A three-dimensional signed small ball inequality

Dmitriy Bilyk, Michael T. Lacey, Ioannis Parissis, and Armen Vagharshakyan

Abstract Let R denote dyadic rectangles in the unit cube $[0,1]^3$ in three dimensions. Let h_R be the L^∞-normalized Haar function whose support is R. We show that for all integers $n \geq 1$ and choices of coefficients $a_R \in \{\pm 1\}$, we have

$$\Big\| \sum_{\substack{|R|=2^{-n} \\ |R_1| \geq 2^{-n/2}}} a_R \, h_R \Big\|_{L^\infty} \gtrsim n^{9/8}.$$

The trivial L^2 lower bound is n, and the sharp lower bound would be $n^{3/2}$. This is the best exponent known to the authors. This inequality is motivated by new results on the star-discrepancy function in all dimensions $d \geq 3$.

1 Introduction

We are motivated by the classical question of irregularities of distribution [2] and recent results which give new lower bounds on the star-discrepancy in all dimensions $d \geq 3$ [4, 5]. We recall these results.

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

Michael T. Lacey
School of Mathematics, Georgia Institute of Technology,
Atlanta GA 30332, USA, e-mail: lacey@math.gatech.edu

Ioannis Parissis
Institutionen för Matematik, Kungliga Tekniska Högskolan,
SE 100 44, Stockholm, Sweden, e-mail: ioannis.parissis@gmail.com

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

Given integer N, and selection $\mathcal{P}$ of N points in the unit cube $[0,1]^d$, we define a *Discrepancy Function* associated to $\mathcal{P}$ as follows. At any point $x \in [0,1]^d$, set

$$D_N(x) = \sharp(\mathcal{P} \cap [0,x)) - N|[0,x)|.$$

Here, by $[0,x)$ we mean the d-dimensional rectangle with left-hand corner at the origin, and right-hand corner at $x \in [0,1]^d$. Thus, if we write $x = (x_1, \dots, x_d)$ we then have

$$[0,x) = \prod_{j=1}^{d} [0, x_j).$$

At point x we are taking the difference between the actual and the expected number of points in the rectangle. Traditionally, the dependence of D_N on the selection of points $\mathcal{P}$ is only indicated through the number of points in the collection $\mathcal{P}$. We mention only the main points of the subject here, and leave the (interesting) history of the subject to references such as [2].

The result of Klaus Roth [7] gives a definitive average case lower bound on the discrepancy function.

K. Roth's Theorem *For any dimension $d \geq 2$, we have the following estimate*

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

The same lower bound holds in all L^p, $1 < p < \infty$, as observed by Schmidt [8]. But, the L^∞ infinity estimate is much harder. In dimension $d = 2$ the definitive result was obtained by Schmidt again [9].

Schmidt's Theorem *We have the estimates below, valid for all collections $\mathcal{P} \subset [0,1]^2$:*

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

The L^∞ estimates are referred to as star-discrepancy bounds. Extending and greatly simplifying an intricate estimate of J. Beck [1], some of these authors have obtained a partial extension of Schmidt's result to all dimensions $d \geq 3$.

Theorem 1.3 ([4, 5]**)** *For dimensions $d \geq 3$ there is an $\eta = \eta(d) > 0$ for which we have the inequality*

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

That is, there is an η improvement in the Roth exponent.

As explained in these references, the analysis of the star-discrepancy function is closely related to other questions in probability theory, approximation theory, and harmonic analysis. We turn to one of these, the simplest to state question, which is central to all of these issues. We begin with the definition of the Haar functions.

In one dimension, the dyadic intervals of the real line $\mathbb{R}$ are given by

$$\mathcal{D} = \left\{ \left[j2^k, (j+1)2^k \right) \,:\, j,k \in \mathbb{Z} \right\}.$$

Any interval I is a union of its left and right halves, denoted by $I_{\text{left/right}}$, which are also dyadic. The *Haar function h_I associated to I*, or simply *Haar function* is

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

Note that for dyadic intervals $J \subsetneq I$, the Haar function h_J is completely supported on a set where h_I is constant. This basic property leads to far-reaching implications that we will exploit in these notes.

In higher dimensions $d \geq 2$, we take the dyadic rectangles to be the tensor product of dyadic intervals in dimension d:

$$\mathcal{D}^d = \left\{ R = R_1 \times \cdots \times R_d \,:\, R_1, \ldots, R_d \in \mathcal{D} \right\}.$$

The *Haar function* associated to $R \in \mathcal{D}_d$ is likewise defined as

$$h_R(x_1, \ldots, x_d) = \prod_{j=1}^{d} h_{R_j}(x_j), \qquad R = R_1 \times \cdots \times R_d. \tag{1.4}$$

While making these definitions on all of $\mathbb{R}^d$, we are mainly interested in local questions, thus rectangles $R \subset [0,1]^d$ are always dyadic rectangles $R \in \mathcal{D}^d$. Namely, we are mainly interested in the following conjectural *reverse triangle inequality* for sums of Haar functions on L^∞:

Conjecture: The Small Ball Inequality *For dimensions $d \geq 3$, there is a constant C_d so that for all integers $n \geq 1$, and constants $\{a_R \,:\, |R| = 2^{-n},\ R \subset [0,1]^d\}$, we have*

$$n^{(d-2)/2} \Bigg\| \sum_{\substack{|R| \geq 2^{-n} \\ R \subset [0,1]^d}} a_R \cdot h_R \Bigg\|_\infty \geq C_d 2^{-n} \sum_{\substack{|R| = 2^{-n} \\ R \subset [0,1]^d}} |a_R|. \tag{1.5}$$

We are stating this inequality in its strongest possible form. On the left, the sum goes over all rectangles with volume *at least* 2^{-n}, while on the right, we only sum over rectangles with volume *equal to* 2^{-n}. Given the primitive state of our knowledge of this conjecture, we will not insist on this distinction below.

