On Littlewood–Offord problem for arbitrary distributions

T. Juškevičius*, V. Kurauskas

Vilnius University

July 4, 2020

Abstract

Let $X_1, \ldots, X_n$ be independent identically distributed discrete random vectors in $\mathbb{R}^d$. We consider upper bounds on $\sup_x \mathbb{P}(a_1X_1 + \cdots + a_nX_n = x)$ under various restrictions on X_i and 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 to 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 $\mathbb{P}(a_1X_1 + \cdots + a_nX_n = x)$ under various assumptions on the weights a_i and the distributions of X_i.

*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).

A special case when X_i are Rademacher random variables, i.e., $\mathbb{P}(X_i = \pm 1) = \frac{1}{2}$, is known as the Littlewood–Offord problem; the classical result of Erdős [3, 19] is that for non-zero real weights a_i we have

$$\mathbb{P}(a_1X_1 + \cdots + a_nX_n = x) \leq \mathbb{P}(X_1 + \cdots + X_n \in \{0, 1\}) = \frac{\binom{n}{\lfloor \frac{n}{2} \rfloor}}{2^n} = \sqrt{\frac{2}{\pi n}} + O(n^{-\frac{3}{2}}). \quad (1)$$

Kleitman [13] proved that the latter result remains true for Rademacher random variables and $a_i \in \mathbb{R}^d$. That is, linear combinations with equal weights exhibit the worst case behaviour. The same problem with restrictions on the arithmetical structure of the weights a_i was considered in [4, 7, 20]. Fairly recently the Littlewood–Offord problem was considered in certain matrix groups in [22] and sharp results for arbitrary groups were obtained in [12].

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 out 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 n be a positive even integer and let $X_1, \ldots, X_n$ be independent discrete random vectors in $\mathbb{R}^d$. Then there is $j \in \{1, 2, \ldots, n\}$ such that for all $x \in \mathbb{R}^d$*

$$\mathbb{P}(X_1 + \cdots + X_n = x) \leq \mathbb{P}(Y_1 - Y_2 + \cdots + Y_{n-1} - Y_n = 0)$$

where $Y_1, \ldots, Y_n$ are iid copies of X_j. The inequality is strict unless $\sum X_i - x$ and $\sum (-1)^{i+1} Y_i$ have the same distribution.

Intuitively, this lemma says that the probability for a random walk with an even number of steps from some class of distributions to hit a particular value x is never greater than the probability to hit the origin by repeatedly 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. *For even n and any $x \in \mathbb{R}^d$ we have*

$$\mathbb{P}(a_1X_1 + \cdots + a_nX_n = x) \leq \mathbb{P}(X_1 - X_2 + \cdots + X_{n-1} - X_n = 0)$$

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

(b) for iid discrete 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 ± 1s, regardless of the distribution of the random variables X_i. Therefore 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 Lévy 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^α a random variable such that $\mathbb{P}(U^\alpha = l) = \alpha$ for $l = 0, 1, \ldots, \lfloor \frac{1}{\alpha} \rfloor - 1$ 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 $\alpha = \frac{1}{k}$ with $k \in \mathbb{N}$ this random variable has the uniform distribution on $\{0, \ldots, k-1\}$ and for $\alpha \in [\frac{1}{2}, 1)$ it has the Bernoulli distribution with parameter $1 - \alpha$. We then establish the following inequality.

Theorem 1. *Let n be a positive even integer and let $\alpha \in (0,1)$. Let $X_1, \ldots, X_n$ be independent random vectors in $\mathbb{R}^d$ such that for all $i \in \{1, \ldots, n\}$ we have*

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

Then

$$\mathbb{P}(X_1 + \cdots + X_n = x) \leq \mathbb{P}(U_1^\alpha - U_2^\alpha + \cdots + U_{n-1}^\alpha - U_n^\alpha = 0),$$

where the random variables U_i^α are iid copies of U^α.

