

A Survey of Khintchine Type Inequalities for Random Variables

Alex Havrilla

March 22, 2021

Abstract

We present a collection of Khintchine type inequalities for random signs, uniform distributions, gaussian mixtures, and the ultra sub-Gaussian class. We also present new results for symmetric discrete distributions and the class of Type \mathcal{L} random variables.

Contents

1	Introduction	3
1.1	Notation	3
1.2	Khintchine Inequalities	3
2	Khintchine Inequalities for Various Distributions	5
2.1	Random Signs - The Classical Story	5
2.2	Uniform Random Variables	8
2.3	Ultra Sub-Gaussian Random Variables	10
2.4	Gaussian Mixtures	12
2.5	Random Vectors Uniform on \mathbb{S}^{n-1}	13
3	Type \mathcal{L} Random Variables	16
3.1	Connections to Ultra Sub-Gaussianity	16
3.2	Type \mathcal{L} Random Variables with "Enough Gaussanity"	18
3.2.1	Examples of Type \mathcal{L} Random Variables	20
4	Discrete Symmetric Distributions	21
4.1	Connections to Ultra Sub-Gaussianity	21
4.2	Small mass at 0	21
4.3	Large mass at 0	28

1 Introduction

1.1 Notation

\mathbb{E} denotes expectation. ϵ will refer to a random sign uniformly distributed on $\{-1, 1\}$. $\|X\|_p$ denotes the p th moment of random variable X i.e. $\|X\|_p = (\mathbb{E}|X|^p)^{1/p}$. $a = (a_i)_{i=1}^\infty$ will be a sequence. g will be used as a normal random variable. B_q^n is the n -dimensional unit ball in the q -norm. \mathcal{L} will denote the class of type \mathcal{L} random variables. \mathbb{L} will refer to class of random variables uniformly distributed on $[L] \cup -[L]$ for $L \in \mathbb{Z}^+$ with some mass at 0 which we denote ρ_0 .

1.2 Khintchine Inequalities

The study of Khintchine inequalities, initiated by Aleksandr Khinchin in [19], seeks to find constants $C_{p,q}$ comparing p th and q th moments of linear combinations of independent identically distributed (i.i.d) random variables.

The classical Khintchine inequality compares for $0 < p, q < \infty$ the p th and q th moments of sums of iid random signs with coefficients $a_i \in \mathbb{R}$

$$S = \sum_{i=1}^n a_i \epsilon_i$$

where we recall $\mathbb{P}(\epsilon = 1) = \mathbb{P}(\epsilon = -1) = \frac{1}{2}$.

Theorem 1. *Khintchine's Inequalities:*

$\forall p, q > 0 \exists C_{p,q} > 0$ and $\forall n, a \in \mathbb{R}^n$ we have

$$\|S\|_p \leq C_{p,q} \|S\|_q$$

Of particular interest is finding sharp constants $C_{p,q}$. This is easy when $p < q$, as then $C_{p,q} = 1$. However the case $p > q$ is difficult. In general we only know comparisons for arbitrary p and $q = 2$ and for even p, q . A natural generalization of this inequality is to consider sums of random variables other than random signs. In some cases, particularly with type \mathcal{L} random variables, we are even able to relax the independence between terms in the sum.

We can summarize what is known about optimal constants for the distributions we present in the following table.

- Random Signs: Optimal $C_{2,p}$ and $C_{p,2}$ known for $p \in [0, 2]$ and $p \in [2, \infty)$. Optimal constants also known for even integer $C_{p,q}$ via ultra sub-Gaussianity.
- Uniform distributions on line: Optimal $C_{p,2}$ known for $p \in [2, \infty)$ and $C_{2,p}$ known for $p \in [1, 2]$.
- Uniform distributions on B_q^n for $q \in (0, 2]$: Optimal constants $C_{p,2}$ and $C_{2,p}$ for $p \in [2, \infty)$ and $p \in (-1, 2)$.
- Ultra sub-Gaussian class: Optimal $C_{p,q}$ for all even integer pairs p, q .
- Type \mathcal{L} class: Optimal $C_{p,q}$ for all even integer pairs p, q via ultra sub-Gaussianity. $C_{p,2}$ known for $p \geq 3$.

- The \mathbb{L} class: Depends on ρ_0 the mass at 0. All even p, q for small enough ρ_0 . $C_{p,2}$ for $p \geq 3$ when $\rho_0 = 0$. $C_{p,2}$ for $p \geq 3$ when $L = 1$ and ρ_0 small. $C_{1,2}$ for large ρ_0 .

This list is by no means comprehensive and we shall attempt to paint a more complete albeit shallower picture in the next section.

2 Khintchine Inequalities for Various Distributions

This section contains a collection of Khintchine-type results with optimal constants for variously distributed random variables. We start with the progenitor case of random signs.

2.1 Random Signs - The Classical Story

In this section we deal strictly with sums of random signs ϵ_i . First we establish constants $C_{p,2}$ and $C_{2,p}$ do exist for $\|S\|_p \leq C_{p,2}\|S\|_2$, $\|S\|_2 \leq C_{2,p}\|S\|_p$ where $S = \sum a_i \epsilon_i$. We argue as in [8]. This is a natural starting point since second moment of sums is easily computable:

$$\|S\|_2 = \sqrt{\mathbb{E}S_a^2} = \sqrt{\mathbb{E} \sum_i a_i^2 \epsilon_i + \mathbb{E} \sum_{i,j} a_i a_j \epsilon_i \epsilon_j} = \sqrt{\sum_i a_i^2}$$

ie. the second moment of the sum is just the sum of squares of the coefficients. Often without loss of generality we will choose a_i s.t. $\sum_i a_i^2 = 1$.

We adopt the convention $A_p = C_{2,p}$ and $B_p = C_{p,2}$. First we establish the existence of such constants.

Theorem 2. *Khinchin's Inequality:*

For $0 < p < \infty$, $\exists A_p, B_p > 0$ s.t. for $a \in \mathbb{R}^n$,

$$A_p \sqrt{\sum_i a_i^2} \leq \|S_a\|_p \leq B_p \sqrt{\sum_i a_i^2}$$

We argue as in [8].

Proof. Via homogeneity we may suppose $\sum_i a_i^2 = 1$. Then via layer cake and chebyshev we may write:

$$\mathbb{E} \left| \sum_i \epsilon_i a_i \right|^p = \int_0^\infty \mathbb{P} \left(\left| \sum_i \epsilon_i a_i \right| > t \right) dt^p \leq 2 \int_0^\infty e^{-t^2/2} dt^p = B_p^p$$

For the other side write

$$1 = \mathbb{E} \left(\sum_i \epsilon_i a_i \right)^2 \leq (\mathbb{E} \left| \sum_i \epsilon_i a_i \right|^p)^{2/3} (\mathbb{E} \left| \sum_i \epsilon_i a_i \right|^{6-2p})^{1/3} \leq (\mathbb{E} \left| \sum_i \epsilon_i a_i \right|^p)^{2/3} B_{6-2p}^{2-2p/3}$$

at which point we are done. \square

So we know such comparisons are possible but we seek to find optimal constants. The case for $p \geq 3$ turns out to be relatively easy via some convexity arguments by Young in [15]. Note this in fact first shown by Whittle in [20].

Theorem 3. *Young:*

For $3 \leq p < \infty$ we have $B_p = 2^{1/2}(\Gamma((p+1)/2)/\sqrt{\pi})^{1/p}$

The proof crucially relies on the convexity of a certain function h for $p \geq 3$ which inspires relevant ideas later:

Lemma 4. *Define for $z \in \mathbb{C}$*

$$h_p(z, t) = |1 - tz|^{p-2} + |1 + tz|^{p-2}$$

For $p \geq 3$ h_p is convex in t .

