University of Warsaw
Faculty of Mathematics, Informatics and Mechanics

Marta Strzelecka

Estimates for moments of random vectors

PhD dissertation

Supervisor
prof. dr hab. Rafał Latała
Institute of Mathematics
University of Warsaw

Auxiliary Supervisor
dr Piotr Nayar
Institute of Mathematics
University of Warsaw

January 2019

Author's declaration:
I hereby declare that this dissertation is my own work.

January 31, 2019 ....

Marta Strzelecka

Supervisor's declaration:
The dissertation is ready to be reviewed

January 31, 2019 ....

prof. dr hab. Rafał Latała

Abstract

This dissertation is devoted to estimates of moments of norms of random vectors. It consists of four main results.

In the first part we show that for $p \geq 1$ and $r \geq 1$ the p-th moment of the ℓ_r-norm of a log-concave random vector is comparable to the sum of the first moment and the weak p-th moment up to a constant proportional to r. This extends the previous result of Paouris concerning Euclidean norms.

The second main result states that for $p \geq 1$, the p-th moments of suprema of linear combinations of independent centered random variables are comparable with the sum of the first moment and the weak p-th moment provided that the $2q$-th and q-th integral moments of these variables are comparable for all $q \geq 2$. The latter condition turns out to be necessary in the i.i.d. case.

In the next part we show that every symmetric random variable with log-concave tails satisfies the convex infimum convolution inequality with an optimal cost function (up to scaling). As a result, we obtain nearly optimal comparison of weak and strong moments for symmetric random vectors having independent coordinates with log-concave tails.

The last main result is an estimate of $\mathbb{E}\|X\|_{\ell_{p'} \to \ell_q}$ for $p, q \geq 2$, where X is a random matrix, which entries are of the form $a_{ij} Y_{ij}$, where Y has i.i.d. isotropic log-concave rows. This generalises the result of Guédon, Hinrichs, Litvak, and Prochno for Gaussian matrices with independent entries. Our estimate is optimal up to logarithmic factors.

2010 Mathematics Subject Classification. Primary: 60E15. Secondary: 26A51, 26B25.

Keywords and phrases: log-concave vectors, comparison of weak and strong moments, infimum convolution, norms of random matrices.

Streszczenie

Ta rozprawa poświęcona jest oszacowaniom momentów norm wektorów losowych. Składa się ona z czterech głównych wyników.

W pierwszej części pokazujemy, że dla $p \geq 1$ i $r \geq 1$, p-ty moment normy ℓ_r log-wklęsłego wektora losowego jest porównywalny z sumą pierwszego momentu i słabego p-tego momentu, z dokładnością do stałej proporcjonalnej do r. Jest to uogólnienie uzyskanego wcześniej przez Paourisa oszacowania dla norm euklidesowych.

Drugi główny wynik orzeka, że dla $p \geq 1$, p-ty moment supremów liniowych kombinacji niezależnych scentrowanych zmiennych losowych jest porównywalny z sumą pierwszego momentu i słabego p-tego momentu, o ile $2q$-te i q-te momenty całkowe tych zmiennych są porównywalne dla każdego $q \geq 2$. Ten drugi warunek okazuje się być konieczny w przypadku wektorów o współrzędnych niezależnych o jednakowych rozkładach.

W kolejnej części wykazujemy, że każda symetryczna zmienna losowa o log-wklęsłych ogonach spełnia wypukłą nierówność splotu infimum z optymalną (z dokładnością do skalowania) funkcją kosztu. Jako wniosek otrzymujemy niemal optymalne porównywanie słabych i silnych momentów dla symetrycznych wektorów losowych o niezależnych współrzędnych o log-wklęsłych ogonach.

Ostatnim głównym wynikiem jest oszacowanie $\mathbb{E}\|X\|_{\ell_{p'} \to \ell_q}$ dla $p, q \geq 2$ i macierzy losowej X, której wyrazy mają postać $a_{ij}Y_{ij}$, gdzie Y jest macierzą o niezależnych wierszach o tym samym izotropowym i log-wklęsłym rozkładzie. Uogólnia to wynik Guédona, Hinrichsa, Litvaka i Prochny dla macierzy gaussowskich o niezależnych wyrazach. Nasze oszacowanie jest optymalne z dokładnością do czynników logarytmicznych z wymiaru.

Klasyfikacja tematyczna. 60E15; 26A51, 26B25.

Słowa kluczowe: wektory log-wklęsłe, porównywanie słabych i silnych momentów, splot infimum, normy macierzy losowych.

Contents

Chapter 1

Introduction

This dissertation is devoted to estimates of norms of some natural classes of random vectors in $\mathbb{R}^n$. Dimension-free bounds are of most interest, since they may be generalised to infinite-dimensional spaces. However, if the dependence on the dimension is mild (especially if an estimate depends only on the logarithm of the dimension), a bound is useful too and gives us a better understanding of the behaviour of the class of random vectors we investigate. Let us describe three types of estimates we are dealing with in this thesis.

In convex geometry the class of log-concave vectors is often investigated. One of the classical theorems concerning this class is the Paouris inequality from [29], which gives estimates of the standard Euclidean norm of any isotropic log-concave random vector, and in a version from [1] also of arbitrary log-concave vector (or, equivalently, it provides estimates for any Euclidean norm). It is natural to ask if this result can be generalised to other norms or any wider class of vectors. In the first part of this dissertation we partially answer this question.

Our first main result says that an analogue of the Paouris inequality holds for the ℓ_r-norm of any log-concave vector, with a constant depending linearly on r. This comes from the joint work with Rafał Latała [22], which is presented in Sections 2.1.1 and 2.2. If the constant in such an estimate depended linearly on r^γ (instead of r), this would imply a non-asymptotic bound for any norm, with a constant Cn^γ (an estimate with the constant $C\sqrt{n}$ may be gained easily). Moreover, if the dependence of r was lost, then the bound with a universal constant would hold. However, our bound with a constant Cr yields strong corollaries too (see Section 2.1.1) – among others we use it in Chapter 4 to obtain almost optimal estimates for log-concave random matrices.

The second main result comes from another joint work with Rafał Latała [23] and is presented in details and proved in Sections 2.1.2 and 2.3. We characterise all centred random variables X_1, for which every vector $X = (X_1, \ldots, X_n)$ with i.i.d. coordinates satisfies the generalisation of Paouris inequality for any norm in

$\mathbb{R}^n$. The equivalent condition may be expressed easily in the language of growth of integral moments of X_i. Moreover, we provide the same estimate for any X with independent coordinates satisfying the same moments growth condition.

Another important inequality in high dimensional probability is the infimum convolution inequality and the convex infimum convolution inequality (convex ICI for short). They appear naturally in the research connected to the concentration of measure and the theory of optimal transport. Their concentration counterparts gives some estimates for norms of vectors, as may be seen in the second part of this dissertation in the case of convex ICI with optimal cost function, which implies a strong enough concentration to provide a Paouris inequality-like estimate. Our third main result, based on the joint work with Michał Strzelecki and Tomasz Tkocz [35], is that vectors with independent coordinates with log-concave tails satisfy the convex ICI with the optimal cost function (i.e. in a sense the optimal possible convex ICI). The content of [35] may be found in Chapter 3.

A special type of norms are operator norms of matrices (an $m \times n$-dimensional vector may be treated as an $m \times n$ matrix). We are interested in estimating the expected value of the operator norm from ℓ_p^n to ℓ_q^m of certain random matrices. Most results concerning this quantity deal with the spectral norm only (i.e. the operator norm from ℓ_2^n to ℓ_2^m). Moreover, in the vast majority of known results one has to assume the independence of entries of the matrix. Chapter 4, which is part of a work in progress [34] by the author, provides an estimate, which is optimal up to logarithmic factors and is valid for weighted matrices with i.i.d. isotropic log-concave rows. In particular, we do not require the entries of the matrix to be independent. To obtain the results from Chapter 4 we use theorems from the first three chapters of this thesis.

1.1 Notation

By C we denote universal constants. If a constant C depends on a parameter α, we express it as $C(\alpha)$. The value of $C, C(\alpha)$ may differ at each occurrence. Whenever we want to fix the value of an absolute constant we use letters $C_1, C_2, \ldots$. We may always assume that $C_i \geq 1$.

For a random variable X we denote by $\|X\|_p$ the p-th integral norm of X, i.e. the quantity $(\mathbb{E}|X|^p)^{1/p}$. For a vector $x \in \mathbb{R}^n$ (in particular for a random vector X) and $r \geq 1$, by $\|x\|_r$ we denote the ℓ_r-norm of x, i.e. $\|x\|_r := (\sum_{i=1}^n |x_i|^r)^{1/r}$. For $r = 2$ we shall also write $|\cdot|$ instead of $\|\cdot\|_2$. It will be always clear from the context, what $\|X\|_q$ means for a random object X, so the double meaning of $\|\cdot\|_q$ will not lead to any misunderstanding. For an $m \times n$ matrix A by $\|A\|_{p,q}$ we denote its norm from ℓ_p^n to ℓ_q^m.

For a given norm $\|\cdot\|$, $B_{\|\cdot\|}$ denotes the unit ball in this norm and $\|\cdot\|^*$ denotes the dual norm of $\|\cdot\|$. Recall that the space dual to $(\mathbb{R}^n, \|\cdot\|)$ is isomorphic to $(\mathbb{R}^n, \|\cdot\|^*)$ (we may identify a functional $\varphi \in (\mathbb{R}^n, \|\cdot\|)^*$ with a vector $y \in (\mathbb{R}^n, \|\cdot\|^*)$ via the scalar product, such that for any $x \in \mathbb{R}^n$ we have $\varphi(y) = \langle y, x \rangle$).