Note that the latter inequality is optimal as the random variables U_i^α satisfy the condition of the theorem. The result for $\alpha = \frac{1}{k}$ with $k \in \mathbb{N}$ was established by Rogozin [17] and also follows from the results of Leader and Radcliffe [15]. Bounds for arbitrary α were obtained by Ushakov [23], but they were not optimal when $\alpha < \frac{1}{2}$. We postpone the detailed discussion regarding a more complete history of this problem to Section 3.

Corollary 2. *In the setting of Theorem 1 for all n (even or odd) we get*

$$\mathbb{P}(X_1 + \cdots + X_n = x) \leq (2\pi n \, Var(U^\alpha))^{-\frac{1}{2}} (1 + o(1)), \quad \text{where} \tag{2}$$
$$Var(U^\alpha) = \frac{1}{12} \lfloor \alpha^{-1} \rfloor (\lfloor \alpha^{-1} \rfloor + 1) \alpha (2 + 4\lfloor \alpha^{-1} \rfloor - 3\alpha \lfloor \alpha^{-1} \rfloor - 3\alpha \lfloor \alpha^{-1} \rfloor^2).$$

When α^{-1} is integer, this simplifies to $\mathrm{Var}(U^\alpha) = \frac{1-\alpha^2}{12\alpha^2}$.

Recently Fox, Kwan and Sauermann [6] have posed the following question (we rephrase it slightly).

Question 1. *Let $a_1, \ldots, a_n$ be non-zero real numbers and let $X_1, \ldots, X_n$ 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 $\alpha = 1 - p$. The situation for odd n seems to be much more involved. We were only able to prove that optimal a_i must be ± 1 when n is large enough (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 [23], communicated by Prokhorov in the early 80s, already contains the asymptotically sharp bound

$$\mathbb{P}(X_1 + \cdots + X_n = x) \leq (2\pi np(1-p))^{-\frac{1}{2}}(1 + O(n^{-\frac{1}{2}})).$$

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 and Corollary 1 are proved in Section 2. Theorem 1 is proved and the history of the problem is discussed in Section 3. Section 4 is devoted to the Bernoulli case and results for odd values of n. Finally, we present and discuss some open problems in Section 5.

2 Proof of the balancing lemma

We shall use the notation $X \sim Y$ to denote the fact that the random vectors 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 It is enough to prove the lemma for $x = 0$, otherwise we can redefine X_1 as $X_1 - x$.

Let us split the sum into two halves:

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

Let S' and T' be independent copies of S and T respectively. By the inequality of arithmetic and geometric means

$$\begin{aligned}\mathbb{P}(S+T=0) &= \sum_x \mathbb{P}(S=x)\mathbb{P}(-T=x) \le \sum_x \frac{\mathbb{P}(S=x)^2+\mathbb{P}(-T=x)^2}{2}\\ &= \frac{1}{2}\sum_x \mathbb{P}(S=x)^2 + \frac{1}{2}\sum_x \mathbb{P}(-T=x)^2\\ &\le \max\left\{\sum_x \mathbb{P}(S=x)^2, \sum_x \mathbb{P}(T=x)^2\right\}\\ &= \max\left\{\mathbb{P}(S-S'=0), \mathbb{P}(T-T'=0)\right\}. \end{aligned} \tag{3}$$

Note that for non-negative p and q, $pq \le \frac{p^2+q^2}{2}$ and we have an equality if and only if $p=q$. Therefore (3) is equality if and only if $T \sim -S$.

We will say that random vectors X and Y have the same *type* if either Y or $-Y$ has the same distribution as X. Consider the different equivalence classes (types) of $\{X_1,\dots,X_n\}$ defined by the above equivalence relation. Note that $S-S'$ is a sum of n independent random vectors whose terms preserve the types of S, furthermore, each term in S is matched by a term with an opposite sign in S' and similarly for $T-T'$. Thus if all the random variables have the same type, the proof follows by (3).

