aligned}$$

Now we prove the claim. First, we estimate the distribution function of $\phi(Y_k)$ as follows:

$$\begin{aligned}
\mathbb{P}(\phi(Y_k) > \lambda) =& \mathbb{P}(Y_k > \phi^{-1}(\lambda)) \\
=& \mathbb{P}(Y_k > \sqrt{2\log\lambda}) \\
=& \mathbb{P}(|Z_k| > \mathbb{E}|X_k| A \sqrt{2\log\lambda}) \\
\le& \mathbb{P}(|Z_k| > A C_d^{-1} n^{\frac{d-1}{2}} \sqrt{2\log\lambda}).
\end{aligned}$$

Therefore, using Bernstein's inequality, stated in Lemma 4.2, we get

$$\begin{aligned}\mathbb{P}(\phi(Y_k) > \lambda) \leq& 2\exp\left(\frac{n^{d-1}\log(\lambda^{-A^2C_d^{-2}2})}{4\#\mathbb{H}_n^d}\right)\\ \leq& 2\exp\left(\frac{n^{d-1}\log(\lambda^{-A^2C_d^{-2}2})}{4n^{d-1}C_d'}\right)\\ \lesssim& \frac{1}{\lambda^{\frac{A^2C_d^{-2}}{2C_d'}}}\\ \lesssim& \frac{1}{\lambda^2},\end{aligned}$$

for A chosen large enough. Now, applying the distribution identity,

$$\mathbb{E}(\phi(Y_k)) = \int_0^\infty \mathbb{P}(\phi(Y_k) > \lambda)d\lambda,$$

we have that

$$\mathbb{E}(\phi(Y_k)) \lesssim \int_0^1 d\lambda + \int_1^\infty \lambda^{-2}d\lambda \leq C_d''.$$

□

Chapter 7

Orlicz space bounds for special classes of hyperbolic sums

As we have already mentioned, the L^p-bounds of the signed hyperbolic sums, $H_n = \sum_{|R|=2^{-n}} \alpha_R h_R$ with $\{\alpha_R\} \subset \{-1, 1\}$, are equivalent to the L^2-bounds. Indeed, applying Proposition 5.6, there exist positive constants $C_i(2, p, d)$, $i = 1, 2$, such that

$$C_1(2, p, d)\|H_n\|_2 \leq \|H_n\|_p \leq C_1(2, p, d)\|H_n\|_2, \ 0 < p < \infty.$$

Using the orthogonality of Haar functions, we obtain $\|H_n\|_2 = \|S_d(H_n)\|_2 \simeq n^{\frac{d-1}{2}}$. Therefore, $\|H_n\|_p \simeq n^{\frac{d-1}{2}}$ for all p in $(0, \infty)$, where the implicit constants depend on the value of p and the dimension d.

In this chapter, we explore lower bounds of signed hyperbolic sums in exponential Orlicz spaces. We have chosen to work with these spaces because, for $1 \leq p < \infty$, exponential Orlicz spaces can be considered as intermediate function spaces between L^p and L^∞, and the L^∞-lower bounds of signed hyperbolic sums remain conjectured for all $d \geq 3$ whereas the L^p-lower bounds are well-understood. Also, lower bounds of the discrepancy function have already been studied in exponential Orlicz spaces in dimension $d = 2$ [6]. We are interested in obtaining analogous lower bounds in the case of signed hyperbolic sums. In dimension $d = 2$, we will adjust Temlyakov's proof of the two-dimensional small ball inequality for signed hyperbolic sums to the framework of exponential Orlicz spaces, and show that the lower bounds obtained for the discrepancy function can also be realized for signed hyperbolic sums. In higher dimensions, we will obtain lower bounds of signed hyperbolic sums in exponential Orlicz spaces under the additional assumption that the coefficients α_R are from the set $\mathcal{A}_{\text{split}}$.

Before proving our main results, we first present the background material on Orlicz spaces needed to follow the proofs of the main theorems of this chapter, Theorem 1.6 and Theorem 7.4.

7.1 Orlicz spaces

For $p \geq 1$, Orlicz spaces are natural generalizations of L^p-spaces. To see this, note that $|x|^p$ is a convex function for $p \geq 1$, and a function f is in $L^p(X, \mu)$ if

$$\int_X |f|^p \, d\mu < \infty,$$

whereas, roughly speaking, f is in the Orlicz space associated to a convex function $\phi : \mathbb{R} \to [0, \infty)$ if

$$\int_X \phi(f) \, d\mu < \infty.$$

More precisely, let $\phi : \mathbb{R} \to [0, \infty]$ be a convex, even function with $\phi(0) = 0$ and $\phi \not\equiv 0$. Given a finite measure space $(X, \mathcal{M}, \mu)$, one defines the Orlicz space associated to the function ϕ as

$$L_\phi(X, \mu) = \left\{ f : X \to \mathbb{R} \text{ measurable} : \exists a > 0, \int_X \phi(af) d\mu < \infty \right\}.$$

We endow Orlicz spaces with the so-called Luxembourg norm: for any measurable function $f \in L_\phi(X, \mu)$, the norm of f is defined by

$$\|f\|_\phi := \inf \left\{ t > 0 : \int_X \phi(f/t) d\mu \leq 1 \right\}.$$

This is precisely the Minkowski functional associated to the set

$$V = \Big\{f \;\; \text{measurable} : \int_X \phi(f)d\mu \leq 1\Big\}.$$

Note that

$$L_\phi(X,\mu) = \Big\{f \text{ measurable} : \|f\|_\phi < \infty\Big\}.$$

Such a function ϕ is increasing on $\mathbb{R}_+ := [0,\infty)$. Indeed, let $0 \leq s_1 < s_2$, then the convexity of ϕ implies that

$$\phi(s_1) \leq \frac{s_2 - s_1}{s_2}\phi(0) + \frac{s_1}{s_2}\phi(s_2) \leq \phi(s_2).$$

From the definition we can see that, to understand the composition of $L_\phi(X,\mu)$, we only need to know how $\phi(x)$ grows as $|x| \to \infty$. More specifically, if $\phi \sim h$ as $|x| \to \infty$, *i.e.*,

$$C_1 h \leq \phi \leq C_2 h, \;\; |x| \geq M,$$

for some positive constants C_1, C_2, M, then the two quantities $\|f\|_\phi$ and $\|f\|_h$ are equivalent and we see that it suffices to know the asymptotic behaviour of ϕ.

