

SHARP KHINCHIN-TYPE INEQUALITIES FOR SYMMETRIC DISCRETE UNIFORM RANDOM VARIABLES*

BY

ALEX HAVRILLA

School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA
e-mail: ahavrilla3@gatech.edu

AND

TOMASZ TKOCZ

Department of Mathematical Sciences, Carnegie Mellon University
Pittsburgh, PA 15213, USA
e-mail: ttkocz@math.cmu.edu

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 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.

1. Introduction

The classical Khinchin inequality asserts that all moments of weighted sums of independent random signs are comparable (see [13]). 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

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

* Research supported in part by the Collaboration Grants from the Simons Foundation and NSF grant DMS-1955175.

Received January 2, 2020 and in revised form October 15, 2020

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 [18]. They have found numerous applications in classical results in analysis (for example in the proof of the Littlewood–Payley 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 [21], [23]). One of the major challenges is to find the best constants $C_{p,q}$, which has attracted considerable 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.

For results concerning the best constant $C_{p,q}$ in the classical Khinchin inequality (1), we mention in passing works [7, 11, 14, 19, 24, 25, 26, 30, 32, 34, 36, 37], highlighting only that the optimal value of $C_{p,q}$ is known when $p < q$ (trivial), either p or q is 2, or both p and q are even. 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, 15, 16, 20]), mixtures of centred Gaussians (see [1, 8]), the so-called exponential family given by the density $e^{-|x|^\alpha}$ and uniform distributions on unit ℓ_α^n balls (see [4, 8, 9]), dependent random signs (see [31, 33]), as well as general random variables via their spectral properties (see [17, 29]).

This paper concerns 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 the 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 [20].

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

$$\|G\|_p = 2^{1/2}(\pi^{-1/2}\Gamma((p+1)/2))^{1/p},$$

where Γ stands for the gamma function.

ACKNOWLEDGEMENTS. We are indebted to Krzysztof Oleszkiewicz for his help and valuable feedback.

2. Results

Given $\rho_0 \in [0, 1]$ and a positive integer L , consider a random variable X with

$$(2) \quad \mathbb{P}(X=0)=\rho_0 \quad \text{and} \quad \mathbb{P}(X=j)=\mathbb{P}(X=-j)=\frac{1-\rho_0}{2L} \quad \text{for } j=1, \dots, L.$$

For $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ and $p \geq 1$, we let

$$(3) \quad N_p(a) = \left\| \sum_{i=1}^n a_i X_i \right\|_p,$$

where $X_1, X_2, \dots$ are i.i.d. copies of X . Throughout, G stands for a standard Gaussian random variable. We refer to the classical monograph [12], or to [6] for a concise exposition of majorisation and Schur-convexity. Our main results are as follows.

THEOREM 1: *Let $\rho_0 \in [0, \frac{1}{2}]$ and $L = 1$. If $p \geq 3$, then the function*

$$(a_1, \dots, a_n) \mapsto N_p(\sqrt{a_1}, \dots, \sqrt{a_n})$$

is Schur-concave on $[0, +\infty)^n$.

As an immediate corollary, we obtain the best constants in Khinchin inequalities (it can be done as, for instance, in the proof of Corollary 25 from [8]).

COROLLARY 2: *Under the assumptions of Theorem 1, the best constant C_p such that the inequality $N_p(a) \leq C_p N_2(a)$ holds for all n and $a \in \mathbb{R}^n$ is $C_p = \|G\|_p$.*

Our next result concerns Khinchin inequalities for $p \geq 3$ for arbitrary L and $\rho_0 = 0$.

THEOREM 3: *Let $\rho_0 = 0$ and $L \geq 1$ be an integer. If $p \geq 3$, then the best constant C_p such that the inequality $N_p(a) \leq C_p N_2(a)$ holds for all n and $a \in \mathbb{R}^n$ is $C_p = \|G\|_p$.*

Finally, in the presence of large mass at 0 and arbitrarily many atoms L , we obtain a sharp $L_1 - L_2$ inequality, which holds in a greater generality.

THEOREM 4: *Let $\rho_0 \in [\frac{1}{2}, 1)$ and let $Y, Y_1, Y_2, \dots$ be i.i.d. symmetric random variables with $\mathbb{P}(Y = 0) = \rho_0$. Define $N_p(a) = \|\sum_{i=1}^n a_i Y_i\|_p$. The best constant c_1 such that the inequality $N_1(a) \geq c_1 N_2(a)$ holds for all n and $a \in \mathbb{R}^n$ is $c_1 = \|Y\|_1 / \|Y\|_2$.*

Some restrictions on ρ_0 in our theorems are needed, however our specific ones may not be optimal. We defer a discussion to the last section.

