Probability Mass of Rademacher Sums Beyond One Standard Deviation

Vojtěch Dvořák[1] and Ohad Klein[2]

[1]Department of Pure Maths and Mathematical Statistics, University of Cambridge, UK
[2]Department of Mathematics, Bar Ilan University, Ramat Gan, Israel
Email addresses: [1]vd273@cam.ac.uk, [2]ohadkel@gmail.com

August 1, 2022

Abstract

Let $a_1, \dots, a_n \in \mathbb{R}$ satisfy $\sum_i a_i^2 = 1$, and let $\varepsilon_1, \dots, \varepsilon_n$ be independent uniformly random $\pm$ signs and $X = \sum_{i=1}^n a_i \varepsilon_i$. It is conjectured that $X = \sum_{i=1}^n a_i \varepsilon_i$ has $\Pr[X \geq 1] \geq 7/64$. The best lower bound so far is $1/20$, due to Oleszkiewicz [12]. In this paper we improve this to $\Pr[X \geq 1] \geq 6/64$.

Keywords: Rademacher sums; combinatorial probability; anti-concentration

1 Introduction

1.1 Background

Tail inequalities characterize the possible values of $\Pr[X \geq t]$ for various thresholds t and random variables X with mean 0. We consider the case of Rademacher sums $X = \sum_{i \in [n]} a_i \varepsilon_i$ for real numbers a_i and independently and uniformly distributed signs $\varepsilon_i \sim \{-1, 1\}$. We further focus on lower bounds to $\Pr[X \geq t]$.

If $t > \sqrt{\mathrm{Var}(X)}$ we may have $\Pr[X \geq t] = 0$. If $t \leq 0$, clearly $\Pr[X \geq t] \geq \frac{1}{2}$ because of the symmetry, and if $0 < t < \sqrt{\mathrm{Var}(X)}$, the Paley-Zygmund inequality gives

$$\Pr[X \geq t] \geq \Pr[X > t] = \frac{1}{2} \Pr[X^2 > t^2] \geq \frac{1}{2}(1 - \frac{t^2}{\mathrm{Var}(X)})^2 \frac{\mathrm{Var}(X)^2}{\mathbb{E}[X^4]} > 0.$$

What happens when $t = \sqrt{\operatorname{Var} X}$? This case was studied in 1967 by Burkholder [3] with the conclusion that if $C_s = \inf_X \Pr[X \geq s\sqrt{\operatorname{Var}(X)}]$, where the infimum is taken over all Rademacher sums, then $C_1 > 0$. It was then improved by Hitczenko and Kwapień [7] to $C_1 \geq e^{-4}/8$, and then in 1996 by Oleszkiewicz [12] to $C_1 \geq 1/20$. Hitczenko and Kwapień [7] conjectured that $C_1 = 7/64$, having the tightness example $a_1 = \cdots = a_6 > 0$.

We point out that this problem is a natural counterpart to the Tomaszewski's problem [6], which in the same setting of Rademacher sums, is concerned with the value of $\inf_X \Pr[|X| \leq \sqrt{\operatorname{Var}(X)}]$. This problem attracted wide attention over the years before it was finally settled recently by Keller and the second author [9] – the value is exactly $\frac{1}{2}$ (and henceforth, $C_{-1} = 3/4$).

1.2 Our results

The main result of our paper is the following.

Theorem 1.1. *Any Rademacher sum $X = \sum_i a_i \varepsilon_i$ has*

$$\Pr[X \geq \sqrt{\operatorname{Var}(X)}] \geq 6/64.$$

This theorem improves on the previously best known bound by Oleszkiewicz [12], who derived an analogous result with the constant $\frac{1}{20} = 0.05$ instead of our constant $\frac{6}{64} = 0.09375$. We believe that our tools could be useful in order to prove the conjectured optimal bound of $\frac{7}{64}$. We make some progress toward this goal by handling certain difficult, near-extremal, classes of Rademacher sums. See further Section 1.4.

While already $\Pr[X > \sqrt{\operatorname{Var}(X)}]$ might be 0, as demonstrated by $X = 1 \cdot \varepsilon_1$, the aforementioned proof by Oleszkiewicz [12] in fact shows that $\Pr[X > \sqrt{\operatorname{Var}(X)}] \geq 1/20$ whenever X is not of the form $a_i \varepsilon_i$. This bound is quite tight due to the example $a_1 = \cdots = a_4 > 0$ having $\Pr[X > \sqrt{\operatorname{Var}(X)}] = 1/16$. We show that this is indeed the extremal case.

Theorem 1.2. *Any Rademacher sum $X = \sum_i a_i \varepsilon_i$ with $a_1, a_2 > 0$ has*

$$\Pr[X > \sqrt{\operatorname{Var}(X)}] \geq 1/16.$$

Another inequality in this vein was conjectured by Lowther [10] to be $C_{1/\sqrt{7}} = 1/4$, which is saturated by $a_1 = \cdots = a_7 > 0$. We prove the following slightly weaker result.

Theorem 1.3. *Any Rademacher sum $X = \sum_i a_i \varepsilon_i$ has*

$$\Pr[X > 0.35\sqrt{\operatorname{Var}(X)}] \geq 1/4.$$

In the paper of Ben-Tal, Nemirovski and Roos [1], the higher-dimensional analogue of the $C_1 = 7/64$ problem first appeared. In this setting, $X = \sum_i a_i \varepsilon_i$ with $a_i \in \mathbb{R}^d$ and we are concerned with the probability $P(X) := \Pr\left[\, \|X\|_2^2 \geq \mathbb{E}[\|X\|_2^2]\right]$. The best result in this framework is due to Veraar [16] who showed that $P(X) \geq (\sqrt{12} - 3)/15 \approx 0.031$. We remark that the following holds.

Theorem 1.4. *Any $X = \sum_i a_i \varepsilon_i$ with $a_i \in \mathbb{R}^d$ (for any $d \geq 1$) has*

$$\Pr\left[\|X\|_2^2 \geq \mathbb{E}[\|X\|_2^2]\right] \geq \frac{1 - \sqrt{1 - 1/e^2}}{2} > 0.035.$$

Interestingly, we are not aware of any example that would demonstrate that the constant in Theorem 1.4 could not be as large as $\frac{7}{32}$ (which is the best one could hope for, since the result does not hold for any constant larger than that even when we only consider the case $d = 1$, as commented previously).

1.3 Overview of techniques

A prevalent method for understanding the distribution of Rademacher sums is to partition their weights $\{a_i\}$ into two parts ($X = L + S$): large weights and small weights. Such partitioning is efficient, as the Rademacher sum having small weights is easy to analyze using quantitative versions of the Central Limit Theorem, while the Rademacher sum having large weights can be analyzed by enumeration over all the possibilities. In high level, this is the approach we take, but let us dive a little further into the details.

Consider a Rademacher sum X with $\mathrm{Var}(X) = 1$. The problem addressed in Theorem 1.1 concerns with lower bounding $\Pr[X \geq 1]$. It turns out to be instructive to generalize this problem in two different ways:

- Enable a more flexible threshold t, and not only $t = 1$.
- Impose a restriction on the weights: $|a_i| \leq a$ for a parameter $a \leq 1$.

Denote by $G(a, t)$ the answer to this more general problem: the infimum of $\Pr[X \geq t]$, assuming $|a_i| \leq a$ ($a \in (0, 1]$, $t \in \mathbb{R}$). Ultimately, Theorem 1.1 is encapsulated in the statement $G(1, 1) \geq 6/64$, but we study $G(a, t)$ for all parameters a, t at once.

The crucial point is that using the decomposition of our Rademacher sum to its large and small parts $X = L + S$, we can lower bound $G(a, t)$ by

$$G(a, t) \geq \inf_L \mathop{\mathbb{E}}_{l \sim L} [G(a'/\sigma, (t - l)/\sigma)] \tag{1}$$

where the infimum is taken over all possible values of L induced by decompositions $X = L + S$ (for example, if we decompose $X = L + S$ with $L = a_1\varepsilon_1 + a_2\varepsilon_2$ whenever $a_1 + a_2 \geq 1$ and $L = a_1\varepsilon_1$ otherwise, the infimum is taken over all $L = a_1\varepsilon_1 + a_2\varepsilon_2$ with $a_1 + a_2 \geq 1$ and with $a' = \min(a_2, \sqrt{1 - a_1^2 - a_2^2})$ and $L = a_1\varepsilon_1$ with $a' = \min(a_1, 1 - a_1)$), the expectation is taken over l being a realization of the random variable L, σ is the standard deviation of S (that is, $\sqrt{1 - \mathrm{Var}(L)}$), and a' is an upper bound on the weights of S (whose value depends on the notion of how we decompose $X = L + S$).

Equation (1) enables one to recursively compute lower bounds on $G(a, t)$, and ultimately on $G(1, 1)$. Roughly speaking, considering the decompositions $X = L + S$ with L containing at most the three largest weights of X, we almost deduce Theorem 1.1. However, using solely this method, we run into the following problem: In order to concretely define $G(a, t)$ through the recursive (1), we have to propose an initial lower estimate for $G(a, t)$. The initial estimate we use is 'continuous' in nature (the Berry-Esseen inequality), and is unable to differentiate between bounds on $\Pr[X \geq t]$ and on $\Pr[X > t]$. However, there are various instances X, detailed in Section 1.4, for which the stronger bound $\Pr[X > 1] \geq \frac{7}{64}$ (or even the bound $\Pr[X > 1] \geq \frac{6}{64}$, that we prove) does not hold! (e.g. the aforementioned $a_1 = \cdots = a_4 > 0$.)

To handle these more tight cases, we take a completely different approach toward lower bounding $\Pr[X \geq 1]$ (i.e. Theorem 1.1). That is, we *upper bound* $\Pr[X \in (-1, 1)]$ (recall that X is symmetric). To do that, we take the advantage of the following trade-off that usually arises. The collections $\{a_1, \ldots, a_n\}$ that either contain large mass of their variance in the small weights, or have their large weights very non-uniform, are harder to describe precisely, but are nevertheless easy to analyze, since usually stronger bounds hold for these. And the collections $\{a_1, \ldots, a_n\}$ that contain only very small mass of their variance in the small weights and have their large weights quite uniform are easier to describe precisely, so despite only more tight bounds being true for these, we can derive those bounds.

In various tight cases that arise, we commonly want to upper bound $\Pr[X \in I]$ for some particular interval $I \subset \mathbb{R}$. To do that, we use a chain lemma, and a few related observations.

In the chain lemma, we assume X has some weights $a_1, \ldots, a_l$ which are 'large' compared to the length of I and consider the signed sums $\pm a_1 \pm \ldots \pm a_l$ – ignoring the remaining 'small' weights. We then associate the set of these 2^l signed sums with a hypercube graph in a natural way and then use a famous result of Erdős [5] to show that these sums are not very tightly concentrated. That in turn implies an upper bound on $\Pr[X \in I]$.

Occasionally, we have to consider the case when I is a very short interval (much smaller than $(-1, 1)$). In such a case we divide the small weights into disjoint parts (a method introduced by Montgomery-Smith [11]), so that each part has a substantial probability to be large compared to I, and apply the chain lemma on these 'large' parts to deduce that $\Pr[X \in I]$ is small enough.

1.4 Difficult cases

As described in the previous subsection, similarly to Tomaszewski's problem [9], the particular difficulty we are facing when trying to prove the conjecture $C_1 = 7/64$, are the cases when $\Pr[X > 1] < 7/64$ despite $\Pr[X \geq 1] \geq 7/64$ (and their 'neighborhoods', i.e. the collections with the few largest weights being roughly of the same sizes as in these cases). Notably, we have

- for $a_1 = 1$, $\Pr[X > 1] = 0$;
- for $a_1 = \ldots = a_4 = \frac{1}{2}$, $\Pr[X > 1] = \frac{1}{16}$;
- for $a_1 = \ldots = a_9 = \frac{1}{3}$, $\Pr[X > 1] = \frac{23}{256} \approx 0.0898\ldots < \frac{6}{64}$;
- for $a_1 = \frac{2}{3}, a_2 = \ldots = a_6 = \frac{1}{3}$, $\Pr[X > 1] = \frac{6}{64}$;
- for $a_1 = a_2 = \frac{1}{2}, a_3 = \ldots = a_{10} = \frac{1}{4}$, $\Pr[X > 1] = \frac{55}{512} < \frac{7}{64}$.

We have to deal with the first three cases even when proving our bound of $6/64$, and the last two cases are further hurdles on the way to the optimal bound.

In our proof of the $6/64$ bound, big part of the argument is spent dealing with a subcase presented in Section 4.1.4, which corresponds to the collections 'close to' the third case from above (which is the most intricate of the first three 'barriers').

In Section 6, we discuss these difficulties in more detail and make progress toward proving the $7/64$ bound, by proving it for families corresponding to the 'neighbourhoods' of all the cases above except the third one.

1.5 Organization

In Section 2, we introduce notation, and define a certain type of a useful random process. In Section 3, we describe our main tools and prove Theorem 1.3. We then use these tools in Section 4 to prove Theorem 1.1, the main result of the paper. Section 5 contains the proof of Theorem 1.2. In Section 6, we discuss the deficiency of our 6/64 proof and propose how to advance toward 7/64, proving the result in two out of three 'difficult' cases. In Section 7, we discuss the high dimensional version of the problem as well as of the problem of Tomaszewski and prove Theorem 1.4. Finally in Section 8, we summarize the open problems arising in the paper.

Some of the more technical proofs from various parts of the paper are in Appendix A and Appendix B.

2 Background and definitions

In this section, we describe our setting, notation and assumptions that we are working with.