For $p \in [1, \infty]$, B_p^n denotes the (closed) unit ball in the norm $\|\cdot\|_p$ in $\mathbb{R}^n$. We will usually denote the Hölder conjugate of p by p' (with the convention $\infty' = 1$ and $1' = \infty$), i.e. p' satisfies $1 = \frac{1}{p} + \frac{1}{p'}$. If E is a normed linear space, then by $\|\cdot\|_E$ we denote the norm on E, and by B_E we denote the closed unit ball in this norm.

By $|I|$ we denote the cardinality of a finite set I. For an n-dimensional random vector Z and $a \in \mathbb{R}^n$ we write aZ for the vector $(a_i Z_i)_i$. Observe that $\mathbb{E}\|aZ\|_2^2 = \sum_i a_i^2 \mathbb{E} Z_i^2$.

The symbol $\sim$ denotes either equal distributions of two random variables or the comparability of two positive quantities (i.e. $a \sim b$ if there exist an absolute constant C such that $aC^{-1} \le b \le Ca$).

For a given sequence $(x_i)_{i=1}^n$ of real numbers we denote by $(x_i^*)_{i=1}^n$ the non-increasing rearrangement of the sequence $(|x_i|)_{i=1}^n$.

1.2 Preliminaries

We say that $K \subset \mathbb{R}^n$ is a *convex body* if K is convex, compact and has nonempty interior.

A measure μ on a locally convex linear space F is called *logarithmically concave* (*log-concave* in short) if for any compact nonempty sets $K, L \subset F$ and $\lambda \in [0,1]$,

$$\mu(\lambda K + (1-\lambda)L) \ge \mu(K)^\lambda \mu(L)^{1-\lambda}.$$

A random vector with values in F is called log-concave if its distribution is logarithmically concave. The class of log-concave measures is closed under linear transformations, convolutions and weak limits. By the result of Borell [5] a d-dimensional vector with a full dimensional support is log-concave if and only if it has a log-concave density, i.e. a density of the form e^{-h}, where h is a convex function with values in $(-\infty, \infty]$.

We say that a vector X in $\mathbb{R}^n$ is *isotropic* if $\operatorname{Cov} X = \operatorname{Id}$ (recall that $\operatorname{Cov} X$ is the $n \times n$ matrix with entries $\operatorname{Cov}(X_i, X_j)$). If X is a log-concave random vector in $\mathbb{R}^n$ with full dimensional support, then there exists a linear transformation T such that $\operatorname{Cov}(TX) = \operatorname{Id}$ – then we say that TX is *an isotropic position* of X.

The class of log-concave measures is a natural generalization of uniform measures over convex bodies (these measures are log-concave, since they have log-concave densities). Moreover, any log-concave measure can be obtained as a weak limit of projections of uniform measures over (higher dimensional) convex bodies (see e.g. [2]). On the other hand, Ball (in [3]) introduced bodies $K_p(f)$ associated with a measurable function f such that $f(0) > 0$ (we skip the definition and the details, since we will not need them in further chapters). For a convex body K we know that $K_p(\mathbf{1}_K) = K$ for all $p > 0$. Moreover for a log-concave function f the body $K_p(f)$ is convex and $K_{n+2}(f)$ has the isotropic constant comparable with the isotropic constant of f (i.e. the quantity $L_f := (f(\mathbb{E}X))^{1/n}(\det \mathrm{Cov}(X))^{1/2n}$, where X has the density f). In particular it suffices to estimate the isotropic constant for convex bodies in order to investigate the isotropic constant conjecture[1]. Other links between log-concave measures and convex bodies are described in [2], and other results and conjectures about log-concave measures are discussed in the recently published monograph [6].

We say that a random vector X is unconditional if it has the same distribution as ηX for every $\eta \in \{-1,1\}^n$ or equivalently, if X has the same distribution as εX, where ε_i are i.i.d. symmetric Bernoulli random variables (i.e. $\mathbb{P}(\varepsilon_i = 1) = \mathbb{P}(\varepsilon_i = -1) = \frac{1}{2}$). Similarly we say that a subset T of $\mathbb{R}^n$ is unconditional, if it is symmetric with respect to all coordinates axes, i.e. $t \in T$ if and only if $\eta t \in T$ for every $\eta \in \{-1,1\}^n$. We also say that a norm $\|\cdot\|$ on $\mathbb{R}^n$ is unconditional, if its unit ball is unconditional or equivalently if $\|x\| = \|\eta x\|$ for every $x \in \mathbb{R}^n$ and every $\eta \in \{-1,1\}^n$.

Note that the unconditionality is a much stronger property than the isotropicity in a sense that there may not exist any linear transformation of a given random vector X that makes it unconditional, even if X has a full dimensional support.

We will also consider random variables *with log-concave tails*, i.e. variables X for which the function

$$t \mapsto N(t) := -\ln \mathbb{P}(|X| \geq t), \quad t \geq 0,$$

is convex. Note that the definition of log-concavity implies that log-concave variables have log-concave tails[2].

Let us recall a few basic facts about log-concave vectors and vectors with log-concave tails.

Definition 1.1. We say that a random variable Z is *α-regular* (for $\alpha \geq 1$) if

$$\|Z\|_q \leq \alpha \frac{q}{p} \|Z\|_p \quad \text{for all } q \geq p \geq 2.$$

Then we also say that *the moments of Z grow α-regularly.*

Remark 1.2. If X is a symmetric random variable with log-concave tails, then its moments grow 1-regularly (this classical fact follows for instance from Proposition 5.5 from [13] and the proof of Proposition 3.8 from [26]).

[1]The isotropic constant conjecture states that the isotropic constant of any log-concave isotropic vector is bounded by an absolute constant.

[2]Moreover, the class of variables with log-concave tails is *strictly* larger than the class of log-concave random variables.

The above remark implies that a log-concave symmetric random variable Z is 1-regular. Thus if Z is a log-concave centred (i.e. $\mathbb{E}Z = 0$) random variable and Z' is its independent copy, then for $q \geq p \geq 2$ we have

$$\|Z\|_q = \|Z - \mathbb{E}Z'\|_q \leq \|Z - Z'\|_q \leq \frac{q}{p}\|Z - Z'\|_p \leq 2\frac{q}{p}\|Z\|_p$$

so Z is 2-regular. Moreover, if Z is an arbitrary log-concave random variable, we have

$$\|Z\|_q \leq \|Z - \mathbb{E}Z'\|_q + \mathbb{E}|Z'| \leq (q+1)\|Z\|_2,$$

so we get by Chebyshev's inequality $\mathbb{P}(|Z| \geq e(p+1)\|Z\|_2) \leq e^{-p}$ for $p \geq 2$. Thus

$$\mathbb{P}(\|Z\| \geq t) \leq \exp\Big(2 - \frac{t}{2e\|Z\|_2}\Big) \quad \text{for } t > 0. \tag{1.1}$$

It is easy to see that the definition of log-concavity implies that if an n-dimensional symmetric random vector X is log-concave, then for any $t \in \mathbb{R}^n$ the variable $\langle t, X\rangle$ is also log-concave and symmetric, so it is also $1 - regular$.

Moreover, if $f : \mathbb{R}^n \to \mathbb{R}$ is a seminorm,

$$(\mathbb{E}f(Z)^p)^{1/p} \leq C_1\frac{p}{q}(\mathbb{E}f(Z)^q)^{1/q} \tag{1.2}$$

for $p \geq q \geq 1$ (see [6, Theorem 2.4.6]).

If K is a convex body in $\mathbb{R}^n$, X is a log-concave vector in $\mathbb{R}^n$ such that $\mathbb{P}(X \in K) > 0$, and $A := \{X \in K\}$ then the vector Y defined by

$$\mathbb{P}(Y \in B) = \frac{\mathbb{P}(A \cap \{X \in B\})}{\mathbb{P}(A)} = \frac{\mathbb{P}(X \in B \cap K)}{\mathbb{P}(X \in K)},$$

is log-concave. It follows immediately by the definition of log-concavity. We say that Y is distributed as X conditioned on K.

We will also need the inequality $(\frac{n}{k})^k \leq \binom{n}{k} \leq (\frac{en}{k})^k$, valid for $1 \leq k \leq n$.

For a non-decreasing function $g : \mathbb{R} \to \mathbb{R}$ we define its *generalized inverse* $g^{-1} : \mathbb{R} \to (-\infty, \infty]$ by a formula $g^{-1}(y) := \inf\{x : g(x) \geq y\}$. Note that if g is continuous, then $g(g^{-1}(y)) = y$ for all $y \in g(\mathbb{R})$.

Let us make a remark which will be used multiple times in next chapters.

Remark 1.3. Let us justify that for any nonempty set T and any random vector X we have

$$\lim_{p\to\infty} \sup_{t\in T} \|\langle t, X\rangle\|_p = \sup_{t\in T} \operatorname{ess\,sup} |\langle t, X\rangle|. \tag{1.3}$$

If moreover we can control the growth of integral moments of $\langle t, X\rangle$ (for example if $\|\langle t, X\rangle\|_{2p} \leq \alpha\|\langle t, X\rangle\|_p$ for every $p \geq p_0$ and $t \in T$), then $p \mapsto \sup_{t\in T} \|\langle t, X\rangle\|_p$ is continuous for $p \geq p_0$.

Indeed, by Hölder inequality we get

$$\mathbb{E}|\langle t,X\rangle|^{p+\varepsilon} \le \left(\mathbb{E}|\langle t,X\rangle|^{p}\right)^{\frac{p-\varepsilon}{p}} \left(\mathbb{E}|\langle t,X\rangle|^{2p}\right)^{\frac{\varepsilon}{p}} \le \alpha^{2\varepsilon}\|\langle t,X\rangle\|_p^{p+\varepsilon},$$

so $\sup_{t\in T}\|\langle t,X\rangle\|_{p+\varepsilon} \le \alpha^{2\varepsilon/p}\sup_{t\in T}\|\langle t,X\rangle\|_p \le \alpha^{2\varepsilon/p}\sup_{t\in T}\|\langle t,X\rangle\|_{p+\varepsilon}$ for $p \ge p_0$.