For the case of $d = 2$, (1.5) holds, and is a theorem of Talagrand [10]. (Also see [6, 8, 11].)

The special case of the Small Ball Inequality when all the coefficients a_R are equal to either -1 or $+1$ we refer to as the 'Signed Small Ball Inequality.' Before stating this conjecture, let us note that we have the following (trivial) variant of Roth's Theorem in the Signed case:

$$\Bigg\| \sum_{\substack{|R| = 2^{-n} \\ R \subset [0,1]^d}} a_R \cdot h_R \Bigg\|_\infty \gtrsim n^{(d-1)/2}, \qquad a_R \in \{\pm 1\}.$$

The reader can verify this by noting that the left-hand side can be written as about n^{d-1} orthogonal functions, by partitioning the unit cube into homothetic copies of dyadic rectangles of a fixed volume. The Signed Small Ball Inequality asserts a 'square root of n' gain over this average case estimate.

Conjecture: The Signed Small Ball Inequality *For coefficients $a_R \in \{\pm 1\}$,*

$$\Big\| \sum_{\substack{|R|=2^{-n} \\ R\subset[0,1]^d}} a_R \cdot h_R \Big\|_\infty \geq C'_d n^{d/2}. \tag{1.6}$$

Here, C'_d is a constant that only depends upon dimension.

We should emphasize that random selection of the coefficients shows that the power on n on the right is sharp. Unfortunately, random coefficients are very far from the 'hard instances' of the inequality, so do not indicate a proof of the conjecture.

The Signed Small Ball Conjecture should be easier, but even this special case eludes us. To illustrate the difficulty in this question, note that in dimension $d=2$, each point x in the unit square is in $n+1$ distinct dyadic rectangles of volume 2^{-n}. Thus, it suffices to find a *single* point where all the Haar functions have the same sign. This we will do explicitly in § 2 below.

Passing to three dimensions reveals a much harder problem. Each point x in the unit cube is in about n^2 rectangles of volume 2^{-n}, but in general we can only achieve a $n^{3/2}$ supremum norm. Thus, the task is to find a single point x where the number of pluses is more than the number of minuses by $n^{3/2}$. In percentage terms this represents only a $n^{-1/2}$-percent imbalance over equal distribution of signs.

The main theorem of this note is Theorem 4.1 below, which gives the best exponent we are aware of in the Signed Small Ball Inequality. The method of proof is also the simplest we are aware of. (In particular, it gives a better result than the more complicated argument in [3].) Perhaps this argument can inspire further progress on this intriguing and challenging question.

The authors thank the anonymous referee whose attention to detail has helped greater clarity in our arguments.

Dedication to Walter Philipp One of us was a PhD student of Walter Philipp, the last of seven students. Walter was very fond of the subject of this note, though the insights he would have into the recent developments are lost to us. As a scientist, he held himself to high standards in all his areas of study. As a friend, he was faithful, loyal, and took great pleasure in renewing contacts and friendship. He is very much missed.

2 The two-dimensional case

This next definition is due to Schmidt, refining a definition of Roth. Let $\vec r \in \mathbb{N}^d$ be a partition of n, thus $\vec r = (r_1, \dots, r_d)$, where the r_j are non-negative integers and

$|\vec{r}| := \sum_{t=1}^{d} r_t = n$. Denote all such vectors at $\mathbb{H}_n$. ('$\mathbb{H}$' for 'hyperbolic.') For vector $\vec{r}$, let $\mathcal{R}_{\vec{r}}$ be all dyadic rectangles R such that for each coordinate $1 \le t \le d$, $|R_t| = 2^{-r_t}$.

Definition 2.1 We call a function f an r-*function with parameter* $\vec{r}$ if

$$f = \sum_{R \in \mathcal{R}_{\vec{r}}} \varepsilon_R \, h_R, \qquad \varepsilon_R \in \{\pm 1\}. \tag{2.2}$$

We will use $f_{\vec{r}}$ to denote a generic r-function. A fact used without further comment is that $f_{\vec{r}}^2 \equiv 1$.

Note that in the Signed Small Ball Inequality, one is seeking lower bounds on sums $\sum_{|\vec{r}|=n} f_{\vec{r}}$.

There is a trivial proof of the two-dimensional Small Ball Inequality.