Throughout, we will consider $X = \sum_{i=1}^{n} a_i \varepsilon_i$, where ε_i are independent Rademacher random variables (i.e. independent random variables such that $\Pr\left[\varepsilon_i = +1\right] = \Pr\left[\varepsilon_i = -1\right] = \frac{1}{2}$) and a_i are real numbers with $\sum_{i=1}^{n} a_i^2 = 1$. Moreover, we will always, without loss of generality, assume that

$$a_1 \geq a_2 \geq \ldots \geq a_n > 0.$$

Sometimes, we will work with variables $\{b_i\}$ or $\{c_i\}$ instead of $\{a_i\}$. For these, we do not assume any conditions on their ordering unless so stated.

At some points, we will also write $\mathbf{a}$ to denote $\{a_1, \ldots, a_n\}$.

Our central aim will be to lower bound

$$\Pr\left[X \geq 1\right] = \frac{1}{2} \Pr\left[|X| \geq 1\right]. \tag{2}$$

At some points, we will work with $\Pr\left[X \geq 1\right]$, while at other points, we will work with $\Pr\left[|X| \geq 1\right]$. As expressed by (2), working with these two forms is of course equivalent and the entire proof could be rewritten using just one of these. We use both quantities in order to streamline the proof.

The function $D(a, x) : (0, 1] \times \mathbb{R} \to \mathbb{R}$ appears repeatedly throughout the proof. This is a particular function that we construct in subsection 3.3 and it has a property that for any $a \in (0, 1], x \in \mathbb{R}$, if we have $a_1 \leq a$, then $\Pr\left[X \geq x\right] \geq D(a, x)$. While its computation is computer-aided, we emphasize that by writing 'D', we always refer to its exact value, and not to its approximation.

3 Tools

3.1 Stopped random walks and chain argument

We start with an observation (following trivially from a well known result of Erdős [5]) which we will use repeatedly.

Observation 3.1. *Let $b_1 \geq b_2 \geq \ldots \geq b_t > 0$ be such that $b_{t-k+1} + \ldots + b_t \geq \alpha$ for some $\alpha > 0$ and $0 < k \leq t$. Then, for any x and any $b_{t+1}, \ldots, b_s$, we have*

$$\Pr\big[\sum_{i=1}^{s} b_i\varepsilon_i \in (x-\alpha, x+\alpha)\big] \leq f(k,t)/2^t$$

where $f(k,t)$ denotes the sum of k largest binomial coefficients of the form $\binom{t}{i}$ for some i, $0 \leq i \leq t$.

Proof. If the probability was more than $f(k,t)/2^t$ for some fixed x, then in particular we can choose signs $\varepsilon_{t+1} = \varepsilon'_{t+1}, \ldots, \varepsilon_s = \varepsilon'_s$ in such a way that at least $f(k,t)+1$ of the sums

$$\pm b_1 \pm \ldots \pm b_t + b_{t+1}\varepsilon'_{t+1} + \ldots + b_s\varepsilon'_s$$

are within less than 2α of each other. Let

$$T = \{\pm b_1 \pm \ldots \pm b_t + b_{t+1}\varepsilon'_{t+1} + \ldots + b_s\varepsilon'_s\}.$$

Consider the bijection $g : T \to Q_t \simeq \{\pm 1\}^t$ given by

$$b_1\varepsilon_1 + \ldots + b_t\varepsilon_t + b_{t+1}\varepsilon'_{t+1} + \ldots + b_s\varepsilon'_s \to (\varepsilon_1, \ldots, \varepsilon_t).$$

Let $S \subset T$ be the set of $f(k,t)+1$ elements of T that are all within 2α of each other. Then by the result of Erdős [5, Theorem 5], $g(S)$ contains an chain of length at least k. But that contradicts the assumption that $b_{t-k+1} + \ldots + b_t \geq \alpha$. ☐

Some times, we will only check the stronger condition that (in the cases $k = 2, 3$) no two out of the sums $x_0 \pm b_1 \pm \ldots \pm b_k$ are within less than 2δ of each other, which in particular implies no two hit any interval of the form $(x-\delta, x+\delta)$. For the special cases we need, we will use the following two straightforward observations to verify that.

Observation 3.2. *Fix $\delta > 0$ and $b_1, b_2 \geq \delta$ such that $|b_1 - b_2| \geq \delta$. Then for any x and any $b_3, \ldots, b_l$, we have*

$$\Pr\big[\sum_{i=1}^{l} b_i\varepsilon_i \in (x-\delta, x+\delta)\big] \leq \frac{1}{4}.$$

Proof. If the probability was more than $\frac{1}{4}$ for some fixed x, then in particular we can choose signs $\varepsilon_3 = \varepsilon'_3, \ldots, \varepsilon_l = \varepsilon'_l$ in such a way that at least two of the four sums

$$\pm b_1 \pm b_2 + b_3\varepsilon'_3 + \ldots + b_l\varepsilon'_l$$

are within less than 2δ of each other. Looking at differences of this set, it can only happen if the set

$$D = \{b_1 + b_2,\, b_1,\, b_2,\, |b_1 - b_2|\}$$

contains some element smaller than δ, and our assumptions guarantee that can not happen. $\square$

Observation 3.3. *Fix $\delta > 0$ and $c_1 \geq c_2 \geq c_3 \geq \delta$ such that $c_1 - c_2, c_2 - c_3 \geq \delta$, $|c_1 - c_2 - c_3| \geq \delta$. Then for any x and any $c_4, \ldots, c_m$, we have*

$$\Pr\Big[\sum_{i=1}^{m} c_i\varepsilon_i \in (x - \delta, x + \delta)\Big] \leq \frac{1}{8}.$$

Proof. If the probability was more than $\frac{1}{8}$ for some fixed x, then in particular we can choose signs $\varepsilon_4 = \varepsilon'_4, \ldots, \varepsilon_m = \varepsilon'_m$ in such a way that at least two of the eight sums

$$\pm c_1 \pm c_2 \pm c_3 + c_4\varepsilon'_4 + \ldots + c_m\varepsilon'_m$$

are within less than 2δ of each other. Looking at differences of this set, it can only happen if the set

$$D = \{c_1,\, c_2,\, c_3,\, c_1 \pm c_2,\, c_1 \pm c_3,\, c_2 \pm c_3,\, c_1 + c_2 \pm c_3,\, c_1 - c_2 + c_3,\, |c_1 - c_2 - c_3|\}$$

contains some element smaller than δ; our assumptions guarantee it is impossible. $\square$

In the easy cases, we are already given enough large weights as a part of our collection $\{a_i\}$ and can use these weights in the anti-concentration observations above. But if that is not true and we instead have a lot of very small weights, we can 'generate' larger weights from them, as described in the subsection that follows.

3.2 The random process $W(S; x)$ and its success probability

For a set of real numbers $S = \{d_1, \ldots, d_n\}$ and a real number $x > 0$, we denote by $W(S; x)$ (or by $W(d_1, \ldots, d_n; x)$) the following random process. We first fix a permutation $(i_1, \ldots, i_n)$ of $\{1, \ldots, n\}$ which maximizes the probability that the process is successful (what it means for this process to be successful will be defined in due course). Next, we set $W_0 = 0$. After choosing W_j for some $j < n$, if $|W_j| \geq x$, we set

$$W_{j+1} = \ldots = W_n = W_j.$$

While if $|W_j| < x$, we let $\varepsilon_{i_{j+1}}$ be Rademacher random variable independent of the previous part of the process, and set

$$W_{j+1} = W_j + d_{i_{j+1}} \varepsilon_{i_{j+1}}.$$

We denote by $r(S; x)$ (or by $r(d_1, \ldots, d_n; x)$) the final value of this process, i.e. W_n. We call it *successful* if $|r(S; x)| \geq x$, and *unsuccessful* otherwise.

We denote by $p(S; x)$ (or by $p(d_1, \ldots, d_n; x)$) the probability that the process is successful. In particular, if we have $|d_i| \geq x$ for any $i \in \{1, \ldots, n\}$, clearly the corresponding process will always be successful because of our condition on ordering.

The following lemma is crucial for us when working with such random processes.

Lemma 3.4. *Assume we have positive reals $b_1, \ldots, b_k$ such that $\sum_{i=1}^k b_i^2 \geq c\alpha^2$ for some fixed $c > 1$ and fixed $\alpha > 0$. Then*

$$p(b_1, \ldots, b_k; \alpha) \geq \frac{c-1}{c+3}.$$

Moreover, if for some $\eta \in (0, 1)$, we have $b_1, \ldots, b_k \in (0, \eta\alpha] \cup [\alpha, \infty)$, then

$$p(b_1, \ldots, b_k; \alpha) \geq \frac{c-1}{c + \eta^2 + 2\eta}.$$

Proof. If any term out of $b_1, \ldots, b_k$ has size at least α, then clearly $p(b_1, \ldots, b_k; \alpha) = 1$. So further assume none of the terms has size at least α.

Run the random process $W(b_1, \ldots, b_k; \alpha)$. Without loss of generality (and for notational convenience), we can assume that the ordering $b_1, \ldots, b_k$ maximizes the probability that the process is successful. We define the stopping time T as follows. Let T be the first time i such that $|W_i| \geq \alpha$ if this time is at most k, and let $T = k$ otherwise. Let $p = p(b_1, \ldots, b_k; \alpha)$ be the probability that the process $W(b_1, .., b_k; \alpha)$ is successful, i.e. that it hits absolute value at least α.

Now we will lower and upper bound $\mathbb{E}\big[W_T^2\big]$.

Clearly $|W_T| \leq 2\alpha$ (as every term has size at most α and T is the first time we reach absolute value at least α), and $|W_T| \leq \alpha$ in the case when we never hit absolute value at least α. This gives

$$\mathbb{E}\big[W_T^2\big] \leq 4p\alpha^2 + (1-p)\alpha^2. \tag{3}$$

But also, writing $A = b_1\varepsilon_1 + \ldots + b_T\varepsilon_T$ and $B = b_{T+1}\varepsilon_{T+1} + \ldots + b_k\varepsilon_k$ (setting $B = 0$ if $T = k$), we collect the following easy observations. Firstly

$$\mathbb{E}\big[AB\big] = \sum_{T_0,x} \Pr\big[T = T_0, A = x\big]\mathbb{E}\big[AB|T = T_0, A = x\big] = 0, \tag{4}$$

since for any T_0, x, we have

$$\mathbb{E}\big[AB|T = T_0, A = x\big] = x\mathbb{E}\big[b_{T_0+1}\varepsilon_{T_0+1} + \ldots + b_k\varepsilon_k\big] = 0.$$

Furthermore, noting that if $T = k$, then $B = 0$, we obtain

$$\mathbb{E}\big[B^2\big] = \sum_{i=1}^{k-1} \Pr\big[T = i\big]\mathbb{E}\big[B^2|T = i\big] = \sum_{i=1}^{k-1} \Pr\big[T = i\big]\big(\textstyle\sum_{j=i+1}^{k} b_j^2\big) \leq p\displaystyle\sum_{i=1}^{k} b_i^2. \tag{5}$$

Using (4) we conclude

$$\sum_{i=1}^{k} b_i^2 = \mathbb{E}\big[(A+B)^2\big] = \mathbb{E}\big[A^2\big] + \mathbb{E}\big[B^2\big] + 2\mathbb{E}\big[AB\big] = \mathbb{E}\big[A^2\big] + \mathbb{E}\big[B^2\big]. \tag{6}$$

Overall, combining (5) and (6) we conclude

$$\mathbb{E}\big[A^2\big] \geq (1-p)\sum_{i=1}^{k} b_i^2 \geq (1-p)c\alpha^2. \tag{7}$$

Combining (3) and (7), we obtain

$$(1-p)c\alpha^2 \leq \mathbb{E}\big[W_T^2\big] \leq 4p\alpha^2 + (1-p)\alpha^2.$$

Rearranging gives the first result.

For the second result, just note that with our additional condition $b_1, \ldots, b_k \in (0, \eta\alpha]$, we can replace the inequality

$$\mathbb{E}\big[W_T^2\big] \leq 4p\alpha^2 + (1-p)\alpha^2$$

by the stronger inequality

$$\mathbb{E}\big[W_T^2\big] \leq p(1+\eta)^2\alpha^2 + (1-p)\alpha^2,$$

and conclude in exactly the same way as before. □

3.3 Dynamic Programming bound

Denote by $\widetilde{G}(a_1, x)$ the quantity $\inf_X \Pr[X > x]$ where the infimum is taken over all Rademacher sums X with $\mathrm{Var}(X) = 1$, and whose largest weight is at most a_1.

For the proof, it is useful to understand the function $\widetilde{G}$. Evaluating the function $\widetilde{G}(a_1, x)$ is in general harder than the problem we are concerned with in Theorem 1.1; the latter is, nonrigorously, encapsulated in $\widetilde{G}(1, 1-\epsilon)$.

The goal of the dynamic-programming approach is to derive a lower bound on $\widetilde{G}$ by first obtaining some lower bound on $\widetilde{G}(a_1, x)$ for many values of a_1, x, and then using an iterative procedure to improve this bound further. The key tool enabling us to iterate is elimination of the largest weight (see Section 3.4 for more details about elimination).

3.3.1 Prawitz's smoothing Inequality

We will use a smoothing inequality of Prawitz [15]. This inequality is a useful tool, providing bounds on the values of the cumulative distribution function of a random variable, in terms of a partial information regarding its characteristic function. Specifically, given the characteristic function of a random variable, it is possible to determine its distribution via the Gil-Pelaez formula. In the case of a Rademacher sum $X = \sum_i a_i \varepsilon_i$, we have the characteristic function $\varphi_X(t) = \prod_i \cos(a_i t)$. Assuming that we know the largest weight a_1, it is possible to estimate the value of $\varphi_X(t)$ for $t \ll 1/a_1$. Although for $t \gg 1/a_1$, we have no information regarding $\varphi_X(t)$, Prawitz' inequality is still capable of providing a decent estimate for the cumulative distribution function of X.

