On Littlewood-Offord theory for arbitrary distributions

T. Juškevičius, V. Kurauskas

Vilnius University

December 19, 2019

Abstract

Let $X_1, \ldots, X_n$ be independent identically distributed random vectors in $\mathbb{R}^d$. We consider upper bounds on $\max_x \mathbb{P}(a_1 X_1 + \cdots + a_n X_n = x)$ under various restrictions on $X_i$ and the weights $a_i$. When $\mathbb{P}(X_i = \pm 1) = \frac{1}{2}$, this corresponds to the classical Littlewood-Offord problem. We prove that in general for identically distributed random vectors and even values of $n$ the optimal choice for $(a_i)$ is $a_i = 1$ for $i \leq \frac{n}{2}$ and $a_i = -1$ for $i > \frac{n}{2}$, regardless of the distribution of $X_1$. Applying these results for Bernoulli random variables answers a recent question of Fox, Kwan and Sauermann.

Finally, we provide sharp bounds for concentration probabilities of sums of random vectors under the condition $\sup_x \mathbb{P}(X_i = x) \leq \alpha$, where it turns out that the worst case scenario is provided by distributions on an arithmetic progression that are in some sense as close to the uniform distribution as possible.

Unlike much of the literature on the subject we use neither methods of harmonic analysis nor those from extremal combinatorics.

1 Introduction

Let $X_1, \ldots, X_n$ be independent identically distributed (iid) discrete random vectors in $\mathbb{R}^d$. In this paper we shall be interested in bounding probabilities $\mathbb{P}(a_1 X_1 + \cdots + a_n X_n = x)$ under various assumptions on the weights $a_i$ and the distributions of $X_1$. The special case

*This project has received funding from European Social Fund (project No 09.3.3-LMT-K-712-02-0151) under grant agreement with the Research Council of Lithuania (LMTLT).
when $P(X_i = \pm 1) = \frac{1}{2}$ is known as the Littlewood-Offord problem. The classical result of Erdős [4] is that for non-zero real weights $a_i$ we have

$$P(a_1 X_1 + \cdots + a_n X_n = x) \leq P(X_1 + \cdots + X_n \in \{0, 1\}) = \left(\frac{n}{2}\right)^n = \frac{2}{\pi n} + O(n^{-3/2}). \quad (1)$$

Kleitman [12] proved that the latter result remains true for $a_i \in \mathbb{R}^d$. That is, linear combinations with equal weights exhibit the worst case behaviour. Perhaps the most general results in with restrictions of the arithmetical structure of the weights $a_i$ were obtained in the classical work of Halász [8]. Other types of restrictions were considered in [5, 19].

The first goal of the present work is to extend these problems to random variables with an arbitrary distribution in $\mathbb{R}^d$. It turns our that for even values of $n$ there is a unique choice of weights $a_i$ that is optimal for arbitrary distributions. The case of odd values of $n$ is discussed later on in the paper. Let us state a result that is essential for all other results in the paper.

**Lemma 1.** Let $X_1, \ldots, X_{2k}$ be independent random vectors in $\mathbb{R}^d$. Then there is $j \in \{1, 2, \ldots, 2k\}$ such that for all $x \in \mathbb{R}^d$

$$P(X_1 + \cdots + X_{2k} = x) \leq P(Y_1 - Y_2 + \cdots + Y_{2k-1} - Y_{2k} = 0)$$

where the random variables $Y_i$ are independent and distributed as $X_j$.

The inequality is strict unless $\sum X_i$ and $\sum (-1)^{i+1} Y_i$ have the same distribution.

Intuitively, this lemma says that the probability for a random walk with steps from some class of distributions to hit a particular value $x$ is never greater than the probability to hit the origin by going back and forth according to some specific distribution from the class. Its proof is very simple: it merely uses multiple applications of the comparison between the arithmetic and geometric means.

Two straightforward consequences of Lemma 1 are

**Corollary 1.** We have

$$P(a_1 X_1 + \cdots + a_{2k} X_{2k} = x) \leq P(X_1 - X_2 + \cdots + X_{2k-1} - X_{2k} = 0)$$

(a) for iid real random variables $X_i$ and any non-zero $a_i \in \mathbb{R}^d$; and

(b) for iid real random vectors $X_i$ in $\mathbb{R}^d$ and any non-zero $a_i \in \mathbb{R}$.

In other words, for even values of $n$ the worst case scenario in the latter two situations is provided by the balanced collection of $\pm 1$’s, regardless of the distribution of the random variables $X_i$. We shall refer to Lemma 1 as the “balancing lemma”.
Remark 1. Bounds for even values of \( n \) also give bounds for odd values since by conditioning on \( X_{n+1} \) (or by monotonicity of the Levy concentration function) for any \( n \geq 1 \)

\[
\max_x \mathbb{P}(X_1 + \cdots + X_{n+1} = x) \leq \max_x \mathbb{P}(X_1 + \cdots + X_n = x).
\]

The second part of our work has a bit different flavour. Instead of linear combinations of random vectors with given distributions we consider sums of independent random vectors \( X_i \) in \( \mathbb{R}^d \) such that no \( X_i \) takes a particular value with too large a probability. For \( \alpha \in (0, 1) \) we shall denote by \( U^\alpha \) (possibly supplied with a subscript) a random variable such that \( \mathbb{P}(U^\alpha = k) = \alpha \) for \( k = 0, 1, \ldots, \lfloor \frac{1}{\alpha} \rfloor - 1 \) and \( \mathbb{P}(U^\alpha = \lfloor \frac{1}{\alpha} \rfloor) = 1 - \mathbb{P}(U^\alpha \in \{0, \ldots, \lfloor \frac{1}{\alpha} \rfloor - 1\}) \). For \( \alpha = \frac{1}{m} \) with \( m \in \mathbb{N} \) this random variable has the uniform distribution on \( \{0, \ldots, m - 1\} \) and for \( \alpha \in (0, \frac{1}{2}) \) it has the Bernoulli distribution with parameter \( 1 - \alpha \). We then establish the following inequality.

**Theorem 1.** Let \( X_1, \ldots, X_{2k} \) be independent random vectors in \( \mathbb{R}^d \) such that for some \( \alpha \in (0, 1) \) and all \( i \) we have

\[
\sup_{x \in \mathbb{R}^d} \mathbb{P}(X_i = x) \leq \alpha.
\]

Then

\[
\mathbb{P}(X_1 + \cdots + X_{2k} = x) \leq \mathbb{P}(U_1^\alpha - U_2^\alpha + \cdots + U_{2k-1}^\alpha - U_{2k}^\alpha = 0),
\]

where the random variables \( U_i^\alpha \) are iid with distribution \( U^\alpha \).

Note that the latter inequality is optimal as the random variables \( U_i^\alpha \) satisfy the condition of the theorem.

The latter result for \( \alpha = \frac{1}{m} \) with \( m \in \mathbb{N} \) was established by Rogozin \[16\] and also follows from the results of Leader and Radcliffe \[13\]. Bounds for arbitrary \( \alpha \) were obtained by Ushakov \[20\], but they were not optimal in general. We postpone the detailed discussion regarding a more complete history of this problem to Section 3.

In the setting of Theorem 1 for all \( n \) (even or odd) we get

\[
\mathbb{P}(X_1 + \cdots + X_n = x) \leq c_\alpha n^{-1/2}(1 + o(1)) \quad \text{where}
\]

\[
c_\alpha = \left(\frac{\pi}{6} \alpha [\alpha^{-1}]([\alpha^{-1}] + 1)((4 - 3\alpha)[\alpha^{-1}] + 2 - 3\alpha[\alpha^{-1}]^2)\right)^{-1/2}.
\]

When \( \alpha^{-1} \) is integer, this simplifies to \( c_\alpha = \frac{\sqrt{\pi \alpha}}{\sqrt{(1 - \alpha^2)\pi}} \).

Very recently Fox, Kwan and Sauermann \[17\] have posed the following question (we rephrase it slightly).