Proposition 2.3 *The random variables* $f_{(j,n-j)}$, $0 \le j \le n$ *are independent.*

Proof. The sigma-field generated by the functions $\{f_{(k,n-k)} : 0 \le k < j\}$ consists of dyadic rectangles $S = S_1 \times S_2$ with $|S_1| = 2^{-j}$ and $|S_2| = 2^{-n}$. On each line segment $S_1 \times \{x_2\}$, $f_{(j,n-j)}$ takes the values ± 1 in equal measure, so the proof is finished. □

We then have

Proposition 2.4 *In the case of two dimensions,*

$$\mathbb{P}\Big(\sum_{k=0}^{n} f_{(k,n-k)} = n+1 \Big) = 2^{-n-1}.$$

Proof. Note that

$$\mathbb{P}\Big(\sum_{k=0}^{n} f_{(k,n-k)} = n+1 \Big) = \mathbb{P}\big(f_{(k,n-k)} = 1 \ \forall 0 \le k \le n \big) = 2^{-n-1}.$$

□

It is our goal to give a caricature of this argument in three dimensions. See § 5 for a discussion.

3 Elementary lemmas

We recall some elementary lemmas that we will need in our three-dimensional proof.

Paley–Zygmund Inequality *Suppose that* Z *is a positive random variable with* $\mathbb{E}Z = \mu_1$, $\mathbb{E}Z^2 = \mu_2^2$. *Then,*

$$\mathbb{P}(Z \ge \mu_1/2) \ge \tfrac{1}{4} \frac{\mu_2^2}{\mu_1^2}. \tag{3.1}$$

Proof.

$$\begin{aligned}\mu_1 = \mathbb{E}Z &= \mathbb{E}Z\mathbf{1}_{Z<\mu_1/2} + \mathbb{E}Z\mathbf{1}_{Z\geq\mu_1/2}\\ &\leq \mu_1/2 + \mu_2\mathbb{P}(Z\geq\mu_1/2)^{1/2}.\end{aligned}$$

Now solve for $\mathbb{P}(Z\geq\mu_1/2)$. □

Second Paley–Zygmund Inequality *For all* $\rho_1>1$ *there is a* $\rho_2>0$ *so that for all random variables Z which satisfy*

$$\mathbb{E}Z=0, \qquad \|Z\|_2\leq\|Z\|_4\leq\rho_1\|Z\|_2 \tag{3.2}$$

we have the inequality $\mathbb{P}(Z>\rho_2\|Z\|_2)>\rho_2$.

Proof. Let $Z_+ := Z\mathbf{1}_{Z>0}$ and $Z_- := -Z\mathbf{1}_{Z<0}$, so that $Z=Z_+-Z_-$. Note that $\mathbb{E}Z=0$ forces $\mathbb{E}Z_+=\mathbb{E}Z_-$. And,

$$\begin{aligned}\sigma_2^2 &:= \mathbb{E}Z^2 = \mathbb{E}Z_+^2+\mathbb{E}Z_-^2,\\ \sigma_4^4 &:= \mathbb{E}Z^4 = \mathbb{E}Z_+^4+\mathbb{E}Z_-^4.\end{aligned}$$

Suppose that the conclusion is not true. Namely $\mathbb{P}(Z>\rho_2\sigma_2)<\rho_2$ for a very small ρ_2. It follows that

$$\begin{aligned}\mathbb{E}Z_+ &\leq \mathbb{E}Z_+\mathbf{1}_{Z_+\leq\rho_2\sigma_2}+\mathbb{E}Z_+\mathbf{1}_{Z_+>\rho_2\sigma_2}\\ &\leq \rho_2\sigma_2+\mathbb{P}(Z>\rho_2\sigma_2)^{1/2}\sigma_2\leq 2\rho_2^{1/2}\sigma_2,\end{aligned}$$

for $\rho_2<1$. Hence $\mathbb{E}Z_-=\mathbb{E}Z_+\leq 2\rho_2^{1/2}\sigma_2$. It is this condition that we will contradict below.

We also have

$$\begin{aligned}\mathbb{E}Z_+^2 &\leq \mathbb{E}Z_+^2\mathbf{1}_{\{Z_+\leq\rho_2\sigma_2\}}+\mathbb{E}Z_+^2\mathbf{1}_{\{Z_+>\rho_2\sigma_2\}}\\ &\leq \rho_2^2\sigma_2^2+\rho_2^{1/2}\sigma_4^2\\ &\leq 2\rho_2^{1/2}\rho_1^2\sigma_2^2.\end{aligned}$$

So for $\rho_2<(4\rho_1)^{-4}$, we have $\mathbb{E}Z_+^2\leq\frac{1}{2}\sigma_2^2$.

It follows that we have $\mathbb{E}Z_-^2\geq\frac{1}{2}\sigma_2^2$, and $\mathbb{E}Z_-^4\leq\rho_1\sigma_2^4$. So by (3.1), we have

$$\mathbb{P}(Z_->\rho_3\sigma_2)>\rho_3$$

where ρ_3 is only a function of ρ_1. But this contradicts $\mathbb{E}Z_-\leq 2\rho_2^{1/2}\sigma_2$, for small ρ_2, so finishes our proof. □

