Stochastic Dominance and Weak Concentration for Sums of Independent Symmetric Random Vectors

Witold Bednorz[1] and Tomasz Tkocz [iD],[2],[*]

[1]University of Warsaw and [2]Carnegie Mellon University

[*]*Correspondence to be sent to: e-mail: ttkocz@math.cmu.edu*

Kwapień and Woyczyński asked in their monograph (1992) whether their notion of superstrong domination is inherited when taking sums of independent symmetric random vectors (one vector dominates another if, essentially, tail probabilities of any norm of the two vectors compare up to some scaling constants). We answer this question positively. As a by-product of our methods, we establish that a certain notion of weak concentration is also preserved by taking sums of independent symmetric random vectors.

1 Introduction

Stochastic orderings quantitatively capture the notion of one random variable being *greater* than another one. Common examples include $\mathcal{U}$-stochastic orderings. If $\mathcal{U}$ is a family of real valued functions defined on, say a real separable Banach space E, we say for E-valued random vectors X and Y that X is $\mathcal{U}$-dominated by Y, written $X \prec_{\mathcal{U}} Y$, if $\mathbb{E}f(X) \leq \mathbb{E}f(Y)$, for all functions f in $\mathcal{U}$. For instance, taking $\mathcal{U}$ to be the family of nonnegative convex functions on E results in the usual convex stochastic ordering, or considering $\mathcal{U}$ to be the family of the exponents of bounded linear functionals, $\exp\{x^*(\cdot)\}$, $x^* \in E^*$, can be used to define sub-Gaussian random vectors, just to name several important examples. An inductive argument shows that these two orderings are

inherited for sums: if $X_i \prec_{\mathcal{U}} Y_i$, for $i = 1, \ldots, n$ and the X_i and Y_i are independent, then $\sum_{i=1}^n X_i \prec_{\mathcal{U}} \sum_{i=1}^n Y_i$. This is a significant property of a stochastic ordering as it allows to compare sums in presence of the comparison for independent summands. The main goal of this article is to establish such a *tensorisation* property for symmetric random vectors of a stochastic ordering called *superstrong domination*, which we shall now define.

Let X and Y be symmetric random vectors with values in a real separable Banach space E. We say in this paper that Y dominates X with constants $\kappa, \lambda \geq 1$ ((κ, λ)-dominates, in short) if for every closed convex and symmetric set K in E we have

$$\mathbb{P}\left(X \notin K\right) \leq \kappa \mathbb{P}\left(\lambda Y \notin K\right).$$

We sometimes write $X \prec_{(\kappa, \lambda)} Y$. Equivalently, X is (κ, λ)-dominated by Y if for every continuous norm $\| \cdot \|$ on E, we have

$$\mathbb{P}\left(\|X\| > 1\right) \leq \kappa \mathbb{P}\left(\lambda \|Y\| > 1\right).$$

(The inequality for convex sets clearly implies the inequality for norms. Conversely, given a closed convex and symmetric set K in E, take $K_{\delta, R}$ to be the δ-enlargement of K, $\{x \in E, \text{dist}(x, K) \leq \delta\}$ intersected with the closed ball of radius R in E. The Minkowski functional of $K_{\delta, R}$ defines a continuous norm on E. Letting δ go to 0 and R to ∞ finishes the argument.)

This notion appears as *superstrong domination* in the monograph by Kwapień and Woyczyński (see [10], Chapters 3.2 and 3.6). It can be viewed as a less restrictive version of the $\mathcal{U}$-stochastic ordering for the family $\mathcal{U}$ comprising the indicators of complements of convex symmetric sets. The special case, $\kappa = \lambda = 1$ and $E = \mathbb{R}^d$ is known as *peakedness* and was first introduced by Birnbaum (univariate case, $d = 1$, see [6]), by Sherman (multivariate case, $d \geq 1$, see [14]), and by Kanter (see [9]), which has found numerous applications in convex geometry (see, for instance, [2], [13], [15]). In this case, by considering symmetric strips, if $X \prec_{(1,1)} Y$, then necessarily $\mathbb{E}|\langle t, X \rangle|^2 \leq \mathbb{E}|\langle t, Y \rangle|^2$, for all vectors t in $\mathbb{R}^d$. For symmetric Gaussian random vectors X and Y, this simple necessary condition is also sufficient! (Since the matrix $[\mathbb{E}(Y_i Y_j - X_i X_j)]_{i,j}$ is positive semi-definite, there is an independent symmetric Gaussian random vector Z such that $Y = X + Z$ and Anderson's inequality, see [1], finishes the argument.)

Kwapień and Woyczyński posed a question whether superstrong domination is preserved by taking sums of independent symmetric random vectors. They remarked

that the answer is positive for vectors taking values in one-dimensional subspaces, crediting this result to Jain and Marcus (see [7] and Theorem 3.2.1 in [10]). Kanter's result says that the peakedness of log-concave measures in $\mathbb{R}^d$ tensorises (see Corollary 3.2 in [9]), supporting the "yes" answer in this case. Our 1st main result provides the positive answer in full generality.

Theorem 1.   Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be independent symmetric random vectors with values in separable Banach space. Suppose that X_i is (κ, λ)-dominated by Y_i for each $i = 1, \ldots, n$. Then the sum $X_1 + \ldots + X_n$ is $(16\alpha^{-1}\lceil \kappa \rceil, (1+\alpha)\lceil \kappa \rceil \lambda)$-dominated by the sum $Y_1 + \ldots + Y_n$ for any $0 < \alpha \leq 1$.

