A Survey of Khintchine Type Inequalities for Random Variables

Alex Havrilla

April 30, 2021

Abstract

We survey a collection of Khintchine type inequalities in a great deal of cases and present in detail those 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.

Thesis Supervisor: Tomasz Tkocz
Title: Assistant Professor, Department of Mathematical Sciences

Acknowledgements

My deepest gratitude for the time and guidance Professor Tkocz has given me over the last two years. Working with him has been truly inspirational. I would also like to thank Professors Wesley Pegden and Agoston Pisztora for serving on my thesis committee and Professors Ian Tice and Robert Pego for their excellent classes and guidance.

Contents

1 Introduction

1.1 Overview

The study of Khintchine inequalities, initiated by Aleksandr Khintchine in [15], seeks to find constants $C_{p,q}$ comparing pth and qth moments of linear combinations of independent identically distributed(i.i.d) random variables.

Let $\epsilon_1, ..., \epsilon_n$ be i.i.d random signs defined such that $\mathbb{P}(\epsilon_i = -1) = \mathbb{P}(\epsilon_i = 1) = \frac{1}{2}$ for $i \in \{1, ..., n\}$. The classical Khintchine inequality compares for $0 < p, q < \infty$ the pth and qth moments of sums of these i.i.d random signs with coefficient vector $a \in \mathbb{R}^n$

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

Throughout we write the pth moment of random variable S as $||S||_p = (\mathbb{E}\,|S|^p)^{1/p}$.

Proposition 1.1.1 (Khintchine's Inequalities)**.** *Let S be defined as above. Then for all $p, q > 0$ there exists $C_{p,q} > 0$ such that for all integers $n \geq 0$ and all coefficient vectors $a \in \mathbb{R}^n$ 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 $C_{p,q} = 1$ due to Hölder's inequality and the resulting monotonicity of moments. Note the sharpness of this is verified by the case $n = 1$. However the case $p > q$ is more involved. In general we only know comparisons with sharp constants for arbitrary p and $q = 2$ and for even p, q. Further a natural generalization is to consider sums of random variables other than random signs. We will survey such generalizations to a number of distributions in the introduction and present in depth those with close connections to the new results we present in the last sections.

Besides being of interest for their own sake, such inequalities have shown themselves useful in Banach space theory, where we consider sums of vectors and their norms. The inequalities play a particularly notable role in characterizing Hilbert spaces as those Banach spaces with *type* 2 and *cotype* 2, the definitions of can be viewed as robust versions of the parallelogram identity. For a full presentation of these ideas we refer readers to [18]. Interest in the optimal constants has also been fruitful in the study of convex geomtry. For example the $C_{2,1}$ optimal constant is used in proofs establishing the maximum volume projections of the n-dimensional cross-polytope onto $n-1$ dimensional subspaces(see [3]). Further there is much to be said for understanding these optimal constants on their own as a means of appreciating the background structure making them so. Somehow finding these optimal constants often leads to "deeper understanding" and suggests a "natural" techniques in the course of the proof. Other applications include important results in analysis such as the proof of Littlewood-Paley decomposition and Grothendieck's inequality.

1.2 Some History

The first interesting constants for random signs were found by Whittle in [38] who claimed the $C_{p,2}$ optimal constant for $p > 2$ is $(\mathbb{E}\,|G|^p)^{1/p}$ where G is a standard Gaussian. However his proof is only valid for $p \geq 3$. Interested in the applications in Banach Space theory, Young in [39] proved this again for $p \geq 3$. Independently Stechkin [34] showed optimality of the

$C_{p,2} = (\mathbb{E}\,|G|^p)^{1/p}$ for even integer moments. Szarek in [36] showed a lower bound $C_{2,1} = \sqrt{2}$ for $p = 1$. This left Haagerup to find sharp constants in the remaining cases in [10], addressing $p \in (0, 2)$ and $p \in (2, 3)$ i.e. the "hard regime".

A natural yet powerful technique showing Khintchine type inequalities is showing the Schur convexity of certain functions associated with expectations of classes of random variables. Komorowski in [15] building on the work of Eaton [7] showed expectations of sums of random signs are Schur convex for $p \geq 3$, establishing a Khintchine type inequality. This also naturally treats vector valued random variable generalizations. For example Peskir generalizes results for random signs to complex sums $\sum_{i=1}^n z_i e^{i\epsilon_i}$ (often called Steinhaus random variables) in [28].

Moving to other distributions, Latała and Oleszkiewicz provide a Khintchine type inequality for random variables uniformly distibuted on $[-1, 1]$ with a Schur convexity approach. We cover this is more detail in section 2.1. Culverhouse and Baerstein, inspired by the previous work, obtained in [1] sharp inequalities for independent sums of random vectors uniformly distributed on the unit sphere or unit ball in $\mathbb{R}^n$ via Schur convexity results for expectations of bisubharmonic functions. Towards general comparisons of $C_{p,q}$ where $q \neq 2$ Czerwiński in an unpublished thesis showed sharp constants for random signs when p is divisible by q and both even. In [21] Nayar and Oleskiewicz remove the need for $q \mid p$ while retaining optimal $C_{p,q}$ when both even. Further these results are generalized to the class of ultra sub-Gaussian random variables, which are dicussed in section 2.2.

Yet another generalization is to consider sums of random variables which are not independent but slightly dependent. From [27] Pass and Spektor consider sums of k-wise independent random signs and find bounds on the optimal constants when $k < p$ using interpolation arguments, with stonger results when $k = 2$ and $k = 3$. However we note the results are not optimal. Similar work is continued by Spektor in [33] considering sums of random signs conditioned on summing to 0. In fact this is inspired by results from O'Rourke [26] working in random matrix theory. We conclude by noting the very general results of [16] which achieve $C_{p,2p}$ optimal constants for certain classes of symmetric random variables, notably including the beta distribution, by developing spectral methods and approaches utilizing Poincaré-type inequalities.

For convencience the preceding discussion is summarized in the following list. For definitions of each class of distributions we refer readers to the relevant section or referenced paper.

- Random Signs via [10]. Optimal $C_{2,p}$ and $C_{p,2}$ known for $p \in [0, 2]$ and $p \in [2, \infty)$.
- Uniform distributions on B_q^n for $q \in (0, 2]$ via [8].
- Ultra sub-Gaussian class via [21].
- Type $\mathcal{L}$ class via [12].
- Discrete symmetric uniform class via [11].
- Random variable with exponential densities $e^{-|x|^\alpha}$ with $2 \leq \alpha < \infty$ via [9].
- Steinhaus random variables (uniform on the complex unit circle) via [28].
- Euclidean spheres and balls via [1].

2 Khintchine Inequalities for Various Distributions

This section contains a selected collection of Khintchine type results with optimal constants. These examples are chosen for their illustration of common techniques used in proof of Khintchine type inequalities and their close relation to the new results we present in sections 3 and 4. 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 independent random signs $\epsilon_1, ..., \epsilon_n$. First we establish constants $C_{p,2}$ and $C_{2,p}$ do exist independent of the coefficient vector $a \in \mathbb{R}^n$ such that

$$||S||_p \leq C_{p,2} ||S||_2$$
$$||S||_2 \leq C_{2,p} ||S||_p$$

where $S = \sum a_i \epsilon_i$ for arbitrary coefficient vector $a \in \mathbb{R}^n$. This is a natural starting point since the second moment of sums is easily computable.

$$||S||_2 = \sqrt{\mathbb{E}S^2} = \sqrt{\mathbb{E}\sum_{i=1}^n a_i^2 \epsilon_i^2 + \mathbb{E}\sum_{i \neq j \in [n]} a_i a_j \epsilon_i \epsilon_j} = \sqrt{\sum_{i=1}^n a_i^2}$$

i.e. 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 such that $\sum_i a_i^2 = 1$.

We adopt the convention $A_p = C_{2,p}$ and $B_p = C_{p,2}$. We now establish the existence of constants A_p, B_p independent of $a \in \mathbb{R}^n$.