Remark 5: When p is a positive even integer, Theorem 3 can be deduced from Newman's results from [27] (see also [28]).

Remark 6: Using [25], Theorem 3 can be extended to a sharp moment comparison between all even moments with an *optimal* restriction on ρ_0 , which will appear elsewhere.

Remark 7: When $L = 1$ and $Y = X$, Theorem 4 follows from general results of Oleszkiewicz from [29] concerning arbitrary symmetric random variables and coefficients in Banach space (see Corollary 2.4 therein).

We finish this section with a few words on proofs. Our proof of Theorem 1 follows a direct approach from Eaton's work [7], combined with techniques (used, for instance, in [10], or [9]) exploiting linearity and allowing to reduce verification of certain inequalities needed for averages of power functions $|\cdot|^p$ to simple (piecewise linear) functions. To prove Theorem 3, we employ an inductive argument (on n) which crucially uses independence and convexity of certain functions and is based on swapping the X_i one by one with independent Gaussians. For Theorem 4, we extend Haagerup's short proof from [11] of Szarek's result from [34] saying that the best constant $C_{2,1}$ in (1) is $\sqrt{2}$ (for the latter, see also [19, 22, 35]). We rely on an integral representation for the first moment used by Haagerup, combined with convexity arguments allowing to handle more atoms.

3. Proofs

3.1. A SCHUR-CONVEXITY RESULT: PROOF OF THEOREM 1. We 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 8: Let $q \geq 2$, $w \geq 0$ and

$$\phi_w(x) = \operatorname{sgn}(x+w)|x+w|^q + \operatorname{sgn}(x-w)|x-w|^q, \quad x \in \mathbb{R}.$$

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

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

Taking one more derivative gives

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

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 $\phi(x) = (1+x)^q - (1-x)^q$, we find that

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

If $q > 2$, this is positive for $0 < x < 1$. Then in this case, consequently,

$$x^3 r''(x) > x^3 r''(x)|_{x=0} = 0,$$

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. ■

LEMMA 9: *The best constant D such that the inequality*

$$(4) \quad D \cdot \left[\frac{\phi(a+b) - \phi(b-a)}{2a} - \frac{\phi(a+b) + \phi(b-a)}{2b} \right] \geq \left[\frac{\phi(b)}{b} - \frac{\phi(a)}{a} \right]$$

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 (4) 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 (4) 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

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

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

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

Proof of Theorem 1. Fix $p \geq 3$ and let $F(x) = |x|^p$. We would like to show 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 $\frac{\partial \Phi}{\partial a_1} \geq \frac{\partial \Phi}{\partial a_2}$, $a_1 < a_2$, 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 $W = \sum_{i>2} \sqrt{a_i}X_i$. We take the expectation with respect to X_1 and X_2 . Suppose $\rho_0 < 1$. Since F' is odd and W is symmetric, we get,

$$-\mathbb{E}F'(-\sqrt{a_1} + W) = \mathbb{E}F'(\sqrt{a_1} + W)$$

and similarly for the other terms that show up. Consequently, the inequality can be equivalently rewritten as

$$\begin{aligned} & 3 \frac{1}{\sqrt{a_1}} (2\rho_0 \mathbb{E} F'(\sqrt{a_1} + W) + (1 - \rho_0) \mathbb{E} [F'(\sqrt{a_1} + \sqrt{a_2} + W) \\ & \quad - F'(-\sqrt{a_1} + \sqrt{a_2} + W)]) \\ & \geq \frac{1}{\sqrt{a_2}} (2\rho_0 \mathbb{E} F'(\sqrt{a_2} + W) + (1 - \rho_0) \mathbb{E} [F'(\sqrt{a_2} + \sqrt{a_1} + W) \\ & \quad + F'(\sqrt{a_2} - \sqrt{a_1} + W)]). \end{aligned}$$

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 $0 < a < b$,

$$(5) \quad (\rho_0^{-1} - 1) \left[\frac{\phi(a+b) - \phi(b-a)}{2a} - \frac{\phi(a+b) + \phi(b-a)}{2b} \right] \geq \left[\frac{\phi(b)}{b} - \frac{\phi(a)}{a} \right].$$

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 8 and 9, inequality (5) 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 $\rho_0 \rightarrow 0+$. ■

3.2. SECOND, THIRD AND HIGHER MOMENTS: PROOF OF THEOREM 3. The value of C_p 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 $N_p(a) \leq C_p N_2(a)$, we set

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

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 $N_p(a) \leq C_p N_2(a)$ 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 all reals a, b , we have

$$(7) \quad \mathbb{E}|a + bX_1|^p \leq \mathbb{E}|a + bG_1|^p.$$

This will follow from the following claim.

