Resilience for the Littlewood–Offord Problem

Afonso S. Bandeira [*]   Asaf Ferber[†]   Matthew Kwan[‡]

August 2, 2017

Abstract

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. Rademacher random variables (that is, $\Pr[\xi_i = 1] = \Pr[\xi_i = -1] = 1/2$). The classical Littlewood–Offord problem asks for the best possible upper bound on the concentration probabilities $\Pr[X = x]$. In this paper we study a resilience version of the Littlewood–Offord problem: how many of the ξ_i is an adversary typically allowed to change without being able to force concentration on a particular value? We solve this problem asymptotically, 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) Rademacher random variables $\boldsymbol{\xi} = (\xi_i)_{i=1}^{n}$ (meaning $\Pr[\xi_i = 1] = \Pr[\xi_i = -1] = 1/2$), define the random sum

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

Sums of this form are ubiquitous in probability theory. For example, X can be interpreted as the outcome of an unbiased random walk with step sizes given by $\boldsymbol{a}$. The central limit theorem asserts that if the a_i are all equal then X asymptotically has a normal distribution. More flexible variants of the central limit theorem allow the a_i to differ to an extent, and give quantitative control over the distribution of X. An important example is the Berry–Esseen theorem [2, 9], which gives an estimate for the probability that X lies in a given interval, comparing it to the corresponding probability for an appropriately scaled normal distribution (we give a precise statement, adapted to our context, later in the paper). The Berry–Esseen theorem is effective when the a_i are of the same order of magnitude, in which case it can be used to easily deduce the estimate

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

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

Over half a century ago, in connection with their study of random polynomials, Littlewood and Offord [13] considered anti-concentration in the general setting where no assumption is made on $\boldsymbol{a}$, other than that its entries being nonzero. The classical result of Littlewood and Offord [13] strengthened by Erdős [6] states that no matter the choice of $\boldsymbol{a} \in (\mathbb{R} \setminus \{0\})^n$, for all $x \in \mathbb{R}$ we have

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

which is sharp for the sequence $\boldsymbol{a} = (1, 1, \ldots, 1)$. This result is particularly remarkable due to the fact that if one does not assume anything about the a_i, then the distribution of X may be far from normal and Berry–Esseen type bounds may no longer be meaningful.

Erdős' proof of the above inequality was combinatorial and extremely simple, as follows. First, we can assume that each a_i is positive, because changing the sign of some a_i does not affect the distribution of X. Then, observe that a sign vector $\boldsymbol{\xi} \in \{-1, 1\}^n$ can be identified with the subset $\{i : \xi_i = 1\}$ of $\{1, \ldots, n\}$, and under this identification each fiber $X^{-1}(x) = \{\boldsymbol{\xi} : X(\boldsymbol{\xi}) = x\}$ corresponds to a *Sperner family*[1]. It then suffices to apply a classical bound for the maximal size of a Sperner family.

Since the Littlewood–Offord problem was first introduced, many variants of it have been addressed; 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 [17] considered the case where the a_i are all distinct, and showed that the stronger bound $\Pr[X = x] = O(n^{-3/2})$ holds. Halász [11] gave even stronger bounds for sequences which are "arithmetically unstructured" in an appropriate sense. More recently, Tao and Vu [20, 21] and Nguyen and Vu [15] investigated the inverse problem of characterizing the arithmetic structure of $\boldsymbol{a}$ given the concentration probability $\max_x \Pr[X = x]$.

Many fruitful connections have been found between Littlewood–Offord-type problems and various areas of mathematics. In particular, Littlewood–Offord-type theorems are essential tools in some of the landmark results in random matrix theory (see for example [19, 20]). In particular, 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 [4, Section 14.2]) to bound the probability that a Bernoulli random matrix is singular.

1.1 Our Results

In this paper we are interested in studying a "resilience" version of the Littlewood–Offord problem. Given a sequence $\boldsymbol{a} \in (\mathbb{R} \setminus \{0\})^n$ and a real number $x \in \mathbb{R}$, we know that most sequences $\boldsymbol{\xi} \in \{-1, 1\}^n$ do not satisfy the event $\{X(\boldsymbol{\xi}) = x\}$. We are interested in understanding whether most sequences $\boldsymbol{\xi}$ are "far" from this event. In order to make this question precise we need a few definitions. Given two sequences $\boldsymbol{\xi}, \boldsymbol{\xi}' \in \{-1, 1\}^n$ we define $d(\boldsymbol{\xi}, \boldsymbol{\xi}')$ to be the *Hamming distance* between $\boldsymbol{\xi}$ and $\boldsymbol{\xi}'$ (that is, $d(\boldsymbol{\xi}, \boldsymbol{\xi}')$ denotes the number of coordinates in which $\boldsymbol{\xi}$ and $\boldsymbol{\xi}'$ differ). If $S \subset \{-1, 1\}^n$ is a subset of the hypercube we further define $d(\boldsymbol{\xi}, S)$ as the minimum Hamming distance from $\boldsymbol{\xi}$ to a point in

---

[1]A Sperner family is a collection of subsets of $[n]$ in which no subset is included in any other. For more details on Sperner families, the reader is referred to the book of Bollobás [3].

S. Finally, for a fixed sequence $\boldsymbol{a}$ of non-zero reals and $\xi \in \{-1, 1\}^n$, let us define

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

which is the minimum number of signs one needs to change in $\boldsymbol{\xi}$ in order to satisfy $X = x$. (For completeness, if $X = x$ is impossible then we set $R_x(\boldsymbol{\xi}) = \infty$). We refer to $R_x(\boldsymbol{\xi})$ as the *resilience* of $\boldsymbol{\xi}$ with respect to the event $\{X \neq x\}$, and if $R_x > k$ we say $\boldsymbol{\xi}$ is *k-resilient*.