Orlicz spaces coincide with L^p spaces for certain choices of ϕ. If $\phi(x) = |x|^p$ $(1 \leq p < \infty)$, then $L_\phi(X,\mu) = L^p(X,\mu)$ and $\|f\|_\phi = \|f\|_p$, and if

$$\phi(x) = \begin{cases} 0, & \text{if } |x| \leq 1 \\ \infty, & \text{otherwise} \end{cases},$$

then $L_\phi(X,\mu) = L^\infty(X,\mu)$ and the $\|f\|_\phi = \|f\|_\infty$. We will be interested in the Orlicz spaces which are associated with functions of the form

$$\phi_a(x) = \begin{cases} e^{|x|^a} - 1, & |x| \geq 1 \\ (e-1)|x|, & \text{otherwise} \end{cases},$$

where $a > 0$. These spaces are known as exponential Orlicz spaces and are denoted by $\exp(L^a)$.

To prove our main results, we will need an Orlicz space version of Hölder's inequality, *i.e.*, given ϕ as in the definition of L_ϕ, we want to find $C > 0$ and a function ϕ^* such that

$$\int |fg| d\mu \leq C \|f\|_\phi \|g\|_{\phi^*}, \tag{7.1}$$

for all $f \in L_\phi$ and $g \in L_{\phi*}$. Note that this requires that ϕ^* is an even convex function with $\phi^*(0) = 0$. To prove this generalization of Hölder's inequality, ϕ^* is taken to be the function conjugate to ϕ, *i.e.*, $\phi^* : \mathbb{R} \to [0, \infty]$ is the function defined by

$$\phi^*(y) = \sup\{xy - \phi(x) : x \in \mathbb{R}\}. \tag{7.2}$$

It follows from the definition that ϕ^* is convex, even and satisfies $\phi^*(0) = 0$. Another immediate consequence of this definition is Young's inequality:

$$xy \leq \phi(x) + \phi^*(y)$$

for all x and y. Young's inequality implies that

$$|f(x)g(x)| \leq \phi(f(x)) + \phi^*(g(x)), \text{ for all } \ x \in X. \tag{7.3}$$

We can now see that (7.1) holds for $C = 2$. Indeed, it suffices to show (7.1) for $\|f\|_\phi = \|g\|_{\phi^*} = 1$ and, in such a case, we obtain (7.1) simply by integrating both sides of inequality (7.3).

This generalization of Hölder's inequality will be applied to ϕ_a and ϕ_a^* to prove the main results of the chapter. Note that, by convention, the Orlicz spaces associated to ϕ_a^* are denoted by $L(\log L)^{\frac{1}{a}}$, and it can be verified that $\phi_a^*(y) \sim |y|(\log(1 + |y|))^{\frac{1}{a}}$ as $|y| \to \infty$.

The next proposition deals with estimating Orlicz space norms of an indicator function. A detailed proof can be found in [37].

Proposition 7.1. *Let $E \in \mathcal{M}$. Then*

$$\frac{1}{2}\mu(E)(\phi_a)^{-1}\left(\frac{1}{\mu(E)}\right) \leq \| \mathbb{1}_E \|_{\phi_a^*} \leq \mu(E)(\phi_a)^{-1}\left(\frac{1}{\mu(E)}\right),$$

where $(\phi_a)^{-1}$ *is the inverse of* $\phi_a(x)$, $x \geq 0$.

Applying the above proposition to a probability measure gives us that

$$\| \mathbb{1}_E \|_{L(\log L)^{\frac{1}{a}}} \leq \mathbb{P}(E) \cdot (1 - \log(\mathbb{P}(E))^{\frac{1}{a}}. \tag{7.4}$$

7.2 The proof of Theorem 1.6

Before we proceed to the proof of Theorem 1.6, note that it follows from the definition of exponential Orlicz spaces that, for each $1 < p < \infty$ and $a > 0$, we have continuous embeddings $L^\infty \subset \exp(L^a) \subset L^p$. Therefore,

$$n^{\frac{d-1}{2}} \lesssim \|H_n\|_2 \lesssim \|H_n\|_{\exp(L^a)} \tag{7.5}$$

holds for any choice of coefficients $\{\alpha_R\} \subset \mathcal{A}_{\text{split}}$. In fact, (7.5) holds for any collection $\{\alpha_R\} \subset \{-1, 1\}$. We can see that the lower bound in (7.5) is better than the one given in Theorem 1.6 when $a < 2$, but when $2 \leq a < \infty$, inequality (1.16) from the statement of Theorem 1.6 gives us an improvement over the lower bound in (7.5) by a factor of $n^{\frac{1}{2}-\frac{1}{a}}$ for any collection of coefficients $\{\alpha_R\} \subset \mathcal{A}_{\text{split}}$. It is for this reason that Theorem 1.6 is stated only for $a \geq 2$. As we will soon see, the proof of Theorem 1.6 works for all $a > 0$.

