Anticoncentration of random vectors via the strong perfect graph theorem

Tomas Juškevičius[*] and Valentas Kurauskas

12 August 2023

Abstract

In this paper we give anticoncentration bounds for sums of independent random vectors in finite-dimensional vector spaces. In particular, we asymptotically establish a conjecture of Leader and Radcliffe (1994) and a question of Jones (1978).

The highlight of this work is an application of the strong perfect graph theorem by Chudnovsky, Robertson, Seymour and Thomas (2003) in the context of anticoncentration.

keywords: concentration function; Littlewood-Offord problem; perfect graph

1 Introduction

Let X be a random vector taking values in some metric space $\mathcal{S}$. Given $t > 0$ define

$$\mathcal{Q}(X, t) = \sup_B \mathbb{P}(X \in B) \quad \text{and} \quad \tilde{\mathcal{Q}}(X, t) = \sup_S \mathbb{P}(X \in S),$$

where B is an open ball and S an open set, both having diameter at most t. We clearly have $\mathcal{Q}(X, t) \le \tilde{\mathcal{Q}}(X, t)$.

$\mathcal{Q}(X, \cdot)$ is often referred to as the *Lévy concentration function* of X or just the *concentration function* of X in the literature (sometimes the ball B is considered to be closed), see e.g. [15], [2]. The similar function $\tilde{\mathcal{Q}}(X, \cdot)$ was used by Leader and Radcliffe [17] in their work on generalized Littlewood-Offord type inequalities.

One special case of our results (see Section 2) is as follows.

Theorem 1.1 *Fix an integer $d \geq 2$. Let $X_1, \ldots, X_n$ be independent random vectors in $\mathbb{R}^d$ equipped with a norm $\|\cdot\|$. Assume that for some $k \in \mathbb{N}$ and each $i \in \{1, \ldots, n\}$ we have*

$$\tilde{\mathcal{Q}}(X_i, 1) \leq \frac{1}{k}.$$

Let $U_1, \ldots, U_n$ be i.i.d. random variables distributed uniformly on the set $\{1, 2, \ldots, k\}$. Then

$$
\begin{aligned}
\tilde{\mathcal{Q}}(X_1 + \cdots + X_n, 1) &\leq (1 + o(1))\tilde{\mathcal{Q}}(U_1 + \cdots + U_n, 1) \quad\quad\quad (1) \\
&= (1 + o(1)) \max_{\ell \in \mathbb{Z}} \mathbb{P}(U_1 + \cdots + U_n = \ell).
\end{aligned}
$$

Note that even though $\mathcal{Q}$ and $\tilde{\mathcal{Q}}$ are defined using diameter, which depends on the distance or the norm, our convention will be not to indicate this in the notation. For 1-dimensional random variables like U_i, the usual distance $d(x, y) = |x - y|$ will be assumed, unless mentioned otherwise. For claims about general normed spaces with norm $\|\cdot\|$, like on the left of (1), we implicitly use the induced distance $d(x, y) = \|x - y\|$. Let us also note that the $o(1)$ term above is dependent on k, the norm $\|\cdot\|$ and the dimension d, which is not reflected in the notation.

The inequality (1) is clearly optimal apart from the multiplicative factor $1 + o(1)$. This is because $\tilde{\mathcal{Q}}(U_i, 1) = \frac{1}{k}$ and we can attain the value $\tilde{\mathcal{Q}}(U_1 + \ldots + U_n, 1)$ by X_i that are uniformly distributed on the first k integer multiples for some unit vector $v \in \mathbb{R}^d$.

The case $k = 2$ in Theorem 1.1 corresponds to the classical Littlewood-Offord problem. Leader and Radcliffe [17] conjectured that in this case the inequality (1) holds in arbitrary normed spaces without the multiplicative error factor (Conjecture 13, p. 101). They proved it for $d = 1$. Jones [10] asked whether or not the inequality (1) without the error factor holds when $k > 3$ and each random variable X_i has a uniform distribution on a k-point set A_i in a Banach space such that the pairwise distances in A_i are at least 1 (Question 1 in Section 4). The latter is a special case of our setup. Consequently, Theorem 1.1 establishes both problems asymptotically in finite-dimensional spaces.

The highlight of our work is the use of the strong perfect graph theorem that allows us to represent distributions of random vectors in a useful way so that existing combinatorial methods can be applied to the components of the decomposition. We believe that our approach can be useful in other contexts.

The paper is organized as follows. Notation and important definitions are presented and main results are formulated in Section 2. We outline the proof strategy in Section 3. The subsequent Sections 4–10 deal with different parts of the proof.

2 Main results

For a finite multiset A let $\nu(A)$ be the uniform measure on A. Thus, for example, if $A = \{0, 0, 1\}$ then $\nu(A)$ is the distribution of a Bernoulli random variable with

success probability $p = \frac{1}{3}$. We call a finite set of k points in a metric space a *k-block* if those points are at pairwise distance at least 1. We shall sometimes refer to the latter sets as *blocks* without specifying the particular parameter k.

Let us define a parameterized family of probability measures that will serve as worst-case distributions in the forthcoming results. For $\alpha \in (0, 1]$ set $k = \lfloor \alpha^{-1} \rfloor$ and take the unique $p = p(\alpha) \in (0, 1]$ such that $\frac{p}{k} + \frac{1-p}{k+1} = \alpha$. Define a measure on $\mathbb{R}$ by

$$\nu_\alpha^* := p\nu(\{1, 2, \ldots, k\} - \tfrac{k+1}{2}) + (1 - p)\nu(\{1, 2, \ldots, k+1\} - \tfrac{k+2}{2}). \quad (2)$$

Note that in the special case $\alpha = \frac{1}{k}$ we have $p = 1$ and so in view of (2) the measure ν_α^* is the uniform distribution on a symmetric arithmetic progression on k points with difference 1. It therefore has the same distribution as $U_i - \mathbb{E}\, U_i$ with U_i as in Theorem 1.1. When α is not an inverse of an integer the measure ν_α^* is a non-trivial convex combination of two distinct uniform distributions.

We can now state a more general version of Theorem 1.1, which is one of our main results.

Theorem 2.1 *Fix an integer $d \geq 2$, and $\alpha \in (0, 1]$. Let $X_1, \ldots, X_n$ be independent random vectors in $\mathbb{R}^d$ endowed with a norm $|| \cdot ||$. Let $Y_1, \ldots, Y_n$ be i.i.d. random variables with distribution ν_α^*. Assume that for each i*

$$\tilde{\mathcal{Q}}(X_i, 1) \leq \alpha. \quad (3)$$

We then have

$$\tilde{\mathcal{Q}}(X_1 + \cdots + X_n, 1) \leq (1 + o(1))\tilde{\mathcal{Q}}(Y_1 + \cdots + Y_n, 1), \quad (4)$$

where the $o(1)$ term depends only on α, d and the underlying norm $|| \cdot ||$.

As mentioned above, in the one-dimensional case the sharp inequality (without the $1 + o(1)$ multiplicative factor) was proved by Leader and Radcliffe [17] when α is an inverse of an integer. Answering a question of theirs it was extended to $d = 1$ and all values of α by the first author in [11] and [12].

There are essential differences between the one-dimensional and the higher-dimensional case. This is due to the fact that the class of probability distributions satisfying the condition (3) when $d > 1$ is in some sense much richer in contrast to the case $d = 1$, see Section 3. The crucial tool to analyze the situation in high dimensions, perhaps surprisingly, comes from structural graph theory. To be more precise, we use the strong perfect graph theorem famously proved by Chudnovsky, Seymour, Robertson and Thomas [4]. We discuss this in greater detail in Section 3.

Furthermore, our methods allow us to obtain the optimal inequality in greater generality provided all distributions are close to a single line.

Let $\mathcal{S}$ be a normed space with norm $|| \cdot ||$. The norm induces a distance $d(x, y) = ||x - y||$. For $x \in \mathcal{S}$ and S a set in $\mathcal{S}$ we write $d(x, S) = \inf_{y \in S} d(x, y)$. A line in $\mathcal{S}$ is a set $\{tv + b : t \in \mathbb{R}\}$ where $v, b \in \mathcal{S}$ and $||v|| = 1$.

Theorem 2.2 *Let $X_1, \ldots, X_n$ be independent random vectors with values in a normed space over the reals $\mathcal{S}$ with norm $\|\cdot\|$ and let $\alpha_1, \ldots, \alpha_n \in (0, 1]$. Assume that $\tilde{\mathcal{Q}}(X_1, 1) \leq \alpha_1, \ldots, \tilde{\mathcal{Q}}(X_n, 1) \leq \alpha_n$ and that there is a line L in $\mathcal{S}$ so that $\sup_\omega d(X_i(\omega), L) < \frac{1}{8}$ ($< \frac{\sqrt{3}}{4}$ if $\mathcal{S}$ is Hilbert) for each i. If either*

i) the random vectors X_i have countable supports; or

ii) $\mathcal{S}$ is complete and separable;

we have that

$$\tilde{\mathcal{Q}}(X_1 + \cdots + X_n, 1) \leq \tilde{\mathcal{Q}}(Y_1 + \cdots + Y_n, 1) = \mathbb{P}(Y_1 + \cdots + Y_n \in \{0, \frac{1}{2}\}),$$

where $Y_1, \ldots, Y_n$ are independent and Y_i has distribution $\nu_{\alpha_i}^*$ defined in (2).

The latter inequality is optimal and in the case $\alpha_i = \frac{1}{2}$ it gives the desired result in the Leader-Radcliffe conjecture and when $\alpha_i = \frac{1}{k}$ it also answers the discussed question by Jones in this restricted setting.

Finally, our methods allow us to treat the general situation with a non-uniform condition in (3) when a certain kind of a local limit theorem for the extremal distributions holds. We shall formulate it in the language of triangular arrays. The interested reader can find a general non-asymptotic version of this result in Section 10, see Theorem 10.3.

Theorem 2.3 *Fix an integer $d \geq 2$, and a norm $\|\cdot\|$ on $\mathbb{R}^d$. Then there is a constant $C = C(d, \|\cdot\|)$ such that the following holds.*

For each $l \in \{1, 2, \ldots,\}$ let $X_{l,1}, \ldots, X_{l,n_l}$ be a sequence of independent random variables in $(\mathbb{R}^d, \|\cdot\|)$ such that $\tilde{\mathcal{Q}}(X_{l,1}, 1) \leq \alpha_{l,1}, \ldots, \tilde{\mathcal{Q}}(X_{l,n_l}, 1) \leq \alpha_{l,n_l}$ and $1 \geq \alpha_{l,n_l} \geq \cdots \geq \alpha_{l,1} > 0$. Let $Y_{l,1}, \ldots, Y_{l,n_l}$ be independent random variables with distributions $\nu_{\alpha_{l,1}}^*, \ldots, \nu_{\alpha_{l,n_l}}^*$ respectively. Define a function $\xi_d : \mathbb{R} \to \mathbb{R}$ by setting $\xi_d(x) = x$ if $d = 2$ and $\xi_d(x) = 1$ otherwise. Denote

$$\bar{\alpha}_l = n_l^{-1} \sum_{i=1}^{n_l} \alpha_{l,i} \quad and \quad V_{l,t}^* = \sum_{i=1}^{t} \mathbb{E}\, Y_{l,i}^2.$$

Suppose $V_{l,n_l}^* \to \infty$ as $l \to \infty$,

$$\xi_d(\bar{\alpha}_l)^2 V_{l,n_l}^* = o(n_l^2), \qquad\qquad V_{l,n_l}^* \leq e^{\frac{C^2}{36} n_l^{\frac{1}{2}}},$$

$$\sum_{i=1}^{n_l} \mathbb{E}\, |Y_{l,i}|^3 = o((V_{l,n_l}^*)^{\frac{3}{2}}), \qquad\qquad \xi_d(\bar{\alpha}) \bar{\alpha}_l^2 n_l = o((V_{l,n_l}^*)^{3/2})$$

and

$$\lim_{\epsilon \downarrow 0} \lim_{l \to \infty} \frac{V_{l, \lceil n_l(1-\epsilon) \rceil}^*}{V_{l,n_l}^*} = 1.$$

Then

$$\tilde{\mathcal{Q}}(X_{l,1} + \cdots + X_{l,n_l}, 1) \le \tilde{\mathcal{Q}}(Y_{l,1} + \cdots + Y_{l,n_l}, 1)(1 + o(1))$$

$$= \frac{1}{\sqrt{2\pi V^*_{l,n_l}}}(1 + o(1)) \tag{5}$$

$$\le \sqrt{\frac{6}{\pi}} \frac{\bar{\alpha}_l}{\sqrt{(1 - \bar{\alpha}_l^2)n_l}}(1 + o(1)). \tag{6}$$

Notice that the bound in (5) is like what we would expect from a Gaussian distribution with variance V^*_{l,n_l}. In some simple cases (5) follows by a standard local limit theorem. However, when, for example, part of $\alpha_{l,i}^{-1}$ are odd integers and the remaining ones are even integers, the distribution of $Y_{l,1} + \cdots + Y_{l,n_l}$ is a mixture of distributions on two different lattices and we need a different argument (see Lemma 10.1 below).