In order to prove (1.3) recall that $\lim_{p\to\infty}\|\langle t,X\rangle\|_p = \operatorname{ess\,sup}|\langle t,X\rangle|$. In the case $\sup_{t\in T}\operatorname{ess\,sup}|\langle t,X\rangle| =: A < \infty$ take $t^{(i)} \in T$ such that $\operatorname{ess\,sup}|\langle t^{(i)},X\rangle| \ge A - 1/i$ and then p_i such that $\|\langle t^{(i)},X\rangle\|_{p_i} \ge \operatorname{ess\,sup}|\langle t^{(i)},X\rangle| - 1/i$. Then $A \ge \sup_{t\in T}\|\langle t,X\rangle\|_{p_i} \ge A - 2/i$. In the case when $A = \infty$ we proceed similarilly.

1.3 Acknowledgements

I am grateful to my advisor, professor Rafał Latała for his patience and guidance, for the unbounded amount of time he was ready to give me, for numerous interesting articles and facts he pointed out to me, and for all the interesting problems he shared with me and the discussions around them.

I would also like to thank all the people who were keeping me interested in mathematics, providing a great atmosphere in everyday study and work, and sharing with me a wide range of feellings associated with the beauty, joyfulness and difficulty of mathematics.

The research presented in this thesis was partially supported by the National Science Centre, Poland, via the Opus grant no. 2012/05/B/ST1/00412, the Preludium grant no. 2015/19/N/ST1/02661, and the Etiuda grant no. 2018/28/T/ST1/00001.

Chapter 2

Comparison of weak and strong moments

One of the fundamental properties of log-concave vectors is the Paouris inequality [29] (see also [1] for a shorter proof). It states that for a log-concave vector X in $\mathbb{R}^n$,

$$(\mathbb{E}\|X\|_2^p)^{1/p} \leq C_2 \left((\mathbb{E}\|X\|_2^2)^{1/2} + \sigma_X(p)\right) \quad \text{for } p \geq 1, \tag{2.1}$$

where

$$\sigma_X(p) := \sup_{\|t\|_2 \leq 1} \left(\mathbb{E}\Big|\sum_{i=1}^n t_i X_i\Big|^p\right)^{1/p}$$

is the Euclidean weak $p-$th moment of X. We call the quantity $(\mathbb{E}\|X\|_2^p)^{1/p}$ the p-th strong moment of X (with respect to the Euclidean norm).

It is natural to ask whether inequality (2.1) may be generalized to non-Euclidean norms. In [19] Latała formulated and discussed the following conjecture.

Conjecture 2.1. There exists a universal constant C such that for any log-concave vector X with values in a finite dimensional normed space $(F, \| \ \|)$,

$$(\mathbb{E}\|X\|^p)^{1/p} \leq C\Big(\mathbb{E}\|X\| + \sup_{\varphi \in F^*, \|\varphi\|_* \leq 1} (\mathbb{E}|\varphi(X)|^p)^{1/p}\Big) \quad \text{for } p \geq 1. \tag{2.2}$$

Note that a reverse inequality with the constant $\frac{1}{2}$ always holds, since by the Jensen inequality

$$(\mathbb{E}\|X\|^p)^{1/p} \geq \max\Big\{\mathbb{E}\|X\|, \sup_{\varphi \in F^*, \|\varphi\|_* \leq 1} (\mathbb{E}|\varphi(X)|^p)^{1/p}\Big\}.$$

Therefore we may think that the conjecture above states that weak and strong moments of norms of log-concave vectors are comparable. For a given normed space $(F, \| \ \|)$ by a weak p-th (for $p \geq 1$) moment we mean

$$\sigma_{\|\cdot\|,X}(p) := \sup_{\varphi \in F^*, \|\varphi\|_* \leq 1} \left(\mathbb{E}\varphi(X)^p\right)^{1/p}.$$

Today we only know that Conjecture 2.1 holds in some special cases, and we do not know any possible counterexample. In [19, Section 3] Latała proved that for n-dimensional spaces F inequality (2.2) is true with an additional factor $\log n$ in front of $\mathbb{E}\|X\|$ if we assume additionally that X is unconditional . He also proved there that we can skip $\log n$ if $(\mathbb{R}^n, \| \cdot \|)$ has nontrivial cotype q. In this case C depends on q and the cotype constant T_q^*, and we still have to assume that X is unconditional. Moreover, [19, Corollary 2.4] states that (2.2) holds with a universal constant for log-concave vectors X with independent coordinates. Of course, Conjecture 2.1 is true for norms on $\mathbb{R}^n$, the unit balls of which are ellipsoids, since the Paouris inequality (2.1) holds and the linear transformation of a log-concave vector is a log-concave vector. This was observed in [1, Section 3] – the authors proved the theorem without the unnceccesary assumption that X is in the isotropic position, which appeared in the original work of Paouris.

There are two links between Conjecture 2.1 and other problems in convex and high dimensional probability. Latała proved in [21] that for all vectors X satisfying Sudakov Minoration Principle with constant κ the comparison of weak and strong moments holds for every norm on $\mathbb{R}^d$ up to a factor $\ln(ed/p)/\kappa$ at $\mathbb{E}\|X\|$. Moreover, due to Latała and Wojtaszczyk (see [26]) the optimal concentration (or equivalently the convolution inequality with optimal cost function) of the distribution of X implies (2.2). As Strzelecki, Tkocz, and the author noticed in [35], inequality (2.2) follows even by a weaker *convex* infimum convolution inequality with optimal cost function. We postpone further details and definitions to Chapter 3.

It is also interesting to find more general assumptions than log-concavity under which (2.2) holds in some special cases. Latała and Tkocz proved in [24, Theorem 2.3] that for vectors with independent coordinates we may indeed assume less then the log-concavity for (2.2) to hold. This weaker assumption is the α-regularity of growth of moments of coordinates of X[1] (then the constant C depends on α). However, in the case of dependent coordinates the α-regularity of growth of moments of $\langle t, X\rangle$ (for all $t \in \mathbb{R}^n$) does not imply (2.2) even for the Euclidean norm as the example below shows.

Example 2.2. Let g be a standard Gaussian random variable and let G be a standard n-dimensional Gaussian vector independent of g. Consider $X := gG$. For every $t \in \mathbb{R}^n$ we have $\langle t, G\rangle \sim |t|G_1$, so $\|\langle t, X\rangle\|_p = \|g\|_p \cdot \|\langle t, G\rangle\|_p = |t| \cdot \|g\|_p^2 \sim p|t|$.

[1]In the next section we will see the weaker condition sufficient and – in a sense – necessary for the comparison of moments to hold in the case of independent coordinates.

This means that the moments of $\langle t, X\rangle$ grow C-regularly. On the other hand

$$\big(\mathbb{E}|X|^p\big)^{1/p} = \|g\|_p\big(\mathbb{E}|G|^p\big)^{1/p} \geq \sqrt{n}\|g\|_p \sim \sqrt{np} \qquad \text{for } p \geq 2,$$

and

$$\sigma_X(p) = \sup_{t\in B_2^n} \|\langle t, G\rangle\|_p\|g\|_p \sim p,$$

hence (2.2) cannot hold for $X = gG$ with any constant independent of the dimension n.

In the next section we present new results related to the comparison of weak and strong moments obtained by Latała and the author in [22] and [23]. Further parts of this chapter contain the proofs and some additional observations.

2.1 Main results

2.1.1 Comparison of moments for ℓ_r-norms

Our first main result states that Conjecture 2.1 holds for spaces which may be isometrically embedded in ℓ_r for some $r \geq 1$. This result, as well as its consequences comes from [22].

Theorem 2.3. *Let X be a log-concave vector with values in a normed space $(F, \|\ \|)$ which may be isometrically embedded in ℓ_r for some $r \in [1, \infty)$. Then for $p \geq 1$,*

$$(\mathbb{E}\|X\|^p)^{1/p} \leq Cr\left(\mathbb{E}\|X\| + \sup_{\varphi\in F^*, \|\varphi\|_*\leq 1} (\mathbb{E}|\varphi(X)|^p)^{1/p}\right).$$

This theorem implies the following deviation inequality for $\|X\|$.

Corollary 2.4. *Let X and F be as above. Then*

$$\mathbb{P}(\|X\| \geq 2eCrt\mathbb{E}\|X\|) \leq \exp\left(-\sigma^{-1}_{\|\cdot\|,X}(t\mathbb{E}\|X\|)\right) \qquad \text{for } t \geq 1.$$

We may take C as in Theorem 2.3.

Proof of Corollary 2.4. We will use Remark 1.3. In the case when $t\mathbb{E}\|X\| \geq \sup_{\|u\|_*\leq 1} \operatorname{ess\,sup} |\langle u, X\rangle|$ both sides of the estimate are equal to 0. If $t\mathbb{E}\|X\| < \sup_{\|u\|_*\leq 1} \operatorname{ess\,sup} |\langle u, X\rangle|$, then take $p := \sigma^{-1}_{\|\cdot\|,X}(t\mathbb{E}\|X\|)$. By Chebyshev's inequality and Theorem 2.3 we obtain

$$\mathbb{P}(\|X\| \geq 2eCrt\mathbb{E}\|X\|) \leq \frac{\mathbb{E}\|X\|^p}{(2eCrt\mathbb{E}\|X\|)^p} \leq \left(\frac{1+t}{2t}\right)^p e^{-p} \leq e^{-p}. \qquad \square$$

Since log-concavity is preserved under linear transformations and, by the Hahn-Banach theorem, any linear functional on a subspace of ℓ_r is a restriction of a functional on the whole ℓ_r with the same norm, it is enough to prove Theorem 2.3 for $F = \ell_r$. An easy approximation argument shows that we may consider finite dimensional spaces ℓ_r^n. To simplify the notation for an n-dimensional vector X and $p \geq 1$ we write

$$\sigma_{r,X}(p) := \sup_{\|t\|_{r'} \leq 1} \left(\mathbb{E}\Big|\sum_{i=1}^n t_i X_i\Big|^p \right)^{1/p}.$$

We will prove the following equivalent version of Theorem2.3. A constant C is the same in both theorems.

