Resilience for the Littlewood-Offord Problem

Afonso S. Bandeira * Asaf Ferber† Matthew Kwan‡

September 29, 2016

Abstract

In this paper we study a resilience version of the classical Littlewood-Offord problem. Consider the sum $X(\boldsymbol{\xi}) = \sum_{i=1}^n a_i\xi_i$, where $\boldsymbol{a} = (a_i)_{i=1}^n$ is a sequence of non-zero reals and $\boldsymbol{\xi} = (\xi_i)_{i=1}^n$ is a sequence of i.i.d. random variables with $\Pr[\xi_i = 1] = \Pr[\xi_i = -1] = 1/2$. Motivated by some problems from random matrices, we consider the following question for any given x: how many of the ξ_i is an adversary typically allowed to change without making $X = x$? We solve this problem up to a constant factor and present a few interesting open problems.

1 Introduction

Let $\boldsymbol{a} = (a_i)_{i=1}^n$ be a fixed sequence of nonzero real numbers, and for a sequence of i.i.d. (independent, identically distributed) random variables $\boldsymbol{\xi} = (\xi_i)_{i=1}^n$, define the random sum

$$X = X(\boldsymbol{\xi}) = \sum_{i=1}^n a_i\xi_i.$$

Sums of this form are ubiquitous in probability theory. Most famously, the Central Limit Theorem asserts that if each a_i is the same, then X asymptotically has a normal distribution. More sophisticated variants of the central limit theorem allow the a_i to differ to some extent, and give quantitative control of the distribution. An important example is the Berry-Esseen theorem [1, 8], which gives an estimate for the probability that X lies in a given interval, in terms of the corresponding probability for an appropriate normal distribution. (We give a precise statement, adapted to our context, in Section 2). The Berry-Esseen theorem is effective when the a_is are of the same order of magnitude, in which case it can be used to deduce the estimate

$$\Pr[X = x] = O\left(\frac{1}{\sqrt{n}}\right)$$

for any x. That is to say, X is unlikely to be concentrated on any particular value (it is *anti-concentrated*).

*Department of Mathematics and Center for Data Science, Courant Institute of Mathematical Sciences, NYU. Email: bandeira@cims.nyu.edu. Part of this work was done while ASB was with the Department of Mathematics at the Massachusetts Institute of Technology and supported by NSF Grant DMS-1317308.

†Department of Mathematics, MIT. Email: ferbera@mit.edu.

‡Department of Mathematics, ETH Zürich. Email: matthew.kwan@math.ethz.ch.

In connection with a problem on random polynomials, in 1943 Littlewood and Offord [13] studied anti-concentration in the setting where each ξ_i takes the values ± 1 with equal probability, and where no assumption is made on $\boldsymbol{a}$, other than that its entries are nonzero. In this setting X can be understood as the outcome of an unbiased random walk with step sizes given by $\boldsymbol{a}$. The classical result of Littlewood and Offord [13] strengthened by Erdős [6] states that for all x and $\boldsymbol{a}$ we have

$$\Pr[X = x] \leq \binom{n}{\lfloor n/2 \rfloor} / 2^n = O\left(\frac{1}{\sqrt{n}}\right),$$

which is sharp for the sequence $\boldsymbol{a}$ with $a_i = 1$ for all i. This result was quite surprising; if one does not assume anything about the a_i then the distribution of X may be very far from normal and Berry-Esseen type bounds are no longer useful. To prove the above inequality, Erdős observed that each fiber $X^{-1}(x) = \{\boldsymbol{\xi} : X(\boldsymbol{\xi}) = x\}$ is a Sperner family under a suitable identification of $\{-1, 1\}^n$ with the Boolean hypercube. For more details about Sperner families, the reader is referred to the excellent book of Bollobás [2].

After the Littlewood-Offord problem was first introduced, many variants of the problem have been proposed and solved. One particularly interesting line of research involves the relationship between the structure of $\boldsymbol{a}$ and the resulting concentration probability $\max_x \Pr[X = x]$. Erdős and Moser [7] and Sárközy and Szemerédi [18] considered the case where the a_i are distinct, and showed that the stronger bound $\Pr[X = x] = O(n^{-3/2})$ is valid. Halász [11] gave even stronger bounds for sequences which are "arithmetically unstructured" in a certain sense. More recently, Tao and Vu [21, 22] and Nguyen and Vu [15] introduced and studied the inverse problem of characterizing the arithmetic structure of $\boldsymbol{a}$ given the concentration probability $\max_x \Pr[X = x]$.

Many connections have been found between Littlewood-Offord-type problems and various different areas of mathematics. In particular, Littlewood-Offord-type theorems were essential tools in some of the landmark results in random matrix theory (see for example [20, 21]). Most straightforwardly, the Littlewood-Offord theorem gives an upper bound on the probability that a particular row of a random ± 1 matrix is orthogonal to a given vector, and can thus be used (see for example [3, Section 14.2]) to bound the probability that a random matrix is singular.

1.1 Our setting and results

In this paper we are interested in studying a "resilience" version of the Littlewood-Offord problem. That is to say, we know that most $\boldsymbol{\xi} \in \{-1, 1\}^n$ do not satisfy $X(\boldsymbol{\xi}) = x$, but can we say that moreover most $\boldsymbol{\xi}$ are "far" (in some sense) from the event "$X = x$"?