Write $a_n \gg b_n$ if $b_n = o(a_n)$ as $n \to \infty$.

Remark 2.4 *In the case of the uniform bound for all random variables, i.e. $\alpha_{i,l} = \bar{\alpha}_l$ for all $i \in \{1, \ldots, n_l\}$ the conditions of Theorem 2.3 simplify to*

$$\bar{\alpha}_l \gg n_l^{-\frac{1}{2}} \quad and \quad 1 - \bar{\alpha}_l \gg n_l^{-\frac{1}{3}}.$$

(When $d = 2$ the second condition can be weakened to $\bar{\alpha}_l \ge e^{-\frac{C^2}{71}\sqrt{n}}$.)

It would be tempting to drop the $(1 + o(1))$ factors and the assumptions of our results and formulate the following.

Conjecture 2.5 *Let $X_1, \ldots, X_n$ be independent random vectors with values in a normed space $\mathcal{S}$. Suppose $\tilde{\mathcal{Q}}(X_1, 1) \le \alpha_1, \ldots, \tilde{\mathcal{Q}}(X_n, 1) \le \alpha_n$, where $\alpha_1, \ldots, \alpha_n \in (0, 1]$. Then*

$$\tilde{\mathcal{Q}}(X_1 + \cdots + X_n, 1) \le \mathbb{P}(Y_1 + \cdots + Y_n \in \{0, \frac{1}{2}\})$$

where $Y_1, \ldots, Y_n$ are independent and Y_i has distribution $\nu^*_{\alpha_i}$ defined in (2).

In particular, the case $\alpha_1 = \cdots = \alpha_n = \frac{1}{2}$ corresponds exactly to the conjecture by Leader and Radcliffe and $\alpha_1 = \cdots = \alpha_n = \frac{1}{k}$ is a generalization of the question of Jones.

We now show that this natural conjecture is false, even for $n = 2$ and $\alpha_1 = \alpha_2$. A counterexample for 2-dimensional Euclidean space is when both X_1 and X_2 are chosen uniformly at random from the vertices of a regular octagon with radius $(\sqrt{2} + 2)^{-\frac{1}{2}} \approx 0.5412$. Then $\mathcal{Q}(X_1 + X_2, 1) = \tilde{\mathcal{Q}}(X_1 + X_2, 1) = \tilde{\mathcal{Q}}(X_1, 1) = \frac{3}{8}$. This value is a sharp upper bound when $\alpha_1 = \alpha_2 = \frac{3}{8}$ since $\tilde{\mathcal{Q}}(X_1 + X_2, 1) \le \tilde{\mathcal{Q}}(X_1, 1)$. It is also strictly larger than $\tilde{\mathcal{Q}}(Y_1 + Y_2, 1)$ for $Y_1, Y_2 \sim \nu^*_\alpha$ and $\alpha \in [\frac{3}{8}, \frac{5}{12})$. Non-discrete counterexamples may be obtained when X_1 and X_2 are uniformly distributed on a circle in $\mathbb{R}^2$ of an appropriate radius.

Although we have just demonstrated that the worst case 1-dimensional examples do not in general maximize $\tilde{\mathcal{Q}}(\cdot, 1)$, they are still optimal asymptotically by Theorem 2.3 at least when the conditions of that theorem hold.

We can trivially increase n in our counterexamples by appending any number of α_is that are close to 1. But we have neither non-trivial constructions for $n > 2$ nor ones where each α_i^{-1} is integer. So Conjecture 2.5 might still be true when all α_i^{-1} are integers.

3 Difficulties, the approach and an outline of the proofs

Let $X_1, ..., X_n$ be independent random vectors in a normed space $\mathcal{S}$ equipped with a norm $|| \cdot ||$. Suppose that

$$\tilde{\mathcal{Q}}(X_i, 1) \le \alpha_i, \quad i \in \{1, \ldots, n\}. \tag{7}$$

The first typical step in doing optimization with such random variables in our context is passing to a much smaller class of distributions. The probabilities of the form $\mathbb{P}(X_1 + \cdots + X_n \in A)$ are linear in each distribution. Thus one example of such a class would be the extreme points of the set of measures satisfying (7). (Recall that a point is extreme if it does not lie in the interior of any segment contained in the set.) The extreme measures are often quite simple or are contained in another simple set. If we can prove a desired inequality when each distribution is extreme, then the inequality holds for any distributions satisfying (7).

As it was shown implicitly by Leader and Radcliffe [17], any distribution on $\mathbb{R}$ satisfying (7) with $\alpha_i = \frac{1}{k}$ can be expressed as a convex combination of uniform distributions on k-blocks. This then allowed them to pass to working with just such distributions and use tools from extremal combinatorics which originate from Erdős's well-known approach [7].

Answering a question of Erdős, Kleitman [14] established the sharp upper bound for $\tilde{\mathcal{Q}}$ for sums of random vectors each distributed on some 2-block. Regrettably, we were unable to use the ideas of this short and simple proof, due to the fact that in some sense there are "too many" extreme points and in some sense only the minority are distributions on 2-blocks. It is not difficult to come up with examples of distributions in the plane that are extreme in the set of measures satisfying (7) with $\alpha_i = \frac{1}{2}$ and that cannot be written as convex combinations of distributions on 2-blocks. Consider, for example, the uniform distribution on the vertices of an equilateral triangle with a side length of 1 and the point at its center. It has concentration exactly $\frac{1}{2}$ and it is not difficult to check that it is extreme.

As it turns out, in the latter example it actually makes sense to decompose the measure into a delta measure on the center of the triangle and the uniform distribution on the three vertices of the triangle (a 3-block). The first measure does not satisfy (7) with $\alpha_i = \frac{1}{2}$, but the second distribution satisfies it with

$\alpha_i = \frac{1}{3}$. In this work we show how to decompose arbitrary measures into not necessarily equally sized blocks in an optimal way so that traditional techniques for proving concentration inequalities can be applied.

We start by approximating the underlying distributions with discrete ones. We then examine the properties of their distance graphs (we join points of the support of the distribution if they are at a distance < 1 in the underlying norm). We show that those distributions that are close to some line in $\mathcal{S}$ have distance graphs with no induced odd cycles of length ≥ 5 and the same is true for the complements of such graphs. By the strong perfect graph theorem proved by Chudnovsky, Robertson, Seymour and Thomas [4] such graphs are perfect, i.e., for all their induced subgraphs the chromatic number is equal to the clique number.

The condition (7) provides a bound on a clique number of the distance graph. For instance, if the distribution under consideration is uniform and it satisfies (7) with $\alpha_i = \frac{1}{2}$, it means that the distance graph of its support has no clique larger than half of the number of its vertices. A vertex-colouring of the distance graph provides a decomposition of the distribution into blocks where each block corresponds to a colour class. Since in our case the clique number and the chromatic number coincide by the strong perfect graph theorem, we obtain a decomposition of our distribution into the optimal number of blocks.

This allows us to obtain optimal results in this restrictive setting for arbitrary normed spaces, not just $\mathbb{R}^d$ (see Theorem 2.2).

The remaining case is settled using a classical result by Halász [9]. We use it to show that if $\tilde{\mathcal{Q}}$ of our sum is at least the value we are trying to establish (which is of order $n^{-\frac{1}{2}}$ in the i.i.d. case), then most of the distributions can be made close to the same line by shifting.

The proofs are organized as follows. The relevant discrete approximation which allows us to pass to finitely supported distributions is provided in Section 4. We then turn to establishing properties of distance graphs of finite sets close to a line. For methodical reasons we treat the case of the Euclidean norm first in Section 5 and a generalization is given in Section 7. In Section 6 we present our result on the decomposition of measures. In Section 7 we prove that blocks that are close to a line satisfy, the so called, B.T.K. chain condition, see [3, 10, 17]. In Section 8 we use this property in a similar way as Jones [10] and our decomposition to prove Theorem 2.2. Next, we proceed to the general case of distributions that are not necessarily close to a line. The tools from the work of Halász are presented and adapted to our setting in Section 9. Finally, in Section 10 we prove a version of local limit theorem, Lemma 10.1, for our extremal distributions, and combine all of the above to prove the main result, Theorem 10.3. Theorems 1.1, 2.1 and 2.3 are then obtained as corollaries of Theorem 10.3.

4 Discrete approximation

A natural way to approximate an arbitrary probability distribution μ by a discrete one is to use empirical distributions [21]. In general this is not enough to ensure continuity of the concentration function [6]. For example, let X be distributed uniformly on a unit circle in the plane. The empirical distribution X_n of n i.i.d copies of X has $\tilde{\mathcal{Q}}(X_n, 2) = 1$ almost surely, while $\tilde{\mathcal{Q}}(X, 2) = \frac{1}{2}$. See [6] for a more general result. For concentration a general way out turns out to be sampling from $(1 + \delta)X$ instead.

If a random vector X has distribution μ, it will be sometimes be more convenient to use the notation $\tilde{\mathcal{Q}}(\mu, \cdot)$ for $\tilde{\mathcal{Q}}(X, \cdot)$, and similarly for $\mathcal{Q}$. We write $\mu_n \Rightarrow \mu$ if a sequence of measures (μ_n) converges weakly to a measure μ, see [1].

Lemma 4.1 *Let X be a random vector with values in a complete separable normed space $\mathcal{S}$. Suppose $\tilde{\mathcal{Q}}(X, 1) = \alpha$. Fix $\delta > 0$ and let μ be the distribution of $(1 + \delta)X$. For each n define the empirical measure $M_n = M_{\omega, n}$, such that given a realization ω of i.i.d. random vectors $\xi_1, \xi_2, \ldots$ distributed according to μ, $M_{\omega, n}(A) = n^{-1}|\sum_{i=1}^{n} \mathbf{1}_A(\xi_i(\omega))|$ for each Borel set A. Then almost surely*

$$M_n \Rightarrow \mu \tag{8}$$

and

$$\limsup_{n \to \infty} \tilde{\mathcal{Q}}(M_n, 1) \leq \alpha. \tag{9}$$

Proof TOPROVE 0 □

Lemma 4.2 *Let $X_1, \ldots, X_m$ be independent random vectors with values in a complete separable normed space $\mathcal{S}$. Suppose $\tilde{\mathcal{Q}}(X_i, 1) \leq \alpha$ and write $S = X_1 + \cdots + X_m$. Let $\epsilon > 0$. Then there is $\delta_0 > 0$ such that for any $\delta \in (0, \delta_0)$ there exist random vectors $\{Y_{i,n}, i \in \{1, \ldots, m\}, n \in \{1, 2, \ldots\}\}$ such that $Y_{i,n}$ has the uniform distribution on a multiset of cardinality n, for each n the vectors $Y_{1,n}, \ldots, Y_{m,n}$ are independent, for each $i \in \{1, \ldots, m\}$ $Y_{i,n} \Rightarrow (1+\delta)X_i$, $\limsup_n \tilde{\mathcal{Q}}(Y_{i,n}, 1) \leq \alpha$ as $n \to \infty$ and for $T_n = Y_{1,n} + \cdots + Y_{m,n}$ we have $\liminf_n \tilde{\mathcal{Q}}(T_n, 1) \geq \tilde{\mathcal{Q}}(S, 1) - \epsilon$ as $n \to \infty$.*

Proof TOPROVE 1 □

5 Euclidean distance graphs for points close to a line

We define a *distance graph* $G = G(x_1, \ldots, x_N)$ of a sequence of N points $x_1, \ldots, x_N$ in a metric space $\mathcal{S}$ with distance d by connecting two points if they are at distance less than 1:

$$V(G) = \{x_1, \ldots, x_N\}; \qquad E(G) = \{x_i x_j : i \neq j \text{ and } d(x_i, x_j) < 1\}.$$

We allow $x_i = x_j$ for $i \neq j$, in this case vertices can be distinguished by their index, and similarly for edges. So the graph in all cases is simple.

The distance in the above definition implicitly depends on the metric or the norm induced by the space $\mathcal{S}$; in this section we start with Euclidean spaces with (l_2) norm $\| \cdot \|_2$ and (l_2) distance $d_2(x, y) = \|x - y\|_2$.

We will need the following simple fact.