While the inequality is applicable to all random variables, it was shown in [9] that its specialization to Rademacher sums gives tighter estimates.

Prawitz' bound gives a lower bound on $\widetilde{G}(a_1, x)$, for all parameters $q \in [0, 1],\ T > 0$:

$$\forall q \in [0,1],\, T > 0\colon \qquad \widetilde{G}(a_1, x) \geq F(a_1, x, T, q). \tag{8}$$

Specifically, a formula for F may be derived from [9, Proposition 4.2] (which is derived from [15]):

$$F(a,x,T,q) = 1/2 - \int_0^q |k(u,x,T)|\, g(Tu,a)\mathrm{d}u - \int_q^1 |k(u,x,T)|\, h(Tu,a)\mathrm{d}u \\ - \int_0^q k(u,x,T)\exp(-(Tu)^2/2)\mathrm{d}u, \tag{9}$$

where[1] $k(u,x,T) = \frac{(1-u)\sin(\pi u + Tux)}{\sin(\pi u)} + \frac{\sin(Tux)}{\pi}$,

$$g(v,a) = \begin{cases} \exp(-v^2/2) - \cos(av)^{1/a^2}, & av \le \frac{\pi}{2} \\ \exp(-v^2/2) + 1, & \text{otherwise} \end{cases}, \quad h(v,a) = \begin{cases} \exp(-v^2/2), & av \le \theta \\ (-\cos(av))^{1/a^2}, & \theta \le av \le \pi \\ 1, & \text{otherwise} \end{cases},$$

$Z \sim N(0,1)$ is a standard Gaussian and $\theta = 1.778 \pm 10^{-4}$ is the unique solution of $\exp(-\theta^2/2) = -\cos(\theta)$ in the interval $[0,\pi]$. We note that $F(a,x,T,q)$ is a function (weakly) decreasing in a.

3.3.2 Recursion

Note that as in (1), by considering the two values that the sign of the largest weight can take (see subsection 3.4 for more details), we have

$$\widetilde{G}(a_1,x) \ge \frac{1}{2}\inf_{a\in(0,a_1]}\left(\widetilde{G}\left(\frac{a}{\sqrt{1-a^2}},\frac{x-a}{\sqrt{1-a^2}}\right) + \widetilde{G}\left(\frac{a}{\sqrt{1-a^2}},\frac{x+a}{\sqrt{1-a^2}}\right)\right). \tag{10}$$

Hence, $\widetilde{G}$ is lower bounded by the lowest function satisfying both inequalities (8), (10). Computationally, to obtain a concrete lower bound on $\widetilde{G}$, we iteratively define the functions

$$D_i \colon (0,1)\times\mathbb{R} \to \mathbb{R}$$

by $D_0(a_1,x) = \max(F(a_1,x),\ \mathbb{1}\{x<0\}/2)$ with $F(a_1,x) = \sup_{T,q}\{F(a_1,x,T,q)\}$ and

$$D_{i+1}(a_1,x) = \max\Big(D_i(a_1,x), \frac{1}{2}\inf_{a\in(0,a_1]}\Big(D_i\Big(\frac{a}{\sqrt{1-a^2}},\frac{x-a}{\sqrt{1-a^2}}\Big) + D_i\Big(\frac{a}{\sqrt{1-a^2}},\frac{x+a}{\sqrt{1-a^2}}\Big)\Big)\Big), \tag{11}$$

[1] k can be smoothly continued to the range $u \in \{0,1\}$ by setting $k(0,x,T) = 1 + Tx/\pi$ and $k(1,x,T) = 0$.

and observe that $\widetilde{G}(a_1, x) \geq D_i(a_1, x)$ for all i. Choosing a large I ($I = 10$ suffices) and writing

$$D(a_1, x) = D_I(a_1, x)$$

we derive

$$\forall X \in \mathcal{X}\colon (X = \sum b_i \varepsilon_i \wedge |b_i| \leq a_1) \implies \Pr[X > x] \geq D(a_1, x). \tag{12}$$

Note that D is a function depending on two continuous variables, which cannot be stored programmatically. We compute $D_i(a_1, x)$ for $a_1 \in [0, 1]$ and $x \in [-3, 3]$ with granularity of $\delta = 1/400$ (a_1 starting from 0 and x starting from -3). Correspondingly, we replace (11) with a variant that feeds D_{i+1} with arguments rounded up (to a multiple of δ), hence underestimating D_{i+1}; This enables considering a finite set of $a \in [0, a_1]$ in the infimum at (11). We apply this rounding-up to both the $\frac{a}{\sqrt{1-a^2}}$ and the $\frac{x \pm a}{\sqrt{1-a^2}}$ arguments. Moreover, in any computation of $D(a_1, x)$ we round the arguments up to multiples of δ. When $x < -3$ we round x to -3, and when $x \geq 3$ we round x to ∞ and set $D(a_1, \infty) = 0$. This results in a dynamic-programming method for computing $D_i(a_1, x)$.

Our implementation of this computation can be found at [4].

Several concrete values. Along the paper, we use the following lower bounds for values of D, derived by the described computation.

$$\begin{aligned}
&D(0.35, 0.35) > \frac{1}{4}, && D(0.3, 1) > \frac{3}{32}, \\
&D(0.3/\sqrt{0.51}, 0.3/\sqrt{0.51}) > \frac{3}{16}, && D(0.4, 1) > \frac{1}{12}, \\
&D(0.5, 0.5) > \frac{1}{6}, && D(0.34, 1.42) > 0.04, \\
&D(0.43, 1.42) > 0.03, && D(0.51, 1.01) = \frac{1}{16}.
\end{aligned} \tag{13}$$

Note that $D(0.51, 1.01) = 1/16$ is a precise value (unlike the other values mentioned for which we just have lower bounds). On the one hand, we clearly see that $D(0.51, 1.01) \leq 1/16$, as saturated by the weights $a_1 = \ldots = a_4 = 1/2$. On the other hand, to derive $D(0.51, 1.01) \geq 1/16$, it is crucial that we set $D_0(a_1, x) = \max(F(a_1, x), \mathbb{1}\{x < 0\}/2)$ instead of just using $F(a_1, x)$. Our iterative procedure and the lower bounds on $F(a_1, x)$ are then enough to prove $D(0.51, 1.01) \geq 1/16$.

Precision. As described, the lower bound $D(a, x)$ we numerically get for $\widetilde{G}(a, x)$ is precise. The only detail disregarded so far is the computation of $F(a_1, x)$. Programmatically we replace $F(a_1, x)$ by $F(a_1, x, \pi/a_1, 0.5)$, that is, we do not compute the maximum of $F(a_1, x, T, q)$ over all values of T, q, but set $T = \pi/a_1$ and $q = 0.5$. Since we use $F(a_1, x)$ as a lower bound, this underestimation of $F(a_1, x)$ is valid. We further note that this choice of T, q simplifies the first integrand in $F(a_1, x, T, q)$ to be continuous (specifically, $g(v, a)$ is applied only when $av \leq \pi/2$). Finally, to numerically estimate the integrals appearing in the definition of $F(a, x, T, q)$ we take two approaches.

In the first approach we compute the integrals appearing in (9) verbatim by using the standard Python integrator `scipy.integrate.quad`, and check that the integrator estimates that its error is well below some constant (0.01) that we discount from $F(a, x, T, q)$. We also split the domains of integration so that the integrands are smooth in each subdomain. This evaluation of F is simple, but requires relying on the accuracy of `scipy.integrate.quad`.

In the second approach we compute the integrals with the trapezoid rule, using explicit bounds B on the derivatives of the integrands (more accurately, we use that these are B-lipschitz functions), to get an explicit estimation of the integrals, together with a provable error estimates. The bounds B are computed in [9, Appendix B.2].

While the first approach is neat and simple, the second approach is transparent and reviewable. The accompanied code is available at [4].

3.4 Elimination

Elimination is the process of replacing a probabilistic inequality in $X = \sum_{i=1}^{n} a_i \varepsilon_i$, by an inequality involving $Z = \sum_{i=m}^{n} a_i \varepsilon_i$ with $m > 1$. For example, the inequality

$$\Pr[X \geq 1] \geq 3/32$$

is equivalent to the following inequality, which involves $Z = \sum_{i=2}^{n} a_i \epsilon_i$ (i.e. $m = 2$),

$$\Pr[Z \geq 1 - a_1] + \Pr[Z \geq 1 + a_1] \geq 3/16.$$

via the law of total probability. A more elaborate derivation can be found at [9, Lemma 2.1].

3.5 A $1/\sqrt{7}$-type inequality

Lowther [10] conjectured that $\Pr[|X| \geq 1/\sqrt{7}] \geq 1/2$ is true for all Rademacher sums X with $\mathrm{Var}(X) = 1$. In the proof of Theorem 1.1 we make use of Theorem 1.3, i.e. $\Pr[|X| > 0.35] \geq 1/2$, which we henceforth prove.

We split into two cases. If $a_1 > 0.35$, and $\varepsilon' = (-\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$, then at least one of $X(\varepsilon)$ and $X(\varepsilon')$ has absolute value more than a_1, hence $\Pr[|X| > 0.35] \geq 1/2$. If $a_1 \leq 0.35$, then we conclude using (13) since

$$D(0.35, 0.35) > 1/4.$$

4 Proof of $\Pr[X \geq 1] \geq 3/32$

In this section we show that for any Rademacher sum X with $\mathrm{Var}(X) = 1$,

$$\Pr[X \geq 1] \geq 3/32, \tag{14}$$

that is, Theorem 1.1. The proof splits into two main cases - the case when $a_1 + a_2 + a_3 \leq 1$ and the case when $a_1 + a_2 + a_3 > 1$.

In the case $a_1 + a_2 + a_3 \leq 1$, the tools we have developed in subsections 3.3 and 3.4 enable us to handle most of the subcases. Nevertheless, as discussed before, one can not hope for these tools to work in the subcase $(a_1, a_2, a_3) \approx (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and $a_1 + a_2 + a_3 \leq 1$. Thus, we spend majority of this subsection dealing with the subcase $a_3 \geq 0.325$ and $a_1 + a_2 + a_3 \leq 1$. To do that, we use the tools developed in subsection 3.1. Our strategy is to show that the family of such collections $\{a_i\}$ with $a_3 \geq 0.325$ and $a_1 + a_2 + a_3 \leq 1$ is contained in the union of several subfamilies, for each of which we can obtain the desired bound.

In the case $a_1 + a_2 + a_3 > 1$, the proof is less lengthy. We divide it into several subcases and use the tools from subsections 3.3 and 3.4 and crucially also Theorem 1.3, to resolve these cases.

4.1 Case $a_1 + a_2 + a_3 \leq 1$

4.1.1 Subcase $a_1 \leq 0.3$

Using (13) we have

$$D(0.3, 1) > 3/32$$

implying the assertion (14) through (12).

4.1.2 Subcase $a_1 \geq 0.7$

Using elimination, in order to deduce (14) regarding $X = \sum_{i=1}^{n} a_i \varepsilon_i$ it suffices to check

$$\Pr\left[X' \geq \frac{1 - a_1}{\sqrt{1 - a_1^2}}\right] \geq 3/16 \tag{15}$$

with $X' = \frac{1}{\sqrt{1-a_1^2}} \sum_{i=2}^n a_i \varepsilon_i$ the a_1-eliminated version of X. Using (13) we deduce (15) from

$$D(0.3/\sqrt{0.51}, 0.3/\sqrt{0.51}) > 3/16,$$

since $a_1 + a_2 \leq 1$. This argument does not rely on $a_1 + a_2 + a_3 \leq 1$, but only assumes $a_1 + a_2 \leq 1$ (and $a_1 \geq 0.7$). This is used in Section 4.2.2.

4.1.3 Subcase $a_3 \leq 0.325$ and $a_1 \in [0.3, 0.7]$

Under the conditions $a_1 \geq 0.3$ and $a_3 \leq 0.325$ (and $a_2 \in [a_3, a_1]$), denote $\sigma_2 = \sqrt{1 - a_1^2 - a_2^2}$, and note that $a = \min(1 - a_1 - a_2, a_2, 0.325)$ is an upper bound on a_3. We show in Appendix A.1 that

$$\mathop{\mathbb{E}}_{\varepsilon \in \{-1,1\}^2} \left[D\left(\frac{a}{\sigma_2}, \frac{1 + a_1\varepsilon_1 + a_2\varepsilon_2}{\sigma_2} \right) \right] \geq 3/32, \tag{16}$$

verifying (14) in this case, via elimination of a_1, a_2.

4.1.4 Subcase $a_3 \geq 0.325$

Let $Y = \sum_{i=4}^n a_i \varepsilon_i$ and denote:

$$\begin{aligned}
q_1 &= \Pr\big[|Y| \geq 1 - a_1 - a_2 - a_3\big], & q_2 &= \Pr\big[|Y| \geq 1 - a_1 - a_2 + a_3\big], \\
q_3 &= \Pr\big[|Y| \geq 1 - a_1 + a_2 - a_3\big], & q_4 &= \Pr\big[|Y| \geq 1 + a_1 - a_2 - a_3\big], \\
q_5 &= \Pr\big[|Y| \geq 1 - a_1 + a_2 + a_3\big], & q_6 &= \Pr\big[|Y| \geq 1 + a_1 - a_2 + a_3\big], \\
q_7 &= \Pr\big[|Y| \geq 1 + a_1 + a_2 - a_3\big], & q_8 &= \Pr\big[|Y| \geq 1 + a_1 + a_2 + a_3\big].
\end{aligned}$$

Then using elimination, we have $\Pr[X \geq 1] = \frac{1}{16}(q_1 + \ldots + q_8)$. Hence we are required to show

$$q_1 + \ldots + q_8 \geq 3/2. \tag{17}$$

The key lemma which lets us handle this case, is the following.