Proof of Theorem 1.6. From Lemma 6.4 there is a positive function ψ such that $\|\psi\|_{L^1([0,1]^d)} = 1$ and

$$\int_{[0,1]^d} \psi H_n dx \gtrsim n^{\frac{d}{2}}. \tag{7.6}$$

Also note that this test function can be written as

$$\psi = 2^{n+1} \mathbb{1}_E,$$

where

$$E = \left\{ \vec{x} = (x_1, \vec{x}') \in [0,1]^d : B_r(x_1) \cdot \operatorname{sgn}(F_r(\vec{x}')) = 1 \text{ for all } r = 0, \ldots, n \right\}.$$

Furthermore, $|E| = 2^{-(n+1)}$ since $\int_{[0,1]^d} \psi dx = 1$. Applying Hölder's inequality for Orlicz spaces to (7.6), we get

$$\|2^{n+1}\,\mathbb{1}_E\,\|_{L(\log L)^{\frac{1}{a}}} \|H_n\|_{\exp(L^a)} \gtrsim n^{\frac{d}{2}}. \tag{7.7}$$

Combining (7.4) with the fact that $|E| = \frac{1}{2^{n+1}}$, we obtain

$$\|2^{n+1}\,\mathbb{1}_E\,\|_{L(\log L)^{\frac{1}{a}}} \leq C_d 2^{\frac{1}{a}} n^{\frac{1}{a}},\ n > 1. \tag{7.8}$$

Now, using (7.8), (7.7) becomes

$$\|H_n\|_{\exp(L^a)} \geq (C_d)^{-1} 2^{-\frac{1}{a}} n^{\frac{d}{2}-\frac{1}{a}},$$

which completes the proof. □

From the proof of Theorem 1.6 we can see that the constant $C(d, a)$ appearing in (1.16) must be equal to $(C_d)^{-1}2^{-\frac{1}{a}}$. Taking the limit of both sides of inequality (1.16) as the scale of integrability a approaches infinity, we get the signed small ball inequality when $\{\alpha_R\} \subset \mathcal{A}_{\text{split}}$.

It is worth mentioning that bounds of this type have already appeared in the field of irregularities of distribution of points (see [6]). In particular, Bilyk *et al.* [6] showed that, in $d = 2$, the discrepancy function,

$$D_N(x_1, x_2) := \#\Big(\mathcal{P}_N \cap [0, x_1) \times [0, x_2)\Big) - N x_1 x_2,$$

associated to the N-point set $\mathcal{P}_N \subset [0,1]^2$ satisfies the following lower bound

$$\|D_N\|_{\exp(L^a)} \gtrsim (\log N)^{1-\frac{1}{a}}, \quad 2 \leq a < \infty.$$

They used a duality argument in the setting of exponential Orlicz spaces in their proof which, in fact, provided the motivation for the proof of Theorem 1.6.

One can also use a duality argument in framework of exponential Orlicz spaces to prove Conjecture 1.5 in $d = 2$. Indeed, recall Temlyakov's test function used in the proof of the two-dimensional small ball inequality,

$$\Psi = \prod_{k=0}^{n}(1 + f_{(k,n-k)}) = 2^{n+1}\,\mathbb{1}_E,$$

where

$$E = \{\vec{x} \in [0,1]^2 : f_{(k,n-k)} = 1, \text{ for all } k \in \{0, 1, \dots, n\}\}.$$

Adapting the proof of the two-dimensional small ball inequality to the signed hyperbolic sum $H_n = \sum_{k=0}^{n} f_{(k,n-k)}$, we have

$$\langle \Psi, H_n \rangle \gtrsim n.$$

Applying Hölder's inequality for Orlicz spaces as in the proof of Theorem 1.6, we get $\|H_n\|_{\exp(L^a)} \gtrsim n^{1-\frac{1}{a}}$, $2 \le a < \infty$.

7.3 Study of signed hyperbolic sums

In this section, we are concerned with exploring L^∞ and $\exp(L^a)$-lower bounds of a linear combination of Haar functions with signed coefficients $\alpha_R \in \{-1, 1\}$. To motivate our discussion, we first restate the signed small ball conjecture and Conjecture 1.5 in an equivalent form.

Conjecture 7.2. (Restatement of the signed small ball conjecture) There exists a vector $\vec{\mu} = (\mu_0, \dots, \mu_n) \in \{-1, 1\}^{n+1}$ such that, for any choice of coefficients $\{\beta_R\} \subset \{-1, 1\}$ and any integer $n \ge 1$, we have the inequality

$$\Big\| \sum_{\vec{r} \in \mathbb{H}_n^d} \mu_{r_1} \sum_{R \in \mathcal{R}_{\vec{r}}} \beta_R h_R \Big\|_\infty \gtrsim n^{\frac{d}{2}}, \quad d \ge 3. \tag{7.9}$$

By choosing $\mu_{r_1} = 1$, $r_1 = 0, 1, \dots, n$ and $\beta_R = \alpha_R$, $R \in \mathcal{A}_n^d$, it is immediate that the small ball conjecture implies (7.9). To get the reverse implication, (7.9) implies the small ball conjecture, simply set $\beta_R = \mu_{r_1}\alpha_R$ for $R = R_1 \times R' \in \mathcal{A}_n^d$ with $|R_1| = 2^{-r_1}, r_1 = 0, \dots, n$.

Conjecture 7.3. (Restatement of Conjecture 1.5) There exists a vector $\vec{\mu} = (\mu_0, \ldots, \mu_n) \in \{-1,1\}^{n+1}$ such that, for any choice of coefficients $\{\beta_R\} \subset \{-1,1\}$ and any integer $n > 1$, we have the inequality

$$\Big\| \sum_{\vec{r} \in \mathbb{H}_n^d} \mu_{r_1} \sum_{R \in \mathcal{R}_{\vec{r}}} \beta_R h_R \Big\|_{\exp(L^a)} \gtrsim n^{\frac{d-1}{2} + \frac{1}{2} - \frac{1}{a}}, \quad 2 \le a < \infty, \tag{7.10}$$

for all $d \ge 3$.

The equivalence can be obtained by methods similar to those presented before.