Given $\boldsymbol{a}$ we define

$$q_k(\boldsymbol{a}) = \max_x \Pr[R_x^{\boldsymbol{a}}(\boldsymbol{\xi}) \leq k].$$

as the maximum probability that $\boldsymbol{\xi}$ fails to be k-resilient. We also define $p_k(n)$ as the "worst case" for this probability over all sequences $\boldsymbol{a} \in (\mathbb{R} \setminus \{0\})^n$:

$$p_k(n) = \max_{\boldsymbol{a} \in \{\mathbb{R} \setminus \{0\}\}^n} q_k(\boldsymbol{a})$$

Equivalently, $p_k(n)$ corresponds to the maximum volume of the k-neighbourhood of a suitable "Boolean hyperplane" $X^{-1}(x)$ 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 $p_k(n)$ as $n \to \infty$?*

The Erdős–Littlewood–Offord bound trivially gives

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

Understanding the case $k = 1$ already has non-trivial implications. In fact, Füredi, Kahn and Kleitman [10] showed that there are Sperner families whose 1-neighbourhood comprises a constant proportion of the hypercube, while we will see in Section 4.2 that $p_1(n) \to 0$. 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 $k = k(n)$ does $p_k(n) \to 0$ as $n \to \infty$?*

In other words, 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 influential survey of Sudakov and Vu [18]), and in view of questions asked by Vu [22, 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 our initial motivation to investigate the questions treated in this paper.

Before stating our results, we compute the typical resilience for a few simple illustrative specific choices of $\boldsymbol{a}$ and x.

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

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

Standard binomial estimates show that with (say) 99% certainty we have $|X| = \Theta(\sqrt{n})$. Noting that $R_0 = |X|/2$, we typically have $R_0 = \Theta(\sqrt{n})$.

3

Example 1.4. Let us next consider the sequence $\boldsymbol{a} = (1, 2, \ldots, n)$. Since all the a_i are distinct, it follows from the result of Sárközy and Szemerédi [17] that $q_0(\boldsymbol{a}) = O(n^{-3/2})$. Moreover, changing k signs of $\boldsymbol{\xi}$ can increase or decrease X by no more than kn, so there are at most $2kn + 1$ ways to affect X by changing k signs. Therefore, as long as $kn = o(n^{3/2})$ (that is, $k = o(n^{1/2})$), the union bound shows that for any x, typically $R_x \geq k$.

Example 1.5. Take $\boldsymbol{a} = (1, 2, 4, \ldots, 2^{n-1})$. Note that X can take 2^n different values (the odd integers between -2^n and 2^n). This of course leads to the minimum possible concentration probability $\max_x \Pr[X = x] = 2^{-n}$. Each x in the support of X can be obtained by exactly one $\boldsymbol{\xi}$, so R_x has the binomial distribution $\text{Bin}(n, 1/2)$ and 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 three examples that the type of additive structure influencing the concentration probability does contribute somewhat to the typical resilience. However, the following example shows that the typical resilience can be much more strongly influenced by small subsequences of $\boldsymbol{a}$.

Example 1.6. 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 $\boldsymbol{\xi}$, modifying at most k coordinates 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).

Somewhat surprisingly, there is a sequence which typically results in significantly lower resilience than Example 1.6.

Theorem 1.7. *There exists a sequence $\boldsymbol{a} \in (\mathbb{R} \setminus \{0\})^n$ such that for any fixed $\varepsilon > 0$, a.a.s.[2] $R_0(\boldsymbol{\xi}) \leq (1 + \varepsilon) \log_3 \log n$. (That is to say, for $k \geq (1 + \varepsilon)) \log_3 \log n$, we have $p_k \to 1$).*

The crux of Example 1.6 was the fact that one can form all non-negative integers less than 2^k with sums of subsets of $\{1, 2, 4, \ldots, 2^{k-1}\}$. In other words, $\{1, \ldots, 2^{k-1}\}$ is an *additive basis* of $\{0, 1, 2, \ldots, 2^k - 1\}$. The proof of Theorem 1.7, which we defer to Section 3, involves a more efficient additive basis construction, using an idea from a 1937 paper of Rohrbach [16].

We are also able to prove that Theorem 1.7 is in fact optimal, essentially answering Problem 1.2.

Theorem 1.8. *For any fixed $\varepsilon > 0$, any $\boldsymbol{a} \in (\mathbb{R} \setminus \{0\})^n$ and any $x \in \mathbb{R}$, a.a.s. $R_x(\boldsymbol{\xi}) \geq (1 - \varepsilon) \log_3 \log n$. (That is to say, for $k = (1 - \varepsilon) \log_3 \log n$, we have $p_k \to 0$).*

We prove Theorem 1.8 in Section 2.

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

Theorem 1.9. *We have*
$$p_1 = \Theta\left(n^{-1/6}\right),$$
and for any fixed $k \geq 2$,
$$p_k(n) = n^{-1/(2 \times 3^k)} \log^{O(1)} n.$$

4

1.2 Notation

For a set of indices $I \subseteq [n]$ define

$$X_I(\boldsymbol{\xi}) = \sum_{i \in I} a_i \xi_i$$

to be the "part" of X corresponding to I.

We use standard asymptotic notation throughout. For functions $f = f(n)$ and $g = g(n)$ we write $f = O(g)$ to mean there is a constant C such that $|f| \leq C|g|$, we write $f = \Omega(g)$ to mean there is a constant $c > 0$ such that $f \geq c|g|$, we write $f = \Theta(g)$ to mean that $f = O(g)$ and $f = \Omega(g)$, and we write $f = o(g)$ or $g = \omega(f)$ to mean that $f/g \to 0$. All asymptotics are taken 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.

1.3 Structure of the paper

The structure of the paper is as follows. In Section 2 we give a lower bound on typical resilience (proving Theorem 1.8), in Section 3 we construct a sequence with low resilience (proving Theorem 1.7), and in Section 4 we estimate the asymptotics of $p_k(n)$ (proving Theorem 1.9).

2 Lower bound for typical resilience

In this section we prove Theorem 1.8. The heart of the proof is the following recurrence relation for $p_k(n)$.

Lemma 2.1. *Let $k \in \mathbb{N}$ and let $f := f(n) \to \infty$ be any function satisfying $(k+1)f^2 \log n < n$. Then, for some constant C,*

$$p_k(n) \leq \sum_{\ell=1}^{k} \big(4(k+1)f^2 \log n\big)^\ell \max_{n'} p_{k-\ell}(n') + C(k/f + 1/n),$$

where the maximum is over all n' satisfying $0 \leq n - n' \leq 4(k+1)f^2 \log n$.