Lemma 4.1. *Let $\mathcal{A}$ be the family of the collections $\mathbf{a} = (a_1, \ldots, a_n)$ with $n \geq 4$, $a_1 \geq \ldots \geq a_n > 0$, $\sum_{i=1}^n a_i^2 = 1$, $a_1 + a_2 + a_3 \leq 1$ and $a_3 \geq 0.325$. Then $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{A}_3$, where $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ are the subsets of $\mathcal{A}$ characterized by the following additional conditions:*

- $\mathcal{A}_1$*:* $a_4 \leq 7/40$,
- $\mathcal{A}_2$*:* $q_1 \geq \frac{793}{1024}$,

- $\mathcal{A}_3$*:* $q_2, q_3 \geq \frac{37}{128}$.

Proof. Firstly, if we had $a_1 + a_2 + a_3 = 1$, then clearly $q_1 = 1$. So further consider only the case $a_1 + a_2 + a_3 < 1$. Write $a_3 = \frac{1}{3} - \delta$, and assume that $a_1 + a_2 + a_3 < 1$ and $0 < \delta \leq \frac{1}{120}$ (which is equivalent to $a_3 \geq 0.325$). Note that

$$1 - a_1 - a_2 - a_3 \leq 1 - 3a_3 = 3\delta, \tag{18}$$

and that

$$1 - a_1 - a_2 + a_3, 1 - a_1 + a_2 - a_3 \leq 1 - a_3 = \frac{2}{3} + \delta. \tag{19}$$

If $a_4 \leq 21\delta \leq 7/40$, we have $\mathbf{a} \in \mathcal{A}_1$. So further assume that $a_4 \geq 21\delta$, in which case we have to show that $\mathbf{a} \in \mathcal{A}_2 \cup \mathcal{A}_3$.

Let k be the smallest integer such that $a_k < 1 - a_1 - a_2 - a_3$ (if $a_n \geq 1 - a_1 - a_2 - a_3$, set $k = n + 1$). Note that $k \geq 5$, since

$$a_4 \geq 21\delta > 3\delta \geq 1 - a_1 - a_2 - a_3,$$

where the last inequality follows by (18).

Claim 4.2. *If* $\sum_{i=k}^{n} a_i^2 \geq 450\delta^2$*, then* $\mathbf{a} \in \mathcal{A}_2$.

Proof of Claim 4.2. Note that $a_k, \ldots, a_n < 3\delta$. We can find disjoint subsets $S, T_1, \ldots, T_4$ of $\{a_k, \ldots, a_n\}$ with the following properties. We have

$$234\delta^2 \geq \sum_{i \in S} a_i^2 \geq 225\delta^2$$

and for $j = 1, \ldots, 4$, we have

$$54\delta^2 \geq \sum_{i \in T_j} a_i^2 \geq 45\delta^2.$$

Now consider the corresponding random processes $W(S; 9\delta)$ and $W(T_j; 3\delta)$ for $j = 1, \ldots, 4$.

We consider three events partitioning our probability space. The first event is the event C_1 that $W(S; 9\delta)$ is successful and also at least one out of $W(T_j; 3\delta)$ for $j = 1, \ldots, 4$ is successful. The second event is the event $C_2 = C_2' \cap C_1^C$, where C_2' is the event that at least one out of

$$W(S; 9\delta),\ W(T_1; 3\delta), \ldots,\ W(T_4; 3\delta)$$

is successful. And the last event is $C_3 = C_1^C \cap C_2^C$.

By independence of the processes $W(S;9\delta), W(T_1;3\delta), \ldots, W(T_4;3\delta)$ and Lemma 3.4, we have

$$\Pr\left[C_1\right] \geq \frac{15}{32}, \qquad \Pr\left[C_3\right] \leq \frac{1}{32}. \tag{20}$$

We start by assessing the probability $\Pr\left[|Y| \geq 1 - a_1 - a_2 - a_3\right]$ conditioned on C_1. We look at

$$\Pr\left[|\sum_{i=4}^{n} a_i\varepsilon_i| < 1 - a_1 - a_2 - a_3 | C_1, x_1, i_1, j, x_2, i_2, r(T_1;3\delta), \ldots, r(T_{j-1};3\delta)\right],$$

for fixed $x_1, i_1, j, x_2, i_2, r(T_1;3\delta), \ldots, r(T_{j-1};3\delta)$, where x_1, i_1, j, x_2, i_2 are reals such that both $|r(S;9\delta)| = x_1 \in [9\delta, 12\delta]$, and the processes $W(T_1;3\delta), \ldots, W(T_{j-1};3\delta)$ are not successful, but the process $W(T_j;3\delta)$ is successful for some fixed j, $1 \leq j \leq 4$, and

$$|r(T_j;3\delta)| = x_2 \in [3\delta, 6\delta].$$

Moreover, for the process $W(S;9\delta)$ it took i_1 terms to be successful, and for the process $W(T_j;3\delta)$ it took i_2 terms to be successful. Note that the value of $W(S;9\delta)$ is $\pm x_1$ with equal probabilities, and the value of $W(T_j;3\delta)$ is $\pm x_2$ with equal probabilities, independently both of each other and of all the other information.

Since $a_4 \geq 21\delta$, we can apply Observation 3.3 with a_4, x_1, x_2 to conclude that

$$\Pr\left[|\sum_{i=4}^{n} a_i\varepsilon_i| < 1 - a_1 - a_2 - a_3 | C_1, x_1, i_1, j, x_2, i_2, r(T_1;3\delta), \ldots, r(T_{j-1};3\delta)\right] \leq \frac{1}{8}. \tag{21}$$

As $x_1, i_1, j, x_2, i_2, r(T_1;3\delta), \ldots, r(T_{j-1};3\delta)$ were arbitrary and we have finitely many possibilities for them, we conclude from (21) that

$$\Pr\left[|\sum_{i=4}^{n} a_i\varepsilon_i| < 1 - a_1 - a_2 - a_3 | C_1\right] \leq \frac{1}{8}. \tag{22}$$

We can furthermore estimate the probability $\Pr\left[|Y| \geq 1 - a_1 - a_2 - a_3\right]$, conditioned on C_2, using Observation 3.2:

$$\Pr\left[|\sum_{i=4}^{n} a_i\varepsilon_i| < 1 - a_1 - a_2 - a_3 | C_2\right] \leq \frac{1}{4}. \tag{23}$$

Analogously, the probability $\Pr\left[|Y| \geq 1 - a_1 - a_2 - a_3\right]$ conditioned on C_3, is significant, as shown by Observation 3.1:

$$\Pr\left[|\sum_{i=4}^{n} a_i\varepsilon_i| < 1 - a_1 - a_2 - a_3 | C_3\right] \leq \frac{1}{2}. \tag{24}$$

Combining (20), (22), (23) and (24), we get $q_1 \geq \frac{205}{256} > \frac{793}{1024}$, and hence $\mathbf{a} \in \mathcal{A}_2$. ☐

We turn to investigating the case $\sum_{i=k}^{n} a_i^2 < 450\delta^2$. We record a property that will repeatedly be used in the sequel

$$\begin{aligned} a_1^2 + a_2^2 + a_3^2 &= (a_1 + a_2 + a_3 - 2a_3)^2 + 2a_3^2 - 2(a_1 - a_3)(a_2 - a_3) \\ &\leq (\frac{1}{3} + 2\delta)^2 + 2(\frac{1}{3} - \delta)^2 = \frac{1}{3} + 6\delta^2. \end{aligned} \tag{25}$$

Claim 4.3. *If $\sum_{i=k}^{n} a_i^2 < 450\delta^2$, then $k \geq 11$.*

Proof of Claim 4.3. Assume that we had $\sum_{i=k}^{n} a_i^2 < 450\delta^2$ and $k \leq 10$. Then using (25), we get

$$1 = \sum_{i=1}^{n} a_i^2 < (\frac{1}{3} + 2\delta)^2 + 8(\frac{1}{3} - \delta)^2 + 450\delta^2 = 1 + 462\delta^2 - 4\delta,$$

being a contradiction, as $1 + 462\delta^2 - 4\delta < 1$ for $\delta \in (0, 1/120]$. ☐

Claim 4.4. *If $k \geq 11$ and $a_8 + a_9 + a_{10} \geq \frac{2}{3} + \delta$, then $\mathbf{a} \in \mathcal{A}_3$.*

Proof of Claim 4.4. Assume that we had $a_8 + a_9 + a_{10} \geq \frac{2}{3} + \delta$ and $k \geq 11$. Then by (19) and Observation 3.1 applied to $a_4, \ldots, a_{10}$, we obtain $q_2, q_3 \geq \frac{37}{128}$. ☐

Claim 4.5. *If $\sum_{i=k}^{n} a_i^2 < 450\delta^2$ and $a_5 - a_{10} \geq 3\delta$, then $\mathbf{a} \in \mathcal{A}_2$.*

Proof of Claim 4.5. Consider the events D_1, D_2, where

$$D_1 = \{\varepsilon_4 = \varepsilon_6 = \varepsilon_7\}$$

and $D_2 = D_1^C$. Note that

$$\Pr\left[D_1\right] = \frac{1}{4} \qquad \Pr\left[D_2\right] = \frac{3}{4}. \tag{26}$$

In the case when D_1 occurs, let $c_1 = a_4 + a_6 + a_7$, $c_2 = a_5$, $c_3 = a_{10}$. Since the conditions of Observation 3.3 hold for c_1, c_2, c_3 (by Claim 4.3, $a_{10} \geq 1 - a_1 - a_2 - a_3$), we deduce that

$$\Pr\big[|\sum_{i=4}^{n} a_i\varepsilon_i| \leq 1 - a_1 - a_2 - a_3 \,|\, D_1\big] \leq \frac{1}{8}. \tag{27}$$

In the case when D_2 occurs, Observation 3.2 applied on $b_1 = a_5$, $b_2 = a_{10}$ implies that

$$\Pr\big[|\sum_{i=4}^{n} a_i\varepsilon_i| \leq 1 - a_1 - a_2 - a_3 \,|\, D_2\big] \leq \frac{1}{4}. \tag{28}$$

Combining (26), (27) and (28), we get

$$q_1 \geq \frac{25}{32} \geq \frac{793}{1024}.$$

□

Claim 4.6. *If $\sum_{i=k}^{n} a_i^2 < 450\delta^2$, $a_5 - a_{10} < 3\delta$ and $a_8 + a_9 + a_{10} < \frac{2}{3} + \delta$, then $k \geq 15$.*

Proof of Claim 4.6. Assume that all of the conditions above hold, yet $k \leq 14$. We clearly have

$$a_4 \leq a_3 = \frac{1}{3} - \delta, \tag{29}$$

and the combination of $a_5 - a_{10} < 3\delta$ and $a_8 + a_9 + a_{10} < \frac{2}{3} + \delta$ gives

$$a_{10}, \ldots, a_{13} \leq \frac{2}{9} + \frac{\delta}{3} \qquad a_5, \ldots, a_9 < \frac{2}{9} + \frac{10}{3}\delta. \tag{30}$$

Using $\sum_{i=k}^{n} a_i^2 < 450\delta^2$, (25), (29) and (30), we get

$$\begin{aligned}
1 &= \sum_{i=1}^{n} a_i^2 \\
&= \sum_{i=1}^{3} a_i^2 + a_4^2 + \sum_{i=5}^{9} a_i^2 + \sum_{i=10}^{k-1} a_i^2 + \sum_{i=k}^{n} a_i^2 \\
&\leq \Big(\frac{1}{3} + 6\delta^2\Big) + \Big(\frac{1}{3} - \delta\Big)^2 + 5\Big(\frac{2}{9} + \frac{10}{3}\delta\Big)^2 + 4\Big(\frac{2}{9} + \frac{\delta}{3}\Big)^2 + 450\delta^2 \\
&= \frac{8}{9} + 513\delta^2 + \frac{22}{3}\delta,
\end{aligned}$$

being a contradiction, as the ultimate expression is strictly smaller than 1 for any $\delta \in (0, 1/120]$. Hence $k \geq 15$. $\square$

Claim 4.7. *If $k \geq 15$, then $\mathbf{a} \in \mathcal{A}_2$.*

Proof of Claim 4.7. Applying Observation 3.1 with $a_4, \ldots, a_{14}$ gives $q_1 \geq \frac{793}{1024}$, as required. $\square$

The combination of the above claims concludes the proof of Lemma 4.1. $\square$

We are now ready to complete the proof of (14) in the case $a_3 \geq 0.325$ and $a_1 + a_2 + a_3 \leq 1$; that is, we verify (17).

We note that combining $\delta \leq \frac{1}{120}$ with (25), we get

$$a_1^2 + a_2^2 + a_3^2 \leq \frac{801}{2400}. \tag{31}$$

First, consider the family $\mathcal{A}_1$ with $a_4 \leq \frac{7}{40}$. In this case, (31) implies

$$\frac{a_4}{\sqrt{\sum_{i=4}^n a_i^2}} \leq \frac{\frac{7}{40}}{\sqrt{\frac{1599}{2400}}} < 0.216. \tag{32}$$

Moreover, using (18):

$$\frac{1 - a_1 - a_2 - a_3}{\sqrt{\sum_{i=4}^n a_i^2}} \leq \frac{\frac{3}{120}}{\sqrt{\frac{1599}{2400}}} < 0.032. \tag{33}$$

Finally, (32) and (33) imply

$$q_1 \geq 2D(0.216, 0.032). \tag{34}$$

Analogously to (34), we have

$$\begin{aligned} q_2 &\geq 2D(0.216, 0.828), & q_3 &\geq 2D(0.216, 0.828), \\ q_4 &\geq 2D(0.216, 0.858), & q_5 &\geq 2D(0.216, 1.634), \\ q_6 &\geq 2D(0.216, 1.654), & q_7 &\geq 2D(0.216, 1.654), \end{aligned}$$
$$q_8 \geq 2D(0.216, 2.452).$$

Using the following estimate,

$$D(0.216, 0.032) + D(0.216, 0.828) + D(0.216, 0.828) + D(0.216, 0.858) + D(0.216, 1.634) + D(0.216, 1.654) + D(0.216, 1.654) + D(0.216, 2.452) \geq \frac{3}{4}$$

we deduce (17) for any $\mathbf{a} \in \mathcal{A}_1$.