Lemma 5.1 *Let $r = (r_1, \ldots, r_d)^T$ and $s = (s_1, \ldots, s_d)^T$ be two points in $\mathbb{R}^d$ such that $\|s - r\|_2 \geq 1$, and both r and s are at l_2 distance at most $\frac{\sqrt{3}}{4}$ from the the line L spanned by the vector $e^1 = (1, 0, \ldots, 0)^T$ (the "x axis"). Denote by $r^* = \langle r, e^1 \rangle e^1$ and $s^* = \langle s, e^1 \rangle e^1$ their projections onto L. Then*

$$|r_1 - s_1| = \|r^* - s^*\|_2 \geq \frac{1}{2}.$$

Proof TOPROVE 2 □

Lemma 5.2 *Let G be a distance graph of points $x_1, \ldots, x_N$ in $\mathbb{R}^d$, and assume each of the points is of distance at most $\frac{\sqrt{3}}{4}$ from the line generated by the vector $(1, 0, \ldots, 0)^T$ (the "x axis"). Then $\bar{G}$ does not contain an induced cycle of odd length 5 or more.*

Proof TOPROVE 3 □

Lemma 5.3 *Let G be a distance graph of points $x_1, \ldots, x_N$ in $\mathbb{R}^d$, and assume each of the points is of distance at most $\frac{\sqrt{3}}{4}$ from the line generated by the vector $(1, 0, \ldots, 0)^T$ (the "x axis"). Then G does not contain an induced cycle of length 5 or more.*

Proof TOPROVE 4 □

Corollary 5.4 *A distance graph G of a sequence of points in $\mathbb{R}^d$ with Euclidean distance at most $\frac{\sqrt{3}}{4}$ from a line is a perfect graph.*

Proof TOPROVE 5 □

By adding $(0, -\frac{\sqrt{3}}{4})$ to each of the points $(-1, 0)$, $(0, 0)$, $(1, 0)$, $(\frac{1}{2}, \frac{\sqrt{3}}{2})$, $(-\frac{1}{2}, \frac{\sqrt{3}}{2})$ and slightly perturbing them it is easy to see that the constant $\frac{\sqrt{3}}{4}$ is best possible in Lemma 5.3 and Corollary 5.4.

In Section 7 we extend the corollary to arbitrary normed spaces over the reals.

Lemma 5.5 *Let $\mathcal{S}$ be a normed space over the reals. A distance graph G of a sequence of points in $\mathcal{S}$ with distance at most $\frac{1}{8}$ (at most $\frac{\sqrt{3}}{4}$ if $\mathcal{S}$ is Hilbert) from a line in $\mathcal{S}$ is a perfect graph.*

9

6   Decomposition to uniform measures

Let $S = (s_1, \ldots, s_N)$ be a finite sequence in a metric space. If we ignore the order, S can be viewed as a multiset, so we denote $|S| = N$ and define the uniform distribution ν on S by $\nu(\{x\}) = |S|^{-1}|\{i \in \{1, \ldots, |S|\} : x_i = x\}|$.

Lemma 6.1 *Let $\mathcal{S}$ be a metric space with distance d. Let X be a random vector distributed uniformly on a finite sequence S in $\mathcal{S}$. Let G be the distance graph of S. Then $\tilde{\mathcal{Q}}(X, 1) = \frac{\omega(G)}{|S|}$ where $\omega(G)$ is the clique number of G.*

Proof TOPROVE 6 □

For (multi-)sets A and B denote by $A \uplus B$ the operation of multiset union, so the multiplicity of x in $A \uplus B$ is the sum of multiplicities of x in A and B respectively.

Lemma 6.2 *Let $\mathcal{S}$ be a normed space over the reals with norm $\|\cdot\|$. Let X be a random vector distributed uniformly on a finite multiset $S \subset \mathcal{S}$.*
Suppose $\tilde{\mathcal{Q}}(X, 1) \leq \alpha$ and there is a line L in $\mathcal{S}$ such that S is at distance at most $\frac{1}{8}$ (at most $\frac{\sqrt{3}}{4}$ in the case $\mathcal{S}$ is Hilbert) from L.
Then there exist $K \leq \alpha|S|$ and sets $S_1, \ldots, S_K$ that partition S $(S = \uplus_{j=1}^{K} S_j)$ such that for each $j \in \{1, \ldots, K\}$ and each $x, y \in S_j$, $x \neq y$ we have $\|x - y\| \geq 1$.

Proof TOPROVE 7 □

Recalling that $\nu(S)$ denotes a uniform measure on a (multi-)set S, the lemma says that $X \sim \nu(S)$ and $\nu(S) = \frac{\sum_{j=1}^{K} |S_j| \nu(S_j)}{|S|}$, that is ν is a mixture of uniform measures on blocks S_j.

7   Arbitrary norms and the B. T. K. chain condition for sets close to a line

The main result of this section is Lemma 7.3 which extends a result of Jones [10] for $d = 1$ to general normed spaces and distributions close to a line.

We will need the following concept, which was also used in [17]. Let v be a unit vector in a normed space $\mathcal{S}$ over the reals. We say that f is a *supporting functional* for v if f is a bounded linear functional on $\mathcal{S}$ such that $|f(x)| \leq \|x\|$ for all $x \in \mathcal{S}$ and $f(v) = 1$.

Note that f as above satisfies $|f(x)| = \|x\|$ for $x \in L_v$ where $L_v = \{tv : t \in \mathbb{R}\}$. The Hahn–Banach theorem can be used to obtain a supporting functional for any unit vector v in any normed space $\mathcal{S}$ over the reals: first define $f(x) = t$ if $x = tv$ for all $x \in L_v$, then extend f to whole $\mathcal{S}$ while preserving its norm $\|f\| = 1$.

In the case when the space is Hilbert, we can easily generalize Lemma 5.1 and allow a larger distance from the line.

Lemma 7.1 *Let r and s be two points in a Hilbert space $\mathcal{S}$ with norm $\|\cdot\|$ and inner product $\langle\cdot,\cdot\rangle$. Let v be a unit vector in $\mathcal{S}$. Suppose $\|s - r\| \geq 1$, and both r and s are at distance at most $\frac{\sqrt{3}}{4}$ from the the line L_v generated by v. Denote by $r^* = \langle r, v\rangle v$ and $s^* = \langle s, v\rangle v$ their projections onto L. Then*

$$|\langle r, v\rangle - \langle s, v\rangle| = \|r^* - s^*\| \geq \frac{1}{2}.$$

Proof TOPROVE 8 □

The next lemma extends this for general normed spaces.

Lemma 7.2 *Let $\mathcal{S}$ be a normed space over the reals with norm $\|\cdot\|$. Let $v \in \mathcal{S}$ with $\|v\| = 1$, let L_v be the line generated by v and let f be a supporting functional for v. If $x, y \in \mathcal{S}$ have distance at most $\frac{1}{8}$ from L_v and $\|x - y\| \geq 1$ then $|f(x) - f(y)| \geq \frac{1}{2}$.*

Furthermore, if $\mathcal{S}$ is Hilbert with inner product $\langle\cdot,\cdot\rangle$, we can take $f(x) = \langle x, v\rangle$ for all $x \in \mathcal{S}$ and the statement still holds for x, y with distance at most $\frac{\sqrt{3}}{4}$ from L_v.

Proof TOPROVE 9 □

l_1 norm in the plane can be used to show that $\frac{1}{8}$ in Lemma 7.2 is optimal. We can now generalize the proof of Corollary 5.4.

Proof of Lemma 5.5 Let $v \in \mathcal{S}$ be a unit vector such that all points are at distance at most $\frac{1}{8}$ (at most $\frac{\sqrt{3}}{4}$ when $\mathcal{S}$ is Hilbert) from the line L_v generated by v.

Lemma 5.2 and Lemma 5.3 (with $(1, 0, \ldots, 0)^T$ replaced with v) can be generalized to the Hilbert setting using Lemma 7.1 and to the general setting using Lemma 7.2 instead of Lemma 5.1. In the Hilbert case the projection to the first coordinate $a \mapsto a_1$ can be replaced with the mapping to the inner product $a \mapsto \langle a, v\rangle$, while in the general case it can be replaced with $a \mapsto f_v(a)$ where f_v is the supporting functional for v. Analogous inequalities as those used in the proofs of Lemma 5.2 and Lemma 5.3 remain correct by the definition of the supporting functional.

Finally use the same argument as in Corollary 5.4. □

Let $\mathcal{S}$ be a normed space with norm $\|\cdot\|$. Following [10, 17], we say that sets $A, B \subset \mathcal{S}$ satisfy *B.T.K. chain condition* if the set $A \times B$ can be decomposed into n sets $S_{m-n+1}, S_{m-n+3}, \ldots, S_{m+n-1}$, $m = \max(|A|, |B|)$, $n = \min(|A|, |B|)$ such that $|S_t| = t$ and for each distinct $(x, y), (u, v) \in S_t$ we have $\|(x+y) - (u+v)\| \geq 1$, that is, each S_t is a *t-block*. When we refer to $S \subseteq A \times B$ as a set of *points* we view the pairs $(x, y) \in S$ as elements $x + y \in \mathcal{S}$ like Jones [10].

Lemma 7.3 *Let $A = \{x_1, \ldots, x_a\}$ and $B = \{y_1, \ldots, y_b\}$ be non-empty sets in a normed space $\mathcal{S}$ over the reals with norm $\|\cdot\|$. Suppose that for each pair of distinct points $x, y \in A$ or $x, y \in B$ we have $\|x - y\| \geq 1$.*

Let υ be a unit vector in $\mathcal{S}$ and let f be a supporting functional for υ. Assume that $f(x_1) \leq \cdots \leq f(x_a)$, $f(y_1) \leq \cdots \leq f(y_b)$ and $|f(x) - f(y)| \geq \frac{1}{2}$ for distinct $x, y \in A$ or $x, y \in B$.

Then A and B satisfy the B.T.K. chain condition and furthermore for each set S_t in the decomposition and any distinct $x, y \in S_t$ we have

$$|f(x) - f(y)| \geq \frac{1}{2}. \tag{10}$$

Proof TOPROVE 10 ☐

8 Concentration for sums of random vectors close to a line

The next lemma extends the results of Jones [10] in the case when random variables are close to a line.

Lemma 8.1 *Let $\mathcal{S}$ be a normed space over the reals with norm $\| \cdot \|$. Let $X_1, \ldots, X_n$ be independent random vectors, such that X_i is distributed uniformly on a k_i-block $S_i \subset \mathcal{S}$ (that is, $|S_i| = k_i$ and each pair of distinct points $x, y \in S_i$ satisfies $\|x - y\| \geq 1$).*

Suppose there is a line L in $\mathcal{S}$ such that for each $i \in \{1, \ldots, n\}$ each point $x \in S_i$ is at distance at most $\frac{1}{8}$ from L (at most $\frac{\sqrt{3}}{4}$ if $\mathcal{S}$ is Hilbert).

Then

$$\tilde{\mathcal{Q}}(X_1 + \cdots + X_n, 1) \leq \mathbb{P}(Y_1 + \cdots + Y_n \in \{0, \frac{1}{2}\})$$

where $Y_1, \ldots, Y_n$ are independent random variables and Y_i is uniformly distributed on $\{1, \ldots, k_i\} - \frac{k_i+1}{2}$, i.e., in our notation $Y_i \sim \nu^*_{k_i^{-1}}$.

Proof TOPROVE 11 ☐

For a sequence $\alpha_1, \ldots, \alpha_n \in (0, 1]$ define $t(\alpha_1, \ldots, \alpha_n) := \mathbb{P}(Y_1 + \cdots + Y_n \in \{0, \frac{1}{2}\})$, where $Y_i \sim \nu^*_{\alpha_i}$ and $Y_1, \ldots, Y_n$ are independent.

Recall that $\mu_1 * \mu_2$ denotes the convolution of μ_1 and μ_2 and $*_{i=1}^n \mu_i$ denotes $\mu_1 * \cdots * \mu_n$.

Proof of Theorem 2.2 We first prove ii). Let $c_\mathcal{S}$ be $\frac{\sqrt{3}}{4}$ if $\mathcal{S}$ is Hilbert and $\frac{1}{8}$ otherwise. Let $t = \max_i \sup_\omega d(X_i(\omega), L)$. Fix $\epsilon \in (0, \frac{\min_i \alpha_i}{2})$. By Lemma 4.2 (and the construction given in Lemma 4.1) there exist δ, $0 < \delta < \frac{c_\mathcal{S}}{t} - 1$ and random variables Z_i, $i \in \{1, \ldots, n\}$ such that

(a) Z_i is uniformly distributed on a finite multiset S_i which lies in the support of $(1+\delta)X_i$, so $\sup_\omega d(Z_i, (1+\delta)L) \leq c_\mathcal{S}$,

(b) $\tilde{\mathcal{Q}}(Z_i, 1) \leq \alpha_i + \epsilon$ and

(c) $\tilde{\mathcal{Q}}(Z_1 + \cdots + Z_n, 1) \geq \tilde{\mathcal{Q}}(X_1 + \cdots + X_n, 1) - \epsilon$.