Using the techniques developed in the previous section, we will show the existence of $(n+1)$-dimensional vectors $\vec{\mu}$ and $\vec{\nu} \in \{-1,1\}^{n+1}$ which satisfy inequalities (7.9) and (7.10) respectively, though these vectors may depend on the choice of coefficients $\{\alpha_R\} \subset \{-1,1\}$. Specifically, we show:

Theorem 7.4. *For any integer $d \ge 3$ and any choice of coefficients $\{\alpha_R\} \subset \{-1,1\}$, there exist $(n+1)$-dimensional vectors $\vec{\mu} = (\mu_0, \ldots, \mu_n) \in \{-1,1\}^{n+1}$, $n \ge 1$, and $\vec{\nu} = (\nu_0, \ldots, \nu_n) \in \{-1,1\}^{n+1}$, $n > 1$, such that*

$$\Big\| \sum_{\vec{r} \in \mathbb{H}_n^d} \mu_{r_1} \sum_{R \in \mathcal{R}_{\vec{r}}} \alpha_R h_R \Big\|_{\infty} \gtrsim n^{\frac{d}{2}}, \tag{7.11}$$

and

$$\Big\| \sum_{\vec{r} \in \mathbb{H}_n^d} \nu_{r_1} \sum_{R \in \mathcal{R}_{\vec{r}}} \alpha_R h_R \Big\|_{\exp(L^a)} \gtrsim n^{\frac{d-1}{2} + \frac{1}{2} - \frac{1}{a}}, \tag{7.12}$$

for $2 \le a < \infty$.

This theorem will be proved by employing a duality argument which uses a test function in the form of a Riesz product. Instead of working with the function

$$\sum_{\vec{r} \in \mathbb{H}_n^d} \mu_{r_1} \sum_{R \in \mathcal{R}_{\vec{r}}} \alpha_R h_R,$$

we let $\mu_{r_1} : \Omega \to \{-1,1\}$ be iid random variables with $\mathbb{P}(\mu_{r_1} = \pm 1) = \frac{1}{2}$, $r_1 =$

$0, 1, \ldots, n$ and consider the function

$$H_{\vec{\mu}}(\omega, \vec{x}) = \sum_{\vec{r} \in \mathbb{H}_n^d} \mu_{r_1}(\omega) \sum_{R \in \mathcal{R}_{\vec{r}}} \alpha_R h_R(\vec{x})$$

defined on the probability space $\big(\Omega \times [0,1)^d, \mathbb{P} \times |\cdot|\big)$. Then, for a particular choice of $\omega \in \Omega$, $H_{\vec{\mu}}(\omega, \cdot)$ takes the form of the function inside the L^∞ and $\exp(L^a)$-norm in (7.9) and (7.10) respectively. The usage of these random variables combined with (5.33) will be crucial in getting the lower bounds in (7.11) and (7.12).

The idea of introducing random coefficients for proving existence results is not new; as we saw in Chapter 4, it was already used to prove the sharpness of the conjectured lower bound in the small ball inequality for all $d \geq 3$. However, the randomization used in Chapter 4 required each coefficient in the collection $\{\alpha_R\}$ to be an iid ± 1-valued random variable, and there were about $n^{d-1} 2^n$ iid random coefficients needed for that proof, whereas the randomization used here is considerably milder, employing just $n+1$ iid random variables.

Just as we proved the lower bound appearing in Theorem 1.3 separately in Lemma 6.4, we will prove the lower bound appearing in Theorem 7.4 in the following lemma.

Lemma 7.5. *Let $d \geq 3$ and $n \geq 1$. For each collection of coefficients $\alpha_R \in \{-1, 1\}$, there exists a positive test function $\psi_{\vec{\mu}} \in L^1(\Omega \times [0,1]^d)$ such that*

$$\|\psi_{\vec{\mu}}\|_{L^1([0,1]^d \times \Omega)} = 1 \tag{7.13}$$

and

$$\mathbb{E}_{\vec{\mu}} \mathbb{E}(\psi_{\vec{\mu}} H_{\vec{\mu}}) \gtrsim n^{\frac{d}{2}}, \tag{7.14}$$

where $\mathbb{E}_{\vec{\mu}}$ is the expected value with respect to the probability measure $\mathbb{P}$ and $\mathbb{E}(\psi_{\vec{\mu}} H_{\vec{\mu}}) = \int_{[0,1]^d} (\psi_{\vec{\mu}} H_{\vec{\mu}}) dx$ is the expected value with respect to the Lebesgue measure.

Proof. We want to find a test function $\psi_{\vec{\mu}} \in L^1(\Omega \times [0,1]^d)$ such that

$$\|\psi_{\vec{\mu}}\|_{L^1(\Omega \times [0,1]^d)} = 1 \tag{7.15}$$

and

$$\mathbb{E}_{\vec{\mu}}\mathbb{E}(\psi_{\vec{\mu}}H_{\vec{\mu}}) \gtrsim n^{\frac{d}{2}}. \tag{7.16}$$

Note first that

$$H_{\vec{\mu}}(\omega,\vec{x}) = \sum_{r_1=0}^{n} \mu_{r_1}(\omega)A_{r_1}(\vec{x}),$$

where

$$A_{r_1} = \sum_{\vec{r}'\in\mathbb{H}^{d-1}_{n-r_1}} \sum_{R\in\mathcal{R}_{\vec{r}}} \alpha_R h_R,$$

with $(r_1,\vec{r}') \in \mathbb{H}^d_n$. Consider the function $\psi_{\vec{\mu}}$ defined by

$$\psi_{\vec{\mu}}(\omega,\vec{x}) = \prod_{r=0}^{n} \left(1+\mu_r(\omega)\operatorname{sgn}(A_r(\vec{x}))\right), \tag{7.17}$$

where we have replaced r_1 with r for our convenience. Observe that, for all $(\omega,\vec{x}) \in \Omega\times[0,1]^d$, $\psi_{\vec{\mu}}(\omega,\vec{x}) \geq 0$, since each factor in the product is nonnegative. Expanding the product in (7.17), we get

$$\psi_{\vec{\mu}} = 1+\psi_1+\psi_2, \tag{7.18}$$

with

$$\psi_1(\omega,\vec{x}) = \sum_{r=0}^{n} \mu_r(\omega)\operatorname{sgn}(A_r(\vec{x})), \tag{7.19}$$

and

$$\psi_2(\omega,\vec{x}) = \sum_{k=2}^{n} \sum_{0\leq s_1<\cdots<s_k\leq n} \prod_{j=1}^{k} \mu_{s_j}(\omega)\operatorname{sgn}(A_{s_j}(\vec{x})). \tag{7.20}$$