If there are more than two types, let $\mathcal{X}_1 = \{X_{i_1},\dots,X_{i_k}\}$ and $\mathcal{X}_2 = \{X_{j_1},\dots,X_{j_l}\}$ be different classes other than the largest equivalence class (break ties arbitrarily). Clearly $k \le \frac{n}{2}$ and $l \le \frac{n}{2}$. Rearrange the variables so that all the variables in $\mathcal{X}_1$ are in S and all the variables in $\mathcal{X}_2$ are in T. Applying (3) yields a new sequence of random variables $X_1',\dots,X_n'$ with $X_{2k}' \sim -X_{2k-1}'$, $k \in \{1,\dots,\frac{n}{2}\}$ which has at least one less type and

$$\mathbb{P}(X_1+\cdots+X_n=0) \le \mathbb{P}(X_1'+\cdots+X_n'=0).$$

By repeating this argument at most n times, we reduce the number of types to one or two. It remains to consider the case when there are exactly two types among $X_1,\dots,X_n$. 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 the first cycle (S_1,T_1), (S_2,T_2), …, (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. This procedure is well defined because the number of possible configurations for S and T is finite. Using (3) and additionally ordering the variables by their "sign" we can even ensure that the total number of steps is at most $n-1$. Suppose we still have two types in the end. Then

$$\mathbb{P}(X_1+\cdots+X_n=0) \le \cdots \le \mathbb{P}(S_1+T_1=0) \le \cdots \le \mathbb{P}(S_k+T_k=0),$$

which implies $\mathbb{P}(S_1+T_1=0)=\cdots=\mathbb{P}(S_k+T_k=0)$, and so $T_1 \sim -S_1$, see our observation on the equality in (3).

Taking j such that X_j has the prevailing type completes the proof of the stated inequality. By the above observation, if we applied (3) at least once where $T \sim -S$ does not hold, this inequality is strict. Otherwise we must have $\sum X_i \sim \sum (-1)^{i+1} Y_i$. ■

Proof of Corollary 1 Apply Lemma 1 to the independent random vectors $a_i X_i$. ■

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 [16].

3 Random variables with bounded concentration

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

$$\sup_{x \in \mathbb{R}^d} \mathbb{P}(X_i = x) \le \alpha \in (0, 1).$$

The bounds on the concentration probability $\mathbb{P}(S_n = x)$ were studied by many authors. Let us just mention the work of Esseen [5], Rogozin [17] and Gamkrelidze [10]. It was proved by Rogozin that when $d = 1$ and $\alpha = \frac{1}{k}$ for $k \in \mathbb{N}$, the probability $\mathbb{P}(S_n = x)$ is maximized when all X_i are iid uniform random variables in the set $\{0, \dots, k-1\}$. This result also follows from more general bounds obtained by Leader and Radcliffe [15]. To our knowledge the sharpest known bounds for $\alpha \in [\frac{1}{2}, 1)$ and all d were obtained by Ushakov [23]. Such α are especially interesting as they cover all Bernoulli distributions. Ushakov established the inequality

$$\mathbb{P}(S_n = x) \le (2\pi(n+1)\alpha(1-\alpha))^{-\frac{1}{2}} \left(1 + (2(n+1)\alpha(1-\alpha))^{-\frac{1}{2}}\right),$$

which is asymptotically sharp: this can be seen by Lemma 9 in [11] (with the correct second order term), or alternatively by using the Local Limit Theorem.

Let us give a short description 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 a result of Ushakov [23] 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 [9] which we have been fortunate to find in the classical monograph of Hardy, Littlewood and Pólya [8]. 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.

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 μ 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 functional. Also note 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 in S_α 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 α^{-1} is integer it is equal to $|A|$, in this case y becomes a dummy parameter: $\mu_{\alpha,A,y}$ 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 μ and ν on X is *non-trivial* if $0 < p < 1$.