We remark that Lemma 2.1 is also used in the proof of Theorem 1.9.

2.1 Proof of Lemma 2.1

Before giving the details of the proof of Lemma 2.1, we give a brief outline of the ideas. Intuitively, we expect X to typically have order of magnitude about its standard deviation (which is $\sqrt{\sum_i a_i^2}$). If this is much larger than any individual a_i then we expect the resilience R_0 to be large, as flipping a sign in $\boldsymbol{\xi}$ has a relatively small impact on X. Therefore (as already suggested by Example 1.6), it is important to distinguish those a_i that are "abnormally large", and consider them separately.

So, the proof of Lemma 2.1 starts by isolating "large" a_i such that a_i^2 is almost as large as the sum of the squares of all $a_j \leq a_i$ (here "almost as large" is parameterized by the function f). If there are many such a_i, then for similar reasons as in Example 1.5 the resilience is very likely to be high. We can therefore assume that there are a small quantity of such a_i; we need to give an upper bound on the probability of being able to make $X = x$ with up to k sign changes.

First consider the case where $\ell \geq 1$ of the k changes are made on the "large" numbers. Because there are few such numbers, it is not too wasteful to take the union bound over each possible way to make these changes. Then, we can recursively bound the probability that we can make $X = x$ with at most $k - \ell$ further changes to the "small" numbers.

Otherwise, if none of the sign changes are made on "large" numbers, then as we have already explained, the typical size of X is larger than one can "cancel out" without making a large number of sign flips, so the resilience is high. We will rigorously establish this fact using the Berry–Esseen theorem, as follows (this version of the Berry–Esseen theorem immediately follows from the statement in [9]).

Theorem 2.2. *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,*

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

Now we give the details of the proof of Lemma 2.1.

Proof of Lemma 2.1. Fix $k > 0$ and $\boldsymbol{a}$. Note that we may assume that all the a_i are non-negative, as changing signs of any subset of the a_i does not change the distribution of X. Moreover, by relabeling if necessary, we can assume that

$$0 \leq a_1 \leq \cdots \leq a_n.$$

We denote partial sums of squares as follows:

$$\sigma_i^2 = \sum_{j=1}^{i} a_j^2.$$

Now, let $i_1 := n > i_2 > \ldots > i_t$ be a longest subsequence of indices for which the following properties hold for all $j < t$:

1. $a_{i_j} \geq 2a_{i_{j+1}}$, and

2. for all $i > i_{j+1}$ we have $a_{i_j} < 2a_i$.

Note that Property 1 forces all possible signed sums of the a_{i_j} to be distinct (that is, $X_{\{i_1,\ldots,i_t\}}$ takes 2^t different values). Maximality and Property 2 imply that $a_i > a_{i_t}/2$ for all $i \in [n]$.

If t is large, then the atom probabilities are small, and therefore the resilience is high. We summarize this in the following claim.

Claim 2.3. *If $t > (k+1)\log n$ then $q_k(\boldsymbol{a}) < 1/n$.*

Proof. Let $I = \{i_j : j \leq t\}$ and condition on the outcomes of the ξ_j, $j \notin I$. The random variable X can then take 2^t different values, each occurring with probability 2^{-t}. This means that, unconditionally, the probability that X is equal to any particular value is at most 2^{-t}. Now, there are at most n^k ways to change up to k of the ξ_i, and given a particular choice of indices at which to perform changes, the resulting sequence $\boldsymbol{\xi}'$ has the same distribution as $\boldsymbol{\xi}$. Therefore, the probability that $X(\boldsymbol{\xi}')$ is equal to any particular value after this change is still at most 2^{-t}, so by the union bound $q_k(\boldsymbol{a}) \leq n^k 2^{-t} < 1/n$ as desired. $\square$

From now on we assume that $t \leq (k+1)\log n$. Let τ be the first j for which $a_{i_j} \leq \sigma_{i_j}/f$. (If there is no such j, we set $\tau = \infty$). This condition defining τ is chosen so that we will later be able to control $X_{[i_\tau]}$ via the Berry–Esseen theorem. In the following claim we show that $\tau < \infty$ and moreover that $[i_\tau]$ comprises most of $[n]$.

Claim 2.4. *We have $\tau \leq t$ and $n - i_\tau \leq 4(k+1)f^2\log n$.*

Proof. Note that for any j with $a_{i_j} > \sigma_{i_j}/f$ we have

$$\left| \left\{ i : \frac{a_{i_j}}{2} < a_i \leq a_{i_j} \right\} \right| \leq 4f^2.$$

Indeed, otherwise we would have the contradiction

$$\sigma_{i_j}^2 > 4f^2 \left(\frac{a_{i_j}}{2} \right)^2 = (f\, a_{i_j})^2 > \sigma_{i_j}^2.$$

If we were to have $\tau = \infty$ this would mean $a_{i_j} > \sigma_{i_j}/f$ for all $j \leq t$. Therefore, this would lead to the contradiction

$$n = \left| \left\{ i : a_i > \frac{a_{i_t}}{2} \right\} \right| = \sum_{j=1}^{t} \left| \left\{ i : \frac{a_{i_j}}{2} < a_i \leq a_{i_j} \right\} \right| \leq 4f^2 t \leq 4f^2(k+1)\log n < n.$$

Similarly, we have

$$n - i_\tau = \left| \left\{ i : a_i > \frac{a_{i_{\tau-1}}}{2} \right\} \right| = \sum_{t=1}^{\tau-1} \left| \left\{ i : \frac{a_{i_t}}{2} < a_i \leq a_{i_t} \right\} \right| \leq 4f^2 t \leq 4(k+1)f^2\log n. \qquad \square$$

Now, let $n' = i_\tau$, let $J = [n']$ and let $I = [n] \backslash J$. For each $0 \leq \ell \leq k$ we will consider the case where we change exactly ℓ elements of $\boldsymbol{\xi}|_I$, and we will then take a union bound over all ℓ.