Hence,

$$\|\psi_{\vec{\mu}}\|_{L^1(\Omega\times[0,1]^d)} = \mathbb{E}\mathbb{E}_{\vec{\mu}}\psi_{\vec{\mu}} = 1+\mathbb{E}\mathbb{E}_{\vec{\mu}}\psi_1+\mathbb{E}\mathbb{E}_{\vec{\mu}}\psi_2.$$

Therefore, to complete the proof of (7.15), it suffices to show that, for every $\vec{x} \in [0,1]^d$,

$$\mathbb{E}_{\vec{\mu}}\psi_1 = \mathbb{E}_{\vec{\mu}}\psi_2 = 0,$$

but this follows from the fact that $\{\mu_r\}_{r=0}^n$ are iid ± 1-valued random variables.

Now, we'll show that $\psi_{\vec{\mu}}$ satisfies (7.16). Using (7.18), we have

$$\mathbb{E}_{\vec{\mu}}\mathbb{E}(\psi_{\vec{\mu}}H_{\vec{\mu}}) = \mathbb{E}_{\vec{\mu}}\mathbb{E}H_{\vec{\mu}} + \mathbb{E}_{\vec{\mu}}\mathbb{E}(\psi_1 H_{\vec{\mu}}) + \mathbb{E}_{\vec{\mu}}\mathbb{E}(\psi_2 H_{\vec{\mu}}), \tag{7.21}$$

and, since $\mathbb{E}_{\vec{\mu}}(\mu_r) = 0$, $r = 0, \ldots, n$, we get that $\mathbb{E}_{\vec{\mu}}H_{\vec{\mu}} = 0$ for every $x \in [0,1]^d$. Therefore, it suffices to show that

$$\mathbb{E}_{\vec{\mu}}\mathbb{E}(\psi_1 H_{\vec{\mu}}) \gtrsim n^{\frac{d}{2}}, \tag{7.22}$$

and

$$\mathbb{E}_{\vec{\mu}}\mathbb{E}(\psi_2 H_{\vec{\mu}}) = 0. \tag{7.23}$$

Appealing to the fact that $\{\mu_r\}_{r=0}^n$ are iid random variables and using (5.33), we obtain (7.22) as follows:

$$\begin{aligned}\mathbb{E}_{\vec{\mu}}\mathbb{E}(\psi_1 H_{\vec{\mu}}) =& \mathbb{E}\sum_{r,r'=0}^{n} \operatorname{sgn}(A_r)A_{r'}\mathbb{E}_{\vec{\mu}}\Big(\mu_r\mu_{r'}\Big) \\ =& \sum_{r=0}^{n} \|A_r\|_{L^1([0,1]^d)} \\ \gtrsim& \sum_{r=0}^{n} \|A_r\|_{L^2([0,1]^d)} \\ \gtrsim& \sum_{r=0}^{n} (n-r)^{\frac{d-2}{2}} \gtrsim n^{\frac{d}{2}}.\end{aligned} \tag{7.24}$$

Now, using the definition of the functions ψ_2 and $H_{\vec{\mu}}$, and, again, the fact that $\{\mu_r\}_{r=0}^n$ are iid random variables, it is clear that we have

$$\begin{aligned}\mathbb{E}_{\vec{\mu}}\mathbb{E}(\psi_2 H_{\vec{\mu}}) =& \mathbb{E}\sum_{r=0}^{n}\sum_{k=2}^{n}\sum_{0\le s_1<\cdots<s_k\le n} A_r\prod_{j=1}^{k}\operatorname{sgn}(A_{s_j})\mathbb{E}_{\vec{\mu}}\prod_{j=1}^{k}\Big(\mu_{s_j}\mu_r\Big) \\ =& 0,\end{aligned} \tag{7.25}$$

which gives (7.23) and completes the proof. ☐

Proof of Theorem 7.4. Applying Hölder's inequality to (7.16) we get

$$\|H_{\vec{\mu}}\|_{L^\infty(\Omega\times[0,1]^d)} \gtrsim n^{\frac{d}{2}}.$$

Therefore, there exists $\omega_0 \in \Omega$ such that $\|H_{\vec{\mu}}(\omega_0,\cdot)\|_{L^\infty([0,1]^d)} \gtrsim n^{\frac{d}{2}}$ and we see that (7.11) holds.

To prove (7.12), we note first that the test function can also be written as

$$\psi_{\vec{\mu}} = 2^{n+1}\,\mathbb{1}_E,$$

where

$$E = \Big\{(\omega,\vec{x}) \in \Omega \times [0,1]^d : \mu_r(\omega)\cdot \mathrm{sgn} A_r(\vec{x}) = 1 \text{ for all } r = 0,\dots,n\Big\}.$$

Since $\mathbb{E}_{\vec{\mu}}\mathbb{E}\psi_{\vec{\mu}} = 1$, it is clear that

$$(\mathbb{P}\times|\cdot|)(E) = 2^{-(n+1)}.$$

In addition, from (7.4) we have that $\|\psi_{\vec{\mu}}\|_{L(\log L)^{\frac{1}{a}}} \le 2^{\frac{1}{a}} n^{\frac{1}{a}}$. Now, applying Hölder's inequality for Orlicz spaces gives

$$\|H_{\vec{\mu}}\|_{\exp(L^a_{\omega,\vec{x}})} \ge C(d,a) n^{\frac{d}{2}-\frac{1}{a}}, \tag{7.26}$$

where $C(d,a)$ is a positive constant depending on the dimension d and the scale of integrability a. Here, $\|\cdot\|_{\exp(L^a_{\omega,\vec{x}})}$ denotes the exponential Orlicz space norm with respect to the measure $\mathbb{P}\times|\cdot|$. We claim that there exists $\omega_0 \in \Omega$ such that $\|H_{\vec{\mu}}(\omega_0,\cdot)\|_{\exp(L^a)} \ge \frac{1}{2}C(d,a)n^{\frac{d}{2}-\frac{1}{a}}$, where $\|\cdot\|_{\exp(L^a)}$ denotes the exponential Orlicz space norm with respect to Lebesgue measure only. Indeed, using the definition of exponential Orlicz space, (7.26) implies that

$$\int_\Omega \int_{[0,1]^d} \phi_a \left(\frac{H_{\vec{\mu}}}{(2)^{-1} C(d,a) n^{\frac{d}{2}-\frac{1}{a}}} \right) d\mathbb{P} dx > 1.$$