12

Let $\mu_i \sim \nu(S_i)$ be the distribution of Z_i. By our decomposition, Lemma 6.2, we have for each $i \in \{1, \ldots, n\}$

$$\mu_i = w_{i,1}\mu_{i,1} + \ldots + w_{i,K_i}\mu_{i,K_i}$$

where $w_{i,j} = \frac{k_{i,j}}{N_i}$, $N_i = |S_i|$, $K_i \leq (\alpha_i + \epsilon)N_i$ and $\mu_{i,j}$ is the uniform measure on a *set* $S_{i,j}$ of size $k_{i,j}$ with each pair of points at distance at least 1 (a $k_{i,j}$-block).
We have

$$\mu := *_{i=1}^n \mu_i = \sum \left(\prod_{i=1}^n w_{i,j_i} \right) *_{i=1}^n \mu_{i,j_i}.$$

where the sum is over all $(j_1, \ldots, j_n) \in \{1, \ldots, K_1\} \times \cdots \times \{1, \ldots, K_n\}$.

Since each μ_{i,j_i} is uniform on a k_{i,j_i}-block, by (a) and Lemma 8.1 for any open set B of diameter at most 1:

$$(*_{i=1}^n \mu_{i,j_i})(B) \leq (*_{i=1}^n \nu^1_{k_{i,j_i}})(\{0, \frac{1}{2}\})$$

where ν^1_s is a uniform measure on $\{1, \ldots, s\} - \frac{s+1}{2}$. So

$$\mu(B) \leq (*_{i=1}^n \mu_i^1)(\{0, \frac{1}{2}\})$$

where

$$\mu_i^1 = w_{i,1}\nu^1_{k_{i,1}} + \cdots + w_{i,K_i}\nu^1_{k_{i,K_i}}.$$

Now the concentration of μ_i^1 satisfies

$$\tilde{\mathcal{Q}}(\mu_i^1, 1) \leq \sum_{j=1}^{K_i} w_{i,j} \frac{1}{k_{i,j}} = \sum_{j=1}^{K_i} \frac{k_{i,j}}{N_i} \frac{1}{k_{i,j}} = \frac{K_i}{N_i} \leq \alpha_i + \epsilon.$$

We have reduced the problem to the case $\mathcal{S} = \mathbb{R}^1$. By Theorem 2.1.3 of Juškevičius [12] for any $n \geq 1$:

$$\tilde{\mathcal{Q}}(X_1 + \cdots + X_n, 1) \leq \tilde{\mathcal{Q}}(*_{i=1}^n \mu_i^1, 1) \leq t(\alpha_1 + \epsilon, \ldots, \alpha_n + \epsilon).$$

Since t is continuous and our inequality holds for any $\epsilon \in (0, \frac{\min_i \alpha_i}{2})$, the claim ii) follows.

Finally, to prove the theorem in the case i), notice that in this case the support of each random variable as well as the support of their sum is contained in a countable set. The closure of a countable set is separable. So i) follows from ii). $\square$

9 A corollary of a result of Halász

In this section we temporarily switch to the convention of Halász [9] and consider concentration for diameter 2 rather than 1. Results in this section concern only $\mathbb{R}^d$ with the Euclidean norm $\| \cdot \|_2$ and distance $d_2(x, y) = \|x - y\|_2$.

Theorem 9.1 *(Halász (1977), Theorem 4 in [9]) Let X'_k be independent of X_k but of the same distribution, $X^*_k = X_k - X'_k$ and denote by $F_k(x)$ the distribution function of X^*_k. Let*

$$F(x) = \sum_{k=1}^{n} F_k(x).$$

Introducing the notation

$$a_* = \min(a, 1) \quad (a \geq 0)$$

let

$$D = \inf_{|e|=1} \int_{\mathbb{R}^d} |\langle x, e \rangle|^2_* dF(x)$$

and

$$\mu = \sup_{y \in \mathbb{R}^d} \sum_{k=1}^{n} \mathbb{P}(\|X^*_k - y\|_2 < 1).$$

We have

$$\mathcal{Q}(\sum_{k=1}^{n} X_k, 2) \leq c_{H,d}\, \mu n^{-1} D^{-d/2} \quad (n \geq 8).$$

The constant $c_{H,d}$ depends on d only.

Lemma 9.2 *Let $d \geq 2$ be an integer. Let $\xi_2(x) = x$ and let $\xi_d(x) = 1$ for $d > 2$.*
Let $c \in (0,1)$, $\delta > 0$, $\epsilon > 0$ and $n \geq 8$. If $X_1, \ldots, X_n$ are independent random vectors in $\mathbb{R}^d$ with $\mathcal{Q}(X_i, 2) \leq \alpha_i$, $\alpha_i \in (0,1]$, and $\frac{\sum \alpha_i}{n} = \bar{\alpha}$ then either $\tilde{\mathcal{Q}}(\sum_{i=1}^{n} X_i, 2) < \delta$ or there exists a line L in $\mathbb{R}^d$ such that L contains the origin and for all but at most

$$9(d-1)^3 c_H c^{-2} \delta^{-1} \epsilon^{-1} \xi_d(\bar{\alpha})$$

indices $i \in \{1, \ldots, n\}$ we have $\mathbb{P}(d_2(X_i, L + a_i) \geq c) \leq \epsilon$ with some $a_i \in \mathbb{R}^d$. Here $c_H = c_{H,2}$ from Theorem 9.1.

Halász [9] notes after his theorem: *"D measures to what extent our distributions are d-dimensional. d-dimensionality can be measured even in R^{d_1}, $d_1 > d$: $|e| = 1$ should be replaced by subspaces of dimension $d_1 - d + 1$ and $\langle x, e \rangle$ by projection of x on the subspace.".* This remark implies that Lemma 9.2 should be true with $\bar{\alpha}$ in place of ξ_d for all d and some other d-dependent constant. The proof of the remark is not given in [9], and it seems to require a certain technical effort to produce that proof, hence we state our results with $\xi_d(\bar{\alpha})$ rather than $\bar{\alpha}$, which is likely suboptimal for $d \geq 3$.

Proof of Lemma 9.2 Assume $n \geq 8$. Let us first prove the claim for $d = 2$. A disk of radius 2 can be easily covered by 9 open disks of radius 1. To see this, center each of the smaller disks on a point in the 3x3 grid with step size, for

example, $\sqrt{2} - 0.01$. So $\tilde{\mathcal{Q}}(\cdot, 2) \leq \mathcal{Q}(\cdot, 4) \leq 9\mathcal{Q}(\cdot, 2)$. Using Theorem 9.1 and the notation given in its statement

$$\mathcal{Q}(\sum_{i=1}^{n} X_i, 2) \leq c_H \bar{\alpha} D^{-1}. \tag{11}$$

Write

$$D_i = \mathbb{E} |\langle X_i^*, e \rangle|_*^2$$

for the unit vector e that achieves the minimum in the definition of D, so that $D = \sum_{i=1}^{n} D_i$. If $D > 9c_H \delta^{-1}\bar{\alpha}$ then $\mathcal{Q}(\sum_{i=1}^{n} X_i, 2) < \delta/9$ and the claim follows.

Otherwise there can be at most $9c_H c^{-2}\delta^{-1}\epsilon^{-1}\bar{\alpha}$ indices i such that

$$D_i \leq c^2\epsilon \tag{12}$$

does not hold.

Let L be the line containing the origin and orthogonal to e. Recall that $X_i^* \sim X_i - X_i'$ where X_i' is an independent copy of X_i. By conditioning and the fact that $|x|_*^2 \geq \min(c^2, 1)\mathbb{I}_{|x| \geq c} = c^2 \mathbb{I}_{|x| \geq c}$

$$D_i = \mathbb{E} |\langle X_i^*, e \rangle|_*^2$$
$$\geq \inf_{a' \in \mathbb{R}^2} \mathbb{E} |\langle X_i - a', e \rangle|_*^2$$
$$= \mathbb{E} d_2(X_i, L + a_i)_*^2$$
$$\geq \mathbb{P}(d_2(X_i, L + a_i) \geq c)c^2$$

for a vector $a_i \in \mathbb{R}^d$ that realises the infimum. Thus for any i such that (12) holds we have

$$\mathbb{P}(d_2(Y_i, L + a_i) \geq c) \leq c^{-2}D_i \leq \epsilon,$$

which completes the proof for the case $d = 2$.

Let us now prove the claim for $d > 2$. If $\tilde{\mathcal{Q}}(X_1 + \cdots + X_n, 2) < \delta$ we are done. Otherwise, choose any plane $P_1 \subset \mathbb{R}^d$ containing the origin and let π_1 be the projection of $\mathbb{R}^d$ to P_1. For a random vector X in $\mathbb{R}^d$ we have $\tilde{\mathcal{Q}}(\pi_1 X, 2) \geq \tilde{\mathcal{Q}}(X, 2)$. Thus $\tilde{\mathcal{Q}}(\pi_1(X_1 + \cdots + X_n), 2) \geq \delta$. For the summands we only have the trivial upper bound $\tilde{\mathcal{Q}}(\pi_1 X_i, 2) \leq 1$.

Let $m = 9(d-1)^2 c_H c^{-2}\delta^{-1}\epsilon^{-1}$. By the already proved case $d = 2$ for $j = 1$ there is a unit vector $\tilde{e}_j \in P_j$ and vectors $a_{j,1}, \ldots, a_{j,n} \in P_j$ such that such that for all but at most m indices $i \in \{1, \ldots, n\}$

$$\mathbb{P}(|\langle X_i - a_{j,i}, \tilde{e}_j \rangle| \geq \frac{c}{\sqrt{d-1}}) \leq \frac{\epsilon}{d-1}. \tag{13}$$

Let us do similarly for each $j = 2, \ldots, d-1$: at step j let $\hat{e}_{j-1}$ be a unit vector in P_{j-1} orthogonal to $\tilde{e}_{j-1}$. Since $j \leq d-1$ we can pick a unit vector $e_j \in \mathbb{R}^d$ orthogonal to $\tilde{e}_1, \ldots, \tilde{e}_{j-1}$ and $\hat{e}_{j-1}$. Consider a plane $P_j \subset \mathbb{R}^d$ generated by e_j and $\hat{e}_{j-1}$. Again there is a unit vector $\tilde{e}_j$ and vectors $a_{j,1}, \ldots, a_{j,n}$ in P_j such

that for all but at most m indices i (13) holds. Finally, define $\tilde{e}_d$ to be a unit vector orthogonal to $\tilde{e}_1, \ldots, \tilde{e}_{d-1}$.

We can assume that $\tilde{e}_1, \ldots, \tilde{e}_d$ are the standard unit vectors $(1, 0, \ldots, 0)^T$, $\ldots, (0, \ldots, 0, 1)^T$. Define $a_i = ((a_{1,i})_1, \ldots, (a_{d-1,i})_{d-1}, 0)^T$. Now in (13) we can replace $a_{j,i}$ by a_i and in the new coordinate system it is equivalent to

$$\mathbb{P}(|(X_i - a_i)_j| \geq \frac{c}{\sqrt{d-1}}| \leq \frac{\epsilon}{d-1}.$$

It follows by the union bound that if L is the line generated by $\tilde{e}_d$, then for all but at most $(d-1)m$ indices

$$\mathbb{P}(d_2(X_i - a_i, L) \geq c) \leq \epsilon.$$

$\square$

10 The general result and its proof

When α^{-1} is integer ν_α^* is uniform and its variance is $f(\alpha)$, where $f(x) = \frac{1}{12}(x^{-2} - 1)$. From (2) we get $\alpha = \frac{1}{k+1} + \frac{p}{k(k+1)}$, where $k = \lfloor \alpha^{-1} \rfloor$. So for any $\alpha \in (0, 1]$ the variance V_α^* of ν_α^* is a piecewise linear function of α which agrees with f at points k^{-1}, $k \in \{1, 2, \ldots\}$ and is its linear interpolation between subsequent such points. It can also be written as

$$V_\alpha^* = \frac{1}{12}\lfloor \alpha^{-1} \rfloor (1 + \lfloor \alpha^{-1} \rfloor)(3 - \alpha - 2\alpha\lfloor \alpha^{-1} \rfloor). \tag{14}$$

In order to complete our proofs we need precise asymptotics of the concentration of sums of independent variables with distributions $\nu_{\alpha_i}^*$. A standard method for this is the local limit theorem. However, the existing variants we are aware of are difficult to apply when some α_i^{-1} are close to an integer and some are not. Therefore we prove our own lemma using a central limit theorem (Berry–Esseen) and the properties of $\nu_{\alpha_i}^*$.

Write $x = y \pm \epsilon$ as a shorthand for $x \in [y - \epsilon, y + \epsilon]$.

