

Sharp Khinchin-type inequalities for symmetric discrete uniform random variables

Alex Havrilla* and Tomasz Tkocz†

31th December 2019

Abstract

We establish several optimal moment comparison inequalities (Khinchin-type inequalities) for weighted sums of independent identically distributed symmetric discrete random variables which are uniform on sets of consecutive integers. Specifically, we obtain sharp constants for even moments (using ultra subgaussianity introduced by Nayar and Oleszkiewicz) as well as for the second moment and any moment of order at least 3 (using convex dominance by Gaussian random variables). In the case of only 3 atoms, we also establish a Schur-convexity result. For moments of order less than 2, we get sharp constants in two cases by exploiting Haagerup's arguments for random signs.

2010 Mathematics Subject Classification. Primary 60E15; Secondary 26D15.

Key words. Khinchin inequality, moment comparison, ultra sub-Gaussianity, convex ordering, majorisation, Schur convexity

1 Introduction

The classical Khinchin inequality asserts that all moments of weighted sums of independent random signs are comparable (see [15]). More specifically, if we consider independent random signs $\varepsilon_1, \varepsilon_2, \dots$, the probability of each ε_i taking the value ± 1 is a half and form a weighted sum $S = \sum_{i=1}^n a_i \varepsilon_i$ with real coefficients a_i , then for every $p, q > 0$, there is a positive constant $C_{p,q}$ independent of n and the a_i such that

$$\|S\|_p \leq C_{p,q} \|S\|_q. \quad (1)$$

As usual, $\|X\|_p = (\mathbb{E}|X|^p)^{1/p}$ denotes the p -th moment of a random variable X . Moment comparison inequalities like this one are well understood up to universal constants in a great generality due to Latała's formula from [20]. They have found numerous applications in classical results in analysis (for example in the proof of the Littlewood-Paley decomposition or Grothendieck's inequality) and, especially their extensions to vector valued settings (Kahane's inequalities), have been widely used in (local) theory of Banach spaces (see [23], [26]). One of the major challenges is to find the best constants $C_{p,q}$, which has attracted considerable

*Carnegie Mellon University; Pittsburgh, PA 15213, USA. Email: alumhavr@andrew.cmu.edu

†Carnegie Mellon University; Pittsburgh, PA 15213, USA. Email: ttkocz@math.cmu.edu. Research supported in part by the Collaboration Grants from the Simons Foundation.

attention and has important applications (for instance in geometry, $C_{2,1}$ is directly linked with the maximum volume projections of the n -dimensional cross-polytope onto $n - 1$ dimensional subspaces, see [3, 5]). Besides, attacking sharp inequalities forces us to uncover often deep and effective mechanisms explaining *bigger pictures* and providing insights as to why certain inequalities are true.

Plainly, since for any random variable X , the function $p \mapsto \|X\|_p$ is nondecreasing, the best value of $C_{p,q}$ in (1) when $p < q$ equals 1. Since $\|S\|_2$ is explicit in terms of the weights a_i , that is $\|S\|_2 = \sqrt{\sum a_i^2}$, the most important are $C_{p,2}$ when $p > 2$ and $C_{2,q}$ when $q < 2$. In the case being discussed of symmetric random signs, the values of these constants have been known since the work of Haagerup [13]. We mention in passing works [39, 9, 37] which had made important partial contributions preceding Haagerup's result. Papers [27, 29] provide great simplifications and deeper understanding of technical parts in Haagerup's proofs. Paper [21] establishes in a slick way that $C_{2,1} = \sqrt{2}$ in a general setting of norm space-valued coefficients a_i (for recent results concerning this setting see also [33]). We refer to [28] for historical accounts and beautiful recent results for even moments. The constants $C_{p,2}$, $p > 2$ are attained in the asymptotic case when the number of summands n tends to infinity with weights a_i being chosen all equal. Consequently, by the central limit theorem, the value of $C_{p,2}$ is given by the p -th moment of a standard Gaussian. This phenomenon is in some sense universal – for distributions other than random signs where such results are known, the same case is extremal. The behaviour of the optimal value of $C_{2,q}$, $q < 2$, is more involved: as q decreases, the worst case changes at $q = q_0 = 1.847..$ from the asymptotic one just described to the one given by $n = 2$ and equal weights $a_1 = a_2$ (see [13]).

There have been only a handful of results concerning random variables other than random signs. They involve continuous random variables uniformly distributed on symmetric intervals and generalisations for random vectors uniformly distributed on Euclidean spheres and balls (see [2, 17, 18, 22]), as well as mixtures of centred Gaussians (see [1, 10]). Papers [19, 32] establish moment comparison inequalities for quite general random variables (based on their *spectral properties*, introducing differential inequalities techniques), which additionally yield sharp constants in certain cases. In recent works [10, 11], Eskenazis, Nayar and the second author have settled most of the cases for random variables with densities proportional to $e^{-|x|^\alpha}$ when $0 < \alpha < \infty$ (the so-called *exponential family*). This, combined with results from [4], yields sharp constants in Khinchin inequalities for linear forms based on vectors uniformly distributed on unit balls $B_\alpha^n = \{x \in \mathbb{R}^n, |x_1|^\alpha + \dots + |x_n|^\alpha \leq 1\}$ of ℓ_α spaces (previously these constants were known up to constant factors – see [4]). This is particularly interesting because the summands of such linear forms are not independent. For results concerning dependent random signs, see [34, 36] (moment comparison is obtained with constants of the right order, but their optimal values in most cases do not seem to be known).

This paper initiates the study of Khinchin-type inequalities with sharp constant for symmetric discrete random variables, generalising random signs by allowing more than just two atoms. Specifically, in the simplest case, let L be a positive integer and let X be uniform on the set $\{-L, \dots, -1\} \cup \{1, \dots, L\}$. What are best constants in moment comparison inequalities for weighted sums of independent copies of X ? Note that the following two *extreme* cases

have been understood: when $L = 1$, X is a symmetric random sign discussed above, whereas when $L \rightarrow \infty$, X/L converges in distribution to a random variable uniform on $[-1, 1]$, the case analysed in [22].

We present our results in the next section and then proceed with their proofs in their order of statement. We say that a random variable X is *symmetric* if $-X$ has the same distribution as X , equivalently εX and $\varepsilon|X|$ have the same distribution as X , where ε is an independent *symmetric random sign*, that is $\mathbb{P}(\varepsilon = -1) = \mathbb{P}(\varepsilon = 1) = \frac{1}{2}$. We usually denote by G a standard Gaussian random variable, that is a real-valued random variable with density $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$. For $p > 0$, we have $\mathbb{E}|G|^p = \frac{2^{p/2}\Gamma(\frac{p+1}{2})}{\sqrt{\pi}}$, where Γ stands for the gamma function. If p is a positive even integer, then $\mathbb{E}|G|^p = \frac{p!}{2^{p/2}(p/2)!} = 1 \cdot 3 \cdot \dots \cdot (p-1) = (p-1)!!$ (the double factorial of $p-1$). A nonnegative sequence $(a_n)_{n=1}^\infty$ is called *log-concave*, if it is supported on a contiguous set, that is the set $\{n \geq 1, a_n > 0\}$ is of the form $\{a, a+1, \dots, b\}$ for some $1 \leq a \leq b \leq \infty$, and $a_n^2 \geq a_{n-1}a_{n+1}$ for $n = 2, 3, \dots$. Sometimes we write x_+ which is $\max\{x, 0\}$.

Acknowledgements. We are indebted to Piotr Nayar for his suggestions regarding ultra sub-Gaussianity without which Theorem 2 would not have been discovered. We also thank Krzysztof Oleszkiewicz for his help and valuable feedback.

2 Results

2.1 Even moments

Nayar and Oleszkiewicz introduced in [28] the following notion of ultra sub-Gaussianity (as well as its multidimensional analogue): a random variable X is *ultra sub-Gaussian* if it is symmetric, has all moments finite and the sequence $(a_m)_{m=0}^\infty$ defined by $a_0 = 1$, $a_m = \frac{\mathbb{E}|X|^{2m}}{2^m m!}$, $m \geq 1$, is log-concave, that is $a_{m-1}a_{m+1} \leq a_m^2$ for every $m \geq 1$. This means that we have

$$\mathbb{E}|X|^4 \leq 3(\mathbb{E}|X|^2)^2 \quad (2)$$

($m = 1$) and reverse Cauchy-Schwarz estimates hold

$$\frac{2m-1}{2m+1}\mathbb{E}|X|^{2m-2}\mathbb{E}|X|^{2m+2} \leq (\mathbb{E}|X|^{2m})^2, \quad \text{for all } m \geq 2. \quad (3)$$

The normalisation is chosen such that if X is a standard Gaussian random variable G , then $a = (1, 1, \dots)$ is a constant sequence (since $\mathbb{E}|G|^{2m} = 1 \cdot 3 \cdot \dots \cdot (2m-1)$). For example, a symmetric random sign is ultra sub-Gaussian. Results from [28] (see Lemma 2 and Theorem 2 therein) assert that sums of independent sub-Gaussian random variables are sub-Gaussians and, consequently, it leads to Khinchin-type inequalities with sharp constants for even moments.

Theorem 1 (Nayar and Oleszkiewicz, [28]). *Let $2 \leq p < q$ be even integers and let X_1, \dots, X_n be independent ultra sub-Gaussian random variables. Then $X_1 + \dots + X_n$ is ultra sub-Gaussian and*

$$\left(\mathbb{E} \left| \sum_{i=1}^n X_i \right|^q \right)^{1/q} \leq \frac{(\mathbb{E}|G|^q)^{1/q}}{(\mathbb{E}|G|^p)^{1/p}} \left(\mathbb{E} \left| \sum_{i=1}^n X_i \right|^p \right)^{1/p}, \quad (4)$$

where G is a standard Gaussian random variable, so $\frac{(\mathbb{E}|G|^q)^{1/q}}{(\mathbb{E}|G|^p)^{1/p}} = \frac{[1 \cdot 3 \cdots (q-1)]^{1/q}}{[1 \cdot 3 \cdots (p-1)]^{1/p}}$.