**Question 1.** Let \( a_i \) be non-zero real number and let \( X_i \) be independent Bernoulli random variables with parameter \( 0 < p \leq \frac{1}{2} \). What upper bounds (in terms of \( n \) and \( p \)) can we give on the maximum point probability

\[
\max_{x \in \mathbb{R}} \mathbb{P}(a_1 X_1 + \cdots + a_n X_n = x)?
\]
Taking Bernoulli random variables in either Corollary 1(a) or 1(b) we obtain the optimal bound for the probability in question for even values of \( n \) under more general conditions. Alternatively, it is a special case of Theorem 1 applied with with \( \alpha = 1 - p \). The situation for odd \( n \) seems to be much more involved. We were only able to prove that optimal \( a_i \) must be \( \pm 1 \) (up to a constant) when \( n \) (see Section 4) and get some partial results illustrating why this case is more difficult.

For the Bernoulli case, or \( \alpha \geq \frac{1}{2} \) in Theorem 1 Ushakov’s paper [20] communicated by Prokhorov in the early 80s gives an asymptotically optimal bound \((2\pi np(1 - p))^{-1/2}(1 + o(1))\).

**Remark 2.** A solution to Question 1 has been very recently and independently obtained by Singhal [18] using different methods with stronger results than ours in the case when \( n \) is odd.

The paper is organized as follows. Lemma 1 with its corollaries are proved in Section 2. Theorem 1 is proved and the history of the problem discussed in Section 3. Section 4 is devoted to the discussion of the conjecture of Fox, Kwan and Sauermann and results for odd values of \( n \). Finally, we present and discuss some open problems in Section 5.

## 2 Balancing lemma and its consequences

We shall use the notation \( X \sim Y \) to denote the fact that the random variables \( X \) and \( Y \) have the same distribution. We shall now proceed with an elementary proof of the balancing lemma, which also works for random summands taking values in a countable subset of an Abelian group.

**Proof of Lemma 1** Let us split the sum into two halves:

\[
S = \sum_{i=1}^{n} X_i \quad \text{and} \quad T = \sum_{i=n+1}^{2n} X_i.
\]

By the inequality of arithmetic and geometric means

\[
\mathbb{P}(S + T = 0) = \sum_{x} \mathbb{P}(S = x) \mathbb{P}(-T = x) \leq \sum_{x} \frac{\mathbb{P}(S = x)^2 + \mathbb{P}(-T = x)^2}{2} \\
= \frac{1}{2} \sum_{x} \mathbb{P}(S = x)^2 + \frac{1}{2} \sum_{x} \mathbb{P}(-T = x)^2 \\
\leq \max\{\sum_{x} \mathbb{P}(S = x)^2, \sum_{x} \mathbb{P}(T = x)^2\} \\
= \max\{\mathbb{P}(S - S' = 0), \mathbb{P}(T - T' = 0)\}.
\]

(3)
Here $S'$ and $T'$ are independent copies of $S$ and $T$ respectively. Note that for non-negative $p$ and $q$, $pq \leq \frac{p^2 + q^2}{2}$ and we have an equality if and only if $p = q$. Therefore (3) is equality if and only if $S \sim -T$.

Given a sequence of random variables $(Z_i, i \in \{1, \ldots, 2n\})$ consider equivalence classes (types) defined by the equivalence relation $Z_i \sim Z_j$ or $Z_i \sim -Z_j$. If all of $\{X_1, \ldots, X_{2n}\}$ have the same type, the proof follows by (3).

If there are more than two types, let $X_1 = \{X_{i_1}, \ldots, X_{i_k}\}$ and $X_2 = \{X_{j_1}, \ldots, X_{j_l}\}$ be different classes other than the largest equivalence class (break ties arbitrarily). Clearly $k \leq n$ and $l \leq n$. Rearrange the variables so that all the variables in $X_1$ are in $S$ and all the variables $X_2$ are in $T$. Applying (3) yields a new sequence of random variables $Y_1, \ldots, Y_{2n}$, $Y_{2k} \sim -Y_{2k-1}$, $k \in \{1, \ldots, n\}$ which has at least one less type and

$$P(X_1 + \cdots + X_{2n} = 0) \leq P(Y_1 + \cdots + Y_{2n} = 0).$$

By repeating this argument at most $2n$ times, we reduce the number of types to one or two.

It remains to consider the case when there are exactly two types among $X_1, \ldots, X_{2n}$. Repeatedly apply (3) by rearranging the sequence so that the first half $S$ contains only the variables of the largest type. Stop when either a single type remains or after at most $n$ steps the first cycle $(S_1, T_1), (S_2, T_2), \ldots, (S_k, T_k)$ is formed, i.e. the two halves $(S_k, T_k)$ after some step have the same distribution as the two halves $(S_1, T_1)$ in a previous step. We have

$$P(X_1 + \cdots + X_{2n} = 0) \leq P(S_1 + T_1 = 0) \leq \cdots \leq P(S_k + T_k = 0),$$

which implies $P(S_1 + T_1 = 0) = \cdots = P(S_k + T_k = 0)$, and so $T_1 \sim -S_1$.

If we applied (3) at least once where $S \sim -T$ does not hold, the final inequality is strict. Otherwise we have that $\sum X_i \sim \sum (-1)^{i+1} Y_i$. ■

**Proof of Corollary 1** Apply Lemma 1 to independent random variables $a_i X_i$. ■

Part of the early inspiration for Lemma 1 came from a simple observation of a *math.stackexchange* user André Nicolas about simple symmetric random walks [15].

### 3 Random variables with bounded concentration

Let $X_1, \ldots, X_n$ be independent random vectors in $\mathbb{R}^d$ and denote their sum by $S_n$. Assume that for all $i$ we have

$$\sup_{x \in \mathbb{R}^d} P(X_i = x) \leq \alpha < 1.$$

The bounds on the concentration probability $P(S_n = x)$ were studied by many authors. Let us just mention the work of Esseen [6], Rogozin [16] and Gamerklidze [11]. It was proved by Rogozin that when $d = 1$ and $\alpha = \frac{1}{k}$ for some integer $k$ that $P(S_n = x)$ is
maximized when all $X_i$ are iid uniform random variables in the set $\{0, \ldots, k - 1\}$. This result also follows from more general bounds obtained by Leader and Radcliffe [13]. To our knowledge the sharpest known bounds for $\alpha \in \left(\frac{1}{2}, 1\right)$ and all $d$ were obtained by Ushakov [20]. Such $\alpha$ are especially interesting as it covers all Bernoulli distributions. Ushakov established the inequality
\[
P(S_n = x) \leq \left(2\pi(n + 1)\alpha(1 - \alpha)\right)^{-\frac{1}{2}}(1 + \left(\pi(n + 1)\alpha(1 - \alpha)\right)^{-\frac{1}{2}}
\]
which is asymptotically sharp: this can be seen by Lemma 9 below (with the correct second order term), or alternatively by using the Local Central Limit Theorem.

Let us give a short description strategy of the proof of Theorem 1. Firstly, we characterize the extremal points of the convex set of distributions with a bound on their maximal probability. We then make use of the result of Ushakov [20] to reduce the problem from high dimensions to integer valued random variables. Having narrowed down the class of distributions we use the balancing lemma. The latter step produces a sum of symmetric distributions and we then proceed by using an old rearrangement inequality for convolutions of sequences proved by Gabriel [10] which we fortunate to find in the classical monograph of Hardy, Littlewood and Pólya [9]. After the latter operation the random variables under consideration become symmetric and unimodal. The final touch is to use a discrete analogue of Birnbaum’s result from [1] on peakedness of symmetric unimodal random variables which intuitively compresses the mass of the underlying distributions to the center as much as it is possible. This discrete version of Birnbaum’s result is probably known, but we could not find it in the literature, so we provide a simple proof for the readers convenience. Having outlined the strategy we shall step by step introduce the relevant notions and results until we can then combine them and finish the proof in a few lines.