Lemma 10.1 *Let $1 \geq \alpha_1 \geq \cdots \geq \alpha_n > 0$. Let $Y_1, \ldots, Y_n$ be independent random variables with distributions $\nu_{\alpha_1}^*, \ldots, \nu_{\alpha_n}^*$ respectively. Let*

$$V_t^* := \mathbb{E}\, Y_1^2 + \cdots + \mathbb{E}\, Y_t^2 = V_{\alpha_1}^* + \cdots + V_{\alpha_t}^*; \quad V^* := V_n^* > 0.$$

If there is $c \in (0, 1)$ such that

$$V_{\lceil n(1-c) \rceil}^* \geq \frac{1}{2}V^* \tag{15}$$

and $\delta' \in (0, 1)$ such that

$$\mathbb{E}\, |Y_1|^3 + \cdots + \mathbb{E}\, |Y_n|^3 \leq \delta'(V^*)^{3/2} \tag{16}$$

16

then assuming

$$\epsilon' := 405(\delta')^{\frac{1}{2}}c^{-\frac{3}{4}} \leq \frac{1}{2}. \tag{17}$$

we have

$$t(\alpha_1, \ldots, \alpha_n) = \mathbb{P}(Y_1 + \cdots + Y_n \in \{0, \frac{1}{2}\}) = \frac{1 \pm \epsilon'}{\sqrt{2\pi V^*}}.$$

Since V_α^* is convex, Jensen's inequality implies that for a fixed $\bar{\alpha}$ the main asymptotic term $\frac{1}{\sqrt{2\pi V^*}}$ is maximized when $\alpha_1 = \cdots = \alpha_n = \bar{\alpha}$:

Remark 10.2 *Let $Y_1, \ldots, Y_n$, V^* and $\bar{\alpha}$ be as in Lemma 10.1. Then*

$$\frac{1}{\sqrt{2\pi V^*}} \leq \frac{1}{\sqrt{2\pi n V_{\bar{\alpha}}}} \leq \sqrt{\frac{6}{\pi n}} \frac{\bar{\alpha}}{\sqrt{1 - \bar{\alpha}^2}}.$$

Proof TOPROVE 12 □

The next theorem is the most general and most detail one in this paper. It combines all of the results above.

Theorem 10.3 *For each integer $d \geq 2$ and each norm $\| \cdot \|$ on $\mathbb{R}^d$ there is a constant $C = C(d, \| \cdot \|)$ such that the following holds.*

Let $X_1, \ldots, X_n$ be independent random vectors in $(\mathbb{R}^d, \| \cdot \|)$; $n \geq 8$. Suppose $\tilde{Q}(X_i, 1) \leq \alpha_i$ where $1 \geq \alpha_1 \geq \cdots \geq \alpha_n > 0$. Let $Y_1, \ldots, Y_n$ be independent random variables with the extremal distributions $\nu_{\alpha_1}^*, \ldots, \nu_{\alpha_n}^*$ respectively. Write

$$\bar{\alpha} = n^{-1} \sum_{i=1}^n \alpha_i.$$

Let V_t^*, V^* be as in Lemma 10.1. Suppose there is $c \in (0, \frac{1}{3})$ such that

$$V_{\lceil (1-c)n \rceil}^* \geq \frac{3}{4} V^* \tag{18}$$

and conditions (16) and (17) hold for some positive δ' and ϵ' as defined in Lemma 10.1. Suppose further that $\epsilon' \leq \frac{3}{16}$,

$$\xi_d(\bar{\alpha})\bar{\alpha}^2 n \leq (V^*)^{3/2}\gamma \text{ for some } \gamma \leq (10C)^{-2} \tag{19}$$

where $\xi_2(x) = x$ and $\xi_d(x) = 1$ for $d > 2$, and

$$m := C\xi_d(\bar{\alpha})^{\frac{1}{2}} t(\alpha_1, \ldots, \alpha_n)^{-\frac{1}{2}} n^{\frac{1}{2}} < \frac{cn}{5}. \tag{20}$$

Then

$$\tilde{Q}(X_1 + \cdots + X_n, 1) \leq \frac{1 + 6\epsilon' + 4mn^{-1} + C\sqrt{\gamma}}{\sqrt{2\pi V_{n-\lfloor m \rfloor}^*}} + e^{-\frac{m^2}{9n}}.$$

17

Let us prove Lemma 10.1 first.

Proof of Lemma 10.1 Let $S = Y_1 + \cdots + Y_n$.

Write $k_i = \lfloor \alpha_i^{-1} \rfloor$. Then p_i (corresponding to p in (2)) is $p_i = k_i(\alpha_i(k_i + 1) - 1)$.

Let $A = \{i : Y_i \in \frac{1}{2} + \mathbb{Z}\}$. For each fixed subset of indices A' such that $\mathbb{P}(A = A') > 0$ let $S_{A'}$ be distributed as S given the event $A = A'$. Note that $S_{A'}$ has either support on integers (when $|A'|$ is even) or on half-integers (when $|A'|$ is odd). By [13] it is unimodal on its support and log-concave: $\mathbb{P}(S_{A'} = x)^2 \geq \mathbb{P}(S_{A'} = x - 1)\mathbb{P}(S_{A'} = x + 1)$ for all x.

Consequently, S conditioned on the event that $|A|$ is odd (even) is symmetric, unimodal and distributed on half-integers (integers). So S itself is symmetric and unimodal when restricted to half-integers or to integers. Therefore for $a \in \mathbb{Z} \cup (\frac{1}{2} + \mathbb{Z})$ and $t \in \mathbb{Z}$, with $a \geq 0$, $t > 0$

$$\mathbb{P}(S \in \{a, a + \frac{1}{2}\}) \geq \frac{\mathbb{P}(S \in [a, a + t - \frac{1}{2}])}{t} \geq \mathbb{P}(S \in \{a + t - 1, a + t - \frac{1}{2}\}).$$

We will next see that the above monotonicity argument together with a central limit theorem easily yields a useful lower bound. For the upper bound we will additionally need to condition on A, slightly more (but standard) calculations and, crucially, the log-concavity property of $S_{A'}$.

Let us continue the proof. The assumption (16) allows us to apply the Berry–Esseen theorem. This theorem is known to hold with a constant smaller than 1, see, e.g., [20], that is, for any $z \in \mathbb{R}$ we have

$$|\mathbb{P}(\frac{S}{\sqrt{V_n^*}} \leq z) - \Phi(z)| \leq \delta'$$

so

$$\mathbb{P}(S \in (x\sqrt{V^*}, y\sqrt{V^*}]) = \Phi(y) - \Phi(x) \pm 2\delta'.$$

Therefore for any $\tilde{\epsilon} > 0$

$$\mathbb{P}(S \in \{0, \frac{1}{2}\}) \geq \frac{\mathbb{P}(S \in [0, \tilde{\epsilon}\sqrt{V^*}])}{\tilde{\epsilon}\sqrt{V^*} + 1}$$

$$\geq \frac{1}{\tilde{\epsilon}\sqrt{V^*} + 1}(\Phi(\tilde{\epsilon}) - \Phi(0) - 2\delta').$$

By Taylor's theorem for any real x and $h > 0$ we have $\Phi(x + h) = \Phi(x) + h\Phi'(x) + \frac{\Phi''(\xi)}{2}h^2$ for some $\xi \in [x, x + h]$. Since $\Phi'(0) = \frac{1}{\sqrt{2\pi}}$, $|\Phi''(x)| \leq \frac{1}{\sqrt{2e\pi}}$ for all $x \in \mathbb{R}$ and $(1 + h)^{-1} \geq 1 - h$ for $h > 0$ and $(1 - h_1)(1 - h_2) \geq 1 - h_1 - h_2$ for any $h_1, h_2 \geq 0$

$$\mathbb{P}(S \in \{0, \frac{1}{2}\}) \geq \frac{1}{\sqrt{V^*} + \tilde{\epsilon}^{-1}}\left(\frac{1}{\sqrt{2\pi}} - \frac{\tilde{\epsilon}}{2\sqrt{2e\pi}} - 2\delta'\tilde{\epsilon}^{-1}\right)$$

$$\geq \frac{1}{\sqrt{2\pi V^*}}\left(1 - \frac{\tilde{\epsilon}}{2\sqrt{e}} - 2\delta'\tilde{\epsilon}^{-1}\sqrt{2\pi} - \frac{1}{\tilde{\epsilon}\sqrt{V^*}}\right). \tag{21}$$

18

Deferring the choice of $\tilde{\epsilon}$ to the end of the proof, we now turn to the upper bound.

For any $\alpha \in (0,1]$ let Y_α be a random variable with distribution ν_α^* and recall that $V_\alpha^* = \mathbb{E} Y_\alpha^2$. Clearly if $k = \lfloor \alpha^{-1} \rfloor$ then $V_{k^{-1}}^* \leq V_\alpha^* \leq V_{(k+1)^{-1}}^*$. Furthermore, if $k \geq 2$ we have

$$1 \leq \frac{V_{(k+1)^{-1}}^*}{V_{k^{-1}}^*} = \frac{(k+1)^2 - 1}{k^2 - 1} \leq \frac{8}{3}. \tag{22}$$

It is also easy to see that $\mathbb{E} |Y_{k^{-1}}|^3 \leq \mathbb{E} |Y_\alpha|^3 \leq \mathbb{E} |Y_{(k+1)^{-1}}|^3$. We claim that for $k \geq 2$

$$\mathbb{E} |Y_{(k+1)^{-1}}|^3 \leq 8\mathbb{E} |Y_{k^{-1}}|^3. \tag{23}$$

To see this, note that if k is even then $|Y_{k^{-1}}| + \frac{1}{2}$ stochastically dominates $|Y_{(k+1)^{-1}}|$ and $|Y_{k^{-1}}| \geq \frac{1}{2}$, so (23) follows. If k is odd (so $k \geq 3$) we may obtain $\nu_{(k+1)^{-1}}^*$ from $\nu_{k^{-1}}^*$ by shifting the mass of each of the $k - 1$ non-zero atoms x of $\nu_{k^{-1}}^*$ to the atoms $\{x - \frac{1}{2}, x + \frac{1}{2}\}$, so that the mass at each of the new atoms is $\frac{k-1}{k(k+1)}$ and multiplying the measure by $\frac{k}{k-1}$. This implies that $\frac{\mathbb{E} |Y_{(k+1)^{-1}}|^3}{\mathbb{E} |Y_{k^{-1}}|^3} \leq \max_{x \geq 1} \frac{k(x+\frac{1}{2})^3}{x^3(k-1)} \leq \frac{3}{2}(\frac{3}{2})^3 < 8$.

Let t be the largest index such that $k_1 = \cdots = k_t = 1$. Define $V_{t,A}^* := \sum_{i=1}^t \mathbb{E} (Y_i^2 | A)$. $4V_{t,A}^*$ is equal to $|\{1, \ldots, t\} \setminus A|$, which is a sum of t independent Bernoulli random variables with parameters $1 - p_1, \ldots, 1 - p_t$ and $\mathbb{E} V_{t,A}^* = V_t^* = \sum_{i=1}^t \mathbb{E} Y_i^2 = \frac{1}{4} \sum_{i=1}^t (1 - p_i)$. By well known concentration inequalities (e.g. Theorem 2.3 of [18])

$$\mathbb{P}(|V_{t,A}^* - V_t^*| > 0.5V_t^*) \leq 2e^{-\frac{3}{28} V_t^*}. \tag{24}$$

Also since $k_i = 1$ for $i \in \{1, \ldots, t\}$

$$\mathbb{E} \left(\sum_{i=1}^t |Y_i|^3 | A \right) = \frac{1}{2} V_{t,A}^*. \tag{25}$$

Let ϵ be small, but not too small:

$$\epsilon > 2e^{-\frac{3V^*}{112}}. \tag{26}$$

Write $V_A^* = \sum_{i=1}^n \mathbb{E} (Y_i^2 | A)$. Let C be the event that

$$\frac{3}{16} \leq \frac{V_A^*}{V^*} \leq \frac{8}{3c} \quad \text{and} \quad \mathbb{E} \left(\sum_{i=1}^n |Y_i|^3 | A \right) \leq \frac{8}{c} \mathbb{E} \left(\sum_{i=1}^n |Y_i|^3 \right). \tag{27}$$

Here the constant c is from (15). We claim that

$$\mathbb{P}(\bar{C}) \leq \frac{\epsilon}{\sqrt{V^*}}. \tag{28}$$

19

First suppose that the right side of (24) is at most $\frac{\epsilon}{\sqrt{V^*}}$. Then on the event $|V_{t,A}^* - V_t^*| \le 0.5 V_t^*$ we have, using (22)

$$V_A^* = V_{t,A}^* + \sum_{i=t+1}^{n} \mathbb{E}\left(Y_i^2 | A\right) \ge \min\left(\frac{1}{2}, \frac{3}{8}\right) V^* \ge \frac{3V^*}{8},$$

and similarly $V_A^* \le \frac{8}{3} V^*$.

Similarly bounding the contribution of $i \le t$ using (25) and the contribution of $i > t$ using (23), we get that on the event $|V_{t,A}^* - V_t^*| \le 0.5 V_t^*$ we have $\sum_{i=1}^{n} \mathbb{E}\left(|Y_i|^3 | A\right) \le 8 \sum_{i=1}^{n} \mathbb{E}|Y_i|^3$, thus this event implies C and (28) follows.