In [2] Haagerup concluded the search for random signs by treating $0 < p < 2$ and $2 < p < 3$ and which establishes the following constants:

$$A_p = \begin{cases} 2^{1/2-1/p} & 0 < p \leq p_0 \\ 2^{1/2}(\Gamma((p+1)/2)/\sqrt{\pi})^{1/p} & p_0 < p < 2 \\ 1 & 2 \leq p < \infty \end{cases}$$

and

$$B_p = \begin{cases} 1 & 0 < p \leq 2 \\ 2^{1/2}(\Gamma((p+1)/2)/\sqrt{\pi})^{1/p} & 2 < p < \infty \end{cases}$$

where p_0 is the solution to $\Gamma((p+1)/2) = \sqrt{\pi}/2$ in $[1, 2]$. Szarek [14] showed $A_1 = 1/\sqrt{2}$. Haagerup extended this result for $0 < p < p_0$ and established optimality of B_p for $2 < p < 3$. Note this is seen to be sharp by way of Central Limit Theorem, as the sum $\sum_i \frac{1}{\sqrt{n}} \epsilon_i \rightarrow g \sim N(0, 1)$ which achieves the B_p in the limit. Note the $1/\sqrt{2}$ is achieved by the sum of two random signs.

While it would be nice to give complete proofs of these sharp constants, Haagerups arguments in particular are quite technical and not entirely enlightening. Instead we elect to present a proof via Nazarov, F. Podkorytov [10] which relies on crucial elements of Haagerups proof while using distribution functions to "simplify" some more technical details. This treats the $0 < p < 2$ case. In [9] Mordhorst treats $2 < p < 3$ using the same technique.

In preparation consider a lemma on distribution functions:

Lemma 5. *Let $Y > 0$, $f, g : \mathcal{M} \rightarrow [0, Y]$ be any two measurable functions on (\mathcal{M}, μ) . Let F_* and G_* be their modified distribution functions. Assume both $F_*(y)$ and $G_*(y)$ are finite for every $y \in (0, Y)$. Assume also $\exists! y_0$ s.t. $F_* - G_* = 0$. Furthermore at y_0 we need a change in sign from $+$ to $-$. Let $S = \{s > 0 : f^s - g^s \in L^1(\mathcal{M}, \mu)\}$. Then*

$$\phi(s) = \frac{1}{sy_0^s} \int_{\mathcal{M}} f^s - g^s d\mu$$

is monotone increasing on S .

We will use this to show an integral inequality in the following proof. Here we present a sketch of Nazarov and Podkorytov's approach:

Proof. First we reduce the case $0 < p < p_0$ to $p = p_0$. Let $S = \sum_i a_i \epsilon_i$ and suppose $\mathbb{E}|S|^p \geq 2^{p_0-2/2}$ where $\sum_i a_i^2 = 1$. Via holder we have

$$\mathbb{E}|S|^{p_0} \leq (\mathbb{E}|S|^p)^{2-p_0/2-p} (\mathbb{E}|S|^2)^{p_0-p/2-p} = (\mathbb{E}|S|^p)^{2-p_0/2-p}$$

since $\mathbb{E}|S|^2 = 1$ by assumption. So we only consider now $p_0 \leq p \leq 2$.

As in Haagerup's argument, we crucially rely on an integral representation of these moments

$$\mathbb{E}|S|^p = C_p \int_0^\infty \frac{1 - \prod_{k=1}^n \cos(a_k u)}{u^{p+1}}$$

where we note the product of cosines the product of characteristic functions of the scaled random signs.

Using this representation we can reduce to question to an integral inequality. We have the comparison via AM-GM

$$\prod_{k=1}^n \cos(a_k u) \leq \prod_{k=1}^n |\cos(a_k u)| \leq \sum_{k=1}^n a_k^2 |\cos(a_k u)|^{q/a_k^2} = 1 - \sum_{k=1}^n a_k^2 (1 - |\cos(a_k u)|^{1/a_k^2})$$

Define

$$I_p(s) = C_p \int_0^\infty (1 - |\cos(\frac{u}{\sqrt{s}})|^2) \frac{du}{u^{p+1}}$$

Then we can write

$$\mathbb{E}|S|^p \geq \sum_{k=1}^n a_k^2 I_p(\frac{1}{a_k^2})$$

so it suffices to treat I_p . We know for $I_p(2) = 2^{p-2/2}$ via the integral representation. For $\lim_{s \rightarrow \infty} I_p(s)$ we may apply DCT to see get $C_p \int_0^\infty (1 - e^{-u^2/2}) \frac{du}{u^{p+1}}$ as well $\lim_{s \rightarrow \infty} I_p(s) = 2^{p-2/2} \frac{\Gamma(p+1/2)}{\Gamma(3/2)}$ via a CLT arugment. Then showing $I_p(s) \geq I_p(\infty)$ would allow us to conclude. Ie write

$$H(p, s) = \int_0^\infty (e^{-sx^2/2} - |\cos(x)|^s) \frac{dx}{x^{p+1}} \geq 0$$

It is this kind of integral inequality to which we can apply the distribution function lemma. We must compute

$$F_*(y) = \mu\{x > 0 : e^{-x^2/2} < y\} = \frac{1}{p} 2 \ln(1/y)^{-p/2}$$

$$G_*(y) = \mu\{x > 0 : |\cos(x)| < y\} = \frac{1}{p} \sum_{k=0}^\infty \frac{1}{(\pi k + \arccos(y))^p} - \frac{1}{(\pi k + \pi - \arccos(y))^p}$$

and show the difference only changes sign once on the interval $(0, \pi/2)$. This is shown using an extension of the difference to the complex plane.

After this it remains to treat the case of large coefficient, ie. one of the $a_k > \frac{1}{2} \sum_{k=1}^n a_k^2$. We must show

$$R_p(a) = \int_0^1 |\sum_{k=1}^n a_k r_k(t)|^p dt \geq A_p (1 + a_2^2 + \dots + a_n^2)^{p/2}$$

where we define $\phi_p(x) = (1+x)^{p/2}$ with $A_p = 2^{p-2/2} \frac{\Gamma(p+1/2)}{\Gamma(3/2)}$ for $p \in [p_0, 2)$.

To do this we go by induction. The observation

$$\mathbb{E}|S|^p(a) = \frac{\mathbb{E}|S|^p(a^+) + \mathbb{E}|S|^p(a^-)}{2}$$

where $a^+ = (a_1, \dots, a_{n-2}, a_{n-1} + a_n)$ and a^- analogously allows us to absorb large coefficients into smaller term expressions. For $n \geq 3$ we have the identity

$$\sum_{j=2}^n a_j^2 = \frac{1}{2} ((a_2^-)^2 + \dots + (a_{n-1}^-)^2 + (a_2^+)^2 + \dots + (a_{n-1}^+)^2)$$

Denoting $x = \sum_{j=2}^n a_j^2$ and correspondingly x^+, x^- we could write

$$R_p(a) = \frac{R_p(a^+) + R_p(a^-)}{2} \geq A_p \frac{\phi_p(x^+) + \phi_p(x^-)}{2} \geq A_p \phi_p\left(\frac{x^+ + x^-}{2}\right) = A_p \phi_p(x)$$

however ϕ_p is concave for $p < 2$. So instead we find a convex function Φ_p that dominates ϕ and show estimate using this instead. We find such a function by modifying ϕ_p on $[0, 1]$:

$$\Phi_p(x) = \begin{cases} \phi_p(x) & x \geq 1 \\ 2\phi_p(1) - \phi_p(2-x) & 0 \leq x \leq 1 \end{cases}$$

Here we have convexity for $x', x'' \geq 0$ with $\frac{x' + x''}{2} \leq 1$.

Let $n \geq 3$ and assume the induction hypothesis. Set $a_1 = 1$ with x as before. If a_1 is the largest coefficient with $x \geq 1$ then we are just in the small coefficient regime. If a_1 is largest with $x < 1$ we can apply the convexity trick since $x = \frac{x^- + x^+}{2} < 1$. If a_1 is not the largest coefficient then $x > 1$. We may wlog rearrange and renormalize the coefficients so $a_1 = 1$ and is the largest, falling into one of the two previous cases.

So it suffices to finish the base case. We want to show $R_p(1, a_2) \geq A_p \Phi_p(a_2^2)$. Wlog suppose $a_1 = 1$ is largest of coefficients. It is enough to prove the pointwise estimate

$$\frac{(1 + \sqrt{x})^p + (1 - \sqrt{x})^p}{2} \geq 2^{p-1} \left(2 - \left(\frac{3-x}{2}\right)^{p/2}\right)$$

After some rewriting we seek to show

$$a^p + b^p(1 + 2ab)^{p/2} \leq 2$$

for $p \leq 2, a, b \geq 0$ with $a + b = 1$. We have equality at $p = 2$ and notice we have convexity in p so we simply need to show we are decreasing at $p = 2$. This can be verified after straightforward computation and we are done. \square

Moving forward it is good to keep these sharp constants in mind, and the techniques used to prove them. Via the CLT arguments we see in B_p does not change if we change the distribution of the terms being summed. Similarly convexity arguments seem to continue to work well for $p \geq 3$.