In particular, in this elegant and slick way, Nayar and Oleszkiewicz obtained sharp constants in the classical Khinchin inequalities for even moments. We extend this result to symmetric random variables on consecutive integers by verifying that they are ultra sub-Gaussian.

Theorem 2. *Let $\rho_0 \in [0, 1]$ and let L be a positive integer. Let X_1, X_2, \dots be i.i.d. copies of a random variable X with $\mathbb{P}(X = 0) = \rho_0$ and $\mathbb{P}(X = -j) = \mathbb{P}(X = j) = \frac{1-\rho_0}{2L}$, $j = 1, \dots, L$. Then X is ultra sub-Gaussian if and only if $\rho_0 = 1$, or*

$$\rho_0 \leq 1 - \frac{2}{5} \frac{3L^2 + 3L - 1}{(L+1)(2L+1)}. \quad (5)$$

If this holds, then, consequently, for positive even integers $q > p \geq 2$, every $n \geq 1$ and reals a_1, \dots, a_n , we have

$$\left(\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^q \right)^{1/q} \leq C_{p,q} \left(\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^p \right)^{1/p} \quad (6)$$

with $C_{p,q} = \frac{[1 \cdot 3 \cdots (q-1)]^{1/q}}{[1 \cdot 3 \cdots (p-1)]^{1/p}}$ which is sharp.

Remark 3. It can be checked that the right hand side of (5) as a function of L is strictly decreasing. It converges to $1 - \frac{2}{5} \cdot \frac{3}{2} = \frac{2}{5}$ as $L \rightarrow \infty$, so (5) holds for every positive integer L as long as $\rho_0 \leq \frac{2}{5}$. This is precisely the condition under which a random variable θU is ultra sub-Gaussian, where θ is a Bernoulli random variable with parameter $1 - \rho_0$ and U is an independent random variable uniformly distributed on $[-1, 1]$ (which is not surprising because X/L converges in distribution to θU).

As explained in the proof, it is condition (2) that imposes the restriction (5) on ρ_0 , the mass put at 0. It turns out that condition (3) holds for every integers $L \geq 1$ and $m \geq 2$ (regardless ρ_0) and our proof proceeds by induction on L .

2.2 Second, third and higher moments

Here we first need to recall the classical notions of majorisation and Schur-convexity. Given two nonnegative sequences $(a_i)_{i=1}^n$ and $(b_i)_{i=1}^n$, we say that $(b_i)_{i=1}^n$ majorises $(a_i)_{i=1}^n$, denoted $(a_i) \prec (b_i)$ if

$$\sum_{i=1}^n a_i = \sum_{i=1}^n b_i \quad \text{and} \quad \sum_{i=1}^k a_i^* = \sum_{i=1}^k b_i^* \quad \text{for all } k = 1, \dots, n,$$

where $(a_i^*)_{i=1}^n$ and $(b_i^*)_{i=1}^n$ are nonincreasing permutations of $(a_i)_{i=1}^n$ and $(b_i)_{i=1}^n$ respectively. For example, $(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}) \prec (a_1, a_2, \dots, a_n) \prec (1, 0, \dots, 0)$ for every nonnegative sequence (a_i) with $\sum_{i=1}^n a_i = 1$. A function $\Psi: [0, \infty)^n \rightarrow \mathbb{R}$ which is symmetric (with respect to permuting the coordinates) is said to be *Schur-convex* if $\Psi(a) \leq \Psi(b)$ whenever $a \prec b$ and *Schur-concave* if $\Psi(a) \geq \Psi(b)$ whenever $a \prec b$. For instance, a function of the form $\Psi(a) = \sum_{i=1}^n \psi(a_i)$ with $\psi: [0, +\infty) \rightarrow \mathbb{R}$ being convex is Schur-convex. We refer to the classical monograph [14], or to [6] for a concise exposition of majorisation.

Let $p > 0$, let X_1, \dots, X_n be i.i.d. copies of a symmetric random variable X with finite p -th moment and consider the function $\Psi: [0, \infty)^n \rightarrow [0, +\infty)$ defined as

$$\Psi(a_1, \dots, a_n) = \mathbb{E} \left| \sum_{i=1}^n \sqrt{a_i} X_i \right|^p.$$

When X is uniform on $[-1, 1]$, Latała and Oleszkiewicz showed in [22] that Ψ is Schur-concave when $p \geq 2$ and Schur-convex when $1 \leq p \leq 2$ (see also [11] for a different proof). Such results give extremal sequences in Khinchin inequalities for any fixed number of summands n and in particular yield that optimal values of constants valid *for all* n are Gaussian. Suppose now that X is a symmetric random sign. Based on Eaton's criterion from [9], Komorowski showed in [16] that when $p \geq 3$ (the *easy regime*), Ψ is Schur-concave (which gives sharp constants in Khinchin inequalities; they were first found by Young in Theorem 9 in [40] and can be easily deduced from Corollary 2.5 from Pinelis' work [35]). For $p < 3$ (the *hard regime*), the Schur-convexity/concavity of Ψ fails and its behaviour is much more complicated (it is worth mentioning here the tantalizing *Zinn's doubling-conjecture* discussed in [41]: for $p \in (2, 3)$, $n \geq 1$, $a_1, \dots, a_n \geq 0$, we have $\Psi(\frac{a_1}{2}, \frac{a_1}{2}, \dots, \frac{a_n}{2}, \frac{a_n}{2}) \geq \Psi(a_1, \dots, a_n, 0, \dots, 0)$).

2.2.1 No atom at 0

If X is uniformly distributed on the set $\{-L, \dots, -1\} \cup \{1, \dots, L\}$, then we confirm for the easy regime that the Gaussian case yields sharp constants in Khinchin inequalities.

Theorem 4. *Let L be a positive integer. Let X_1, X_2, \dots be i.i.d. copies of a random variable X with $\mathbb{P}(X = -j) = \mathbb{P}(X = j) = \frac{1}{2L}$, $j = 1, \dots, L$. For every $n \geq 1$, reals a_1, \dots, a_n and $p \geq 3$, we have*

$$\left(\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^p \right)^{1/p} \leq C_p \left(\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^2 \right)^{1/2} \quad (7)$$

with $C_p = \sqrt{2} \left(\frac{\Gamma(\frac{p+1}{2})}{\sqrt{\pi}} \right)^{1/p}$ which is sharp.

Our inductive argument (on n) crucially uses independence and convexity of certain functions and is based on swapping the X_i one by one with independent Gaussians (the same yields a short proof of the classical Khinchin inequalities with sharp constants – see Theorem 1.1. in [12] as a nice illustration of such an approach; see also [5] and [11] where the same was used).

Remark 5. When p is a positive even integer, then of course Theorem 4 follows from Theorem 2. In this special case, it can be deduced from the main result of [30] (see also [31]). Indeed, after Newman, a random variable X of *type L* if $\mathbb{E}e^{zX}$, $z \in \mathbb{C}$ is well defined, possibly vanishes only if z is purely imaginary and there is a positive constant C such that $\mathbb{E}e^{zX} \leq e^{Cz^2}$ for all real z . Newman's result asserts that if the X_i are independent, each of type L, then (7) holds for every even integer p . It can be checked that X from Theorem 4 is of type L.

2.2.2 A Schur-convexity result for 3 atoms

We make an incremental progress for Schur-convexity in the easy regime for more general symmetric distributions than random signs by allowing an atom at zero.

Theorem 6. Let $\rho_0 \in [0, \frac{1}{2}]$. Let X_1, X_2, \dots be i.i.d. copies of a random variable X with $\mathbb{P}(X = 0) = \rho_0$ and $\mathbb{P}(X = -1) = \mathbb{P}(X = 1) = \frac{1-\rho_0}{2}$. Let $p \geq 3$. For every $n \geq 1$ and reals $a_1, \dots, a_n, b_1, \dots, b_n$ such that $(a_i^2)_{i=1}^n \prec (b_i^2)_{i=1}^n$, we have

$$\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^p \geq \mathbb{E} \left| \sum_{i=1}^n b_i X_i \right|^p. \quad (8)$$

Our proof follows a direct approach from Eaton's work [9], combined with rather standard techniques (used for instance in [12], or [11]) exploiting linearity and allowing to reduce verification of certain inequalities needed for averages of power functions $|\cdot|^p$ to *simple* (piecewise linear) functions.

As an immediate corollary, we obtain best constants in Khinchin inequalities (it can be done as, for instance, in the proof of Corollary 25 from [10]).

Corollary 7. Under the assumptions of Theorem 6 for every $n \geq 1$ and reals a_1, \dots, a_n , we have

$$\left(\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^p \right)^{1/p} \leq C_p \left(\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^2 \right)^{1/2} \quad (9)$$

with $C_p = \sqrt{2} \left(\frac{\Gamma(\frac{p+1}{2})}{\sqrt{\pi}} \right)^{1/p}$ which is sharp.

2.3 First and second moments

Littlewood posed in [24] the conjecture that the sharp constant in the classical Khinchin inequality for the first and second moment ($C_{2,1}$ in (1)) is attained in the case of exactly two nonzero equal weights ($n = 2, a_1 = a_2$). Not until 45 years after it had been stated, was Littlewood's conjecture proved, by Szarek in [37]. His argument was simplified by Tomaszewski in [38]. Haagerup, using integral representations for power functions, gave a different, much shorter proof in his seminal work [13] on sharp constants in Khinchin inequality. We show here that his argument is robust enough to cover certain cases for more atoms.

Theorem 8. Let $\rho_0 \in [\frac{1}{2}, 1]$ and let L be a positive integer. Let X_1, X_2, \dots be i.i.d. copies of a random variable X with $\mathbb{P}(X = 0) = \rho_0$ and $\mathbb{P}(X = -j) = \mathbb{P}(X = j) = \frac{1-\rho_0}{2L}, j = 1, \dots, L$. For every $n \geq 1$ and reals a_1, \dots, a_n , we have

$$\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right| \geq c_1 \left(\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^2 \right)^{1/2} \quad (10)$$

with $c_1 = \frac{\mathbb{E}|X|}{\sqrt{\mathbb{E}|X|^2}} = \sqrt{\frac{3(1-\rho_0)L(L+1)}{2(2L+1)}}$ which is sharp.