Theorem 2.1.1 (Khintchine's Inequality,[18])**.** *Let $0 < p < \infty$. Then there exist constants $A_p, B_p > 0$ dependent only on p such that for any $a \in \mathbb{R}^n$ we have*

$$A_p \sqrt{\sum_i a_i^2} \leq ||S||_p \leq B_p \sqrt{\sum_i a_i^2} \tag{1}$$

Proof. Via homogeneity we may suppose $\sum_i a_i^2 = 1$. Recall Bernstein's inequality which tells us $\mathbb{P}\left(|\sum_i a_i \epsilon_i| > t\right) \leq e^{-t^2/2}$. Then with layer cake and Chebyshev we may write:

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

The reverse inequality follows from Hölder when $0 < p < 2$.

$$\begin{aligned} 1 = \mathbb{E}\left(\sum_{i=1}^n \epsilon_i a_i\right)^2 &= \mathbb{E}\left(\left|\sum_{i=1}^n \epsilon_i a_i\right|^{2p/3} \left|\sum_{i=1}^n \epsilon_i a_i\right|^{2-2p/3}\right) \\ &\leq \left(\mathbb{E}\left|\sum_{i=1}^n \epsilon_i a_i\right|^p\right)^{2/3} \left(\mathbb{E}\left|\sum_{i=1}^n \epsilon_i a_i\right|^{6-2p}\right)^{1/3} \leq \left(\mathbb{E}\left|\sum_{i=1}^n \epsilon_i a_i\right|^p\right)^{2/3} B_{6-2p}^{2-2p/3} \end{aligned}$$

□

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 [39]. Note this in fact was first shown by Whittle in [38].

Theorem 2.1.2 (Young, [39]). *For $3 \leq p < \infty$ we have*

$$B_p = 2^{1/2}(\Gamma((p+1)/2)/\sqrt{\pi})^{1/p} = (\mathbb{E}\,|G|)^{1/p}$$

Young's interest in the optimal constant is motivated by arguments from Banach space theory which give a nice application of these Khintchine type inequalities. Crucially Young notes a particular function is nonnegative when $p \geq 3$, which we establish in the following lemma.

Lemma 2.1.3 (Young,[39]). *For $p \geq 3$ and $a, b \in \mathbb{R}^n$ let*

$$f : (a, b, p) \to \left|a - \sqrt{2}b\right|_2^p + 2\,|a|_2^p + \left|a + \sqrt{2}b\right|_2^p - 2\,|a-b|_2^p - 2\,|a+b|_2^p\,.$$

Then $f \geq 0$.

Proof of 2.1.3. Set

$$g_p(x) = \left|1 - \sqrt{2}x\right|^p + 2 + \left|1 + \sqrt{2}x\right|^p - 2\,|1-x|^p - 2\,|1+x|^p$$

Note $g_p(0) = 0$ and $g_p'(0) = 0$. Set $z \in \mathbb{C}$ and

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

When $p \geq 3$ we have $h_p(z, \cdot)$ positive, even and convex and therefore monotonically increasing away from the origin, $g_p''(x) = 2p(p-1)\{h_p(x, \sqrt{2}) - h_p(x,1)\} \geq 0$ so g_p nonnegative on $\mathbb{R}^+$.

Now define

$$k_p(x,y) = g_p(x+iy)$$

Then $k_p(x,0) = g_p(x) \geq 0$ and $k_p'(x,y) = 2py\{h_p(z,\sqrt{2}) - h_p(z,1)\} \geq 0$ and then $f(a,b,p) = k_p(x,y) \geq 0$. Which establishes the claim. □

Now we may finish the proof of optimal constants for $p \geq 3$.

Proof of 2.1.2. Define $c(p) = \sqrt{2}(\frac{\Gamma((p+1)/2)}{\Gamma(1/2)})^{1/p}$ i.e. the pth Gaussian moment. Note first for $p > 0$ we must have both $A_p \leq c(p)$ and $c(p) \leq B_p$. Indeed by the Central Limit Theorem we know $\sum_{i=1}^n \frac{1}{\sqrt{n}}\epsilon_i \to G \sim N(0,1)$ in distribution. And in the limit as $n \to \infty$ the inequality must hold as well. This follows from a standard argument which shows convergence of expectation from convergence in distribution and uniform integrabilty which we outline in Section 2.1 and prove as 5.1.3 in the appendix.

Young's argument proves the claim in more general setting of L_p spaces. We focus the argument on the one-dimensional case while using some of the notation from Young. It suffices to show for coefficient vector sequences $a \in \mathbb{R}^n$ for $n \geq 0$ that

$$\left(\mathbb{E}\left|\sum_{n=1}^{n} a_i\epsilon_i\right|^p\right)^{1/p} \leq c(p)\,||a||_2$$

where we may via homogeneity assume $||a||_2 = 1$. Set for $a \in \mathbb{R}^n$ the sequences $a_i^{(1)} := a_1$ for $1 \leq i \leq n$ and $a_{2i-1}^{(m+1)} := a_{2i}^{(m+1)} = \frac{1}{\sqrt{2}}a_i^{(m)}$ for $1 \leq i \leq n2^{m-1}$. Then for arbitrary m we have $||a||_2 = \left|\left|a^{(m)}\right|\right|_2 = 1$ since we are simply spreading the square mass into two terms. Further by our lemma 2.1.3 we know $m \to \left(\mathbb{E}\left|\sum_{i=1}^{n2^{m-1}} a_i^{(m)}\epsilon_i\right|^p\right)^{1/p}$ is non-decreasing. We see this from writing our expression as

$$\begin{aligned}
\mathbb{E}\left|\sum_{i=1}^{n2^{m-1}} a_i^{(m)}\epsilon_i\right|^p &= \mathbb{E}_{\{\epsilon_2,\ldots,\epsilon_{n2^{m-1}}\}}\mathbb{E}_{\{\epsilon_1,\epsilon_2\}}\left|\sum_{i=1}^{n2^{m-1}} a_i^{(m)}\epsilon_i\right|^p \\
&= \mathbb{E}_{\{\epsilon_2,\ldots,\epsilon_{n2^{m-1}}\}}\frac{1}{2}\left|a_1^{(m)}+\sum_{i=2}^{n2^{m-1}} a_i^{(m)}\epsilon_i\right|^p + \frac{1}{2}\left|-a_1^{(m)}+\sum_{i=2}^{n2^{m-1}} a_i^{(m)}\epsilon_i\right|^p \\
&= \mathbb{E}_{\{\epsilon_3,\ldots,\epsilon_{n2^{m-1}}\}}\frac{1}{4}\left|a_1^{(m)}+a_2^{(m)}+\sum_{i=3}^{n2^{m-1}} a_i^{(m)}\epsilon_i\right|^p + \frac{1}{4}\left|-a_1^{(m)}+a_2^{(m)}+\sum_{i=3}^{n2^{m-1}} a_i^{(m)}\epsilon_i\right|^p \\
&+ \mathbb{E}_{\{\epsilon_3,\ldots,\epsilon_{n2^{m-1}}\}}\frac{1}{4}\left|a_1^{(m)}-a_2^{(m)}+\sum_{i=3}^{n2^{m-1}} a_i^{(m)}\epsilon_i\right|^p + \frac{1}{4}\left|-a_1^{(m)}-a_2^{(m)}+\sum_{i=3}^{n2^{m-1}} a_i^{(m)}\epsilon_i\right|^p \\
&= \mathbb{E}_{\{\epsilon_3,\ldots,\epsilon_{n2^{m-1}}\}}\frac{1}{4}\left|\frac{2}{\sqrt{2}}a_1^{(m-1)}+\sum_{i=3}^{n2^{m-1}} a_i^{(m)}\epsilon_i\right|^p + \frac{1}{4}\left|-\frac{2}{\sqrt{2}}a_1^{(m-1)}+\sum_{i=3}^{n2^{m-1}} a_i^{(m)}\epsilon_i\right|^p \\
&+ \frac{1}{2}\mathbb{E}_{\{\epsilon_3,\ldots,\epsilon_{n2^{m-1}}\}}\left|\sum_{i=3}^{n2^{m-1}} a_i^{(m)}\epsilon_i\right|^p
\end{aligned}$$

from which we take a difference with the $(m-1)$ sequence and apply the lemma to pointwise. Notice that as $m \to \infty$ via the Central Limit Theorem $\sum_{i=1}^{n2^{m-1}} a_i^{(m)}\epsilon_i$ goes in distribution to a standard Gaussian. Hence via non-decreasingness we have

$$\mathbb{E}\left|\sum_{i=1}^{n2^{m-1}} a_i^{(m)}\epsilon_i\right|^p \leq \mathbb{E}\,|G|^p$$

where G is a standard Gaussian and we are done. □

The most important takeaway from the above proof is the crucial nonnegativity of the "smoothing function" f which arises as a result of the square mass decomposition of the sequence. Often similar results on inequalities for simple functions such as these prove useful, for example particularly in Section 4. However this nice property is lost for $2 < p < 3$ and so some more technical arguments are required. In [10] Haagerup concluded the search for random signs by treating $0 < p < 2$ and $2 < p < 3$. He established the constants

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

and

$$B_p = \begin{cases} 1 & 0 < p \leq 2 \\ 2^{1/2}(\Gamma((p+1)/2)/\sqrt{\pi})^{1/p} & 2 < p < \infty \end{cases} \tag{3}$$

where p_0 is the solution to $\Gamma((p+1)/2) = \sqrt{\pi}/2$ in $[1,2]$. Szarek showed in [36] that $A_1 = 1/\sqrt{2}$. Haagerup extended this result for $0 < p < p_0$ and 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 $\sum_i \frac{1}{\sqrt{n}}\epsilon_i \to G \sim N(0,1)$ which achieves B_p in the limit, whereas $2^{1/2-1/p}$ is achieved by the sum of two random signs.