In order to put everything in a formal setting, let $d(\boldsymbol{\xi}, \boldsymbol{\xi}')$ be the *Hamming distance* between $\boldsymbol{\xi}$ and $\boldsymbol{\xi}'$ (that is, the number of coordinates in which $\boldsymbol{\xi}$ and $\boldsymbol{\xi}'$ differ). Then, define the *resilience* of $\boldsymbol{\xi}$ with respect to the event $\{X = x\}$, denoted by $R_x(\boldsymbol{\xi})$, to be the minimum number of entries in which one should change $\boldsymbol{\xi}$ in order to obtain $X = x$. That is,

$$R_x(\boldsymbol{\xi}) = d(\boldsymbol{\xi}, X^{-1}(x)).$$

Moreover, let us denote the maximum probability that the resilience is at most k by

$$p_k(n) = \max_{\boldsymbol{a}, x} \Pr[R_x(\boldsymbol{\xi}) \leq k].$$

Equivalently, $p_k(n)$ is the maximum volume of the k-neighbourhood of a "Boolean line" $X^{-1}(x)$ in the hypercube. An immediate natural question is as follows:

Problem 1.1. *Given a non-negative integer k, what is the asymptotic behavior of $p_k(n)$?*

The Erdős-Littlewood-Offord bound clearly gives

$$p_0(n) = \Theta\big(1/\sqrt{n}\big).$$

Note that even the case $k = 1$ is already quite interesting on account of a discovery by Füredi, Kahn and Kleitman [10] that there are Sperner families whose 1-neighbourhood comprises a non-negligible proportion of the hypercube. If $p_1(n) \to 0$ then this demonstrates a special structural property of "arithmetic" Sperner families of the form $X^{-1}(x)$. More generally, we believe an especially interesting question is to understand the qualitative behaviour of $p_k(n)$ as a function of k.

Problem 1.2. *For which k does $p_k(n) \to 0$?*

Stated another way, we are asking for which k we can expect a typical $\boldsymbol{\xi} \in \{-1, 1\}^n$ to be k-resilient, regardless of the choice of x and $\boldsymbol{a}$. This question is especially compelling in view of the recent popularity of resilience problems for random graphs (see for example the excellent survey of Sudakov and Vu [19]), and in view of some questions asked by Vu [23, Conjectures 7.4-5] concerning the resilience of the singularity of random matrices. Specifically, Vu asked how many entries of a random ± 1 matrix one has to change ("globally" or "locally") to make it singular; due to the connection between the Littlewood-Offord problem and singularity of random matrices, these conjectures were actually our initial motivation for this paper.

Before stating our results, in order to give the reader some feeling for the setting we give some simple examples computing the typical resilience for specific choices of $\boldsymbol{a}$ and x.

Example 1.3. Consider the case $\boldsymbol{a} = (1, \dots, 1)$, and for simplicity assume n is even. One can easily derive that for all even x we have

$$\Pr[X = x] = \binom{n}{\frac{n+x}{2}} 2^{-n}.$$

Therefore, by a direct calculation, one can show that a.a.s.[1] we have $|X| = \Theta(\sqrt{n})$. Noting that $R_0 = |X|/2$, we typically have $R_0 = \Theta(\sqrt{n})$.

Example 1.4. Here we consider the case $\boldsymbol{a} = \big(1, 2, 4, \dots, 2^{n-1}\big)$. Note that X can take 2^n different values (the odd integers between -2^n and 2^n). This of course leads to the minimal possible concentration probability $\Pr[X = 1] = 2^{-n}$. Each x in the support of X can be obtained by exactly one $\boldsymbol{\xi}$, so R_x has the binomial distribution $\mathrm{Bin}(n, 1/2)$ and R_x is tightly concentrated around $n/2$ by a large deviation inequality for the binomial distribution (see for example [12, Theorem 2.1]).

We can see from the above two examples that the type of additive structure influencing the concentration probability does contribute to the typical resilience, to some extent. However, the following example shows that the typical resilience is much more strongly influenced by small subsequences of $\boldsymbol{a}$.

Example 1.5. Let k be the minimal integer such that $k \geq \log_2 n$ and $n - k$ is odd. Define $\boldsymbol{a}$ by $a_1 = \cdots = a_{n-k} = 1$ and $a_{n-k+i} = 2^{i-1}$. For any $\boldsymbol{\xi}$, with at most k modifications we can make $\sum_{i=1}^{k} \xi_{n-k+i} a_{n-k+i}$ equal to any odd number between $-n$ and n, so in particular we can make it equal to $-\sum_{i=1}^{n-k} \xi_i a_i$, so that $X = 0$. This means $R_0 = O(\log n)$ (with probability 1).

[1] By "asymptotically almost surely", or "a.a.s.", we mean that the probability of an event is $1 - o(1)$. Here and for the rest of the paper, asymptotics are as $n \to \infty$.

A priori, it seems plausible that Example 1.5 demonstrates essentially the least robust anti-concentration possible. Surprisingly this is not true.

Theorem 1.6. *There exists a sequence $\boldsymbol{a}$ such that a.a.s. $R_0(\boldsymbol{\xi}) = O(\log\log n)$.*

Note that the construction in Example 1.5 was effective because one can form all nonnegative integers less than 2^k with sums of subsets of $\{1, 2, 4, \ldots, 2^{k-1}\}$. That is to say, $\{1, \ldots, 2^{k-1}\}$ is an *additive basis* of $\{0, 1, 2, \ldots, 2^k - 1\}$. In order to prove Theorem 1.6 we construct a more optimized additive basis, using an idea that goes back to Rohrbach [17].