Lemma 2. *Let $\alpha \in (0,1)$ and let μ be a measure in S_α. Then μ can be written as a non-trivial convex combination of two distinct measures in S_α if and only if it is not a measure $\mu_{\alpha,A,y}$ for some $A \subseteq X$ and $y \in X \setminus A$.*

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 α^{-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 $\mu_2\{x\} = 1 - \lfloor \alpha^{-1} \rfloor \alpha < \alpha$. So $p\mu_1\{x\} + (1-p)\mu_2\{x\} < \alpha = \mu\{x\}$, also a contradiction.

Now assume that μ is not of the form $\mu_{\alpha,A,y}$. Since $\mu \in S_\alpha$, its support is of size at least $\lfloor \alpha^{-1} \rfloor + 1$. Let A be the set of $\lfloor \alpha^{-1} \rfloor$ largest atoms of μ and let y be its largest atom outside A. Thus $\mu\{x\} > 0$ for each $x \in A \cup \{y\}$. Let $\mu_2 = \mu_{\alpha,A,y}$. Fix a positive ϵ small enough that $(1+\epsilon)\mu\{x\} - \epsilon\mu_2\{x\} \geq 0$ for $x \in A \cup \{y\}$ and $1 + \epsilon \leq \alpha(\lfloor \alpha^{-1} \rfloor + 1)$. Define $\mu_1 = (1+\epsilon)\mu - \epsilon\mu_2$. We have $\mu = p\mu_1 + (1-p)\mu_2$ with $p = \frac{1}{1+\epsilon}$. Since $\mu_1 = \mu + \epsilon(\mu - \mu_2)$ and $\mu \neq \mu_2$, μ_1 and μ_2 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 $x \in X \setminus A$, $\mu\{x\} \leq \mu\{y\} \leq (|A|+1)^{-1}$. Thus $\mu_1\{x\} \leq (1+\epsilon)\mu\{x\} \leq (1+\epsilon)(|A|+1)^{-1} \leq \alpha$ for $x \in X \setminus A$. ■

In his work on the problem of this section Ushakov [23] proved a couple of reduction lemmas that allow switching from distributions in Hilbert spaces to distributions on the integers. We shall state the one we require here.

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

$$Q(\mu_i) \le \alpha.$$

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

$$Q(\mu_1 * \ldots * \mu_n) \le Q(\nu_1 * \ldots * \nu_n),$$

where $*$ stands for convolution.

In two important parts of the proof we shall use rearrangement results from [9] (see also [8] page 273, Theorem 374). First let us define certain special rearrangements of a finite sequence of non-negative numbers $(a) = (a_{-k}, \ldots, a_k)$ indexed by integers. The rearrangement $({}^+a)$ is defined by inequalities ${}^+a_0 \ge {}^+a_{-1} \ge {}^+a_1 \ge {}^+a_{-2} \ge \ldots \ge {}^+a_k$. Analogously, the rearrangement (a^+) is defined by inequalities $a_0^+ \ge a_1^+ \ge a_{-1}^+ \ge a_2^+ \ge \ldots \ge a_{-k}^+$. Finally, if in the sequence (a) all values except the largest one appear an even number of times, we define the symmetric decreasing rearrangement (a^*) by the inequalities $a_0^* \ge a_1^* = a_{-1}^* \ge a_2^* = a_{-2}^* \ge \ldots \ge a_{-k}^* = a_k^*$. 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 a finite collection of finite sequences of non-negative numbers such that all collections except maybe (a) and (b) have a symmetric decreasing rearrangement. Then*

$$\sum_{r+s+t+u+\cdots=0} a_r b_s c_t d_u \ldots \le \sum_{r+s+t+u+\cdots=0} {}^+a_r b_s^+ c_t^* d_u^* \ldots.$$

The final tool we shall require is a discrete counterpart of Birnbaum's [1] result on the peakedness of symmetric unimodal distributions. This result might be known, but we could not find it in the literature, so we provide a simple proof for the readers' convenience.