This concludes our discussion of Khintchine type inequalities for random signs. We move on to discussing generalizations to other distributions.

2.2 Uniform Random Variables

Here we present results from [12] leading to Khintchine-type results for uniform variables t_i distributed over $[-1, 1]$. We get optimal upper bounds in the $p \geq 2$ case and lower bounds in $p \in [1, 2]$ case which are gaussian moments for comparisons between uniform random variables. This is made precise by the following.

Theorem 6. *Let $a = (a_1, \dots, a_n)$ and $b = (b_1, \dots, b_n)$ be two sequences of real numbers s.t. $(a_i^2) \prec (b_i^2)$ and t_1, \dots, t_n be a sequence of independent random variables uniformly distributed on $[-1, 1]$. Then*

$$(\mathbb{E} |\sum_{i=1}^n a_i t_i|^p)^{1/p} \leq (\mathbb{E} |\sum_{i=1}^n b_i t_i|^p)^{1/p}$$

for $p \in [1, 2]$ and

$$(\mathbb{E}|\sum_{i=1}^n a_i t_i|^p)^{1/p} \geq (\mathbb{E}|\sum_{i=1}^n b_i t_i|^p)^{1/p}$$

for $p \geq 2$

We establish some lemmas. First an alternative way of writing densities of symmetric unimodal distributions. We recall a random variable is *symmetric unimodal* if its density is symmetric and nonincreasing on $[0, \infty)$.

Lemma 7. *A real random variable X is symmetric unimodal \iff there exists a probability measure μ on $[0, \infty)$ s.t. density g of X is*

$$g(x) = \int_0^\infty \frac{1}{2t} \chi_{[-t, t]}(x) d\mu(t)$$

Proof. Define measure ν on $[0, \infty)$ via $\nu([x, \infty)) = g(x)$ where g is a density of some symmetric unimodal random variable. Set $\mu(t) = 2t\nu(t)$. For $x > 0$

$$g(x) = \int_0^\infty \chi_{[-t, t]}(x) d\nu(t) = \int_0^\infty \frac{1}{2t} \chi_{[-t, t]}(x) d\mu(t)$$

Compute

$$\int_0^\infty d\mu(t) = \int_0^\infty 2t d\nu(t) = \int_0^\infty \int \chi_{[-t, t]}(x) dx d\nu(t) = \int g(x) dx = 1$$

so μ probability measure. \square

Lemma 8. *If $X = \sum_{i=1}^n X_i$ and X_i are independent symmetric unimodal random variables then X is symmetric unimodal. In particular if $X = \sum_{i=1}^n a_i t_i$ where t_i are independent and uniformly distributed on $[-1, 1]$ then X symmetric unimodal.*

Proof. We simply show closure under sum. Let X_1, X_2 be independent symmetric unimodal with densities g_1, g_2 and measures μ_1, μ_2 as above. Compute density g of $X_1 + X_2$ as

$$g(x) = \int_0^\infty \int_0^\infty \frac{1}{4ts} \chi_{[-t, t]} * \chi_{[-s, s]}(x) d\mu(t) d\mu(s)$$

via lemma 1 representation. Clearly g is symmetric and nonincreasing on $[0, \infty)$. \square

We now present a key technical lemma.

Lemma 9.

$$G(t) = \begin{cases} (p+2) \frac{(t+1)^{p+1} - (t-1)^{p+1}}{t^2} - \frac{(t+1)^{p+2} - (t-1)^{p+2}}{t^3} & t \geq 1 \\ (p+2) \frac{(1+t)^{p+1} + (1-t)^{p+1}}{t^2} - \frac{(1+t)^{p+2} - (1-t)^{p+2}}{t^3} & 0 < t < 1 \end{cases}$$

Then G is nondecreasing on $(0, \infty)$ if $p \geq 2$ and nonincreasing for $1 \leq p \leq 2$.

For proof we refer readers to lemma 3 of [12]. We present the final lemma which shall directly imply the desired result.

Lemma 10. *If t_1, t_2, t_3 are independent random variables uniformly distributed on $[-1, 1]$ and $a, b, c, d > 0$ with $a^2 + b^2 = c^2 + d^2$ and $c \geq a \geq b \geq d$ then*

$$\begin{aligned} \mathbb{E}|t_1 + at_2 + bt_3|^p &\leq \mathbb{E}|t_1 + ct_2 + dt_3|^p & p \in [1, 2] \\ \mathbb{E}|t_1 + at_2 + bt_3|^p &\geq \mathbb{E}|t_1 + ct_2 + dt_3|^p & p \geq 2 \end{aligned}$$

Proof. We have the observation

$$|x|^p = \frac{d^3}{dx^3} \left(\frac{x^3 |x|^p}{(p+1)(p+2)(p+3)} \right)$$

integrating by parts then gives

$$\begin{aligned} \mathbb{E}|t_1 + at_3 + bt_3|^p &= \frac{1}{8} \int_{-1}^1 \int_{-1}^1 \int_{-1}^1 |x_1 + ax_2 + bx_3|^p dx_1 dx_2 dx_3 = \\ &= c_p \left(\frac{(a+b+1)^3 |a+b+1|^p + (a-b-1)^3 |a-b-1|^p}{ab} \right. \\ &\quad \left. - \frac{(a-b+1)^3 |a-b+1|^p + (a+b-1)^3 |a+b-1|^p}{ab} \right) \end{aligned}$$

□

Proof. Proof of the theorem

give proof

□

Note the schur concavity result immediately shows optimal constant khintchine inequalities via central limit theorem. Indeed for the case $p \in [1, 2]$ we have

$$\mathbb{E} \left| \sum_{i=1}^n a_i t_i \right|^p \geq \left(\sum_{i=1}^n a_i^2 \right)^{p/2} \mathbb{E} \left| \sum_{i=1}^n \frac{1}{\sqrt{n}} t_i \right|^p$$

from which we just apply CLT.

The technique of finding the stronger schur-concavity result is often useful and used repeatedly in what follows to show Khintchine-type inequalities.

2.3 Ultra Sub-Gaussian Random Variables

This section examines a result via Nayar and Oleszkiewicz in [6] which significantly generalizes the class of random variables considered and achieves comparisons for all even integer moments via log-concavity.

First we discuss the notions of Ultra Sub-Gaussianity and Strong Log Concavity of a Random Variable. Recall

Definition 11. A sequence $(a_i)_{i=0}^\infty$ of non-negative real numbers is called *log-concave* if $a_i^2 \geq a_{i-1}a_{i+1}$.

Then we can define the notion of *Ultra Sub-Gaussianity* for \mathbb{R}^n valued random vectors from [6].

Definition 12. \mathbb{R}^n valued X is *Ultra sub-Gaussian* if $X = 0$ or X is rotation invariant, has finite moments, and has gaussian log-concave even moments ie. $a_i = \mathbb{E}||X||^{2i} / \mathbb{E}||G||^{2i}$ are log-concave.

The results crucially use Walkup's Theorem concerning preservation of log-concavity under convolution:

Theorem 13. Let $(a_i)_{i=0}^\infty$ and $(b_i)_{i=0}^\infty$ be two log-concave sequences of positive real numbers. We define:

$$c_n = \sum_{i=0}^n \binom{n}{i} a_i b_{n-i}$$

Then the $(c_n)_{n=0}^\infty$ is log-concave

This tells us the collection **ULCS** of ultra log-concave sequences is closed under convolution:

Definition 14. $(a_i)_{i=0}^\infty$ is ultra log-concave if $(i!a_i)_{i=0}^\infty$ is log-concave.

Given X Ultra Sub-Gaussian we can extract a khintchine type inequality of the following form.

Theorem 15. *Let n, d positive integers and $p > q \geq 2$ even integers. If X_1, \dots, X_n independent \mathbb{R}^n valued random vectors are ultra sub-Gaussian then*

$$(\mathbb{E}|S|^p)^{1/p} \leq \frac{(\mathbb{E}|G|^p)^{1/p}}{(\mathbb{E}|G|^q)^{1/q}} \mathbb{E}(|S|^q)^{1/q}$$

where $S = \sum_i X_i$

The proof rests on some lemmas we highlight now.