We are also able to prove that Theorem 1.6 is in fact optimal, essentially answering Problem 1.2.

Theorem 1.7. *For any $\boldsymbol{a} \in (\mathbb{R} \setminus \{0\})^n$ and $x \in \mathbb{R}$, a.a.s. $R_x(\boldsymbol{\xi}) = \Omega(\log\log n)$.*

As for Problem 1.1, we are able to find the asymptotics of each $p_k(n)$ up to a polylogarithmic factor, as stated in the next theorem.

Theorem 1.8. *For $k = 1$ we have*

$$p_1 = \Theta\Big(n^{-1/6}\Big)$$

and for any fixed $k \geq 2$ we have

$$p_k(n) = n^{-1/(2\times 3^k)} \log^{O(1)} n.$$

1.2 Notation

We use standard asymptotic notation throughout. For functions $f = f(n)$ and $g = g(n)$:

- $f = O(g)$ means there is a constant C such that $|f| \leq C|g|$,
- $f = \Omega(g)$ means there is a constant $c > 0$ such that $f \geq c|g|$,
- $f = \Theta(g)$ means that $f = O(g)$ and $f = \Omega(g)$,
- $f = o(g)$ means that $f/g \to 0$,
- $f = \omega(g)$ means that $f/g \to \infty$,

where all asymptotics are as $n \to \infty$. Also, for a real number x, the floor and ceiling functions are denoted $\lfloor x \rfloor = \max\{i \in \mathbb{Z} : i \leq x\}$ and $\lceil x \rceil = \min\{i \in \mathbb{Z} : i \geq x\}$. For a positive integer i, we write $[i]$ for the set $\{1, 2, \ldots, i\}$. Finally, all logarithms are base 2, unless specified otherwise.

2 Auxiliary results and proof outlines

In this section we introduce some tools and auxiliary results to be used in the proofs of our main results.

First, we state a version of the Berry-Esseen theorem (as mentioned in the introduction), which will be a key ingredient in our proofs. This version immediately follows from the version in [8].

Theorem 2.1. *For $X = \sum_{i=1}^n a_i\xi_i$ as in the introduction, let $\sigma^2 = \sum_{i=1}^n a_i^2$ be the variance of X, and let $\rho = \sum_{i=1}^n |a_i^3|$. Let Φ be the cumulative distribution function of the standard normal distribution. Then,*

$$\left|\Pr\left[\frac{X}{\sigma} \leq x\right] - \Phi(x)\right| = O\Big(\frac{\rho}{\sigma^3}\Big).$$

2.1 Additive bases

An order-h *additive basis* of $[n]$ is a subset $B \subseteq [n]$ such that for each $x \in [n]$, there are distinct $b_1, \dots, b_h \in B$ with $x = b_1 + \cdots + b_h$. In Example 1.5, we used an additive basis of $[n]$ which was of order $\log n$, and to construct a sequence with typical resilience $O(\log\log n)$ to prove Theorem 1.6, it makes sense to search for an additive basis of order $O(\log\log n)$. However, a critical issue is that our additive basis must be part of the sequence $\boldsymbol{a}$ itself, and therefore affects the behaviour of the typical sum. This issue was circumvented in Example 1.5 because the size of the basis was equal to its order: we were able to control each element in the basis with our $k = \log n$ changes.

In order to overcome this difficulty, we will use an additive basis consisting of two parts. The first (main) part will be an additive basis designed to be as "bottom-heavy" as possible, so that it does not cause the typical sum to be too large. The second part will be a small (size $O(\log\log n)$) sequence of the form $r, 2r, 4r, \dots, 2^q r$, which we will have total control over. For the first part, we will need the following lemma. Let $v_h(n)$ be the minimum sum of squares of an order-h additive basis of $[n]$.

Lemma 2.2. *For $h \geq 1$ we have*

$$v_h(n) = O\Big(n^{2+2/(3^h-1)}\Big).$$

Our proof of Lemma 2.2 uses an inductive construction closely resembling a construction of Rohrbach [17].

Proof. We prove by induction on h that $v_h(n) \leq Cn^{2+2/(3^h-1)}$ for some large C to be determined. For $h = 1$ we can take $\boldsymbol{a} = (1, \dots, n)$. So, assume $v_{h-1}(n) \leq Cn^{2+2/(3^{h-1}-1)}$ for all n. Then let $m = \left\lceil Dn^{2\times 3^{h-1}/(3^h-1)} \right\rceil$ for some constant $D \geq 1$ and consider an order-$(h-1)$ additive basis B' of $[m/n]$ with sum of squares $v_{h-1}(n/m)$.

Now, let $m \cdot B' = \{mb : b \in B'\}$, and note that $B = [m] \cup (m \cdot B')$ is an order-h additive basis of $[n]$. Indeed, for any $x = mq + r$, there are $b_1, \dots, b_{h-1} \in B'$ with $b_1 + \cdots + b_{h-1} = q$. Then, note that each $mb_i \in B$, and $r \in B$, so we can write $x = mb_1 + \cdots + mb_{h-1} + r$.