Lemma 5. *Let X, Y and Y' be independent symmetric unimodal integer random variables. Suppose $\mathbb{P}(Y \in [-k, k]) \le \mathbb{P}(Y' \in [-k, k])$ for any integer $k \ge 0$.*

Then for any integer $k \ge 0$

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

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 $|Y'| \le |y|$.

Since X is symmetric and unimodal, for any $y, y' \in \mathbb{Z}$ if $0 \le y' \le y$ or $y \le y' \le 0$ we have $\mathbb{P}(X \in [y - k, y + k]) \le \mathbb{P}(X \in [y' - k, y' + k])$ for any integer $k \ge 0$. Therefore

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

■

The class of symmetric unimodal distributions is closed under convolution (see e.g. [8] Theorem 375). Applying Lemma 5 with the ith term of the sum and the rest of the sum for each $i \in \{1, \dots, n\}$ we get:

Corollary 3. *Let $X_1, \dots, X_n, Y_1, \dots, Y_n$ be independent symmetric unimodal integer random variables. Suppose for $i \in \{1, \dots, n\}$ and any integer $k \geq 0$ we have $\mathbb{P}(X_i \in [-k, k]) \leq \mathbb{P}(Y_i \in [-k, k])$. Then for any integer $k \geq 0$ 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 Let $k = \frac{n}{2}$. 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 the distribution of each random variable X_i is finitely supported; the general case follows by approximating with truncated random variables. The Krein–Milman theorem [14] tells us that the convex set of distributions μ 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. The extreme points of a convex set A are the points that do not lie in the interior of any segment in A. The extreme points for our situation are described by Lemma 2: they are exactly the measures of the form $\mu_{\alpha,A,y}$ for some $A \subset \mathbb{Z}$ with $|A| = \lfloor \alpha^{-1} \rfloor$ and $y \in \mathbb{Z} \setminus A$. For each i let us define $g(t) = \mathbb{P}(X_1 + \cdots + X_{2k} - X_i + t = x)$. We have $\mathbb{P}(X_1 + \ldots + X_{2k} = x) = \mathbb{E}\, g(X_i)$ and so if the distribution of X_i is a linear combination of some collection of distributions, then $\mathbb{E}\, g(X_i)$ is a linear combination of expectations of g with respect to each of these distributions. This means that we can assume that the maximum of $\mathbb{P}(X_1 + \ldots + X_{2k} = x)$ is attained when each X_i has distribution μ_{α,A_i,y_i} for some $A_i \subset \mathbb{Z}$ and $y_i \in \mathbb{Z} \setminus A_i$. Applying Lemma 1 we obtain

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

where Y_i are iid random variables distributed as some X_j. Let us denote the distribution of that particular X_j by $\mu_{\alpha,A,y}$ (dropping the subscripts of A_i and y_i). The random variables $Z_i = Y_{2i-1} - Y_{2i}$ are iid and symmetric. Let $(Z_i^*, i \in \{1, \dots, k\})$ be a sequence of iid random variables where the distribution of Z_i^* is obtained from the symmetric decreasing rearrangement of the distribution of Z_1. Applying Lemma 4 we obtain

$$\begin{aligned}
\mathbb{P}(Y_1 - Y_2 + \ldots + Y_{2k-1} - Y_{2k} = 0) &= \mathbb{P}(Z_1 + \ldots + Z_k = 0) \\
&= \sum_{m_1 + \cdots + m_k = 0} \mathbb{P}(Z_1 = m_1) \ldots \mathbb{P}(Z_k = m_k) \\
&\leq \sum_{m_1 + \cdots + m_k = 0} \mathbb{P}(Z_1 = m_1)^* \ldots \mathbb{P}(Z_k = m_k)^* \\
&= \mathbb{P}(Z_1^* + \ldots + Z_k^* = 0).
\end{aligned}$$