Now suppose the opposite

$$2e^{-\frac{3}{28} V_t^*} > \frac{\epsilon}{\sqrt{V^*}}.$$

We will show that C always holds. The last inequality is equivalent to

$$-\frac{3V_t^*}{28} > \ln \frac{\epsilon}{2} - \frac{1}{2} \ln V^*;$$
$$V_t^* < \frac{28}{3} \ln \frac{2}{\epsilon} + \frac{14}{3} \ln V^*.$$

We can rewrite (26) as $\frac{28}{3} \ln \frac{2}{\epsilon} < \frac{V^*}{4}$. We can assume $V^* \ge 83$, see the end of the proof. Thus we have $\frac{14}{3} \ln V^* \le \frac{V^*}{4}$, so

$$\frac{28}{3} \ln \frac{2}{\epsilon} + \frac{14}{3} \ln V^* < \frac{V^*}{2}.$$

Combining the previous two inequalities we get $V_t^* < \frac{V^*}{2}$. Therefore by the assumption (15) of the lemma $t \le \lceil (1-c)n \rceil - 1 \le (1-c)n$. Also by (22)

$$V_A^* \ge \sum_{i=t+1}^{n} \mathbb{E}\left(Y_i^2 | A\right) \ge \frac{3}{8} \sum_{i=t+1}^{n} \mathbb{E}\, Y_i^2 = \frac{3}{8}(V^* - V_t^*) \ge \frac{3}{16} V^*.$$

As $V_{\lceil \alpha^{-1} \rceil - 1}^* \le V_{\lfloor \beta^{-1} \rfloor - 1}^*$ for $1 \ge \alpha \ge 0.5 \ge \beta > 0$, we get

$$V_{t,A}^* \le (1-c)V_A^*;$$
$$V_{t,A}^* \le \frac{1-c}{c}(V_A^* - V_{t,A}^*).$$

So using (22)

$$V_A^* \le \frac{8}{3} V^* \left(1 + \frac{1-c}{c}\right) = \frac{8V^*}{3c}.$$

Similarly using (25) and (23) we get

$$\mathbb{E}\left(|Y_1|^3 + \cdots + |Y_n|^3 | A\right) \le \frac{8}{c} \mathbb{E}\left(|Y_1|^3 + \cdots + |Y_n|^3\right),$$

which completes the proof of (28).

Now fix a set $A' \subseteq \{1, \ldots, n\}$ such that the event $A = A'$ implies C. Let $V_{A'}^* = Var(S_{A'})$. By (27) and (16)

$$\mathbb{I}_C \mathbb{E}\left(|Y_1|^3 + \cdots + |Y_n|^3 \mid A\right) \leq \delta' c' \mathbb{I}_C (V_A^*)^{3/2},$$

$c' = \frac{8}{c}\left(\frac{16}{3}\right)^{3/2}$. Thus similarly as above by the Berry–Esseen theorem for $z \in \mathbb{R}$

$$\left| \mathbb{P}\left(\frac{S_{A'}}{\sqrt{V_{A'}^*}} \leq z \right) - \Phi(z) \right| \leq \delta' c'.$$

Set $x = \frac{1}{2}$ if $|A'|$ is odd and $x = 0$ otherwise. Fix small ϵ_1 and ϵ_2 such that $0 \leq \epsilon_1 < \epsilon_2 < 1$ but

$$(\epsilon_2 - \epsilon_1)\sqrt{V^*} \geq 2. \tag{29}$$

Define $I_{\epsilon_1,\epsilon_2} = (x + \mathbb{Z}) \cap (\epsilon_1\sqrt{V^*}, \epsilon_2\sqrt{V^*}]$. Let $r_{A'} = \frac{V_{A'}^*}{V^*}$. We have

$$\mathbb{P}(S_{A'} \in I_{\epsilon_1,\epsilon_2}) = \Phi(r_{A'}^{-\frac{1}{2}}\epsilon_2) - \Phi(r_{A'}^{-\frac{1}{2}}\epsilon_1) \pm 2\delta' c'$$

Let $m = \min I_{\epsilon_1,\epsilon_2}$ Using a similar argument as above, since S_A' is unimodal on $x + \mathbb{Z}$, and $-(2e\pi)^{-1/2} \leq \Phi^{(2)}(z) \leq 0$ for $z \geq 0$, using Taylor's theorem twice there is $\xi \in [0, r_{A'}^{-\frac{1}{2}}\epsilon_2]$ such that

$$\mathbb{P}(S_{A'} = m) \geq \frac{\Phi(r_{A'}^{-\frac{1}{2}}\epsilon_2) - \Phi(r_{A'}^{-\frac{1}{2}}\epsilon_1) - 2\delta' c'}{|I_{\epsilon_1,\epsilon_2}|}$$

$$\geq \frac{\Phi(r_{A'}^{-\frac{1}{2}}\epsilon_2) - \Phi(r_{A'}^{-\frac{1}{2}}\epsilon_1) - 2\delta' c'}{(\epsilon_2 - \epsilon_1)\sqrt{V^*} + 1}$$

$$\geq \frac{\Phi'(0)r_{A'}^{-\frac{1}{2}}(\epsilon_2 - \epsilon_1) + 2^{-1}\Phi^{(2)}(\xi)r_{A'}^{-1}(\epsilon_2^2 - \epsilon_1^2) - 2\delta' c'}{(\epsilon_2 - \epsilon_1)\sqrt{V^*} + 1}$$

$$\geq \frac{1}{\sqrt{2\pi V^*}}\left(r_{A'}^{-\frac{1}{2}} - \frac{r_{A'}^{-1}(\epsilon_2^2 - \epsilon_1^2)}{2\sqrt{e}(\epsilon_2 - \epsilon_1)} - \frac{2\sqrt{2\pi}\delta' c'}{\epsilon_2 - \epsilon_1}\right)\left(1 - \frac{1}{(\epsilon_2 - \epsilon_1)\sqrt{V^*}}\right)$$

$$\mathbb{P}(S_{A'} = m) \geq \frac{1}{\sqrt{2\pi r_{A'} V^*}}\left(1 - \frac{\epsilon_1 + \epsilon_2}{2\sqrt{er_{A'}}} - \frac{2\sqrt{2\pi r_{A'}}\delta' c'}{(\epsilon_2 - \epsilon_1)} - \frac{1}{(\epsilon_2 - \epsilon_1)\sqrt{V^*}}\right). \tag{30}$$

Similarly for $M = \max I_{\epsilon_1,\epsilon_2}$, since

$$\frac{1 + z_1}{1 - z_2} \leq 1 + 2z_1 + 2z_2 \quad \text{for } z_2 \in \left(0, \frac{1}{2}\right) \text{ and } z_1 > 0, \tag{31}$$

using (29) we get

$$\mathbb{P}(S_{A'} = M) \leq \frac{1}{\sqrt{2\pi r_{A'} V^*}}\left(1 + \frac{4\sqrt{2\pi r_{A'}}\delta' c'}{(\epsilon_2 - \epsilon_1)} + \frac{2}{(\epsilon_2 - \epsilon_1)\sqrt{V^*}}\right). \tag{32}$$

21

Fix $\tilde{\epsilon} > 2(V^*)^{-1/2}$. Denote $\tau = \lfloor \tilde{\epsilon}\sqrt{V^*} \rfloor$. Applying (30) with $\epsilon_1 = 2\tilde{\epsilon}$ and $\epsilon_2 = 3\tilde{\epsilon}$, then using (27) we get

$$\mathbb{P}(S_{A'} = x + 2\tau) \geq \frac{1}{\sqrt{2\pi r_{A'} V^*}}(1 - h_1);$$

$$h_1 = \frac{10\tilde{\epsilon}}{\sqrt{3e}} + 8\sqrt{\frac{\pi}{3c}}\delta' c' \tilde{\epsilon}^{-1} + (\tilde{\epsilon}\sqrt{V^*})^{-1}.$$

Similarly applying (32) with $\epsilon_1 = 0$ and $\epsilon_2 = \tilde{\epsilon}$, then using (27) we get

$$\mathbb{P}(S_{A'} = x + \tau) \leq \frac{1}{\sqrt{2\pi r_{A'} V^*}}(1 + h_2);$$

$$h_2 = 16\sqrt{\frac{\pi}{3c}}\delta' c' \tilde{\epsilon}^{-1} + 2(\tilde{\epsilon}\sqrt{V^*})^{-1}.$$

Now since $S_{A'}$ is unimodal and log-concave a 'telescoping' argument implies that

$$\mathbb{P}(S_{A'} = x)\mathbb{P}(S_{A'} = x + 2\tau) \leq \mathbb{P}(S_{A'} = x + \tau)^2. \tag{33}$$

So

$$\mathbb{P}(S_{A'} = x)(1 - h_1) \leq \mathbb{P}(S_{A'} = x + \tau)(1 + h_2).$$

For all A' (even or odd size) this can be written as

$$\mathbb{P}(S_{A'} \in \{0, \tfrac{1}{2}\})(1 - h_1) \leq \mathbb{P}(S_{A'} \in \{\tau, \tau + \tfrac{1}{2}\})(1 + h_2).$$

Averaging over all A' such that C holds on $A = A'$ this becomes

$$\mathbb{P}(S \in \{0, \tfrac{1}{2}\} \cap C)(1 - h_1) \leq \mathbb{P}(S \in \{\tau, \tau + \tfrac{1}{2}\} \cap C)(1 + h_2).$$

Assume

$$\mathbb{P}(\bar{C}) + h_1 + h_2 < \frac{1}{4}. \tag{34}$$

For any event A, $\mathbb{P}(A) \leq \mathbb{P}(A \cap C) + \mathbb{P}(\bar{C})$ so using (31) and (34)

$$\mathbb{P}(S \in \{0, \tfrac{1}{2}\}) \leq \mathbb{P}(S \in \{\tau, \tau + \tfrac{1}{2}\})(1 + 2h_1 + 2h_2) + \mathbb{P}(\bar{C}). \tag{35}$$

By a simple calculation:

$$2h_1 + 2h_2 = \frac{20}{\sqrt{3e}}\tilde{\epsilon} + 48\sqrt{\frac{\pi}{3c}}\frac{8}{c}\left(\frac{16}{3}\right)^{\frac{3}{2}}\frac{\delta'}{\tilde{\epsilon}} + \frac{6}{\tilde{\epsilon}\sqrt{V^*}} \leq 8\tilde{\epsilon} + \frac{4840\delta'}{c^{\frac{3}{2}}\tilde{\epsilon}} + \frac{6}{\tilde{\epsilon}\sqrt{V^*}}.$$

Similarly as in (21) and (32) we get (since $\tilde{\epsilon}\sqrt{V^*} > 2$)

$$\mathbb{P}(S \in \{\tau, \tau + \tfrac{1}{2}\}) \leq \frac{1}{\sqrt{2\pi V^*}}\left(1 + 4\delta'\tilde{\epsilon}^{-1}\sqrt{2\pi} + \frac{2}{\tilde{\epsilon}\sqrt{V^*}}\right) \tag{36}$$

22

Combining (35) and (36) and using (21), (28) and (34) yields

$$\mathbb{P}(S \in \{0, \frac{1}{2}\})$$
$$\leq \frac{1}{\sqrt{2\pi V^*}} \left(1 + 4\delta'\tilde{\epsilon}^{-1}\sqrt{2\pi} + \frac{2}{\tilde{\epsilon}\sqrt{V^*}}\right)(1 + 2h_1 + 2h_2) + \mathbb{P}(\bar{C})$$
$$\leq \frac{1}{\sqrt{2\pi V^*}}(1 + 2h_1 + 2h_2 + 6\delta'\tilde{\epsilon}^{-1}\sqrt{2\pi} + \frac{3}{\tilde{\epsilon}\sqrt{V^*}}) + \frac{\epsilon}{\sqrt{V^*}}.$$
$$\leq \frac{1}{\sqrt{2\pi V^*}}(1 + h);$$
$$h = 8\tilde{\epsilon} + \frac{4840\delta'}{c^{\frac{3}{2}}\tilde{\epsilon}} + \frac{9}{\tilde{\epsilon}\sqrt{V^*}} + 6\delta'\tilde{\epsilon}^{-1}\sqrt{2\pi} + \sqrt{2\pi}\epsilon.$$

We have

$$h \leq \frac{1}{\sqrt{V^*}}\frac{9}{\tilde{\epsilon}} + 8\tilde{\epsilon} + \frac{4856\delta'}{c^{\frac{3}{2}}\tilde{\epsilon}} + \sqrt{2\pi}\epsilon.$$

Put $\epsilon = \frac{2}{\sqrt{V^*}}$. Since $V^* \geq 83$, (26) is satisfied. The minimum x_0 of $f(x) = \frac{a}{x} + bx$ with $a, b > 0$ is $x_0 = \sqrt{\frac{a}{b}}$ with $f(x_0) = 2\sqrt{ab}$. Hence if we put $\tilde{\epsilon} = 8^{-1/2}(\frac{9}{\sqrt{V^*}} + \frac{4856\delta'}{c^{3/2}})^{1/2}$, we get

$$h \leq 2 \times \sqrt{8\left(\frac{9}{\sqrt{V^*}} + 4856\delta'c^{-\frac{3}{2}}\right)} + \frac{2\sqrt{2\pi}}{\sqrt{V^*}}$$
$$\leq 2 \times \sqrt{8\left(18\delta' + 4856\delta'c^{-\frac{3}{2}}\right)} + 4\sqrt{2\pi}\delta'$$
$$\leq 2 \times \sqrt{8 \cdot 4874\delta'c^{-\frac{3}{2}}} + 4\sqrt{2\pi}(\delta')^{\frac{1}{2}}$$
$$\leq 405(\delta')^{\frac{1}{2}}c^{-\frac{3}{4}} = \epsilon'. \tag{37}$$