Lemma 16. *Let $\pi : \mathbb{R}^n \rightarrow \mathbb{R}$ be projection to first coordinate. For $p > 0$ assume X rotation invariant random vector on \mathbb{R}^n with finite p th moment. Then*

$$\frac{\mathbb{E}|X|^p}{\mathbb{E}|G|^p} = \frac{\mathbb{E}\|X\|^p}{\mathbb{E}\|G\|^p}$$

Notably the class of ultra sub-gaussian random variables is closed under sums.

Lemma 17. *If $X, Y \in \mathbf{USG}$ are independent random vectors then $X + Y \in \mathbf{USG}$ Uniform.*

Proof. We may assume X and Y nonzero constant. Setting

$$\begin{aligned} a_i &= \mathbb{E}\|X\|^{2i} / \mathbb{E}\|G\|^{2i} = \mathbb{E}(\prod X)^{2i} / \mathbb{E}G^{2i} \\ b_i &= \mathbb{E}\|Y\|^{2i} / \mathbb{E}\|G\|^{2i} = \mathbb{E}(\prod Y)^{2i} / \mathbb{E}G^{2i} \\ b_i &= \mathbb{E}\|X + Y\|^{2i} / \mathbb{E}\|G\|^{2i} = \mathbb{E}(\prod (X + Y))^{2i} / \mathbb{E}G^{2i} \end{aligned}$$

and then note

$$c_n = \frac{1}{(2n-1)!!} \sum_{i=0}^n \binom{2n}{2i} \mathbb{E}(X)^{2i} \mathbb{E}(Y)^{2n-2i} = \sum_{i=0}^n \frac{(2n)!!}{(2i)!!(2n-2i)!!} a_i b_{n-i} = \sum_{i=0}^n \binom{n}{i} a_i b_{n-i}$$

and we conclude with Walkup's theorem. \square

We must also note distributions producting with nonnegative random variables to a standard gaussian must ultra sub-gaussian.

Lemma 18. *Assume an \mathbb{R}^d valued random vector X and non-negative random variable R are independent and that $R \cdot X$ has distribution $N(0, Id_d)$. Then X is ultra sub-gaussian*

Now we establish the Khintchine-type inequality for ultra sub-gaussian random variables.

Proof. Note first the constant is optimal as indicated by the central limit theorem.

If we show S ultra log-concave then we know the sequence $a_k = \mathbb{E}\|S\|^{2k} / \mathbb{E}\|G\|^{2k}$ is log-concave which gives us our desired result since then we know $(a_s^{1/s})_{s=1}^\infty$ is non-increasing and in particular $a_{p/2}^{2/p} \leq a_{q/2}^{2/q}$. But S is clearly ultra log-concave as it is the sum of ultra log-concave random variables. So we are done. \square

These results in turn relate to our discussion of type \mathcal{L} random variables, where we discover every Type \mathcal{L} random variable is ultra sub-gaussian.

2.4 Gaussian Mixtures

First we look at [11]. Gaussian measures very important in convex geometry and local banach space theory.

Recall what a gaussian mixture is:

Definition 19. A random variable X is a *Gaussian mixture* if there exists a positive random variable Y and standard Gaussian Z independent from Y s.t. X has the same distribution as YZ .

Example 20. A random variable with density $f(x) = \sum_{j=1}^m p_j \frac{1}{\sqrt{2\pi}\sigma_j} e^{-x^2/2\sigma_j^2}$ is a gaussian mixture

We also recall definitions of majorization:

Definition 21. $a \in \mathbb{R}^n$ is majorized by $b \in \mathbb{R}^n$, written $a \prec b$, if for $k = 1, \dots, n-1$ we have

$$\sum_{j=1}^k a_j^* \leq \sum_{j=1}^k b_j^*$$

,where a^* is the decreasing rearrangement, and we have $\sum_{j=1}^n a_j = \sum_{j=1}^n b_j$.

Definition 22. Then a function preserving \prec is Schur convex

This allows us to state:

Theorem 23. Let X be a Gaussian mixture and X_1, \dots, X_n be independent copies of X . For two vectors $a, b \in \mathbb{R}^n$ with $p \geq 2$ we have

$$(a_1^2, \dots, a_n^2) \prec (b_1^2, \dots, b_n^2) \implies \left\| \sum_{i=1}^n a_i X_i \right\|_p \leq \left\| \sum_{i=1}^n b_i X_i \right\|_p$$

The proof relies on Bernstein's theorem:

Theorem 24. $g : (0, \infty) \rightarrow \mathbb{R}$ is completely monotonic, ie. $(-1)^n g^{(n)} \geq 0$, if and only if \exists non-negative Borel measure μ on $[0, \infty)$ s.t.

$$f(x) = \int_0^\infty e^{-tx} d\mu(t)$$

Bernstein in turn can be used to show the following:

Theorem 25. A symmetric random variable X with density f is a Gaussian mixture if and only if $x \rightarrow f(\sqrt{x})$ is completely monotonic for $x > 0$.

Proof. Let X be a symmetric random variable with density f with $x \rightarrow f(\sqrt{x})$ is completely monotonic. By Bernstein there exists a non-negative Borel measure μ supported on $[0, \infty)$ s.t. $f(\sqrt{x}) = \int_0^\infty e^{-tx} d\mu(t)$. We can now compute

$$P(X \in A) = \int_A \int_0^\infty e^{-tx^2} d\mu(t) dx = \int_0^\infty \int_A e^{-tx^2} dx d\mu(t) = \int_0^\infty \int_{\sqrt{2t}A} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

and then simplify to $\int_0^\infty \gamma_n(\sqrt{2t}A) d\nu(t)$ by writing $d\nu(t) = \sqrt{\frac{\pi}{t}} d\mu(t)$. We note this is a probability measure.

Now let V be distributed via ν . V is positive and hence we can write $Y = \frac{1}{\sqrt{2V}}$. Then

$$P(YZ \in A) = P\left(\frac{1}{2V} Z \in A\right) = \int_0^\infty \gamma_n(\sqrt{2t}A) d\nu(t) = P(X \in A)$$

which concludes the proof. The converse follows via Bernstein and noting X has density $f(x) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-x^2/2y^2} \frac{d\nu(y)}{y}$ □

Note this is a la [12]. Similarly inspired by [13] we have a result for vectors uniformly distributed on l_p balls:

Theorem 26. Fix $q \in (0, 2]$ and let $X = (X_1, \dots, X_n)$ a random vector uniformly distributed on B_q^n . For two vectors (a_1, \dots, a_n) and (b_1, \dots, b_n) in \mathbb{R}^n with $p \geq 2$ we have

$$(a_1^2, \dots, a_n^2) \prec (b_1^2, \dots, b_n^2) \implies \|\sum_{i=1}^n a_i X_i\|_p \leq \|\sum_{i=1}^n b_i X_i\|_p$$

whereas for $p \in (-1, 2)$ the second inequality is reversed

2.5 Random Vectors Uniform on \mathbb{S}^{n-1}

Here we present the results from [18].

Let Δ denote the laplace operator. Recall a function is subharmonic if for any ball $B(x, r)$ for $x \in \mathbb{R}^n, r > 0$ and continuous function h harmonic on $B(x, r)$ s.t. $\phi \leq h$ on $\partial B(x, r)$ then $\phi \leq h$ in $B(x, r)$ or equivalently $\Delta \phi \geq 0$. Then a function is superharmonic if its negation is subharmonic. We say continuous $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$ is bisubharmonic if $\Delta \Phi$ equals a subharmonic in distribution. Equivalently if $\Phi \in C^4(\mathbb{R}^n; \mathbb{R})$ then Φ satisfies $\Delta \Delta \Phi \geq 0 \iff$ bisubharmonic.

Theorem 27. Let $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$ bisubharmonic and continuous on \mathbb{R}^n . Set $X = \sum_{j=1}^n a_j U_j, Y = \sum_{j=1}^n b_j U_j$. If $(b_i^2) \prec (a_i^2)$ then

$$\mathbb{E} \Phi(X) \leq \mathbb{E} \Phi(Y)$$

So for fixed square sum more spread out mass on the coefficients results in larger moments. Set

$$I(r) = I(r, f) = \int_{\mathbb{S}^{d-1}} f(rx) dx, \quad r \geq 0$$

where the integral over \mathbb{S}^{n-1} denotes the normalized uniform measure on \mathbb{S}^{n-1} . Before we prove the theorem we establish some lemmas.