Haagerup's original argument is quite technical so we instead elect to present a proof via Nazarov, Podkorytov [22] which relies on crucial elements of Haagerups proof while using distribution functions to simplify the more technical details. Note however Nazarov and Podkorytov only treat the $0 < p < 2$ case. In [20] Mordhorst treats $2 < p < 3$ using the same technique.

In preparation we state a lemma on distribution functions.

Lemma 2.1.4 (Nazarov and Podkorytov, [22]). *Let $Y > 0$, $f, g : \mathcal{M} \to [0, Y]$ be any two measurable functions on $(\mathcal{M}, \mu)$. 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 there exists unique y_0 such that $F_* - G_* = 0$. 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*

$$\phi(s) = \frac{1}{s y_0^s} \int_{\mathcal{M}} f^s - g^s d\mu$$

is monotone increasing on S.

We include the proof in Section 2 of the appendix.

Proof of 2 and 3. First we reduce the case $0 < p < p_0$ to $p = p_0$. Let $S = \sum_{i=1}^n a_i \epsilon_i$ and suppose $\mathbb{E}|S|^p \geq 2^{p_0-2/2}$ where $\sum_{i=1}^n a_i^2 = 1$. Via Hölder 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

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

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,

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

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 $\lim_{s\to\infty} I_p(s)$ we may apply DCT to see get $C_p \int_0^\infty (1 - e^{-u^2/2}) \frac{du}{u^{p+1}}$ as well $\lim_{s\to\infty} I_p(s) = 2^{p-2/2} \frac{\Gamma(p+1/2)}{\Gamma(3/2)}$ via a Central Limit Theorem arugment. Then showing $I_p(s) \geq I_p(\infty)$ would allow us to conclude. So 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 a rather intricate which we omit.

After this it remains to treat the case of large coefficient, i.e. one of the $a_j^2 > \frac{1}{2}\sum_{k=1}^n a_k^2$ for some $1 \leq j \leq n$. We must show

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

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

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

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

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

however ϕ_p is concave for $p < 2$. So instead we find a function Φ_p that dominates ϕ and convex on $(0, 1)$. We then 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 without loss of generality 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 $R_p(1, a_2) \geq A_p\Phi_p(a_2^2)$. Without loss of generality suppose $a_1 = 1$ is largest of coefficients. It is enough to prove the following pointwise estimate.

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

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 Central Limit Theorem arguments we know often 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 $p \geq 3$, whereas the $2 < p < 3$ case often remains less accessible. This concludes our discussion of Khintchine type inequalities for random signs. We move on to discussing other distributions.

2.2 Uniform Random Variables

Here we present results from [17] leading to Khintchine-type results for uniform variables U_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.

Before stating the result 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

$$\sum_{i=1}^n a_i = \sum_{i=1}^n b_i \qquad \text{and} \qquad \sum_{i=1}^k a_i^* = \sum_{i=1}^k b_i^* \;\text{ for all }\; k = 1, \ldots, n,$$

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, $(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}) \prec (a_1, a_2, \ldots, a_n) \prec (1, 0, \ldots, 0)$ for every nonnegative sequence (a_i) with $\sum_{i=1}^n a_i = 1$. A function $\Psi\colon [0, \infty)^n \to \mathbb{R}$ 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 $\psi\colon [0, +\infty) \to \mathbb{R}$ being convex is Schur convex. Now we may state the result.

Theorem 2.2.1 (Latała and Oleszkiewicz, [17])**.** *Let $a = (a_1, ..., a_n)$ and $b = (b_1, ..., b_n)$ be two sequences of real numbers such that $(a_1^2, ..., a_n^2) \prec (b_1^2, ..., b_n^2)$ and $U_1, ..., U_n$ be a sequence of independent random variables uniformly distributed on $[-1, 1]$. Then*

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

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

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

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 2.2.2. *A real random variable X is symmetric unimodal if and only if there exists a probability measure μ on $[0, \infty)$ such that density g of X is*

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

Proof of 2.2.2. 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

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

so μ is a probability measure. □

Lemma 2.2.3. *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 $X = \sum_{i=1}^n a_iU_i$ where U_i are independent and uniformly distributed on $[-1,1]$ then X symmetric unimodal.*

Proof of 2.2.3. We 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 our Lemma representation. Clearly g is symmetric and nonincreasing on $[0,\infty)$. □

We now present a key technical lemma.

Lemma 2.2.4.

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

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 [17]. We present the final lemma which shall directly imply the desired result.

Lemma 2.2.5. *If U_1, U_2, U_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*

$$\mathbb{E}|U_1 + aU_2 + bU_3|^p \leq \mathbb{E}|U_1 + cU_2 + dU_3|^p \qquad p \in [1,2]$$
$$\mathbb{E}|U_1 + aU_2 + bU_3|^p \geq \mathbb{E}|U_1 + cU_2 + dU_3|^p \qquad p \geq 2$$

Proof of 2.2.5. 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 then gives

$$\mathbb{E}|U_1 + aU_3 + bU_3|^p = \frac{1}{8}\int_{-1}^1\int_{-1}^1\int_{-1}^1 |x_1 + ax_2 + bx_3|^p dx_1dx_2dx_3 =$$
$$c_p\Big(\frac{(a+b+1)^3|a+b+1|^p + (a-b-1)^3|a-b-1|^p}{ab} - \frac{(a-b+1)^3|a-b+1|^p + (a+b-1)^3|a+b-1|^p}{ab}\Big)$$

where $c_p = \frac{1}{4(p+1)(p+2)(p+3)}$

Setting $k = a^2 + b^2$ and $s = 2ab$ we then reduce $f(s) = \mathbb{E}|U_1 + aU_2 + bU_3| = 2c_p\frac{g(s)}{s}$ after appropriate substitutions. We can show f nondecreasing for $p \geq 2$ and nonincreasing if $p \in [1,2]$. This follows from direct computation. □

This allows us to prove a corollary giving us the theorem for free.

Corollary 2.2.6. *If X, U_1, U_2 are independent random variables, U_1, U_2 are uniformly distibuted on $[-1,1]$ and X symmetric unimodal with $a^2+b^2 = c^2+d^2$ and $c \geq a \geq b \geq d$ then*

$$\mathbb{E}|X+aU_1+bU_2|^p \leq \mathbb{E}|X+cU_1+dU_2|^p \qquad p\in[1,2]$$
$$\mathbb{E}|X+aU_1+bU_2|^p \geq \mathbb{E}|X+cU_1+dU_2|^p \qquad p\geq 2$$

Proof of 2.2.6. Let g be the density of X and μ the corresponding measure via our first 2.2.2. We have

$$\begin{aligned}\mathbb{E}|X+aU_1+bU_2|^p &= \int_{-\infty}^{\infty}\mathbb{E}|x+aU_1+bU_2|^p g(x)dx = \int_0^{\infty}\frac{1}{2s}\int_{-s}^{s}\mathbb{E}|t+aU_1+bU_2|^p dt d\mu(s)\\ &= \int_0^{\infty}\mathbb{E}|sU_3+aU_1+bU_2|^p d\mu(s) \leq \int_0^{\infty}\mathbb{E}|sU_3+cU_1+dU_2|^p d\mu(s) = \mathbb{E}|X+cU_1+dU_2|^p\end{aligned}$$

where we get the inequality from our last lemma. □

Now we finish the theorem.

Proof of 2.2.1. Via a lemma from [19] it suffices to prove inequalities in the case $a_i^2 = b_i^2$ for $i \neq j,k$ and $a_j^2 = tb_j^2 + (1-t)b_k^2$ with $a_k^2 = tb_k^2+(1-t)b_j^2$. Via symmetry $a_i, b_i \geq 0$. So our proposition follows directly from the corollary by setting $X = \sum_{i\neq j,k} a_iU_i$. □

Note the Schur concavity result immediately shows optimal constant Khintchine inequalities via a standard argument using the Central Limit Theorem. Indeed for the case $p \in [1,2]$ we have

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

and a reversal of the inequality for $p \geq 2$. Via Central Limit Theorem we know $\sum_{i=1}^n \frac{1}{\sqrt{n}}U_i \to g \sim N(0,1)$ is distribution. Then we have a convergence of the moments via the following lemma. Further $\mathbb{E}\left|\sum_{i=1}^n \frac{1}{\sqrt{n}}U_i\right|^p$ is decreasing in n since $(\frac{1}{\sqrt{n-1}},...,\frac{1}{\sqrt{n-1}},0)$ majorizes $(\frac{1}{\sqrt{n}},...,\frac{1}{\sqrt{n}})$.