Next, we consider an $\mathbf{a}$ in the families $\mathcal{A}_2, \mathcal{A}_3$. Using $a_4 \leq \frac{1}{3}$ and (31), we obtain

$$\frac{a_4}{\sqrt{\sum_{i=4}^{n} a_i^2}} \leq \frac{\frac{1}{3}}{\sqrt{\frac{1599}{2400}}} < 0.41, \tag{35}$$

and we note that (33) still holds. Using (33) and (35), we obtain

$$q_1 \geq 2D(0.41, 0.032). \tag{36}$$

Analogously to (36), we derive

$$\begin{aligned} q_2 &\geq 2D(0.41, 0.828), & q_3 &\geq 2D(0.41, 0.828), \\ q_4 &\geq 2D(0.41, 0.858), & q_5 &\geq 2D(0.41, 1.634), \\ q_6 &\geq 2D(0.41, 1.654), & q_7 &\geq 2D(0.41, 1.654), \end{aligned}$$
$$q_8 \geq 2D(0.41, 2.452).$$

Note that we only mention the bound for q_8 above for the sake of completeness, since we have $D(0.41, 2.452) = 0$.

When $\mathbf{a} \in \mathcal{A}_2$ we can easily verify that

$$\frac{793}{2048} + D(0.41, 0.828) + D(0.41, 0.828) + D(0.41, 0.858) + D(0.41, 1.634) + \\ D(0.41, 1.654) + D(0.41, 1.654) + D(0.41, 2.452) \geq \frac{3}{4}$$

and hence (17) follows for all $\mathbf{a} \in \mathcal{A}_2$. For the family $\mathcal{A}_3$ we can verify that

$$D(0.41, 0.032) + \frac{37}{256} + \frac{37}{256} + D(0.41, 0.858) + D(0.41, 1.634) + \\ D(0.41, 1.654) + D(0.41, 1.654) + D(0.41, 2.452) \geq \frac{3}{4}$$

and hence (17) follows for all $\mathbf{a} \in \mathcal{A}_3$. Proof of this subcase is thus finished.

4.2 Case $a_1 + a_2 + a_3 > 1$

4.2.1 Subcase $a_1 + a_2 \geq 1$

Using Observation 3.1, we have $\Pr[|X| \geq 1] \geq 1/4$.

4.2.2 Subcase $a_1 \geq 0.7$ and not previous subcase

The proof is the same as in Section 4.1.2.

4.2.3 Setting for the rest of the subcases

Assume $a_1 + a_2 < 1$ and $a_1 + a_2 + a_3 > 1$. The required inequality (14), involves $\Pr[|X| \geq 1]$, and may be re-written using elimination in terms of $Y = \sum_{i=4}^{n} a_i \varepsilon_i$ as

$$\frac{2}{8}\Pr[|Y| \leq a_1 + a_2 + a_3 - 1] + \frac{1}{8}\Pr[|Y| > a_1 + a_2 + a_3 - 1] + \frac{1}{8}\Pr[|Y| \geq 1 - a_1 - a_2 + a_3] + \frac{1}{8}\Pr[|Y| \geq 1 - a_1 + a_2 - a_3] + \frac{1}{8}\Pr[|Y| \geq 1 + a_1 - a_2 - a_3] + \frac{1}{8}\Pr[|Y| \geq 1 - a_1 + a_2 + a_3] + \frac{1}{8}\Pr[|Y| \geq 1 + a_1 - a_2 + a_3] + \frac{1}{8}\Pr[|Y| \geq 1 + a_1 + a_2 - a_3] + \frac{1}{8}\Pr[|Y| \geq 1 + a_1 + a_2 + a_3] \geq 3/16.$$

Denote

$$L_1, L_2, L_3, L_4 = a_1+a_2+a_3-1, 1-a_1-a_2+a_3, 1-a_1+a_2-a_3, 1+a_1-a_2-a_3.$$

The inequality we are proving follows by rearranging and multiplying the following inequality by 1/8:

$$\Pr[|Y| \in (L_1, L_2)] \leq 1/2 + \Pr[|Y| \geq L_3] + \Pr[|Y| \geq L_4]. \tag{37}$$

Write $\sigma_j^2 = 1 - \sum_{i=1}^{j} a_i^2$. Recall the variance of Y is σ_3^2 and its largest weight is a_4.

4.2.4 Subcase $a_4 \geq 1 - a_1 - a_3$ and (either $a_4 \notin (L_1, L_2)$ or $\max(L_2 - a_4, a_4 - L_1) \leq 0.35\sigma_4$) and not previous subcases

Let us prove (37), i.e. $\Pr[|Y| \in (L_1, L_2)] \leq 1/2 + \Pr[|Y| \geq L_3] + \Pr[|Y| \geq L_4]$.

Since this inequality is symmetric with respect to Y, we may assume without loss of generality that $\varepsilon_4 = 1$, in which case it is clearly sufficient to prove

$$\Pr[Y \in (L_1, L_2) | \varepsilon_4 = 1] + \Pr[Y \in (-L_2, -L_1) | \varepsilon_4 = 1] \leq 1/2 + \Pr[Y > L_3 | \varepsilon_4 = 1].$$

To this end, note that $\Pr[Y \in (-L_2, -L_1) | \varepsilon_4 = 1] \leq \Pr[Y > L_3 | \varepsilon_4 = 1]$, which follows by (recall $L_3 - a_4 \leq a_4 + L_1$ by assumption):

$$\Pr[Y' + a_4 < -L_1] = \Pr[Y' > L_1 + a_4] \leq \Pr[Y' > L_3 - a_4]$$

with $Y' = Y - a_4\varepsilon_4$.

Hence our task is to verify $\Pr[Y \in (L_1, L_2)|\varepsilon_4 = 1] \le 1/2$. There are two subcases. If $a_4 \le L_1$ or $a_4 \ge L_2$, then we conclude with a general $\Pr[Y' > 0] \le 1/2$ bound. If $a_4 \in [L_1, L_2]$, we conclude with the inequality from Section 3.5, recalling that $\max(L_2 - a_4, a_4 - L_1) \le 0.35\sigma_4$.

4.2.5 Subcase not previous cases

Note that (37) follows from

$$\Pr[|Y| > L_1] \le 1/2 + \Pr[|Y| \ge L_2] + \Pr[|Y| \ge L_3] + \Pr[|Y| \ge L_4].$$

As the left hand side is a probability, it is sufficient we show the right hand side is at least 1. This in turn follows from (see Appendix A.2)

$$D(a_4/\sigma_3, L_2/\sigma_3) + D(a_4/\sigma_3, L_3/\sigma_3) + D(a_4/\sigma_3, L_4/\sigma_3) \ge 1/4. \tag{38}$$

5 Proof of $\Pr[X > 1] \ge 1/16$ unless $X = \varepsilon_1$

In this section, we prove Theorem 1.2 (which is the best possible). Note that unlike for Theorem 1.1 where significant further work was required, most of the work toward proving Theorem 1.2 was done when we developed our tools in 3.3 and now we can just conclude pretty easily.

5.1 Case $a_1 + a_2 + a_3 > 1$

Clearly,

$$\Pr[X > 1] \ge \Pr\left[\sum_{i=1}^{3} a_i\varepsilon_i > 1 \wedge \sum_{j=4}^{n} a_j\varepsilon_j \ge 0\right] \ge 1/8 \cdot 1/2 = 1/16.$$

5.2 Case $a_1 + a_2 + a_3 \le 1$

In this case we actually show $\Pr[X > 1] \ge 1/12$, and the proof is analogous to that of Section 4.1.1.

5.2.1 Subcase $a_1 \le 0.4$

We conclude using (12) and (13) with $D(0.4, 1) > \frac{1}{12}$.

5.2.2 Subcase $a_1 \geq 0.6$

Let $a = 0.6$. We conclude using elimination, (12) and (13) with

$$D\left(\frac{1-a}{\sqrt{1-a^2}}, \frac{1-a}{\sqrt{1-a^2}}\right) = D(1/2, 1/2) > 1/6.$$

Notice that in this case a_1 might be 1, which forbids elimination by a_1. This is where the assumption $X \neq \varepsilon_1$ is used.

5.2.3 Subcase $a_1 \in [0.4, 0.6]$

We write $\sigma_2 = \sqrt{1 - a_1^2 - a_2^2}$ and recall that a_3 is upper bounded by $a = \min(a_2, 1 - a_1 - a_2)$, so that $\Pr[X > 1] \geq 1/12$ follows from (see Appendix A.3):

$$\mathop{\mathbb{E}}_{\varepsilon \in \{-1,1\}^2}\left[D\left(\frac{a}{\sigma_2}, \frac{1 + a_1\varepsilon_1 + a_2\varepsilon_2}{\sigma_2}\right)\right] \geq 1/12. \tag{39}$$

6 Toward the $7/64$ bound

We strongly believe that $C_1 = 7/64$. Further to the brief discussion in subsection 1.4, we will comment in this section what the next steps would be and what hurdles one would face if we try to continue further to this bound using the methods of this paper, i.e. combining lower bounds of the type 3.3 with separate arguments for some difficult cases. While somewhat tedious, we note that similar approach was recently used by Keller and the second author to resolve the problem of Tomaszewski [9]. Nevertheless, the tools needed here would be rather different than the ones used in the proof of Tomaszewski's conjecture, since we are now dealing with an anti concentration inequality instead of a concentration one.

Continuing further to the 7/64 bound using our methods (or similar ones), there are two particular classes of the collections $\{a_i\}$ one has to be very careful about.

First such class are the collections $\{a_i\}$ for which we have precisely $\Pr\left[X \geq 1\right] = \frac{7}{64}$ and thus we can not afford to obtain any suboptimal bound. As an example of the collection in the first class, one can consider $a_1 = \ldots = a_6 = \frac{1}{\sqrt{6}}$. For this particular collection, the bound follows trivially from Observation 3.1, since $a_3 + a_4 + a_5 \geq 1$. We suspect that in fact all the collections in this class satisfy $a_3 + a_4 + a_5 \geq 1$, making it not too difficult to handle.

Second such class are the collections $\{a_i\}$ with

$$\Pr\left[X > 1\right] < \frac{7}{64},$$

since for these one can't verify the conjecture by only assuming that the few largest weights lie in some, however narrow, ranges. Five examples of the collections in the second class are mentioned in the subsection 1.4 and we believe these are only such examples.

The collections 'close to' $a_1 = 1$ are not a big problem for us, since Lemma 3.4 allows us to show that the bound of 7/64 holds for collections with a_1 large.

Proposition 6.1. *If* $a_1 \geq \frac{14}{15}$*, then* $\Pr\left[X \geq 1\right] \geq \frac{7}{64}$.

Proof. Note that it is enough to argue that $p(a_2, \ldots, a_n; 1 - a_1) \geq \frac{7}{8}$. For that, by Lemma 3.4 we know that it suffices if $1 - a_1^2 \geq 29(1 - a_1)^2$. We can easily check that this is satisfied whenever $a_1 \geq \frac{14}{15}$. □

'Neighbourhoods' of the remaining problematic collections are more difficult (though luckily note that the family $F_1(\delta)$ below covers the 'neighbourhood' of both the second and the fifth collection). For fixed $\delta > 0$, consider the families

$$F_1(\delta) = \{a_1 + a_2 < 1; a_2 \geq \frac{1}{2} - \delta\},$$

$$F_2(\delta) = \{a_1 + a_2 < 1; |a_1 - \frac{2}{3}|, |a_2 - \frac{1}{3}| \leq \delta\},$$

$$F_3(\delta) = \{a_1 + a_2 + a_3 < 1; a_3 \geq \frac{1}{3} - \delta\}.$$

If we want to verify that $C_1 = 7/64$ with the help of computational methods similar to the ones used in this paper, we must be able to find some $\delta > 0$ for which we can verify by different means that the conjecture holds for all the collections in $F_1(\delta), F_2(\delta), F_3(\delta)$. Hope is this could be done in somewhat similar way as the proof of 6/64 bound within $F_3(\frac{1}{120})$ in 4.1.4 when proving Theorem 1.1.

We make a progress in that direction by using stopped random walks and chain arguments to prove the following.

Proposition 6.2. *For* $\delta_0 = 10^{-9}$*, we have* $\Pr\left[X \geq 1\right] \geq \frac{7}{64}$ *for all collections* $\{a_i\}$ *in* $F_1(\delta_0), F_2(\delta_0)$.

Our value δ_0 is extremely small, but that is because we have not tried to optimize it at all (as that would result in an even more tedious argument). We believe with some effort, our solution could be improved to work for much larger value of δ which could actually be used in practice.

The arguments for $F_1(\delta_0)$ and $F_2(\delta_0)$ are rather similar in style and are somewhat tedious. Hence in this section, we only include the argument for the family $F_1(\delta_0)$ and the argument for the family $F_2(\delta_0)$ is placed in Appendix B.

Surprisingly, while we were able to improve the bound closer to $\frac{7}{64}$ in that case too, we were not able to prove the bound of 7/64 for the family $F_3(\delta)$ for any $\delta > 0$, so we pose this as an open problem to the reader. We believe even verifying the conjecture just in this narrow range of parameters would be of interest.

In subsection 6.1 and in Appendix B, we sometimes sketch the proofs instead of going through all the details of the calculations. That is because the calculations would otherwise be very long and it is easy to see that the sketch could indeed be turned into a rigorous proof.