For $\ell > 0$, there are at most $\left(4(k+1)f^2\log n \right)^\ell$ ways to modify ℓ elements of $\boldsymbol{\xi}|_I$. For each such possibility, we can condition on the modified value of $\boldsymbol{\xi}|_I$ (therefore on $X_I(\boldsymbol{\xi})$), and for any x the probability that we will be able to make $X_J = x - X_I$ with our remaining $k - \ell$ modifications is at most $p_{k-\ell}(n')$ by induction. Therefore, the probability we can make $X = x$ while modifying at least one element of $\boldsymbol{\xi}|_I$ is at most

$$\sum_{\ell=1}^{k} \left(4f^2(k+1)\log n \right)^\ell p_{k-\ell}(n').$$

It remains to consider the possibility that we do not modify $\boldsymbol{\xi}|_I$ at all. Again, condition on $\boldsymbol{\xi}|_I$ (therefore on X_I). Note that $\sum_{i \in [i_\tau]} a_i^3 \leq \sigma_{i_\tau}^2 a_{i_\tau}$, so by the Berry–Esseen theorem (Theorem 2.2), with Z having the standard normal distribution,

$$\begin{aligned} \Pr[|X_J + X_I - x| \leq k\sigma_{i_\tau}/f] &= \Pr[|Z + (X_I - x)/\sigma_{i_\tau}| \leq k/f] + O(a_{i_\tau}/\sigma_{i_\tau}) \\ &\leq \Pr[|Z| \leq k/f] + O(a_{i_\tau}/\sigma_{i_\tau}) \\ &= O(k/f). \end{aligned}$$

Note that by changing k elements in $\boldsymbol{\xi}|_J$ we can change the value of X by at most ka_{i_τ}, which is not greater than $k\sigma_{i_\tau}/f$ by the choice of τ. So, the probability that we can make $X = x$ without modifying $\boldsymbol{\xi}|_I$ at all is $O(k/f)$. By combining all the above bounds, we obtain the desired result. $\quad\square$

7

2.2 Proof of Theorem 1.8

Finally, we show how to deduce Theorem 1.8 from Lemma 2.1.

Proof of Theorem 1.8. Let $\delta > 0$ be a small constant and let $c = 3 + \delta$. We prove that $p_k \le n^{-c^{-k-1}}$ for $k \le \log_{3+2\delta} \log n$ and sufficiently large n, from which the theorem statement will follow. (In this section all asymptotics are uniform over $k \le \log_{3+2\delta} \log n$). We prove our desired bound on p_k by induction on k. For $k = 0$, as mentioned in the introduction, the Erdős–Littlewood–Offord theorem gives

$$p_0 = \Theta(n^{-1/2}) \le n^{-c^{-1}}.$$

Next, consider some $0 < k \le \log_{3+2\delta} \log n$ and suppose $p_{k'} \le n^{-c^{-k'-1}}$ for all $k' < k$. Observe that

$$c^k \le (3 + \delta)^{\log_{3+\delta} \log n / \log_{3+\delta}(3+2\delta)} = (\log n)^{1-a},$$

for some constant $0 < a < 1$ depending on δ, and let $f = n^{c^{-k}/3} \ge \exp((\log n)^a/3)$. For some $n' \ge n - 4(k+1)f^2 \log n = n - o(n)$, Lemma 2.1 says that

$$p_k \le \sum_{\ell=1}^{k} \left(4(k+1)f^2 \log n\right)^\ell p_{k-\ell}(n') + O(k/f + 1/n).$$

Observe that $k/f + 1/n = o\left(n^{-c^{-k-1}}\right)$ and $\log(n - o(n)) = \log n + o(1)$, so it follows that

$$p_k \le \sum_{\ell=1}^{k} \left(n^{2c^{-k}/3} \log^2 n\right)^\ell \exp\left(-c^{-k+\ell-1}(\log n + o(1))\right) + o\left(n^{-c^{-k-1}}\right).$$

Now, recalling that $c^k \le (\log n)^{1-a}$, for $1 \le \ell \le k$ we have

$$
\begin{aligned}
&\left(n^{2c^{-k}/3} \log^2 n\right)^\ell \exp\left(-c^{-k+\ell-1}(\log n + o(1))\right) \\
&= \exp\left(c^{-k-1} \log n \left(\frac{2c}{3}\ell + \frac{2c^{k+1} \log \log n}{\log n}\ell - c^\ell(1 + o(1/\log n))\right)\right) \\
&= \exp\left(-c^{-k-1} \log n \left(c^\ell - \frac{2c}{3}\ell + o(1)\right)\right) \\
&= \exp\left(-c^{-k-1} \log n \left(\frac{c}{3} + o(1)\right)\right).
\end{aligned}
$$

(We have used the fact that $c^\ell - (2c/3)\ell \ge c - (2c/3) = c/3$ for $\ell \ge 1$ and $c \ge 2$). Consequently,

$$
\begin{aligned}
p_k &\le k \exp\left(-\left(\frac{c}{3} + o(1)\right)c^{-k-1} \log n\right) + o\left(n^{-c^{-k-1}}\right) \\
&= o\left(n^{-c^{-k-1}}\right).
\end{aligned}
$$

This concludes the proof of the desired bound on p_k, and it follows that if $k = \log_{3+2\delta} \log n$ then

$$p_k \le n^{-c^{-k-1}} = \exp\left(-c^{-1}c^{-k} \log n\right) = o(1).$$

In particular, since δ is arbitrary it follows that for any $\varepsilon > 0$, $\boldsymbol{a} \in (\mathbb{R} \setminus \{0\})^n$ and $x \in \mathbb{R}$, a.a.s. $R_x > (1 - \varepsilon) \log_3 \log n$.

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\Box$

3 A sequence with low typical resilience

In this section we prove Theorem 1.7 by constructing a sequence $\boldsymbol{a}$ such that a.a.s. $R_0 = (1 + o(1)) \log_3 \log n$.