Here in the second line we used

$$\delta' \geq \frac{1}{2\sqrt{V^*}}, \tag{38}$$

which follows since by (16) and the definition of ν_α^*

$$\frac{1}{2}V^* = \frac{1}{2}\sum_{i=1}^{n}\mathbb{E}\,Y_i^2 \leq \sum_{i=1}^{n}\mathbb{E}\,|Y_i|^3 \leq \delta'(V^*)^{3/2}.$$

From (21) $\mathbb{P}(S \in \{0, \frac{1}{2}\}) \geq \frac{1}{\sqrt{2\pi V^*}}(1 - h)$. Finally note that all of our used assumptions, i.e. (34), $V^* \geq 83$, $\tilde{\epsilon} > 2(V^*)^{-1/2}$ and $\mathbb{P}(\bar{C}) \leq \frac{1}{2}$ follow from (17), (37) and (38). $\square$

Proof of Theorem 10.3 Let $\underline{c}$, $\bar{c}$ be positive numbers, guaranteed by the norm equivalence for $\mathbb{R}^d$, such that $\underline{c}\|z\| \leq \|z\|_2 \leq \bar{c}\|z\|$ for all $z \in \mathbb{R}^d$. Set $c_\mathcal{S} = \frac{\sqrt{3}}{4}$ if $\mathcal{S} = (\mathbb{R}^d, \|\cdot\|)$ is Hilbert and $c_\mathcal{S} = \frac{1}{8}$ otherwise. Let $\overline{C}$ be the minimum number

of open sets of l_2-diameter one that can cover any open set of l_2-diameter $\bar{c}$. As such a set is contained in an open l_2-ball of diameter $2\bar{c}$, the number $\overline{C}$ is finite (although exponential in d). For the Euclidean space we have $\underline{c} = \bar{c} = \overline{C} = 1$ and $c_{\mathcal{S}} = \frac{\sqrt{3}}{4}$.

Let $\tilde{\mathcal{Q}}_2$ be the concentration w.r.t. $\|\cdot\|_2$. Then $\tilde{\mathcal{Q}}_2(X_1 + \cdots + X_n, 1) \geq \overline{C}^{-1}\tilde{\mathcal{Q}}(X_1 + \cdots + X_n, 1)$.

Let $c_1 = 9c_H(d-1)^3$ be the constant from Lemma 9.2. We will prove the theorem with

$$C = \sqrt{2c_1\overline{C}}\underline{c}^{-1}c_{\mathcal{S}}^{-1}. \tag{39}$$

Recall that $d(x,y) = \|x-y\|$ and $d_2(x,y) = \|x-y\|_2$. Rescale the space by multiplying by 2, apply Lemma 9.2 with $c = 2\underline{c}c_{\mathcal{S}} = 2\sqrt{2c_1\overline{C}}C^{-1}$, $\delta = \overline{C}^{-1}t(\alpha_1,\ldots,\alpha_n)$ and

$$\epsilon = \frac{m}{4n} = \frac{C}{4}\xi_d(\bar{\alpha})^{1/2}t(\alpha_1,\ldots,\alpha_n)^{-1/2}n^{-1/2},$$

then rescale back by dividing by 2. We get that either $\tilde{\mathcal{Q}}(X_1 + \cdots + X_n, 1) < t(\alpha_1,\ldots,\alpha_n)$ or there is a line L and vectors $a_1,\ldots,a_n$ in $\mathbb{R}^d$ such that for all but at most b,

$$b := \frac{m}{2} = \frac{C}{2}\xi_d(\bar{\alpha})^{1/2}t(\alpha_1,\ldots,\alpha_n)^{-1/2}n^{1/2},$$

indices in $i \in \{1,\ldots,n\}$ we have $\mathbb{P}(d(X_i - a_i, L) \geq c_{\mathcal{S}}) \leq \mathbb{P}(d_2(X_i - a_i, L) \geq \underline{c}c_{\mathcal{S}}) \leq \epsilon$. In the former case we are done by Lemma 10.1, so let us assume the latter one. We may assume that $a_i = 0$ for all i. Let $I = \{i \in \{1,\ldots,n\} : \mathbb{P}(d(X_i, L) \geq c_{\mathcal{S}}) \leq \epsilon\}$, so $|I| \geq n - \lfloor b \rfloor$. Let $\mathbb{I}_i$ be the indicator of the event $d(X_i, L) \geq c_{\mathcal{S}}$. It is a Bernoulli random variable with parameter p_i and if $i \in I$ then $p_i \leq \epsilon$. Note that (20) and $c < \frac{1}{3}$ implies $\epsilon \leq \frac{1}{60}$.

Define a random variable $B = \sum_{i \in I} \mathbb{I}_i$. We have

$$\mathbb{P}(B \geq 2\epsilon n) \leq \mathbb{P}(B - \mathbb{E}\,B \geq \epsilon n) \leq \mathbb{P}\left(B - \mathbb{E}\,B \geq \epsilon|I|\right) \leq e^{-2\epsilon^2|I|} \leq e^{-\frac{29}{15}\epsilon^2 n}.$$

Here the fourth line is a Hoeffding's concentration inequality, see, e.g. [18] and the last line follows since by (20) $|I| \geq n - \frac{cn}{2\cdot 5} \geq \frac{29n}{30}$.

For $i \in \{1,\ldots,n\}$ define $\alpha_i' = \min(1, (1 + \frac{60}{59}\epsilon)\alpha_i)$. For $i \in I$ let X_i' be distributed as X_i conditioned on $\mathbb{I}_i = 0$. Since $\epsilon \leq \frac{1}{60}$ we have

$$\tilde{\mathcal{Q}}(X_i', 1) \leq \min\left(1, \frac{\alpha_i}{1-\epsilon}\right) = \min\left(1, \alpha_i\left(1 + \frac{\epsilon}{1-\epsilon}\right)\right) \leq \alpha_i'.$$

Let $Y_1',\ldots,Y_n'$ be independent random variables, $Y_i' \sim \nu_{\alpha_i'}^*$. Applying the exact result of Juškevičius [12] for dimension 1 for any set of indices $1 \leq i_1 < \cdots < i_k \leq n$ we get $t(\alpha_{i_1}',\ldots,\alpha_{i_k}') \leq t(\alpha_1',\ldots,\alpha_k')$.

So by one of our main results, Theorem 2.2, we have

$$\tilde{\mathcal{Q}}(X_1 + \cdots + X_n, 1) \leq \tilde{\mathcal{Q}}(\sum_{i \in I} X_i', 1)$$

$$\leq \max_{I', I' \subseteq I, |I'| \geq |I| - 2\epsilon n} \tilde{\mathcal{Q}}(\sum_{i \in I'} X_i', 1) + \mathbb{P}(B \geq 2\epsilon n)$$

$$\leq \max_{I', I' \subseteq I, |I'| \geq |I| - 2\epsilon n} \tilde{\mathcal{Q}}(\sum_{i \in I'} Y_i', 1) + \mathbb{P}(B \geq 2\epsilon n)$$

$$\leq \max_{I', I' \subseteq \{1, \ldots, n\}, |I'| \geq n - b - 2\epsilon n} \tilde{\mathcal{Q}}(\sum_{i \in I'} Y_i', 1) + \mathbb{P}(B \geq 2\epsilon n)$$

$$\leq t(\alpha_1', \ldots, \alpha_{\bar{n}}') + e^{-\frac{29}{15}\epsilon^2 n} \tag{40}$$

where

$$\bar{n} := n - \lfloor b + 2\epsilon n \rfloor = n - \lfloor m \rfloor.$$

Let $\alpha \in (\frac{1}{k+1}, \frac{1}{k})$. Then from (14) $-\frac{\partial V_\alpha^*}{\partial \alpha} = \frac{1}{12}k(k+1)(2k+1)$. For $k = 1$ the right side is equal to $\frac{1}{2}$, while for $k \geq 2$ it is at most $6^{-1}(k^2 - 1)(2k+1) \leq 2V_\alpha^*(2\alpha^{-1} + 1)$. Therefore, since V_α is convex on $\alpha \in (0, 1]$, for any such α and $h > 0$, if $\alpha(1 + h) \leq 1$ then

$$V_{\alpha(1+h)}^* \geq V_\alpha^* - \alpha h \max(2V_\alpha^*(2\alpha^{-1} + 1), \frac{1}{2}) \geq V_\alpha^*(1 - 6h) - \frac{\alpha h}{2}.$$

So for $t \in \{1, \ldots, n\}$

$$V_t' := \sum_{i=1}^{t} V_{\alpha_i'}^* \geq V_t^*(1 - 7\epsilon) - \frac{30}{59} t\bar{\alpha}\epsilon. \tag{41}$$

By Lemma 10.1 and the conditions of the theorem

$$t(\alpha_1, \ldots, \alpha_n) = \frac{1 \pm \epsilon'}{\sqrt{2\pi V^*}}.$$

As $\epsilon' < \frac{1}{2}$ we have $t(\alpha_1, \ldots, \alpha_n) \geq \frac{1}{2\sqrt{2\pi V^*}}$ and by the definition of ϵ, (18) and (19) for any $t \geq (1 - c)n$

$$\frac{30 t\bar{\alpha}\epsilon}{59 V_t^*} \leq \frac{30 \cdot 4}{59 \cdot 3} \frac{\bar{\alpha}\epsilon n}{V^*} \leq \frac{10C(8\pi)^{1/4}\sqrt{\gamma}}{59} \leq \frac{1}{8}.$$

Therefore we have

$$\epsilon_V := 7\epsilon + \frac{40}{59} \frac{\bar{\alpha}\epsilon n}{V^*} \leq 7\epsilon + 0.38C\sqrt{\gamma} \leq \frac{1}{8} + \frac{1}{8} = \frac{1}{4} \tag{42}$$

and by (41) for any $t \geq (1 - c)n$

$$V_t' \geq V_t^*(1 - \epsilon_V). \tag{43}$$

25

We shall now check that Lemma 10.1 can be applied to the sequence $Y'_1, \ldots, Y'_{\bar{n}}$. Since (18) holds for $Y_1, \ldots, Y_n$ and by (20)

$$b + 2\epsilon n = C\xi_d(\bar{\alpha})^{\frac{1}{2}} t(\alpha_1, \ldots, \alpha_n)^{-\frac{1}{2}} n^{\frac{1}{2}} \leq \frac{cn}{5} < \frac{cn}{2}, \tag{44}$$

using (18) and (43) we get that (15) also holds for $Y'_1, \ldots, Y'_{\bar{n}}$ with the constant $\frac{c}{2}$ instead of c, i.e.:

$$V'_{\lceil(1-c/2)\bar{n}\rceil} \geq (1 - \epsilon_V)V^*_{\lceil(1-c/2)\bar{n}\rceil} \geq (1 - \epsilon_V)\frac{3}{4}V^* \geq \frac{9}{16}V^*_{\bar{n}} \geq \frac{9}{16}V'_{\bar{n}} > \frac{1}{2}V'_{\bar{n}}.$$

The absolute moments of ν^*_α decrease as α increase. So since (16) holds for $Y_1, \ldots, Y_n$ we have that (16) holds also for $Y'_1, \ldots Y'_{\bar{n}}$ with δ'' instead of δ' where $\delta'' = (4/3)^{3/2}\delta'(1 - \epsilon_V)^{-3/2} \leq (4/3)^3\delta'$. Here we again used (18) and (43).

We shall apply Lemma 10.1 to the sequence $Y'_1, \ldots, Y'_{\bar{n}}$ with parameters

$$(c/2, (4/3)^3\delta', 8\epsilon'/3)$$

instead of (c, δ', ϵ').

It remains to check that (17) holds for this shorter sequence. By (18) and (44) we have

$$V^*_{\bar{n}} \geq V'_{\bar{n}} \geq (1 - \epsilon_V)V^*_{\bar{n}} \geq \frac{3}{4}(1 - \epsilon_V)V^* \geq \frac{9}{16}V^*.$$

Thus replacing δ' with $(4/3)^3\delta'$ and replacing c with $\frac{c}{2}$ increases the left side of (17) by $2^{3/4}(4/3)^{3/2} \leq \frac{8}{3}$, so (17) holds for $Y'_1, \ldots, Y'_{\bar{n}}$ with the claimed parameters by our stronger requirement that $\epsilon' \leq \frac{3}{16}$.