Theorem 2.5. *Let X be a finite dimensional log-concave vector and $r \in [1, \infty)$. Then*

$$(\mathbb{E}\|X\|_r^p)^{1/p} \leq Cr\left(\mathbb{E}\|X\|_r + \sigma_{r,X}(p)\right) \qquad \textit{for } p \geq 1.$$

To show the above theorem we follow the approach from [20] and establish the following cut version of the above inequality.

Theorem 2.6. *Suppose that $r \in [1, \infty)$ and X is a log-concave n-dimensional random vector. Let*

$$d_i := (\mathbb{E}X_i^2)^{1/2}, \quad d := \left(\sum_{i=1}^n d_i^r\right)^{1/r}. \tag{2.3}$$

Then for $p \geq r$,

$$\mathbb{E}\left(\sum_{i=1}^n |X_i|^r \mathbf{1}_{\{|X_i| \geq t d_i\}}\right)^{p/r} \leq (C_3 r \sigma_{r,X}(p))^p \qquad \textit{for } t \geq C_4 r \log\left(\frac{d}{\sigma_{r,X}(p)}\right). \tag{2.4}$$

Let us show how Theorem 2.6 implies Theorem 2.5.

Proof of Theorem 2.5. Since by (1.2) we have $(\mathbb{E}\|X\|_r^p)^{1/p} \leq C_1 p \mathbb{E}\|X\|_r$, we may assume that $p \geq r$. Let d_i and d be as in Theorem 2.6. Then

$$d = \|(\mathbb{E}X_i^2)_i^{1/2}\|_r \leq 2C_1 \|(\mathbb{E}|X_i|)_i\|_r \leq 2C_1 \mathbb{E}\|X\|_r.$$

In particular, if $d \geq \sup_{\|t\|_* \leq 1} \operatorname{ess\,sup} \left|\sum_{i=1}^n t_i X_i\right|$, then

$$(\mathbb{E}\|X\|_r^p)^{1/p} = \left(\mathbb{E} \sup_{\|t\|_* \leq 1} \Big|\sum_{i=1}^n t_i X_i\Big|^p\right)^{1/p} \leq d \leq 2C_1 \mathbb{E}\|X\|_r.$$

If otherwise $d < \sup_{\|t\|_*\le 1} \operatorname{ess\,sup}\left|\sum_{i=1}^n t_iX_i\right|$, set

$$\tilde{p} := \inf\{q \ge p\colon\ \sigma_{r,X}(q) \ge d\} \ge p.$$

Theorem 2.6 applied with $\tilde{p}$ instead of p and $t = 0$, and Remark 1.3 yield

$$\begin{aligned}(\mathbb{E}\|X\|_r^p)^{1/p} &\le (\mathbb{E}\|X\|_r^{\tilde{p}})^{1/\tilde{p}} \le C_3 r\sigma_{r,X}(\tilde{p}) = C_3 r \max\{d, \sigma_{r,X}(p)\}\\ &\le Cr(\mathbb{E}\|X\|_r + \sigma_{r,X}(p)).\end{aligned}$$ □

Remark 2.7. Any finite dimensional space embeds isometrically in ℓ_∞, so to show Conjecture 2.1 it is enough to establish Theorem 2.3 (with a universal constant in place of Cr) for $r = \infty$. Such an estimate holds for isotropic log-concave vectors (see [21, Corollary 3.8]). However a linear image of an isotropic vector does not have to be isotropic, so to establish the conjecture we need to consider either isotropic vectors and an arbitrary norm or vectors with a general covariance structure and the standard ℓ_∞-norm.

Remark 2.8. An n-dimensional space embeds isometrically in ℓ_∞^N, where $N \sim e^n$. Moreover, in $\mathbb{R}^N$ we have $e^{-1}\|\cdot\|_{\log N} \le \|\cdot\|_\infty \le \|\cdot\|_{\log N}$. Therefore Theorem 2.5 implies (2.2) with $C \sim \log N \sim n$. If Theorem 2.5 held with Cr^γ instead of Cr, then (2.2) would hold with $C \sim n^\gamma$, what is unknown for any $\gamma < \frac{1}{2}$.

2.1.2 Comparison of moments in the independent case

Let us now present results obtained in [23]. We may look at the comparison of moments in a slightly different way than the one presented before. For an n-dimensional random vector X instead of taking the moments of norms of X we may considering the moments of $\sup_{t\in T}\left|\sum_{i=1}^n t_iX_i\right|$ – if T is a unit ball of the dual norm of $\|\cdot\|$, then this quantity coincides with $\|X\|$. This approach will be useful in the proof of our second main result concerning the comparison of weak and strong moments, which generalises the aforementioned result of [24, Theorem 2.3] for vectors with independent regular coordinates.

Theorem 2.9. *Let $X_1, \ldots, X_n$ be independent mean zero random variables with finite moments such that*

$$\|X_i\|_{2p} \le \alpha\|X_i\|_p \qquad \text{for every } p \ge 2 \text{ and } i = 1, \ldots, n, \tag{2.5}$$

where α is a finite positive constant. Then for every $p \ge 1$ and every nonempty set $T \subset \mathbb{R}^n$ we have

$$\left(\mathbb{E}\sup_{t\in T}\left|\sum_{i=1}^n t_iX_i\right|^p\right)^{1/p} \le C(\alpha)\left[\mathbb{E}\sup_{t\in T}\left|\sum_{i=1}^n t_iX_i\right| + \sup_{t\in T}\left(\mathbb{E}\left|\sum_{i=1}^n t_iX_i\right|^p\right)^{1/p}\right], \tag{2.6}$$

where $C(\alpha)$ is a constant which depends only on α.

It turns out that Theorem 2.9 may be reversed in the i.i.d. case (see the theorem below). Therefore one cannot weaken assumption (2.5) in Theorem 2.9.

Theorem 2.10. *Let $X_1, X_2, \ldots$ be i.i.d. random variables. Assume that there exists a constant L such that for every $p \geq 1$, every n and every nonempty set $T \subset \mathbb{R}^n$ we have*

$$\Big(\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p} \leq L\bigg[\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big| + \sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p}\bigg]. \tag{2.7}$$

Then

$$\|X_1\|_{2p} \leq \alpha(L)\|X_1\|_p \qquad \text{for } p \geq 2, \tag{2.8}$$

where $\alpha(L)$ is a constant which depends only on $L \geq 1$.

It will be clear from the proof of Theorem 2.10 that it suffices to assume (2.7) for $T = \{\pm e_j : j \in \{1, \ldots, n\}\}$ only, where $\{e_1, \ldots, e_n\}$ is the canonical basis of $\mathbb{R}^n$.

The comparison of weak and strong moments (2.6) yields also a deviation inequality for $\sup_{t\in T}|\sum_{i=1}^n t_iX_i|$.

Corollary 2.11. *Assume $X_1, X_2, \ldots$ satisfy the assumptions of Theorem 2.9. Then for any $u \geq 0$ and any nonempty set T in $\mathbb{R}^n$,*

$$\mathbb{P}\bigg(\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big| \geq C_1(\alpha)\Big[u + \mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big|\Big]\bigg) \leq C_2(\alpha)\sup_{t\in T}\mathbb{P}\bigg(\Big|\sum_{i=1}^n t_iX_i\Big| \geq u\bigg), \tag{2.9}$$

where constants $C_1(\alpha)$ and $C_2(\alpha)$ depend only on the constant α in (2.5).

Another consequence of Theorem 2.10 is the following Khintchine-Kahane type inequality.

Corollary 2.12. *Assume X_i, $1 \leq i \leq n$ satisfy the assumptions of Theorem 2.9. Then for any $p \geq q \geq 2$ and any nonempty set T in $\mathbb{R}^n$ we have,*

$$\Big(\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p} \leq C(\alpha)\Big(\frac{p}{q}\Big)^{\max\{1/2,\log_2\alpha\}}\Big(\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big|^q\Big)^{1/q}$$

where a constant $C(\alpha)$ depends only on the constant α in (2.5).

The rest of this subsection will be dedicated to present a bunch of remarks related to the above results.

Remark 2.13. Exponent $\max\{1/2, \log_2\alpha\}$ in Corollary 2.12 is optimal.

Indeed, since $\|g\|_p \sim \sqrt{p/e}$ as $p \to \infty$ one cannot go below $1/2$ by the central limit theorem.

To see that $\log_2 \alpha$ term cannot be improved it is enough to consider $\alpha > \sqrt{2}$. Let $r = 1/\log_2 \alpha \in (0,2)$ and let X be a symmetric random variable given by $\mathbb{P}(|X| \geq t) = e^{-t^r}$ (with $2 > r > 0$), i.e. $X = |\mathcal{E}|^{1/r} \operatorname{sgn} \mathcal{E}$, where $\mathcal{E}$ has the symmetric exponential distribution. By Stirling's formula $\Gamma(x+1) = (\frac{x}{e})^x \sqrt{2\pi x} e^{f(x)}$ with $f(x) \in (0, 1/12)$ for $x \geq 1$, so for $p \geq 2$,

$$\frac{\|X\|_{2p}}{\|X\|_p} = \frac{\Gamma\big(\frac{2p}{r}+1\big)^{1/(2p)}}{\Gamma\big(\frac{p}{r}+1\big)^{1/p}} \leq 2^{1/r}\left(\frac{r}{\pi p}\right)^{1/(4p)} e^{1/(24p)} \leq 2^{1/r} = \alpha.$$

Moreover, $\|X\|_p \sim (\frac{p}{er})^{1/r}$ for $p \to \infty$, so the assertion of Corollary 2.12 cannot hold with any exponent better than $\log_2 \alpha$.