Lemma 2.2.7. *Suppose $X_n \to X$ in distribution. If $\{X_n\}$ is uniformly integrable then $\mathbb{E}|X| < \infty$ and $\mathbb{E}(X_n) \to \mathbb{E}(X)$ and $\mathbb{E}|X_n| \to \mathbb{E}|X|$. Recall a sequence $\{X_n\}$ uniformly integrable if $\sup_n |X_n| < \infty$ and for all $\epsilon > 0$ we have $\delta > 0$ such that when for some event A $P(A) < \delta$ then $P(|X_n| \in A) < \epsilon$.*

For a proof we refer readers to the appendix. 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 [21] which significantly generalizes the class of random variables considered and achieves comparisons for all even integer moments via log-concavity. This is particularly notable for us since for much time only comparisons to the 2nd moment were known, and often for narrow classes of distributions. Via Ultra Sub-Gaussianity, Nayar and Oleszkiewicz manage to generalize both the class considered and the range of moments compared, albeit only for even integer.

Recall a sequence $(a_i)_{i=0}^{\infty}$ of non-negative real numbers is called *log-concave* if it is supported on an interval and for all i we have $a_i^2 \geq a_{i-1}a_{i+1}$. Then we say $\mathbb{R}^n$-valued X is an *Ultra sub-Gaussian* random variable 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. Note $||\cdot||$ denotes the euclidean norm.

The results crucially use Walkup's Theorem which establishes the preservation of log-concavity under binomial convolution.

Theorem 2.3.1 (Walkup,[37])**.** *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 $(c_n)_{n=0}^{\infty}$ is log-concave

We say $(a_i)_{i=0}^{\infty}$ is *ultra log-concave* if and only if $(i!a_i)_{i=0}^{\infty}$ is log-concave. Walkup's theorem tells us the collection of ultra log-concave sequences is closed under convolution. Given X Ultra Sub-Gaussian we can extract a Khintchine type inequality of the following form.

Theorem 2.3.2 (Nayar and Oleskiewicz,[21])**.** *Let n,d positive integers and $p > q \geq 2$ even integers. If $X_1, ..., X_n$ are independent $\mathbb{R}^d$ 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 $S = \sum_{i=1}^{n} X_i$

The proof rests on lemmas we highlight now. The first shows ratios of expecatations of n-dimensional Ultra Sub-Gaussian random variables against Gaussians are the same as in the 1-dimensional case. So once we have argued in the 1-dimensional case we can easily tensorize our results.

Lemma 2.3.3. *Let $\Pi : \mathbb{R}^d \to \mathbb{R}$ be projection to first coordinate. For $p > 0$ assume X rotation invariant random vector on $\mathbb{R}^d$ with finite pth moment. Then where $G \sim N(0, Id_d)$*

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

Proof of 2.3.3. Let θ be a random vector uniformly distributed on the Euclidean unit sphere $(\mathbb{R}^d, ||\cdot||)$ and independent of X. Since X is rotation invariant it has the same distribution as $||X|| \cdot \theta$ and thus ΠX has the same distribution as $||X|| \cdot \Pi\theta$. So we compute

$$\mathbb{E}\,|\Pi X|^p = \mathbb{E}\,||X||^p\,\mathbb{E}\,|\Pi\theta|^p$$

and also

$$\mathbb{E}\,|\Pi G|^p = \mathbb{E}\,||G||^p\,\mathbb{E}\,|\Pi\theta|^p$$

so taking the ratio we are done. ∎

Next we show the class of Ultra Sub-Gaussian random variables is closed under independent sums.

Lemma 2.3.4. *If X, Y are Ultra Sub-Gaussian and independent random vectors then $X+Y$ is Ultra Sub-Gaussian.*

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

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

and then noting

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

we conclude with Walkup's theorem. □

We are now ready to present the proof of Khintchine-type inequalities for all even moments of Ultra Sub-Gaussian random variables.

Proof of 2.3.2. Recall $S = \sum_{i=1}^{n} X_i$ i.e. a sum of Ultra Sub-Gaussian random variables. So then by repeated application of Lemma 2.3.4 we know S is Ultra Sub-Gaussian. But then we know the sequence $a_i = \mathbb{E}\,||S||^{2i}/\mathbb{E}\,||G||^{2i}$, where G is a standard multi-dimensional Gaussian, is log-concave. In turn we know $a_k^{2k} \geq a_{k-1}^k a_{k+1}^k$ for $k \geq 1$ therefore the sequence $(a_s^{1/s})_{s=1}^{\infty}$ is non-increasing. So in particular $a_{p/2}^{2/p} \leq a_{q/2}^{2/q}$ which directly implies the desired result.

Note the constant is optimal via the standard Central Limit Theorem argument. For unfamiliar readers we refer to 5.1.3 in the appendix. □

These results in turn relate to our discussion of type $\mathcal{L}$ random variables, where in section 3 we discover every Type $\mathcal{L}$ random variable is Ultra Sub-Gaussian, and classes of symmetric discrete uniform random variables with sufficiently small mass at 0.

2.4 Gaussian Mixtures

Here we look at results from [8] addresing the class of Gaussian Mixtures. Recall a random variable X is a *Gaussian mixture* if there exists a positive random variable Y and standard Gaussian Z independent from Y such that X has the same distribution as YZ. We note the 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.

First we present a characterization of densities of Gaussian mixture measures as completely monotonic functions when composed with square root as a means of identifying Gaussian-mixtures.

Theorem 2.4.1 (Eskenazis, Nayar and Tkocz, [8])**.** *A symmeric random variable X with density f is a Gaussian mixture if and only if $x \to f(\sqrt{x})$ is completely monotonic for $x > 0$.*

The proof of this relies on the classical Bernstein's theorem which characterizes every completely monotonic function as the Laplace transform of a non-negative Borel measure.

Theorem 2.4.2 (Bernstein)**.** *A $C^\infty(\mathbb{R}^n)$ function $g : (0,\infty) \to \mathbb{R}$ is completely monotonic, ie. $(-1)^n g^{(n)} \geq 0$, if and only if there exists non-negative Borel measure μ on $[0,\infty)$ such that*

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

Now we give a proof of Theorem 2.4.1.

Proof of Theorem 2.4.1. We begin by establishing some facts about Gaussian mixtures. Let X be such a mixture which has the same distribution as YZ with Y positive and Z an independent standard Gaussian. Let ν be the law of Y.

First note X is symmetric as Z is symmetric. For Borel set $A \subseteq \mathbb{R}$ we have

$$\mathbb{P}(X \in A) = \mathbb{P}(YZ \in A) = \int_0^\infty \mathbb{P}(yZ \in A) d\nu(y) = \int_A \int_0^\infty \frac{1}{\sqrt{2\pi}y} e^{-\frac{x^2}{2y^2}} d\nu(y) dx \tag{4}$$

which gives X the density

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{x^2}{2y^2}} \frac{d\nu(y)}{y}. \tag{5}$$

Now let X be a symmetric random variable with density f such that $x \to f(\sqrt{x})$ is completely monotonic. Thus by Bernstein's theorem we can find a non-negative Borel measure μ on $[0,\infty)$ with

$$f(\sqrt{x}) = \int_0^\infty e^{-tx} d\mu(t). \tag{6}$$

Then for $A \subseteq \mathbb{R}$ we have

$$\mathbb{P}(X \in A) = \int_A \int_0^\infty e^{-tx^2} d\mu(t) dx = \int_0^\infty \int_A e^{-tx^2} dx d\mu(t) \tag{7}$$

$$= \int_0^\infty \int_{\sqrt{2t}A} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \sqrt{\frac{\pi}{t}} d\mu(t) = \int_0^\infty \gamma_n(\sqrt{2t}A) d\nu(t) \tag{8}$$

where we interchange order of integration via Tonelli's theorem. Note $d\nu(t) = \frac{\sqrt{\pi}}{\sqrt{t}} d\mu(t)$. Setting $A = \mathbb{R}$ we see ν is a probability measure.

Let ν distributed according to ν. Set $Y = \frac{1}{\sqrt{2V}}$ and Z be a standard Gaussian random variable indepenent from Y. We can compute via (7).

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

So X has the same distribution as YZ and is therefore a Gaussian mixture.

In the converse direction we use 5 and Bernstein. Suppose X is a Gaussian mixture. Then we know the density f is of the form

$$f(\sqrt{x}) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{x}{2y^2}} \frac{d\nu(y)}{y}$$

for some non-negative probability measure ν. Bernstein then finishes the proof. □

We now state the main Schur convexity result leading to Khintchine-type inequalities for Gaussian mixtures.