The crux of our argument is to devise a *proxy*, a quantity that mimics tail probabilities $\mathbb{P}(\|X\| > 1)$, but, as opposed to them, gives rise to inequalities that are easy to tensorise. With the aid of the proxy as well as several tools for random signs, we first establish the tensorisation of $(1, 1)$-domination. Then we show how to deduce the theorem for arbitrary κ and λ from the case $\kappa = 1 = \lambda$.

Using similar tools and techniques, we derive a tensorisation property for a certain notion of weak concentration, which can be of independent interest. We say that a symmetric random vector X with values in a separable Banach space E satisfies the *weak Borell* inequality (or the weak concentration) with constants $C \geq 1$, $\delta > 0$, and $0 < \theta < 1$ (WB(C, δ, θ) for short), if for every continuous norm $\|\cdot\|$ on E such that $\mathbb{P}(\|X\| > 1) < \theta$, we have

$$\mathbb{P}(\|X\| > \lambda) \leq C\lambda^{-\delta}\mathbb{P}(\|X\| > 1), \qquad \lambda \geq 1. \tag{WB}$$

Again, this is the same as saying that for $\lambda \geq 1$ and every closed symmetric convex subset K of E such that $\mathbb{P}(X \notin K) < \theta$, we have $\mathbb{P}(X \notin \lambda K) \leq C\lambda^{-\delta}\mathbb{P}(X \notin K)$.

For instance, if X is a κ-concave random vector in $\mathbb{R}^d$, $\kappa < 0$, and $\theta < 1/2$, then X satisfies WB$(C, -1/\kappa, \theta)$ with C dependent only on κ and θ (see [4] and [5]). Even though in general κ-concavity with negative κ is not preserved by taking sums of independent vectors, the WB inequality is (modulo a slight change of constants), which is our 2nd main result.

Theorem 2.   Suppose that symmetric random vectors $X_1, \ldots, X_n$ in separable Banach space are independent and each satisfies WB(C, δ, θ). Then $X_1 + \ldots + X_n$ satisfies WB(C', δ, θ'), where $C' = 12 \cdot 9^\delta C$ and $\theta' = \min\{\theta/2, (96C \cdot 9^\delta)^{-1}\}$.

As an application of this result, we establish superstrong domination for weighted sums of i.i.d. symmetric random vectors satisfying the weak Borell inequality when the sequences of weights are comparable in terms of majorisation. Recall that one sequence of real numbers $a = (a_1, \ldots, a_n)$ is majorised by another one $b = (b_1, \ldots, b_n)$, usually denoted $a \prec b$, if the nonincreasing rearrangements $a_1^* \geq \ldots \geq a_n^*$ and $b_1^* \geq \ldots \geq b_n^*$ of a and b satisfy the inequalities

$$\sum_{j=1}^{k} a_j^* \leq \sum_{j=1}^{k} b_j^* \text{ for each } 1 \leq k \leq n-1 \text{ and } \sum_{j=1}^{n} a_j = \sum_{j=1}^{n} b_j.$$

Equivalently, a is a convex combination of the permutations $(b_1', \ldots, b_n')$ of b (see for example Theorem II.1.10 in [3]).

Theorem 3. Let $X_1, X_2, \ldots$ be i.i.d. symmetric random vectors in separable Banach space. Assume that for some $C > 0$, $0 < \theta < 1$, and $\delta > 1$, each X_i satisfies $\mathrm{WB}(C, \delta, \theta)$. Let $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$ be sequences of real numbers such that a is majorised by b. Then

$$\sum_{i=1}^{n} a_i X_i \prec_{(\kappa, \lambda)} \sum_{i=1}^{n} b_i X_i$$

with $\kappa = \max\{2\theta^{-1}, 96C \cdot 9^\delta, 12C \cdot 9^\delta (\delta - 1)^{-1}\}$ and $\lambda = 2$.

This theorem does not hold under the weaker assumption that the X_i satisfy the weak concentration with $\delta < 1$. To see that, fix $\delta \in (0,1)$ and take X_i to be independent real valued symmetric δ-stable random variables. Then $\mathbb{P}(|X_1| > t) \sim t^{-\delta}$, for large t, hence, the X_i satisfy $\mathrm{WB}(C, \delta', \theta)$ if and only if $\delta' \leq \delta$. Consider the sequences $a = (1/n, \ldots, 1/n)$ and $b = (1, 0, \ldots, 0)$. Then $a \prec b$ and $\sum a_i X_i$ has the same distribution as $n^{1/\delta - 1} X_1$, so $\sum_{i=1}^{n} a_i X_i \prec_{(\kappa, \lambda)} \sum_{i=1}^{n} b_i X_i$ would particularly imply that $\mathbb{P}(|X_1| > 1) \leq \kappa \mathbb{P}(\lambda |X_1| > n^{1/\delta - 1})$, which is not true for large n. We suspect that our assumption of $\delta > 1$ can be weakened to $\delta \geq 1$.

2   Auxiliary Results

In this section, we collect several well-known inequalities that will be needed in our proofs. We begin with three results for random signs. Here and throughout, $\varepsilon_1, \varepsilon_2, \ldots$ are independent random signs each taking the value ± 1 with probability $1/2$. Let $v_1, \ldots, v_n$ be vectors in a separable Banach space $(E, \|\cdot\|)$. Kahane's inequality (see [8] or

Proposition 1.4.1. in [10]) says that for $s, t > 0$, we have

$$\mathbb{P}\left(\left\|\sum \varepsilon_i v_i\right\| > s + t\right) \leq 4\mathbb{P}\left(\left\|\sum \varepsilon_i v_i\right\| > s\right)\mathbb{P}\left(\left\|\sum \varepsilon_i v_i\right\| > t\right). \tag{1}$$