Lemma 28. Suppose $f \in C^2(\mathbb{R}^n; \mathbb{R})$. Then

$$\frac{d}{dt} I(t^{1/2}, f) = \frac{1}{2n} \int_{\mathbb{B}^d} \Delta f(t^{1/2} x) dx, \quad 0 \leq t < \infty$$

Proof. We can compute using divergence theorem

$$I'(r) = \frac{r}{n} \int_{\mathbb{B}^d} \Delta f(rx) dx$$

with the change of variables $s = nb$. Note $\mathbb{B}^n = \{x \in \mathbb{R}^n : |x| < 1\}$. □

Lemma 29. Suppose $f \in C(\mathbb{R}^n; \mathbb{R})$. Then for $\alpha, \beta \in \mathbb{R}$ we have

$$\int_{\mathbb{S}^{n-1} \times \mathbb{S}^{d-1}} f(\alpha x + \beta y) dx dy = \int_{\mathbb{S}^{d-1}} I(|\alpha e_1 + \beta y|, f) dy$$

Proof. Fix $n \geq 2$ with $S = \mathbb{S}^{n-1}$ and dg the Haar measure on $G = SO(n)$. For continuous F on \mathbb{R}^n we have

$$\int_G F(gz) dg = \int_S F(|z|x) dx$$

$$\begin{aligned}
\int_{S \times S} f(\alpha x + \beta y) dx dy &= \int_{G \times G} f(\alpha g_1 e_1 + \beta g_2 e_1) dg_1 dg_2 = \int_G dg_1 \int_G f(g_1(\alpha e_1 + \beta g_1^{-1} g_2 e_1)) dg_2 \\
&= \int_G dg_1 \int_G f(g_1(\alpha e_1 + \beta g_2 e_1)) dg_2 = \int_G dg_2 \int_G f(g_1(\alpha e_1 + \beta g_2 e_1)) dg_1 \\
&= \int_G dg_2 \int_S f(|\alpha e_1 + \beta g_2 e_1| y) dy = \int_G I(|\alpha e_1 + \beta g_2 e_1|, f) dg_2 \\
&= \int_S I(|\alpha e_1 + \beta y|, f) dy
\end{aligned}$$

□

Proof. Proof of the theorem Now we are prepared to present a proof of the theorem. We may suppose $a_i, b_i \geq 0$ and via approximation arguments that $\Phi \in C^4$ (if Φ not in C^4 we just get arbitrarily close with a C^4 Ψ). Set \mathbb{R}_+^d to be the strictly positive section of \mathbb{R}^d . We then let $Q : \mathbb{R}_+^d \rightarrow \mathbb{R}$ via

$$Q(s) = \mathbb{E} \Phi \left(\sum_{j=1}^n s_j^{1/2} U_i \right) = \int_{\prod_{j=1}^n \mathbb{S}^{d-1}} \Phi \left(\sum_{j=1}^n s_j^{1/2} x_j \right) dx_1 \dots dx_n$$

So we seek to show if $b \prec a$ then $Q(a) \leq Q(b)$. Write for $b \prec a$

$$Q(b) - Q(a) = \int_0^1 \frac{d}{dt} Q((1-t)a + tb) dt = \int_0^1 \sum_{j=1}^n \partial_j Q((1-t)a + tb) (b_j - a_j) dt$$

We obtain the expression

$$\partial_i Q(s) = \frac{1}{2d} \int_{\Omega_i} \Delta \Phi \left(\sum_{j=1}^n s_j^{1/2} x_j \right) dx_1 \dots dx_n$$

by integrating Q over $\mathbb{S}^{d-1} \times \dots \times \mathbb{B}^d \times \dots \times \mathbb{S}^{d-1}$ where \mathbb{B}^d the i th factor first in x_i the differentiate in s_i and applying the first lemma to $x_i \rightarrow \Phi(x_i + \sum_{j \neq i} s_j^{1/2} x_j)$.

Note $\int_{\mathbb{B}^d} F(x) dx = \int_0^1 dr r^{d-1} dr = \int_{\mathbb{S}^{d-1}} F(rx) dx$

So applying this to the above we get

$$\partial_1 Q(s) = \frac{1}{2} \int_{\prod_{j=1}^{k-2} \mathbb{S}^{d-1}} dx_2 \dots dx_k \dots \int_0^1 r^{n-1} dr \int_{\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}} \Delta \Phi(s_1^{1/2} r x_1 + s_2^{1/2} x_2 + \sum_{j=3}^n s_j^{1/2} x_j) dx_1 dx_2$$

Set $f(z) = \Delta \Phi(z + \sum_{j=3}^n s_j^{1/2} x_j)$. By our second lemma we know the above $dx_1 dx_2$ integral term equals $I(|r s_1^{1/2} + s_2^{1/2}|, f)$. We can apply a similar analysis to $\partial_2 Q(s)$.

Since $f = \Delta \Phi$ subharmonic we know $I(r, f)$ increases as r increases. So then we have shown if $s_1 \geq s_2$ then $\partial_1 Q(s) \leq \partial_2 Q(s)$.

Clearly Q permutation invariant so we may assume components of a, b decrease as i increases. So components of $(1-t)a + tb$ also decrease and hence $\partial_j Q((1-t)a + tb)$ increases. We know $b \prec a$ and $\lambda = (\lambda_1, \dots, \lambda_k)$ has increasing terms, then $b \prec a$ implies $\sum \lambda_j b_j \geq \sum \lambda_j a_j$. Thus we know $\sum_{j=1}^n \partial_j Q((1-t)a + tb) (b_j - a_j)$ is nonnegative and we are done. □

Note we crucially use the bisubharmonicity of Φ to have $f = \Delta \Phi$ subharmonic and thus $I(r, f)$ increasing as r increasing.

Recall if Φ radial then $\Phi(x) = R(|x|)$ for $x \in \mathbb{R}$. Interestingly if Φ radial then a converse holds as well:

Theorem 30. *Let $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$ be continuous and radial. Then if $\mathbb{E}\Phi(X) \leq \mathbb{E}\Phi(Y)$ for all pairs of a, b with $(b_i^2) \prec (a_i^2)$ then Φ is bisubharmonic.*

For proof we refer the reader to [18].

3 Type \mathcal{L} Random Variables

Here we cover results from [5] giving new types of Khintchine inequalities for Type \mathcal{L} random variables.

First defined by Newman in [16], we define random variable X to be of type \mathcal{L} if we have sub-gaussianity i.e. $\exists A, B$ s.t. $|\mathbb{E}e^{zX}| \leq Ae^{B|z|^2}$ and the characteristic is even with strictly real zeroes or equivalently $\mathbb{E}e^{zX}$ is even with pure imaginary zeroes. If the distribution is not symmetric we say instead $X \in \mathcal{L}'$.

Note clearly \mathcal{L} closed under sum since the characteristic of the sum $X + Y$ is the product of the characteristics, which preserves sub-gaussianity and pure imaginary zeroes.

Basic examples of type \mathcal{L} random variables include random signs, arithmetic progressions, and the gaussian. Indeed clearly all these have characteristics which are subgaussian and have strictly real zeroes.

3.1 Connections to Ultra Sub-Gaussianity

Newman showed in [16] a class of $2, p$ khintchine inequalities:

Theorem 31. *If the $\{X_j\}_{j=1}^N$ are independent random variables of type L , then for any real a_j with $X = \sum_j a_j X_j$ and even m we have*

$$\mathbb{E}|X|^{2m} \leq \frac{(2m)!}{2^m m!} (\mathbb{E}|X|^2)^m$$

The proof is importantly uses Hadamard's factorization theorem for the characteristics of our random variables, allowing us to compare moments of with terms in the exponential power series. We generalize this approach below. Now we present the main Khintchine type result for type \mathcal{L} random variables.

Theorem 32. *Let X be a random variable of type \mathcal{L}' . Then for every even integers $2 \leq p \leq q$, we have*

$$\|X\|_q \leq \frac{\|G\|_q}{\|G\|_p} \|X\|_p,$$

where G is a standard Gaussian random variable.