We have now achieved an inequality for the probability in question in terms of symmetric unimodal distributions to which Corollary 3 applies. All that is left to prove is the stochastic domination condition $\mathbb{P}(|Z_i^*| \leq l) \leq \mathbb{P}(|U_{2i-1}^\alpha - U_{2i}^\alpha| \leq l)$ for all integers l. We shall actually show that

$$\mathbb{P}(|Z_i^*| \leq l) = \max_{B\subset\mathbb{Z},|B|=2l+1} \mathbb{P}(Z_i \in B) \leq \mathbb{P}(|U_{2i-1}^\alpha - U_{2i}^\alpha| \leq l).$$

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 on the set $\{-x : x \in B\}$ which is independent of the previously defined random variables. In the case $B = \{-l, \dots, l\}$ we shall denote this random variable by U. Recalling that $Y_1, \dots, Y_{2k}$ have distribution $\mu_{\alpha,A,y}$ and using Lemma 4 we obtain

$$\begin{aligned}
&(2l+1)\mathbb{P}(Z_i \in B) = \mathbb{P}(Y_{2i-1} - Y_{2i} + U(-B) = 0)\\
=& \sum_{r+s+t=0} \mathbb{P}(Y_{2i-1} = r)\mathbb{P}(-Y_{2i} = s)\mathbb{P}(U(-B) = t)\\
\leq& \sum_{r+s+t=0} {}^+\mathbb{P}(Y_{2i-1} = r)\, \mathbb{P}(-Y_{2i} = s)^+\, \mathbb{P}(U(-B) = t)^*\\
=& \sum_{r+s+t=0} \mathbb{P}(U_{2i-1}^\alpha - \lfloor\lfloor\alpha^{-1}\rfloor/2\rfloor = r)\mathbb{P}(-U_{2i}^\alpha + \lfloor\lfloor\alpha^{-1}\rfloor/2\rfloor = s)\mathbb{P}(U = t)\\
=& \mathbb{P}(U_{2i-1}^\alpha - U_{2i}^\alpha + U = 0)\\
=& (2l+1)\mathbb{P}(U_{2i-1}^\alpha - U_{2i}^\alpha \in \{-l, \dots, l\})
\end{aligned}$$

and we are done. ■

Proof of Corollary 2. Let n be even. The Local Limit Theorem, e.g. Theorem 1 of [2], applied to the sum of $\frac{n}{2}$ iid random variables $U_{2i}^\alpha - U_{2i-1}^\alpha$ gives

$$\begin{aligned}
&\mathbb{P}(U_1^\alpha - U_2^\alpha + \dots + U_{n-1}^\alpha - U_n^\alpha = 0) = \left(2\pi\frac{n}{2}\mathrm{Var}\,(U_1^\alpha - U_2^\alpha)\right)^{-\frac{1}{2}}(1 + o(1))\\
&= (2\pi n \mathrm{Var}\,(U^\alpha))^{-\frac{1}{2}}(1 + o(1)).
\end{aligned}$$

Clearly the conditions of [2] are satisfied since the support of $U_1^\alpha - U_2^\alpha$ is an interval.

Let $k = \lfloor\alpha^{-1}\rfloor$. By a simple calculation using the fact that U^α is a mixture of $U^{\frac{1}{k+1}}$ with probability $\lambda = (k+1)(1-k\alpha)$ and $U^{\frac{1}{k}}$ with probability $1 - \lambda$ it follows that

$$\mathrm{Var}\,(U^\alpha) = \frac{1}{12}k(k+1)\alpha(2 + 4k - 3\alpha k^2 - 3\alpha k).$$

Applying Theorem 1 completes the proof for even n. For odd n the same asymptotics follow by Remark 1. ■

4 Bernoulli distributions

Recall that by Corollary 1(a) for even n, the optimal solution to Question 1 is to take $x = 0$, set $\frac{n}{2}$ of the coefficients a_i to 1 (or any other constant c), and the remaining coefficients to -1 (or $-c$). Lemma 1 says that no other assignment can achieve this maximum. For odd n an asymptotically optimal upper bound in Question 1 follows by Remark 1. However, surprisingly, determining which collections of coefficients are extremal becomes a much harder and more subtle problem. It turns out, that for odd n, the answer depends non-trivially on p.