6.1 Family $F_1(\delta_0)$

In this subsection, we prove the following result.

Proposition 6.3. *For $\delta_0 = 10^{-9}$, we have $\Pr\left[X \geq 1\right] \geq \frac{7}{64}$ for all collections $\{a_i\}$ in $F_1(\delta_0)$.*

Together with Proposition B.1, this implies Proposition 6.2.

Assume $a_1 + a_2 < 1$ and $a_2 = \frac{1}{2} - \delta$ for some $\delta \leq 10^{-9}$. Also assume our collection $\{a_1, \ldots, a_n\}$ has $\Pr\left[|X| \geq 1\right] < \frac{7}{32}$. We will derive a contradiction.

Note that $1 - a_1 - a_2 \leq 2\delta$. Denote $Y = \sum_{i=3}^{n} a_i \varepsilon_i$ and

$$p_1 = \Pr\left[|Y| \geq 1 - a_1 - a_2\right], \qquad p_2 = \Pr\left[|Y| \geq 1 - a_1 + a_2\right], \qquad p_3 = \Pr\left[|Y| \geq 1 + a_1 - a_2\right].$$

Then, in particular, we have

$$\Pr\left[|X| \geq 1\right] \geq \frac{1}{4}(p_1 + p_2 + p_3).$$

So, it is enough to show

$$p_1 + p_2 + p_3 \geq \frac{7}{8}. \tag{40}$$

We can also assume that

$$a_3 + a_4 + a_5 < 1, \tag{41}$$

else we would be done by Observation 3.1. We will make consecutive claims about $\{a_1, \ldots, a_n\}$, characterizing it more and more precisely until we are ready to obtain a contradiction.

Call a_i big if $a_i \geq 1 - a_1 - a_2$, and small otherwise. So in particular if $a_i \geq 2\delta$, it must be big. Let k be the smallest integer such that $a_k < 1 - a_1 - a_2$ (if $a_n \geq 1 - a_1 - a_2$, set $k = n + 1$).

Claim 6.4. *Let k be the smallest integer such that $a_k < 1 - a_1 - a_2$. Then we have $\sum_{i=k}^{n} a_i^2 \leq 240000\delta^2 < \frac{\delta}{1000}$.*

Proof. Assume for contradiction that this is not true. Then we can take disjoint subsets $S_1, \ldots, S_{10000}$ of $\{a_k, \ldots, a_n\}$ with

$$24\delta^2 \geq \sum_{i \in S_j} a_i^2 \geq 20\delta^2$$

for $j = 1, \ldots 10000$. Now considering the random processes $W(S_i; 2\delta)$ for $i = 1, \ldots, 10000$, with probability at least $\frac{99}{100}$, at least 1000 of these are successful, and conditional on that, we obtain $\Pr\left[|Y| \geq 1 - a_1 - a_2\right] \geq \frac{19}{20}$ by Observation 3.1. Hence we overall get

$$p_1 \geq \frac{1881}{2000} > \frac{7}{8},$$

and (40) holds. □

Claim 6.5. *a_5 and a_6 are big terms, that is, $a_6 \geq 1 - a_1 - a_2$.*

Proof. Assume for contradiction that a_6 is a small term (i.e. that $k \leq 6$). Combining Claim 6.4 with (41), we arrive at a contradiction for all sufficiently small $\delta > 0$:

$$1 = \sum_{i=1}^{5} a_i^2 + \sum_{i=6}^{n} a_i^2 \leq a_1^2 + a_2 \sum_{i=2}^{5} a_i + \frac{\delta}{1000} \leq (1/2+\delta)^2 + (1/2-\delta)(3/2-\delta) + \frac{\delta}{1000} = 1 - \frac{999}{1000}\delta + 2\delta^2.$$

□

At this point, we split our proof into two cases, the uniform and the non-uniform one, both of which we handle separately.

6.1.1 The uniform case - $a_3 - a_{k-1} \le 20\delta$

Claim 6.6. *Let k be the smallest integer such that $a_k < 1 - a_1 - a_2$. Then we have $k \le 11$.*

Proof. Assume we had $k \ge 12$. Note that $a_3 - a_{k-1} \le 20\delta$ would then in particular imply

$$a_3 \le \sqrt{0.5/9} + O(\delta) < 0.24, \tag{42}$$

and we also know

$$\sum_{i=3}^{n} a_i^2 \ge 0.4999 \tag{43}$$

and

$$1 + a_1 - a_2 \le 1.00001. \tag{44}$$

Using Observation 3.1 for $a_3, \dots, a_{11}$, we get

$$p_1 \ge \frac{386}{512} > \frac{3}{4}.$$

Combining (42), (43) and (44), and using (13), we get

$$p_2, p_3 \ge 2D(0.34, 1.42) > 0.08 > \frac{1}{16}$$

and hence (40) holds. □

The next corollary follows by combining Chebyshev inequality with Claim 6.4, using that δ is small.

Corollary 6.7. *Let k be the smallest integer such that $a_k < 1 - a_1 - a_2$. Then we have* $\Pr\left[|\sum_{i=k}^{n} a_i\varepsilon_i| \ge 0.0001\right] < \frac{1}{1000}$.

We now sketch how we finish our argument in the subcase $a_3 - a_{k-1} \le 20\delta$, using Corollary 6.7. We consider five separate cases depending on the particular value of k which we know is at least 7 and at most 11 (and in fact, we can rule out the case $k = 7$ as then we would have $a_3 + a_4 + a_5 \ge 1$). Due to our restrictions on the value of δ and Corollary 6.7, we know that $\sum_{i=3}^{n} a_i\varepsilon_i$ behaves 'essentially' like $\sum_{i=3}^{k-1} \varepsilon_i \frac{1}{\sqrt{2k-6}}$. So for instance in the case $k = 8$, we argue that $p_1 \ge \frac{999}{1000}$, as due to our restrictions on $a_3, \dots, a_7$, we know we can only have $|\sum_{i=3}^n a_i\varepsilon_i| < 2\delta$ if $|\sum_{i=8}^n a_i\varepsilon_i| \geq 0.0001$; further, in this case $k=8$, we analogously argue that $p_2, p_3 \geq \frac{999}{1000} \cdot \frac{1}{32}$.

Similarly, in the case $k=9$, we argue that $p_1 \geq \frac{11}{16}$, $p_2, p_3 \geq \frac{999}{1000} \cdot \frac{7}{32}$.

The reader can easily verify that such arguments indeed work in all the cases considered. □

6.1.2 The non-uniform case - $a_3 - a_{k-1} > 20\delta$

In this case, we first notice that Observation 3.2 applied to a_3, a_{k-1} immediately implies the following.

Claim 6.8. *We have $p_1 \geq \frac{3}{4}$.*

Next we obtain.

Claim 6.9. *We have $a_3 + a_4 + a_5 + a_6 < 1 + 2\delta$.*

Proof. Assume not. Then by Observation 3.1, we have $p_2, p_3 \geq \frac{1}{16}$, and combining this with Claim 6.8 gives (40). □

Claim 6.10. *a_7 is a big term.*

Proof. Assume for contradiction that a_7 is a small term (i.e. that $k \leq 7$), and recall that $a_3 + a_4 + a_5 + a_6 < 1 + 2\delta$ and $\sum_{i=7}^n a_i^2 < 240000\delta^2 < \frac{\delta}{1000}$. Write $A = a_3 + a_4$, and arrive at a contradiction for all sufficiently small $\delta > 0$:

$$\begin{aligned}
1 = \sum_{i=1}^{6} a_i^2 + \sum_{i=7}^{n} a_i^2 &\leq a_1^2 + a_2^2 + a_2(a_3 + a_4) + a_5(a_5 + a_6) + \frac{\delta}{1000} \\
&\leq \Big(\frac{1}{2} + \delta\Big)^2 + \Big(\frac{1}{2} - \delta\Big)^2 + \Big(\frac{1}{2} - \delta\Big)A + \frac{1 + 2\delta}{3}(1 + 2\delta - A) + \frac{\delta}{1000} \\
&= \frac{1}{2} + 2\delta^2 + A\Big(\frac{1}{6} - \frac{5}{3}\delta\Big) + \frac{1}{3}(1 + 2\delta)^2 + \frac{\delta}{1000} \\
&\leq \frac{1}{2} + 2\delta^2 + (1 - 2\delta)\Big(\frac{1}{6} - \frac{5}{3}\delta\Big) + \frac{1}{3}(1 + 2\delta)^2 + \frac{\delta}{1000} \\
&= 1 - \frac{2}{3}\delta + \frac{20}{3}\delta^2 + \frac{\delta}{1000} < 1.
\end{aligned}$$

where we used the estimates $a_5 + a_6 \leq 1 + 2\delta - A$, and $a_5 \leq (1 + 2\delta)/3$ and $A \leq 2a_2 \leq 1 - 2\delta$.

□

Claim 6.11. *We have $a_4 \geq 0.07$.*

Proof. If not, we can use Claim 6.4 to argue that we have at least 44 big terms, otherwise we would have

$$\sum_{i=3}^{k-1} a_i^2 < 0.49.$$

But Observation 3.1 then implies $p_1 \geq \frac{7}{8}$, and (40) follows. ◻

Claim 6.12. *We have $p_2, p_3 \geq \frac{3}{64}$.*

Proof. We consider two cases. If $a_3 \leq 0.3$, the result follows using the bounds (43) and (44) as well as (13) by

$$p_2, p_3 \geq 2D(0.43, 1.42) > 0.06 > \frac{3}{64}.$$

If on the other hand $a_3 > 0.3$, we may argue (using Claim 6.11 and argument much along the same lines as the proofs of Claim 6.5 and Claim 6.10) that

$$a_3 + a_4 + \sqrt{\sum_{i=5}^{n} a_i^2} \geq 1 + a_1 - a_2.$$

But then let ε' be a sign of $\sum_{i=5}^{n} a_i \varepsilon_i$, and consider the events

$$A = \{\varepsilon \colon \varepsilon_3 = \varepsilon_4 = \varepsilon'\}, \qquad B = \left\{\varepsilon \colon |\sum_{i=5}^{n} a_i \varepsilon_i| \geq \Big(\sum_{i=5}^{n} a_i^2\Big)^{1/2}\right\}.$$

We have $\Pr\big[A \cap B\big] \geq \frac{3}{64}$ (using our bound from the previous sections), and clearly

$$|\sum_{i=3}^{n} a_i \varepsilon_i| \geq 1 + a_1 - a_2$$

whenever event $A \cap B$ occurs. The result follows. ◻

Claim 6.13. *Let k be the smallest integer such that $a_k < 1 - a_1 - a_2$. Then we have $a_4 - a_{k-1} \leq 2\delta$.*

Proof. Assume for contradiction that $a_4 - a_{k-1} \geq 2\delta$. Then $a_3 + a_5 + a_6$ is not within 2δ neither from a_4 nor from $a_4 + a_{k-1}$. Using Observation 3.2 for a_4, a_{k-1} in the case when we do not have $\varepsilon_3 = \varepsilon_5 = \varepsilon_6$, and Observation

3.3 for $a_4, a_{k-1}, a_3 + a_5 + a_6$ in the case when we have $\varepsilon_3 = \varepsilon_5 = \varepsilon_6$ (which happens with probability $\frac{1}{4}$) gives

$$p_1 \geq \frac{3}{4} \cdot \frac{3}{4} + \frac{1}{4} \cdot \frac{7}{8} = \frac{25}{32}. \tag{45}$$

Combining Claim 6.12 with (45) gives (40). □

Now we are ready to reach the contradiction. First, if $a_3 \notin (2a_{k-1} - 8\delta, 2a_{k-1} + 8\delta)$, let $f_1 = a_3 + a_{k-1}$ and $f_2 = a_4 + a_5$. Let $A_1 = \{\varepsilon_3 = \varepsilon_{k-1}\}$ and $A_2 = \{\varepsilon_4 = \varepsilon_5\}$. Then conditional on $A_1 \cap A_2$, we have $\Pr\left[|Y| \geq 1 - a_1 - a_2\right] \geq \frac{7}{8}$ by Observation 3.3 for f_1, f_2, a_6; conditional on $A_1 \cap A_2^C$, we have $\Pr\left[|Y| \geq 1 - a_1 - a_2\right] \geq \frac{3}{4}$ by Observation 3.2 for $a_3 + a_{k-1}, a_6$; and conditional on A_1^C, we have $\Pr\left[|Y| \geq 1 - a_1 - a_2\right] \geq \frac{3}{4}$ by Observation 3.2 for $a_3 - a_{k-1}, a_6$. So we conclude $p_1 \geq \frac{25}{32}$, and hence (40) holds.

So next assume $a_3 \in (2a_{k-1} - 8\delta, 2a_{k-1} + 8\delta)$. Here, we observe that we can assume $k \leq 15$, else we could conclude $p_1 \geq \frac{25}{32}$ from Observation 3.1. But now, we proceed analogously to how we did at the end of the argument for the uniform case, again using Corollary 6.7 and detailed analysis of each of the several cases we have depending on the value of k. Carrying out such analysis is made possible by Claim 6.13.

So the proof of Proposition 6.3 is complete. □

7 The high-dimensional version of the problem

The following (non-tight) result constitutes a high-dimensional variant of Tomaszewski's problem as well as of the problem studied in this paper. The result is merely a consequence of the combination of [16, Proposition 2.2] and [8, Theorem 2]. Nevertheless, for the sake of completeness, we prove it here.