Remark 9. When $L = 1$ and $\rho_0 = \frac{1}{2}$, then $c_1 = \frac{1}{\sqrt{2}}$. Note that the X_i have the same distribution as $\frac{\varepsilon_i + \varepsilon'_i}{2}$, where $\varepsilon_1, \varepsilon'_1, \dots$ are i.i.d. symmetric random signs. Consequently, (10) follows directly from Szarek's result,

$$\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right| = \mathbb{E} \left| \sum_{i=1}^n a_i \frac{\varepsilon_i + \varepsilon'_i}{2} \right| \geq \frac{1}{\sqrt{2}} \left(\mathbb{E} \left| \sum_{i=1}^n a_i \frac{\varepsilon_i + \varepsilon'_i}{2} \right|^2 \right)^{1/2} = c_1 \left(\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^2 \right)^{1/2}.$$

The sharpness of c_1 can be seen by taking $n = 1$.

Remark 10. A substantial generalisation of (10) to arbitrary symmetric random variables and coefficients in Banach space is provided by Corollary 2.4 of [32]. The value of the constant c_1 obtained therein is sharp for three-valued random variables. Thus in this special case, that is of $L = 1$, it recovers Theorem 8.

3 Proofs

3.1 Even moments: Proof of Theorem 2

Let X be a random variable as in the statement of Theorem 2. If X is ultra sub-Gaussian, then so are $a_i X_i$, thus (6) follows directly from (4). The sharpness of $C_{p,q}$ can be seen by taking $a_1 = \dots = a_n = \frac{1}{\sqrt{n}}$, letting $n \rightarrow \infty$ and invoking the central limit theorem.

Showing that X is ultra sub-Gaussian amounts to verifying (2) and (3). Note that (2) is necessary for (6), as seen by taking there $n = 1$, $a_1 = 1$, $p = 2$ and $q = 4$. We can write the low order moments explicitly,

$$\begin{aligned} E|X|^2 &= \frac{1-\rho_0}{L} \sum_{k=1}^L k^2 = \frac{1-\rho_0}{6} (L+1)(2L+1), \\ \mathbb{E}|X|^4 &= \frac{1-\rho_0}{L} \sum_{k=1}^L k^4 = \frac{1-\rho_0}{30} (L+1)(2L+1)(3L^2+3L-1). \end{aligned}$$

As a result, (2) becomes

$$1 - \rho_0 \geq \frac{2}{5} \frac{3L^2 + 3L - 1}{(L+1)(2L+1)}$$

which is (5).

Condition (3) is equivalent to the following: for all integers $m \geq 2$

$$\frac{2m-1}{2m+1} \sum_{k=1}^L k^{2m-2} \sum_{k=1}^L k^{2m+2} \leq \left(\sum_{k=1}^L k^{2m} \right)^2.$$

It turns out to be true for every integers $L \geq 1$ and $m \geq 2$. We only have a rather lengthy cumbersome proof by induction on L . It turns out that the sequence $b_m = m \sum_{k=1}^L k^{m-1}$, $m \geq 1$, is log-concave, that is $b_{m-1} b_{m+1} \leq b_m^2$ for all $m \geq 2$ (Lemma 12 below, see also Remark 13). Then, for all $m \geq 2$, we have $b_{2m}^2 \geq b_{2m-1} b_{2m+1} \geq \sqrt{b_{2m-2} b_{2m}} \sqrt{b_{2m} b_{2m+2}}$, thus $b_{2m}^2 \geq b_{2m-2} b_{2m+2}$, which gives the above with a slightly better constant $\frac{(2m-1)(2m+3)}{(2m+1)^2}$ in place of the required $\frac{2m-1}{2m+1}$. Before showing the log-concavity of (b_m) , which will occupy the rest of this section, we make a remark about the i.i.d. assumption in Theorem 2.

Remark 11. Since Nayar and Oleszkiewicz's Theorem 1 does not require the X_i to be identically distributed, but only independent, we can drop that assumption in Theorem 2 and consider the X_i there to be independent (not necessarily identically distributed). We stated it in the i.i.d. case for simplicity.

As indicated, the log-concavity of (b_m) follows from the following lemma.

Lemma 12. For integers $q \geq 2$ and $n \geq 1$, we have

$$\frac{q(q+2)}{(q+1)^2} \sum_{k=1}^n k^{q-1} \sum_{k=1}^n k^{q+1} \leq \left(\sum_{k=1}^n k^q \right)^2.$$

Proof. Let

$$S_n(q) = \sum_{k=1}^n k^q.$$

By induction on n we show that for every $q \geq 2$, we have

$$q(q+2)S_n(q-1)S_n(q+1) \leq (q+1)^2 S_n(q)^2.$$

The statement is clearly true for $n = 1$. Assume the statement holds for some $n \geq 1$. For $n + 1$, using the inductive hypothesis, we have

$$\begin{aligned} & q(q+2)S_{n+1}(q-1)S_{n+1}(q+1) \\ &= q(q+2) \left(S_n(q-1) + (n+1)^{q-1} \right) \left(S_n(q+1) + (n+1)^{q+1} \right) \\ &\leq (q+1)^2 S_n(q)^2 + q(q+2)(n+1)^{q-1} \left(S_n(q+1) + (n+1)^2 S_n(q-1) + (n+1)^{q+1} \right). \end{aligned}$$

It suffices to show that this is at most

$$\begin{aligned} (q+1)^2 S_{n+1}(q)^2 &= (q+1)^2 (S_n(q) + (n+1)^q)^2 \\ &= (q+1)^2 S_n(q)^2 + 2(q+1)^2 (n+1)^q S_n(q) + (q+1)^2 (n+1)^{2q} \end{aligned}$$

which is equivalent to showing that for $q \geq 2$ and $n \geq 1$, we have

$$q(q+2) \left(S_n(q+1) + (n+1)^2 S_n(q-1) \right) \leq 2(q+1)^2 (n+1) S_n(q) + (n+1)^{q+1}.$$

We shall do this inductively on n . The base case

$$5q(q+2) \leq 4(q+1)^2 + 2^{q+1}, \quad q \geq 2 \tag{11}$$

is verified later. By the inductive hypothesis, for $n \geq 1$, we have

$$\begin{aligned} & q(q+2) \left(S_{n+1}(q+1) + (n+2)^2 S_{n+1}(q-1) \right) \\ &\leq 2(q+1)^2 (n+1) S_n(q) + (n+1)^{q+1} \\ &\quad - q(q+2)(n+1)^2 S_n(q-1) \\ &\quad + q(q+2) \left((n+1)^{q+1} + (n+2)^2 S_{n+1}(q-1) \right) \\ &= 2(q+1)^2 (n+1) S_n(q) + q(q+2)(2n+3) S_n(q-1) \\ &\quad + (q+1)^2 (n+1)^{q+1} + q(q+2)(n+2)^2 (n+1)^{q-1}. \end{aligned}$$

It suffices to show that this is at most

$$2(q+1)^2 (n+2) S_{n+1}(q) + (n+2)^{q+1}$$

which is equivalent to showing that for $q \geq 2$ and $n \geq 1$, we have

$$\begin{aligned} q(q+2)(2n+3)S_n(q-1) + (q+1)^2(n+1)^{q+1} + q(q+2)(n+2)^2(n+1)^{q-1} \\ \leq 2(q+1)^2S_n(q) + 2(q+1)^2(n+2)(n+1)^q + (n+2)^{q+1}. \end{aligned}$$

Writing $(n+2)^2 = (n+1)^2 + 2(n+1) + 1$ as well as $n+2 = (n+1) + 1$ and simplifying gives

$$\begin{aligned} q(q+2)(2n+3)S_n(q-1) + q(q+2)(n+1)^{q-1} \\ \leq 2(q+1)^2S_n(q) + (n+2)^{q+1} + (n+1)^{q+1} + 2(n+1)^q. \end{aligned}$$

We show this again by induction on n . The base case

$$5q(q+2) + q(q+2)2^{q-1} \leq 2(q+1)^2 + 3^{q+1} + 2^{q+2}, \quad q \geq 2 \quad (12)$$

is verified later. By the inductive hypothesis, for $n \geq 1$, we have

$$\begin{aligned} 2(q+1)^2S_{n+1}(q) &\geq q(q+2)(2n+3)S_n(q-1) + q(q+2)(n+1)^{q-1} \\ &\quad - (n+2)^{q+1} - (n+1)^{q+1} - 2(n+1)^q \\ &\quad + 2(q+1)^2(n+1)^q. \end{aligned}$$

It suffices to show that this is at least

$$q(q+2)(2n+5)S_{n+1}(q-1) + q(q+2)(n+2)^{q-1} - (n+3)^{q+1} - (n+2)^{q+1} - 2(n+2)^q$$

which after simplifying is equivalent to showing that for $q \geq 2$ and $n \geq 1$, we have

$$\begin{aligned} (n+3)^{q+1} + 2(n+2)^q \\ \geq 2q(q+2)S_{n+1}(q-1) + q(q+2)(n+2)^{q-1} + (n+1)^{q+1}. \end{aligned}$$

We show this again by induction on n . The base case

$$4^{q+1} + 2 \cdot 3^q \geq 2q(q+2)(1+2^{q-1}) + q(q+2)3^{q-1} + 2^{q+1}, \quad q \geq 2 \quad (13)$$

is verified later. By the inductive hypothesis, for $n \geq 1$, we have

$$\begin{aligned} 2q(q+2)S_{n+2}(q-1) &\leq 2q(q+2)(n+2)^{q-1} + (n+3)^{q+1} + 2(n+2)^q \\ &\quad - q(q+2)(n+2)^{q-1} - (n+1)^{q+1}. \end{aligned}$$