So,

$$\begin{aligned} v_h(n) &\leq m^3 + m^2 v_{h-1}(n/m) \\ &\leq 2D^3 n^{2\times 3^h/(3^h-1)} + 2C/D^2 n^{2\times 3^h/(3^h-1)} \end{aligned}$$

If C is large and say $D = C^{1/4}$ then this is at most

$$Cn^{2\times 3^h/(3^h-1)} = Cn^{2+2/(3^h-1)}.$$

This completes the proof. □

The proof of Theorem 1.6 using Lemma 2.2 appears in Section 4. Lemma 2.2 is also used to prove the lower bound in Theorem 1.8, in Section 5.

2.2 A recurrence relation

The following lemma is our main tool for giving upper bounds on $p_k(n)$. Indeed, Theorem 1.7 (proved in Section 3) and the upper bound for $k \geq 2$ in Theorem 1.8 (proved in Section 5.1) are immediate inductive consequences.

Lemma 2.3. *For each $k \in \mathbb{N}$, and any function $f = f(n) \to \infty$ satisfying $kf^2 \log n = o(n)$, we have*

$$p_k(n) \le \sum_{i=1}^{k} \big(O\big(kf^2 \log n\big)\big)^i p_{k-i}(n - o(n)) + O(k/f + 1/n).$$

As suggested by Example 1.5, it is important to consider small subsets of large a_i, and the proof of Lemma 2.3 indeed follows this path. The argument starts by isolating a_i which are "abnormally large" in the sense that a_i^2 is almost as large as the sum of the squares of all $a_j < a_i$ (here "almost as large" is parameterized by the function f). If there are many such a_i, then for similar reasons as in example 1.4. the resilience is very likely to be high. Now, assuming that there are a small amount of such a_i, we use strong induction to proceed: of the k changes to be made, if $i > 1$ are made on "large" numbers and $k - i$ on "small" numbers, we apply the union bound over all possible choices of i changes on the large a_i (since we know there are few of them) and induct on the subsequence of small a_i. The case where all changes are to be made on small a_i can be addressed using the Berry-Esseen theorem; showing that a typical sum of such a_i is larger than one can "cancel out" by changing just k signs.

Proof of Lemma 2.3. Fix $k > 0$ and $\boldsymbol{a}$. Without loss of generality assume $a_1 \le \cdots \le a_n$. Let $\sigma_i = \sqrt{\sum_{j=1}^{i} a_j^2}$ and $\rho_i = \sum_{j=1}^{i} a_j^3$ and for $I \subseteq [n]$ let $X_I(\boldsymbol{\xi}) = \sum_{i \in I} \xi_i a_i$.

Let $i_1 = n$, and for $t > 1$ with $a_1 \le a_{i_{t-1}}/2$ let $i_t = \max\big\{i : a_i \le a_{i_{t-1}}/2\big\}$. Let $t = t^*$ be the point where this stops (when $a_1 > a_{i_t}/2$). Let τ be the first t such that $a_{i_t} \le \sigma_{i_t}/f$, or $\tau = \infty$ if this never happens.

First, we consider the case where t^* is large

Claim 2.4. *If $t^* = \omega(k \log n)$ then $p_k \le 1/n$.*

Proof. Let $I = \{i_t : t \le t^*\}$ and let $J = [n] \backslash I$. Note that every possible sum of the a_{i_t} is distinct, so if we condition on $\boldsymbol{\xi}|_J$, then X can take 2^{t^*} different values, each occuring with probability 2^{-t^*}. There are n^k possible modifications we can make to go from $\boldsymbol{\xi}$ to $\boldsymbol{\xi}'$, so $p_k \le n^k 2^{-t^*}$. □

So, from now on we assume that $t^* = O(k \log n)$.

Claim 2.5. *We never have $\tau = \infty$ (so $\tau \le t^*$). Also, $n - i_\tau = O\big(kf^2 \log n\big)$.*

Proof. Consider t with $a_{i_t} > \sigma_{i_t}/f(n)$. Note that $|\{i : a_{i_t}/2 < a_i \le a_{i_t}\}| \le 4f^2$. Indeed, otherwise we would have the contradiction $\sigma_{i_t}^2 \ge 4f^2(n)\big(a_{i_t}^2/4\big) > \sigma_{i_t}^2$.

If we were to have $\tau = \infty$ then $a_{i_t} > \sigma_{i_t}/f(n)$ for all $t \le t^*$ so

$$\big|\big\{i : a_i > a_{i_{t^*}}/2\big\}\big| = \sum_{t=1}^{t^*} |\{i : a_{i_t}/2 < a_i \le a_{i_t}\}| = O\big(kf^2 \log n\big) = o(n)$$

and thus $\big\{i : a_i \le a_{i_{t^*}}/2\big\} \ne \varnothing$, which is a contradiction. Similarly,

$$n - i_\tau = \big|\big\{i : a_i > a_{i_{\tau-1}}/2\big\}\big| = \sum_{t=1}^{\tau-1} |\{i : a_{i_t}/2 < a_i \le a_{i_t}\}| = O\big(kf^2 \log n\big).$$ □

Now, let $I = [i_\tau]$ and $J = [n]\backslash I$. We use the union bound over all possibilities for modifying $\boldsymbol{\xi}|_J$.