For any probability measure $\mu$ on a finite set $X \subset \mathbb{R}^n$ define its concentration to be the quantity
\[
Q(\mu) = \max_{x \in X} \mu\{x\}.
\]
Notice that $Q$ is a convex function and that the set of measures $S_\alpha = \{\mu|Q(\mu) \leq \alpha\}$ is convex. Given $\alpha \in (0, 1)$, a set $A \subseteq X$ with $|A| = \lfloor \alpha^{-1} \rfloor$ and $y \in X \setminus A$, let us denote by $\mu_{\alpha,A,y}$ the probability measure on $X$ such that
\[
\mu_{\alpha,A,y}\{x\} = \begin{cases} 
\alpha, & \text{for } x \in A, \\
1 - \lfloor \alpha^{-1} \rfloor \alpha, & \text{for } x = y, \\
0, & \text{otherwise.}
\end{cases}
\]
(When $\alpha^{-1}$ is integer it is equal to $|A|$ and the parameter $y$ becomes dummy variable: $\mu_{\alpha,A,y}$ is the uniform measure on $A$ for any $y \in X \setminus A$.) We shall say that a convex combination $p\mu + (1 - p)\nu$ of two distinct measures $\mu$ and $\nu$ on $X$ is non-trivial if $0 < p < 1$. 
Lemma 2. Let \( \alpha \in (0, 1) \) and let \( \mu \) be a measure in \( S_\alpha \). Then \( \mu \) can be written as a non-trivial convex combination of two distinct measures in \( S_\alpha \) unless it is a measure \( \mu_{\alpha, A,y} \) for some \( A \subseteq X \) and \( y \in X \).

Proof. First let us show that if \( \mu = \mu_{\alpha, A,y} \), then it cannot be decomposed. Assume the contrapositive: that \( \mu = p\mu_1 + (1-p)\mu_2 \) for distinct measures \( \mu_1, \mu_2 \in S_\alpha \) and \( 0 < p < 1 \). It follows that both measures have support on \( \{ A \} \cup \{ y \} \) if \( \alpha^{-1} \) is not integer and on \( \{ A \} \) otherwise. In the latter case all measures are equal, a contradiction. In the former case there is \( x \in A \) such that \( \mu \{ x \} = \mu_1 \{ x \} = \alpha \) and \( \mu_2 \{ x \} = 1 - |\alpha^{-1}| \alpha < \alpha \). So \( p\mu_1 \{ x \} + (1-p)\mu_2 \{ x \} < \alpha = \mu \{ x \} \), also a contradiction. Now assume that \( \mu \) is not of the form \( \mu_{\alpha, A,y} \). Let \( A \) be the set of \( \lfloor \alpha^{-1} \rfloor \) largest atoms of \( \mu \) and let \( y \) be any of its largest atoms outside \( A \). Since \( \mu \in S_\alpha \), its support is of size at least \( |\alpha^{-1}| + 1 \), in particular it has a non-zero mass on each \( x \in A \cup \{ y \} \). Let \( \mu_2 = \mu_{\alpha, A,y} \). Fix a positive \( \varepsilon \) small enough that \( (1+\varepsilon)\mu \{ x \} \geq \varepsilon((1+\varepsilon)\mu - \epsilon\mu_2) \) for \( x \in A \cup \{ y \} \) and \( (1+\varepsilon) \leq \alpha(\lfloor \alpha^{-1} \rfloor + 1) \). Define \( \mu_1 = (1+\varepsilon)\mu - \epsilon\mu_2 \). We have \( \mu = p\mu_1 + (1-p)\mu_2 \) with \( p = \frac{1}{\epsilon} \). Since \( \mu_1 = \mu + \epsilon(\mu - \mu_2) \) and \( \mu \neq \mu_2 \), \( \mu_1 \) and \( \mu_2 \) are distinct, and so \( \mu_1 \) and \( \mu_2 \) also must be distinct. Let us now check that \( \mu_1 \in S_\alpha \). For \( x \in A \) we have \( \mu_1 \{ x \} = \mu \{ x \} + \epsilon(\mu \{ x \} - \alpha) \leq \mu \{ x \} \leq \alpha \). By the choice of \( A \) and \( y \), for each \( x \in X \setminus A \), \( \mu \{ x \} \leq \mu \{ y \} \leq (|A|+1)^{-1} < \alpha \). Thus \( \mu_1 \{ x \} \leq (1+\varepsilon)\mu \{ x \} \leq (1+\varepsilon)(|A|+1)^{-1} \leq \alpha \) for \( x \in X \setminus A \). 

In his work on the problem of this section Ushakov [20] proved a couple reduction lemmas that allows to switch from distributions in Hilbert spaces to distributions on the integers. We shall state the relevant one we require here.

Lemma 3. Let \( \mu_1, \ldots, \mu_n \) be probability distributions in some Hilbert space such that

\[
Q(\mu_i) \leq \alpha.
\]

Then there exist probability distributions \( \nu_1, \ldots, \nu_n \) on \( \mathbb{Z} \) such that \( Q(\nu_i) \leq \alpha \) and

\[
Q(\mu_1 \ast \ldots \ast \mu_n) \leq Q(\nu_1 \ast \ldots \ast \nu_n),
\]

where \( \ast \) stands for convolution.

In two important parts of the proof we shall use rearrangement results from [10] (see also [9] page 273, inequality 374). First let us define certain special rearrangements of a collection of non-negative numbers \( \{a_{-k}, \ldots, a_k\} \) indexed by integers. The rearrangement \( (+a) \) is defined by inequalities \( a_0 \geq a_{-1} \geq a_1 \geq a_{-2} \geq \ldots \geq a_k \). Analogously, the rearrangement \( (a^+) \) is defined by inequalities \( a_0 \geq a_1 \geq a_{-1} \geq a_2 \geq a_{-2} \geq \ldots \geq a_k \). Finally, if in the sequence \( (a) \) all values except the largest one appear in pairs, we define the symmetric decreasing rearrangement \( (a^*) \) by the inequalities \( a_0^* \geq a_1^* = a_{-1}^* \geq a_2^* \geq a_{-2}^* \geq \ldots \geq a_{-k}^* = a_k^* \). When \( a_i = \mathbb{P}(X = i) \) for a random variable \( X \), we will write for brevity \( \mathbb{P}(X = i)^+ = a_i^+ \), etc.
Lemma 4. Let \((a), (b), (c), (d), \ldots\) be collections of non-negative numbers such that all collections except maybe \((a)\) and \((b)\) have a symmetric decreasing rearrangement. Then

\[
\sum_{r+s+t+u+\ldots=0} a_r b_s c_t d_u \ldots \leq \sum_{r+s+t+u+\ldots=0} \left( a_r b_s^* c_t^* d_u^* \right) \ldots.
\]

The final tool we shall require is a discrete counterpart of Birnbaum’s [1] result on the peakedness of symmetric unimodal distributions.

Lemma 5. Let \(X, Y\) and \(Y'\) be independent symmetric unimodal integer random variables. Suppose \(\mathbb{P}(Y \in [-k, k]) \leq \mathbb{P}(Y' \in [-k, k])\) for any integer \(k\). Then for any integer \(k\)

\[
\mathbb{P}(X + Y \in [-k, k]) \leq \mathbb{P}(X + Y' \in [-k, k])
\]

Proof. \(|Y'|\) is stochastically dominated by \(|Y|\), so let us assume \(Y'\) and \(Y\) are coupled so that conditioned on \(Y = y\), \(Y'\) is zero or of the same sign as \(y\) and \(|Y'|\) \(\leq |y|\).

Since \(X\) is symmetric and unimodal, if \(0 \leq y' \leq y\) or \(y \leq y' \leq 0\)

\[
\mathbb{P}(X \in [y - k, y + k]) \leq \mathbb{P}(X \in [y' - k, y' + k]).
\]

Therefore

\[
\mathbb{P}(X + Y \in [-k, k]) = \mathbb{P}(X - Y \in [-k, k]) = \mathbb{P}(X \in [Y - k, Y + k]) = \mathbb{E} \mathbb{E} (\mathbb{1}_{X \in [Y - k, Y + k]} | Y) \leq \mathbb{E} \mathbb{E} (\mathbb{1}_{X \in [Y' - k, Y' + k]} | Y) = \mathbb{P}(X + Y' \in [-k, k]).
\]

Since the class of symmetric unimodal distributions is closed under convolution (see e.g. [9]) the latter result easily carries over to an arbitrary number of random variables.