Let $X = \sum_{i=1}^{n} a_i \xi_i$ as in the introduction. To construct a sequence $\boldsymbol{a}$ that results in low typical resilience, we are looking to improve on the idea of Example 1.6. We start with the "nicely behaved" sequence $\boldsymbol{a} = (1, 1, \ldots, 1)$, and we look to "plant" a small subset B in $\boldsymbol{a}$ which allows us to "cancel out" the typical outcomes of X. This leads us to consider the following notion.

3.1 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 $b_1, \ldots, b_t \in B$, $t \leq h$, with $x = b_1 + \cdots + b_t$. As an easy example, the reader may note that the key part of the sequence in Example 1.6 was the additive basis $\{1, 2, 2^2 \ldots, 2^{\lceil \log n \rceil - 1}\}$ of $[n]$, which is of order $\lceil \log n \rceil$. In order to improve on Example 1.6 and prove Theorem 1.7, we wish to include a lower-order additive basis in our sequence $\boldsymbol{a}$.

The critical issue with this idea is that our additive basis must be part of the sequence $\boldsymbol{a}$ itself, and therefore it contributes to the behaviour of the typical sum. For example, if we define $\boldsymbol{a}$ by taking a sequence of n' "1"s and combining it with a low-order additive basis of $[n']$, then due to the extra "weight" of the additive basis, X can take values (much) larger than n', which are not "covered" by the additive basis. This issue was circumvented in Example 1.6 because the size of the basis was equal to its order: we were able to control each element in the basis with our $k = \Theta(\log n)$ changes.

In order to minimize the impact of including an additive basis in $\boldsymbol{a}$, we need an additive basis with small sum of squares. (Recall that the variance of X is $\sum_{i=1}^{n} a_i^2$, and this controls the typical size of $|X|$). Let $v_h(n)$ be the minimum sum of squares of an order-h additive basis of $[n]$. That is,

$$v_h(n) = \min \left\{ \sum_{b \in B} b^2 \mid B \text{ is an order-}h \text{ additive basis of } [n] \right\}.$$

In the following lemma we provide an upper bound on $v_h(n)$.

Lemma 3.1. *For $h \geq 1$ we have*
$$v_h(n) \leq 10^h n^{2 + 2/(3^h - 1)}.$$

Our proof of Lemma 3.1 uses an inductive construction closely resembling a construction of Rohrbach [16].

Proof. The proof is by induction on h. For the base case, $h = 1$, one can take $B = \{1, \ldots, n\}$. Note that indeed we have
$$\sum_{i=1}^{n} i^2 \leq n^3 \leq 10^1 n^{2 + 2/(3^1 - 1)}.$$

Next, consider $h > 1$ and assume that for all n we have
$$v_{h-1}(n) \leq 10^{h-1} n^{2 + 2/(3^{h-1} - 1)}.$$

For what follows it will be convenient to use the identity
$$2 + 2/(3^h - 1) = 2 \times 3^h/(3^h - 1). \tag{3.1}$$

9

Set
$$m = \left\lceil n^{2\times 3^{h-1}/(3^h - 1)} \right\rceil$$

and consider an order-$(h-1)$ additive basis B' of $[\lfloor n/m \rfloor]$ with sum of squares $v_{h-1}(\lfloor n/m \rfloor)$.

Now, let us define $m \cdot B' = \{mb : b \in B'\}$, and note that $B = [m] \cup (m \cdot B')$ is an order-h additive basis of $[n]$. Indeed, for any $x = mq + r$ (with $q \le \lfloor n/m \rfloor$ and $1 \le r \le m$), there are $b_1, \ldots, b_t \in B'$ with $t \le h - 1$ and $b_1 + \cdots + b_t = q$. Then, note that each $mb_i \in B$, and $r \in B$, so we can write $x = mb_1 + \cdots + mb_t + r$, which is a sum of at most h elements. So, we have

$$v_h(n) \le m^3 + m^2 v_{h-1}(\lfloor n/m \rfloor).$$

Now, observe that $n \le m^{(3^h - 1)/(2\times 3^{h-1})}$. Using (3.1),

$$\left(\frac{3^h - 1}{2\times 3^{h-1}} - 1\right)\left(2 + \frac{2}{3^{h-1} - 1}\right) = \left(\frac{3^h - 1}{2\times 3^{h-1}} - 1\right)\left(\frac{2\times 3^{h-1}}{3^{h-1} - 1}\right)$$
$$= \frac{3^h - 1 - 2\times 3^{h-1}}{3^{h-1} - 1} = 1$$

so the induction hypothesis gives $v_{h-1}(\lfloor n/m \rfloor) \le 10^{h-1} m$. Therefore, $v_h(n) \le (10^{h-1} + 1)m^3$. Noting that $\lceil x \rceil \le 2x$ for $x \ge 1$, and again using (3.1),

$$v_h(n) \le 8(10^{h-1} + 1)n^{2\times 3^h/(3^h - 1)} \le 10^h n^{2 + 2/(3^h - 1)}.$$

This completes the proof. □

3.2 Proof of Theorem 1.7

Recall that the key idea for our construction is to "plant" an additive basis of an appropriate order, with low sum-of-squares, in the all-1 sequence. Note that for $k = \log_3 \log n$ we can use Lemma 3.1 to find an order-k additive basis of $[n]$ with sum-of-squares $O(10^k n^2) = n^2 \log^{O(1)} n$. A variance bound of $\sigma^2 = n^2 \log^{O(1)} n$ is enough to prove that a.a.s. $|X| \le n \log^{O(1)} n$, but is not quite enough to prove that a.a.s. $|X| \le 2n$, which we need for the additive basis of $[n]$ to be effective. We can address this issue by additionally including a very small number of large powers of 2 in our sequence; by modifying the corresponding signs we will be able to make $|X| \le 2n$. A second consideration is the fact that changing a sign increases X if the sign was negative and decreases X if the sign was positive. In order to guarantee that we can a.a.s. use our additive basis to adjust X in either direction, we can include many repetitions of the elements of our basis (so that a.a.s. there will be a copy of each element with a positive sign and with a negative sign). These basic ideas are enough for a sequence with typical resilience $O(\log \log n)$, but to optimize our construction for the asymptotically lowest possible resilience requires some additional technical details. In particular we include in our sequence two different additive bases of different orders, each with different amounts of repetition.