It suffices to show that this is at most

$$(n+4)^{q+1} + 2(n+3)^q - q(q+2)(n+3)^{q-1} - (n+2)^{q+1}$$

which after simplifying is equivalent to showing that for $q \geq 2$ and $n \geq 1$, we have

$$\begin{aligned} q(q+2)(n+3)^{q-1} + q(q+2)(n+2)^{q-1} + (n+3)^{q+1} + 2(n+2)^q + (n+2)^{q+1} \\ \leq (n+4)^{q+1} + 2(n+3)^q + (n+1)^{q+1}. \end{aligned}$$

Setting $x = n + 1$ and swapping q for $q + 1$, we see that it is enough to show that for every $q \geq 1$ the function

$$f_q(x) = (x+3)^{q+2} - (x+2)^{q+2} - (x+1)^{q+2} + x^{q+2} + 2(x+2)^{q+1} - 2(x+1)^{q+1} \\ - (q+1)(q+3)\left((x+2)^q + (x+1)^q\right)$$

is nonnegative for $x \geq 2$. We show in fact that it is nonnegative for $x \geq 0$. From now on we use that q is an integer and apply the binomial formula (if it was not, we could proceed by writing Taylor's expansion instead, but we would need to verify that the $[q]$ derivative of f_q is nonnegative). The coefficients at x^{q+2} and x^{q+1} vanish and we have

$$f_q(x) = \sum_{k=0}^q \left[\binom{q+2}{k} (3^{q+2-k} - 2^{q+2-k} - 1) + 2 \binom{q+1}{k} (2^{q+2-k} - 1) \right. \\ \left. - (q+1)(q+3) \binom{q}{k} (2^{q-k} + 1) \right] x^k.$$

It suffices to show that for every $0 \leq k \leq q$, we have

$$\binom{q+2}{k} (3^{q+2-k} - 2^{q+2-k} - 1) + 2 \binom{q+1}{k} (2^{q+2-k} - 1) \geq (q+1)(q+3) \binom{q}{k} (2^{q-k} + 1)$$

or dividing by $\binom{q}{k}$ and simplifying,

$$\frac{(q+2)}{(q+2-k)(q+1-k)} (3^{q+2-k} - 2^{q+2-k} - 1) + \frac{2}{q+1-k} (2^{q+2-k} - 1) \geq (q+3)(2^{q-k} + 1).$$

Setting $l = q - k$ and multiplying through by $\frac{(l+1)(l+2)}{q+3}$, it becomes

$$\frac{q+2}{q+3} (3^{l+2} - 2^{l+2} - 1) + \frac{1}{q+3} 2(l+2)(2^{l+2} - 1) \geq (l+1)(l+2)(1+2^l). \quad (14)$$

We show this for every integers $0 \leq l \leq q$ in the following steps.

Step 1. We check that (14) for $l = 0, 1$ becomes equality and for $l = 2, 3$, it becomes respectively $\frac{4(q+1)}{q+3} \geq 0$, $\frac{30(q+1)}{q+3} \geq 0$, so it holds true for $l \leq 3$.

Step 2. For integers $q \geq l \geq 4$ we bound the left hand side below by

$$\frac{6}{7} (3^{l+2} - 2^{l+2} - 1)$$

and verify that

$$\frac{6}{7} (3^{l+2} - 2^{l+2} - 1) \geq (l+1)(l+2)(2^l + 1), \quad l \geq 4. \quad (15)$$

To finish the proof, we shall now show the omitted inductive base inequalities (11), (12), (13) as well as final estimate (15). \square

Proof of (11). The right hand side minus the left hand side is

$$2^{q+1} + 4(q+1)^2 - 5q(q+2) = 2^{q+1} - q^2 - 2q + 4 = 2^{q+1} - (q+1)^2 + 5.$$

This is nonnegative for $q = 2$. For $q \geq 3$, we use that $2^x \geq x^2$ for $x \geq 4$ (which is easy to check). \square

Proof of (12). The right hand side minus the left hand side is

$$3^{q+1} + 2^{q+2} + 2(q+1)^2 - 2^{q-1}q(q+2) - 5q(q+2) = 3^{q+1} - 2^{q-1}q(q+2) + 2^{q+2} - 3q(q+2) + 2.$$

We check directly that this is nonnegative for $q = 2, 3, 4$. For $q \geq 5$, easy inductive arguments show that $3^{q+1} \geq 2^{q-1}q(q+2)$ and $2^{q+2} \geq 3q(q+2)$. \square

Proof of (13). We check the inequality directly for $q = 2, 3, \dots, 10$. For $q \geq 11$, easy inductive arguments show that $4^{q+1} \geq 3^{q-1}q(q+2)$ and $3^q \geq (1 + 2^{q-1})q(q+2) + 2^q$. Multiplying the second inequality and adding to the first one gives (13). \square

Proof of (15). We verify the inequality for $l = 4$. Then, by induction, for $l \geq 4$, we have

$$\frac{6}{7}3^{l+3} \geq 3(l+1)(l+2)(2^l+1) + 3 \cdot \frac{6}{7}2^{l+2} + 3 \cdot \frac{6}{7}.$$

It remains to check that this is at least $(l+2)(l+3)(2^{l+1}+1) + \frac{6}{7}2^{l+3} + \frac{6}{7}$. The difference is

$$2^l \left(l^2 - l - \frac{18}{7} \right) + 2l^2 + 4l + 2 \cdot \frac{6}{7}$$

which is clearly positive for $l \geq 4$. \square

Remark 13. It is natural to ask what other symmetric discrete random variables are ultra sub-Gaussian. We pose the following question: *is it true that for every positive integer L and every positive monotone log-concave sequence $(x_n)_{n=1}^L$ of length L , the function $F(t) = \log \left[t \sum_{k=1}^L x_k^t \right]$ is concave on $(0, \infty)$?* This would imply that a symmetric discrete random variable X with $\mathbb{P}(X=0) = \rho_0$ and $\mathbb{P}(X=-x_k) = \mathbb{P}(X=x_k) = \frac{1-\rho_0}{2L}$, $k = 1, \dots, L$, for some $\rho_0 \in [0, 1]$ satisfies (3), hence X would be ultra sub-Gaussian if and only if it satisfies (2). When $x_k = k$, Lemma 12 implies that the sequence $(F(t))_{t=1}^\infty$ is concave. This question also naturally appears in a different context (see [25]). Moreover, it is known that if $\gamma > 0$ and $f: (a, b) \rightarrow (0, +\infty)$ is such that f^γ is concave on (a, b) , then $t \mapsto \log \left((t+\gamma) \int_a^b f(x)^\gamma dx \right)$ is concave on $(-\gamma, \infty)$ (see [7], [8]). It is therefore tempting to ask for a stronger statement: for $\gamma > 0$ and a positive monotone concave sequence $(y_n)_{n=1}^N$, is the function

$$t \mapsto \log \left((t+\gamma) \sum_{n=1}^N y_n^{t/\gamma} \right)$$

concave on $(-\gamma, \infty)$? As pointed to us by Melbourne (also see [25]), the examples of sequences $x = (\frac{1}{4}, \frac{1}{2}, 1, \frac{1}{2}, \frac{1}{4})$ and $y = (1, 2, 3, 2, 1)$ show that the assumption of monotonicity is needed in both questions.

Remark 14. The question from Remark 13 has the affirmative answer for $N = 3$. In this case, the assumption of the log-concavity of (x_n) is not needed (the assertion does not depend on the order of the x_n and given 3 numbers x_1, x_2, x_3 we can always order them to form a concave sequence by choosing x_2 to be $\max\{x_1, x_2, x_3\}$). Thus, we claim that for every positive numbers s, t, a, b, c , we have

$$\frac{s+t}{2} \left(a^{\frac{s+t}{2}} + b^{\frac{s+t}{2}} + c^{\frac{s+t}{2}} \right) \geq \sqrt{s(a^s + b^s + c^s) \cdot t(a^t + b^t + c^t)}$$

which is equivalent to the conjecture being true when $N = 3$. To show the above, by homogeneity, it is enough to consider $s + t = 2$, that is to prove that for every positive numbers a, b, c and $s \in (0, 2)$, we have

$$a + b + c \geq \sqrt{s(2-s)(a^s + b^s + c^s)(a^{2-s} + b^{2-s} + c^{2-s})},$$

or, after squaring and rearranging,

$$(s-1)^2(a^2 + b^2 + c^2) + 2(ab + bc + ca) \geq s(2-s) \left(a^s b^{2-s} + a^{2-s} b^s + b^s c^{2-s} + b^{2-s} c^s + c^s a^{2-s} + c^{2-s} a^s \right).$$

This holds if we show that for every positive a, b and $s \in (0, 2)$, we have

$$\frac{(s-1)^2}{2}(a^2 + b^2) + 2ab \geq s(2-s)(a^s b^{2-s} + a^{2-s} b^s).$$

This follows from the following claim (divide through by ab and set $e^x = \frac{a}{b}$, $\theta = 1 - s$).

Claim. Let $\theta \in [-1, 1]$ and $x \in \mathbb{R}$. Then

$$2 + \theta^2 \cosh x \geq 2(1 - \theta^2) \cosh(\theta x).$$

Proof of the claim. Expanding into a power series yields,

$$\begin{aligned} S = 2 + \theta^2 \cosh x - 2(1 - \theta^2) \cosh(\theta x) &= 2 + \theta^2 \left(1 + \frac{x^2}{2} + \sum_{k \geq 2} \frac{x^{2k}}{(2k)!} \right) \\ &\quad - 2(1 - \theta^2) \left(1 + \frac{\theta^2 x^2}{2} + \sum_{k \geq 2} \frac{\theta^{2k} x^{2k}}{(2k)!} \right) \\ &= \theta^2 \left[3 + \frac{2\theta^2 - 1}{2} x^2 + \sum_{k \geq 2} \frac{1 - 2\theta^{2k-2}(1 - \theta^2)}{(2k)!} x^{2k} \right]. \end{aligned}$$

Note that for $k \geq 2$, we have $1 - 2\theta^{2k-2}(1 - \theta^2) \geq 1 - 2\theta^2(1 - \theta^2) \geq \frac{1}{2} > 0$, so if $2\theta^2 - 1 \geq 0$, then S is clearly positive. If $2\theta^2 - 1 < 0$, then using $1 - 2\theta^{2k-2}(1 - \theta^2) \geq 1 - 2\theta^2$, we get

$$\begin{aligned} S &\geq \theta^2 \left[3 + \frac{2\theta^2 - 1}{2} x^2 + (1 - 2\theta^2) \sum_{k \geq 2} \frac{x^{2k}}{(2k)!} \right] \\ &= \theta^2 \left[3 + (1 - 2\theta^2) \left(-\frac{x^2}{2} + \sum_{k \geq 2} \frac{x^{2k}}{(2k)!} \right) \right]. \end{aligned}$$

It remains to observe that $-\frac{x^2}{2} + \sum_{k \geq 2} \frac{x^{2k}}{(2k)!} \geq -\frac{x^2}{2} + \frac{x^4}{24} = \frac{(x^2-6)^2}{24} - \frac{3}{2} \geq -\frac{3}{2}$, thus $S \geq \theta^2(3 - \frac{3}{2}(1 - 2\theta^2)) = \theta^2(\frac{3}{2} + 3\theta^2) \geq 0$. \square

Combining Remarks 13 and 14 yields the following corollary.