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

$$(8) \quad \mathbb{E}h(X_1^2) \leq \mathbb{E}h(G_1^2).$$

Indeed, (7) for $b = 0$ is clear. Assuming $b \neq 0$, by homogeneity, (7) 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

$$(9) \quad h_a(x) = |a + \sqrt{x}|^p + |a - \sqrt{x}|^p, \quad x \geq 0$$

(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 [10]).

LEMMA 10: Let $p \geq 3$, $a \in \mathbb{R}$. Then h_a defined in (9) 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

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

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 $(0, \infty)$. 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 $(0, \infty)$. $\blacksquare$

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 (6). 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

$$\begin{aligned}
 f'(a) &= -\sqrt{\frac{2}{\pi\sigma^2}} \int_{\sqrt{a}}^{\infty} e^{-x^2/2\sigma^2} dx - \frac{1}{L} \sum_{k=b+1}^L (-1) \\
 &= -\sqrt{\frac{2}{\pi\sigma^2}} \int_{\sqrt{a}}^{\infty} e^{-x^2/2\sigma^2} dx + \frac{L-b}{L}.
 \end{aligned}
 \tag{10}$$

Bounding $b < \sqrt{a}$ yields

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

Let $\tilde{g}(a)$ denote the right-hand side. 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

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

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 (10)) 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. ■

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 11: We can drop the assumption in Theorem 3 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 (8), hence (7), holds for each X_i .

3.3. FIRST AND SECOND MOMENTS: PROOF OF THEOREM 4. For $a_1 = 1$, $a_2 = \dots = a_n = 0$, inequality $N_1(a) \geq c_1 N_2(a)$ becomes equality, so the value of the constant c_1 is sharp. To prove the inequality, we shall closely follow Haagerup's approach from [11]. Note that Y has the same distribution as $\theta \varepsilon R$, where θ is a Bernoulli random variable with parameter $1 - \rho_0$, ε is a symmetric random sign, R is a positive random variable and θ, ε and R are independent (the law of R is the same as the law of $|X|$ conditioned on $X \neq 0$). Let $\phi_Y(t) = \mathbb{E}e^{itY}$ be the characteristic function of Y . We have

$$\phi_Y(t) = \rho_0 + (1 - \rho_0)\mathbb{E} \cos(tR) \geq \rho_0 - (1 - \rho_0) = 2\rho_0 - 1 \geq 0.$$

We also define

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

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 [11] and independence,

$$N_1(a) = \mathbb{E} \left| \sum_j a_j Y_j \right| = \frac{2}{\pi} \int_0^\infty \left[1 - \prod_j \phi_Y(a_j t) \right] \frac{dt}{t^2}.$$

By the AM–GM inequality,

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

thus $N_1(a) \geq \sum_j a_j^2 F(a_j^{-2})$. If we show that

$$(11) \quad F(s) \geq F(1), \quad s \geq 1,$$

then

$$N_1(a) \geq \sum_j a_j^2 F(1) = F(1) = \frac{F(1)}{\sqrt{\mathbb{E}|Y|^2}} N_2(a).$$

Since ϕ_Y is nonnegative, using again Lemma 1.2 from [11], we have

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

so the proof of $N_1(a) \geq c_1 N_2(a)$ with $c_1 = \|Y\|_1 / \|Y\|_2$ is finished.

It remains to show (11). For a fixed $s \geq 1$, the left-hand side

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

is concave as a function of ρ_0 , whereas the right-hand side

$$F(1) = \mathbb{E}|Y| = (1 - \rho_0) \mathbb{E}R$$

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

$$\frac{2}{\pi} \int_0^\infty \left[1 - \left| \frac{1}{2} + \frac{1}{2} \mathbb{E} \cos \left(\frac{tR}{\sqrt{s}} \right) \right|^s \right] \frac{dt}{t^2} \geq \frac{1}{2} \mathbb{E}R.$$

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:

$$\frac{2}{\pi} \int_0^\infty \left[1 - \left| \mathbb{E} \cos^2 \left(\frac{tR}{2\sqrt{s}} \right) \right|^s \right] \frac{dt}{t^2} \geq \mathbb{E} \frac{2}{\pi} \int_0^\infty \left[1 - \left| \cos \left(\frac{tR}{2\sqrt{s}} \right) \right|^{2s} \right] \frac{dt}{t^2}.$$

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

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

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 Lemmas 1.3 and 1.4 in [11]). 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 12: Thanks to Remark 2.5 from [11], the same proof also works if we replace the first moment by the 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.