For $i > 0$, there are $O\Big(\big(kf^2 \log n\big)^i\Big)$ ways to modify i elements of $\boldsymbol{\xi}|_J$, and then we have $k - i$ modifications left to use on $\boldsymbol{\xi}|_I$. For each possibility, we can condition on $\boldsymbol{\xi}|_J$ (therefore on $X_J\big(\boldsymbol{\xi}'\big)$), and the probability that we will be able to make $X_I = x - X_J$ with our remaining $k-i$ modifications is at most $p_{k-i}(i_\tau)$ by induction. Therefore, the probability we can make $X = x$ while modifying at least one element of $\boldsymbol{\xi}|_J$ is at most

$$\sum_{i=1}^{k} \big(O\big(kf^2 \log n\big)\big)^i p_{k-i}(n - o(n))$$

We also need to consider the possibility that we do not modify $\boldsymbol{\xi}|_J$ at all. In this case, condition on $\boldsymbol{\xi}|_J$ (therefore on $X_J\big(\boldsymbol{\xi}'\big) = X_J(\boldsymbol{\xi})$). Note that $\rho_{i_\tau} \le \sigma_{i_\tau} a_{i_\tau}$, so by the Berry-Esseen Theorem (theorem 2.1), with Z having the standard normal distribution,

$$\Pr\big[\big|X_I - X_J\big(\boldsymbol{\xi}'\big)\big| \le k\sigma_{i_\tau}/f(n)\big] = \Pr[|Z| \le k/f(n)] + O(a_{i_\tau}/\sigma_{i_\tau}) = O(k/f(n)).$$

Note that by changing $\boldsymbol{\xi}|_I$ we can change the value of X by at most ka_{i_τ}, which is not greater than $k\sigma_{i_\tau}/f(n)$. So the probability we can make $X = x$ without modifying $\boldsymbol{\xi}|_J$ is at most $O(k/f)$. This completes the proof. □

3 Proof of Theorem 1.7

In this section we deduce Theorem 1.7 from Lemma 2.3.

Proof. Let $\varepsilon > 0$ be any constant and let $c = 3 + \varepsilon$. We will prove that $p_k \le n^{-c^{-k-1}}$ for $k \le \log_{(3+2\varepsilon)} \log n$ and sufficiently large n. This will imply that $p_{\log_{3+2\varepsilon} \log n} \to 0$, and since ε is arbitrary this then implies that for any $\boldsymbol{a}, x$, a.a.s. $R_x > (1 + o(1)) \log_3 \log n$.

We proceed by induction on k. So, consider some k and suppose $p_{k'} \le n^{-c^{-k'-1}}$ for all $k' < k$. Let $f = n^{c^{-k}/3}$. Then,

$$\begin{aligned}
p_k &\le \sum_{i=1}^{k} \big(O\big(kf^2 \log n\big)\big)^i p_{k-i}(n - o(n)) + O(k/f + 1/n) \\
&\le \sum_{i=1}^{k} \Big(n^{2c^{-k}/3} \log^2 n\Big)^i \exp\Big(-c^{-k+i-1}(\log n + o(1))\Big) + o\Big(n^{-c^{-k-1}}\Big) \\
&= \sum_{i=1}^{k} \exp\left(-c^{-k-1} \log n \left(c^i - \frac{2c}{3} i + o(1)\right)\right) + o\Big(n^{-c^{-k-1}}\Big) \\
&\le \exp\Big(-(c/3 + o(1))c^{-k-1} \log n\Big) + o\Big(n^{-c^{-k-1}}\Big) \\
&= o\Big(n^{-c^{-k-1}}\Big) = o(1),
\end{aligned}$$

where the last inequality holds by the choice of k. This completes the proof. □

4 Proof of Theorem 1.6

In this section we prove Theorem 1.6 by constructing a sequence $\boldsymbol{a}$ such that a.a.s. $R_0 = O(\log\log n)$.

Proof. For convenience, suppose $n = 2^r$ is a power of 2. (A factor of 2 will make no difference in the asymptotics of $\log\log n$, and as we will see, the constructed sequence $\boldsymbol{a}$ will consist mostly of "1"s, which we can trim without changing the proof). Let $\varepsilon > 0$ and let

$$h = \log_{3-\varepsilon}\log n,\ h' = \log_{2-\varepsilon}\log n.$$

We will construct a sequence $\boldsymbol{a}$ such that a.a.s. $R_0 \le h + h'$.

Fix an order-h additive basis B of $\left[n/2^{h'}\right]$ with sum of squares

$$\sum_{b\in B} b^2 = O\Big(\Big(n/2^{h'}\Big)^{2+2/\left(3^h-1\right)}\Big) = o\big(n^2/\log n\big).$$

Define $\boldsymbol{a}$ by combining $\log_{2-\varepsilon} n$ copies of each $b \in B$, and the numbers $n, n/2, \dots, n/2^{h'-1}$, and padding the remaining $n - h' - |B|\log_{2-\varepsilon} n$ entries with "1"s. For convenience we assume that $\sum_{i=1}^n a_i$ is even (this can be ensured by adding a superfluous even number to B, if necessary).