Corollary 15. Let $\rho_0 \in [0, 1)$ and let x_1, x_2, x_3 be positive. Let X_1, X_2, \dots be i.i.d. copies of a random variable X with $\mathbb{P}(X = 0) = \rho_0$ and $\mathbb{P}(X = -x_j) = \mathbb{P}(X = x_j) = \frac{1-\rho_0}{6}$, $j = 1, 2, 3$. Then X is ultra sub-Gaussian if and only if $x_1^4 + x_2^4 + x_3^4 \leq (1 - \rho_0)(x_1^2 + x_2^2 + x_3^2)^2$. Moreover in this case, (6) holds.

3.2 Second, third and higher moments, no atom at 0: Proof of Theorem 4

The value of the constant C_p equals the p -th moment of a standard Gaussian random variable and is seen to be sharp by taking $a_1 = \dots = a_n = \frac{1}{\sqrt{n}}$, letting $n \rightarrow \infty$ and applying the central limit theorem.

To establish (7), we shall follow an inductive argument exploiting independence based on swapping the X_i one by one with independent Gaussians (similar ideas have appeared e.g. in [5], [11] or [12]). An appropriate normalisation of the Gaussians is crucial and we shall choose them to have the same variance as the X_i .

Let

$$\sigma = \sqrt{\mathbb{E}|X_1|^2} = \left(\frac{(L+1)(2L+1)}{6} \right)^{1/2} \quad (16)$$

and let G_1, G_2, \dots be i.i.d. centred Gaussian random variables with variance σ^2 . Since

$$C_p^p \left(\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^2 \right)^{p/2} = C_p^p \left(\sum_{i=1}^n a_i^2 \right)^{p/2} \sigma^{p/2} = \mathbb{E} \left| \sum_{i=1}^n a_i G_i \right|^p,$$

inequality (7) is equivalent to

$$\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^p \leq \mathbb{E} \left| \sum_{i=1}^n a_i G_i \right|^p.$$

By independence and induction, it suffices to show that for every reals a, b , we have

$$\mathbb{E}|a + bX_1|^p \leq \mathbb{E}|a + bG_1|^p. \quad (17)$$

This will follow from the following claim.

Claim. For every convex nondecreasing function $h: [0, +\infty) \rightarrow [0, +\infty)$, we have

$$\mathbb{E}h(X_1^2) \leq \mathbb{E}h(G_1^2). \quad (18)$$

Indeed, (17) for $b = 0$ is clear. Assuming $b \neq 0$, by homogeneity, (17) is equivalent to

$$\mathbb{E}|a + X_1|^p \leq \mathbb{E}|a + G_1|^p.$$

Using the symmetry of X_1 , we can write

$$2\mathbb{E}|a + X_1|^p = \mathbb{E}|a + |X_1||^p + \mathbb{E}|a - |X_1||^p = \mathbb{E}h_a(X_1^2),$$

where

$$h_a(x) = |a + \sqrt{x}|^p + |a - \sqrt{x}|^p, \quad x \geq 0 \quad (19)$$

(and similarly for G_1). The convexity of h_a is established in the following standard lemma (see also e.g. Proposition 3.1 in [12]).

Lemma 16. *Let $p \geq 3$, $a \in \mathbb{R}$. Then h_a defined in (19) is convex nondecreasing on $[0, \infty)$.*

Proof. The case $a = 0$ is clear (and the assertion holds for $p \geq 2$). The case $a \neq 0$ reduces by homogeneity to, say $a = 1$. We have

$$h'_1(x) = \frac{p}{2\sqrt{x}} \left[1 + \sqrt{x}|^{p-1} + \operatorname{sgn}(\sqrt{x} - 1)|\sqrt{x} - 1|^{p-1} \right]$$

and it suffices to show that the function $g(y) = \frac{|1+y|^{p-1} + \operatorname{sgn}(y-1)|y-1|^{p-1}}{y}$ is nondecreasing on $(0, \infty)$. Call the numerator $f(y)$. Since $g(y) = \frac{f(y)-f(0)}{y-0}$, it suffices to show that f is convex on $(0, \infty)$. We have $f'(y) = (p-1)(|1+y|^{p-2} + |y-1|^{p-2})$ which is convex on \mathbb{R} for $p \geq 3$, hence nondecreasing on $(0, \infty)$ (as being even). This justifies that h'_1 is nondecreasing, hence h_1 is convex. Since $h'_1(0) = f'(0) = 2(p-1) > 0$, we get $h'_1(x) \geq h'_1(0) > 0$, so h_1 is increasing on $(0, \infty)$. \square

Thus $2\mathbb{E}|a + X_1|^p = \mathbb{E}h_a(X_1^2) \leq \mathbb{E}h_a(G_1^2) = 2\mathbb{E}|a + G_1|^p$ by the claim, as desired. It remains to prove the claim.

Proof of the claim. When $L = 1$, the claim follows immediately because $X_1^2 = 1$ and by Jensen's inequality, $\mathbb{E}h(G_1^2) \geq h(\mathbb{E}G_1^2) = h(1) = \mathbb{E}h(X_1^2)$. We shall assume from now on that $L \geq 2$.

By standard approximation arguments, it suffices to show that the claim holds for $h(x) = (x - a)_+$ for every $a > 0$. Here and throughout $x_+ = \max\{x, 0\}$. Note that

$$\mathbb{E}(X_1^2 - a)_+ = \frac{1}{2L} \sum_{k=-L}^L (k^2 - a)_+ = \frac{1}{L} \sum_{k=\lceil\sqrt{a}\rceil}^L (k^2 - a)$$

and

$$\mathbb{E}(G_1^2 - a)_+ = \int_{-\infty}^{\infty} (x^2 - a)_+ \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2} dx = \sqrt{\frac{2}{\pi\sigma^2}} \int_{\sqrt{a}}^{\infty} (x^2 - a) e^{-x^2/2\sigma^2} dx$$

with σ (depending on L) defined by (16). Fix an integer $L \geq 2$ and set for nonnegative a ,

$$f(a) = \sqrt{\frac{2}{\pi\sigma^2}} \int_{\sqrt{a}}^{\infty} (x^2 - a) e^{-x^2/2\sigma^2} dx - \frac{1}{L} \sum_{k=\lceil\sqrt{a}\rceil}^L (k^2 - a).$$

Our goal is to show that $f(a) \geq 0$ for every $a \geq 0$. This is clear for $a > L^2$ because then the second term is 0. Note that f is continuous (because $x \mapsto x_+$ is continuous). For $a \in (b^2, (b+1)^2)$ with $b \in \{0, 1, \dots, L-1\}$ our expression becomes

$$f(a) = \sqrt{\frac{2}{\pi\sigma^2}} \int_{\sqrt{a}}^{\infty} (x^2 - a) e^{-x^2/2\sigma^2} dx - \frac{1}{L} \sum_{k=b+1}^L (k^2 - a),$$

is differentiable and

$$\begin{aligned} f'(a) &= -\sqrt{\frac{2}{\pi\sigma^2}} \int_{\sqrt{a}}^{\infty} e^{-x^2/2\sigma^2} dx - \frac{1}{L} \sum_{k=b+1}^L (-1) \\ &= -\sqrt{\frac{2}{\pi\sigma^2}} \int_{\sqrt{a}}^{\infty} e^{-x^2/2\sigma^2} dx + \frac{L-b}{L}, \quad a \in (b^2, (b+1)^2). \end{aligned} \quad (20)$$

Bounding $b < \sqrt{a}$ yields

$$\begin{aligned} f'(a) &\geq -\sqrt{\frac{2}{\pi\sigma^2}} \int_{\sqrt{a}}^{\infty} e^{-x^2/2\sigma^2} dx + \frac{L - \sqrt{a}}{L} \\ &= -\sqrt{\frac{2}{\pi}} \int_{\sqrt{a}/\sigma}^{\infty} e^{-x^2/2} dx + \left(1 - \frac{\sqrt{a}}{L}\right). \end{aligned}$$

Call the right hand side $\tilde{g}(a)$,

$$\tilde{g}(a) = -\sqrt{\frac{2}{\pi}} \int_{\sqrt{a}/\sigma}^{\infty} e^{-x^2/2} dx + \left(1 - \frac{\sqrt{a}}{L}\right).$$