4. Necessity of the restrictions on ρ_0

We use the notation from (2) and (3). We derive some necessary conditions on ρ_0 , justifying to some extent our restrictions on ρ_0 made in Theorems 1, 3 and 4.

Remark 13: For Theorem 1 to hold, we necessarily have $\frac{d}{d\lambda} N_3^3(\sqrt{\lambda}, \sqrt{1-\lambda}) \geq 0$ for $\lambda \in (0, \frac{1}{2})$. Letting $\lambda \rightarrow 0+$ yields $(1 - \rho_0)(1 - 2\rho_0) \geq 0$, hence $\rho_0 \leq \frac{1}{2}$.

Remark 14: In Theorem 3, a necessary condition on ρ_0 is

$$\rho_0 \leq 1 - \frac{27\pi}{128} = 0.33732\dots$$

This follows from $N_3(1) \leq \|G\|_3 N_2(1)$ with $L \rightarrow \infty$.

Remark 15: In Theorem 4, a necessary condition on ρ_0 is $\rho_0 \geq \sqrt{2} - 1$. This follows from $N_1(1, 1) \geq c_1 N_2(1, 1)$ applied to $Y = X$, which is equivalent to

$$(1 - \rho_0) \frac{2L + 1}{3L} \geq 2 - \sqrt{2},$$

so $L = 1$ gives $\rho_0 \geq \sqrt{2} - 1$.

Thus the restriction in Theorem 1 is sharp, while those in Theorems 3 and 4 are by-products of our proofs and can perhaps be improved. We believe the optimal ones are indicated above (for the following reasons: one can check that the case $n = 2$ and $Y = X$ of Theorem 4 holds for $\rho_0 \in [\sqrt{2} - 1, 1)$; moreover, in the context of Theorem 3, $N_4(1) \leq \|G\|_4 N_2(1)$ with $L \rightarrow \infty$ is a sufficient condition for $N_q(a) \leq \frac{\|G\|_q}{\|G\|_p} N_p(a)$ to hold for all $2 \leq p < q$ even integers—see Remark 6).

References

- [1] R. Averkamp and C. Houdré, *Wavelet thresholding for non-necessarily Gaussian noise: Idealism*, *Annals of Statistics* **31** (2003), 110–151.
- [2] A. Baernstein, II and R. C. Culverhouse, *Majorization of sequences, sharp vector Khinchin inequalities, and bisubharmonic functions*, *Studia Mathematica* **152** (2002), 231–248.
- [3] K. Ball, *Mahler's conjecture and wavelets*, *Discrete & Computational Geometry* **13** (1995), 271–277.
- [4] F. Barthe, O. Guédon, S. Mendelson and A. Naor, *A probabilistic approach to the geometry of the ℓ_p^n -ball*, *Annals of Probability* **33** (2005), 480–513.
- [5] F. Barthe and A. Naor, *Hyperplane projections of the unit ball of ℓ_p^n* , *Discrete & Computational Geometry* **27** (2002), 215–226.
- [6] R. Bhatia, *Matrix Analysis*, Graduate Texts in Mathematics, Vol. 169, Springer, New York, 1997.
- [7] M. L. Eaton, *A note on symmetric Bernoulli random variables*, *Annals of Mathematical Statistics* **41** (1970), 1223–1226.
- [8] A. Eskenazis, P. Nayar and T. Tkocz, *Gaussian mixtures: entropy and geometric inequalities*, *Annals of Probability* **46** (2018), 2908–2945.
- [9] A. Eskenazis, P. Nayar and T. Tkocz, *Sharp comparison of moments and the log-concave moment problem*, *Advances in Mathematics* **334** (2018), 389–416.