Theorem 2.4.3 (Eskenazis, Nayar and Tkocz, [8]). *Let X be a Gaussin mixture and $X_1, ..., X_n$ be independent copies of X. For two vectors $a, b \in \mathbb{R}^n$ with $p \geq 2$ we have*

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

To show this we must establish a result from on the Schur convexity of certain functions.

Proposition 2.4.4. *Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a convex function, symmetric under permutations of its n arguments. Let $X_1, ..., X_n$ be interchangeable random variables i.e. those whose joint distribution is invariant under permuation of coordinates. Then for $a, b \in \mathbb{R}^n$ we have*

$$a \prec b \implies \mathbb{E}\phi(a_1 X_1, ..., a_n X_n) \leq \mathbb{E}\phi(b_1 X_1, ..., b_n X_n) \tag{9}$$

So $a = (a_1, ..., a_n) \to \mathbb{E}\phi(a_1 X_1, ..., a_n X_n)$ is Schur convex.

We present the proof for Theorem 2.4.3.

Proof of Theorem 2.4.3. Fix $p > -1$ with $p \neq 0$. Let X be a Gaussian mixture and $X_1, ..., X_n$ be independent copies of X. Let X_i have same distribution as $Y_i Z_i$ which are i.i.d copies of a non-negative random variable Y and standard gaussian Z. Via independence we have

$$\mathbb{E}|\sum_{i=1}^{n} a_i X_i|^p = \mathbb{E}|\sum_{i=1}^{n} a_i Y_i Z_i|^p = \mathbb{E}|(\sum_{i=1}^{n} a_i^2 Y_i^2)^{1/2}|^p = \gamma_p^p \mathbb{E}|\sum_{i=1}^{n} a_i^2 Y_i^2|^{p/2}$$

where $\gamma_p = (\mathbb{E}|Z|^p)^{1/p}$.

Then since $t \to t^{p/2}$ is convex for $p \in (-1, 0) \cup [2, \infty)$ and concave for $p \in (0, 2)$ we can apply our lemma and be done.

□

As we come to expect from Schur convexity results this leads easily to a Khintchine type inequality.

Corollary 2.4.5 (Eskenazis, Nayar and Tkocz, [8]). *Let X be a Gaussian mixture and $X_1, ..., X_n$ be independent copies of X. Then, for every $p \in (-1, \infty)$ and $a_1, ..., a_n \in \mathbb{R}$ we have*

$$A_p \left|\left|\sum_{i=1}^{n} a_i X_i\right|\right|_2 \leq \left|\left|\sum_{i=1}^{n} a_i X_i\right|\right|_p \leq B_p \left|\left|\sum_{i=1}^{n} a_i X_i\right|\right|_2 \tag{10}$$

where

$$A_p = \begin{cases} \frac{||X||_p}{||X||_2} & p \in (-1, 2) \\ \gamma_p & p \in [2, \infty) \end{cases} \qquad B_p = \begin{cases} \gamma_p & p \in (-1, 2) \\ \frac{||X||_p}{||X||_2} & p \in [2, \infty) \end{cases} \tag{11}$$

where $\gamma_p = \sqrt{2}(\frac{\Gamma(\frac{p+1}{2})}{\sqrt{\pi}})^{1/p}$ is the pth moment of a standard Gaussian random variable. Further these constants are sharp

Proof of 2.4.5. Without loss of generality assume $(a_1, ..., a_n)$ is unit norm. Let $p \geq 2$. Schur convexity gives us

$$\left|\left|\frac{X_1 + ... + X_n}{\sqrt{n}}\right|\right|_p \leq \left|\left|\sum_{i=1}^{n} a_i X_i\right|\right|_p \leq ||X_1||_p \tag{12}$$

Central Limit Theorem then implies

$$\gamma_p \left|\left|X\right|\right|_2 \leq \left|\left|\sum_{i=1}^{n} a_i X_i\right|\right|_p \leq \left|\left|X\right|\right|_p$$

where we implicitly use convergence of expectations since moments are uniformly bounded. Notice taking $a_1 = ... = a_{n-1}^{-1/2}$ and $a_n = 0$ we have decreasingness in n. □

Using results from [2] as a corollary we have a result on for random vectors distributed on the n-dimensional closed unit ball in the qth norm $B_q^n = \{x \in \mathbb{R}^n : \sum_{i=1}^n |x_i|^q \leq 1\}$ for $q \in (0, 2]$.

Corollary 2.4.6 (Eskenazis, Nayar and Tkocz, [8])**.** *Fix $q \in (0,2]$ and let $X = (X_1, ..., X_n)$ a random vector uniformly distributed on B_q^n. For two vectors $(a_1, ..., a_n)$ and $(b_1, ..., b_n)$ in $\mathbb{R}^n$ with $p \geq 2$ we have*

$$(a_1^2, ..., a_n^2) \prec (b_1^2, ..., b_n^2) \implies ||\sum^n a_i X_i||_p \leq ||\sum^n b_i X_i||_p$$

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

Proof of Corollary 2.4.6. First we recall some probabilistic results about B_q^n. Let $Y_1, ..., Y_n$ be i.i.d random variables distributed according to μ_q and write $Y = (Y_1, ..., Y_n)$. Let $S = (\sum_{i=1}^n |Y_i|^q)^{1/q}$. Let $\mathcal{E}$ be an exponetial random variable independent of the Y_i. Then by a result from [2] the random vector

$$\left(\frac{Y_1}{(S^q + \mathcal{E})^{1/q}}, ..., \frac{Y_n}{(S^q + \mathcal{E})^{1/q}}\right)$$

is uniformly distributed on B_q^n. Further a result from [31] and independently [32] establishes that S and $\frac{Y}{S}$ are independent.

Let $X = (X_1, ..., X_n)$ be a random vector uniformly distributed on B_q^n. Let $Y_1, ..., Y_n, S$ and $\mathcal{E}$ be as above. We compute using our representation and the independence S and $\frac{Y}{S}$

$$\mathbb{E}\left|\sum_{i=1}^n a_i X_i\right|^p = \mathbb{E}\left|\frac{1}{(S^q + \mathcal{E})^{1/q}} \sum_{i=1}^n a_i Y_i\right|^p = \left|\frac{S}{(S^q + \mathcal{E})^{1/q}}\right|^p \mathbb{E}\left|\sum_{i=1}^n a_i \frac{Y_i}{S}\right|^p$$

Again via independence we have

$$\mathbb{E}\left|\sum_{i=1}^n a_i X_i\right|^p = \frac{1}{\mathbb{E}\left|S\right|^p}\mathbb{E}\left|\frac{S}{(S^q + \mathcal{E})^{1/q}}\right| \mathbb{E}\left|\sum_{i=1}^n a_i Y_i\right|^p = c(p,q,n)\mathbb{E}\left|\sum_{i=1}^n a_i Y_i\right|^p \tag{13}$$

where the constant $c(p, q, n)$ is independent of the vector of coefficients a. Then because the $Y_1, ..., Y_n$ are i.i.d Gaussian mixtures and we have a Schur convexity result for these, we are done. Notice this also readily implies Khintchine inequalites as above. □

3 Type $\mathcal{L}$ Random Variables

Here we cover results from [12] giving new types of Khintchine inequalities for Type $\mathcal{L}$ random variables. First defined by Newman in [24], we say that a random variable X is of type $\mathcal{L}$ if we have constants $A, B \in \mathbb{R}$ such that $|\mathbb{E}e^{zX}| \leq Ae^{B|z|^2}$ and the characteristic function of X is even with strictly real zeroes or equivalently $\mathbb{E}e^{zX}$ is even with pure imaginary zeroes. If the evenness assumption is broken we say instead $X \in \mathcal{L}'$.

Note $\mathcal{L}$ is closed under sums since the characteristic function of the sum $X+Y$ is the product of the characteristic functions of X and Y, which preserves the pure imaginary zeroes. Basic examples of type $\mathcal{L}$ random variables include random signs, arithmetic progressions, uniform distributions on symmetric intervals, and the Gaussian. To see this note all have characteristic functions which dominated by the square exponentials and have strictly real zeroes.

3.1 Connections to Ultra Sub-Gaussianity

Newman showed in [24] a class of Khintchine inequalities for the second and even moments.

Theorem 3.1.1 (Newman, [24])**.** *If the $\{X_j\}_{j=1}^N$ are independent random variables of type $\mathcal{L}$, 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 importantly uses Hadamard's factorization theorem for the characteristics of our random variables, allowing us to compare moments of with terms in the expontial power series. We state this now for completeness. For proof we refer the reader to [35] chapter 5.