We also recall the optimal $L_1 - L_2$ moment comparison due to Latała and Oleszkiewicz (see [11]), that is

$$\mathbb{E}\left\|\sum \varepsilon_i v_i\right\|^2 \leq 2\left(\mathbb{E}\left\|\sum \varepsilon_i v_i\right\|\right)^2. \tag{2}$$

This, combined with the Paley–Zygmund inequality yields that for any $\theta \in (0, 1)$, we have

$$\mathbb{P}\left(\left\|\sum \varepsilon_i v_i\right\| > \theta \mathbb{E}\left\|\sum \varepsilon_i v_i\right\|\right) \geq \frac{1}{2}(1 - \theta)^2. \tag{3}$$

The contraction principle (see for instance Theorem 4.4 in [12]) in particular asserts that for two sequences of real numbers $(a_i)_{i=1}^n$ and $(b_i)_{i=1}^n$ such that $|a_i| \leq |b_i|$ for each $i \leq n$, we have

$$\mathbb{E}\left\|\sum \varepsilon_i a_i v_i\right\| \leq \mathbb{E}\left\|\sum \varepsilon_i b_i v_i\right\|. \tag{4}$$

Let us recall several classical inequalities for sums of symmetric independent random vectors $X_1, \ldots, X_n$ with values in the separable Banach space $(E, \|\cdot\|)$. Denote as usual $S_j = X_1 + \ldots + X_j$, $j \leq n$, $X_n^* = \max_{j \leq n} \|X_j\|$, and $S_n^* = \max_{j \leq n} \|S_j\|$. The Lévy inequality says that

$$\mathbb{P}\left(S_n^* > t\right) \leq 2\mathbb{P}\left(\|S_n\| > t\right), \qquad t \geq 0. \tag{5}$$

Moreover, we have

$$\mathbb{P}\left(X_n^* > t\right) \leq 2\mathbb{P}\left(\|S_n\| > t\right), \qquad t \geq 0. \tag{6}$$

The Hoffmann–Jørgensen inequality asserts that

$$\mathbb{P}\left(S_n^* > s + t + u\right) \leq \mathbb{P}\left(X_n^* > s\right) + 2\mathbb{P}\left(S_n^* > t\right)\mathbb{P}\left(\|S_n\| > u\right), \qquad s, t, u \geq 0. \tag{7}$$

Lastly, even without the symmetry of the X_i, we have

$$\sum_{j=1}^n \mathbb{P}\left(\|X_j\| > t\right) \leq \frac{\mathbb{P}\left(X_n^* > t\right)}{1 - \mathbb{P}\left(X_n^* > t\right)}, \qquad t \geq 0. \tag{8}$$

(All of these inequalities can be found for instance in Chapter 1 of [10]).

3   Proof of Theorem 1.1

The goal of the 1st three sections is to show Theorem 1 when $\kappa = 1 = \lambda$. In the last section, we show how to deduce the general case.

3.1   Conditional convexity and a proxy

We start with a simple lemma that lies at the heart of our tensorisation argument.

Lemma 4.   Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be independent symmetric random vectors with values in a separable Banach space E such that X_i is $(1,1)$-dominated by Y_i for each $i = 1, \ldots, n$. Let $\varphi \colon E^n \to [0, \infty)$ be a continuous function, convex with respect to each coordinate. Then for $t \geq 0$, we have

$$\mathbb{P}\left(\mathbb{E}_\varepsilon \varphi(\varepsilon_1 X_1, \ldots, \varepsilon_n X_n) > t\right) \leq \mathbb{P}\left(\mathbb{E}_\varepsilon \varphi(\varepsilon_1 Y_1, \ldots, \varepsilon_n Y_n) > t\right).$$

Proof.   We condition on $X_2, \ldots, X_n$ and define the set

$$K = \left\{x \in E,\ \mathbb{E}_\varepsilon \varphi(\varepsilon_1 x, \varepsilon_2 X_2, \ldots, \varepsilon_n X_n) \leq t\right\},$$

which is closed, convex, and symmetric. Using $X_1 \prec_{(1,1)} Y_1$, we get $\mathbb{P}\left(X_1 \notin K\right) \leq \mathbb{P}\left(Y_1 \notin K\right)$, which means that

$$\mathbb{P}_{X_1}\left(\mathbb{E}_\varepsilon \varphi(\varepsilon_1 X_1, \varepsilon_2 X_2, \ldots, \varepsilon_n X_n) > t\right) \leq \mathbb{P}_{Y_1}\left(\mathbb{E}_\varepsilon \varphi(\varepsilon_1 Y_1, \varepsilon_2 X_2, \ldots, \varepsilon_n X_n) > t\right),$$

so taking the expectation of both sides against $X_2, \ldots, X_n$ gives

$$\mathbb{P}\left(\mathbb{E}_\varepsilon \varphi(\varepsilon_1 X_1, \varepsilon_2 X_2, \ldots, \varepsilon_n X_n) > t\right) \leq \mathbb{P}\left(\mathbb{E}_\varepsilon \varphi(\varepsilon_1 Y_1, \varepsilon_2 X_2, \ldots, \varepsilon_n X_n) > t\right).$$

Similarly, we condition on $Y_1, X_3, \ldots, X_n$ to swap X_2 for Y_2, etc. and finally arrive at the desired inequality.   ∎

Note that the function $u \mapsto (u-1)_+ = \max\{u-1, 0\}$ is convex and nondecreasing. Therefore, for a normed space $(E, \|\cdot\|)$ the function $\varphi \colon E^n \to [0, \infty)$ defined by $\varphi(x_1, \ldots, x_n) = (\|\sum_{i=1}^n x_i\| - 1)_+$ is convex (and continuous). From Lemma 4 we thus get the following corollary.