- [10] T. Figiel, P. Hitzchenko, W. B. Johnson, G. Schechtman and J. Zinn, *Extremal properties of Rademacher functions with applications to the Khintchine and Rosenthal inequalities*, Transactions of the American Mathematical Society **349** (1997), 997–1027.
- [11] U. Haagerup, *The best constants in the Khintchine inequality*, Studia Mathematica **70** (1981), 231–283.
- [12] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge University Press, Cambridge, 1952.
- [13] A. Khintchine, *Über dyadische Brüche*, Mathematische Zeitschrift **18** (1923), 109–116.
- [14] R. Komorowski, *On the best possible constants in the Khintchine inequality for $p \geq 3$* , Bulletin of the London Mathematical Society **20** (1988), 73–75.
- [15] H. König, *On the best constants in the Khintchine inequality for Steinhaus variables*, Israel Journal of Mathematics **203** (2014), 23–57.
- [16] H. König and S. Kwapien, *Best Khintchine type inequalities for sums of independent, rotationally invariant random vectors*, Positivity **5** (2001), 115–152.
- [17] S. Kwapien, R. Latała and K. Oleszkiewicz, *Comparison of moments of sums of independent random variables and differential inequalities*, Journal of Functional Analysis **136** (1996), 258–268.
- [18] R. Latała, *Estimation of moments of sums of independent real random variables*, Annals of Probability **25** (1997), 1502–1513.
- [19] R. Latała and K. Oleszkiewicz, *On the best constant in the Khinchin-Kahane inequality*, Studia Mathematica **109** (1994), 101–104.
- [20] R. Latała and K. Oleszkiewicz, *A note on sums of independent uniformly distributed random variables* Colloquium Mathematicum **68** (1995), 197–206.
- [21] M. Ledoux and M. Talagrand, *Probability in Banach Spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. 23, Springer, Berlin, 1991.
- [22] J. E. Littlewood, *On a certain bilinear form*, Quarterly Journal of Mathematics **1** (1930), 164–174.
- [23] V. Milman and G. Schechtman, *Asymptotic Theory of Finite-dimensional Normed Spaces*, Lecture Notes in Mathematics, Vol. 1200, Springer, Berlin, 1986.
- [24] O. Mordhorst, *The optimal constants in Khintchine’s inequality for the case $2 < p < 3$* , Colloquium Mathematicum **147** (2017), 203–216.
- [25] P. Nayar and K. Oleszkiewicz, *Khinchine type inequalities with optimal constants via ultra log-concavity*, Positivity **16** (2012), 359–371.
- [26] F. L. Nazarov and A. N. Podkorytov, *Ball, Haagerup, and distribution functions*, in *Complex Analysis, Operators, and Related Topics*, Operator Theory: Advances and Applications, Vol. 113, Birkhäuser, Basel, 2000, pp. 247–267.
- [27] C. M. Newman, *An extension of Khintchine’s inequality*, Bulletin of the American Mathematical Society **81** (1975), 913–915.
- [28] C. M. Newman, *Inequalities for Ising models and field theories which obey the Lee–Yang theorem*, Communications in Mathematical Physics **41** (1975), 1–9.
- [29] K. Oleszkiewicz, *Comparison of moments via Poincaré-type inequality*, in *Advances in Stochastic Inequalities (Atlanta, GA, 1997)*, Contemporary Mathematics, Vol. 234, American Mathematical Society, Providence, RI, 1999, pp. 135–148.

- [30] K. Oleszkiewicz, *Precise moment and tail bounds for Rademacher sums in terms of weak parameters*, Israel Journal of Mathematics **203** (2014), 429–443.
- [31] B. Pass and S. Spektor, *On Khintchine type inequalities for k -wise independent Rademacher random variables*, Statistics & Probability Letters **132** (2018), 35–39.
- [32] I. Pinelis, *Extremal probabilistic problems and Hotelling's T^2 test under a symmetry condition*, Annals of Statistics **22** (1994), 357–368.
- [33] S. Spektor, *Restricted Khinchine inequality*, Canadian Mathematical Bulletin **59** (2016), 204–210.
- [34] S. Szarek, *On the best constant in the Khintchine inequality*, Studia Mathematica **58** (1976), 197–208.
- [35] B. Tomaszewski, *A simple and elementary proof of the Khintchine inequality with the best constant*, Bulletin des Sciences Mathématiques **111** (1987), 103–109.
- [36] P. Whittle, *Bounds for the moments of linear and quadratic forms in independent random variables*, Theory of Probability and its Applications **5** (1960), 302–305.
- [37] R. M. G. Young, *On the best possible constants in the Khintchine inequality*, Journal of the London Mathematical Society **14** (1976), 496–504.