Proposition 7.1. *Let $v_1, \ldots, v_n \in \mathbb{R}^d$ be vectors with $\sum_i \|v_i\|_2^2 = 1$. The random variable $X = \sum v_i \varepsilon_i$ with $\varepsilon_i \sim \{-1, 1\}$ uniformly and independently distributed, satisfies*

$$\Pr[\|X\|_2 \geq 1] \geq \frac{1 - \sqrt{1 - 1/e^2}}{2} > 0.035, \qquad \Pr[\|X\|_2 \leq 1] \geq \frac{1 - \sqrt{1 - 1/e^2}}{2}.$$

Proof. The function $f(\varepsilon) = \|X(\varepsilon)\|_2^2 - 1 = \sum_{i,j} \varepsilon_i \varepsilon_j \langle v_i, v_j \rangle$ is a *homogenuous* polynomial of degree 2 in the ε_i's. We wish to lower bound the probabilities $\Pr[f(\varepsilon) \geq 0]$ and $\Pr[f(\varepsilon) \leq 0]$. Recall [8, Theorem 2]:

$$\|f\|_2 \leq e \|f\|_1 . \tag{46}$$

Since $\mathbb{E}[f] = 0$, we can derive (see below)

$$\|f\|_1^2 \leq 4 \Pr[f > 0] \Pr[f \leq 0] \|f\|_2^2 . \tag{47}$$

Plugging (46) into (47) we get

$$\|f\|_1^2 \leq 4e^2 \Pr[f > 0] \Pr[f \leq 0] \|f\|_1^2 .$$

When $f \equiv 0$, we have $\Pr[f = 0] = 1$. Otherwise, dividing by $\|f\|_1^2$ we obtain

$$\Pr[f > 0] \Pr[f \leq 0] \geq e^{-2}/4,$$

which means both $\Pr[f > 0]$ and $\Pr[f \leq 0]$ are at least $\frac{1-\sqrt{1-1/e^2}}{2}$, through $\Pr[f > 0] + \Pr[f \leq 0] = 1$.

To see (47), notice that by the Cauchy-Schwarz inequality,

$$\|f\|_2^2 = \mathbb{E}[f^2 \cdot \mathbb{1}\{f > 0\}] + \mathbb{E}[f^2 \cdot \mathbb{1}\{f \leq 0\}] \geq \Pr[f > 0] \, \mathbb{E}\left[f \mid f > 0\right]^2 + \Pr[f \leq 0] \, \mathbb{E}\left[f \mid f \leq 0\right]^2 . \tag{48}$$

As $\mathbb{E}[f] = 0$, we have

$$\mathbb{E}[f \cdot \mathbb{1}\{f > 0\}] = -\mathbb{E}[f \cdot \mathbb{1}\{f \leq 0\}] = \frac{1}{2} \|f\|_1 .$$

Likewise, we may assume $\Pr[f > 0]$ and $\Pr[f \leq 0]$ are both positive as otherwise (47) trivially holds. Under this assumption, (48) yields

$$\|f\|_2^2 \geq \frac{1}{4} \|f\|_1^2 \left(\frac{1}{\Pr[f > 0]} + \frac{1}{\Pr[f \leq 0]} \right),$$

being (47), using again $\Pr[f > 0] + \Pr[f \leq 0] = 1$. □

Denote by T_d the maximum constant for which $\Pr[\|X\|_2 \leq 1] \geq T_d$ for all X of dimension d as in Proposition 7.1, and denote by O_d the maximum constant for which $\Pr[\|X\|_2 \geq 1] \geq O_d$ for all X of dimension d as in Proposition 7.1. Clearly, T_d and O_d are non-increasing in d. We know $T_1 = \frac{1}{2}$ [9], while this paper proves that $\frac{6}{32} \leq O_1 \leq \frac{7}{32}$. Proposition 7.1 establishes that $T_d, O_d \geq 0.035$ for any d. There are two directions for further research here.

The first is to find tighter bounds for T_d, O_d for small values of $d > 1$. We know that $T_2 \leq \frac{1}{4}$, as demonstrated by

$$_1 = (\frac{1}{\sqrt{3}}, 0), \qquad v_2 = (-\frac{1}{2\sqrt{3}}, \frac{1}{2}), \qquad v_3 = (-\frac{1}{2\sqrt{3}}, -\frac{1}{2}),$$

and $T_3 \leq \frac{3}{16}$, as demonstrated by

$$v_1 = (\sqrt{\frac{7}{30}}, \frac{1}{3}, \frac{1}{5}), \quad v_2 = (\sqrt{\frac{7}{30}}, -\frac{1}{3}, -\frac{1}{5}), \quad v_3 = (0, \frac{1}{3}, -\frac{1}{5}), \quad v_4 = (0, 0, \frac{1}{5}), \quad v_5 = (0, 0, \frac{1}{5}).$$

Interestingly, we have not been able to find any examples demonstrating that $O_2 < \frac{7}{32}$ (or even that $O_{d_0} < \frac{7}{32}$ for any $d_0 \in \mathbb{N}$), and hence we pose this as a problem to a reader.

The second possible direction is to investigate how T_d, O_d behave for large d, and in particular to find better bounds for $\inf_d T_d$ and $\inf_d O_d$. It appears that Proposition 7.1 is far from being tight. Also, as just mentioned, it does not seem completely unthinkable that $O_d = \frac{7}{32}$ for every $d \in \mathbb{N}$ could hold.

8 Conclusion

As mentioned previously, we would hope that mixed with some new ideas, the methods developed in this paper could be used to prove the conjectured optimal bound of 7/64 in Theorem 1.1. That is the main open problem left, and even some progress toward that (like improving Theorem 1.1 to hold for some constant between 6/64 and 7/64) would be of interest.

Another, perhaps easier step one could take in this direction would be to prove the bound of 7/64 for the 'difficult' family $F_3(\delta_0)$ for some $\delta_0 > 0$. The significance of this is discussed in more detail in Section 6.

In a bit different direction, it is likely that one could improve the multiplicative factor in front of $\sqrt{\mathrm{Var}(X)}$ in Theorem 1.3 from 0.35 to the optimal conjectured [10] value of $1/\sqrt{7}$. That would not only be of interest on its own, but as demonstrated by this paper and our use of Theorem 1.3 when deriving Theorem 1.1, also a useful tool when attacking similar problems.

Finally, let us mention two interesting generalizations of our main problem that one can consider.

Firstly, same as Keller and the second author [9], we ask what is the behaviour of the function

$$F(x) = \sup_X \Pr[X > x],$$

where the supremum is taken over all the Rademacher sums with variance 1. Theorem 1.1 establishes that $F(-1) \leq \frac{58}{64}$. We know some asymptotic results about the behaviour of $F(x)$ [14] and we also know the precise value of $F(x)$ for some x [2, 9, 14], but much remains to be understood. It would be tempting to conjecture that $F(x) = F^{=}(x)$, where for $F^{=}(x)$, we take the supremum over all the the Rademacher sums with variance 1 and all the weights equal. Nevertheless, this conjecture turns out not to be true, see [13].

Further, one can also study the various multi-dimensional questions that arise, as discussed in Section 7. We find it especially intriguing that we have not managed to find any $d_0 \in \mathbb{N}$ for which we could show that $O_{d_0} < 7/32$. If there is no such d_0, that would be a beautiful generalization of the result of the one dimensional version of the problem.

Acknowledgements

We profoundly thank Béla Bollobás, Nathan Keller, Peter van Hintum, Marius Tiba and the anonymous referees for fruitful discussions and suggestions.

The first author was supported by EPSRC (grant no. 2260624). The second author was supported by the Clore Scholarship Programme, and by the Israel Science Foundation (grant no. 1612/17).

References

[1] A. Ben-Tal, A. Nemirovski, and C. Roos. Robust solutions of uncertain quadratic and conic-quadratic problems. *SIAM Journal on Optimization*, 13(2):535–560, 2002.

[2] V. K. Bentkus, D. Dzindzalieta, et al. A tight Gaussian bound for weighted sums of Rademacher random variables. *Bernoulli*, 21(2):1231–1237, 2015.

[3] D. Burkholder. Independent sequences with the Stein property. *The Annals of Mathematical Statistics*, 39(4):1282–1288, 1968.

[4] V. Dvořák and O. Klein. `https://github.com/IamPoosha/oleszkiewicz-problem/blob/main/`

[5] P. Erdös. On a lemma of Littlewood and Offord. *Bulletin of the American Mathematical Society*, 51(12):898–902, 1945.

[6] R. K. Guy. Any answers anent these analytical enigmas? *The American Mathematical Monthly*, 93(4):279–281, 1986.

[7] P. Hitczenko and S. Kwapień. On the Rademacher series. In *Probability in Banach Spaces, 9*, pages 31–36. Springer, 1994.

[8] P. Ivanisvili and T. Tkocz. Comparison of moments of Rademacher chaoses. *Arkiv för Matematik*, 57(1):121–128, 2019.

[9] N. Keller and O. Klein. Proof of Tomaszewski's conjecture on randomly signed sums. *arXiv preprint arXiv:2006.16834*, 2020.

[10] G. Lowther. `https://mathoverflow.net/questions/366894/a-rademacher-root-7-anti-con`

[11] S. J. Montgomery-Smith. The distribution of Rademacher sums. *Proceedings of the American Mathematical Society*, 109(2):517–522, 1990.

[12] K. Oleszkiewicz. On the Stein property of Rademacher sequences. *Probability and Mathematical Statistics-PWN*, 16(1):127–130, 1995.

[13] I. Pinelis. On the supremum of the tails of normalized sums of independent Rademacher random variables. *Statistics & Probability Letters*, 99:131–134, 2015.

[14] I. Pinelis et al. An asymptotically Gaussian bound on the Rademacher tails. *Electronic Journal of Probability*, 17, 2012.

[15] H. Prawitz. Limits for a distribution, if the characteristic function is given in a finite domain. *Scandinavian Actuarial Journal*, 1972(2):138–154, 1972.

[16] M. Veraar. A note on optimal probability lower bounds for centered random variables. In *Colloquium Mathematicum*, volume 113(2), pages 231–240. Institute of Mathematics Polish Academy of Sciences, 2008.

Appendix A Proofs of real numbers inequalities

A.1 Proof of (16)

We consider only these a_1, a_2 with $a_1 + a_2 \le 1$ and $a_2 \le a_1 \in [0.3, 0.7]$. We denote $a = \min(1 - a_1 - a_2, a_2, 0.325)$ (being an upper bound on a_3), and $\sigma_2 = \sqrt{1 - a_1^2 - a_2^2}$. We note that both a/σ_2 and $(1 + a_1\varepsilon_1 + a_2\varepsilon_2)/\sigma_2$ for any choice of $\varepsilon_1, \varepsilon_2 \in \{-1, 1\}^n$ are 10-Lipschitz in our domain (e.g., by checking that all partial derivatives $< \sqrt{50}$ in absolute value), so it suffices we check

$$\mathop{\mathbb{E}}_{\varepsilon \in \{-1,1\}^2} \left[D\left(\frac{a}{\sigma_2} + \delta, \frac{1 + a_1\varepsilon_1 + a_2\varepsilon_2}{\sigma_2} + \delta \right) \right] \ge 3/32 \tag{49}$$

on a mesh of $\{(a_1,a_2)\,|\,a_1+a_2\le 1, a_2\le a_1\in[0.3,0.7]\}$ of granularity $\delta/10$ in both axes. Inequality (49) can easily be verified [4] for $\delta=0.005$ on such a mesh.

A.2 Proof of (38)

In order to verify (38) for all relevant a_1,a_2,a_4,a_4, we confirm

$$D(a_4/\sigma_3+\delta, L_2/\sigma_3+\delta)+D(a_4/\sigma_3+\delta, L_3/\sigma_3+\delta)+D(a_4/\sigma_3+\delta, L_4/\sigma_3+\delta)\ge 1/4, \tag{50}$$

on a fine enough mesh of a_1,a_2,a_3 (which induce an upper bound on a_4). Notice the other subcases in the proof handle cases in which $a_4\ge 1-a_1-a_3$ and

$$L_2-L_1\le 0.35\sqrt{1-a_1^2-a_2^2-2a_3^2}. \tag{51}$$

All expressions L_i/σ_3 and a_3/σ_3 and $(1-a_1-a_3)/\sigma_3$ have partial derivatives <10, hence considering a mesh of $\{(a_1,a_2,a_3)\,|\,a_3\le a_2\le a_1\le 0.7, a_1+a_2+a_3\ge 1, a_1+a_2\le 1\}$, with granularity $\delta/15$ in every axis, we may verify (38) by checking (50) on the mesh points. One detail is that on the mesh points we bound a_4 by $1-a_1-a_3$ (instead of $\min(a_3,\sigma_3)$) only if $L_2-L_1+\delta/2<0.35\sqrt{1-a_1^2-a_2^2-2a_3^2}$, ensuring that if (51) is not satisfied for a point, then its nearest mesh point will not use the improved bound $a_4\le 1-a_1-a_3$ (introducing 'discontinuity'); this behavior is overridden to the points $(a_1,a_2,a_3)=(0.5\pm0.02, 0.5\pm0.02, 0.5\pm0.02)$, since there (51) is always satisfied. Choosing $\delta=0.03$, (50) can be verified [4] to all the described mesh points.

A.3 Proof of (39)

Instead of checking (39), we will check that

$$\mathop{\mathbb{E}}_{\varepsilon\in\{-1,1\}^2}\left[D\left(\frac{\min(a_2,1-a_1-a_2)}{\sigma_2}+\delta, \frac{1+a_1\varepsilon_1+a_2\varepsilon_2}{\sigma_2}+\delta\right)\right]\ge 1/12 \tag{52}$$

with $\sigma_2=\sqrt{1-a_1^2-a_2^2}$ is satisfied on a mesh of points in $\{(a_1,a_2)\,|\,a_1+a_2\le 1, a_2\le a_1\in[0.4,0.6]\}$. Since all the involved arguments fed to D are 10-Lipschitz, it suffices we verify (52) on a mesh with $\delta/10$ granularity in every axis. Verification [4] can be done with $\delta=0.01$.