Remark 2.14. If the variables X_i are symmetric then the term $\mathbb{E} \sup_{t \in T} |\sum_{i=1}^n t_i X_i|$ in (2.6) may be replaced by $\mathbb{E} \sup_{t \in T} \sum_{i=1}^n t_i X_i$.

Proof. Let s be any point in T. Then $T \subset T - T + s$, so by the triangle inequality

$$\left(\mathbb{E} \sup_{t \in T} \Big|\sum_{i=1}^n t_i X_i\Big|^p\right)^{1/p} \leq \left(\mathbb{E} \sup_{t \in T-T} \Big|\sum_{i=1}^n t_i X_i\Big|^p\right)^{1/p} + \left(\mathbb{E}\Big|\sum_{i=1}^n s_i X_i\Big|^p\right)^{1/p}.$$

Estimate (2.6) applied to the set $T - T$ yields

$$\left(\mathbb{E} \sup_{t \in T-T} \Big|\sum_{i=1}^n t_i X_i\Big|^p\right)^{1/p} \leq C(\alpha)\left[\mathbb{E} \sup_{t \in T-T} \Big|\sum_{i=1}^n t_i X_i\Big| + \sup_{t \in T-T} \left(\mathbb{E}\Big|\sum_{i=1}^n t_i X_i\Big|^p\right)^{1/p}\right].$$

The set $T - T$ is symmetric, so

$$\mathbb{E} \sup_{t \in T-T} \Big|\sum_{i=1}^n t_i X_i\Big| = \mathbb{E} \sup_{t \in T-T} \sum_{i=1}^n t_i X_i \leq 2\mathbb{E} \sup_{t \in T} \sum_{i=1}^n t_i X_i,$$

where the last estimate follows, since $(X_i)_{i=1}^n$ and $(-X_i)_{i=1}^n$ are equally distributed. Moreover,

$$\sup_{t \in T-T} \left(\mathbb{E}\Big|\sum_{i=1}^n t_i X_i\Big|^p\right)^{1/p} \leq 2 \sup_{t \in T} \left(\mathbb{E}\Big|\sum_{i=1}^n t_i X_i\Big|^p\right)^{1/p},$$

what finishes the proof of the remark. □

Remark 2.15. If the variables X_i are not centred then (2.6) holds provided that the assumption (2.5) is replaced by

$$\|X_i - \mathbb{E}X_i\|_{2p} \leq \alpha \|X_i - \mathbb{E}X_i\|_p \qquad \text{for } p \geq 2 \text{ and } i = 1, \ldots, n.$$

Proof. We have

$$\Big(\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p} \le \Big(\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_i(X_i-\mathbb{E}X_i)\Big|^p\Big)^{1/p} + \sup_{t\in T}\Big|\sum_{i=1}^n t_i\mathbb{E}X_i\Big|.$$

Theorem 2.9 applied to centred variables $X_i-\mathbb{E}X_i$, $i=1,\ldots,n$, yields

$$\Big(\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_i(X_i-\mathbb{E}X_i)\Big|^p\Big)^{1/p}$$
$$\le C(\alpha)\Big[\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_i(X_i-\mathbb{E}X_i)\Big| + \sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_i(X_i-\mathbb{E}X_i)\Big|^p\Big)^{1/p}\Big].$$

To conclude it is enough to observe that

$$\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_i(X_i-\mathbb{E}X_i)\Big| \le \mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big| + \sup_{t\in T}\Big|\sum_{i=1}^n t_i\mathbb{E}X_i\Big|,$$

$$\sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_i(X_i-\mathbb{E}X_i)\Big|^p\Big)^{1/p} \le \sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p} + \sup_{t\in T}\Big|\sum_{i=1}^n t_i\mathbb{E}X_i\Big|,$$

and

$$\sup_{t\in T}\Big|\sum_{i=1}^n t_i\mathbb{E}X_i\Big| \le \mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big|.$$ □

2.2 Proof in the case of ℓ_r-norm

By (1.2) for any log-concave vector X and any r,

$$\sigma_{r,X}(\lambda p) \le C_1\lambda\sigma_{r,X}(p) \quad \text{for } \lambda\ge 1,\ p\ge 2.$$

As in Corollary 2.4, the Paouris inequality (2.1) together with Chebyshev's inequality imply

$$\mathbb{P}\left(\|X\|_2 \ge eC_2\left((\mathbb{E}\|X\|_2^2)^{1/2}+\sigma_X(p)\right)\right) \le e^{-p} \quad \text{for } p\ge 1. \tag{2.10}$$

We will always assume, without loss of generality, that d_i defined in Theorem 2.6 are non-zero.

The next proposition generalizes Proposition 4 from [20].

Proposition 2.16. *Let X, r, d_i, and d be as in Theorem 2.6 and $A := \{X \in K\}$, where K is a convex set in $\mathbb{R}^n$ satisfying $0 < \mathbb{P}(A) \leq 1/e$. Then*
(i) for every $t \geq r$,

$$\sum_{i=1}^{n} \mathbb{E}|X_i|^r \mathbf{1}_{A\cap\{X_i \geq td_i\}} \leq {C_5}^r \mathbb{P}(A) \left(r^r \sigma_{r,X}^r(-\log(\mathbb{P}(A))) + (dt)^r e^{-t/C_6} \right). \quad (2.11)$$

(ii) for every $t > 0$, $u \geq 1$,

$$\sum_{k=0}^{\infty} 2^{kr} \sum_{i=1}^{n} d_i^r \mathbf{1}_{\{\mathbb{P}(A\cap\{X_i \geq 2^k td_i\}) \geq e^{-u}\mathbb{P}(A)\}}$$
$$\leq \frac{(C_7 u)^r}{t^r} \left(\sigma_{r,X}^r(-\log(\mathbb{P}(A))) + d^r \mathbf{1}_{\{t \leq uC_8\}} \right). \quad (2.12)$$

Proof. Let Y be a random vector defined by

$$\mathbb{P}(Y \in B) = \frac{\mathbb{P}(A \cap \{X \in B\})}{\mathbb{P}(A)} = \frac{\mathbb{P}(X \in B \cap K)}{\mathbb{P}(X \in K)},$$

i.e. Y is distributed as X conditioned on A. Clearly, for every measurable set B one has $\mathbb{P}(X \in B) \geq \mathbb{P}(A)\mathbb{P}(Y \in B)$. Recall that Y is log-concave.

To simplify the notation set

$$p_A := -\log \mathbb{P}(A) \quad \text{and} \quad c_i := (\mathbb{E}Y_i^2)^{1/2},\ i = 1, \dots, n.$$

Let

$$I = I(v) := \{i \leq n \colon\ \mathbb{E}Y_i^2 \geq v^2 d_i^2\},$$

where v is an absolute constant to be chosen later. Let us also fix a sequence $(a_i)_{i \leq n}$.

Put $S = \sum_{i\in I} |a_i| c_i^{-1} Y_i^2$. Observe that $S = \|((|a_i|/c_i)^{1/2} Y_i)_{i\in I}\|_2^2$, hence by the log-concavity of Y and (1.2), $\mathbb{E}S^2 \leq (2C_1)^4(\mathbb{E}S)^2$, and the Paley-Zygmund inequality yields

$$\mathbb{P}\left(\sum_{i\in I} |a_i| c_i^{-1} Y_i^2 \geq \frac{1}{2} \sum_{i\in I} |a_i| c_i \right) = \mathbb{P}\left(S \geq \frac{1}{2}\mathbb{E}S \right) \geq \frac{1}{4} \frac{(\mathbb{E}S)^2}{\mathbb{E}S^2} \geq \frac{1}{(2\sqrt{2}C_1)^4}. \quad (2.13)$$

We have $\mathbb{E}Y_i^4 \leq (2C_1 c_i)^4$, so by Chebyshev's inequality we get

$$\mathbb{P}\left(\sum_{i\in I} |a_i| c_i^{-3} Y_i^4 \geq (2C_1)^4 s \sum_{i\in I} |a_i| c_i \right) \leq \frac{1}{s} \quad \text{for } s > 0. \quad (2.14)$$

Combining (2.13) and (2.14) we conclude that there exist constants C_9 and $c \in (0,1)$ such that

$$\mathbb{P}\left(\sum_{i\in I}|a_i|c_i^{-1}Y_i^2 \geq \frac{1}{2}\sum_{i\in I}|a_i|c_i,\ \sum_{i\in I}|a_i|c_i^{-3}Y_i^4 \leq C_9\sum_{i\in I}|a_i|c_i\right) \geq c$$

and therefore

$$\mathbb{P}\left(\sum_{i\in I}|a_i|c_i^{-1}X_i^2 \geq \frac{1}{2}\sum_{i\in I}|a_i|c_i,\ \sum_{i\in I}|a_i|c_i^{-3}X_i^4 \leq C_9\sum_{i\in I}|a_i|c_i\right) \geq c\mathbb{P}(A) \geq e^{-C_{10}p_A}.$$

Let $\tilde{X}$ be the vector $(|a_i|^{1/2}c_i^{-1/2}X_i)_{i\in I}$ conditioned on the set

$$B := \left\{\sum_{i\in I}|a_i|c_i^{-3}X_i^4 \leq C_9\sum_{i\in I}|a_i|c_i\right\}.$$

Then

$$\mathbb{P}\left(\|\tilde{X}\|_2^2 \geq \frac{1}{2}\sum_{i\in I}|a_i|c_i\right) \geq \frac{1}{\mathbb{P}(B)}e^{-C_{10}p_A} \geq e^{-C_{10}p_A}. \tag{2.15}$$

The random vector $\tilde{X}$ is log-concave and by the Markov inequality we have $\mathbb{P}(B) \geq 1/2$ if v is a sufficiently large universal constant (since $\mathbb{E}X_i^4 \leq Cd_i^4 \leq Cv^{-4}c_i^4$ for $i \in I$). Thus

$$\mathbb{E}\|\tilde{X}\|_2^2 = \frac{1}{\mathbb{P}(B)}\mathbb{E}\left(\sum_{i\in I}|a_i|c_i^{-1}X_i^2\mathbf{1}_B\right) \leq 2\sum_{i\in I}|a_i|c_i^{-1}d_i^2 \leq 2v^{-2}\sum_{i\in I}|a_i|c_i. \tag{2.16}$$