Now, consider some $b \in B$ and let I_b be the set of indices corresponding to the copies of b in $\boldsymbol{a}$. Note that

$$\Pr[\boldsymbol{\xi}|_{I_b} = (1,\dots,1)] = \Pr[\boldsymbol{\xi}|_{I_b} = (-1,\dots,-1)] = 2^{-\log_{2-\varepsilon} n} = o(1/n)$$

for large C. So, by the union bound, a.a.s. for each $b \in B$ there is at least one copy of b associated with a negative sign and one associated with a positive sign. Assume this holds.

Now, let $J = \{i : a_i = 1\}$ and $I = \bigcup_{b\in B} I_b$ and note $\sum_{i\in I\cup J} a_i^2 \le n + o\big(n^2\big)$ so $\operatorname{Var} X_{I\cup J} = o(n^2)$ and by Chebyshev's inequality, a.a.s. $|X_{I\cup J}| \le 2n$. If this holds, by modifying the signs associated with $n, n/2, \dots, n/2^{h'-1}$ we can make $|X| \le n/2^{h'-1}$. Then, there are $b_1, \dots b_h \in B$ and $\xi_1, \dots, \xi_h \in \{-1, 1\}$ with $\sum_{i=1}^h \xi_i b_i = |X/2|$, and we can therefore make $X = 0$ by changing a further h signs in $\boldsymbol{\xi}|_I$. This completes the proof. □

5 Proof of Theorem 1.8

In this section we prove Theorem 1.8. Similarly to Theorems 1.6 and 1.7, the upper and lower bounds are proved completely differently.

5.1 Upper bounds

The upper bound $p_k(n) \le n^{-1/\left(2\times 3^k\right)} \log^{O(1)} n$ follows immediately from Theorem 2.3, using a similar (but much simpler) induction argument to the one used to prove Theorem 1.7, as follows.

Proof. Suppose $p_{k'} \le n^{-1/\left(2\times 3^{k'}\right)} \log^{O(1)} n$ for $k' < k$. Let $f = n^{1/\left(2\times 3^k\right)}$. Then,

$$p_k \le \sum_{i=1}^{k} \big(O\big(kf^2\log n\big)\big)^i p_{k-i}(n - o(n)) + O(k/f + 1/n)$$

$$\leq \sum_{i=1}^{k} f^{2i} n^{-1/\left(2\times 3^{k-i}\right)} \log^{O(1)} n$$
$$\leq n^{-1/\left(2\times 3^{k-i}\right)} \log^{O(1)} n.$$

This completes the proof. □

For the upper bound on $p_1(n)$ we will use Sárközy and Szemerédi's theorem which asserts that if $\boldsymbol{a}$ has distinct elements, then

$$\Pr[X = x] = O\Big(n^{-3/2}\Big).$$

Proof of the upper bound on $p_1(n)$. Fix any $\boldsymbol{a}, x$. Suppose there are q distinct values in $\boldsymbol{a}$, so the union bound gives

$$\Pr[R_x \leq 1] = O\Big(qn^{-1/2}\Big).$$

Alternatively, let $a_{i_1}, \ldots, a_{i_q}$ give a representative for each distinct value and let $I = \{i_1, \ldots, i_q\}$. Conditioning on $\boldsymbol{\xi}|_{[n]\setminus I}$ and using the union bound,

$$\Pr[R_x \leq 1] = O\Big(q \times q^{-3/2}\Big) = O\Big(q^{-1/2}\Big).$$

No matter the value of q, one of these gives

$$\Pr[R_x \leq 1] = O\Big(n^{-1/6}\Big),$$

which completes the proof. □

5.2 Lower bounds

First we prove the general lower bound $p_k(n) \geq n^{-1/\left(2\times 3^k\right)} \log^{O(1)} n$.

Proof. Let

$$\sigma_I = \sqrt{\sum_{i\in I} a_i^2},\ \rho_I = \sum_{i\in I} a_i^3,$$

with $\sigma = \sigma_{[n]}$ and $\rho = \rho_{[n]}$. The proof proceeds in a similar way to Theorem 1.6, as follows. Fix an order-k additive basis B of $[n]$ with sum of squares

$$\sum_{b\in B} b^2 = O\Big(q^{2+2/\left(3^k-1\right)}\Big),$$

for some q to be determined. Define $\boldsymbol{a}$ by combining $\log n$ copies of each $b \in B$ (let I be the corresponding set of indices in $\boldsymbol{a}$), and padding the remaining $n - |B| \log n$ entries with "1"s. As in Section 4, we can assume that $\sum_{i=1}^{n} a_i$ is even, and we can show that a.a.s. for each $b \in B$ there is at least one copy of b associated with a negative sign and one associated with a positive sign. Assume this holds.

Now, $\sigma_I^2 = O\Big(q^{2+2/\left(3^k-1\right)} \log n\Big)$, $\sigma^2 = n - q + \sigma_I^2$, and $\rho \leq n - q + q\sigma_I^2$. Choose q such that $\sigma_I^2 = \varepsilon n$ for some small constant $\varepsilon > 0$ to be determined. This implies

$$q = n^{1/(2+2/(3^k-1))} \log^{O(1)} n = n^{1/2-1/(2\times 3^k)} \log^{O(1)} n.$$

By the Berry-Esseen Theorem,

$$\Pr[|X/2| \le q] = \Theta\Big(\frac{q}{\sigma}\Big) + O\Big(\frac{\rho}{\sigma}\Big) = \Theta\bigg(\frac{q}{\sqrt{n}}\bigg) + O\bigg(\frac{\varepsilon^2 q}{\sqrt{n}}\bigg) = n^{-1/(2\times 3^k)} \log^{O(1)} n.$$

Now, if $|X/2| \le q$ then there are $b_1, \dots b_k \in B$ and $\xi_1, \dots, \xi_k \in \{-1, 1\}$ with $\sum_{i=1}^k \xi_i b_i = |X/2|$, and we can therefore make $X = 0$ by changing a further k signs in $\boldsymbol{\xi}|_I$. This completes the proof. □

Finally, we prove the sharp bound $p_1(n) = \Omega\big(n^{-1/6}\big)$.