Corollary 5.   Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be independent symmetric random vectors with values in a separable Banach space E such that X_i is $(1, 1)$-dominated by Y_i for each $i = 1, \ldots, n$. Let $\| \cdot \|$ be a continuous norm on E. Then for $t \geq 0$, we have

$$\mathbb{P}\left(\mathbb{E}_{\varepsilon}\left(\left\|\sum \varepsilon_i X_i\right\| - 1\right)_+ > t\right) \leq \mathbb{P}\left(\mathbb{E}_{\varepsilon}\left(\left\|\sum \varepsilon_i Y_i\right\| - 1\right)_+ > t\right).$$

In particular,

$$\int_0^1 \mathbb{P}\left(\mathbb{E}_{\varepsilon}\left(\left\|\sum \varepsilon_i X_i\right\| - 1\right)_+ > t\right) dt \leq \int_0^1 \mathbb{P}\left(\mathbb{E}_{\varepsilon}\left(\left\|\sum \varepsilon_i Y_i\right\| - 1\right)_+ > t\right) dt. \tag{9}$$

For a nonnegative random variable Y, we plainly have

$$\mathbb{E}\min\{Y, 1\} = \int_0^{\infty} \mathbb{P}\left(\min\{Y, 1\} > t\right) dt = \int_0^{\infty} \mathbb{P}\left(Y > t, 1 > t\right) dt = \int_0^1 \mathbb{P}\left(Y > t\right) dt.$$

Therefore, in view of this corollary, the following quantity

$$\int_0^1 \mathbb{P}\left(\mathbb{E}_{\varepsilon}\left(\left\|\sum \varepsilon_i X_i\right\| - 1\right)_+ > t\right) dt = \mathbb{E}\min\left\{\mathbb{E}_{\varepsilon}\left(\left\|\sum \varepsilon_i X_i\right\| - 1\right)_+, 1\right\}$$

tensorises as well. This is our proxy and we will show that it is comparable to $\mathbb{P}\left(\|\sum X_i\| > 1\right)$. This is where the assumption of symmetry and the aforementioned tools for random signs come into play.

3.2   Upper and lower bounds for the proxy

Lemma 6.   Suppose that $X_1, \ldots, X_n$ are independent symmetric random vectors in a normed space $(E, \| \cdot \|)$. Then for $0 < \alpha \leq 1$, we have

$$\mathbb{E}\min\left\{\mathbb{E}_{\varepsilon}\left(\left\|\sum \varepsilon_i X_i\right\| - 1\right)_+, 1\right\} \geq \alpha \mathbb{P}\left(\left\|\sum X_i\right\| > 1 + \alpha\right).$$

Proof.   Denote $U = \mathbb{E}_{\varepsilon}(\|\sum \varepsilon_i X_i\| - 1)_+$. Notice that for a positive parameter α, we have

$$U \geq \mathbb{E}_{\varepsilon}\left(\left\|\sum \varepsilon_i X_i\right\| - 1\right)_+ \mathbf{1}_{\{\|\sum \varepsilon_i X_i\| > 1 + \alpha\}} \geq \alpha \mathbb{P}_{\varepsilon}\left(\left\|\sum \varepsilon_i X_i\right\| > 1 + \alpha\right).$$

Thus,

$$\mathbb{E}\min\{U, 1\} \geq \mathbb{E}\min\{\alpha \mathbb{P}_{\varepsilon}\left(\left\|\sum \varepsilon_i X_i\right\| > 1 + \alpha\right), 1\}.$$

When $\alpha \leq 1$, the last expression becomes $\alpha \mathbb{P}\left(\|\sum X_i\| > 1 + \alpha\right)$.   ∎

Lemma 7.   Let $X_1, \ldots, X_n$ be independent symmetric random vectors in a normed space $(E, \|\cdot\|)$. Then we have

$$\mathbb{E} \min \left\{ \mathbb{E}_\varepsilon \left(\left\| \sum \varepsilon_i X_i \right\| - 1 \right)_+, 1 \right\} \le 16 \mathbb{P} \left(\left\| \sum X_i \right\| > 1 \right).$$

Proof.   For $p \in (0, 1)$, define the event

$$A_p = \left\{ \mathbb{P}_\varepsilon \left(\left\| \sum \varepsilon_i X_i \right\| > 1 \right) > p \right\}.$$

Clearly, we have

$$\mathbb{E} \min \left\{ \mathbb{E}_\varepsilon \left(\left\| \sum \varepsilon_i X_i \right\| - 1 \right)_+, 1 \right\} \le \mathbb{E} \mathbf{1}_{A_p} + \mathbb{E} \mathbb{E}_\varepsilon \left(\left\| \sum \varepsilon_i X_i \right\| - 1 \right)_+ \mathbf{1}_{A_p^c}.$$

We handle the 1st term directly by Markov's inequality,

$$\mathbb{P}(A_p) \le \frac{1}{p} \mathbb{E} \mathbb{P}_\varepsilon \left(\left\| \sum \varepsilon_i X_i \right\| > 1 \right) = \frac{1}{p} \mathbb{P} \left(\left\| \sum X_i \right\| > 1 \right).$$

To deal with the 2nd term, first notice that by Kahane's inequality (1), we have

$$\mathbb{E}_\varepsilon \left(\left\| \sum \varepsilon_i X_i \right\| - 1 \right)_+ = \int_0^\infty \mathbb{P}_\varepsilon \left(\left\| \sum \varepsilon_i X_i \right\| > 1 + t \right) \mathrm{d}t \le 4 \mathbb{P}_\varepsilon \left(\left\| \sum \varepsilon_i X_i \right\| > 1 \right) \mathbb{E}_\varepsilon \left\| \sum \varepsilon_i X_i \right\|.$$