Corollary 2. Let \(X_1, \ldots, X_n, Y_1, \ldots, Y_n\) be independent symmetric unimodal integer random variables so that for \(i \leq n\) and all integers \(k\) we have \(\mathbb{P}(X_1 \in [-k, k]) \leq \mathbb{P}(Y \in [-k, k])\). Then for all integers \(k\) we have

\[
\mathbb{P}(X_1 + \cdots + X_n \in [-k, k]) \leq \mathbb{P}(Y_1 + \cdots + Y_n \in [-k, k]).
\]

Proof of Theorem 1. Lemma 3 tells us that in order to maximize \(\mathbb{P}(X_1 + \ldots + X_{2k} = x)\) it is sufficient to consider integer random variables \(X_i\) such that \(\mathbb{P}(X_i = x) \leq \alpha\) for all \(x \in \mathbb{Z}\). We can without loss of generality also assume that all distributions of the random variables \(X_i\) are finitely supported; the general case follows by approximating with truncated random variables. The Krein-Milman theorem tells us that the convex set of distributions \(\mu\) on a finite set \(\mathcal{X} \subset \mathbb{Z}\) such that \(Q(\mu) \leq \alpha\) is the closure of the convex hull of its extreme points. These extreme points are described by Lemma 2. Since the probability \(\mathbb{P}(X_2 + \ldots + X_{2k} = x)\) is linear in each distribution, we can assume that
the maximum is attained when each \( X_i \) the a distribution \( \mu_{\alpha,A,y_i} \) for some \( A \subset \mathbb{Z} \) and \( y_i \in \mathbb{Z} \setminus A \). Applying the balancing lemma we thus obtain that

\[
\mathbb{P}(X_1 + \ldots + X_{2k} = x) \leq \mathbb{P}(Y_1 - Y_2 + \ldots - Y_{2k-1} + Y_{2k} = 0),
\]

where \( Y_j \) are iid random variables distributed as some \( X_i \). Let us denote the distribution of that particular \( X_i \) by \( \mu_{\alpha,A,y_i} \) (dropping the subscript of \( y_i \)). The random variables \( Z_j = Y_{2j-1} - Y_{2j} \) are iid and have a symmetric distributions. Denote by \( Z_i^* \) a sequence of iid random variables that have distributions on the integers obtained by a symmetric decreasing rearrangement of the probabilities \( \mathbb{P}(Z_i = m) \). Applying Lemma 4 we obtain

\[
\mathbb{P}(Y_1 - Y_2 + \ldots - Y_{2k-1} + Y_{2k} = 0) = \mathbb{P}(Z_1 + \ldots + Z_k = 0) = \sum_{m_1,\ldots,m_k=0} \mathbb{P}(Z_1 = m_1) \ldots \mathbb{P}(Z_k = m_k) \leq \sum_{m_1,\ldots,m_k=0} \mathbb{P}(Z_1 = m_1)^* \ldots \mathbb{P}(Z_k = m_k)^* = \mathbb{P}(Z_1^* + \ldots + Z_k^* = 0).
\]

We have now achieved an inequality for the probability in question in terms of symmetric unimodal distributions to which Lemma 5 applies. All that is left to prove is the stochastic domination condition \( \mathbb{P}(|Z_i^*| \leq l) \leq \mathbb{P}(|U_{2j-1} - U_{2j}| \leq l) \) for all integer \( l \). We shall actually show that for all \( B \subset \mathbb{Z} \) with \( |B| = 2l + 1 \) we have

\[
\mathbb{P}(|Z_j^*| \leq l) = \mathbb{P}(Z_j \in B) \leq \mathbb{P}(|U_{2j-1} - U_{2j}| \leq l).
\]

The first equality follows from the definition of the symmetric decreasing rearrangement. For the inequality we will use Lemma 4 again.

Let us denote by \( U(-B) \) a uniform random variable that is independent of the collection \( (Z_i) \). In case \( B = \{-l, \ldots, l\} \) we shall denote this random variable by \( U \). Having in mind that \( X_i \) have distribution \( \mu_{\alpha,A,y_i} \) and using Lemma 4 we obtain

\[
(2l + 1)\mathbb{P}(Z_j \in B) = \mathbb{P}(X_{2i-1} - X_{2i} + U(-B) = 0) = \sum_{r+s+t=0} \mathbb{P}(X_{2i-1} = r)\mathbb{P}(X_{2i} = s)\mathbb{P}(U(-B) = t) \leq \sum_{r+s+t=0} \mathbb{P}(X_{2i-1} = r)\mathbb{P}(X_{2i} = s)^*\mathbb{P}(U(-B) = t)^* = \sum_{r+s+t=0} \mathbb{P}(U_{2i-1}^\alpha - \lfloor (l-1)/2 \rfloor = r)\mathbb{P}(-U_{2i}^\alpha + \lfloor (l-1)/2 \rfloor = s)\mathbb{P}(U = t) = \mathbb{P}(U_{2i-1}^\alpha - U_{2i}^\alpha + U = 0) = (2l + 1)\mathbb{P}(U_{2i-1}^\alpha - U_{2i}^\alpha \in \{-l, \ldots, l\})
\]

and we are done.
4 The Littlewood-Offord problem for Bernoulli distributions

In this section we consider the Littlewood-Offord problem in the case $a_1, \ldots, a_n \in \mathbb{R}^d \setminus \{0\}$ and $X_1, \ldots, X_n$ are iid Bernoulli with parameter $p$ ($p = 1 - \alpha$), a positive integer $d$ and $p \in (0, \frac{1}{2})$ will be fixed.

Define
\begin{equation}
T_n = T_{n,p} = X_1 - X_2 + \cdots + (-1)^{n+1}X_n.
\end{equation}

Note that $T_n \sim \text{Binom}(\lfloor \frac{n}{2} \rfloor, p) - \text{Binom}(\lfloor \frac{n}{2} \rfloor, p)$, where by a difference of distributions we denote the distribution of the difference of independent random variables from the corresponding distributions.

By Corollary 1(a) for even $n$ we have
\[ \mathbb{P}(\sum a_i X_i = x) \leq \mathbb{P}(T_n = 0). \]

The situation for odd $n$ is more subtle. We are still able to prove the following.

Lemma 6. Let $p \in (0, \frac{1}{2})$. For all $n$ large enough the following holds. If $X_1, X_2, \ldots, X_n$ are independent Bernoulli random variables with parameter $p$, $a_1, \ldots, a_n \in \mathbb{R}^d \setminus \{0\}$ and $x \in \mathbb{R}^d$ then
\begin{equation}
\mathbb{P}(\sum a_i X_i = x) \leq \max_{0 \leq k \leq n/2} \mathbb{P}(B_{n-k} - B_k' = 0),
\end{equation}
where $B_{n-k}, B_k'$ are independent binomial random variables with parameters $(n-k, p)$ and $(k, p)$.

By Corollary 1(a) for even $n$ the only value that maximizes the right side of (5) is $k = \frac{n}{2}$, and this holds for any $n$. For odd $n$ we cannot explicitly describe the optimal $k$ in (5). As $B_{n-k} - B_k'$ has the same order asymptotic growth for any $k = k(n)$, the answer requires maximizing the second order asymptotic term over all $k$. It seems that for a given $n$ each $k \in \{0, \ldots, \frac{n-1}{2}\}$ can be optimal depending on $p$ in a complicated way (we confirmed this using distr package of R [17] for, e.g., $n \leq 31$).

We only prove that neither of $k \in \{0, \frac{n-3}{2}, \frac{n-1}{2}\}$ dominates for all $p \in (0,1)$. For simplicity we only consider integer $np$.

Proposition 2. Let $n$ be a positive odd integer, $n \geq 3$ and $p \in (0, \frac{1}{2})$ such that $np$ is integer. Consider $T_n = T_{n,p}$ as in (4), $Q_n \sim \text{Binom}((n+3)/2, p) - \text{Binom}((n-3)/2, p)$ and $B_n \sim \text{Binom}(n, p)$. Let $p_1 = (13 - \sqrt{15})/44 \approx 0.21$, $p_2 = (13 + \sqrt{15})/44 \approx 0.38$.

For all $n$ large enough, the first, the second and the third is strictly largest among $\mathbb{P}(T_n = 0)$, $\mathbb{P}(Q_n = 1)$ and $\mathbb{P}(B_n = np)$ if $p$ is in the interval $(0, \frac{1}{4})$, $(\frac{1}{4}, p_2)$ and $(p_2, \frac{1}{2})$ respectively. $\{0\}$, $\{1\}$ and $\{np\}$ is the largest atom of $T_n$, $Q_n$, and $B_n$ respectively in the corresponding interval.