Proof. The construction is exactly the same as above (with $k = 1$), but we include only one copy of each element in B. Recalling the base case for the induction in Section 4, this means $\boldsymbol{a}$ is defined by

$$a_1 = \dots = a_{n-q} = 1,\ a_{n-q+i} = i.$$

By exactly the same argument as above, by the Berry-Esseen Theorem,

$$\Pr[|X/2| \le q] = \Theta\Big(n^{-1/6}\Big)$$

for some $q = \Theta\big(n^{1/3}\big)$. Similarly, with $I = [n-q]$ and $J = [n] \setminus I$, we can use the Berry-Esseen theorem on X_I and X_J, to show that for large C and any $x \in \mathbb{R}$,

$$\Pr\big[X_J > C\sqrt{n}\big] \le 1/C,$$
$$\Pr[|X_I + x| \le 2q] = O\Big(n^{-1/6}\Big).$$

So,

$$\begin{aligned}\Pr\big[|X/2| \le q \text{ and } |X_J| > C\sqrt{n}\big] &= \sum_{x:|x|>C\sqrt{n}} \Pr[|X_I + x| \le 2q] \Pr[X_J = x] \\ &= O\bigg(\frac{n^{-1/6}}{C}\bigg).\end{aligned}$$

For large enough C, we therefore have

$$\Pr\big[|X/2| \le q \text{ and } |X_J| \le C\sqrt{n}\big] = \Theta\Big(n^{-1/6}\Big) - O\bigg(\frac{n^{-1/6}}{C}\bigg) = \Theta\Big(n^{-1/6}\Big).$$

Now, with $N = n - q$, for any $x \le 2C\sqrt{n}$ we have

$$\begin{aligned}\Pr[X_I = x] &= \binom{N}{(N+x)/2} / 2^N \\ &= \frac{\Theta(1)}{\sqrt{N}(1 + x/N)^{(N+x)/2}(1 - x/N)^{(N-x)/2}} \\ &= \frac{\Theta(1)}{\sqrt{N}(1 - x^2/N^2)^{N/2}(1 + O(x/N))^{x/2}} \\ &= \frac{\Theta(1)}{\sqrt{N}(1 - O(1/n))^{O(n)}(1 + O(1/x))^{x/2}}\end{aligned}$$

$$= \Theta\Big(\frac{1}{\sqrt{n}}\Big).$$

That is to say, the probabilities $\Pr[X_I = x]$ differ from each other by at most a constant factor.

Let $s(a) = \mathrm{sign}(\xi_{n-q+a})$. Conditioning on any choice of $\boldsymbol{\xi}|_J$ such that $X_J(\boldsymbol{\xi}) \le C\sqrt{n}$, we have

$$\begin{aligned}\Pr\big[|X/2| \le q \text{ and } \mathrm{sign}(X) = \mathrm{sign}\big(\xi_{n-q+|X/2|}\big)\big] &= \sum_{a:0\le a\le q} \Pr[X_I = 2s(a)a - X_J] \\ &= \Theta\Bigg(\sum_{a:0\le a\le q} \Pr[X_I = -2s(a)a - X_J]\Bigg) \\ &= \Theta\big(\Pr\big[|X/2| \le q \text{ and } \mathrm{sign}(X) \ne \mathrm{sign}\big(\xi_{n-q+|X/2|}\big)\big]\big).\end{aligned}$$

(note that $X_I = X_J \pmod 2$ so X is divisible by 2). So,

$$\begin{aligned}&\Pr\big[|X/2| \le q \text{ and } \mathrm{sign}(X) = \mathrm{sign}\big(\xi_{n-q+|X/2|}\big) \text{ and } X_J(\boldsymbol{\xi}) \le C\sqrt{n}\big] \\ &\quad = \Theta\big(\Pr\big[|X/2| \le q \text{ and } X_J(\boldsymbol{\xi}) \le C\sqrt{n}\big]\big) \\ &\quad = \Omega\Big(n^{-1/6}\Big).\end{aligned}$$

But if $|X/2| \le q$ and $\mathrm{sign}(X) = \mathrm{sign}\big(\xi_{n-q+|X/2|}\big)$ then we can modify $\xi_{n-q+|X/2|}$ to make $X = 0$. This completes the proof. □

6 Concluding remarks and open problems

In this paper we have investigated the resilience of the anti-concentration in the Littlewood-Offord problem. We hope that results and ideas of the type in this paper can be applied to other problems, in particular to the resilience questions for random matrices raised by Vu [23]. There are several very interesting open questions that remain.

- Most obviously, there is the question of removing the polylogarithmic factor in Theorem 1.8. This problem is analogous to the situation in the Erdos-Moser problem, where Sárközy and Szemerédi [18] removed a polylogarithmic factor in Erdős and Moser's original bound. Indeed, it is due to Sárközy and Szemerédi's theorem that we could get a tight result for $p_1(n)$.