The Paley–Zygmund inequalities require a higher moment, and in application we find it convenient to use the Littlewood–Paley inequalities to control this higher moment. Let $\mathcal{F}_0,\mathcal{F}_1,\dots,\mathcal{F}_T$ a sequence of increasing sigma-fields generated by dyadic intervals, and let d_t, $1\leq t\leq T$ be a martingale difference sequence, namely $\mathbb{E}(d_t:\mathcal{F}_{t-1})=0$ for all $t=1,2,\dots,T$. Set $f=\sum_{t=1}^T d_t$. The martingale square function of f is $\mathrm{S}(f)^2:=\sum_{t=1}^T d_t^2$. The instance of the Littlewood–Paley inequalities we need are:

Lemma 3.3 *With the notation above, suppose that we have in addition that the distribution of d_t is conditionally symmetric given $\mathcal{F}_{t-1}$. By this we mean that on each atom A of $\mathcal{F}_{t-1}$, the the distribution of $d_t\mathbf{1}_A$ is equal to that of $-d_t\mathbf{1}_A$. Then, we have*

$$\|f\|_4 \simeq \|\mathrm{S}(f)\|_4. \tag{3.4}$$

Proof. The case of the Littlewood–Paley for even integers can be proved by expansion of the integral, an argument that goes back many decades, and our assumption of being conditionally symmetric is added just to simplify this proof. Thus,

$$\|f\|_4^4 = \sum_{1\le t_1,t_2,t_3,t_4\le T} \mathbb{E}\prod_{u=1}^4 d_{t_u}.$$

We claim that unless the integers $1 \le t_1,t_2,t_3,t_4 \le T$ occur in pairs of equal integers, the expectation on the right above is zero. This claim shows that

$$\|f\|_4^4 = \sum_{1\le t_1,t_2\le T} \mathbb{E} d_{t_1}^2 \cdot d_{t_2}^2.$$

It is easy to see that this proves the lemma, namely we would have

$$\|\mathrm{S}(f)\|_4^4 \le \|f\|_4^4 \le \tfrac{4!}{2}\|\mathrm{S}(f)\|_4^4.$$

Let us suppose $t_1 \le t_2 \le t_3 \le t_4$. If we have t_3 strictly less than t_4, then

$$\mathbb{E}\prod_{u=1}^4 d_{t_u} = \mathbb{E}\prod_{u=1}^3 d_{t_u}\cdot \mathbb{E}(d_{t_4} \,:\, \mathcal{F}_{t_3}) = 0.$$

If we have $t_1 < t_2 = t_3 = t_4$, then by conditional symmetry, $\mathbb{E}(d_{t_2}^3 \,:\, \mathcal{F}_{t_1}) = 0$, and so we have

$$\mathbb{E}\prod_{u=1}^4 d_{t_u} = \mathbb{E} d_{t_1}\cdot \mathbb{E}(d_{t_2}^3 \,:\, \mathcal{F}_{t_1}) = 0.$$

If we have $t_1<t_2<t_3=t_4$, the conditional symmetry again implies $\mathbb{E}(d_{t_2}\cdot d_{t_3}^2 : \mathcal{F}_{t_1}) = 0$, so that

$$\mathbb{E}\prod_{u=1}^4 d_{t_u} = \mathbb{E} d_{t_1}\cdot \mathbb{E}(d_{t_2}\cdot d_{t_3}^2 \,:\, \mathcal{F}_{t_1}) = 0.$$

Thus, the claim is proved. □

We finish this section with an elementary, slightly technical, lemma.

Lemma 3.5 *Let $\mathcal{F}_0, \mathcal{F}_1, \dots, \mathcal{F}_q$ a sequence of increasing sigma-fields. Let $A_1,\dots,A_q$ be events, with $A_t \in \mathcal{F}_t$. Assume that for some $0<\gamma<1$,*

$$\mathbb{E}\big(\mathbf{1}_{A_t} \,:\, \mathcal{F}_{t-1}\big) \ge \gamma, \qquad 1\le t\le q. \tag{3.6}$$

We then have that

$$\mathbb{P}\Big(\bigcap_{t=1}^{q} A_t\Big) \geq \gamma^q. \tag{3.7}$$

More generally, assume that

$$\mathbb{P}\Big(\bigcup_{t=1}^{q}\big\{\mathbb{E}\big(\mathbf{1}_{A_t} : \mathcal{F}_{t-1}\big) \leq \gamma\big\}\Big) \leq \tfrac{1}{2}\cdot\gamma^q. \tag{3.8}$$

Then,

$$\mathbb{P}\Big(\bigcap_{t=1}^{q} A_t\Big) \geq \tfrac{1}{2}\cdot\gamma^q. \tag{3.9}$$

Proof. To prove (3.7), note that by assumption (3.6), and backwards induction we have

$$\begin{aligned}
\mathbb{P}\Big(\bigcap_{t=1}^{q} A_t\Big) &= \mathbb{E}\prod_{t=1}^{q}\mathbf{1}_{A_t} \\
&= \mathbb{E}\prod_{t=1}^{q-1}\mathbf{1}_{A_t} \times \mathbb{E}\big(\mathbf{1}_{A_q} : \mathcal{F}_{q-1}\big) \\
&\geq \gamma\mathbb{E}\prod_{t=1}^{q-1}\mathbf{1}_{A_t} \\
&\;\;\vdots \\
&\geq \gamma^q.
\end{aligned}$$