Furthermore, if $p \in (p_1, p_2)$ and $n$ is large enough then $\mathbb{P}(Q_n = 1) > \mathbb{P}(B_n = np)$. 

10
It was conjectured in the first version of [7] that with an additional restriction \( x \neq 0 \)

\[
\mathbb{P}(\sum_{i=1}^{n} a_i X_i = x) \leq \max_{x \neq 0} \max_{k \neq 0} \mathbb{P}(B_n = x) \quad \text{where} \quad B_n \sim \text{Binom}(n, p).
\]

Darroch [2] showed that a finite sum of independent Bernoulli random variables with mean \( \mu \) has mode equal to \( \lfloor \mu \rfloor \) or \( \lceil \mu \rceil \). Thus Proposition 2 shows that the above conjecture is not true for any \( p \in (p_1, p_2) \) and all large enough odd \( n \) such that \( np \) is integral. There are many more counterexamples, one of the small ones is \( n = 4, a_1 = a_2 = a_3 = -a_4 = 1, x = 1 \) and \( p \in (\sqrt{2}/4, 1/2) \).

Lemma [6] follows from a slightly stronger result, Lemma [7] which we prove next. Recall random vectors \( X \) and \( Y \) have the same type if and only if either \( Y \) or \( -Y \) has the same distribution as \( X \). Similarly we say that \( x, y \in \mathbb{R}^d \) have the same type if and only if \( x \in \{ -y, y \} \).

Lemma 7. Let \( p \in (0, \frac{1}{2}) \). Let \( X_1, \ldots, X_n \) be independent Bernoulli random variables with parameter \( p \). Let \( T_n \) be as in [7].

There is a sequence \( \delta_n = o(n^{-1}) \) such that for any \( x_1, \ldots, x_n \in \mathbb{R}^d \) with at least two types and any \( x \in \mathbb{R}^d \)

\[
\mathbb{P}(\sum_{i=1}^{n} x_i X_i = x) \leq \mathbb{P}(T_n = 0)(1 - (2n)^{-1} + \delta_n).
\]

We will need asymptotics for small deviations of a binomial random variable convolved with its negation.

Lemma 8. Let \( p \in (0, \frac{1}{2}) \). Let \( k \) be a positive integer. Let \( T_{2n} = T_{2n,p} \sim \text{Binom}(n, p) - \text{Binom}(n, p) \). Then

\[
\frac{\mathbb{P}(T_{2n} = k)}{\mathbb{P}(T_{2n} = 0)} = 1 - \frac{k^2}{4p(1-p)n}(1 + o(1)).
\]

Proof. We have

\[ T_{2n} = X_1 + \ldots X_n - X'_1 - \ldots - X'_{n}, \]

where \( X_j, X'_j, j \in \{1, \ldots, n\} \) are independent Bernoulli with parameter \( p \).

On the other hand \( X_j - X'_j \) is distributed as a mixture of 0 (with probability \( 1 - p' = (1 - 2p)^2 \)) and \( \tilde{X}_{2j} + \tilde{X}_{2j+1} - 1 \) (with probability \( p' = 4p(1-p) \)). Here \( \tilde{X}_{2j} \) and \( \tilde{X}_{2j+1} \) are two independent Bernoulli random variables with parameter \( \frac{1}{2} \).

Thus

\[ T_{2n} = -N + B_{2N} \quad \text{where} \quad N \sim \text{Binom}(n, 4p(1-p)) \]

and conditioned on \( N = t \), \( B_{2N} \sim \text{Binom}(2t, \frac{1}{2}) \).
Conditioned on $N = t$, $T_{2n}$ is unimodal and symmetric with mode at 0. For any positive integer $k$

$$
\mathbb{P}(T_{2n} = k | N = t) = \binom{2n}{t+k} \frac{(t - k + 1) \ldots t}{(t + 1) \ldots (t + k)} = \frac{\prod_{j=0}^{k-1}(1 - j/t)}{\prod_{j=1}^{k}(1 + j/t)}
$$

$$
= (1 - \sum_{j=1}^{k-1} \frac{j}{t} + O(t^{-2}))(1 - \sum_{j=1}^{k} \frac{j}{t} + O(t^{-2})) = 1 - \frac{k^2}{t} + O(t^{-2}).
$$

Fix $C$ large enough so that by the concentration of the binomial random variable $N \mathbb{P}(\bar{A}_C) = o(n^{-3/2})$ (see, e.g., [14]). Here $A_C$ is the event $|N - np'| \leq C\sqrt{n \ln n}$.

For any $t$ such that $A_C$ holds on $N = t$

$$
\mathbb{P}(T_{2n} = 2k | N = t) = \mathbb{P}(T_{2n} = 0 | N = t) (1 - \frac{k^2}{np}(1 + o(1)))
$$

where the constant in $o()$ depends only on $k$ and $C$. By Stirling’s approximation we have

$$
\mathbb{P}(T_{2n} = 0) \geq \mathbb{P}(T_{2n} = 0 | A_C) \mathbb{P}(A_C) = \Omega(n^{-\frac{1}{2}}).
$$

So

$$
\mathbb{P}(T_{2n} = k) = \mathbb{E} \mathbb{E}(X_{A_C}X_{T_{2n} = k} | N) + \mathbb{E} \mathbb{E}(X_{A_C}X_{T_{2n} = k} | N)
$$

$$
\mathbb{E} \mathbb{E}(X_{A_C}X_{T_{2n} = k} | N) \leq \mathbb{P}(\bar{A}_C) = o(n^{-1}\mathbb{P}(T_{2n} = 0));
$$

$$
\mathbb{E} \mathbb{E}(X_{A_C}X_{T_{2n} = k} | N) \geq \mathbb{E} \mathbb{E}(X_{A_C}X_{T_{2n} = 0} | N) \left(1 - \frac{k^2}{np} + o(n^{-1})\right)
$$

$$
\geq (\mathbb{P}(T_{2n} = 0) - \mathbb{P}(\bar{A}_C)) \left(1 - \frac{k^2}{np} + o(n^{-1})\right)
$$

$$
= \mathbb{P}(T_{2n} = 0) \left(1 - \frac{k^2}{np} + o(n^{-1})\right)
$$

which gives

$$
\mathbb{P}(T_{2n} = k) = \mathbb{P}(T_{2n} = 0) \left(1 - \frac{k^2}{np}(1 + o(1))\right).
$$

Proof of Lemma 7 Let $S_n = \sum_{i=1}^{n} x_i X_i$. Then

$$
\mathbb{P}(S_n = x) = (1 - p)\mathbb{P}(S_{n-1} = x) + p\mathbb{P}(S_{n-1} = x - x_n) \leq \mathbb{P}(S_{n-1} = x')
$$

for some $x' \in \{x, x-x_n\}$. Let $m = \lfloor \frac{n}{2} \rfloor$ and consider $S_{2m}$. Let $n_a = |\{x_i : i \leq 2m \text{ and } x_i \in \{-a,a\}\}|$. Let $c$ maximize $n_a$ over $a$. Suppose first $n_c \leq 2m - 2$. 

12
Let \( x' \in \mathbb{R}^d \) be arbitrary. By (3) applied with \( X_i = x_iX_i \), \( i = 1, \ldots, 2m - 1 \) and \( X_{2m} = x_{2m}X_{2m} - x' \) placing exactly \( \left\lceil \frac{n}{c} \right\rceil \) of the terms with \( x_i \in \{-c, c\} \) into the first half, we get:

\[
P(S_{2m} = x') \leq P(\sum_{i=1}^{2m} x_i'X_i = 0)
\]

where each constant \( \pm a \) occurs an even number of times \( n' a \) in \( \{x_i'\} \) and \( 2 \leq n_a \leq 2m - 2 \), so there are at least two equivalence classes (types) of random variables in the resulting sum.

We claim that we can keep applying (3) as in the proof of Lemma 1 and stop when exactly two types of random variables remain.

If there are at least four equivalence classes, place half of the \( n_1 \) variables from the largest equivalence class into the first half of the sum, and the remaining \( \frac{n_1}{2} \) variables into the second half of the sum. Note that we can then place the variables from the smallest and the second smallest equivalence classes into different halves, so that after an application of (3) the number of classes is reduced by at least one and remains at least two.