- Let $R = \min_x R_x$. We showed that for $k \le (1 - o(1)) \log_3 \log n$, a.a.s. $R > k$ for any $\boldsymbol{a}$, and for $k \ge (1 + \log_2(3) + o(1)) \log_3 \log n$ there is $\boldsymbol{a}$ such that a.a.s. $R \le k$. It remains open what the behaviour is when k is in the narrow interval between these values. Is there a "sharp threshold" k in the sense that a.a.s. $R > k$ for any $\boldsymbol{a}$, but there is $\boldsymbol{a}$ such that a.a.s. $R \le k + 2$?

- The constructions used to prove Theorem 1.6 had a very special structure consisting of "layered" additive bases. The proof of the lower bound in Theorem 1.7 seems to indicate that this type of structure is necessary for the typical resilience to be small. It would be interesting to formalize this idea in an inverse theorem of some kind. An inverse theorem for Theorem 1.8 would also be interesting: fixing k, what can be said about the structure of $\boldsymbol{a}$ given $\max_x \Pr[R_x \le k]$?

- We have considered the setting where X is a linear combination of independent Rademacher random variables. As suggested to us by Van Vu, we can consider more generally the setting

where X is a low-degree polynomial. The anti-concentration problem in this setting was initated by Costello, Tao and Vu [5] in order to study symmetric random matrices, and was further developed by many authors, most recently by Meka, Nguyen and Vu [14]. Resilience problems in this setting appear to be much more difficult than for the ordinary Littlewood-Offord problem, and are likely to require new ideas.

References

[1] A. C. Berry, The accuracy of the Gaussian approximation to the sum of independent variates, Transactions of the American Mathematical Society 49 (1941), no. 1, 122–136.

[2] B. Bollobás, Combinatorics: set systems, hypergraphs, families of vectors, and combinatorial probability, Cambridge University Press, 1986.

[3] B. Bollobás, Random Graphs, Academic Press, 1985.

[4] J. Bourgain, V. H. Vu, and P. M. Wood, On the singularity probability of discrete random matrices, Journal of Functional Analysis 258 (2010), no. 2, 559–603.

[5] K. P. Costello, T. Tao, V. Vu, et al., Random symmetric matrices are almost surely nonsingular, Duke Mathematical Journal **135** (2006), no. 2, 395–413.

[6] P. Erdős, On a lemma of Littlewood and Offord, Bulletin of the American Mathematical Society 51 (1945), no. 12, 898–902.

[7] P. Erdős, Extremal problems in number theory, 1965 Proc. Sympos. Pure Math., Vol. VIII, 181-189 AMS, Providence, R.I.

[8] C.-G. Esseen, On the Liapounoff limit of error in the theory of probability, Arkiv för Matematik, Astronomi och Fysik A28 (1942), no. 9, 1–19.

[9] W. Feller, An introduction to probability and its applications, (1971) Vol. II. Wiley, New York.

[10] Z. Füredi, J. Kahn, and D. J. Kleitman, Sphere coverings of the hypercube with incomparable centers, Discrete mathematics **83** (1990), no. 1, 129–134.

[11] G. Halász, Estimates for the concentration function of combinatorial number theory and probability, Periodica Mathematica Hungarica 8 (1977), 197-211.

[12] S. Janson, T. Łuczak, and A. Ruciński, Random graphs, Cambridge University Press, 2000.

[13] J. E. Littlewood and A. C. Offord, On the number of real roots of a random algebraic equation. III, Math. Mat. Sbornik N.S. 12 (1943), no. 3, 277–286.

[14] R. Meka, O. Nguyen, and V. Vu, Anti-concentration for polynomials of Rademacher random variables and applications in complexity theory, arXiv preprint arXiv:1507.00829 (2015).

[15] H. Nguyen and V. Vu, Optimal inverse Littlewood-Offord theorems, Advances in Mathematics 226 (2011), no. 6, 5298–5319.

[16] H. Nguyen and V. Vu, Small probability, inverse theorems, and applications, on the occasion of Paul Erdos' 100th anniversary, Bolyai Society Mathematical Studies, Vol. 25 (2013).

[17] H. Rohrbach, Ein Beitrag zur additiven Zahlentheorie, Mathematische Zeitschrift **42** (1937), no. 1, 1–30.

[18] A. Sárközy and E. Szeméredi, Über ein Problem von Erdős und Moser, Acta Arithmetica 11 (1965), 205-208.

[19] B. Sudakov and V. H. Vu, Local resilience of graphs, Random Structures & Algorithms **33** (2008), no. 4, 409–433.

[20] T. Tao and V. Vu, From the Littlewood-Offord problem to the circular law: universality of the spectral distribution of random matrices, Bulletin of the American Mathematical Society 46 (2009), no. 3, 377–396.

[21] T. Tao and V. H. Vu, Inverse Littlewood-Offord theorems and the condition number of random discrete matrices, Annals of Mathematics (2009), 595–632.

[22] T. Tao and V. H. Vu, A sharp inverse Littlewood-Offord theorem, Random Structures & Algorithms 37 (2010), no. 4, 525–539.

[23] V. Vu, Random discrete matrices, Horizons of combinatorics, Springer, 2008, 257–280.