This shows that there exists $\omega_0 \in \Omega$ such that

$$\int_{[0,1]^d} \phi_a \left(\frac{H_\mu(\omega_0, \cdot)}{(2)^{-1} C(d,a) n^{\frac{d}{2}-\frac{1}{a}}} \right) dx > 1,$$

so $\|H_{\vec{\mu}}(\omega_0, \cdot)\|_{\exp(L^a)} \geq \frac{1}{2} C(d,a) n^{\frac{d}{2}-\frac{1}{a}}$ which completes the proof of Theorem 7.4. $\square$

Chapter 8

Conclusions and future projects

In Chapter 6 we gave two proofs of the signed small ball inequality under the additional assumption that the coefficients are in $\mathcal{A}_{\text{split}}$. Imposing the splitting property allowed us to exploit the independence of the random variables $\{\sum_{|R_1|=2^{-r_1}} \alpha_{R_1} h_{R_1}, r_1 = 0, \ldots, n\}$ defined on the probability space $([0,1), |\cdot|)$. This, together with the Littlewood-Paley inequalities for Haar functions, enabled us to show that the signed small ball inequality holds in all dimensions $d \geq 3$. In particular, it was shown that

$$\|H_n\|_\infty \gtrsim \sum_{r_1=0}^{n} \|F_{r_1}\|_1 \gtrsim \sum_{r_1=0}^{n} \|F_{r_1}\|_2 \gtrsim \sum_{r_1=0}^{n} (n-r_1)^{\frac{d-2}{2}} \gtrsim n^{\frac{d}{2}}.$$

It is natural to wonder if the signed small ball inequality might also hold for coefficients which satisfy a different splitting property. That is, consider the collection of coefficients

$$\mathcal{A}'_{\text{split}} = \{\alpha_R : R \in \mathcal{A}_n^d, \alpha_R = \alpha_{R_1 \times R_2} \cdot \alpha_{R_3 \times \cdots \times R_d} \text{ with } \alpha_{R_1 \times R_2}, \alpha_{R_3 \times \cdots \times R_d} \in \{-1, 1\}\}.$$

Problem 1: Given any $n \geq 1$ and $d \geq 3$, does the signed small ball inequality hold for every collection of coefficients $\{\alpha_R\} \subset \mathcal{A}'_{\text{split}}$?

The new splitting property elevates the complexity of the problem. The paths which were taken in the proofs of Theorem 1.3 can not be followed here since, in this case, the corresponding random variables generated by the new splitting property,

$$\sum_{|R'|=2^{-r_1-r_2}} \alpha_{R'} h_{R'}, \quad r_1 \in \{0, \ldots, n\} \text{ and } r_2 \in \{0, \ldots, n-r_1\}$$

with $R' = R_1 \times R_2$, are not independent. This can be seen in the following way. First, observe that the signed small ball inequality is sharp when the coefficients are restricted to the set $\mathcal{A}'_{\text{split}}$. In other words, there exists a constant $C_d > 0$ and coefficients $\{\alpha_R\} \subset \mathcal{A}'_{\text{split}}$ such that

$$\left\| \sum_{|R|=2^{-n}} \alpha_{R_1 \times R_2} \alpha_{R_3 \times \cdots \times R_d} h_R \right\|_\infty \le C_d n^{d/2}. \tag{8.1}$$

This can be shown using the same argument as that given in Theorem 1.4. If the aforementioned random variables were independent, then, as in the proof of Theorem 1.3, we would get that

$$\|H_n\|_\infty \ge \sum_{r_1=0}^{n} \sum_{r_2=0}^{n-r_1} \|F_{r_1,r_2}\|_{L^1([0,1]^{d-2})}, \tag{8.2}$$

where $F_{r_1,r_2} = \sum_{|R''|=2^{-(n-r_1-r_2)}} \alpha_{R''} h_{R''}$ with $R'' = R_3 \times \cdots \times R_d$. Since Proposition 5.6 is also applicable to F_{r_1,r_2}, $\|F_{r_1,r_2}\|_1 \gtrsim \|F_{r_1,r_2}\|_2$, and this together with (8.2) shows that, for any collection of coefficients $\{\alpha_R\} \subset \mathcal{A}'_{\text{split}}$,

$$\|H_n\|_\infty \ge \sum_{r_1=0}^{n} \sum_{r_2=0}^{n-r_1} \|F_{r_1,r_2}\|_{L^2([0,1]^{d-2})} \gtrsim \sum_{r_1=0}^{n} \sum_{r_2=0}^{n-r_1} (n-r_1-r_2)^{(d-3)/2} \gtrsim n^{(d+1)/2}.$$

However, the above inequality contradicts the validity of (8.1).