If there are three equivalence classes, arrange the sum so that the \( n_2 \) terms from the second biggest class go first, then the \( n_1 \) terms from the biggest class and finally the \( n_3 \) terms from the smallest class. We have \( n_1 + n_2 \geq 2\frac{2m}{3} \). As \( n_1, n_2, n_3 \geq 2 \), \( n_2 \leq (2m - n_3)/2 \leq m - 1 \). Hence the first half contains all \( n_2 \) elements from the second biggest class, and at least one element from the biggest class, while the second half contains \( n_1 + n_2 - m \geq \frac{4m}{3} \) elements of the biggest class and all \( n_3 \) elements from the smallest class. Applying (3) one more time, exactly two equivalence classes remain. Hence after at most \( n \) applications of (3) in total we get:

\[
P(S_{2m} = x') \leq \pi(a, b, k) := P(aT_{2k} + bT'_{2(m-k)} = 0)
\]

for \( a, b \in \mathbb{R}^d \setminus \{0\}, a \not\in \{-b, b\}, k, m - k \geq 1 \) and \( T_{2k}, T'_{2(m-k)} \) are independent, \( T'_{2(m-k)} \sim T_{2(m-k)} \).

Denote by \( S'_{2m} \) the sum of \( 2m \) Bernoulli\((p)\) random variables with multipliers \( a, -a, b \) and \( -b \) corresponding to the right side of (6). We claim that

\[
\pi(a, b, k) \leq \min(\pi(a, b, 1), \pi(a, b, m - 1)).
\]

To see this, keep applying (3) as in the proof of Lemma 1 starting with \( S'_{2m} \) by taking exactly one term in the smaller equivalence class in the first half \( S \). Either at some point we get that \( \pi(a, b, k) \leq P(S - S' = 0) \), or we obtain a cycle in a finite number of applications, where again as in the proof of Lemma 1 we get that \( T \sim -S \) so that \( \pi(a, b, k) \leq P(S - S' = 0) \).

\(^1\)Here the \( d \)-dimensional zero vector is also denoted as \( 0 \).
Thus by possibly swapping \(a\) and \(b\)

\[
\Pr(S_{2m} = x') \leq \Pr(aT_{2(m-1)} + bT'_2 = 0),
\]

Let \(X = T_{2(m-1)}\) and \(Y = T'_2\). Note that whenever \(b \neq ra\) for some \(r \in \mathbb{Z}\), \(aX + bY = 0\) if and only if \(X = 0\) and \(Y = 0\), whereas if \(b = 2a\) and \(m \geq 3\) there are additional possibilities for \(aX + bY = 0\), so we can assume \(b = ra\) for \(r \in \mathbb{Z} \setminus \{-1, 0, 1\}\), which reduces the right part of the last inequality to the one-dimensional case \(a = 1\) and \(b = r\).

Furthermore, \(X\) is symmetric and (strongly) unimodal with mode 0 \([2]\). Therefore, if \(r \in \mathbb{Z}\) and \(|r| > 2\), we have

\[
\Pr(X + rY = 0) = \sum_k \Pr(X = -rk)\Pr(Y = k)
\leq \sum_k \Pr(X = -2k)\Pr(Y = k) = \Pr(X + 2Y = 0).
\]

In each case

\[
\Pr(S_n = x) \leq \max_{x'} \Pr(S_{2m} = x') \leq \Pr(X + 2Y = 0).
\]

By Lemma \([3]\)

\[
\Pr(X = k) = \Pr(X = -k) = \Pr(X = 0) \left(1 - \frac{k^2}{2p(1-p)n} + o(n^{-1})\right).
\]

For odd \(n\) let us now compare \(\Pr(X + 2Y = 0)\) with \(\Pr(X + Y' = 0)\) where \(Y' = X'_1 - X'_2 + X'_3\) and \(X'_1, X'_2, X'_3\) are independent Bernoulli\((p)\) random variables independent of \(X\). Note that \(X + Y' \sim T_n\).

\(2Y\) is symmetric and distributed on \([-2, 0, 2]\) with \(\Pr(2Y = 0) = p^2 + (1-p)^2\). Therefore by \([3]\)

\[
\frac{\Pr(X + 2Y = 0)}{\Pr(X = 0)} = \left(2p(1-p)(1 - \frac{4}{2p(1-p)n} + o(n^{-1})) + (p^2 + (1-p)^2)\right)
= (1 - \frac{4}{n} + o(n^{-1})).
\]

\(Y'\) is distributed on \([-1, 0, 1, 2]\) with probabilities \(p(1-p)^2,\ 2p^2(1-p) + (1-p)^3,\ 2p(1-\
\[ p^2 + p^3 \text{ and } p^2(1 - p) \text{ respectively. Therefore by (9) and symmetry of } X \]
\[
\mathbb{P}(X + Y' = 0) = \mathbb{P}(X = 1)\mathbb{P}(Y' = -1) + \mathbb{P}(X = 0)\mathbb{P}(Y' = 0) + \mathbb{P}(X = -1)\mathbb{P}(Y' = 1) + \mathbb{P}(X = -2)\mathbb{P}(Y' = 2) = \\
\left((3p(1 - p)^2 + p^3)(1 - \frac{1}{2p(1 - p)n}) + 2p^2(1 - p) + (1 - p)^3 + \\
p^2(1 - p)(1 - \frac{4}{2p(1 - p)n}) + o(n^{-1})\right)\mathbb{P}(X = 0) \\
= \left(1 - \frac{3 - 2p}{2n(1 - p)} + o(n^{-1})\right)\mathbb{P}(X = 0).
\]

Finally, combining (8) and the last two bounds
\[
\mathbb{P}(S_n = x) \leq (1 - \frac{4}{n} + o(n^{-1}))\mathbb{P}(X = 0) = (1 - \frac{4}{n} + o(n^{-1}))\times \\
\mathbb{P}(T_n = 0) \left(1 - \frac{3 - 2p}{2n(1 - p)} + o(n^{-1})\right)^{-1} \\
= \mathbb{P}(T_n = 0) \left(1 - \frac{5 - 6p}{2n(1 - p)} + o(n^{-1})\right).
\]

If \( n_c = 2m - 1 \) and \( x_n \) is not in the largest class (type), we can exchange \( x_n \) with a constant from the largest type and apply the proof for \( n_c = 2m - 2 \). Finally, if all but one \( \{x_i\} \) is of the same type
\[
\mathbb{P}(S_n = x) \leq \max\{\mathbb{P}(T_{2m} + 2X_1' = 0), \mathbb{P}(T_{2m} + 2X_1' = 1)\}
\]

By Lemma (8) and similar calculations as above
\[
\mathbb{P}(T_{2m} + 2X_1' = 0) = \mathbb{P}(T_n = 0) \left(1 - \frac{3}{2(1 - p)n} + o(n^{-1})\right); \\
\mathbb{P}(T_{2m} + 2X_1' = 1) = \mathbb{P}(T_n = 0) \left(1 - \frac{1}{2pn} + o(n^{-1})\right).
\]

Now assume \( n \) is even. If \( n_c < n - 1 \), we have \( \mathbb{P}(X + X_1' - X_2' = 0) \geq \mathbb{P}(X + Y' = 0) \) so \( \mathbb{P}(S_n = x) \leq \mathbb{P}(T_n = 0)(1 - \frac{5 - 6p}{2n(1 - p)} + o(n^{-1})) \) by the above calculation. The case \( n \) is even and \( n_c = n - 1 \) remains.
\[
\mathbb{P}(S_n = x) \leq \max_x \mathbb{P}(aT_{n-1} + bX_1' = x) \leq \max_{y, r \in \mathbb{Z} \setminus \{-1, 0, 1\}} \mathbb{P}(T_{n-1} + rX_1' = y)
\]

Since \( T_{n-2} \) is unimodal, it is easy to see that \( T_{n-1} \) is also unimodal and
\[
\mathbb{P}(T_{n-1} = 0) > \mathbb{P}(T_{n-1} = 1) > \mathbb{P}(T_{n-1} = -1) > \mathbb{P}(T_{n-1} = 2) > \ldots \quad (10)
\]
This implies that the maximum is attained with \( r = -2 \) and \( y = 0 \) or \( r = 2 \) and \( y = 1 \).