Theorem 3.1.2 (Hadamard)**.** *Suppose f is entire and has growth order ρ_0. Let k be the integer so that $k \leq \rho_0 < k+1$. If $a_1, a_2, \ldots$ denote the (non-zero) zeros of f, then*

$$f(z) = e^{P(z)} z^m \prod_{n=1}^{\infty} E_k(z/a_n)$$

where P is a polynomial of degree $\leq k$ and m is the order of the zero of f at $z = 0$. Note $E_k(z)$ is the canonical factor defined by

$$\begin{aligned} &E_0(z) = 1 - z \\ &E_k(z) = (1-z)e^{z + z^2/2 + \ldots + z^k/k}, \quad k \geq 1 \end{aligned}$$

We generalize this approach via Hadamard factorization below.

Theorem 3.1.3 (Havrilla, Nayar and Tkocz, [12])**.** *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 of 3.1.3. First note if $X \in \mathcal{L}'$ then we can find $c \in \mathbb{R}$ such that $X + c \in \mathcal{L}$. We have

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

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 $\mathcal{L}$ as it is the sum of two type $\mathcal{L}$ random variables. So X of type $\mathcal{L}'$ has moments equal to some type $\mathcal{L}$ random variable $Y + c\epsilon$ and so it suffices to simply consider $X \in \mathcal{L}$.

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 via Walkup's theorem 2.3.1 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 i.e.

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

where we approximate the infinite sequence $\alpha_1, ...,$ via a finite truncation $\alpha_1, ..., \alpha_n$ and see the inequality holds in the limit for the σ_k. This is exactly what is needed, since clearly then $k!\sigma_k$ also log-concave. □

Corollary 3.1.4. *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 3.1.5. *Let $(H, \|\cdot\|)$ be a separable (real or complex) Hilbert space. If $X_1, \ldots, X_n$ are independent type $\mathcal{L}$ random variables, then for every vectors $v_1, \ldots, v_n$ 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 $\|X\|_p = (\mathbb{E}\|X\|^p)^{1/p}$.*

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 [25].

Corollary 3.1.6. *Let $\mu_1, ..., \mu_n$ be Borel probability measures on $\mathbb{R}$, each one of type $\mathcal{L}$. Suppose that $(X_1, \ldots, X_n)$ is a random vector in $\mathbb{R}^n$ whose law ρ on $\mathbb{R}^n$ is of the form*

$$d\rho(x_1, \ldots, x_n) = Z^{-1} \exp\left(\sum_{j=1}^{n} h_j x_j + \sum_{j,k=1}^{n} J_{jk} x_j x_k \right) d\mu_1(x_1) \ldots d\mu_n(x_n) \qquad (14)$$

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 $a_1, \ldots, a_n$, the sum $X = \sum_{j=1}^{n} a_j X_j$ satisfies (3.1.3) for all positive even integers $p \leq q$.

Proof of 3.1.6. Let $(X_1, \ldots, X_n)$ be a random vector with distribution given by (14) and let ε be an independent Rademacher random variable. By the previous theorem, the vector $(Y_0, Y_1, \ldots, Y_n) = (\varepsilon, \varepsilon X_1, \ldots, \varepsilon X_n)$ has distribution ρ' of the form

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

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 (14) with $h \equiv 0$. Therefore, by Theorem 2 from [24], for every $a_0, a_1, \ldots, a_n \geq 0$, 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 (3.1.3). Hence, taking $a_0 = 0$ yields that $\sum_{j=1}^{n} a_j X_j$ also satisfies (3.1.3). □ □

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

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

Lemma 3.2.1. *For every $b > 0$ and $a_1, a_2... \geq 0$ 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 of 3.2.1. 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. □

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 3.2.2 (Havrilla, Nayar and Tkocz, [12])**.** *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 $\sigma = \sqrt{\mathrm{Var}(X)}$. 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 3.2.3. *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 of 3.2.3. By a direct computation,

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

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

Lemma 3.2.4. *Let $X_1, X_2, \ldots$ be i.i.d. copies of Z_1 and let Y be a symmetric random variable independent of the X_j. Then the function*

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

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

Proof of 3.2.4. We use the technique of interlacing densities (see, e.g. [9] 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

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

where g_λ is as in Lemma 3.2.3. For arbitrary α, β, $\int(\alpha x^2 + \beta)(g_{\lambda_2(x)} - g_{\lambda_1}(x))\mathrm{d}x = 0$, so the desired inequality is equivalent to

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

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. □

Now we present the proof the theorem.

Proof of 3.2.2. 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

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

where we know $Var(Z_1) = 3$.

By the lemma we know with Schur-concavity

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

Sending $n \to \infty$ 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. □

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

3.3 Examples of Type $\mathcal{L}$ Random Variables

We list some basic 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 $\frac{1}{2}p_0\le p_1\le\cdots\le p_n$, 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 [29]). 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 $\mathbb{P}(X=0)\le\frac{1}{n+1}$, 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 $1/w_0$. For instance, if for some $\alpha\ge 1$,

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

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, known as Polyá's criterions, implies that X is of $\mathcal{L}$. See [29]

(i) f is nondecreasing on $(0,1)$.

(ii) f is C^2 with $f'<0$ and $f''<0$ on $(0,1)$.

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

(iii) $f(t)=(2\pi)^{-1/2}e^{-t^2/2}(1-b+bt^2)$, $0\le b\le 1$.

Condition (i) is justified again by the Eneström-Kakeya theorem combined with a limit argument (see, e.g. Problem III.205 in [29]), (ii) is due to Pólya (see, e.g. Problem V.173 in [30]), (iii) is justified by a direct computation of the moment generating function which is $(1+bz^2)e^{z^2/2}$. Moreover, if the density of X is of the form $f(t)=\text{const}\cdot e^{-|t|^\alpha}$ with $\alpha\ge 2$, $\alpha\notin\{2,4,\dots\}$, then its characteristic function has infinitely many non-real zeros, in particular X is not of type $\mathcal{L}$ (see the solution of Problem V.171 in [30]).

4 Discrete Symmetric Distributions

We now consider the generalization of a random sign to a symmetric discrete random variables uniformly distributed outside 0. I.e. consider the generalization to $X \in \{-L, , .., 0, ..., L\}$ with some mass $\mathbb{P}(X = 0) = \rho_0$ and otherwise uniformly distributed on the set $\{-L, ..., -1\} \cup \{1, ..., L\}$. 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 4.1.1 (Havrilla and Tkocz, [11])**.** *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*

$$\rho_0 \leq 1 - \frac{2}{5}\frac{3L^2 + 3L - 1}{(L+1)(2L+1)}. \tag{15}$$

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

$$\left(\mathbb{E}\left|\sum_{i=1}^{n} a_i X_i\right|^q\right)^{1/q} \leq C_{p,q}\left(\mathbb{E}\left|\sum_{i=1}^{n} a_i X_i\right|^p\right)^{1/p} \tag{16}$$

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

We refer readers to [11] 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. Note in fact this class of random variables is type $\mathcal{L}$.

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 4.2.1. *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*

$$\left(\mathbb{E}\left|\sum_{i=1}^{n} a_i X_i\right|^p\right)^{1/p} \leq C_p\left(\mathbb{E}\left|\sum_{i=1}^{n} a_i X_i\right|^2\right)^{1/2} \tag{17}$$

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 $a_1 = \ldots = a_n = \frac{1}{\sqrt{n}}$, letting $n \to \infty$ 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

$$\sigma = \sqrt{\mathbb{E}|X_1|^2} = \left(\frac{(L+1)(2L+1)}{6}\right)^{1/2} \tag{18}$$

and let $G_1, G_2, \ldots$ 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 (17) 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

$$\mathbb{E}|a + bX_1|^p \leq \mathbb{E}|a + bG_1|^p. \tag{19}$$

This will follow from the following claim.

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

$$\mathbb{E}h(X_1^2) \leq \mathbb{E}h(G_1^2). \tag{20}$$

Indeed, (19) for $b = 0$ is clear. Assuming $b \neq 0$, by homogeneity, (19) 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

$$h_a(x) = |a + \sqrt{x}|^p + |a - \sqrt{x}|^p, \qquad x \geq 0 \tag{21}$$

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

Lemma 4.2.2. *Let $p \geq 3$, $a \in \mathbb{R}$. Then h_a defined in* (21) *is convex nondecreasing on* $[0, \infty)$.