Proof. First note If $X \in \mathcal{L}'$ then we can find $c \in \mathbb{R}$ s.t. $X + c \in \mathcal{L}$. Since

$$\mathbb{E}e^{zX} = e^{bz^2/2+cz} \prod_j (1 + \alpha_j z^2/2)$$

for $b \geq 0$ and some $c \in \mathbb{R}$, $\alpha_j > 0$ by hadamard factorization. So if $X \in \mathcal{L}'$ then $Y = X - c \in \mathcal{L}$. We can write $\mathbb{E}|X|^p = \mathbb{E}|Y + c|^p = \mathbb{E}|Y + c\epsilon|^p$ since Y is symmetric. Further this is type L as it is the sum of two type L random variable. So X of type \mathcal{L}' has moments equal to some type L random variable $Y + c\epsilon$ and so it suffices to simply consider $X \in \mathcal{L}$.

So we know we have

$$\mathbb{E}e^{zX} = e^{bz^2/2+cz} \prod_j (1 + \alpha_j z^2/2)$$

so expanding with exponential power series and applying $z = \sqrt{2t}$ yields

$$\sum_{n=0}^{\infty} \frac{\mathbb{E}X^{2n}}{(2n)!} 2^n t^n = \mathbb{E}e^{\sqrt{2t}X} = e^{bt} \prod_j (1 + \alpha_j t)$$

We can write $\prod_j (1 + \alpha_j t) = \sum_{k=0}^{\infty} \sigma_k t^k$ where σ_k are the elementary symmetric functions in α_j . So by equating coefficients we know

$$\frac{\mathbb{E} X^{2n}}{\mathbb{E} G^{2n}} = n! \sum_{k=0}^n \frac{b^{n-k}}{(n-k)!} \sigma_k = \sum_{k=0}^n \binom{n}{k} b^{n-k} \sigma_k k!$$

Thus it suffices to show the sequence $s = (\sigma_k k!)_{k \geq 0}$ is log-concave as then $\frac{\mathbb{E} X^{2n}}{\mathbb{E} G^{2n}}$ is log-concave yielding the desired result.

We show s_k log-concave via newton's inequalities which tell us the elementary symmetric functions are ultra-log concave ie.

$$\frac{\sigma_{k-1}}{\binom{n}{k-1}} \frac{\sigma_{k+1}}{\binom{n}{k+1}} \leq \frac{\sigma_k^2}{\binom{n}{k}^2 s}$$

which is exactly what is needed, since clearly then $k! \sigma_k$ also log-concave. \square

Corollary 33. *If $X \in \mathcal{L}'$ then X is ultra sub-Gaussian.*

We see this immediately from the previous theorem since we showed X 's moments log-concave.

We also have the following generalization to hilbert spaces:

Corollary 34. *Let $(H, \|\cdot\|)$ be a separable (real or complex) Hilbert space. If X_1, \dots, X_n are independent type \mathcal{L} random variables, then for every vectors v_1, \dots, v_n in H , the sum $X = \sum_{j=1}^n X_j v_j$ satisfies $\|X\|_q \leq \frac{\|G\|_q}{\|G\|_p} \|X\|_p$ for all positive even integers $p \leq q$, where we denote $\|X\|_p = (\mathbb{E} \|X\|^p)^{1/p}$.*

The proof follows directly from two applications our result after defining the isometric embedding for v in hilbert H , $\{e_1, \dots\}$ an orthonormal basis and g_j iid standard gaussians:

$$v \rightarrow c_q \sum_j \langle v, e_j \rangle g_j$$

Most of the time independence is assumed for these khintchine type inequalities. However here we also present a result allowing for dependencies between the summed random variables, again inspired by ferromagnetic model considerations from [17], in particular proposition 6 which we present now.

Theorem 35. *Suppose X_1, \dots, X_n have joint distribution ν given by*

$$\nu(x_1, \dots, x_n) = \frac{1}{Z} e^{\sum_{j=1}^n h_j x_j + \sum_{j \leq k} J_{j,k} x_j x_k} \prod_{j=1}^n \rho_j(x_j)$$

with $J_{j,k}, h_j \geq 0$ and ρ_j even sub-Gaussian probability measure. Let δ an independent bernoulli and set $Y_0 = \delta$ and $Y_j = \delta X_j$ then the Y_j have joint distribution

$$\hat{\nu}(y_0, \dots, y_n) = \frac{1}{Z} e^{\sum_{j \leq k} J_{j,k} y_j y_k} \prod_{j=0}^n \rho_j(y_j)$$

Having this we state and prove

Corollary 36. *Let μ_1, \dots, μ_n be Borel probability measures on \mathbb{R} , each one of type \mathcal{L} . Suppose that (X_1, \dots, X_n) is a random vector in \mathbb{R}^n whose law ρ on \mathbb{R}^n is of the form*

$$d\rho(x_1, \dots, x_n) = Z^{-1} \exp \left(\sum_{j=1}^n h_j x_j + \sum_{j,k=1}^n J_{j,k} x_j x_k \right) d\mu_1(x_1) \dots d\mu_n(x_n) \quad (1)$$

with $h_j \geq 0$, $J_{jk} \geq 0$ for all $j, k \leq n$, where Z is the normalising constant. Then for every nonnegative a_1, \dots, a_n , the sum $X = \sum_{j=1}^n a_j X_j$ satisfies (32) for all positive even integers $p \leq q$.

Proof. Let (X_1, \dots, X_n) be a random vector with distribution given by (1) and let ε be an independent Rademacher random variable. By the previous theorem, the vector $(Y_0, Y_1, \dots, Y_n) = (\varepsilon, \varepsilon X_1, \dots, \varepsilon X_n)$ has distribution ρ' of the form

$$d\rho'(x_0, x_1, \dots, x_n) = Z'^{-1} \exp \left(\sum_{j,k=0}^n J'_{jk} x_j x_k \right) d\mu_0(x_0) d\mu_1(x_1) \dots \mu_n(x_n),$$

where μ_0 is the distribution of ε , $J'_{0,0} = 0$, $J'_{0,k} = J'_{k,0} = h_k/2$, $J'_{jk} = J_{jk}$, $j, k \geq 1$, so of the form (1) with $h \equiv 0$. Therefore, by Theorem 2 from [16], for every $a_0, a_1, \dots, a_n \geq 0$, the sum $S = \sum_{j=0}^n a_j Y_j = a_0 \varepsilon + \sum_{j=1}^n a_j \varepsilon X_j$ is of type \mathcal{L} and in particular, S satisfies (32). Hence, taking $a_0 = 0$ yields that $\sum_{j=1}^n a_j X_j$ also satisfies (32). \square \square

3.2 Type \mathcal{L} Random Variables with "Enough Gaussianity"

The above results hold only for even integer moment comparisons. We now turn our attention to inequalities for a restricted class of type \mathcal{L} random variables. In particular, those with "enough gaussianity". The following lemma makes this precise.

Lemma 37. *For every $b > 0$ and $a_1, a_2, \dots \geq 0$ with $\sum a_j \leq b$ we know*

$$e^{-bt^2/2} \prod_{j=1}^n (1 - b_j t^2)$$

is the characteristic of a type \mathcal{L} random variable.

Proof. Define $\phi_a(t) = e^{-t^2/2}(1 - at^2)$ for $a \in [0, 1]$. Taking the inverse fourier transform we have

$$f_a(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_a(t) e^{-itx} dt = (1 - a + ax^2) \frac{e^{-x^2/2}}{\sqrt{2\pi}}$$

which is a density for $0 \leq a \leq 1$. In particular it is type \mathcal{L} since its characteristic has real zeroes and is sub-Gaussian.

Then it suffices to notice

$$e^{-bt^2/2} \prod_j (1 - b_j t^2)$$

is the characteristic of a sum of type \mathcal{L} random variables distributed according to f_{a_j} for some a_j and a gaussian. \square

We define Z_a to be the random variable with density $f_a(x) = (1 - a + ax^2) \frac{e^{-x^2/2}}{\sqrt{2\pi}}$. Note Z_0 is gaussian. Having established this class of random variables as type \mathcal{L} we then can show khitchine-type inequalities $p \geq 3$ and $q = 2$.