Proof of Theorem 1.7. Consider small $\varepsilon > 0$ and let

$$h = \left\lceil \log_{3-\varepsilon} \log n \right\rceil, \ h' = \left\lceil \log_{3-\varepsilon} \log \log n \right\rceil, \ r = \left\lceil \log \log^2 \log n \right\rceil.$$

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

Fix an order-h additive basis B of $\left[\lceil n/\log^2 n\rceil\right]$ with sum of squares

$$\sum_{b\in B} b^2 = O\left(10^h\left(n/\log^2 n\right)^{2+2/\left(3^h-1\right)}\right) = o\left(n^2/\log n\right)$$

(note that $10^h = o(\log^3 n)$ for small ε), and similarly fix an order-h' additive basis B' of $\left[\lceil \log^2 n/\log^2\log n\rceil\right]$ with sum of squares

$$o\left(\log^4 n/\log\log n\right).$$

Note that $|B| = o(n/\log n)$ and $|B'| = o(\log^2(n)/\log\log n)$.

Now, define $\boldsymbol{a}$ by combining:

- $\lceil \log n\rceil$ copies of each $b \in B$ (let I be the corresponding set of indices of $\boldsymbol{a}$);

- $\left\lceil \log\log^2 n\right\rceil$ copies of $\left\lceil n/\log^2 n\right\rceil b$ for each $b \in B'$ (let J be the corresponding set of indices);

- the numbers $m, 2m, \ldots, 2^{r-1}m$ for $m = \left\lceil n/\log^2 n\right\rceil\left\lceil \log^2 n/\log^2\log n\right\rceil$ (let K be the corresponding set of indices);

- $n - r - |B|\lceil \log n\rceil - |B'|\left\lceil \log\log^2 n\right\rceil$ copies of the number "1" (let L be the corresponding set of indices).

Also, if necessary change one of the "1"s in the final bullet point to a "2" to ensure that $\sum_{i=1}^n a_i$ is even. (This guarantees that X is always even).

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

$$\Pr[\boldsymbol{\xi}|_{I_b} = (1, \ldots, 1)] = \Pr[\boldsymbol{\xi}|_{I_b} = (-1, \ldots, -1)] \le 2^{-\log n} = \frac{1}{n} = o\left(\frac{1}{|B|}\right).$$

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. Similarly, a.a.s. for each $b \in B'$ there is a negative and positive copy of $\left\lceil n/\log^2 n\right\rceil b$. In what follows we assume both these properties hold.

Next, note that

$$\sigma_I^2 = O(\log n)o\left(n^2/\log n\right) = o\left(n^2\right),$$
$$\sigma_J^2 = O(\log\log n)O\left(\left(n/\log^2 n\right)^2\right)o\left(\log^4 n/\log\log n\right) = o\left(n^2\right),$$
$$\sigma_L^2 \le n,$$

so $\sigma_{I\cup J\cup L}^2 = o(n^2)$ and by Chebyshev's inequality, a.a.s. $|X_{I\cup J\cup L}| \le 2n$. Assuming this, by modifying $\boldsymbol{\xi}|_K$ we can make $|X| \le 2m$. Then, there are $b_1, \ldots, b_t \in B'$ with $t \le h'$ and $\sum_{i=1}^t \xi_i b_i = \left\lceil |X/2|/\lceil n/\log^2 n\rceil\right\rceil$, and we can therefore make $|X| \le 2\lceil n/\log^2 n\rceil$ by changing a further $t \le h'$ signs in $\boldsymbol{\xi}|_J$. Finally, there are $b_1, \ldots, b_s \in B$ with $s \le h$ and $\sum_{i=1}^h \xi_i b_i = |X/2|$, so we can make $X = 0$ by changing $s \le h$ signs in $\boldsymbol{\xi}|_I$. This completes the proof. $\square$

4 Asymptotics of $p_k(n)$

In this section we prove Theorem 1.9. We stress that throughout this section, k is fixed.

4.1 Upper bounds

The upper bound $p_k(n) \leq n^{-1/(2 \times 3^k)} \log^{O(1)} n$ follows immediately from Theorem 2.1, using a similar (but much simpler) induction argument to the one used to prove Theorem 1.8, as follows.

Proof. For $k = 0$, as mentioned in the introduction, the Erdős–Littlewood–Offord theorem gives

$$p_0 = \Theta(n^{-1/2}) = \Theta(n^{-1/(2 \times 3^0)}).$$

For $k > 0$, suppose $p_{k'} \leq n^{-1/(2 \times 3^{k'})} \log^{O(1)} n$ for $k' < k$. Let $f = n^{1/(2 \times 3^k)}$. Then, using Lemma 2.1,

$$
\begin{aligned}
p_k &\leq \sum_{\ell=1}^{k} \left(4f^2(k+1) \log n\right)^{\ell} p_{k-\ell}(n - o(n)) + O(k/f + 1/n) \\
&\leq \sum_{\ell=1}^{k} n^{\ell/3^k} n^{-1/(2 \times 3^{k-\ell})} \log^{O(1)} n + O(n^{-1/(2 \times 3^k)}) \\
&\leq \sum_{\ell=1}^{k} n^{-(3^\ell - 2\ell)/(2 \times 3^k)} \log^{O(1)} n + O(n^{-1/(2 \times 3^k)}) \\
&\leq n^{-1/(2 \times 3^k)} \log^{O(1)} n.
\end{aligned}
$$

This completes the proof. $\qquad\square$

For the tight upper bound $p_1(n) = O(n^{-1/6})$ we will use Sárközy and Szemerédi's theorem (mentioned in the introduction) which asserts that if $\boldsymbol{a}$ has distinct elements, then