We have obtained $f' \geq \tilde{g}$ on $(0, L^2)$ (except for the points $1^2, 2^2, \dots$). Since f is absolutely continuous and $f(0) = 0$, we can write $f(a) = \int_0^a f'(x) dx$ and consequently

$$f(a) \geq g(a), \quad a \in [0, L^2],$$

where we define

$$g(a) = \int_0^a \tilde{g}(x) dx.$$

Note: $g''(a) = \tilde{g}'(a) = \frac{1}{2\sqrt{a}} \left(\sqrt{\frac{2}{\pi}} \frac{1}{\sigma} e^{-\frac{a}{2\sigma^2}} - \frac{1}{L} \right)$ which changes sign from positive to negative (since $\sqrt{\frac{2}{\pi}} \frac{1}{\sigma} - \frac{1}{L} > 0$ for $L \geq 2$). This implies that g' is first strictly increasing, then strictly decreasing and together with $g'(0) = \tilde{g}(0) = 0$, $g'(\infty) = -\infty$, it gives that g' is first positive, then negative. Consequently, g is first strictly increasing and then strictly decreasing. Since $g(0) = 0$, to conclude that g is nonnegative on $[0, L^2]$ (hence f), it suffices to check that $g(L^2) \geq 0$. We have,

$$\begin{aligned} g(L^2) &= \int_0^{L^2} \left[-\sqrt{\frac{2}{\pi}} \int_{\sqrt{a}/\sigma}^{\infty} e^{-x^2/2} dx + \left(1 - \frac{\sqrt{a}}{L}\right) \right] da \\ &= \int_0^{L^2} \left[\sqrt{\frac{2}{\pi}} \int_0^{\sqrt{a}/\sigma} e^{-x^2/2} dx - \frac{\sqrt{a}}{L} \right] da \\ &= \sqrt{\frac{2}{\pi}} \int_0^{L/\sigma} (L^2 - \sigma^2 x^2) e^{-x^2/2} dx - \frac{2}{3} L^2. \end{aligned}$$

Note that for $t = t(L) = \frac{L^2}{\sigma^2} = \frac{6L^2}{(L+1)(2L+1)}$, the expression $\frac{g(L^2)}{\sigma^2}$ becomes

$$h(t) = \sqrt{\frac{2}{\pi}} \int_0^{\sqrt{t}} (t - x^2) e^{-x^2/2} dx - \frac{2}{3} t.$$

We have,

$$h'(t) = \sqrt{\frac{2}{\pi}} \int_0^{\sqrt{t}} e^{-x^2/2} dx - \frac{2}{3}.$$

For $L \geq 7$, we have $t \geq t_0 = t(7) = \frac{49}{20}$. We check that $h'(t_0) = h'(\frac{49}{20}) > 0.2$ and since h' is increasing, $h'(t)$ is positive for $t \geq t_0$, hence $h(t) \geq h(t_0) = h(\frac{49}{20}) > 0.01$ for $t \geq t_0$. Consequently, $g(L^2) > 0$ for every $L \geq 7$, which completes the proof for $L \geq 7$.

It remains to address the cases $2 \leq L \leq 6$. Here lower-bounding f by g incurs too much loss, so we show that f is nonnegative on $[0, L^2]$ by direct computations. First note that

$f'(a)$ (see (20)) is strictly increasing on each interval $a \in (b^2, (b+1)^2)$, $b \in \{0, 1, \dots, L-1\}$. Clearly $f'(0+) = 0$ and we check that $\theta_{L,b} = f'(b^2+) > 0$ for every $b \in \{1, \dots, L-2\}$ and $3 \leq L \leq 6$ (see Table 1), so $f(a)$ is strictly increasing for $a \in (0, (L-1)^2)$. Since $f(0) = 0$, this shows that $f(a) > 0$ for $a \in (0, (L-1)^2)$. On the interval $((L-1)^2, L^2)$, we use the convexity of f and we lower-bound f by its tangent at $a = (L-1)^2+$ with the slope $\theta_{L,L-1}$ (which is negative), that is $f(a) \geq \theta_{L,L-1}(a - (L-1)^2) + f((L-1)^2)$. It remains to check that $v_L = \theta_{L,L-1}(2L-1) + f((L-1)^2)$, the values of the right hand side at the end point $a = L^2$, are positive. We have, $v_2 > 0.2$, $v_3 > 0.7$, $v_4 > 1.2$, $v_5 > 1.9$, $v_6 > 2.6$. This finishes the proof. \square

Table 1: Lower bounds on the values of the slopes $\theta_{L,b} = f'(b^2+)$.

	$b = 1$	$b = 2$	$b = 3$	$b = 4$
$\theta_{3,b}$	0.02			
$\theta_{4,b}$	0.03	0.03		
$\theta_{5,b}$	0.03	0.05	0.03	
$\theta_{6,b}$	0.03	0.05	0.05	0.02

Remark 17. We can drop the assumption in Theorem 4 of the X_i being identically distributed and only assume their independence (we stated it in the i.i.d. case for simplicity). The proof does not change: we only have to choose the independent Gaussian random variables G_i to be such that $\mathbb{E}|G_i|^2 = \mathbb{E}|X_i|^2$ and then (18), hence (17) holds for each X_i .

3.3 A Schur-convexity result for 3 atoms: Proof of Theorem 6

We need to begin with two technical lemmas. Let \mathcal{C} be the linear space of all continuous functions on \mathbb{R} equipped with pointwise topology. Let $\mathcal{C}_1 \subset \mathcal{C}$ be the cone of all odd functions on \mathbb{R} which are nondecreasing convex on $(0, +\infty)$ and let $\mathcal{C}_2 \subset \mathcal{C}$ be the cone of all even functions on \mathbb{R} which are nondecreasing convex on $(0, +\infty)$. Note that \mathcal{C}_2 is the closure (in the pointwise topology) of the set $\mathcal{S} = \{(|x| - \gamma)_+, \gamma \geq 0\}$.

Lemma 18. *Let $q \geq 2$, $w \geq 0$ and $\phi_w(x) = \text{sgn}(x+w)|x+w|^q + \text{sgn}(x-w)|x-w|^q$, $x \in \mathbb{R}$. Then $\phi_w \in \mathcal{C}_1$. Let $r_w(x) = \frac{\phi_w(x)}{x}$, $x \in \mathbb{R}$ (with the value at $x = 0$ understood as the limit). Then $r_w \in \mathcal{C}_2$.*

Proof. The case $w = 0$ is clear. For $w > 0$, verifying that $\phi_w \in \mathcal{C}_1$ and $r_w \in \mathcal{C}_2$, by homogeneity, is equivalent to doing so for $w = 1$. Let $w = 1$ and denote $\phi = \phi_1$ and $r = r_1$. Suppose we have shown that $r \in \mathcal{C}_2$. Then, plainly, $\phi(x) = xr(x)$ is also nondecreasing on $(0, \infty)$ and $\phi''(x) = (r(x) + xr'(x))' = 2r'(x) + xr''(x)$ is nonnegative on $(0, \infty)$ since r' and r'' are nonnegative on $(0, \infty)$.

It remains to prove that $r \in \mathcal{C}_2$. Plainly $\phi(x)$ is odd and thus $r(x)$ is even. Thus we consider $x > 0$.

Case 1. $x \geq 1$. We have, $\phi(x) = (x+1)^q + (x-1)^q$,

$$r'(x) = \frac{\phi'(x)}{x} - \frac{\phi(x)}{x^2} = q \frac{(x+1)^{q-1} + (x-1)^{q-1}}{x} - \frac{(x+1)^q + (x-1)^q}{x^2}$$

and

$$\begin{aligned} x^3 r''(x) &= x^3 \left[\frac{\phi''(x)}{x} - 2 \frac{\phi'(x)}{x^2} + 2 \frac{\phi(x)}{x^3} \right] \\ &= q(q-1)x^2 \left[(x+1)^{q-2} + (x-1)^{q-2} \right] \\ &\quad - 2qx \left[(x+1)^{q-1} + (x-1)^{q-1} \right] + 2 \left[(x+1)^q + (x-1)^q \right]. \end{aligned}$$

Note that taking one more derivative gives

$$(x^3 r''(x))' = q(q-1)(q-2)x^2 \left[(x+1)^{q-3} + (x-1)^{q-3} \right]$$

which is clearly positive for $x > 1$ since $q \geq 2$. Thus, for $x > 1$, we have

$$x^3 r''(x) > r''(1) = q(q-1) \cdot 2^{q-2} - 2q \cdot 2^{q-1} + 2 \cdot 2^q = 2^{q-2} \left(\left(q - \frac{5}{2} \right)^2 + \frac{7}{4} \right) > 0.$$

Therefore, $r''(x) > 0$ for $x > 1$. Since $r'(1) = q2^{q-1} - 2^q = 2^{q-1}(q-2) \geq 0$, we also get that $r'(x)$ is positive for $x > 1$.

Case 2. $0 < x < 1$. The argument and the computations are very similar to Case 1. We have, $\phi(x) = (1+x)^q - (1-x)^q$,

$$r'(x) = \frac{\phi'(x)}{x} - \frac{\phi(x)}{x^2} = q \frac{(1+x)^{q-1} + (1-x)^{q-1}}{x} - \frac{(1+x)^q - (1-x)^q}{x^2}$$

and

$$\begin{aligned} x^3 r''(x) &= x^3 \left[\frac{\phi''(x)}{x} - 2 \frac{\phi'(x)}{x^2} + 2 \frac{\phi(x)}{x^3} \right] \\ &= q(q-1)x^2 \left[(1+x)^{q-2} - (1-x)^{q-2} \right] \\ &\quad - 2qx \left[(1+x)^{q-1} + (1-x)^{q-1} \right] + 2 \left[(1+x)^q - (1-x)^q \right]. \end{aligned}$$

Taking one more derivative yields

$$(x^3 r''(x))' = q(q-1)(q-2)x^2 \left[(1+x)^{q-3} + (1-x)^{q-3} \right].$$

If $q > 2$, this is positive for $0 < x < 1$. Then in this case, consequently, $x^3 r''(x) > x^3 r''(x) \Big|_{x=0} = 0$, so $r''(x)$ is positive for $0 < x < 1$. As a result, $r'(x) > r'(0+) = 0$ for $0 < x < 1$. If $q = 2$, we simply have $\phi(x) = 4x$ and $r(x) = 4$.

Combining the cases, we see that both r' and r'' are nonnegative on $(0, +\infty)$, which finishes the proof. \square

Lemma 19. *The best constant D such that the inequality*

$$D \cdot \left[\frac{\phi(a+b) - \phi(b-a)}{2a} - \frac{\phi(a+b) + \phi(b-a)}{2b} \right] \geq \left[\frac{\phi(b)}{b} - \frac{\phi(a)}{a} \right] \quad (21)$$

holds for all $0 < a < b$ and every function $\phi(x)$ of the form $xr(x)$, $r \in \mathcal{C}_2$, is $D = 1$.