Now we will estimate $\sigma_{\tilde{X}}(p)$. To this end fix $t \in \mathbb{R}^I$ with $|t| \leq 1$. Let $\alpha, s > 0$ be numbers to be chosen later and

$$J_\alpha := \{i \in I\colon\ |t_i|(|a_i|c_i)^{-1/2} \leq \alpha\}.$$

We have

$$\left\|\sum_{i\in J_\alpha}t_i\tilde{X}_i\right\|_p \leq \mathbb{P}(B)^{-1/p}\left\|\sum_{i\in J_\alpha}t_i(|a_i|c_i)^{-1/2}|a_i|X_i\right\|_p \leq 2\alpha\sigma_{1,aX}(p).$$

Moreover

$$\left\|\sum_{i\notin J_\alpha}t_i\tilde{X}_i\mathbf{1}_{\{|\tilde{X}_i|\leq s(|a_i|c_i)^{1/2}\}}\right\|_p \leq \sum_{i\notin J_\alpha}s|t_i|(|a_i|c_i)^{1/2} = s\sum_{i\notin J_\alpha}\frac{|t_i|^2}{|t_i|(|a_i|c_i)^{-1/2}}$$
$$\leq \frac{s}{\alpha}\sum_{i\in I}t_i^2 \leq \frac{s}{\alpha}.$$

Observe that by the definition of the set B and the vector $\tilde{X}$ we have

$$\sum_{i\in I}(|a_i|c_i)^{-1}\tilde{X}_i^4 \le C_9\sum_{i\in I}|a_i|c_i.$$

Thus

$$\begin{aligned}\Big\|\sum_{i\notin J_\alpha} t_i\tilde{X}_i\mathbf{1}_{\{|\tilde{X}_i|>s(|a_i|c_i)^{1/2}\}}\Big\|_p &\le \Big\|\Big(\sum_{i\in I}\tilde{X}_i^2\mathbf{1}_{\{|\tilde{X}_i|>s(|a_i|c_i)^{1/2}\}}\Big)^{1/2}\Big\|_p \\ &\le \Big\|\frac{1}{s}\Big(\sum_{i\in I}(|a_i|c_i)^{-1}\tilde{X}_i^4\Big)^{1/2}\Big\|_p \le \frac{1}{s}\Big(C_9\sum_{i\in I}|a_i|c_i\Big)^{1/2}.\end{aligned}$$

Combining the above estimates we obtain

$$\Big\|\sum_{i\in I}t_i\tilde{X}_i\Big\|_p \le 2\alpha\sigma_{1,aX}(p)+\frac{s}{\alpha}+\frac{1}{s}\Big(C_{10}\sum_{i\in I}|a_i|c_i\Big)^{1/2}.$$

Taking the supremum over t and optimizing over $\alpha>0$ we get

$$\sigma_{\tilde{X}}(p)\le 4(s\sigma_{1,aX}(p))^{1/2}+\frac{1}{s}\Big(C_9\sum_{i\in I}|a_i|c_i\Big)^{1/2} \qquad \text{for } s>0. \tag{2.17}$$

Paouris' inequality (2.10) (applied to $\tilde{X}$ instead of X) together with (2.16) and (2.17) implies that

$$\begin{aligned}\mathbb{P}\Bigg(\|\tilde{X}\|_2 \ge eC_2\Big[\Big(2v^{-2}\sum_{i\in I}|a_i|c_i\Big)^{1/2}+4\big(s\sigma_{1,aX}(C_{10}p_A)\big)^{1/2}+\frac{1}{s}\Big(C_9\sum_{i\in I}|a_i|c_i\Big)^{1/2}\Big]\Bigg) \\ < e^{-C_{10}p_A}.\end{aligned}$$

Comparing the above with (2.15) we get

$$\begin{aligned}eC_2\Big[\Big(2v^{-2}\sum_{i\in I}|a_i|c_i\Big)^{1/2}+4\big(s\sigma_{1,aX}(C_{10}p_A)\big)^{1/2}+\frac{1}{s}\Big(C_9\sum_{i\in I}|a_i|c_i\Big)^{1/2}\Big] \\ \ge \Big(\frac{1}{2}\sum_{i\in I}|a_i|c_i\Big)^{1/2}.\end{aligned}$$

If we choose s and v to be sufficiently large absolute constants we will get

$$\sum_{i\in I}|a_i|(\mathbb{E}Y_i^2)^{1/2}=\sum_{i\in I}|a_i|c_i\le C\sigma_{1,aX}(C_{10}p_A)\le C\sigma_{1,aX}(p_A).$$

Put $a_i := (\mathbb{E}|Y_i|^2)^{(r-1)/2}\mathbf{1}_{i\in I}$. If $\|t\|_\infty \le 1$, then $(\sum |t_i a_i|^{r'})^{1/r'} \le \|a\|_{r'}$. Thus the previous inequality implies

$$\sum_{i\in I}\left(\mathbb{E}|Y_i|^2\right)^{r/2} \le C\sigma_{1,aX}(p_A) \le C\|a\|_{r'}\sigma_{r,X}(p_A) = C\left(\sum_{i\in I}\left(\mathbb{E}|Y_i|^2\right)^{r/2}\right)^{1/r'}\sigma_{r,X}(p_A).$$

This gives

$$\sum_{i\in I}(\mathbb{E}|Y_i|^2)^{r/2} \le C^r\sigma_{r,X}^r(p_A).$$

Since $\|Y_i\|_r \le \max\{1, C_1 r/2\}\|Y_i\|_2$ we also get

$$\sum_{i\in I}\mathbb{E}|Y_i|^r \le (Cr)^r\sigma_{r,X}^r(p_A).$$

To prove (2.11) note that if $i \notin I$, then by (1.1) we have $\mathbb{P}(|Y_i| \ge sd_i) \le 2e^{-s/C}$ for $s \ge 0$, hence we get by integrating by parts that for $t \ge r$,

$$\mathbb{E}|Y_i|^r\mathbf{1}_{\{Y_i\ge td_i\}} \le (Ctd_i)^r e^{-t/C}$$

and therefore

$$\sum_{i\notin I}\mathbb{E}|Y_i|^r\mathbf{1}_{\{Y_i\ge td_i\}} \le (Ctd)^r e^{-t/C}.$$

Hence

$$\frac{1}{\mathbb{P}(A)}\sum_{i=1}^n\mathbb{E}|X_i|^r\mathbf{1}_{A\cap\{X_i\ge td_i\}} = \sum_{i=1}^n\mathbb{E}|Y_i|^r\mathbf{1}_{\{Y_i\ge td_i\}} \le C^r\left(r^r\sigma_{r,X}^r(-\log(\mathbb{P}(A))) + (dt)^r e^{-t/C}\right).$$

To show (2.12) note first that for every i the random variable Y_i is log-concave, hence for $s \ge 0$ inequality (1.1) implies

$$\frac{\mathbb{P}(A\cap\{X_i\ge s\})}{\mathbb{P}(A)} = \mathbb{P}(Y_i \ge s) \le \exp\left(2 - \frac{s}{2e\|Y_i\|_2}\right).$$

Thus, if $\mathbb{P}(A\cap\{X_i \ge 2^k td_i\}) \ge e^{-u}\mathbb{P}(A)$ and $u \ge 1$, then $\|Y_i\|_2 \ge 2^k td_i/(2e(u+2)) \ge 2^k td_i/(6eu)$. In particular this cannot happen if $i \notin I$, $k \ge 0$ and $u \le t/C_8$ with C_8

large enough. Therefore

$$\begin{aligned}
&\sum_{k=0}^{\infty} 2^{kr} \sum_{i=1}^{n} d_i^r \mathbf{1}_{\{\mathbb{P}(A\cap\{X_i\geq 2^k td_i\})\geq e^{-u}\mathbb{P}(A)\}} \\
&\qquad\qquad \leq \left(\sum_{i\in I} + \mathbf{1}_{\{t\leq uC_8\}} \sum_{i\notin I}\right) d_i^r \sum_{k=0}^{\infty} 2^{kr} \mathbf{1}_{\{(\mathbb{E}Y_i^2)^{1/2}\geq 2^k td_i/(6eu)\}} \\
&\qquad\qquad \leq \left(\sum_{i\in I} + \mathbf{1}_{\{t\leq uC_8\}} \sum_{i\notin I}\right) d_i^r \frac{(Cu)^r}{(td_i)^r} (\mathbb{E}Y_i^2)^{r/2} \\
&\qquad\qquad \leq \frac{(Cu)^r}{t^r} \left(\sum_{i\in I} (\mathbb{E}Y_i^2)^{r/2} + \mathbf{1}_{\{t\leq uC_8\}} \sum_{i\notin I} d_i^r\right) \\
&\qquad\qquad \leq \frac{(Cu)^r}{t^r} \left(\sigma_{r,X}^r(-\log(\mathbb{P}(A))) + d^r \mathbf{1}_{\{t\leq uC_8\}}\right). \qquad \square
\end{aligned}$$

We will also use the following combinatorial lemma (Lemma 11 in [18]).