Using Lemma 8 and simple calculation

\[
\mathbb{P}(T_{n-1} - 2X'_1 = 0) = \mathbb{P}(T_{n-2} = 0)(1 - \frac{5p - 4p^2}{2p(1-p)n} + o(n^{-1}));
\]

\[
\mathbb{P}(T_{n-1} + 2X'_1 = 1) = \mathbb{P}(T_{n-2} = 0)(1 - \frac{1 - p + 4p^2}{2p(1-p)n} + o(n^{-1}));
\]

\[
\mathbb{P}(T_n = 0) = \mathbb{P}(T_{n-2} = 0)(1 - n^{-1} + o(n^{-1})).
\]

We have \( \min(5p - 4p^2, 1 - p + 4p^2) \geq 3p \). Hence

\[
\frac{\mathbb{P}(S_n = x)}{\mathbb{P}(T_n = 0)} \leq \frac{\mathbb{P}(T_{n-2} = 0)(1 - \frac{3p}{2p(1-p)n} + o(n^{-1}))}{\mathbb{P}(T_{n-2} = 0)(1 - n^{-1} + o(n^{-1}))} = 1 - \frac{2p + 1}{2(1-p)n} + o(n^{-1}).
\]

Thus for odd \( n \) \( \mathbb{P}(S_n = x) \leq \mathbb{P}(T_n = 0)(1 - n^{-1} \min(\frac{3}{2(1-p)}, \frac{5-6p}{2p(1-p)})) + o(n^{-1})) \leq \mathbb{P}(T_n = 0)(1 - \frac{1}{2n} + o(n^{-1})) \) and for even \( n \)

\[
\mathbb{P}(S_n = x) \leq \mathbb{P}(T_n = 0) \left( 1 - n^{-1} \min(\frac{2p + 1}{2(1-p)}, \frac{5 - 6p}{2(1-p)}) + o(n^{-1}) \right)
\]

\[
\leq \mathbb{P}(T_n = 0) \left( 1 - \frac{1}{2n} + o(n^{-1}) \right).
\]

To prove the Proposition 2 we need precise asymptotics for \( \mathbb{P}(T_n = 0) \).

**Lemma 9.** Let \( p \in (0, \frac{1}{2}) \) and \( T_n \) be as in Lemma 7. \( \mathbb{P}(T_n = x) \) is maximized at the unique point \( x = 0 \).

For even \( n \)

\[
\mathbb{P}(T_n = 0) = \sum_{k=0}^{n} \mathbb{P}(Binom(n, p) = k)^2
\]

\[
= \frac{1}{\sqrt{2\pi np(1-p)}} \left( 1 + \frac{1}{4n} \left( \frac{1}{2p(1-p)} - 3 \right) + O(n^{-2}) \right).
\]

For odd \( n \)

\[
\mathbb{P}(T_n = 0) = \frac{1}{\sqrt{2\pi np(1-p)}} \left( 1 + \frac{p^2 - 6p + 1}{8np(1-p)} + o(n^{-1}) \right).
\]

**Proof.** For even \( n \) the maximum atom of \( T_n \) is 0 by symmetry and unimodality of \( T_n \).

For odd \( n \) the maximum atom is still zero by (10).

Wagner [21] presented a simple argument using analytic combinatorics and the generating function for the values of Legendre polynomials, that for \( b > 0 \) and \( c > 0 \)

\[
[x^n](x^2 + bx + c)^n = \frac{(b + 2\sqrt{c})^{n+1/2}}{2e^{1/4} \sqrt{\pi n}} \left( 1 + \frac{b - 4\sqrt{c}}{16n\sqrt{c}} + O(n^{-2}) \right).
\]
We have
\[
\mathbb{P}(T_{2n} = 0) = (1 - p)^{2n} \sum \binom{n}{k} \frac{p^{2k}}{(1 - p)^{2k}} = (1 - p)^{2n}[x^n](1 + dx)^n(x + 1)^n
\]
\[
= (1 - p)^{2n}d^n[x^n](x^2 + \frac{d + 1}{d}x + d^{-1})^n
\]
where \( d = \frac{p^2}{(1 - p)^2} \). Putting the estimate from (11) with
\[
a = 1; \quad b = \frac{d + 1}{d} = \frac{2p^2 - 2p + 1}{p^2}; \quad c = d^{-1} = \frac{(1 - p)^2}{p^2};
\]
\[
\sqrt{c} = \frac{1 - p}{p}; \quad b + 2\sqrt{c} = p^{-2}; \quad c^{1/4} = (1 - p)^{1/2}p^{-1/2};
\]
\[
b + 2\sqrt{c} = \frac{2p^2 - 2p + 1}{p^2} + \frac{2(1 - p)}{p} = p^{-2}
\]
\[
b - 4\sqrt{c} = \frac{b + 2\sqrt{c}}{2\sqrt{c}} - 3 = \frac{1}{2p(1 - p)} - 3;
\]
\[
\frac{(b + 2\sqrt{c})^{n+1/2}}{2c^{1/4}2p^{-1/2}(1 - p)^{1/2}} = 2^{-1}p^{-2n}p^{-1/2}(1 - p)^{-1/2}
\]
we get
\[
\mathbb{P}(T_{2n} = 0) = \frac{1}{2\sqrt{\pi np(1 - p)}}\left(1 + \frac{1}{8n}\left(\frac{1}{2p(1 - p)} - 3\right) + O(n^{-2})\right)
\]
For odd \( n \) by Lemma 8
\[
\mathbb{P}(T_n = 0) = p\mathbb{P}(T_{n-1} = -1) + (1 - p)\mathbb{P}(T_{n-1} = 0)
\]
\[
= \mathbb{P}(T_{n-1} = 0)(1 - \frac{p}{2p(1 - p)n} + o(n^{-1}))
\]
\[
= \mathbb{P}(T_{n-1} = 0)(1 - \frac{1}{2(1 - p)n} + o(n^{-1})).
\]
Using the already proved part for even \( n \)
\[
\frac{\mathbb{P}(T_{n-1} = 0)}{\sqrt{2\pi np(1 - p)}} = \left(1 + \frac{1}{4n}\left(\frac{1}{2p(1 - p)} - 3\right) + O(n^{-2})\right)\left(1 + \frac{1}{2n} + O(n^{-2})\right).
\]
Combining the last two estimates
\[
\frac{\mathbb{P}(T_n = 0)}{\sqrt{2\pi np(1 - p)}} = 1 + n^{-1}\left(\frac{1}{8np(1 - p)} - \frac{3}{4} - \frac{1}{2(1 - p)} + \frac{1}{2}\right) + o(n^{-1})
\]
\[
= 1 + \frac{2p^2 - 6p + 1}{8np(1 - p)} + o(n^{-1}).
\]
Proof of Proposition 2 By [2], $S_n$ has a unique mode at $np$ and by Stirling’s approximation:
\[ P(S_n = np)(2\pi p(1 - p)n)^{\frac{1}{2}} = 1 + \frac{p - p^2 - 1}{12np(1 - p)} + o(n^{-1}). \] (12)
By Lemma 4, the maximum atom of $T_n$ is 0, and
\[ P(T_n = 0)(2\pi p(1 - p)n)^{\frac{1}{2}} = 1 + \frac{2p^2 - 6p + 1}{8np(1 - p)} + o(n^{-1}). \] (13)
Let us compute the asymptotics for $Q_n$. We have $\mathbb{E} Q_n = 3p$, so by [2] the maximal atom is $\{0\}$, $\{1\}$ or $\{2\}$. Let $T_{n-3}$ be as in Lemma 7 be independent of $B_3 \sim \text{Binom}(3, p)$. Since $Q_n \sim T_{n-3} + B_3$ and $(n - 3)^{-1} = n^{-1} + o(n^{-1})$ we have by Lemma 8 for any $j \in \{0, 1, 2\}$
\[ P(Q_n = j) = \sum_{i=0}^{3} P(T_{n-3} = -i + j) P(B_3 = i) \]
\[ = P(T_{n-3} = 0) \left( (1-p)^3 \left(1 - \frac{j^2}{2p(1-p)n}\right) + 3p(1-p)^2 \left(1 - \frac{(j-1)^2}{2p(1-p)n}\right) + 3p^2(1-p) \left(1 - \frac{(j-2)^2}{2p(1-p)n}\right) + p^3 \left(1 - \frac{(j-3)^2}{2p(1-p)n}\right) + o(n^{-1}) \right) \]
\[ = P(T_{n-3} = 0) \left( 1 - \frac{j^2 - (6j - 3)p + 6p^2}{2p(1-p)n} + o(n^{-1}) \right) \]
Using Taylor’s approximation and Lemma 9
\[ P(T_{n-3} = 0) = (2\pi np(1-p))^{\frac{1}{2}} \left( 1 + \frac{1}{4n} \left( \frac{1}{2p(1-p)} - 3 \right) + \frac{3}{2n} + O(n^{-2}) \right). \]
Combining the last two estimates
\[ P(Q_n = j)(2\pi np(1-p))^{\frac{1}{2}} = 1 - \frac{4j^2 - 1 - (24j - 6)p + 30p^2}{8p(1-p)n} + O(n^{-2}). \] (14)
Now the proof follows by comparing the coefficients at $n^{-1}$ in (12), (13) and (14) for $j \in \{0, 1, 2\}$. (We used [22] to check simple algebra.)