Second, notice that the quantity $\mathbb{E}_\varepsilon \| \sum \varepsilon_i X_i \|$ is bounded on the event A_p^c. Indeed, suppose that $\mathbb{E}_\varepsilon \| \sum \varepsilon_i X_i \| > 1$ and set $\theta = (\mathbb{E}_\varepsilon \| \sum \varepsilon_i X_i \|)^{-1}$. Then on A_p^c, by (3),

$$p \ge \mathbb{P}_\varepsilon \left(\left\| \sum \varepsilon_i X_i \right\| > 1 \right) = \mathbb{P}_\varepsilon \left(\left\| \sum \varepsilon_i X_i \right\| > \theta \mathbb{E}_\varepsilon \left\| \sum \varepsilon_i X_i \right\| \right) \ge \frac{1}{2} (1 - \theta)^2,$$

so $\theta \ge 1 - \sqrt{2p}$ and provided that $p < 1/2$, we get

$$\mathbb{E}_\varepsilon \left\| \sum \varepsilon_i X_i \right\| \le \frac{1}{1 - \sqrt{2p}}.$$

Putting these together yields

$$\mathbb{E} \mathbb{E}_\varepsilon \left(\left\| \sum \varepsilon_i X_i \right\| - 1 \right)_+ \mathbf{1}_{A_p^c} \le \frac{4}{1 - \sqrt{2p}} \mathbb{E} \mathbb{P}_\varepsilon \left(\left\| \sum \varepsilon_i X_i \right\| > 1 \right) = \frac{4}{1 - \sqrt{2p}} \mathbb{P} \left(\left\| \sum X_i \right\| > 1 \right).$$

Altogether,

$$\mathbb{E}\min\left\{\mathbb{E}_\varepsilon\left(\left\|\sum\varepsilon_i X_i\right\|-1\right)_+,1\right\} \le \left(\frac{1}{p}+\frac{4}{1-\sqrt{2p}}\right)\mathbb{P}\left(\left\|\sum X_i\right\|>1\right).$$

Choosing $p = 1/8$ finishes the proof (the optimal choice $p \approx 0.16$ gives the constant ≈ 15.45). ∎

3.3 Proof in the case $\kappa = 1 = \lambda$

Suppose that $X_1,\ldots,X_n$ and $Y_1,\ldots,Y_n$ are independent symmetric random vectors with values in a separable Banach space E. Let X_i be $(1,1)$-dominated by Y_i for each $i \le n$. Fix a continuous norm $\|\cdot\|$ on E. We would like to show that $\mathbb{P}\left(\|\sum X_i\| > 1\right) \le \kappa\mathbb{P}\left(\lambda\|\sum Y_i\| > 1\right)$ for some universal constants κ and λ. Fix $0 < \alpha \le 1$. Applying consecutively Lemma 6, Corollary 5, and Lemma 7 yield

$$\alpha\mathbb{P}\left(\left\|\sum X_i\right\| > 1+\alpha\right) \le 16\mathbb{P}\left(\left\|\sum Y_i\right\| > 1\right).$$

Rescaling the norm gives the desired bound with $\kappa = 16\alpha^{-1}$ and $\lambda = 1 + \alpha$.

3.4 Reduction to the case $\kappa = 1 = \lambda$

We describe two arguments leading to the conclusion that it suffices to prove Theorem 1 when $\kappa = 1 = \lambda$, thus, finishing the whole proof.

The 1st argument is based on the following lemma whose proof is essentially given in the 2nd step of the proof of Theorem 3.2.1 in [10]. We sketch it for completeness.

Lemma 8.   Suppose that for every $n \ge 1$ and independent symmetric random vectors $X_1,\ldots,X_n$ and $Y_1,\ldots,Y_n$ in separable Banach space, the following is true: "If X_i is $(1,1)$-dominated by Y_i, $i \le n$, then $\sum X_i$ is (κ_0,λ_0)-dominated by $\sum Y_i$."Then for every $\kappa,\lambda \ge 1$, $n \ge 1$ and independent symmetric random vectors $X_1,\ldots,X_n$ and $Y_1,\ldots,Y_n$ in separable Banach space such that X_i is (κ,λ)-dominated by Y_i, $i \le n$, we have that $\sum X_i$ is $(\lceil\kappa\rceil\kappa_0, \lceil\kappa\rceil\lambda\lambda_0)$-dominated by $\sum Y_i$.

Proof.   Suppose that $X_i \prec_{(\kappa,\lambda)} Y_i$. The main idea is to take auxiliary random variables $\delta_{i,k}$, $i \le n$, $k \le \lceil\kappa\rceil$, independent of the X_i such that for each i,k, we have $\mathbb{P}\left(\delta_{i,k}=1\right) = \frac{1}{\lceil\kappa\rceil} = 1 - \mathbb{P}\left(\delta_{i,k}=0\right)$; moreover, $\sum_{k=1}^{\lceil\kappa\rceil}\delta_{i,k} = 1$ for each i, and the variables $\delta_{1,k},\ldots,\delta_{n,k}$ are independent for each k. For instance, we can define them on the probability space $[0,1]^n$

with Lebesgue measure by the formula

$$\delta_{i,k}(t_1, \ldots, t_n) = \mathbf{1}_{\left[\frac{k-1}{\lceil \kappa \rceil}, \frac{k}{\lceil \kappa \rceil}\right]}(t_i).$$