Proof. For $\phi(x) = xr(x)$, $r(x) = |x|$, by homogeneity, inequality (21) is equivalent to: for all $0 < a < 1$, we have

$$D \cdot \left[\frac{(1+a)^2 - (1-a)^2}{2a} - \frac{(1+a)^2 + (1-a)^2}{2} \right] \geq 1-a,$$

that is $D \cdot (1-a^2) \geq (1-a)$ for all $0 < a < 1$, which holds if and only if $D \geq 1$. Now we show that in fact (21) holds with $D = 1$ for every $\phi(x) = xr(x)$, where $r \in \mathcal{C}_2$. Since \mathcal{C}_2 is the closure of \mathcal{S} , by linearity, it suffices to show this for all simple functions $r \in \mathcal{S}$, that is $r(x) = (|x| - \gamma)_+$. By homogeneity, this is equivalent to showing that for all $\gamma \geq 0$ and $0 < a < 1$, we have

$$\begin{aligned} & \frac{(1+a)(1+a-\gamma)_+ - (1-a)(1-a-\gamma)_+}{2a} - \frac{(1+a)(1+a-\gamma)_+ + (1-a)(1-a-\gamma)_+}{2} \\ & \geq (1-\gamma)_+ - (a-\gamma)_+. \end{aligned}$$

Fix $0 < a < 1$. Let $h_a(\gamma)$ be the left hand side minus the right hand side. For $\gamma \geq 1+a$, $h_a(\gamma) = 0$. Since as a function of γ , $h_a(\gamma)$ is piecewise linear, showing that it is nonnegative on $[0, 1+a]$ is equivalent to verifying it at the nodes $\gamma \in \{0, 1, a, 1-a\}$. We have, $h_a(0) = a-a^2 > 0$. Next, $h_a(1) = \frac{(1+a)a}{2a} - \frac{(1+a)a}{2} = \frac{1}{2}(1+a)(1-a) > 0$. Finally, to check $\gamma = a$ and $\gamma = 1-a$, we consider two cases.

Case 1. $a \leq 1-a$, that is $0 < a \leq \frac{1}{2}$. Then,

$$h_a(a) = \frac{(1+a) - (1-a)(1-2a)}{2a} - \frac{(1+a) + (1-a)(1-2a)}{2} - (1-a) = a(1-a) > 0$$

and

$$h_a(1-a) = \frac{(1+a)2a}{2a} - \frac{(1+a)2a}{2} - a = 1 - a^2 - a \geq 1 - \frac{1}{4} - \frac{1}{2} = \frac{1}{4}.$$

Case 2. $a > 1-a$, that is $\frac{1}{2} < a < 1$. Then,

$$h_a(a) = \frac{(1+a)}{2a} - \frac{(1+a)}{2} - (1-a) = \frac{(1-a)^2}{2a} > 0$$

and

$$h_a(1-a) = \frac{(1+a)2a}{2a} - \frac{(1+a)2a}{2} - [a - (2a-1)] = a(1-a) > 0.$$

□

Proof of Theorem 6. Fix $p \geq 3$ and let $F(x) = |x|^p$. Then (8) is equivalent to saying that the function

$$\Phi(a_1, \dots, a_n) = \mathbb{E}F\left(\sum_{i=1}^n \sqrt{a_i}X_i\right)$$

is Schur concave. Since Φ is symmetric, by Ostrowski's criterion (see, e.g., Theorem II.3.14 in [6]), Φ is Schur concave if and only if

$$\frac{\partial \Phi}{\partial a_1} \geq \frac{\partial \Phi}{\partial a_2}, \quad a_1 < a_2,$$

which is equivalent to

$$\frac{1}{\sqrt{a_1}} \mathbb{E}[X_1 F'(S)] \geq \frac{1}{\sqrt{a_2}} \mathbb{E}[X_2 F'(S)],$$

where $S = \sqrt{a_1}X_1 + \sqrt{a_2}X_2 + W$ and $W = \sum_{i>2} \sqrt{a_i}X_i$. After taking the expectation with respect to X_1 and X_2 , it becomes

$$\begin{aligned} & \frac{1}{\sqrt{a_1}} \left(\frac{1-\rho_0}{2} \rho_0 \mathbb{E}[F'(\sqrt{a_1} + W) - F'(-\sqrt{a_1} + W)] \right. \\ & \quad + \left(\frac{1-\rho_0}{2} \right)^2 \mathbb{E}[F'(\sqrt{a_1} + \sqrt{a_2} + W) - F'(-\sqrt{a_1} + \sqrt{a_2} + W) \\ & \quad \quad \quad \left. + F'(\sqrt{a_1} - \sqrt{a_2} + W) - F'(-\sqrt{a_1} - \sqrt{a_2} + W)] \right) \\ & \geq \frac{1}{\sqrt{a_2}} \left(\frac{1-\rho_0}{2} \rho_0 \mathbb{E}[F'(\sqrt{a_2} + W) - F'(-\sqrt{a_2} + W)] \right. \\ & \quad + \left(\frac{1-\rho_0}{2} \right)^2 \mathbb{E}[F'(\sqrt{a_2} + \sqrt{a_1} + W) - F'(-\sqrt{a_2} + \sqrt{a_1} + W) \\ & \quad \quad \quad \left. + F'(\sqrt{a_2} - \sqrt{a_1} + W) - F'(-\sqrt{a_2} - \sqrt{a_1} + W)] \right). \end{aligned}$$

This trivially holds for $\rho_0 = 1$. Suppose $\rho_0 < 1$. Note that F' is odd and W is symmetric. Thus, $-\mathbb{E}F'(-\sqrt{a_1} + W) = \mathbb{E}F'(\sqrt{a_1} + W)$ and similarly for the other terms. Consequently, the inequality is equivalent to

$$\begin{aligned} & \frac{1}{\sqrt{a_1}} \left(2\rho_0 \mathbb{E}F'(\sqrt{a_1} + W) \right. \\ & \quad \left. + (1-\rho_0) \mathbb{E}[F'(\sqrt{a_1} + \sqrt{a_2} + W) - F'(-\sqrt{a_1} + \sqrt{a_2} + W)] \right) \\ & \geq \frac{1}{\sqrt{a_2}} \left(2\rho_0 \mathbb{E}F'(\sqrt{a_2} + W) \right. \\ & \quad \left. + (1-\rho_0) \mathbb{E}[F'(\sqrt{a_2} + \sqrt{a_1} + W) + F'(\sqrt{a_2} - \sqrt{a_1} + W)] \right). \end{aligned}$$

Set $a = \sqrt{a_1}$, $b = \sqrt{a_2}$ and

$$\phi(x) = \mathbb{E}F'(x + W), \quad x \in \mathbb{R}$$

(ϕ is also odd). Suppose $\rho_0 > 0$. Then, the validity of the above inequality is equivalent to the question whether for all $0 < a < b$,

$$(\rho_0^{-1} - 1) \left[\frac{\phi(a+b) - \phi(b-a)}{2a} - \frac{\phi(a+b) + \phi(b-a)}{2b} \right] \geq \left[\frac{\phi(b)}{b} - \frac{\phi(a)}{a} \right]. \quad (22)$$

By the symmetry of W , it has the same distribution as $\varepsilon|W|$, where ε is an independent symmetric random sign, so we can write $\phi(x) = \frac{1}{2} \mathbb{E}\phi_{|W|}(x)$, where for $w \geq 0$, we set $\phi_w(x) = F'(x+w) + F'(x-w)$. By Lemmas 18 and 19, inequality (22) holds for ϕ_w in place of ϕ (for every $w \geq 0$) as long as $\rho_0^{-1} - 1 \geq 1$. Taking the expectation against $|W|$ yields the inequality for ϕ , as desired. For $\rho_0 = 0$, we can for instance argue by taking the limit $\rho_0 \rightarrow 0+$ directly in (8). \square

3.4 First and second moments: Proof of Theorem 8

Note that for $a_1 = 1, a_2 = \dots = a_n = 0$, we have equality in (10), which explains why the value of the constant c_1 is sharp.

We shall closely follow Haagerup's approach from [13]. Let $\phi_X(t) = \mathbb{E}e^{itX}$ be the characteristic function of X . We have

$$\begin{aligned}\phi_X(t) &= \rho_0 + (1 - \rho_0) \frac{1}{L} \sum_{k=1}^L \cos(kt) \\ &\geq \rho_0 - (1 - \rho_0) = 2\rho_0 - 1 \geq 0.\end{aligned}$$

We also define

$$F(s) = \frac{2}{\pi} \int_0^\infty \left[1 - \left| \phi_X \left(\frac{t}{\sqrt{s}} \right) \right|^s \right] \frac{dt}{t^2}, \quad s \geq 1.$$

By symmetry, without loss of generality we can assume that a_1, \dots, a_n are positive with $\sum a_j^2 = 1$. By Lemma 1.2 from [13] and independence,

$$\mathbb{E} \left| \sum_j a_j X_j \right| = \frac{2}{\pi} \int_0^\infty \left[1 - \prod_j \phi_X(a_j t) \right] \frac{dt}{t^2}.$$

As in the proof of Lemma 1.3 from [13], by the AM-GM inequality,

$$\prod \phi_X(a_j t) \leq \sum a_j^2 |\phi_X(a_j t)|^{a_j^{-2}},$$

thus

$$\mathbb{E} \left| \sum_j a_j X_j \right| \geq \sum_j a_j^2 F(a_j^{-2}).$$

If we show that

$$F(s) \geq F(1), \quad s \geq 1, \tag{23}$$

then

$$\mathbb{E} \left| \sum_j a_j X_j \right| \geq \sum_j a_j^2 F(1) = F(1) = \frac{F(1)}{\sqrt{\mathbb{E}|X|^2}} \left(\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^2 \right)^{1/2}.$$

Since ϕ_X is nonnegative, using again Lemma 1.2 from [13], we have

$$F(1) = \frac{2}{\pi} \int_0^\infty [1 - |\phi_X(t)|] \frac{dt}{t^2} = \frac{2}{\pi} \int_0^\infty [1 - \phi_X(t)] \frac{dt}{t^2} = \mathbb{E}|X|,$$

so the proof of (10) is finished.