As shown in Remark 6.3, the techniques employed in the proofs of Theorem 1.3 can be used to show that the d-dimensional small ball inequality holds for a special family of coefficients $\{\alpha_{R_1} \cdot \alpha_{R_2 \times \cdots \times R_d}\}$ with $\alpha_{R_1} \in \{\pm 1\}$ and $c_1 \le |\alpha_{R_2 \times \cdots \times R_d}| \le c_2$, for given positive constants c_1 and c_2. If we weaken the above condition so that we have only an upper bound or only a lower bound, one might wonder if (1.8) still holds. That is, let

$$\mathcal{B}_{\text{split}} = \{\alpha_R : R \in \mathcal{A}_n^d, \alpha_R = \alpha_{R_1} \cdot \alpha_{R_2 \times \cdots \times R_d} \text{ with } \alpha_{R_1} \in \{\pm 1\}\}.$$

Problem 2: Does the small ball inequality hold for all $\alpha_R \in \mathcal{B}_{\text{split}}$ which satisfy $0 < c \le |\alpha_{R_2 \times \cdots \times R_d}|$, or for all $\alpha_R \in \mathcal{B}_{\text{split}}$ which satisfy $|\alpha_{R_2 \times \cdots \times R_d}| \le c$?

Under this weaker restriction on the coefficients, it is not known whether the L^p-norms of such hyperbolic sums are comparable, as this cannot be deduced as a direct consequence of Remark 5.7.

It would also be very interesting to know whether the conjectured lower bound of signed hyperbolic sums in exponential Orlicz spaces described in (1.15) is optimal.

Problem 3: If $d \ge 2$ and $n > 1$, is there a positive constant $C(d,a)$ and coefficients $\{\alpha_R\} \subset \{-1, 1\}$ for which

$$\left\| \sum_{|R|=2^{-n}} \alpha_R h_R \right\|_{\exp(L^a)} \le C(d,a)\ n^{\frac{d-1}{2} + \frac{1}{2} - \frac{1}{a}},\ 2 \le a < \infty\ ? \tag{8.3}$$

We would be able to prove the above inequality if we knew that, for any collection of coefficients $\{\alpha_R\} \subset \{-1, 1\}$ and $d \ge 2$,

$$\left\| \sum_{|R|=2^{-n}} \alpha_R h_R \right\|_{\exp(L^2)} \le C_d\ n^{\frac{d-1}{2}}. \tag{8.4}$$

Indeed, first recall the useful interpolation inequality which says that, for any $0 < b < a < \infty$, we have that

$$\|H_n\|_{\exp(L^a)} \lesssim \|H_n\|_{\exp(L^b)}^{\frac{b}{a}} \|H_n\|_\infty^{1-\frac{b}{a}},$$

where $H_n = \sum_{|R|=2^{-n}} \alpha_R h_R$. The proof of the above interpolation inequality can be found in [9]. From Section 4.2, we know that there exists a collection of coefficients $\{\alpha_R\} \subset \{-1, 1\}$ such that $\|H_n\|_\infty \le C_d n^{d/2}$. For this collection of coefficients, using the interpolation inequality for $a \ge 2$ and $b = 2$ together with (8.4), we get

(8.3). However, (8.4) is not known in general. From [3] we know that, for any collection coefficients $\{\alpha_R\} \subset \{-1,1\}$, we have $\|H_n\|_{\exp(L^b)} \lesssim n^{\frac{d-1}{2}}$, where $b = \frac{2}{d-1}$, and combining this with the interpolation inequality only gives

$$\left\| \sum_{|R|=2^{-n}} \alpha_R h_R \right\|_{\exp(L^a)} \leq C(d,a)\, n^{d-1-\frac{1}{a}},\ 2 \leq a < \infty. \tag{8.5}$$

Note that for $d = 2$ inequalities (8.3) and (8.5) coincide, and this shows that the lower bound in (1.15) is optimal when $d = 2$.

Bibliography

[1] T. W. Anderson, *The integral of a symmetric unimodal function over a symmetric convex set and some probabilities inequalities*, Proc. Amer. Math. Soc., **6** (1955), no. 2, 170-176.

[2] J. Beck, *A two-dimensional van Aardenne-Ehrenfest theorem in irregularities of distribution*, Composition Math., **72** (1989), no. 3, 269-339.

[3] D. Bilyk, M. Lacey, *On the small ball inequality in three dimensions*, Duke Math. J., **143** (2008), no. 1, 81-115.

[4] D. Bilyk, M. Lacey and A. Vagharshakyan, *On the small ball inequality in all dimensions*, J. Funct. Anal., **254** (2008), no. 9, 2470-2505.

[5] D. Bilyk, M. Lacey and A. Vagharshakyan, *On the signed small ball inequality*, Online J. of Anal. Comb., **3** (2008), Art 6,7 pp.

[6] D. Bilyk, M. Lacey, I. Parissis and A. Vagharshakyan, *Exponential squared integrability of the discrepancy function in two dimensions*, Matematika, **55** (2009), no. 1-2, 1-27.

[7] D. Bilyk, M. Lacey, I. Parissis and A. Vagharshakyan, *A three-dimensional small ball inequality*, Dependence in probability, analysis and number theory, 73-87, Kendrick Press, Heber City, UT, (2010).

[8] D. Bilyk, *On Roth's orthogonal function method in discrepancy theory*, Unif. Distrib. Theory, **6** (2011), no. 1, 143-184.

[9] D. Bilyk, L. Markhasin, *BMO and exponential Orlicz space estimates of the discrepancy function in arbitrary dimension*, available at http://arxiv.org/abs/1411.5794.

[10] D. Bilyk, *Roth's orthogonal function method in discrepancy theory and some new connections*, a panorama of discrepancy theory, 71-158, Lecture Notes in Math., 2107, Springer-Verlag, Cham, (2014).

[11] D. Bilyk, N. Feldheim, *The two-dimensional small ball inequality and binary nets*, available at www.arxiv.org/pdf/1511.07326.pdf.

[12] D. L. Burkholder, *Martingale tranforms*, Ann. Math. Statist., **37** (1966), no. 6, 1494-1504.

[13] D. L. Burkholder, *Sharp inequalities for martingales an stochastic integrals*, Astérisque, **157-158** (1988), 75-94.

[14] W. W. L. Chen, *On irregularities of distribution*, Mathematika, **27** (1981), no. 2, 153-170.