We check that for every i and k, we have $\delta_{i,k} X_i \prec_{(1,1)} \lambda Y_i$, so for every k, we obtain the comparison $\sum_i \delta_{i,k} X_i \prec_{(\kappa_0, \lambda_0)} \lambda \sum_i Y_i$ and thus,

$$\mathbb{P}\left(\left\|\sum_i X_i\right\| > 1\right) = \mathbb{P}\left(\left\|\sum_i \sum_k \delta_{i,k} X_i\right\| > 1\right) \leq \sum_{k=1}^{\lceil \kappa \rceil} \mathbb{P}\left(\lceil \kappa \rceil \left\|\sum_i \delta_{i,k} X_i\right\| > 1\right)$$

$$\leq \lceil \kappa \rceil \kappa_0 \mathbb{P}\left(\lceil \kappa \rceil \lambda \lambda_0 \left\|\sum_i Y_i\right\| > 1\right).$$

∎

The 2nd argument is based on the observation that if for some symmetric independent random vectors X_i and Y_i we have $X_i \prec_{(\kappa, \lambda)} Y_i$, $i \leq n$, then taking δ_i to be independent Bernoulli random variables such that $\mathbb{P}(\delta_i = 1) = 1/\kappa$, $\mathbb{P}(\delta_i = 0) = 1 - 1/\kappa$ and defining $X_i' = \delta_i X_i$, $Y_i' = \lambda Y_i$ we get $X_i' \prec_{(1,1)} Y_i'$. To obtain $\sum X_i \prec \sum Y_i$, we first apply (1) from Corollary 5 to the X_i' and Y_i', which gives

$$\int_0^1 \mathbb{P}\left(\mathbb{E}_\varepsilon \left(\left\|\sum \varepsilon_i X_i'\right\| - 1\right)_+ > t\right) dt \leq \int_0^1 \mathbb{P}\left(\mathbb{E}_\varepsilon \left(\left\|\sum \varepsilon_i Y_i'\right\| - 1\right)_+ > t\right) dt.$$

Then we bound the right-hand side by Lemma 7, but before using Lemma 6 for the lower bound of the left-hand side, we would like to pass from the X_i' to X_i. This can be achieved if we have an inequality like the following:

$$\mathbb{P}\left(\mathbb{E}_\varepsilon \left\|\sum \varepsilon_i \delta_i X_i\right\| > u\right) \geq c\mathbb{P}\left(c' \mathbb{E}_\varepsilon \left\|\sum \varepsilon_i X_i\right\| > u\right), \qquad u > 0,$$

where c and c' are some constants. This is possible thanks to a simple lemma that is in the spirit of the Paley–Zygmund inequality.

Lemma 9.   Let $v_1, \ldots, v_n$ be vectors in a separable Banach space $(E, \|\cdot\|)$, $p \in (0, 1]$ and let $\delta_1, \ldots, \delta_n$ be independent Bernoulli random variables with $\mathbb{P}(\delta_i = 1) = p$, $\mathbb{P}(\delta_i = 0) = 1 - p$. Then

$$\mathbb{P}_\delta\left(\mathbb{E}_\varepsilon \left\|\sum \varepsilon_i \delta_i v_i\right\| > 1\right) \geq \frac{p}{4} \mathbf{1}_{\{\mathbb{E}_\varepsilon \|\sum \varepsilon_i v_i\| > 2/p\}}.$$

By virtue of this lemma, we can take above $c = \frac{1}{4\kappa}$ and $c' = \frac{1}{2\kappa}$. After passing through Lemma 6 applied to the X_i, we conclude that $\sum X_i \prec_{(\kappa',\lambda')} \sum Y_i$ with $\kappa' = 64\alpha^{-1}\kappa$ and $\lambda' = 2(1+\alpha)\kappa\lambda$ for every $\alpha \in (0,1]$. We finish this section by showing the lemma.

Proof of Lemma 9.   Obviously, we can assume that $\mathbb{E}_\varepsilon \|\sum \varepsilon_i v_i\| > 2/p$ since, otherwise, there is nothing to prove. By Jensen's inequality,

$$\mathbb{E}_{\delta,\varepsilon} \left\| \sum \delta_i \varepsilon_i v_i \right\| \geq \mathbb{E}_\varepsilon \left\| \mathbb{E}_\delta \sum \delta_i \varepsilon_i v_i \right\| = p\mathbb{E}_\varepsilon \left\| \sum_i \varepsilon_i v_i \right\|,$$

thus $\frac{1}{2}\mathbb{E}_\delta \mathbb{E}_\varepsilon \|\sum \delta_i \varepsilon_i v_i\| \geq \frac{p}{2}\mathbb{E}_\varepsilon \|\sum \varepsilon_i v_i\| > 1$, so

$$\mathbb{P}_\delta \left(\mathbb{E}_\varepsilon \left\| \sum \varepsilon_i \delta_i v_i \right\| > 1 \right) \geq \mathbb{P}_\delta \left(\mathbb{E}_\varepsilon \left\| \sum \varepsilon_i \delta_i v_i \right\| > \frac{1}{2}\mathbb{E}_\delta \mathbb{E}_\varepsilon \left\| \sum \delta_i \varepsilon_i v_i \right\| \right)$$

$$\geq \frac{1}{4} \frac{\left(\mathbb{E}_\delta \mathbb{E}_\varepsilon \left\| \sum \delta_i \varepsilon_i v_i \right\| \right)^2}{\mathbb{E}_\delta \left(\mathbb{E}_\varepsilon \left\| \sum \delta_i \varepsilon_i v_i \right\| \right)^2},$$

where in the last estimate we used the Paley–Zygmund inequality. Using the contraction principle (4) we obtain (ε_i' denote independent copies of ε_i)

$$\mathbb{E}_\delta \left(\mathbb{E}_\varepsilon \left\| \sum \delta_i \varepsilon_i v_i \right\| \right)^2 = \mathbb{E}_\delta \left(\mathbb{E}_\varepsilon \left\| \sum \delta_i \varepsilon_i v_i \right\| \mathbb{E}_{\varepsilon'} \left\| \sum \delta_i \varepsilon_i' v_i \right\| \right)$$

$$\leq \left(\mathbb{E}_\varepsilon \left\| \sum \varepsilon_i v_i \right\| \right) \left(\mathbb{E}_{\delta,\varepsilon} \left\| \sum \delta_i \varepsilon_i v_i \right\| \right) \leq \frac{1}{p} \left(\mathbb{E}_{\delta,\varepsilon} \left\| \sum \delta_i \varepsilon_i v_i \right\| \right)^2.$$