Proof of (2) For even $n$, apply Theorem 1 and use the Local Central Limit Theorem (e.g., as stated in Theorem 1 of [3]). It is easy to see that $U_1^n - U_2^n$ satisfies the conditions of [3]. A simple calculation, which we omit, shows that its variance is $c_n^2 \pi^{-1}$. Thus by the Central Limit Theorem and the Local Central Limit Theorem
\[ P(X_1 + \cdots + X_n = x) \leq P(U_1^n - U_2^n + \cdots + U_{2k-1}^n - U_{2k}^n = 0) = (2\pi)^{-1/2} (2^{-1} \text{Var}(U_1^n - U_2^n)n)^{-1/2} = c_n n^{-1/2} (1 + o(1)). \]
For odd $n$ the same asymptotics follows by Remark 1.
5 Open problems and concluding remarks

We believe that at least for lattice-valued random vectors the following more general result is true.

**Conjecture 1.** Let $X_1, \ldots, X_n$ be iid random vectors in $\mathbb{Z}^d$. Then there exists a choice of non-random weights $w_i \in \{-1, 1\}$ such that for all non-zero $a_i \in \mathbb{R}$ and all $x \in \mathbb{R}^d$ we have

$$
P(a_1 X_1 + \ldots + a_n X_n = x) \leq \max_{k \in \mathbb{Z}^d} P(w_1 X_1 + \ldots + w_n X_n = k).
$$

Of course, in view of Corollary 1(b) one would have to only prove it for odd $n$.

The second conjecture concerns Theorem 1.

**Conjecture 2.** Let $X_1, \ldots, X_n$ be iid random vectors in $\mathbb{R}^d$ such that

$$
\sup_{x \in \mathbb{R}^d} P(X_i = x) \leq \alpha.
$$

Then there exists a choice of non-random weights $w_i \in \{-1, 1\}$ such that for all non-zero $a_i \in \mathbb{R}$ and all $x \in \mathbb{R}^d$ we have

$$
P(X_1 + \ldots + X_n = x) \leq \max_{k \in \mathbb{Z}^d} P(w_1 U_1^\alpha + \ldots + w_n U_n^\alpha = k).
$$

For the simplest case of iid Bernoulli random variables $X_i$ (i.e. $\alpha = \frac{1}{2}$) we saw that this is true for even $n$ and large odd $n$, and Singhal [18] independently proved it for all $n$. But the optimal number $k^*$ of $+1$s as a function of $p$ and $n$ is far from obvious. Can $k^*$ be any value in $\{0, \ldots, n\}$ depending on $p$? Can an explicit formula be obtained?

Tao and Vu proved in [23] that for a collection of non-zero $a_i \in \mathbb{Z}^d$ and independent random variables $X_i$ such that $P(X_i = \pm 1) = \frac{1}{2}$ the probability $P(a_1 X_1 + \ldots + a_n X_n = x)$ is large, then most of the $a_i$ can be covered by a small number of generalized arithmetic progressions. In other words, the collection of weights $a_i$ has strong additive structure. Their work lead important progress in the investigation of random matrices.

**Question 2.** Can inverse statements of Corollary 1(a) be obtained if we additionally assume that $a_i \in \mathbb{Z}^d$?

In the case when the variances of $X_i$ are bounded, we believe that the inverse statements should be analogous to the corresponding ones in [23]. The precise statement might need to be formulated differently in the case when $X_i$s have a heavy-tailed distribution.

References

[1] Z. W. Birnbaum, On random variables with comparable peakedness, *Ann. Math. Stat.*, 19 (1948), 76–81.
[2] J. N. Darroch, On the distribution of the number of successes in independent trials, *Annals of Mathematical Statistics* **35** (1964), 1317–1321.

[3] B. Davis and D. McDonald, An elementary proof of the local central limit theorem, *Journal of Theoretical Probability* **8** (1995), 693–702.

[4] P. Erdős, On a lemma of Littlewood and Offord, *Bull. Amer. Math. Soc.*, **51** (1945), 898–902.

[5] P. Erdős and L. Moser, Elementary problems and solutions: Solutions: E736, *The American Mathematical Monthly* **54** (1947), 229–230.

[6] C. G. Esseen, On the Kolmogorov-Rogozin inequality for the concentration function, *Z. Wahrscheinlichkeitstheor. Verw. Geb.* **5** (1966), 210–216.

[7] J. Fox, M. Kwan, and L. Sauermann, Combinatorial anti-concentration inequalities, with applications, *arXiv preprint arXiv:1905.12142* (2019).

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

[9] G. H. Hardy, J. E. Littlewood, and George Pólya, Inequalities. 2nd ed., Cambridge, Engl.: At the University Press. XII, 324 p. (1952)., 1952.

[10] B. M. Gabriel, The rearrangement of positive Fourier coefficients, *Proc. Lond. Math. Soc. (2)*, **33** (1931), 32–51.

[11] N. G. Gamkrelidze, Estimation of the maximum probability for sums of lattice random variables, *Theory of Probability & Its Applications* **18** (1974), 799–803.

[12] D. J. Kleitman, On a lemma of Littlewood and Offord on the distributions of linear combinations of vectors, *Advances in Math.* **5** (1970), 155–157.

[13] I. Leader and A. J. Radcliffe, Littlewood-Offord inequalities for random variables, *SIAM J. Discrete Math.* **7** (1994), 90–101.

[14] C. McDiarmid, Concentration, in *Probabilistic Methods for Algorithmic Discrete Mathematics*, M. Habib, C. McDiarmid, J. Ramirez-Alfonsin and B. Reed Eds., Springer, New York (1998) 195–248.

[15] A. Nicolas, Probability two symmetric random walks are at the same point after $n$ steps, *Mathematics Stack Exchange*, https://math.stackexchange.com/q/739658 (version: 2014-04-04).
[16] B. A. Rogozin, Inequalities for concentration functions of convolutions of arithmetic distributions and distributions with bounded densities, *Theory of Probability & Its Applications* 32 (1988), 325–329.

[17] P. Ruckdeschel, M. Kohl, T. Stabla, and F. Camphausen, S4 Classes for Distributions, *R News*, 6 (2006), 2–6.

[18] M. Singhal, Erdős-Littlewood-Offord problem with arbitrary probabilities, *arXiv preprint arXiv:1912.02886* (2019).

[19] R. P. Stanley, Weyl groups, the hard lefschetz theorem, and the sperner property, *SIAM Journal on Algebraic Discrete Methods* 1 (1980), 168–184.

[20] N. G. Ushakov, Upper estimates of maximum probability for sums of independent random vectors, *Theory Probab. Appl.*, 30 (1986), 38–49.

[21] S. Wagner, Asymptotics of generalised trinomial coefficients, *arXiv preprint arXiv:1205.5402* (2012).

[22] Wolfram—Alpha, [https://www.wolframalpha.com](https://www.wolframalpha.com), Last visited on 2019-09-20.

[23] T. Tao and Van H. Vu, Inverse Littlewood-Offord theorems and the condition number of random discrete matrices, *Ann. Math. (2)*, 169 (2009) 595–632.