Theorem 38. *Let X be type \mathcal{L} random variable with characteristic function of the form $\phi_X(t) = e^{-bt^2/2} \prod_{j=1}^{\infty} (1 - a_j t^2)$ with $b > 0$, $a_j \geq 0$, $\sum a_j \leq b$. Let $\sigma = \sqrt{\text{Var}(X)}$. Then for every $p \geq 3$,*

$$\mathbb{E}|\sigma Z_1|^p \leq \mathbb{E}|X|^p \leq \mathbb{E}|\sigma Z_0|^p,$$

where Z_0 is a standard Gaussian random variable and Z_1 is a random variable with density $(2\pi)^{-1/2} x^2 e^{-x^2/2}$.

Before giving proof we establish a useful lemma which will allow us to bound moments via schur-concavity.

Lemma 39. For $\lambda \in (0, 1)$, let g_λ be the density of $\sqrt{\lambda}X_1 + \sqrt{1-\lambda}X_2$, where X_1, X_2 are independent copies of Z_1 . Then for every $0 < \lambda_1 < \lambda_2 < \frac{1}{2}$, the function $g_{\lambda_2} - g_{\lambda_1}$ on $(0, +\infty)$ has exactly two zeros and the sign pattern $+-+$.

Proof. By a direct computation,

$$g_\lambda(x) = \left(x^2 + \lambda(1-\lambda)(3-6x^2+x^4) \right) \frac{e^{-x^2/2}}{\sqrt{2\pi}}$$

so $g_{\lambda_2} - g_{\lambda_1}$ has the same sign as $(\lambda_2 - \lambda_1)(1 - \lambda_1 - \lambda_2)(3 - 6x^2 + x^4)$. \square

Lemma 40. Let X_1, X_2, \dots be i.i.d. copies of Z_1 and let Y be a symmetric random variable independent of the X_j . Then the function

$$\Psi(b_1, \dots, b_n) = \mathbb{E}|\sqrt{b_1}X_1 + \dots + \sqrt{b_n}X_n + Y|^p$$

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

Proof. We use the technique of interlacing densities (see, e.g. [21] or [23]). Let $h(x) = |x + 1|^p + |x - 1|^p$. It suffices to show that for every $0 < \lambda_1 < \lambda_2 < \frac{1}{2}$, we have

$$\int_0^\infty h(x)(g_{\lambda_2(x)} - g_{\lambda_1(x)})dx \geq 0,$$

where g_λ is as in Lemma 39. For arbitrary α, β , $\int(\alpha x^2 + \beta)(g_{\lambda_2(x)} - g_{\lambda_1(x)})dx = 0$, so the desired inequality is equivalent to

$$\int_0^\infty \tilde{h}(|x|)(g_{\lambda_2(x)} - g_{\lambda_1(x)})dx \geq 0,$$

with $\tilde{h}(x) = h(x) + \alpha x^2 + \beta$. Let x_1, x_2 be the zeros of $g_{\lambda_2(x)} - g_{\lambda_1(x)}$. Choose α and β such that \tilde{h} has zeros at x_1 and x_2 . Since for $p \geq 3$, $\tilde{h}(\sqrt{x})$ is convex on $(0, +\infty)$, \tilde{h} on $(0, +\infty)$ has no other zeros and the sign pattern $+-+$. Thus the integrand is pointwise nonnegative, hence the result. \square

Now we present the proof the theorem.

Proof. Proof of Theorem

Via approximation we suppose finitely many of the a_j are nonzero. Normalize $\sum_{j=1}^n a_j = 1$ and $b = 1 + c$ so then X is the same in distribution as $\sum_{j=1}^n \sqrt{a_j}Z_1^{(j)} + \sqrt{c}Z_0$. We compute the variance

$$\text{Var}(X) = \sum_{j=1}^n a_j \text{Var}(Z_1) + c \text{Var}(Z_0) = 3 + c$$

where we know $\text{Var}(Z_1) = 3$.

By the lemma we know with schur-concavity

$$\mathbb{E}|X|^p \leq \mathbb{E}\left|\sum_{j=1}^n \frac{1}{\sqrt{n}}X_j + \sqrt{c}Z_0\right|^p$$

Sending $n \rightarrow \infty$ and using the central limit theorem we get the gaussian moment as an upper bound, as desired.

Again via schur-concavity we get our lower bound by shifting all mass onto one term giving

$$\mathbb{E}|X|^p \geq \mathbb{E}|Z_1 + \sqrt{c}Z_0|^p$$

whose density can be directly computed yielding the desired lower bound. \square

This concludes the discussion of Khintchine-type results for type \mathcal{L} random variables.

3.2.1 Examples of Type \mathcal{L} Random Variables

We list some examples of probability distributions of type \mathcal{L} . In what follows, X is a symmetric random variable.

- (a) Let X be integer-valued with $\mathbb{P}(X = 0) = p_0$ and $\mathbb{P}(X = -k) = \mathbb{P}(X = k) = p_k$, $k = 1, \dots, n$ for nonnegative p_0, \dots, p_n with $p_0 + 2 \sum_{k=1}^n p_k = 1$.

If $\frac{1}{2}p_0 \leq p_1 \leq \dots \leq p_n$, then $\mathbb{E} \cos(zX) = p_0 + \sum_{k=1}^n (2p_k) \cos(kz)$ has only real zeros, as it follows from the Eneström-Kakeya theorem (see, e.g. Problem III.204 in [22]). As a result, X is of type \mathcal{L} . In particular, if X is uniform on $\{-n, \dots, 1, 1, \dots, n\}$ with a possible atom at 0 satisfying $\mathbb{P}(X = 0) \leq \frac{1}{n+1}$, then X is of \mathcal{L}

By the symmetry of X , the polynomial $Q(w) = \mathbb{E}w^{X+n}$ is self-inversive (the sequence of its coefficients is a palindrome, in other words, $w^{2n}Q(1/w) = Q(w)$). In particular, all its roots are symmetric with respect to the unit circle, that is if w_0 is a root of Q , then so is $1/\overline{w_0}$. For instance, if for some $\alpha \geq 1$,

$$\frac{1}{2}p_0^\alpha + \sum_{k=1}^{n-1} p_k^\alpha \leq \left(\frac{2}{n-2}\right)^{\alpha-1} p_n^\alpha,$$

where n is the number of nonzero coefficients of Q , then Q has zeros only on the unit circle, so X is of \mathcal{L}

- (b) Let X take values in $[-1, 1]$ and have a density f (which is even). Each of the following conditions implies that X is of \mathcal{L} .

- (i) f is nondecreasing on $(0, 1)$.
- (ii) f is C^2 with $f' < 0$ and $f'' < 0$ on $(0, 1)$.
- (iii) $f(t) = h(t)^\alpha$, where $\alpha > -1$ and h is an entire even function which is real-valued on the real line, $h(1) = 0$, $h(0) > 0$ and $h'(iz)$ is in the Laguerre-Pólya class. In particular, $f(t) = \text{const} \cdot (1 - t^{2m})^\alpha$ for a nonnegative integer m .

Moreover, if X has a density on \mathbb{R} of the following form, then it is of type \mathcal{L} .

- (iv) $f(t) = \text{const} \cdot e^{-t^{2m}}$ for a nonnegative integer m .
- (v) $f(t) = (2\pi)^{-1/2} e^{-t^2/2} (1 - b + bt^2)$, $0 \leq b \leq 1$.

4 Discrete Symmetric Distributions

We now consider the generalization of a random sign to a symmetric discrete random variables uniformly distributed outside 0. Ie. consider the generalization to $X \in \{-L, \dots, 0, \dots, L\}$ with some mass $\mathbb{P}(X = 0) = \rho_0$ and otherwise uniformly distributed on the $\{-L, \dots, -1\} \cup \{1, \dots, L\}$. We have the following results.

4.1 Connections to Ultra Sub-Gaussianity

For small enough mass at 0 the random variables is ultra sub-gaussian and thus we have khintchine type-inequalities for all even moments.

Theorem 41. *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*

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

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

$$\left(\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^q \right)^{1/q} \leq C_{p,q} \left(\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^p \right)^{1/p} \quad (3)$$

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

We refer readers to [4] for the proof which is a technical nested induction argument.

The expression bounding mass at 0 ρ_0 in terms of L suggests a tradeoff between the size of the support and mass. So we break into cases, treating large and small masses at 0 separately.