Lemma 2.17. *Let $l_0 \geq l_1 \geq \ldots \geq l_s$ be a fixed sequence of positive integers and*

$$\mathcal{F} := \{f\colon \{1,2,\ldots,l_0\} \to \{0,1,2,\ldots,s\}\colon\ \forall_{1\leq i\leq s}\ |\{r\colon f(r)\geq i\}| \leq l_i\}.$$

Then

$$|\mathcal{F}| \leq \prod_{i=1}^{s} \left(\frac{el_{i-1}}{l_i}\right)^{l_i}.$$

Proof of Theorem 2.6. Observe that we may assume that $t \geq C_4 r$. Indeed, if $e\sigma_{r,X}(p) \leq d$ then by our assumption $t \geq C_4 r$. If $e\sigma_{r,X}(p) > d$ then

$$\begin{aligned}
&\left(\mathbb{E}\left(\sum_{i=1}^{n} |X_i|^r \mathbf{1}_{\{|X_i|\geq td_i\}}\right)^{p/r}\right)^{1/p} \\
&\qquad \leq C_4 r \left(\sum_{i=1}^{n} d_i^r\right)^{1/r} + \left(\mathbb{E}\left(\sum_{i=1}^{n} |X_i|^r \mathbf{1}_{\{|X_i|\geq \max\{t,C_4 r\}d_i\}}\right)^{p/r}\right)^{1/p} \\
&\qquad \leq eC_4 r\sigma_{r,X}(p) + \left(\mathbb{E}\left(\sum_{i=1}^{n} |X_i|^r \mathbf{1}_{\{|X_i|\geq \max\{t,C_4 r\}d_i\}}\right)^{p/r}\right)^{1/p}.
\end{aligned}$$

Moreover, the vector $-X$ is also log-concave, has the same values of d_i and $\sigma_{r,-X} = \sigma_{r,X}$. Hence it is enough to show that

$$\mathbb{E}\left(\sum_{i=1}^{n} X_i^r \mathbf{1}_{\{X_i\geq td_i\}}\right)^{p/r} \leq (Cr\sigma_{r,X}(p))^p \quad \text{for } t \geq C_4 r \max\left\{1, \log\left(\frac{d}{\sigma_{r,X}(p)}\right)\right\}.$$

Observe that for $l = 1, 2, \ldots,$

$$\mathbb{E}\left(\sum_{i=1}^{n} X_i^r \mathbf{1}_{\{X_i \geq td_i\}}\right)^l \leq \mathbb{E}\left(\sum_{i=1}^{n}\sum_{k=0}^{\infty} 2^{(k+1)r}(td_i)^r \mathbf{1}_{\{X_i \geq 2^k td_i\}}\right)^l$$
$$= (2t)^{rl} \sum_{i_1,\ldots,i_l=1}^{n} \sum_{k_1,\ldots,k_l=0}^{\infty} 2^{(k_1+\ldots+k_l)r} d_{i_1}^r \ldots d_{i_l}^r \mathbb{P}(B_{i_1,k_1\ldots,i_l,k_l}),$$

where

$$B_{i_1,k_1\ldots,i_l,k_l} := \{X_{i_1} \geq 2^{k_1} td_{i_1}, \ldots, X_{i_l} \geq 2^{k_l} td_{i_l}\}.$$

Define a positive integer l by

$$\frac{p}{r} < l \leq 2\frac{p}{r} \quad \text{and} \quad l = 2^M \text{ for some positive integer } M.$$

Then, by (1.2) we get $\sigma_{r,X}(p) \leq \sigma_{r,X}(rl) \leq \sigma_{r,X}(2p) \leq 2C_1\sigma_{r,X}(p)$. Since for any nonnegative random variable Z we have $(\mathbb{E}Z^{p/r})^{r/p} \leq (\mathbb{E}Z^l)^{1/l}$, it is enough to show that

$$m(l) \leq \left(\frac{Cr\sigma_{r,X}(rl)}{t}\right)^{rl} \quad \text{for } t \geq C_4 r \max\left\{1, \log\left(\frac{d}{\sigma_{r,X}(rl)}\right)\right\}, \tag{2.18}$$

where

$$m(l) := \sum_{k_1,\ldots,k_l=0}^{\infty} \sum_{i_1,\ldots,i_l=1}^{n} 2^{(k_1+\ldots+k_l)r} d_{i_1}^r \ldots d_{i_l}^r \mathbb{P}(B_{i_1,k_1,\ldots,i_l,k_l}).$$

We divide the sum in $m(l)$ into several parts. Define sets

$$I_0 := \left\{(i_1, k_1, \ldots, i_l, k_l) \colon\ \mathbb{P}(B_{i_1,k_1,\ldots,i_l,k_l}) > e^{-rl}\right\},$$

and for $j = 1, 2, \ldots,$

$$I_j := \left\{(i_1, k_1, \ldots, i_l, k_l) \colon\ \mathbb{P}(B_{i_1,k_1,\ldots,i_l,k_l}) \in (e^{-rl2^j}, e^{-rl2^{j-1}}]\right\}.$$

Then $m(l) = \sum_{j\geq 0} m_j(l)$, where

$$m_j(l) := \sum_{(i_1,k_1,\ldots,i_l,k_l)\in I_j} 2^{(k_1+\ldots+k_l)r} d_{i_1}^r \ldots d_{i_l}^r \mathbb{P}(B_{i_1,k_1\ldots,i_l,k_l}).$$

To estimate $m_0(l)$ define for $1 \leq s \leq l$,

$$P_s I_0 := \{(i_1, k_1, \ldots, i_s, k_s) \colon (i_1, k_1, \ldots, i_l, k_l) \in I_0 \text{ for some } i_{s+1}, \ldots, k_l\}.$$

We have by (1.1) (since t is assumed to be large)

$$\mathbb{P}(B_{i_1,k_1,\ldots,i_s,k_s}) \le \mathbb{P}(B_{i_1,k_1}) \le \exp(2 - 2^{k_1-1}t/e) \le e^{-1}.$$

Thus for $s = 1, \ldots, l-1$,

$$\sum_{(i_1,k_1,\ldots,i_{s+1},k_{s+1})\in P_{s+1}I_0} 2^{(k_1+\ldots+k_{s+1})r} d_{i_1}^r \ldots d_{i_{s+1}}^r \mathbb{P}(B_{i_1,k_1,\ldots,i_{s+1},k_{s+1}})$$
$$\le \sum_{(i_1,k_1,\ldots,i_s,k_s)\in P_sI_0} 2^{(k_1+\ldots+k_s)r} d_{i_1}^r \cdots d_{i_s}^r F(i_1,k_1,\ldots,i_s,k_s),$$

where

$$\begin{aligned}
F(i_1,k_1,\ldots,i_s,k_s) &:= \sum_{i=1}^n \sum_{k=0}^\infty 2^{kr} d_i^r \mathbb{P}(B_{i_1,k_1,\ldots,i_s,k_s} \cap \{X_i \ge 2^k t d_i\}) \\
&= \sum_{i=1}^n \sum_{k=0}^\infty 2^{kr} d_i^r \sum_{j=k}^\infty \mathbb{P}(B_{i_1,k_1,\ldots,i_s,k_s} \cap \{2^j t d_i > X_i \ge 2^j t d_i\}) \\
&\le \sum_{i=1}^n \sum_{j=0}^\infty d_i^r \mathbb{P}(B_{i_1,k_1,\ldots,i_s,k_s} \cap \{2^j t d_i > X_i \ge 2^j t d_i\}) 2^{jr+1} \\
&\le \sum_{i=1}^n \mathbb{E} 2t^{-r} |X_i|^r \mathbf{1}_{B_{i_1,k_1,\ldots,i_s,k_s} \cap \{X_i \ge t d_i\}} \\
&\le 2t^{-r} C_5^r \mathbb{P}(B_{i_1,k_1,\ldots,i_s,k_s}) \left(r^r \sigma_{r,X}^r(-\log \mathbb{P}(B_{i_1,k_1,\ldots,i_s,k_s})) + (dt)^r e^{-t/C_6} \right),
\end{aligned}$$

where the last inequality follows by (2.11). Note that for $(i_1,k_1,\ldots,i_s,k_s) \in P_sI_0$ we have $\mathbb{P}(B_{i_1,k_1,\ldots,i_s,k_s}) > e^{-rl}$. Moreover, by our assumptions on t (if C_4 is sufficiently large with respect to C_6),

$$(dt)^r e^{-t/C_6} = t^r e^{-t/(2C_6)} d^r e^{-t/(2C_6)} \le r^r \sigma_{r,X}^r(rl).$$

Therefore

$$\sum_{(i_1,k_1,\ldots,i_{s+1},k_{s+1})\in P_{s+1}I_0} 2^{(k_1+\ldots+k_{s+1})r} d_{i_1}^r \ldots d_{i_{s+1}}^r \mathbb{P}(B_{i_1,k_1,\ldots,i_{s+1},k_{s+1}})$$
$$\le 4t^{-r}(C_5 r \sigma_{r,X}(rl))^r \sum_{(i_1,k_1,\ldots,i_s,k_s)\in P_sI_0} 2^{(k_1+\ldots+k_s)r} d_{i_1}^r \ldots d_{i_s}^r \mathbb{P}(B_{i_1,k_1,\ldots,i_s,k_s}).$$

By induction we get

$$\begin{aligned}
m_0(l) &= \sum_{(i_1,k_1,\ldots,i_l,k_l)\in I_0} 2^{(k_1+\ldots+k_l)r} d_{i_1}^r \cdots d_{i_l}^r \mathbb{P}(B_{i_1,k_1,\ldots,i_l,k_l}) \\
&\le \left(\frac{4C_5 r \sigma_{r,X}(rl)}{t} \right)^{r(l-1)} \sum_{(i_1,k_1)\in P_1I_0} 2^{k_1 r} d_{i_1}^r \mathbb{P}(B_{i_1,k_1}).
\end{aligned}$$