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

Proof of the upper bound on $p_1(n)$. Fix any $\boldsymbol{a}, x$. Suppose there are g distinct values in $\boldsymbol{a}$, so there are at most $2g$ different ways to affect X by flipping a sign. Just as in the proof of Claim 2.3, for any particular choice of index at which to perform a flip, the resulting sequence $\boldsymbol{\xi}'$ has the same distribution as $\boldsymbol{\xi}$, so the probability that $X(\boldsymbol{\xi}') = x$ after the change is $O(n^{-1/2})$ by the Erdős–Littlewood–Offord theorem. The union bound over all possible ways to make one flip (or no flips) then gives

$$\Pr[R_x \leq 1] = O\left(gn^{-1/2}\right). \tag{4.1}$$

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

$$\Pr[R_x \leq 1] = O\left(g \times g^{-3/2}\right) = O\left(g^{-1/2}\right). \tag{4.2}$$

No matter the value of g, one of (4.1) or (4.2) gives $\Pr[R_x \leq 1] = O\left(n^{-1/6}\right)$ (if $g \leq n^{1/3}$ then use (4.1), otherwise use (4.2)). This completes the proof. $\qquad\square$

4.2   Lower bounds

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

Proof. Let

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

and define $\sigma = \sigma_{[n]}$ and $\rho = \rho_{[n]}$, for use with the Berry–Esseen theorem (Theorem 2.2). The proof proceeds in a similar way to Theorem 1.7, as follows. Let $g = (\varepsilon n / \log n)^{1/(2+2/(3^k-1))} = n^{1/2-1/(2 \times 3^k)} \log^{O(1)} n$, for some small $\varepsilon > 0$ to be determined (where useful for clarity, asymptotic notation will be uniform over ε). Using Lemma 3.1 fix an order-k additive basis B of $[g]$ with sum of squares

$$\sum_{b \in B} b^2 = \Theta\left(g^{2+2/(3^k-1)}\right).$$

Define $\boldsymbol{a}$ by combining $\lceil 2 \log n \rceil$ copies of each $b \in B$ (let I be the corresponding set of indices in $\boldsymbol{a}$), and padding the remaining $n - |B|\lceil 2 \log n \rceil$ entries with "1"s. As in Section 3, if necessary we can change a "1" to a "2" to ensure 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, we have $\sigma_I^2 = \Theta\left(g^{2+2/(3^k-1)} \log n\right)$ and $\sigma^2 = n - g + \sigma_I^2$, and since each $a_i \leq g$, we also have $\rho = n - g + \rho_I \leq n - g + g\sigma_I^2$. By the definition of g, this means $\sigma_I^2 = \Theta(\varepsilon n)$, so $\sigma^2 = \Theta(n)$ and $\rho = O(\varepsilon g n)$. By the Berry–Esseen theorem (Theorem 2.2), for small enough ε we have

$$\Pr[|X/2| \leq g] = \Theta\left(\frac{g}{\sigma}\right) + O\left(\frac{\rho}{\sigma^3}\right) = \Theta\left(\frac{g}{\sqrt{n}}\right) - O\left(\frac{\varepsilon g}{\sqrt{n}}\right) = n^{-1/(2 \times 3^k)} \log^{O(1)} n.$$

Now, if $|X/2| \leq g$ then there are $b_1, \ldots b_t \in B$, $t \leq k$, with $\sum_{i=1}^{t} b_i = |X/2|$, and we can therefore make $X = 0$ by changing t signs in $\boldsymbol{\xi}|_I$. This completes the proof. $\qquad\square$

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

Proof. The construction is similar to the one given above (with $k = 1$), but we include only one copy of each element in B. Recalling the base case for the induction in the proof of Lemma 3.1, define $\boldsymbol{a}$ by

$$a_1 = \cdots = a_{n-g} = 1, \ a_{n-g+i} = i,$$

where $g = (\varepsilon n)^{1/3}$ for some $\varepsilon > 0$ to be determined. (We will be able to choose an appropriate ε such that $\sum_{i=1}^{n} a_i$ is even, without having to change a "1" to a "2"). Let $J = [n - g]$ and $I = [n] \setminus J$. By the same arguments as above, we have $\sigma_I^2 = \Theta(g^3) = \Theta(\varepsilon n)$, $\sigma^2 = \Theta(n)$, and $\rho_I, \rho = O(\varepsilon g n)$, so using the Berry–Esseen theorem in the same way as in the last proof gives

$$\Pr[|X/2| \leq g] = \Theta\left(\frac{g}{\sqrt{n}}\right) - O\left(\frac{\varepsilon g}{\sqrt{n}}\right) = \Theta\left(n^{-1/6}\right).$$

Similarly, we can use the estimates $\sigma_I^2, \sigma_J^2 = \Theta(n)$, $\rho_I = O(n^{4/3})$ and $\rho_J = O(n)$, and the Berry–Esseen theorem applied to X_I and X_J, to show that for large C and any $x \in \mathbb{R}$,

$$\Pr\left[|X_I| > C\sqrt{n}\right] \leq 1/C,$$

$$\Pr[|X_J + x| \le 2g] = O\left(n^{-1/6}\right).$$

So,

$$\Pr\left[|X/2| \le g \text{ and } |X_I| > C\sqrt{n}\right] = \sum_{x:|x|>C\sqrt{n}} \Pr[|X_J + x| \le 2g]\Pr[X_I = x]$$

$$= O\left(n^{-1/6}\right) \sum_{x:|x|>C\sqrt{n}} \Pr[X_I = x]$$

$$= O\left(\frac{n^{-1/6}}{C}\right).$$

For large enough C, we therefore have

$$\Pr\left[|X/2| \le g \text{ and } |X_I| \le C\sqrt{n}\right] = \Theta\left(n^{-1/6}\right) - O\left(\frac{n^{-1/6}}{C}\right) = \Theta\left(n^{-1/6}\right).$$