Proof of 4.2.2. 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 $(0, \infty)$. 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 $(0, \infty)$. □

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

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

with σ (depending on L) defined by (18). Fix an integer $L \geq 2$ and set for nonnegative a,

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

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 $b \in \{0, 1, \ldots, L-1\}$ 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

$$\begin{aligned} f'(a) &= -\sqrt{\frac{2}{\pi\sigma^2}}\int_{\sqrt{a}}^{\infty}e^{-x^2/2\sigma^2}\mathrm{d}x - \frac{1}{L}\sum_{k=b+1}^{L}(-1) \\ &= -\sqrt{\frac{2}{\pi\sigma^2}}\int_{\sqrt{a}}^{\infty}e^{-x^2/2\sigma^2}\mathrm{d}x + \frac{L-b}{L}, \qquad a \in (b^2, (b+1)^2). \end{aligned} \tag{22}$$

Bounding $b < \sqrt{a}$ yields

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

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

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

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

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

where we define

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

Note: $g''(a) = \tilde{g}'(a) = \frac{1}{2\sqrt{a}}\left(\sqrt{\frac{2}{\pi}}\frac{1}{\sigma}e^{-\frac{a}{2\sigma}} - \frac{1}{L}\right)$ 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,

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

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

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

We have,

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

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 (22)) is strictly increasing on each interval $a \in (b^2, (b+1)^2)$, $b \in \{0, 1, \ldots, L-1\}$. Clearly $f'(0+) = 0$ and we check that $\theta_{L,b} = f'(b^2+) > 0$ for every $b \in \{1, \ldots, L-2\}$ and $3 \leq L \leq 6$, 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. □

Here 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

$$\sum_{i=1}^n a_i = \sum_{i=1}^n b_i \qquad \text{and} \qquad \sum_{i=1}^k a_i^* = \sum_{i=1}^k b_i^* \;\; \text{for all} \;\; k = 1, \ldots, n,$$

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, $(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}) \prec (a_1, a_2, \ldots, a_n) \prec (1, 0, \ldots, 0)$ for every nonnegative sequence (a_i) with $\sum_{i=1}^n a_i = 1$. A function $\Psi\colon [0,\infty)^n \to \mathbb{R}$ 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 $\psi\colon [0,+\infty) \to \mathbb{R}$ 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 4.2.3. *Let $\rho_0 \in [0, \frac{1}{2}]$. Let $X_1, X_2, \ldots$ 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 $a_1, \ldots, a_n, b_1, \ldots, b_n$ such that $(a_i^2)_{i=1}^n \prec (b_i^2)_{i=1}^n$, we have*

$$\mathbb{E}\left|\sum_{i=1}^n a_i X_i\right|^p \geq \mathbb{E}\left|\sum_{i=1}^n b_i X_i\right|^p. \tag{23}$$

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 4.2.4. *Let $q \geq 2$, $w \geq 0$ and $\phi_w(x) = \mathrm{sgn}(x+w)|x+w|^q + \mathrm{sgn}(x-w)|x-w|^q$, $x \in \mathbb{R}$. 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 of 4.2.4. 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

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

Note that taking one more derivative gives

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

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

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

Taking one more derivative yields

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

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. □

Lemma 4.2.5. *The best constant D such that the inequality*

$$D \cdot \left[\frac{\phi(a+b) - \phi(b-a)}{2a} - \frac{\phi(a+b) + \phi(b-a)}{2b} \right] \geq \left[\frac{\phi(b)}{b} - \frac{\phi(a)}{a} \right] \tag{24}$$

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 of 4.2.5. For $\phi(x) = xr(x)$, $r(x) = |x|$, by homogeneity, inequality (24) 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 (24) 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

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

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 4.2.3. Fix $p \geq 3$ and let $F(x) = |x|^p$. Then (23) is equivalent to saying that the function

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

is Schur concave. Since Φ is symmetric, by Ostrowski's criterion (see, e.g., Theorem II.3.14 in [4]), Φ is Schur concave if and only if

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

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

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

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

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

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

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

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

$$(\rho_0^{-1} - 1)\left[\frac{\phi(a+b) - \phi(b-a)}{2a} - \frac{\phi(a+b) + \phi(b-a)}{2b}\right] \geq \left[\frac{\phi(b)}{b} - \frac{\phi(a)}{a}\right]. \tag{25}$$

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 4.2.4 and 4.2.5, inequality (25) 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 $\rho_0 \to 0+$ directly in (23). $\square$

Corollary 4.2.6. *Under the assumptions of Theorem 4.2.3 for every $n \geq 1$ and reals $a_1, \dots, a_n$, we have*

$$\left(\mathbb{E}\left|\sum_{i=1}^n a_i X_i\right|^p\right)^{1/p} \leq C_p \left(\mathbb{E}\left|\sum_{i=1}^n a_i X_i\right|^2\right)^{1/2} \tag{26}$$

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 ameniable to bounds for $p < 2$. Here we consdier the case $p = 1$.

Theorem 4.3.1. *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*

$$\mathbb{E}\left|\sum_{i=1}^n a_i X_i\right| \geq c_1 \left(\mathbb{E}\left|\sum_{i=1}^n a_i X_i\right|^2\right)^{1/2} \tag{27}$$

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.

Proof of 4.3.1. Note that for $a_1 = 1$, $a_2 = \cdots = a_n = 0$, we have equality in (27), which explains why the value of the constant c_1 is sharp.

We shall closely follow Haagerup's approach from [10]. 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

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

By symmetry, without loss of generality we can assume that $a_1, \ldots, a_n$ are positive with $\sum a_j^2 = 1$. By Lemma 1.2 from [10] 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 [10], by the AM-GM inequality,

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

thus

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

If we show that

$$F(s) \ge F(1), \qquad s \ge 1, \tag{28}$$

then

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

Since ϕ_X is nonnegative, using again Lemma 1.2 from [10], we have

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

so the proof of (27) is finished.

It remains to show (28). For a fixed $s \ge 1$, 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

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

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

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

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

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

where $F_{\mathrm{Haa}}(s) = \frac{2}{\pi}\int_0^\infty \left[1 - \left|\cos\left(\frac{t}{\sqrt{s}}\right)\right|^s\right] \frac{dt}{t^2}$ is Haagerup's function (see Lemma 1.3 and 1.4 in [10]). He showed therein that it is increasing, so for $s \ge 1$, we get $F_{\mathrm{Haa}}(2s) \ge F_{\mathrm{Haa}}(2) = \frac{1}{\sqrt{2}}$ and this finishes the proof. $\square$

5 Appendix

Here we present proofs for completeness of standard arguments used but not thoroughly argued above.

5.1 Proof of Convergence of Moments

Lemma 5.1.1. *Suppose $X_n \to X$ in distribution. If $\{X_n\}$ is uniformly integrable then $\mathbb{E}|X| < \infty$ and $\mathbb{E}(X_n) \to \mathbb{E}(X)$ and $\mathbb{E}|X_n| \to \mathbb{E}|X|$. Recall a sequence $\{X_n\}$ uniformly integrable if $\sup_n |X_n| < \infty$ and for all $\epsilon > 0$ we have $\delta > 0$ such that when for some event A, $\mathbb{P}(A) < \delta$, then $\mathbb{P}(|X_n| \in A) < \epsilon$.*

We follow the presentation of Billingsey in [5]. The proof relies on a standard fact from measure theory.

Proposition 5.1.2. *Let Ω be a set with finite measure μ. Let $f_n \to f$ almost everywhere. If the f_n are uniformly integrable then f is integrable and*

$$\int f_n d\mu \to \int f d\mu.$$

Proof of 5.1.2. Via Fatou's Lemma we know $\int |f| d\mu < \infty$. We define

$$f_n^{(\alpha)} = \begin{cases} f_n & |f_n| < \alpha \\ 0 & |f_n| \geq \alpha \end{cases} \qquad f^{(\alpha)} = \begin{cases} f & |f| < \alpha \\ 0 & |f| \geq \alpha \end{cases}$$

as cutoff functions controlling the size of f. Then here we may apply the Dominated Convergence Theorem to see

$$\int f_n^{(\alpha)} d\mu \to \int f^{(\alpha)} d\mu$$

since clearly $f_n^{(\alpha)} \to f^{(\alpha)}$ pointwise and we have the bound of $\left(f_n^{(\alpha)}\right) \leq \alpha$.

Then noting

$$\int f_n d\mu - \int f_n^{(\alpha)} d\mu = \int_{|f_n| \geq \alpha} f_n d\mu$$
$$\int f d\mu - \int f^{(\alpha)} d\mu = \int_{|f| \geq \alpha} f d\mu$$

we have

$$\limsup_{n\to\infty} | \int f_n d\mu - \int f d\mu| \leq \sup_n \int_{|f_n| \geq \alpha} |f_n| d\mu + \int_{|f| \geq \alpha} |f| d\mu$$

But using uniform integrabilty we can send the first term to 0. The second term goes to 0 since f integrable. □

We also need Skorohod's Theorem allowing us to pick pointwise converging random variables with specified distributions.