This, combined with the previous inequality, finishes the proof.   ■

4 Proof of Theorem 2

Suppose $X_1, \ldots, X_n$ are independent symmetric random vectors and each satisfies $WB(C, \delta, \theta)$. Let $S_n = X_1 + \ldots + X_n$. Fix a continuous norm $\|\cdot\|$. We would like to show that

$$\mathbb{P}\left(\|S_n\| > \lambda \right) \leq C'\lambda^{-\delta}\mathbb{P}\left(\|S_n\| > 1 \right), \qquad \lambda \geq 1,$$

provided that $\mathbb{P}\left(\|S_n\| > 1 \right) < \theta'$. (We shall find the values of the constants C' and θ' as the argument goes along.) First observe that if $\theta' \leq \theta/2$, then by (6) we also have that

$$\mathbb{P}\left(\|X_j\| > 1 \right) \leq \mathbb{P}\left(X_n^* > 1 \right) \leq 2\mathbb{P}\left(\|S_n\| > 1 \right) < 2\theta' \leq \theta,$$

where $X_n^* = \max_{j \leq n} \|X_j\|$. This will let us use the WB inequality for $X_j, j = 1, \ldots, n$.

Let $p_k = \mathbb{P}\left(\|S_n\| > 3^k\right)$ for $k = 0, 1, \ldots$. Our 1st goal is to establish that $p_k \leq C' \cdot 3^{-\delta k} p_0$, assuming $p_0 \leq \theta'$. Then, possibly increasing C', we will get that $\mathbb{P}\left(\|S_n\| > \lambda\right) \leq C' \lambda^{-\delta} \mathbb{P}\left(\|S_n\| > 1\right)$ for every $\lambda \geq 1$. We begin with deriving a recursive inequality for the p_k. Fix $k \geq 1$. By (5)–(8) and the union bound, we obtain

$$p_k = \mathbb{P}\left(\|S_n\| > 3 \cdot 3^{k-1}\right) \leq \mathbb{P}\left(X_n^* > 3^{k-1}\right) + 2\mathbb{P}\left(S_n^* > 3^{k-1}\right)\mathbb{P}\left(\|S_n\| > 3^{k-1}\right)$$

$$\leq \sum_{j=1}^n \mathbb{P}\left(\|X_j\| > 3^{k-1}\right) + 4p_{k-1}^2$$

$$\leq C \cdot 3^{-\delta(k-1)} \sum_{j=1}^n \mathbb{P}\left(\|X_j\| > 1\right) + 4p_{k-1}^2$$

$$\leq C \cdot 3^{-\delta(k-1)} \frac{2p_0}{1 - 2p_0} + 4p_{k-1}^2.$$

If we assume additionally that $\theta' \leq 1/3$, then $\frac{1}{1-2p_0} \leq 3$, so

$$p_k \leq 6C \cdot 3^{-\delta(k-1)} p_0 + 4p_{k-1}^2, \qquad k \geq 1.$$

Let us prove inductively that $p_k \leq (12 \cdot 3^\delta C) \cdot 3^{-k\delta} p_0$, $k \geq 0$. For $k = 0$ this is obvious. Suppose it holds for $k - 1$, for some $k \geq 1$. By the recursive inequality,

$$p_k \leq 6C \cdot 3^{-\delta(k-1)} p_0 + 4(12 \cdot 3^\delta C)^2 \cdot 3^{-2\delta(k-1)} p_0^2$$

$$= (12 \cdot 3^\delta C) \cdot 3^{-k\delta} p_0 \cdot \left(\frac{1}{2} + 48C \cdot 3^{-\delta k + 3\delta} p_0\right)$$

$$\leq (12 \cdot 3^\delta C) \cdot 3^{-k\delta} p_0 \cdot \left(\frac{1}{2} + 48C \cdot 3^{2\delta} \theta'\right),$$

and we get the inductive assertion as long as $\theta' \leq (96C \cdot 9^\delta)^{-1}$. Therefore, we set $\theta' = \min\{\theta/2, (96C \cdot 9^\delta)^{-1}\}$. Then, as we have shown,

$$\mathbb{P}\left(\|S_n\| > \lambda\right) \leq (12 \cdot 3^\delta C) \cdot \lambda^{-\delta} \mathbb{P}\left(\|S_n\| > 1\right),$$

for $\lambda = 3^k$, $k \geq 0$. It remains to extend this to any $\lambda \geq 1$. If $1 \leq \lambda < 3$, then trivially

$$\mathbb{P}\left(\|S_n\| > \lambda\right) \leq \mathbb{P}\left(\|S_n\| > 1\right) \leq 3^\delta \lambda^{-\delta} \mathbb{P}\left(\|S_n\| > 1\right).$$

If $3^k \leq \lambda < 3^{k+1}$ for some $k \geq 1$, we get

$$\mathbb{P}\left(\|S_n\| > \lambda\right) \leq \mathbb{P}\left(\|S_n\| > 3^k\right) \leq (12 \cdot 3^\delta C) \cdot 3^{-k\delta} \mathbb{P}\left(\|S_n\| > 1\right)$$

$$\leq (12 \cdot 3^{2\delta} C) \cdot \lambda^{-\delta} \mathbb{P}\left(\|S_n\| > 1\right).$$

We set $C' = 12 \cdot 9^\delta C$ and the proof is complete.