Appendix B Family $F_2(\delta_0)$

In this appendix, we prove the following result, which together with Proposition 6.3, implies Proposition 6.2.

Proposition B.1. *For $\delta_0 = 10^{-9}$, we have $\Pr\left[X \geq 1\right] \geq \frac{7}{64}$ for all collections $\{a_i\}$ in $F_2(\delta_0)$.*

To prove Proposition B.1, take smallest possible $\delta > 0$ such that $a_1 \in \left[\frac{2}{3} - \delta, \frac{2}{3} + \delta\right]$ and $a_2 \in \left[\frac{1}{3} - \delta, \frac{1}{3} + \delta\right]$. Assume $\delta \leq 10^{-9}$. Assume our collection $\{a_1, \ldots, a_n\}$ has $\Pr\left[|X| \geq 1\right] < \frac{7}{32}$. We will derive a contradiction.

Note that

$$1 - a_1 - a_2 \leq 2\delta. \tag{53}$$

Denote

$$p_1 = \Pr\left[|\sum_{i=3}^{n} a_i\varepsilon_i| \geq 1 - a_1 - a_2\right], \quad p_2 = \Pr\left[|\sum_{i=3}^{n} a_i\varepsilon_i| \geq 1 - a_1 + a_2\right], \quad p_3 = \Pr\left[|\sum_{i=3}^{n} a_i\varepsilon_i| \geq 1 + a_1 - a_2\right].$$

Note that $\Pr\left[|X| \geq 1\right] \geq \frac{1}{4}(p_1 + p_2 + p_3)$, so it is enough to show that

$$p_1 + p_2 + p_3 \geq \frac{7}{8}. \tag{54}$$

We can assume

$$a_3 + a_4 + a_5 < 1, \tag{55}$$

else we would be done by Observation 3.1. We will make consecutive claims about the collection $\{a_1, \ldots, a_n\}$, characterizing it more and more precisely until we are ready to obtain a contradiction.

Note that for $\eta = 10^{-5}$, the following two lemmas hold.

Lemma B.2. *Assume $b_1 \geq \ldots \geq b_m > 0$, $\sum_{i=1}^{m} b_i^2 = 1$ and $b_1 \leq \frac{1}{2} + \eta$. Then*

$$\Pr\left[|\sum_{i=1}^{m} b_i\varepsilon_i| \geq 4\delta\right] \geq \frac{5}{8}.$$

Proof. Note that if $b_3 \geq 4\delta$, we are done by Observation 3.1. So we only need to consider the case when $b_3 < 4\delta$. First, we argue that we have

$$\sum_{i=3}^{m} b_i^2 \geq 960\delta^2. \tag{56}$$

Since we know that

$$\sum_{i=3}^{m} b_i^2 \geq 1 - 2(\frac{1}{2} + \eta)^2 = \frac{1}{2} - 2\eta - 2\eta^2,$$

to prove (56) holds, it is enough to show

$$960\delta^2 \leq \frac{1}{2} - 2\eta - 2\eta^2. \tag{57}$$

But (57) trivially holds as $\delta \leq 10^{-9}, \eta = 10^{-5}$.

Now using $4\delta > b_3, \ldots, b_m > 0$ and (56), we know that we can choose 10 disjoint subsets $S_1, \ldots, S_{10}$ of $\{b_3, \ldots, b_m\}$ such that for $1 \leq i \leq 10$, we have

$$96\delta^2 \geq \sum_{b_j \in S_i} b_j^2 \geq 80\delta^2.$$

Then for each of these sets S_i, we consider the random process $W(S_i; 4\delta)$. By Lemma 3.4, each of these is successful with probability at least $\frac{1}{2}$ and independently of the other ones. If for some t, $1 \leq t \leq 10$, we condition on the event E_t that precisely t of these processes are successful, Observation 3.1 ensures that

$$\Pr\left[|\sum_{i=1}^{m} b_i\varepsilon_i| < 4\delta | E_t\right] \leq \binom{t}{\lfloor t/2 \rfloor} 2^{-t}.$$

So we can bound

$$\Pr\left[|\sum_{i=1}^{m} b_i\varepsilon_i| < 4\delta\right] \leq 2^{-10} + 2^{-10}\sum_{t=1}^{10}\binom{10}{t}\binom{t}{\lfloor t/2 \rfloor} 2^{-t} \leq \frac{3}{8}.$$

This now finishes the proof of Lemma B.2. □

Lemma B.3. *Assume* $b_1 \geq \ldots \geq b_m > 0$, $\sum_{i=1}^m b_i^2 = 1$ *and* $b_1 \leq \frac{1}{2} + \eta$. *Then*

$$\Pr\left[|\sum_{i=1}^{m} b_i\varepsilon_i| \geq 1 + 4\delta\right] \geq \frac{1}{8}.$$

Proof. This follows directly using (13) by $D(0.51, 1.01) = \frac{1}{16}$. □

Now we can use these lemmas to prove the following corollary.

Corollary B.4. *We have* $p_1 \geq \frac{5}{8}$ *and* $p_2 \geq \frac{1}{8}$.

Proof. Note that $a_3 \leq a_2 \leq \frac{1}{3} + \delta$ and

$$\sum_{i=3}^{n} a_i^2 \geq 1 - (\frac{2}{3} + \delta)^2 - (\frac{1}{3} - \delta)^2 = \frac{4}{9} - \frac{2}{3}\delta - 2\delta^2.$$

So for $\delta \le 10^{-9}, \eta = 10^{-5}$, we see that we have

$$a_3 \le (\frac{1}{2} + \eta)\sqrt{\sum_{i=3}^{n} a_i^2}.$$

Thus we conclude from Lemma B.2 that

$$\begin{aligned} p_1 &= \Pr\big[|\sum_{i=3}^{n} a_i\varepsilon_i| \ge 1 - a_1 - a_2\big] \\ &\ge \Pr\big[|\sum_{i=3}^{n} a_i\varepsilon_i| \ge 2\delta\big] \\ &\ge \Pr\big[|\sum_{i=3}^{n} a_i\varepsilon_i| \ge 4\delta\sqrt{\sum_{i=3}^{n} a_i^2}\big] \\ &\ge \frac{5}{8}. \end{aligned}$$

Analogously, we conclude from Lemma B.3 that

$$\begin{aligned} p_2 &= \Pr\big[|\sum_{i=3}^{n} a_i\varepsilon_i| \ge 1 - a_1 + a_2\big] \\ &\ge \Pr\big[|\sum_{i=3}^{n} a_i\varepsilon_i| \ge \frac{2}{3} + 2\delta\big] \\ &\ge \Pr\big[|\sum_{i=3}^{n} a_i\varepsilon_i| \ge (1 + 4\delta)\sqrt{\sum_{i=3}^{n} a_i^2}\big] \\ &\ge \frac{1}{8}. \end{aligned}$$

□

Claim B.5. *We have $a_3 > 14\delta$.*

Proof. Assume we had $a_3 \le 14\delta$. By our choice of δ, we can trivially check that

$$\sum_{i=3}^{n} a_i^2 \ge 1 - (\frac{2}{3} + \delta)^2 - (\frac{1}{3} - \delta)^2 \ge 3920\delta^2. \tag{58}$$

Using (58), we can choose 20 disjoint subsets (possibly containing a single element) $S_1, \dots, S_{20}$ of $\{a_3, \dots, a_n\}$ such that for $1 \le i \le 20$, either S_i contains a single element $a_i \ge 2\delta$, or all its elements are smaller than 2δ and we have

$$24\delta^2 \ge \sum_{b_j \in S_i} b_j^2 \ge 20\delta^2.$$

Then for each of these sets S_i, consider the random process $W(S_i; 2\delta)$. By Lemma 3.4, each of these is successful with probability at least $\frac{1}{2}$ and independently of the other ones. If for some t, $1 \le t \le 20$, we condition on the event F_t that precisely t of these processes are successful, Observation 3.1 ensures that

$$\Pr\left[|\sum_{i=3}^{n} a_i \varepsilon_i| < 2\delta | F_t\right] \le \binom{t}{\lfloor t/2 \rfloor} 2^{-t}.$$

So we can bound

$$1 - p_1 \le \Pr\left[|\sum_{i=3}^{n} a_i| < 2\delta\right] \le 2^{-20} + 2^{-20} \sum_{t=1}^{20} \binom{20}{t} \binom{t}{\lfloor t/2 \rfloor} 2^{-t} \le \frac{1}{4}.$$

Combining $p_1 \ge \frac{3}{4}$ with $p_2 \ge \frac{1}{8}$ that we have proven before, this verifies (54). ☐

Let k be an integer such that $a_{k-1} \ge 1 - a_1 - a_2$, but $a_k < 1 - a_1 - a_2$ (if $a_n \ge 1 - a_1 - a_2$, set $k = n + 1$).

Claim B.6. *Let k be the smallest integer such that $a_k < 1 - a_1 - a_2$. Then we have $\sum_{i=k}^{n} a_i^2 < 328\delta^2$.*

Proof. Assume that we had $\sum_{i=k}^{n} a_i^2 \ge 328\delta^2$. Then we can find disjoint subsets $S_1, S_2, T_1, \dots, T_5$ of $\{a_k, \dots, a_n\}$ such that the following holds. For $x = 1, 2$ we have

$$104\delta^2 \ge \sum_{i \in S_x} a_i^2 \ge 100\delta^2$$

and for $y = 1, \dots, 5$, we have

$$24\delta^2 \ge \sum_{i \in T_y} a_i^2 \ge 20\delta^2.$$

Now consider the random processes

$$W(S_1; 6\delta), W(S_2; 6\delta), W(T_1; 2\delta), \dots, W(T_5; 2\delta).$$

By Lemma 3.4, each of these is successful with probability at least $\frac{1}{2}$ and independently of the other ones. We apply Observations 3.2 and 3.3, using a_3 and $r(S_1; 6\delta), \ldots, r(T_5; 2\delta)$, to bound p_1. With probability at least $\frac{93}{128}$, both some process corresponding to S_x and some process corresponding to T_y are successful, and conditional on that we get the lower bound of $\frac{7}{8}$ on $\Pr\left[|Y| \geq 1 - a_1 - a_2\right]$. Further, we get the lower bound of $\frac{3}{4}$ on $\Pr\left[|Y| \geq 1 - a_1 - a_2\right]$ if either some process corresponding to S_x or some process corresponding to T_y are successful, and the lower bound of $\frac{1}{2}$ otherwise (this last case happens at most with probability $\frac{1}{128}$). So overall, we obtain $p_1 > \frac{3}{4}$. Combining that with $p_2 \geq \frac{1}{8}$ that we have proven in Corollary B.4, we verify that (54) holds. □

Claim B.7. *Let k be the smallest integer such that $a_k < 1 - a_1 - a_2$. Then we have $a_3 - a_{k-1} < 1 - a_1 - a_2 < 2\delta$.*

Proof. If we had two terms $a_s, a_t > 1 - a_1 - a_2$ such that $|a_s - a_t| \geq 1 - a_1 - a_2$, Observation 3.2 for a_s, a_t gives $p_1 \geq \frac{3}{4}$. Combining that with $p_2 \geq \frac{1}{8}$ verifies (54). □

Claim B.8. *Let k be the smallest integer such that $a_k < 1 - a_1 - a_2$. Then we have $k < 12$.*

Proof. If we had $k \geq 12$, by Observation 3.1, we have $p_1 \geq \frac{193}{256} > \frac{3}{4}$, and combining that with $p_2 \geq \frac{1}{8}$ verifies (54). □

Claim B.9. *Let k be the smallest integer such that $a_k < 1 - a_1 - a_2$. Then we have $k < 8$.*

Proof. We have already shown that $k \leq 11$. If $11 \geq k \geq 8$, note that using our choice of δ, Claim B.6 and Claim B.7, we get

$$a_5 + a_6 + a_7 \geq \frac{2}{3} + 2\delta \geq 1 - a_1 + a_2.$$

That gives $p_2 \geq \frac{7}{32}$.

Since $k \geq 8$, we also have $p_1 \geq \frac{11}{16}$ by Observation 3.1. Hence we verify (54). □

Claim B.10. *Let k be the smallest integer such that $a_k < 1 - a_1 - a_2$. Then we have $k \geq 7$ (and hence as also $k \leq 7$, we have $k = 7$).*

Proof. By our choice of δ, we have $\sum_{i=6}^{n} a_i^2 \geq 328\delta^2$, and result thus follows by Claim B.6. □

We will now show that $\sum_{i=7}^{n} a_i^2 \geq 328\delta^2$ (which together with Claim B.6 gives a desired contradiction). By our definition of δ and assumption that $a_1 + a_2 < 1$, we either have $a_1 = \frac{2}{3} - \delta$ or $a_2 = \frac{1}{3} - \delta$.

First consider the case $a_2 = \frac{1}{3} - \delta$. Then

$$\sum_{i=1}^{6} a_i^2 \leq (\frac{2}{3} + \delta)^2 + 5(\frac{1}{3} - \delta)^2 = 1 - 2\delta + 6\delta^2,$$

and hence

$$\sum_{i=7}^{n} a_i^2 \geq 2\delta - 6\delta^2 > 328\delta^2$$

for every $0 < \delta < \frac{1}{167}$.

So we can assume that instead $a_1 = \frac{2}{3} - \delta$. But now we use (55) to bound

$$\sum_{i=1}^{6} a_i^2 \leq (\frac{2}{3} - \delta)^2 + (\frac{1}{3} + \delta)^2 + 4(\frac{1}{3})^2 \leq 1 - \frac{2}{3}\delta + 2\delta^2,$$

and hence

$$\sum_{i=7}^{n} a_i^2 \geq \frac{2}{3}\delta + 2\delta^2 > 328\delta^2$$

for every $0 < \delta < \frac{1}{489}$.

Thus we reached a desired contradiction, and the proof of Proposition B.1 is complete. $\square$