Now, with $N = n - g$, for any x with $N + x$ even and $|x| \le 2C\sqrt{n}$ we have

$$\Pr[X_J = x] = \binom{N}{(N+x)/2}/2^N$$

$$= \frac{\Theta(1)}{\sqrt{N}(1 + x/N)^{(N+x)/2}(1 - x/N)^{(N-x)/2}}$$

$$= \frac{\Theta(1)}{\sqrt{N}(1 - x^2/N^2)^{N/2}(1 + O(x/N))^{x/2}}$$

$$= \frac{\Theta(1)}{\sqrt{N}(1 - O(1/n))^{O(n)}(1 + O(1/x))^{x/2}}$$

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

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

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

$$\Pr\left[|X/2| \le g \text{ and } \text{sign}(X) = \text{sign}(\xi_{n-g+|X/2|})\right] = \sum_{a:0\le a\le g} \Pr[X_J = 2s(a)a - X_I]$$

$$= \Theta\left(\sum_{a:0\le a\le g} \Pr[X_J = -2s(a)a - X_I]\right)$$

$$= \Theta\left(\Pr\left[|X/2| \le g \text{ and } \text{sign}(X) \ne \text{sign}(\xi_{n-g+|X/2|})\right]\right).$$

So,

$$\Pr\left[|X/2| \le g \text{ and } \text{sign}(X) = \text{sign}(\xi_{n-g+|X/2|}) \text{ and } |X_I(\boldsymbol{\xi})| \le C\sqrt{n}\right]$$
$$= \Theta\left(\Pr\left[|X/2| \le g \text{ and } |X_I(\boldsymbol{\xi})| \le C\sqrt{n}\right]\right)$$
$$= \Omega\left(n^{-1/6}\right).$$

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

5   Concluding remarks and open problems

In this paper we have investigated the resilience of the anti-concentration in the Littlewood–Offord problem. We hope the results and ideas in this paper can be applied to other problems, in particular to the resilience questions for random matrices raised by Vu [22]. We would like to draw attention to several interesting open questions.

- It would be interesting if the polylogarithmic error term could be removed from Theorem 1.9. This problem is analogous to the situation in the Erdős–Moser problem, where Sárközy and Szemerédi [17] 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 the right order of magnitude for $p_1(n)$.

- We showed that for some $\varepsilon \to 0$, for $k \le (1 - \varepsilon) \log_3 \log n$, a.a.s. $R_x > k$ for any $\boldsymbol{a}, x$, and for $k \ge (1 + \varepsilon) \log_3 \log n$ there is $\boldsymbol{a}$ such that a.a.s. $R_0 \le k$. It remains open what the behaviour is when k is very close to $\log_3 \log n$. Is there a "sharp threshold" $k = k(n)$ in the sense that $p_k \to 0$ but $p_{k+1} \to 1$ (or $p_{k+a} \to 1$ for some fixed a)? This would be analogous to the two-point concentration phenomenon for the chromatic number of random graphs [1]. As pointed out to us by Joel Spencer, there is also the possibility that there is some $f = o(\log \log n)$ such that, if $k = \log_3 \log n + \lambda f(n)$, then p_k depends nontrivially on λ. This would be analogous to the behaviour of the connectivity threshold for random graphs; see [8].

- The constructions used to prove Theorem 1.7 had a very special "layered" structure, and the proof of the lower bound in Theorem 1.8 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, and we suspect such a theorem would be very useful for the random matrix questions of Vu mentioned in the introduction. An inverse theorem for Theorem 1.9 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, one 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 more difficult than for the ordinary Littlewood–Offord problem, and are likely to require new ideas.

We would also like to highlight an alternative construction of a sequence $\boldsymbol{a}$ which results in $\Pr[R_0 \le k] \ge 99\%$ for $k = (1 + o(1)) \log \log n$, due to Svante Janson and Joel Spencer. Let $\boldsymbol{a}$ consist of all "1"s, except $1000 \log(i + 1)$ copies of each $\sqrt{n}/i$ for $1 \le i \le n^{0.2}$, and $10 \log n$ copies of each j for $2 \le j \le n^{0.3}$. (If the sum of all these numbers is odd, change a single "1" to a "2"). We give a sketch proof that this sequence has the claimed property. First observe that

$$\mathrm{Var}(X) = O\left(n + n^{0.9} \log n + \sum_{i=1}^{n^{0.2}} \frac{n}{i^2} \log(i + 1)\right) = O(n),$$

15

so by Chebyshev's inequality, $|X| \leq L\sqrt{n}$ for some L, with probability at least 99.9%. Also, observe that with probability at least 99.9% there is a positive and negative sign associated with each distinct value in $\boldsymbol{a}$. Indeed, the probability that this fails is at most

$$2\left(n^{0.3}n^{-10} + \sum_{i=1}^{n^{0.2}}(i+1)^{-1000}\right) < 0.1\%.$$

Now, consider an outcome of X satisfying both of these properties. By the divergence of the harmonic series, there is B such that $\sum_{i=1}^{B}\sqrt{n}/B \geq L\sqrt{n}$; first make at most B flips among the elements $\sqrt{n}/i$, for $i \leq B$, to obtain $|X| < \sqrt{n}/B$. Then, the key reason we have resilience $(1 + o(1))\log\log n$ is that if $2\sqrt{n}/(i+1) \leq |X| < 2\sqrt{n}/i$ then flipping a sign to add or subtract $2\sqrt{n}/(i+1)$ results in $|X| \leq 2\sqrt{n}/(i(i+1))$. That is to say, if $|X| \approx \sqrt{n}/i$ then with one flip we can make $|X| \approx \sqrt{n}/i^2$, so it takes approximately $\log\log n$ flips to go from $|X| \approx \sqrt{n}/B$ to $|X| \approx n^{0.3}$, after which we can make $X = 0$ with a single flip. We suspect that with some optimization this type of construction could lead to an alternative proof of Theorem 1.7.

Acknowledgements. We warmly thank Svante Janson and Joel Spencer for giving us permission to present their alternative construction. We also thank Van Vu and Joel Spencer for many insightful discussions.

References

[1] N. Alon and M. Krivelevich, The concentration of the chromatic number of random graphs, Combinatorica **17** (1997), no. 3, 303–313.

[2] 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.

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

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

[5] K. P. Costello, T. Tao, V. Vu, 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, Proc. Sympos. Pure Math., Vol. VIII, AMS, Providence, R.I (1965), 181-189.

[8] P. Erdős and A. Rényi, On the evolution of random graphs, Publ. Math. Inst. Hungar. Acad. Sci **5** (1960), 17–61.

[9] 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.

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

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

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

[19] 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.

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

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

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