To prove (3.9), let us consider an alternate sequence of events. Define

$$\beta_t := \big\{\mathbb{E}(A_t : \mathcal{F}_{t-1}) \leq \gamma\big\}.$$

These are the ‘bad’ events. Now define $\widetilde{A}_t := A_t \cup \beta_t$. By construction, the sets $\widetilde{A}_t$ satisfy (3.6). Hence, we have by (3.7),

$$\mathbb{P}\Big(\bigcap_{t=1}^{q} \widetilde{A}_t\Big) \geq \gamma^q.$$

But, now note that by (3.8),

$$\begin{aligned}
\mathbb{P}\Big(\bigcap_{t=1}^{q} A_t\Big) &= \mathbb{P}\Big(\bigcap_{t=1}^{q} \widetilde{A}_t\Big) - \mathbb{P}\Big(\bigcup_{t=1}^{q} \beta_t\Big) \\
&\geq \gamma^q - \tfrac{1}{2}\cdot\gamma^q \geq \tfrac{1}{2}\cdot\gamma^q.
\end{aligned}$$

□

4 Conditional expectation approach in three dimensions

This is the main result of this note.

Theorem 4.1 *For $|a_R| = 1$ for all R, we have the estimate*

$$\Big\| \sum_{\substack{|R|=2^{-n} \\ |R_1|\geq 2^{-n/2}}} a_R h_R \Big\|_{L^\infty} \gtrsim n^{9/8}.$$

We restrict the sum to those dyadic rectangles whose first side has the lower bound $|R_1| \geq 2^{-n/2}$.

Heuristics for our proof are given in the next section. The restriction on the first side lengths of the rectangles is natural from the point of view of our proof, in which the first coordinate plays a distinguished role. Namely, if we hold the first side length fixed, we want the corresponding sum over R to be suitably generic. Let $1 \ll q \ll n$ be integers. The integer q will be taken to be $q \simeq n^{1/4}$. Our 'gain over average case' estimate will be $\sqrt{q} \simeq n^{1/8}$. While this is a long way from $n^{1/2}$, it is much better than the explicit gain of $1/24$ in [3].

We begin the proof. Let $\mathcal{F}_t$ be the sigma field generated by dyadic intervals in [0,1] with $|I| = 2^{-\lfloor tn/q \rfloor}$, for $1 \leq t \leq \frac{1}{2}q$. Let $\mathbb{I}_t := \{\vec{r} : (t-1)n/q \leq r_1 < tn/q\}$. Note that the size $\#\mathbb{I}_t \approx n^2/q$. Let $f_{\vec{r}}$ be the r-functions specified by the choice of signs in Theorem 4.1. Here is a basic observation.

Proposition 4.2 *Let $I \in \mathcal{F}_t$. The distribution of $\{f_{\vec{r}} : \vec{r} \in \mathbb{I}_t\}$ restricted to the set $I \times [0,1]^2$ with normalized Lebesgue measure is that of*

$$\{f_{\vec{s}} : |\vec{s}| = n - \lfloor tn/q \rfloor, 0 \leq s_1 < n/q\},$$

where the $f_{\vec{s}}$ are some r-functions. The exact specification of this collection depends upon the atom in $\mathcal{F}_t$.

Proof. An atom I of $\mathcal{F}_t$ are dyadic intervals of length $2^{-\lfloor tn/q \rfloor}$. For $\vec{r} \in \mathbb{I}_t$, $f_{\vec{r}}$ restricted to $I \times [0,1]^2$, with normalized measure, is an r-function with index

$$(r_1 - \lfloor tn/q \rfloor, r_2, r_3).$$

The statement holds jointly in $\vec{r} \in \mathbb{I}_t$ so finishes the proof. □

Define sum of 'blocks' of $f_{\vec{r}}$ as

$$B_t := \sum_{\vec{r} \in \mathbb{I}_t} f_{\vec{r}}, \tag{4.3}$$

$$\sqcap_t := \sum_{\substack{\vec{r} \neq \vec{s} \in \mathbb{I}_t \\ r_1 = s_1}} f_{\vec{r}} \cdot f_{\vec{s}}. \tag{4.4}$$

The sums $\sqcap_t$ play a distinguished role in our analysis, as revealed by the basic computation of a square function in (4.9) and the fundamental Lemma 4.10. Let us set $\sigma_t^2 = \|B_t\|_2^2 \simeq n^2/q$, for $0 \le t \le q/2$.

We want to show that for q as big as $cn^{1/4}$, we have

$$\mathbb{P}\Big(\sum_{t=1}^{q/2} B_t \gtrsim n\sqrt{q}\Big) > 0. \tag{4.5}$$

In fact, we will show

$$\mathbb{P}\Big(\bigcap_{t=1}^{q/2} \big\{B_t \gtrsim n/\sqrt{q}\big\}\Big) > 0,$$

from which (4.5) follows immediately.

Note that the event $\big\{B_t \gtrsim n/\sqrt{q}\big\}$ simply requires that B_t be of typical size, and positive, that is this event will have a large probability. Clearly, we should try to show that these events are in some sense independent, in which case the lower bound in (4.5) will be of the form e^{-Cq}, for some $C > 0$. Exact independence, as we had in the two-dimensional case, is too much to hope for. Instead, we will aim for some conditional independence, as expressed in Lemma 3.5.