Thus Lemma 10.1 yields:

$$\begin{aligned}
t(\alpha'_1, \ldots, \alpha'_{\bar{n}}) &\leq \frac{1 + \frac{8}{3}\epsilon'}{\sqrt{2\pi V^*_{\bar{n}}(1 - \epsilon_V)}} \\
&\leq \frac{1}{\sqrt{2\pi V^*_{\bar{n}}}}(1 + \frac{2 \cdot 8}{3}\epsilon' + 2\epsilon_V) \\
&\leq \frac{1}{\sqrt{2\pi V^*_{\bar{n}}}}(1 + 6\epsilon' + 14\epsilon + C\sqrt{\gamma}).
\end{aligned}$$

Here in the second line we used $\frac{8}{3}\epsilon' \leq \frac{1}{2}$ and $\epsilon_V \leq \frac{1}{4}$ and in the last line we used (42). The proof is completed by combining the last bound with (40) and using $\epsilon = \frac{m}{4n}$. $\qquad \square$

Now our first two theorems follow easily as simple special cases.

Proof of Theorem 1.1 and Theorem 2.1 It is enough to prove the more general Theorem 2.1. We may assume $\alpha < 1$. Let us apply Theorem 10.3 to each $n \in \{1, 2, \ldots\}$ with $\alpha_1 = \cdots = \alpha_n = \alpha$.

For the i.i.d. summands $Y_1, \ldots, Y_n \sim \nu^*_\alpha$ we have $V^*_{\lfloor(1-c)n\rfloor} \geq (1 - c)V^*_n$ for any $c \in (0, 1)$. So we can take, for example $c = \frac{1}{4}$ so that (18) is satisfied for

any n. Let the constant $C = C(d, \| \cdot \|)$ be as in (39). Since α is fixed, (16) holds for any positive δ' with n sufficiently large. Since $V_n^* \to \infty$ as $n \to \infty$, we can fix a $\delta' \in (0, 1)$ such that for all n large enough, ϵ' defined in (17) satisfies $\epsilon' \leq 0.05$. We can actually make ϵ' arbitrarily small for all large enough n if we choose δ' small enough.

As $\bar{\alpha}$ ($\bar{\alpha} = \alpha$) is constant and V_n^* is linear in n, we can choose an arbitrarily small γ such that (19) is satisfied for all n large enough.

Consequently, it follows by Lemma 10.1 (or a standard local limit theorem) that $t(\alpha_1, \dots, \alpha_n) = \tilde{\mathcal{Q}}(Y_1 + \cdots + Y_n, 1) = (2\pi n V_\alpha)^{-\frac{1}{2}}(1 + o(1))$, so m is of the order $n^{\frac{3}{4}}$. Therefore (20) is satisfied for all n large enough. This also implies that for the error term we have $e^{-m^2/(9n)} = e^{-\Omega(n^{\frac{1}{2}})} = o(t(\alpha_1, \dots, \alpha_n))$.

As $mn^{-1} = o(1)$, $V_{n-\lfloor m \rfloor}^* = V_n^*(1 + o(1))$ and Theorem 10.3 holds with arbitrarily small ϵ' and γ, for all n large enough, we get

$$\tilde{\mathcal{Q}}(X_1 + \cdots + X_n, 1) \leq \frac{1}{\sqrt{2\pi V_n^*}}(1 + o(1))$$

$$= \mathbb{P}(Y_1 + \cdots + Y_n \in \{0, \frac{1}{2}\})(1 + o(1))$$

$$= \tilde{\mathcal{Q}}(Y_1 + \cdots + Y_n, 1)(1 + o(1)).$$

In the setting of Theorem 1.1 just note that the value of the concentration function is not affected if we shift each Y_i by $\frac{k+1}{2}$ to make it distributed as U_i.

$\square$

Finally, Theorem 2.3 follows as an asymptotic version of Theorem 10.3.

Proof of Theorem 2.3 By the assumptions of the theorem, for each $c \in (0, \frac{1}{3})$ we can choose sequences of positive numbers $\{\delta_l'\}$, $\{\gamma_l\}$ $\delta_l' \to 0$, $\gamma_l \to 0$ so that for $l \in \{1, 2, \dots\}$ Theorem 10.3 applies with the absolute constant $C = C(d, \| \cdot \|)$ defined in (39), $\delta' = \delta_l'$ and $\gamma = \gamma_l'$. With this choice the left side of (17) converges to 0 as $l \to \infty$, so we get that $\tilde{\mathcal{Q}}(X_{l,1} + \cdots + X_{l,n}, 1) \leq (2\pi V_{l, \lceil n_l(1-c) \rceil}^*)^{-1/2}(1 + o(1)) + r_l$, where $r_l = \exp(-C^2 \xi_d(\bar{\alpha}_l) t_l^{-1}/9)$ and $t_l = t(\alpha_{l,1}, \dots, \alpha_{l,n_l})$. Assuming that

$$r_l = o((2\pi V_{l,n_l}^*)^{-1/2}) \tag{45}$$

as $l \to \infty$, the proof of the upper bound is completed by considering c arbitrarily small, using the last assumption of the theorem and using Remark 10.2.

By Lemma 10.1 we have that

$$t_l = (2\pi V_{l,n_l}^*)^{-1/2}(1 + o(1)) \to 0. \tag{46}$$

This implies the lower bound in (5).

It remains to show (45). First consider the case $\xi_d(\bar{\alpha}_l) = \Omega(1)$. Then by (46) $r_l = \exp(-\Omega(t_l^{-1})) = o(t_l^{-1})$ and this implies (45).

Now consider the case $\bar{\alpha}_l \to 0$. We can assume $\bar{\alpha}_l \leq \frac{1}{4}$. Using (46) we can see that (45) holds if, for example,

$$\frac{C^2 \bar{\alpha}_l}{9} \geq \frac{2 \ln\left((2\pi V^*_{l,n_l})^{1/2}\right)}{(2\pi V^*_{l,n_l})^{1/2}}$$

which for large enough l follows from

$$\frac{C^2 \bar{\alpha}_l}{9} \geq \frac{2 \ln V^*_{l,n_l}}{(2\pi V^*_{l,n_l})^{1/2}}.$$

Remark 10.2 yields for $\bar{\alpha}_l \leq \frac{1}{2}$

$$\bar{\alpha}_l \geq \left(\frac{3}{4} \cdot \frac{\pi n_l}{6}\right)^{\frac{1}{2}} (2\pi V^*_{l,n_l})^{-\frac{1}{2}} \geq \frac{1}{2} n_l^{\frac{1}{2}} (2\pi V^*_{l,n_l})^{-\frac{1}{2}}$$

Thus a sufficient condition for (45) is

$$\frac{C^2 n_l^{\frac{1}{2}}}{9 \cdot 2(2\pi V^*_{l,n_l})^{1/2}} \geq \frac{2 \ln V^*_{l,n_l}}{(2\pi V^*_{l,n_l})^{1/2}},$$

$$\frac{C^2 n_l^{\frac{1}{2}}}{18} \geq 2 \ln V^*_{l,n_l} \quad \text{or} \quad V^*_{l,n_l} \leq \exp(\frac{1}{36} C^2 n_l^{\frac{1}{2}}).$$

$\square$

Proof of Remark 2.4 The proof follows by Theorem 2.3, the fact that $V_{\bar{\alpha}_l} = \frac{1}{12\bar{\alpha}_l^2}(1 + o(1))$ when $\bar{\alpha}_l \to 0$ and $V_{\bar{\alpha}_l} = \frac{1}{2}(1 - \bar{\alpha}_l)$ when $\bar{\alpha}_l > 0.5$. $\square$

Let us end this section with a comparison of our results to a known result. For the case $d = 1$, Kesten [16] proved a concentration bound of the correct order.

Theorem 10.4 *(Theorem 2 of [16]) There is an absolute constant $C > 0$ such that for independent random variables $X_1, \ldots, X_n$ and any $0 < \lambda_1, \ldots, \lambda_n \leq 2L$ if $S_n = X_1 + \cdots + X_n$ then*

$$\mathcal{Q}(S_n, L) \leq 4\sqrt{2}(1 + 9C)L \frac{\sum_{i=1}^n \lambda_i^2 (1 - \mathcal{Q}(X_i, \lambda_i))\mathcal{Q}(X_i, L)}{\left(\sum_{i=1}^n \lambda_i^2 (1 - \mathcal{Q}(X_i, \lambda_i))\right)^{\frac{3}{2}}}$$

$$\leq 4\sqrt{2}(1 + 9C)L \frac{\max_i \mathcal{Q}(X_i, L)}{\sqrt{\sum_{i=1}^n \lambda_i^2 (1 - \mathcal{Q}(X_i, \lambda_i))}}.$$

Applying it with our conditions $L = \lambda_1 = \cdots = \lambda_n = 1$ and, for example, with $\mathcal{Q}(X_i, 1) = \bar{\alpha} = \bar{\alpha}(n)$ for all i we get

$$\mathcal{Q}(X_1 + \cdots + X_n, 1) \leq 4\sqrt{2}(1 + 9C)\frac{\bar{\alpha}}{\sqrt{(1 - \bar{\alpha})n}}.$$

28

For $\bar{\alpha} \geq \frac{1}{2}$ and $1 - \bar{\alpha} \gg n^{-1/3}$, Theorem 2.3 and (14) show that $\mathcal{Q}(X_1 + \cdots + X_n, 1) \leq \frac{1}{\sqrt{\pi(1-\bar{\alpha})n}}(1+o(1))$. Depending on $\bar{\alpha}$, Kesten's constant is at least 5–10 times larger. For $\bar{\alpha} \in (0, \frac{1}{2})$, $\bar{\alpha} \gg n^{-1/2}$ the constant in our cruder bound (6) is at least 4–5 times smaller than Kesten's. For the case $d = 1$, using Remark 10.2, the precise asymptotics also follow from Theorem 1.2 of Juškevičius [12] without the restrictions on $\bar{\alpha}$ – many of them in Theorem 2.3 seem to be artefacts of our method and are probably not necessary.

References

[1] P. Billingsley, *Convergence of Probability Measures,* 2nd edition, John Wiley & Sons, New York, 1999.

[2] S.G. Bobkov and G. P. Chistyakov, On concentration functions of random variables, *Journal of Theoretical Probability* **28** (2015), 976–988.

[3] N. G. de Bruijn, C. van Ebbenhorst Tengbergen and D. Kruyswijk, On the set of divisors of a number, *Nieuw Arch. Wiskunde* **23** (1951) 191–193.

[4] M. Chudnovsky, N. Robertson, P. Seymour and R. Thomas, The strong perfect graph theorem, *Annals of Mathematics* **164** (2006), 51–229.

[5] R. Durrett, *Probability: Theory and Examples*, 5th edition, 2019, `https://services.math.duke.edu/~rtd/PTE/pte.html`.

[6] W. F. Eddy, and J. A. Hartigan, Uniform convergence of the empirical distribution function over convex sets, *The Annals of Statistics* **5** (1977), 370–374.

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

[8] B. V. Gnedenko, *Theory of Probability*, 4th edition, MIR Publishing, Moscow, 1978.

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

[10] L. Jones, On the distribution of sums of vectors, *SIAM Journal on Applied Mathematics* **34** (1978), 1–6.

[11] T. Juškevičius, Probabilistic inequalities and bootstrap percolation, PhD dissertation, University of Memphis, 2015.

[12] T. Juškevičius, The sharp form of the Kolmogorov-Rogozin inequality and a conjecture of Leader–Radcliffe, *arXiv preprint* arXiv:2201.09861 (2022).

[13] J. Keilson and H. Gerber, Some results for discrete unimodality, *Journal of the American Statistical Association* **66** (1971), 386–389.

[14] D. Kleitman, On a lemma of Littlewood and Offord on the distributions of linear combinations of vectors, *Advances in Mathematics* **5** (1970), 427–429.

[15] P. S. Lévy, Théorie de l'addition des variables aléatoires, *The Mathematical Gazette* **39** (1955), 344.

[16] H. Kesten, A sharper form of the Doeblin–Lévy–Kolmogorov–Rogozin inequality for concentration functions, *Mathematica Scandinavica* **25** (1969), 133–144.

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

[18] 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.

[19] J. R. Munkres, Topology, 2nd edition, Featured Titles for Topology Series, Prentice Hall Inc., Upper Saddle River, NJ, 2008.

[20] I. G. Shevtsova, An improvement of convergence rate estimates in the Lyapunov theorem, *Doklady Mathematics* **82** (2010), 862–864.

[21] V. S. Varadarajan, On the convergence of sample probability distributions, *Sankhyā* **19** (1958) 23–26.

INSTITUTE OF COMPUTER SCIENCE OF THE CZECH ACADEMY OF SCIENCES, POD VODÁRENSKOU VĚŽÍ 271/2 182 00 PRAHA 8, CZECHIA
Email address: tomas.juskevicius@gmail.com

FACULTY OF MATHEMATICS AND INFORMATICS, VILNIUS UNIVERSITY, AKADEMIJOS 4, 08412, VILNIUS, LITHUANIA
Email address: valentas@gmail.com