4.2 Small mass at 0

First we consider a strong claim for the case with no mass at 0, allowing us to access all $p \geq 3$.

Theorem 42. *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*

$$\left(\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^p \right)^{1/p} \leq C_p \left(\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^2 \right)^{1/2} \quad (4)$$

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

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.

We shall follow an inductive argument exploiting independence based on swapping the X_i one by one with independent Gaussians. We normalize the gaussians to have same variance as the X_i .

Let

$$\sigma = \sqrt{\mathbb{E}|X_1|^2} = \left(\frac{(L+1)(2L+1)}{6} \right)^{1/2} \quad (5)$$

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 (4) 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

$$\mathbb{E}|a + bX_1|^p \leq \mathbb{E}|a + bG_1|^p. \quad (6)$$

This will follow from the following claim.

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

$$\mathbb{E}h(X_1^2) \leq \mathbb{E}h(G_1^2). \quad (7)$$

Indeed, (6) for $b = 0$ is clear. Assuming $b \neq 0$, by homogeneity, (6) 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

$$h_a(x) = |a + \sqrt{x}|^p + |a - \sqrt{x}|^p, \quad x \geq 0 \quad (8)$$

(and similarly for G_1). The convexity of h_a is established in the following standard lemma.

Lemma 43. Let $p \geq 3$, $a \in \mathbb{R}$. Then h_a defined in (8) 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 (5). 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

$$\begin{aligned} f'(a) &= -\sqrt{\frac{2}{\pi\sigma^2}} \int_{\sqrt{a}}^{\infty} e^{-x^2/2\sigma^2} dx - \frac{1}{L} \sum_{k=b+1}^L (-1) \\ &= -\sqrt{\frac{2}{\pi\sigma^2}} \int_{\sqrt{a}}^{\infty} e^{-x^2/2\sigma^2} dx + \frac{L-b}{L}, \quad a \in (b^2, (b+1)^2). \end{aligned} \quad (9)$$

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 (9)) 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 ??), 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

We recall the 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.

Now if we restrict our attention to a random sign with some mass added at 0, we can achieve a Schur convexity type result, which in turn yields khintchine type inequalities.

Theorem 44. 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

$$\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^p \geq \mathbb{E} \left| \sum_{i=1}^n b_i X_i \right|^p. \quad (10)$$

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 45. 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 46. *The best constant D such that the inequality*

$$D \cdot \left[\frac{\phi(a+b) - \phi(b-a)}{2a} - \frac{\phi(a+b) + \phi(b-a)}{2b} \right] \geq \left[\frac{\phi(b)}{b} - \frac{\phi(a)}{a} \right] \quad (11)$$

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 (11) 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 (11) 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 44. Fix $p \geq 3$ and let $F(x) = |x|^p$. Then (10) 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 [?]), 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

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

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

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

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 $0 < a < b$,

$$(\rho_0^{-1} - 1) \left[\frac{\phi(a+b) - \phi(b-a)}{2a} - \frac{\phi(a+b) + \phi(b-a)}{2b} \right] \geq \left[\frac{\phi(b)}{b} - \frac{\phi(a)}{a} \right]. \quad (12)$$

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 45 and 46, inequality (12) 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 (10). \square

Corollary 47. *Under the assumptions of Theorem 44 for every $n \geq 1$ and reals a_1, \dots, a_n , we have*

$$\left(\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^p \right)^{1/p} \leq C_p \left(\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^2 \right)^{1/2} \quad (13)$$

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

Now we consider when ρ_0 is "large".

4.3 Large mass at 0

It turns out large mass at 0 is more amenable to bounds for $p < 2$. Here we consider the case $p = 1$.

Theorem 48. *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*

$$\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right| \geq c_1 \left(\mathbb{E} \left| \sum_{i=1}^n a_i X_i \right|^2 \right)^{1/2} \quad (14)$$

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.

Note that for $a_1 = 1$, $a_2 = \dots = a_n = 0$, we have equality in (14), which explains why the value of the constant c_1 is sharp.

We shall closely follow Haagerup's approach from [?]. 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 [?] 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 [?], 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

$$F(s) \geq F(1), \quad s \geq 1, \quad (15)$$

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 [?], 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 (14) is finished.

It remains to show (15). 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 [?]). 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.

References

- [1] Tkocz T. Khinchin Inequalities with Sharp Constants.

http://www.him.uni-bonn.de/fileadmin/him/Lecture_Notes/Khinchin_inequalities_-_slides_3.pdf
- [2] Haagerup U. The Best Constants in the Khintchine Inequality.

<https://eudml.org/doc/218383>
- [3] Szarek H. On the best constants in the Khinchin's inequality.

<https://eudml.org/doc/218071>
- [4] Havrilla. A, Tkocz. T, Sharp Khinchin-type inequalities for symmetric discrete uniform random variables.

<https://arxiv.org/abs/1912.13345>
- [5] Havrilla .A, Nayar. P, Tkocz. T, Khinchin-type inequalities via Hadamard's factorisation.

<https://arxiv.org/abs/2102.09500>
- [6] Nayar. P, Oleszkiewicz. Khinchine type inequalties with optimal constants via ultra log-concavity.

<https://link.springer.com/article/10.1007/s11117-011-0130-z>
- [7] J.-P. Kahane, Sur les sommes vectorielles, C. R. Acad Sci. Paris 259(1964), 2577-2580
- [8] Ledoux. M, Talagrand. M, Probability in Banach Spaces.

<https://link.springer.com/book/10.1007/978-3-642-20212-4>
- [9] Mordhorst, O. The optimal constants in Khintchine's inequality for the case $2 < p < 3^*$.

<https://arxiv.org/pdf/1601.07850.pdf>
- [10] Nazarov, F. Podkorytov, A. Ball, Haagerup, and Distribution Functions.

https://link.springer.com/chapter/10.1007/978-3-0348-8378-8_21
- [11] Eskenazis, A. Nayar, P. Tkocz, T. Gaussian Mixtures: Entropy and Geometric Inequalities.

<https://arxiv.org/pdf/1611.04921.pdf>
- [12] Latała, R. Oleszkiewicz, K. A note on sums of independent uniformly distributed random variables.

- <https://www.mimuw.edu.pl/~rlatala/papers/cm6826.pdf>
- [13] Barthe, F. Guedon, O. Mendelson, S. Naor, A. A Probabilistic Approach To The Geometry of the l_p^N -Ball.
- <https://arxiv.org/pdf/math/0503650.pdf>
- [14] Szarek, S. On the best constants in the Khinchin inequality.
- <https://eudml.org/doc/218071>
- [15] Young, R. On the Best Possible Constants in the Khintchine Inequality.
- <https://academic.oup.com/jlms/article/s2-14/3/496/2939431?login=true>
- [16] Newman, C. An Extension of Khintchines Inequality.
- <https://projecteuclid.org/journals/bulletin-of-the-american-mathematical-society-new-series/v>
- [17] Newman, C. Moment Inequalities for Ferromagnetic Gibbs Distributions
- https://aip.scitation.org/doi/pdf/10.1063/1.522748?casa_token=Z_1AHe4y5o4AAAAA:r88BmLJsWwn8td
- [18] Baernstein, A. Culverhouse, R. Majorization of sequences, sharp vector Khinchin inequalities, and bisubharmonic functions
- <https://www.impan.pl/pl/wydawnictwa/czasopisma-i-serie-wydawnicze/studia-mathematica/all/152/>
- [19] A. Ya. Khinchin, Über dyadische Brüche, Math. Z. 18 (1923), 109-116
- [20] Whittle, P., Bounds for the moments of linear and quadratic forms in independent random variables. *Theory Probab. Appl.* 5, 302-305 (1960)
- [21] Eskenazis, A., Nayar, P., Tkocz, T., Sharp comparison of moments and the log-concave moment problem. *Adv. Math.* 334 (2018), 389–416.
- [22] Pólya, G., Szegő, G., Problems and theorems in analysis. I. Series, integral calculus, theory of functions. Translated from the German by Dorothee Aeppli. Reprint of the 1978 English translation. Classics in Mathematics. Springer-Verlag, Berlin, 1998.
- [23] Nayar, P., Zwara, S., Sharp variance-entropy comparison for nonnegative gaussian quadratic forms. Preprint (2020), arXiv:2005.11705.