There is a crucial relationship between B_t and $\sqcap_t$, which is expressed through the martingale square function of B_t, computed in the first coordinate. Namely, define

$$\mathrm{S}(B_t)^2 := \sum_{j \in J_t} \Big| \sum_{\vec{r}\,:\, r_1 = j} f_{\vec{r}} \Big|^2 \tag{4.6}$$

where $J_t = \{s \in \mathbb{N} : (t-1)n/q \le s < tn/q\}$.

Proposition 4.7 *We have*

$$\mathrm{S}(B_t)^2 = \sigma_t^2 + \sqcap_t, \tag{4.8}$$

$$\mathrm{S}(B_t : \mathcal{F}_t) = \sigma_t^2 + \mathbb{E}(\sqcap_t : \mathcal{F}_t). \tag{4.9}$$

By construction, we have $\sharp\,\mathbb{I}_t \simeq n^2/q$, for $0 \le t < \frac{1}{2}q$.

Proof. In (4.6), one expands the square on the right-hand side. Notice that this shows that

$$\mathrm{S}(B_t)^2 = \sum_{\substack{|\vec{r}|=|\vec{s}|=n \\ r_1 = s_1 \in J_t}} f_{\vec{r}} \cdot f_{\vec{s}}.$$

We can have $\vec{r} = \vec{s}$ for $\sharp\,\mathbb{I}_t$ choices of $\vec{r}$. Otherwise, we have a term that contributes to $\sqcap_t$. The conditional expectation conclusion follows from (4.8) □

The next fact is the critical observation in [3–5] concerning coincidences, assures us that typically on the right in (4.8) that the first term $\sigma_t^2 \simeq n^2/q$ is much larger than the second $\sqcap_t$. See [5, 4.1, and the discussion afterwards].

Lemma 4.10 *We have the uniform estimate*

$$\|\sqcap_t\|_{\exp(L^{2/3})} \lesssim n^{3/2}/\sqrt{q}.$$

Here, we are using standard notation for an exponential Orlicz space.

Remark 4.11 A variant of Lemma 4.10 holds in higher dimensions, which permits an extension of Theorem 4.1 to higher dimensions. We return to this in § 5.

Let us quantify the relationship between these two observations and our task of proving (4.5).

Proposition 4.12 *There is a universal constant* $\tau > 0$ *so that defining the event*

$$\Gamma_t := \left\{ \mathbb{E}(\sqcap_t^2 \,:\, \mathcal{F}_t)^{1/2} < \tau n^2/q \right\} \tag{4.13}$$

we have the estimate

$$\mathbb{P}(B_t > \tau \cdot n/\sqrt{q} \,:\, \Gamma_t) > \tau \mathbf{1}_{\Gamma_t}. \tag{4.14}$$

The point of this estimate is that the events Γ_t will be overwhelmingly likely for $q \ll n$.

Proof. This is a consequence of the Paley–Zygmund Inequalities, Proposition 4.2, Littlewood–Paley inequalities, and (4.9). Namely, by Proposition 4.2, we have $\mathbb{E}(B_t \,:\, \mathcal{F}_t) = 0$, and the conditional distribution of B_t given $\mathcal{F}_t$ is symmetric. By (4.9), we have

$$\mathbb{E}(B_t^2 \,:\, \mathcal{F}_t) = \mathrm{S}(B_t \,:\, \mathcal{F}_t) = \sigma_t^2 + \mathbb{E}(\sqcap_t \,:\, \mathcal{F}_t).$$

We apply the Littlewood–Paley inequalities (3.4) to see that

$$\mathbb{E}(B_t^4 \,:\, \mathcal{F}_t) \lesssim \mathbb{E}(S(B_t \,:\, \mathcal{F}_t)^2 \,:\, \mathcal{F}_t) = \sigma_t^4 + 2\sigma_t^2 \mathbb{E}(\sqcap_t \,:\, \mathcal{F}_t) + \mathbb{E}(\sqcap_t^2 \,:\, \mathcal{F}_t).$$

The event Γ_t gives an upper bound on the terms involving $\sqcap_t$ above. This permits us to estimate, as $\Gamma_t \in \mathcal{F}_t$,

$$\left| \mathbb{E}(B_t^2 \,:\, \Gamma_t)^{1/2} - \sigma_t \mathbf{1}_{\Gamma_t} \right| \le \tau n/\sqrt{q},$$

but $\sigma_t \simeq n/\sqrt{q}$, so we have $\mathbb{E}(B_t^2 \,:\, \Gamma_t)^{1/2} \simeq n/\sqrt{q}$. Similarly,

$$\begin{aligned} \mathbb{E}(B_t^4 \,:\, \Gamma_t)^{1/4} &\lesssim \sigma_t + \sigma_t^{1/2} |\mathbb{E}(\sqcap_t \,:\, \mathcal{F}_t)|^{1/4} + \mathbb{E}(\sqcap_t^2 \,:\, \mathcal{F}_t)^{1/4} \\ &\lesssim (1+\tau)\sigma_t. \end{aligned}$$

Hence, we can apply the Paley–Zygmund inequality (3.2) to conclude the proposition. □