Theorem 5.1.3 (Skorohod's Theorem)**.** *Suppose μ_n and μ probability measures on $(R^1, \mathcal{R}^1)$ with $\mu_n \to \mu$. Then we can find random variables Y_n and Y on probability space $(\Omega, \mathcal{F}, P)$ such that Y_n has distribution μ_n and Y has distribution μ with $Y_n \to Y$ almost surely.*

Now we are able to show the convergence of moments given convergence in distribution and uniform integrability.

Proof of 5.1.1. It suffices to pick $Y_n \to Y$ pointwise with same distributions as $X_n \to X$ in distribution. This can be done with Skorohod's Lemma. But since we still have uniform integrability with pointwise convergence we can apply proposition 5.1.2 and to get the convergence of moments. □

5.2 Proof of Distribution Lemma

We present a proof of the distribution function lemma used in Nazarov and Podkorytovs' proof of optimal Khintchine inequalities for random signs.

Lemma 5.2.1 (Nazarov and Podkorytov, [22])**.** *Let $Y > 0$, $f, g : \mathcal{M} \to [0, Y]$ be any two nonnegative measurable functions on $(\mathcal{M}, \mu)$. Let F and G be their distribution functions. Assume both $F(y)$ and $G(y)$ are finite for every $y > 0$. Assume also there exists unique y_0 such that $F_* - G_* = 0$. 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.

Proof of 5.2.1. Note since we have finiteness for all y then

$$\int_{\mathcal{M}} f - g d\mu = \int_0^\infty F(y) - G(y) dy$$

since letting $h(x) = min(f(x), g(x))$ and with $H(y)$ as the corresponding distribution function we know

$$0 \leq \int_{\mathcal{M}} f - h d\mu = \int_0^\infty F(y) - H(y) dy < \infty$$
$$0 \leq \int_{\mathcal{M}} g - h d\mu = \int_0^\infty G(y) - H(y) dy < \infty$$

via Fubini applied to the Characteristic $\{(x, y) \in \mathcal{M} \times (0, \infty) : h(x) \leq y < f(x)\}$. Then we subtract the two for the desired claim.

Then via properites of distribution functions we know

$$\int_{\mathcal{M}} f^s - g^s d\mu = \int_0^\infty F(y^{1/s}) - G(y^{1/s}) dy = s \int_0^\infty y^{s-1}(F(y) - G(y)) dy$$

so for $s > s_0$ we can write

$$\phi(s) - \phi(s_0) = \frac{1}{y_0} \int_0^\infty ((\frac{y}{y_0})^{s-1} - (\frac{y}{y_0})^{s_0-1})(F(y) - G(y)) dy \geq 0$$

where the get the inequality since both factors in integrand change signs at y_0. □

References

[1] Baernstein, A., II, Culverhouse, Robert C., Majorization of sequences, sharp vector Khinchin inequalities, and bisubharmonic functions. Studia Math. 152 (2002), no. 3, 231–248.

[2] Barthe, F., Guédon, O., Mendelson, S., Naor, A., A probabilistic approach to the geometry of the ℓ_p^n-ball. *Ann. Probab.* 33 (2005), no. 2, 480–513.

[3] Barthe, F., Naor, A., Hyperplane projections of the unit ball of ℓ_p^n. *Discrete Comput. Geom.* 27 (2002), no. 2, 215–226.

[4] Bhatia, R., Matrix analysis. Graduate Texts in Mathematics, 169. *Springer-Verlag, New York*, 1997.

[5] Billingsley, B. *Probability and Measure.*

[6] Czerwinski, W. Khinchine inequality, University of Warsaw, Master Thesis, 2008.

[7] Eaton, M. L., A note on symmetric Bernoulli random variables. *Ann. Math. Statist.* 41 (1970), 1223—1226.

[8] Eskenazis, A., Nayar, P., Tkocz, T., Gaussian mixtures: entropy and geometric inequalities, *Ann. of Prob.* 46(5) 2018, 2908–2945.

[9] Eskenazis, A., Nayar, P., Tkocz, T., Sharp comparison of moments and the log-concave moment problem, *Adv. Math.* 334 (2018) 389–416.

[10] Haagerup, U., The best constants in the Khintchine inequality. *Studia Math.* 70 (1981), no. 3, 231–283.

[11] Havrilla, A., Tkocz, T., Sharp Khinchin-type inequalities for symmetric discrete uniform random variables. Preprint (2019), arXiv:1912.13345, to appear in Israel J. Math.

[12] Havrilla .A, Nayar. P, Tkocz. T, Khinchin-type inequalities via Hadamard's factorisation.

[13] J.-P. Kahane, Sur les sommes vectorielles, C. R. Acad Sci. Paris 259(1964), 2577-2580

[14] Khinchin, A. Über dyadische Brüche, Math. Z. 18 (1923), 109-116

[15] Komorowski, R., On the best possible constants in the Khintchine inequality for $p \geq 3$. *Bull. London Math. Soc.* 20 (1988), no. 1, 73-–75.

[16] Kwapień, S., Latała, R., Oleszkiewicz, K., Comparison of moments of sums of independent random variables and differential inequalities. *J. Funct. Anal.* 136 (1996), no. 1, 258–268.

[17] Latała, R., Oleszkiewicz, K., A note on sums of independent uniformly distributed random variables. *Colloq. Math.* 68 (1995), no. 2, 197–206.

[18] Ledoux, M., Talagrand, M., Probability in Banach spaces. Isoperimetry and processes. *Springer-Verlag*, Berlin, 1991.

[19] Marshall. A, Olkin. I, Arnold. B, *Inequalities: Theory of Majorization and its Applications.* Springer.

[20] Mordhorst, O., The optimal constants in Khintchine's inequality for the case $2 < p < 3$. *Colloq. Math.* 147 (2017), no. 2, 203-–216.

[21] Nayar, P., Oleszkiewicz, K., Khinchine type inequalities with optimal constants via ultra log-concavity. *Positivity* 16 (2012), no. 2, 359–371.

[22] Nazarov, F. L., Podkorytov, A. N., Ball, Haagerup, and distribution functions. *Complex analysis, operators, and related topics*, 247–267, *Oper. Theory Adv. Appl.*, 113, Birkhäuser, Basel, 2000.

[23] Nayar, P., Zwara, S., Sharp variance-entropy comparison for nonnegative gaussian quadratic forms. Preprint (2020), arXiv:2005.11705.

[24] Newman, C. M., An extension of Khintchine's inequality. Bull. Amer. Math. Soc. 81 (1975), no. 5, 913–915.

[25] Newman, C, M., Moment inequalities for ferromagnetic Gibbs distributions. J. Math. Phys. 16 (1975), no. 9, 1956–1959.

[26] O'Rourke, S. A note on the Marchenko-Pastur law for a class of random matrices with dependent entries. Electronic Communications in Probability, Vol. 17, No. 28, pp. 1-13 (2012)

[27] Pass, B., Spektor, S., On Khintchine type inequalities for k-wise independent Rademacher random variables. *Statist. Probab. Lett.* 132 (2018), 35–39.

[28] Peskir, G. Best constants in the Kahane-Khintchine inequalities for complex Steinhaus functions. 1995.

[29] Pólya, G., Szegö, 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.

[30] Pólya, G., Szegö, G., Problems and theorems in analysis. II. Theory of functions, zeros, polynomials, determinants, number theory, geometry. Translated from the German by C. E. Billigheimer. Reprint of the 1976 English translation. Classics in Mathematics. Springer-Verlag, Berlin, 1998.

[31] S. T. Rachev and L. Ruschendorf. *Approximate independence of distributions on spheres and their stability properites.*

[32] G. Schechtman and J. Zinn. On the volume of the intersection of two L_p^n balls. Banach Archive 11/9/89.

[33] Spektor, S., Restricted Khinchine inequality. *Canad. Math. Bull.* 59 (2016), no. 1, 204–210.

[34] S. B. Stechkin, "On best lacunary systems of functions", Izv. Akad. Nauk SSSR Ser. Mat., 25:3 (1961), 357–366

[35] Stein, E. Shakarchi, R. *Princeton Lectures in Analysis Vol II: Complex Analysis.*

[36] Szarek, S., On the best constant in the Khintchine inequality. *Stud. Math.* 58, 197–208 (1976)

[37] Walkup, D.W. Polya sequences, binomial convolution and the union of random sets. J. Appl. Probab. 13, 76-85 (1976)

[38] Whittle, P., Bounds for the moments of linear and quadratic forms in independent random variables. *Theory Probab. Appl. 5, 302-305* (1960)

[39] Young, R. M. G., On the best possible constants in the Khintchine inequality. *J. London Math. Soc.* (2) 14 (1976), no. 3, 496–504.