It remains to show (23). For a fixed $s \geq 1$, the left hand side

$$F(s) = \frac{2}{\pi} \int_0^\infty \left[1 - \left| \rho_0 + (1 - \rho_0) \frac{1}{L} \sum_{k=1}^L \cos \left(\frac{kt}{\sqrt{s}} \right) \right|^s \right] \frac{dt}{t^2}$$

is concave as a function of ρ_0 , whereas the right hand side $F(1) = \mathbb{E}|X| = (1 - \rho_0) \frac{L+1}{2}$ is linear as a function of ρ_0 . Therefore, it is enough to check the cases: 1) $\rho_0 = 1$ which is clear, 2) $\rho_0 = 1/2$ which becomes

$$\frac{2}{\pi} \int_0^\infty \left[1 - \left| \frac{1}{2} + \frac{1}{2} \frac{1}{L} \sum_{k=1}^L \cos \left(\frac{kt}{\sqrt{s}} \right) \right|^s \right] \frac{dt}{t^2} \geq \frac{L+1}{4}.$$

Using $\frac{\cos x + 1}{2} = \cos^2(x/2)$ and then employing convexity, the left hand side can be rewritten and lower bounded as follows

$$\frac{2}{\pi} \int_0^\infty \left[1 - \left| \frac{1}{L} \sum_{k=1}^L \cos^2 \left(\frac{kt}{2\sqrt{s}} \right) \right|^s \right] \frac{dt}{t^2} \geq \frac{1}{L} \sum_{k=1}^L \frac{2}{\pi} \int_0^\infty \left[1 - \left| \cos \left(\frac{kt}{2\sqrt{s}} \right) \right|^{2s} \right] \frac{dt}{t^2}.$$

A change of variables $t = \sqrt{2}t'/k$ allows to write the right hand side as

$$\frac{1}{L} \sum_{k=1}^L \frac{2}{\pi} \int_0^\infty \left[1 - \left| \cos \left(\frac{t'}{\sqrt{2s}} \right) \right|^{2s} \right] \frac{dt'}{t'^2} \frac{k}{\sqrt{2}} = \frac{L+1}{2\sqrt{2}} F_{\text{Haa}}(2s),$$

where $F_{\text{Haa}}(s) = \frac{2}{\pi} \int_0^\infty \left[1 - \left| \cos \left(\frac{t}{\sqrt{s}} \right) \right|^s \right] \frac{dt}{t^2}$ is Haagerup's function (see Lemma 1.3 and 1.4 in [13]). He showed therein that it is increasing, so for $s \geq 1$, we get $F_{\text{Haa}}(2s) \geq F_{\text{Haa}}(2) = \frac{1}{\sqrt{2}}$ and this finishes the proof.

Remark 20. Thanks to Remark 2.5 from [13], the same proof also works if we replace the first moment by p_0 -th one, where $p_0 = 1.847\dots$ is the unique solution to $\Gamma(\frac{p+1}{2}) = \frac{\sqrt{\pi}}{2}$, $p \in (0, 2)$. The cases of other values of $p \in (1, 2)$ have been elusive.

References

- [1] Averkamp, R., Houdré, C., Wavelet thresholding for non-necessarily Gaussian noise: Idealism. *Ann. Statist.* 31 (2003), 110–151.
- [2] Baernstein, A., II, Culverhouse, Robert C., Majorization of sequences, sharp vector Khinchin inequalities, and bisubharmonic functions. *Studia Math.* 152 (2002), no. 3, 231–248.
- [3] Ball, K., Mahler's conjecture and wavelets. *Discrete Comput. Geom.* 13 (1995), no. 3-4, 271–277.
- [4] Barthe, F., Guédon, O., Mendelson, S., Naor, A., A probabilistic approach to the geometry of the ℓ_p^n -ball. *Ann. Probab.* 33 (2005), no. 2, 480–513.
- [5] Barthe, F., Naor, A., Hyperplane projections of the unit ball of ℓ_p^n . *Discrete Comput. Geom.* 27 (2002), no. 2, 215–226.
- [6] Bhatia, R., Matrix analysis. Graduate Texts in Mathematics, 169. *Springer-Verlag, New York*, 1997.
- [7] Borell, C., Complements of Lyapunov's inequality. *Math. Ann.* 205 (1973), 323–331.
- [8] Cohn, J. H. E., Some integral inequalities. *Quart. J. Math. Oxford Ser. (2)* 20 (1969), 347–349.
- [9] Eaton, M. L., A note on symmetric Bernoulli random variables. *Ann. Math. Statist.* 41 (1970), 1223–1226.

- [10] Eskenazis, A., Nayar, P., Tkocz, T., Gaussian mixtures: entropy and geometric inequalities, *Ann. of Prob.* 46(5) 2018, 2908–2945.
- [11] Eskenazis, A., Nayar, P., Tkocz, T., Sharp comparison of moments and the log-concave moment problem, *Adv. Math.* 334 (2018) 389–416.
- [12] Figiel, T., Hitczenko, P., Johnson, W. B., Schechtman, G., Zinn, J., Extremal properties of Rademacher functions with applications to the Khintchine and Rosenthal inequalities. *Trans. Amer. Math. Soc.* 349 (1997), no. 3, 997–1027.
- [13] Haagerup, U., The best constants in the Khintchine inequality. *Studia Math.* 70 (1981), no. 3, 231–283.
- [14] Hardy, G. H., Littlewood, J. E., Pólya, G., Inequalities. 2nd ed. *Cambridge, at the University Press*, 1952.
- [15] Khintchine, A., Über dyadische Brüche. *Math. Z.* 18 (1923), no. 1, 109–116.
- [16] Komorowski, R., On the best possible constants in the Khintchine inequality for $p \geq 3$. *Bull. London Math. Soc.* 20 (1988), no. 1, 73–75.
- [17] König, H., On the best constants in the Khintchine inequality for Steinhaus variables. *Israel J. Math.* 203 (2014), no. 1, 23–57.
- [18] König, H., Kwapien, S., Best Khintchine type inequalities for sums of independent, rotationally invariant random vectors. *Positivity* 5 (2001), no. 2, 115–152.
- [19] Kwapien, S., Latała, R., Oleszkiewicz, K., Comparison of moments of sums of independent random variables and differential inequalities. *J. Funct. Anal.* 136 (1996), no. 1, 258–268.
- [20] Latała, R., Estimation of moments of sums of independent real random variables. *Ann. Probab.*, 25(3):1502–1513, 1997.
- [21] Latała, R., Oleszkiewicz, K., On the best constant in the Khinchin-Kahane inequality. *Studia Math.* 109 (1994), no. 1, 101–104.
- [22] Latała, R., Oleszkiewicz, K., A note on sums of independent uniformly distributed random variables. *Colloq. Math.* 68 (1995), no. 2, 197–206.
- [23] Ledoux, M., Talagrand, M., Probability in Banach spaces. Isoperimetry and processes. *Springer-Verlag*, Berlin, 1991.
- [24] Littlewood, J. E., On a certain bilinear form, *Quart. J. Math. Oxford Ser. 1* (1930), 164–174.
- [25] Melbourne, J., personal communication (Oct, 2019).
- [26] Milman, V., Schechtman, G., Asymptotic theory of finite-dimensional normed spaces. With an appendix by M. Gromov. Lecture Notes in Mathematics, 1200. *Springer-Verlag*, Berlin, 1986.

- [27] Mordhorst, O., The optimal constants in Khintchine’s inequality for the case $2 < p < 3$. *Colloq. Math.* 147 (2017), no. 2, 203–216.
- [28] Nayar, P., Oleszkiewicz, K., Khinchine type inequalities with optimal constants via ultra log-concavity. *Positivity* 16 (2012), no. 2, 359–371.
- [29] Nazarov, F. L., Podkorytov, A. N., Ball, Haagerup, and distribution functions. *Complex analysis, operators, and related topics*, 247–267, *Oper. Theory Adv. Appl.*, 113, Birkhäuser, Basel, 2000.
- [30] Newman, C. M., An extension of Khintchine’s inequality. *Bull. Amer. Math. Soc.* 81 (1975), no. 5, 913–915.
- [31] Newman, C. M., Inequalities for Ising models and field theories which obey the Lee-Yang theorem. *Comm. Math. Phys.* 41 (1975), 1–9.
- [32] Oleszkiewicz, K., Comparison of moments via Poincaré-type inequality. *Advances in stochastic inequalities (Atlanta, GA, 1997)*, 135–148, *Contemp. Math.*, 234, *Amer. Math. Soc.*, Providence, RI, 1999.
- [33] Oleszkiewicz, K., Precise moment and tail bounds for Rademacher sums in terms of weak parameters. *Israel J. Math.* 203 (2014), no. 1, 429–443.
- [34] Pass, B., Spektor, S., On Khintchine type inequalities for k -wise independent Rademacher random variables. *Statist. Probab. Lett.* 132 (2018), 35–39.
- [35] Pinelis, I., Extremal probabilistic problems and Hotelling’s T^2 test under a symmetry condition. *Ann. Statist.* 22 (1994), no. 1, 357–368.
- [36] Spektor, S., Restricted Khinchine inequality. *Canad. Math. Bull.* 59 (2016), no. 1, 204–210.
- [37] Szarek, S., On the best constant in the Khintchine inequality. *Stud. Math.* 58, 197–208 (1976)
- [38] Tomaszewski, B., A simple and elementary proof of the Khintchine inequality with the best constant. *Bull. Sci. Math.* (2) 111 (1987), no. 1, 103–109.
- [39] Whittle, P., Bounds for the moments of linear and quadratic forms in independent random variables. *Theory Probab. Appl.* 5, 302–305 (1960)
- [40] Young, R. M. G., On the best possible constants in the Khintchine inequality. *J. London Math. Soc.* (2) 14 (1976), no. 3, 496–504.
- [41] <https://mathoverflow.net/questions/208349>