By way of explaining the next steps, let us observe the following. If we have

$$\mathbb{E}(\mathbf{1}_{\Gamma_t} : \mathcal{F}_t) \ge \tau \qquad \text{a.s. } (x_1), \qquad 1 \le t \le q/2, \tag{4.15}$$

then (4.14) holds, namely $\mathbb{P}(B_t(\cdot, x_2, x_3) > \tau \cdot n/\sqrt{q} : \Gamma_t) > \tau$ almost surely. Applying Lemma 3.5, and in particular (3.7), we then have

$$\mathbb{P}_{x_1}\Big(\bigcap_{t=1}^{q/2}\{B_t(\cdot, x_2, x_3) > \tau n/\sqrt{q}\}\Big) \ge \tau^{q/2}.$$

Of course there is no reason that such a pair (x_2, x_3) exits. Still, the second half of Lemma 3.5 will apply if we can demonstrate that we can choose x_2, x_3 so that (4.15) holds except on a set, in the x_1 variable, of sufficiently small probability.

Keeping (3.8) in mind, let us identify an exceptional set. Use the sets Γ_t as given in (4.13) to define

$$E := \left\{(x_2, x_3) : \mathbb{P}_{x_1}\left[\bigcup_{t=1}^{q/2}\Gamma_t^c\right] > \exp\big(-c_1(n/q)^{1/3}\big)\right\}. \tag{4.16}$$

Here, $c_1 > 0$ will be a sufficiently small constant, independent of n. Let us give an upper bound on this set.

$$\mathbb{P}_{x_2,x_3}(E) \le \exp\big(c_1(n/q)^{1/3}\big) \cdot \mathbb{P}_{x_1,x_2,x_3}\Big(\bigcup_{t=1}^{q/2}\Gamma_t^c\Big) \tag{4.17}$$

$$\le \exp\big(c_1(n/q)^{1/3}\big)\sum_{t=1}^{q/2}\mathbb{P}_{x_1,x_2,x_3}(\Gamma_t^c) \tag{4.18}$$

$$\lesssim q\exp\big(c_1(n/q)^{1/3}\big)\cdot\exp\bigg(-\Big[\tau(n^2/q)\|\mathbb{E}\big(\sqcap_t^2 : \mathcal{F}_t\big)^{1/2}\|_{\exp(L^{2/3})}^{-1}\Big]^{2/3}\bigg) \tag{4.19}$$

$$\le q\exp\big((c_1 - c_2\tau^{2/3})\cdot(n/q)^{1/3}\big). \tag{4.20}$$

Here, we have used Chebyscheff inequality. And, more importantly, the convexity of conditional expectation and L^2-norms to estimate

$$\|\mathbb{E}\big(\sqcap_t^2 : \mathcal{F}_t\big)^{1/2}\|_{\exp(L^{2/3})} \lesssim n^{3/2}/\sqrt{q},$$

by Lemma 4.10. The implied constant is absolute, and determines the constant c_2 in (4.20). For an absolute choice of c_1, and constant τ', we see that we have

$$\mathbb{P}_{x_2,x_3}(E) \lesssim \exp(-\tau'(n/q)^{1/3}). \tag{4.21}$$

We only need $\mathbb{P}_{x_2,x_3}(E) < \frac{1}{2}$, but an exponential estimate of this type is to be expected.

Our last essential estimate is

Lemma 4.22 *For $0 < \kappa < 1$ sufficiently small, $q \le \kappa n^{1/4}$, and $(x_2, x_3) \notin E$, we have*

$$\mathbb{P}_{x_1}\Big(\bigcap_{t=1}^{q/2} \{B_t(\cdot, x_2, x_3) > \tau n/\sqrt{q}\}\Big) \gtrsim \tau^q.$$

Assuming this lemma, we can select $(x_2, x_3) \notin E$. Thus, we see that there is some (x_1, x_2, x_3) so that for all $1 \le t \le q/2$ we have $B_t(x_1, x_2, x_3) > \tau n/\sqrt{q}$, whence

$$\sum_{t=1}^{q/2} B_t(x_1, x_2, x_3) > \tfrac{\tau}{2} \cdot n\sqrt{q}.$$

That is, (4.5) holds. And we can make the last expression as big as $\gtrsim n^{9/8}$.

Proof. If $(x_2, x_3) \notin E$, bring together the definition of E in (4.16), Proposition 4.12, and Lemma 3.5. We see that (3.9) holds (with $\gamma = \tau$, and the q in (3.9) equal to the current $q/2$) provided

$$\tfrac{1}{2} \cdot \tau^{q/2} > \exp\big(-c_1 (n/q)^{1/3}\big).$$

But this is true by inspection, for $q \le \kappa n^{1/4}$. □

5 Heuristics

In two dimensions, Proposition 2.4 clearly reveals an underlying exponential-square distribution governing the Small Ball Inequality. The average case estimate is $n^{1/2}$, and the set on which the sum is about n (a square root gain over the average case) is exponential in n.