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

[16] W. W. L. Chen, M. M. Skriganov, *Davenport's theorem in the theory of irregularities of point distribution*, J. Math. Sci. (N. Y.), **115** (2003), no. 1, 2076-2084.

[17] E. Csáki, *On small values of the square integral of a multiparameter Wiener process*, Statistics and probability (Visegrád, 1982), 19-26, Reidel, Dordrecht (1984).

[18] T. Dunker, T. Kühn, M. Lifshits, W. Linde, *Metric entropy of the integration operator and small ball probabilities for the Brownian sheet*, C. R. Acad. Sci. Paris Sér. I Math., **326** (1998), no. 3, 347-352.

[19] W. Feller *An Introduction to Probability Theory and its Applications.* Wiley, New York (1966)

[20] K. K. Frolov, *Upper bounds on the discrepancy in L_p, $2 \leq p < \infty$*, Dokl. Akad. Nauk. USSR, **252** (1980), 805-807.

[21] L. Grafakos, *Classical Fourier Analysis*, Third edition,Graduate Springer, New York 2014. ISBN: 978-1-4939-1193-6; 978-1-4939-1194-3.

[22] U. Haagerup, *The best constants in Khintchine's inequality*, Studia Math., **70** (1981), no. 3, 231-283.

[23] J. H. Halton, *On the efficiency of certain quasirandym sequences of points in evaluating multidimensional integrals*, Num. Math., **2** (1960), no. 1, 84-90.

[24] J. M. Hammersley *Monte Carlo methods for solving multivariable problems*, Ann. New York Acad. Sci., **86** (1960), 844-874.

[25] G. Halász, *On Roth's method in the theory of irregularities of point distributions*, Recent progress in analytic number theory, Vol. 2, pp. 79-94. Academic Press, London (1981).

[26] A. Hinrichs, L. Markhasin, *On lower bounds for the L^2 discrepancy*, Journal of Complexity, **27** (2011), no. 2, 127-132.

[27] E. Hlawka, *Funktionen von beschränkter Variation in der Theorie der Gleichverteilung*, Ann. Mat. Pura Appl, **54** (1961), no. 1, 325-333.

[28] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc., **58** (1963), no. 301, 13-30.

[29] D. Karslidis, *On the signed small ball inequality with restricted coefficients*, Indiana Math. J., to appear, available at http://www.iumj.indiana.edu/IUMJ/forthcoming.php.

[30] D. Karslidis, *Orlicz spaces bounds for special classes of hyperbolic sums*, submitted.

[31] J. F. Koksma, *Een algemeene stelling uit de theorie der gelijkmatige verdeeling modulo 1*, Mathematica B, **11** (1943), 7-11.

[32] A. N. Kolmogorov, V. M. Tikhomirov, *ε-entropy and ε-capacity of sets in functional spaces*, Uspekhi Matem. Nauk, **14** (1959), no. 2, 3-86.

[33] J. Kuelbs, W. V. Li, *Metric entropy and the small ball problem for Gaussian measures*, J. Funct. Anal., **116** (1993), no. 1, 133-157.

[34] M. Lacey, *Small Ball and Discrepancy Inequalities* (2008), available at www.arxiv.org/pdf/math/0609816v1.pdf

[35] P. Levy, *Problemes concrets d'analyze fonctionelle*, Gautheir-Villars (1951)

[36] G. G. Lorentz, *Metric entropy and approximation*, Bull. Amer. Math. Soc., **72** (1966), 903-937.

[37] M. M. Rao, Z. D. Ren, *Theory of Orlicz Spaces*, Pure and Applied Mathematics, Inc., Marcel Dekker, New York-Basel-Hong Kong 1991. ISBN: 0-8247-8478-2

[38] S. I. Resnick, *A Probability Path*, Birkhäuser Boston, Inc., Boston, MA, 1999. ISBN: 0-8176-4055-X.

[39] K. F. Roth, *On irregularities of distribution*, Mathimatika, **1** (1954), no. 2, 73-79.

[40] K. F. Roth, *On irregularities of distribution III*, Acta Arith., **35** (1979), no. 4, 373-384.

[41] K. F. Roth, *On irregularities of distribution IV*, Acta Arith., **37** (1980), no. 1, 67-75.

[42] W. M. Schmidt, *Irregularities of distribution VI.* Compositio Math., **24** (1972), no. 1, 63-74.

[43] W. M. Schmidt, *Irregularities of distribution X*, Number theory and algebra, 311-329, Academic Press, New York (1977).

[44] S. Sidon, *Verallgemeinerung Emus Satzes über die absolute Konvergenz von Fourierreihen mit Lücken*, Math. Ann., **97** (1927), 675-676.

[45] S. Sidon, *Ein Satzüber trigonometrische Polyfonem mit Lgücken und seine Anwendung in der Theorie der Fourier Reihen*, J. Reine Angew. Math., **163** (1930), 251-252.

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

[47] M. Talagrand, *The small ball problem for the Brownian sheet*, Ann. Probab., **22** (1994), no. 3, 1331-1354.

[48] C. Thiele, *Lectures on Real Analysis*, Bonn 2010/UCLA 2012, available at math.ucla.edu/ thiele/lecturenotes/bonnlectures/lecture4.pdf

[49] V. N. Temlyakov, *Some inequalities for multivariate Haar polynomials*, East Journal on Approximations, **1** (1995), no. 1, 61-72.

[50] V. N. Temlyakov, *An inequality for trigonometric polynomials and its applications for estimating the entropy numbers*, Journal of Complexity, **11** (1995), no. 2, 293-307.

[51] A. Vagharshakyan, *Lower bounds for L^1 Discrepancy*, Mathematica, **59** (2013), no. 2, 365-379.

[52] G. Wang, *Sharp square-function inequalities for conditionally symmetric martingales* . Trans. Amer. Math. Soc. **328** (1991), no. 1, 393-419.

[53] R. M. G. Young, *On the best possible constants in the Khintchine inequality*, J. London Math. Soc., (2) **14** (1976), no. 3, 496-504.