In the full version of this paper [11] we use a similar technique as in the proof of Lemma 1 to give a partial answer for odd n and several related results. For example, we show that when p is fixed and n is large enough, like for even n, the optimal a_i in Question 1 must be ± 1 (or $\pm c$ for a constant c). Independently, Singhal [18] proves a similar result for all n. For fixed p Singhal [18] also determines the asymptotically optimal proportion of $+1$s and -1s using characteristic functions. This proportion depends non-continuously on the number p and can be very far from balanced.

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 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*

$$\mathbb{P}(a_1X_1 + \ldots + a_nX_n = x) \leq \max_{k \in \mathbb{Z}^d} \mathbb{P}(w_1X_1 + \ldots + w_nX_n = k).$$

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*

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

Then there exists a choice of 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

$$\mathbb{P}(X_1 + \ldots + X_n = x) \leq \max_{k \in \mathbb{Z}} \mathbb{P}(w_1U_1^\alpha + \ldots + w_nU_n^\alpha = k).$$

For the simplest case of iid Bernoulli random variables X_i (i.e. $\alpha \geq \frac{1}{2}$) we saw that this is true for even n and large odd n, and Singhal [18] independently proved it for all n.

What is the optimal number $l = l(n, p)$ of $+1$s? Although Singhal obtains excellent results for fixed p and large n, the complete answer still seems not obvious. For example, for a given n, can l take any value in $\{0, \ldots, n\}$ depending on p?

Tao and Vu proved in [21] that for a collection of non-zero $a_i \in \mathbb{Z}^d$ and independent random variables X_i such that $\mathbb{P}(X_i = \pm 1) = \frac{1}{2}$ if the probability $\mathbb{P}(a_1X_1+\cdots+a_nX_n = x)$ 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 [21]. The precise statement might need to be formulated differently in the case when X_is have a heavy-tailed distribution.

Acknowledgements. We would like to thank the reviewers for their useful remarks. We are also grateful to Matas Šileikis for a careful reading of our manuscript and numerous corrections that he provided.

References

[1] Z. W. Birnbaum, On random variables with comparable peakedness, *Ann. Math. Stat.* **19** (1948), 76–81.

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

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

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

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

[6] J. Fox, M. Kwan, and L. Sauermann, Combinatorial anti-concentration inequalities, with applications, *Math. Proc. Cambridge Philos. Soc.*, to appear; *arXiv preprint* arXiv:1905.12142 (2019).

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

[8] G. H. Hardy, J. E. Littlewood, and George Pólya, *Inequalities*, Cambridge University Press, Cambridge, 1952.

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

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

[11] T. Juškevičius and V. Kurauskas, On Littlewood–Offord problem for arbitrary distributions (full version), *arXiv preprint* arXiv:1912.08770.

[12] T. Juškevičius and G. Šemetulskis, Optimal Littlewood–Offord inequalities in groups, *Combinatorica* **39** (2019) 911–921.

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

[14] M. Krein and D. Milman, On extreme points of regular convex sets, *Studia Mathematica* **9** (1940) 133–138.

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

[16] 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).

[17] 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.

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

[19] E. Sperner, Ein Satz über Untermengen einer endlichen Menge. *Math. Z.* **27** (1928), 544–548.

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

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

[22] Pham H. Tiep and Van H. Vu, Non-abelian Littlewood–Offord inequalities, *Advances in Mathematics* **302** (2016) 1233–1250.

[23] N. G. Ushakov, Upper estimates of maximum probability for sums of independent random vectors, *Theory of Probability & Its Applications* **30** (1986), 38–49.