Let us take it for granted that the same phenomena should hold in three dimensions. Namely, in three dimensions the average case estimate for a signed small ball sum is n, then the event that the sum exceeds $n^{3/2}$ (a square root gain over the average case) is also exponential in n. How could this be proved? Let us write

$$H = \sum_{\substack{|R| = 2^{-n} \\ |R_1| \ge 2^{-n}}} a_R h_R = \sum_{\substack{|\vec{r}| = n \\ r_1 \le n/2}} f_{\vec{r}} = \sum_{j=0}^{n/2} \beta_j \,,$$

$$\beta_j := \sum_{\substack{|\vec{r}| = n \\ r_1 = j}} f_{\vec{r}} \,.$$

Here we have imposed the same restriction on the first coordinate as we did in Theorem 4.1. With this restriction, note that each β_j is a two-dimensional sum, hence by Proposition 2.3, a sum of bounded independent random variables. It follows that we have by the usual Central Limit Theorem,

$$\mathbb{P}(\beta_j > c\sqrt{n}) \ge \tfrac{1}{4},$$

for a fixed constant c. If one could argue for some sort of independence of the events $\{\beta_j > c\sqrt{n}\}$ one could then write

$$\mathbb{P}(H > cn^{3/2}) \geq \mathbb{P}\Big(\bigcap_{j=0}^{n/2} \{\beta_j > c\sqrt{n}\}\Big) \gtrsim \varepsilon^n,$$

for some $\varepsilon > 0$. This matches the 'exponential in n' heuristic. We cannot implement this proof for the β_j, but can in the more restrictive 'block sums' used above.

We comment on extensions of Theorem 4.1 to higher dimensions. Namely, the methods of this paper will prove

Theorem 5.1 *For $|a_R| = 1$ for all R, we have the estimate estimate in dimensions $d \geq 4$:*

$$\Big\| \sum_{\substack{|R|=2^{-n} \\ |R_1| \geq 2^{-n/2}}} a_R h_R \Big\|_{L^\infty} \gtrsim n^{(d-1)/2+1/4d}.$$

We restrict the sum to those dyadic rectangles whose first side has the lower bound $|R| \geq 2^{-n/2}$.

This estimate, when specialized to $d = 3$ is worse than that of Theorem 4.1 due to the fact that the full extension of the critical estimate Lemma 4.10 is not known to hold in dimensions $d \geq 4$. Instead, this estimate is known. Fix the coefficients $a_R \in \{\pm 1\}$ as in Theorem 5.1, and let $f_{\vec{r}}$ be the corresponding r-functions. For $1 \ll q \ll n$, define $\mathbb{I}_t$ as above, namely $\{\vec{r} : |\vec{r}| = n,\ r_1 \in J_t\}$. Define $\sqcap_t$ as in (4.4). The analog of Lemma 4.10 in dimensions $d \geq 4$ are

Lemma 5.2 *In dimensions $d \geq 4$ we have the estimate*

$$\|\sqcap_t\|_{\exp(L^{2/(2d-1)})} \lesssim n^{(2d-3)/2}/\sqrt{q}.$$

See [4, Section 5, especially (5.3)], which proves the estimate above for the case of $q = 1$. The details of the proof of Theorem 5.1 are omitted, since the theorem is at this moment only a curiosity.

It would be quite interesting to extend Theorem 5.1 to the case where, say, one-half of the coefficients are permitted to be zero. This result would have implications for Kolmogorov entropy of certain Sobolev spaces; as well this case is much more indicative of the case of general coefficients a_R. As far as the authors are aware, there is no straight forward extension of this argument to the case of even a small percentage of the a_R being zero.

Acknowledgements The authors are grateful to the Fields Institute, and the American Mathematical Institute for hospitality and support, and to the National Science Foundation for support through the grants DMS-0456538 and DMS-0801036.

References

[1] J. Beck, A two-dimensional van Aardenne–Ehrenfest theorem in irregularities of distribution, *Compositio Math.* **72** (1989), 269–339. MR1032337 (91f:11054)

[2] J. Beck and W. W. L. Chen, *Irregularities of distribution*, Cambridge Tracts in Mathematics 89, Cambridge University Press, Cambridge, 1987, ISBN 0-521-30792-9, MR903025 (88m:11061)

[3] D. Bilyk, M. Lacey, and A. Vagharshakyan, On the signed small ball inequality, *Online J. Anal. Comb.* **3** (2008), Art. 6, 7, MR 2375609

[4] D. Bilyk, M. T. Lacey, and A. Vagharshakyan, On the small ball inequality in all dimensions, *J. Funct. Anal.* **254** (2008), 2470–2502. MR 2409170

[5] D. Bilyk and M. T. Lacey, On the small ball inequality in three dimensions, *Duke Math. J.* **143** (2008), 81–115. MR 2414745

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

[7] K. F. Roth, On irregularities of distribution, *Mathematika* **1** (1954), 73–79. MR 0066435 (16,575c)

[8] W. M. Schmidt, Irregularities of distribution. VII, *Acta Arith.* **21** (1972), 45–50. MR 0319933 (47 #8474)

[9] W. M. Schmidt, Irregularities of distribution. X, in: *Number theory and algebra*, 311–329, Academic Press, New York, 1977. MR 0491574 (58 #10803)

[10] M. Talagrand, The small ball problem for the Brownian sheet, *Ann. Probab.* **22** (1994), 1331–1354. MR 95k:60049

[11] V. N. Temlyakov, An inequality for trigonometric polynomials and its application for estimating the entropy numbers, *J. Complexity* **11** (1995), 293–307. MR 96c:41052