5 Proof of Theorem 3

Since the sequence a is majorised by b, there are nonnegative weights λ_σ adding up to 1 indexed by all permutations σ of the n-element set $\{1, \ldots, n\}$ such that $a = \sum_\sigma \lambda_\sigma b_\sigma$, where $b_\sigma = (b_{\sigma(1)}, \ldots, b_{\sigma(n)})$ is the sequence b permuted according to σ. It easily follows that for every convex function $\varphi \colon E \to \mathbb{R}$ defined on the Banach space E, the X_i take values in, we have

$$\mathbb{E}\varphi\left(\sum a_i X_i\right) = \mathbb{E}\varphi\left(\sum_i \sum_\sigma \lambda_\sigma b_{\sigma(i)} X_i\right) \leq \sum_\sigma \lambda_\sigma \mathbb{E}\varphi\left(\sum_i b_{\sigma(i)} X_i\right) = \mathbb{E}\varphi\left(\sum b_i X_i\right)$$

(provided the expectations exist).

Notice that since each $b_i X_i$ satisfies WB(C, δ, θ), by Theorem 2, the sum $\sum b_i X_i$ satisfies WB(C', δ, θ'), where $C' = 12 \cdot 9^\delta C$ and $\theta' = \min\{\theta/2, (96C \cdot 9^\delta)^{-1}\}$.

Let $\|\cdot\|$ be a continuous norm on E. Denote $S_a = \|\sum a_i X_i\|$ and $S_b = \|\sum b_i X_i\|$. We want to show that $\mathbb{P}\left(S_a > 1\right) \leq \kappa \mathbb{P}\left(\lambda S_b > 1\right)$. If $\mathbb{P}\left(S_b > 1\right) \geq \theta'$, then we trivially get

$$\mathbb{P}\left(S_a > 1\right) \leq 1 = \frac{1}{\theta'}\theta' \leq \frac{1}{\theta'}\mathbb{P}\left(S_b > 1\right).$$

Suppose that $\mathbb{P}\left(S_b > 1\right) < \theta'$. Using the initial observation for $\varphi(x) = (\|x\| - 1)_+$ we get

$$\mathbb{P}\left(S_a > 2\right) = \mathbb{P}\left((S_a - 1)_+ > 1\right) \leq \mathbb{E}(S_a - 1)_+ \leq \mathbb{E}(S_b - 1)_+.$$

By the weak Borell inequality for S_b, we get

$$\mathbb{E}(S_b - 1)_+ = \int_1^\infty \mathbb{P}\left(S_b > \lambda\right) d\lambda \leq \int_1^\infty C' \lambda^{-\delta} \mathbb{P}\left(S_b > 1\right) d\lambda = \frac{C'}{\delta - 1}\mathbb{P}\left(S_b > 1\right).$$

In summary, we have showed that for any continuous norm $\|\cdot\|$ on E, we have

$$\mathbb{P}\left(\left\|\sum a_i X_i\right\| > 2\right) \leq \kappa \mathbb{P}\left(\left\|\sum b_i X_i\right\| > 1\right)$$

with $\kappa = \max\{\frac{1}{\theta'}, \frac{C'}{\delta-1}\}$. Rescaling the norm finishes the proof.

Funding

Acknowledgments

References

[1]   Anderson, T. "The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities." *Proc. Amer. Math. Soc.* 6 (1955): 170–6.

[2]   Ball, K. "Convex geometry and functional analysis." In *Handbook of the Geometry of Banach Spaces*, vol. I, 161–94. Amsterdam: North-Holland, 2001.

[3]   Bhatia, R., *Matrix Analysis. Graduate Texts in Mathematics*, 169. New York: Springer, 1997.

[4]   Borell, Ch. "Convex measures on locally convex spaces." *Ark. Mat.* 12 (1974): 239–52.

[5]   Borell, Ch. "Convex set functions in d-space." *Period. Math. Hungar.* 6, no. 2 (1976): 111–36.

[6]   Birnbaum, Z. W. "On random variables with comparable peakedness." *Ann. Math. Statistics* 19 (1948): 76–81.

[7]   Jain, N. C. and M. B. Marcus. "Integrability of infinite sums of independent vector-valued random variables." *Trans. Amer. Math. Soc.* 212 (1975): 1–36.

[8]   Kahane, J.-P. *Some Random Series of Functions*, 2nd ed. Cambridge Studies in Advanced Mathematics, Vol. 5, Cambridge: Cambridge University Press, 1985.

[9]   Kanter, M. "Unimodality and dominance for symmetric random vectors." *Trans. Amer. Math. Soc.* 229 (1977): 65–85.

[10] Kwapień, S. and W. Woyczyński. *Random Series and Stochastic Integrals: Single and Multiple. Probability and its Applications*. Boston, MA: Birkhäuser Boston, Inc., 1992.

[11] Latała, R. and K. Oleszkiewicz. "On the best constant in the Khinchin-Kahane inequality." *Studia Math.* 109, no. 1 (1994): 101–4.

[12] Ledoux, M. and M. Talagrand. *Probability in Banach Spaces*. Berlin: Springer, 1991.

[13] Meyer M. and A. Pajor. "Sections of the unit ball of L_p^n." *J. Funct. Anal.* 80, no. 1 (1988): 109–23.

[14] Sherman, S. "A theorem on convex sets with applications." *Ann. Math. Statist.* 26 (1955): 763–7.

[15] Vaaler, J. D. "A geometric inequality with applications to linear forms." *Pacific J. Math.* 83, no. 2 (1979): 543–53.