We have

$$\sum_{(i_1,k_1)\in P_1 I_0} 2^{k_1 r} d_{i_1}^r \mathbb{P}(B_{i_1,k_1}) \le \sum_{i_1=1}^{n} d_{i_1}^r \sum_{k_1=0}^{\infty} 2^{k_1 r} e^{2-2^{k_1-1}t/e}$$
$$\le \sum_{i_1=1}^{n} d_{i_1}^r 2e^{2-t/(2e)} \le \left(\frac{Cr\sigma_{r,X}(rl)}{t}\right)^r,$$

where the last two inequalities follow from the assumption that for large enough C_4, $t \ge C_4 r \max\left\{1, \log\left(\frac{d}{\sigma_{r,X}(rl)}\right)\right\}$. Thus

$$m_0(l) \le \left(\frac{Cr\sigma_{r,X}(rl)}{t}\right)^{rl}.$$

Now we estimate $m_j(l)$ for $j > 0$. Fix $j > 0$ and define a positive integer ρ_1 by

$$r2^{\rho_1 - 1} < \frac{t}{C_8} \le r2^{\rho_1},$$

where the constant C_8 comes from part (ii) of Proposition 2.16. Now for all $(i_1, k_1, \dots, i_l, k_l) \in I_j$ define a function $f_{i_1,k_1,\dots,i_l,k_l} \colon \{1, \dots, \ell\} \to \{0, 1, \dots\}$ by

$$f_{i_1,k_1,\dots,i_l,k_l}(s) := \begin{cases} 0 & \text{if } \frac{\mathbb{P}(B_{i_1,k_1,\dots,i_s,k_s})}{\mathbb{P}(B_{i_1,k_1,\dots,i_{s-1},k_{s-1}})} > e^{-r}, \\ \rho & \text{if } e^{-r2^{\rho}} < \frac{\mathbb{P}(B_{i_1,k_1,\dots,i_s,k_s})}{\mathbb{P}(B_{i_1,k_1,\dots,i_{s-1},k_{s-1}})} \le e^{-r2^{\rho-1}}, \ \rho \ge 1. \end{cases}$$

Note that for every $(i_1, k_1, \dots, i_l, k_l) \in I_j$ one has

$$1 = \mathbb{P}(B_\emptyset) \ge \mathbb{P}(B_{i_1,k_1}) \ge \mathbb{P}(B_{i_1,k_1,i_2,k_2}) \ge \dots \ge \mathbb{P}(B_{i_1,k_1,\dots,i_l,k_l}) > \exp(-rl2^j).$$

Denote

$$\mathcal{F}_j := \{f_{i_1,k_1,\dots,i_l,k_l} \colon \ (i_1, k_1, \dots, i_l, k_l) \in I_j\}.$$

Then for $f = f_{i_1,k_1,\dots,i_l,k_l} \in \mathcal{F}_j$ and $\rho \ge 1$ one has

$$\exp(-r2^j l) < \mathbb{P}(B_{i_1,k_1,\dots,i_l,k_l}) = \prod_{s=1}^{\ell} \frac{\mathbb{P}(B_{i_1,k_1,\dots,i_s,k_s})}{\mathbb{P}(B_{i_1,k_1,\dots,i_{s-1},k_{s-1}})}$$
$$\le \exp(-r2^{\rho-1}|\{s\colon \ f(s) \ge \rho\}|).$$

Hence for every $\rho \ge 1$ one has

$$|\{s\colon \ f(s) \ge \rho\}| \le \min\{2^{j+1-\rho} l, l\} =: l_\rho. \tag{2.19}$$

In particular f takes values in $\{0,1,\ldots,j+1+\lfloor\log_2 l\rfloor\}$. Clearly, $\sum_{\rho\geq 1} l_\rho=(j+2)l$ and $l_{\rho-1}/l_\rho\leq 2$, so by Lemma 2.17

$$|\mathcal{F}_j|\leq \prod_{\rho=1}^{j+1+\lfloor\log_2 l\rfloor}\left(\frac{el_{\rho-1}}{l_\rho}\right)^{l_\rho}\leq e^{2(j+2)l}.$$

Now fix $f\in\mathcal{F}_j$ and define

$$I_j(f):=\{(i_1,k_1,\ldots,i_l,k_l)\colon\ f_{i_1,k_1,\ldots,i_l,k_l}=f\}$$

and for $s\leq l$,

$$I_{j,s}(f):=P_sI_j(f)=\{(i_1,k_1,\ldots,i_s,k_s)\colon\ f_{i_1,k_1,\ldots,i_l,k_l}=f \text{ for some } i_{s+1},k_{s+1}\ldots,i_l,k_l\}.$$

Recall that for $s\geq 1$, $\mathbb{P}(B_{i_1,k_1,\ldots,i_s,k_s})\leq e^{-1}$. Moreover for $s\leq l$ and any $(i_1,k_1,\ldots,i_l,k_l)\in I_j$, we get by (1.2) that

$$\sigma_X(-\log\mathbb{P}(B_{i_1,k_1,\ldots,i_s,k_s}))\leq\sigma_X(-\log\mathbb{P}(B_{i_1,k_1,\ldots,i_l,k_l}))\leq\sigma_X(rl2^j)\leq C_12^j\sigma_X(rl).$$

Hence estimate (2.12) applied with $u=r2^{f(s+1)}$ implies for $1\leq s\leq l-1$,

$$\begin{aligned}\sum_{(i_1,k_1,\ldots,i_{s+1},k_{s+1})\in I_{j,s+1}(f)} & 2^{(k_1+\ldots+k_{s+1})r}d_{i_1}^r\ldots d_{i_{s+1}}^r\mathbb{P}(B_{i_1,k_1,\ldots,i_{s+1},k_{s+1}})\\ &\leq g(f(s+1))\sum_{(i_1,k_1,\ldots,i_s,k_s)\in I_{j,s}(f)}2^{(k_1+\ldots+k_s)r}d_{i_1}^r\ldots d_{i_s}^r\mathbb{P}(B_{i_1,k_1,\ldots,i_s,k_s}),\end{aligned}$$

where

$$g(\rho):=\begin{cases}(C_1C_7r)^rt^{-r}2^{jr}\sigma_{r,X}^r(rl) & \text{for } \rho=0,\\ (C_1C_7r)^rt^{-r}2^{r(\rho+j)}\sigma_{r,X}^r(rl)\exp(-r2^{\rho-1}) & \text{for } 1\leq\rho<\rho_1,\\ (C_1C_7r)^rt^{-r}2^{r\rho}\big(2^{rj}\sigma_{r,X}^r(rl)+d^r\big)\exp(-r2^{\rho-1}) & \text{for } \rho\geq\rho_1.\end{cases}$$

Suppose that $(i_1,k_1)\in I_1(f)$ and $f(1)=\rho$. Then by (1.1) we have

$$\exp(-r2^\rho)\leq\mathbb{P}(X_{i_1}\geq 2^{k_1}td_{i_1})\leq\exp(2-2^{k_1-1}t/e),$$

hence $2^{k_1}t\leq er2^{\rho+3}$. We may assume without loss of generality that $C_9>8e$. Then $\rho\geq\rho_1$. Moreover, $2^{rk_1}\leq(8er)^r2^{r\rho}t^{-r}$, hence

$$\sum_{(i_1,k_1)\in I_{j,1}(f)}2^{rk_1}d_{i_1}^r\mathbb{P}(B_{i_1,k_1})\leq d^r(16er)^rt^{-r}2^{r\rho}\exp(-r2^{\rho-1})\leq g(\rho)=g(f(1)),$$

since without loss of generality $C_1C_7 \geq 16e$. Thus the induction shows that

$$\begin{aligned} m_j(f) &:= \sum_{(i_1,k_1,\ldots,i_l,k_l)\in I_j(f)} 2^{(k_1+\ldots+k_l)r} d_{i_1}^r \ldots d_{i_l}^r \mathbb{P}(B_{i_1,k_1,\ldots,i_l,k_l}) \\ &\leq \prod_{s=1}^{l} g(f(s)) = \prod_{\rho=0}^{\infty} g(\rho)^{n_\rho}, \end{aligned}$$

where $n_\rho := |f^{-1}(\rho)|$.

Observe that

$$e^{-rl2^{j-1}} \geq \mathbb{P}(B_{i_1,k_1,\ldots,i_l,k_l}) = \prod_{s=1}^{l} \frac{\mathbb{P}(B_{i_1,k_1,\ldots,i_s,k_s})}{\mathbb{P}(B_{i_1,k_1,\ldots,i_{s-1},k_{s-1}})} \geq e^{-rl} \prod_{s\colon f(s)\geq 1} e^{-r2^{f(s)}}.$$

Therefore

$$r\sum_{\rho=1}^{\infty} n_\rho 2^{\rho-1} = \frac{r}{2} \sum_{s\colon f(s)\geq 1} 2^{f(s)} \geq \frac{r}{2} l(2^{j-1}-1).$$

Moreover

$$\sum_{\rho\geq 1} \rho n_\rho \leq (j+1)l + \sum_{\rho\geq j+2} \rho l_\rho = (2j+5)l.$$

Thus

$$\prod_{\rho=0}^{\infty} g(\rho)^{n_\rho} \leq \left(\frac{C_1C_7 r 2^j \sigma_{r,X}(rl)}{t}\right)^{rl} 2^{rl(2j+5)} \left(1+\frac{d^r}{\sigma_{r,X}(rl)^r}\right)^{m} \exp\left(-\frac{rl}{2}(2^{j-1}-1)\right),$$

where $m = \sum_{\rho\geq\rho_1} n_\rho \leq l_{\rho_1} \leq 2^{j+1-\rho_1} l$. By the assumption on t we have $1 + d^r/\sigma_{r,X}^r(rl) \leq 2\exp(t/C_4) \leq \exp(r2^{\rho_1-4})$ if C_4 is large enough (with respect to C_8). Hence

$$m_j(l) \leq |\mathcal{F}_j| \left(\frac{\sqrt{e}C_1C_7 2^{(3j+5)} r\sigma_{r,X}(rl)}{t}\right)^{rl} \exp(-rl2^{j-3}).$$

We get

$$m(l) = \sum_{j=0}^{\infty} m_j(l) \leq \left(\frac{Cr\sigma_{r,X}(rl)}{t}\right)^{rl} + \sum_{j=1}^{\infty} \left(\frac{C2^{3j} r\sigma_{r,X}(rl)}{t}\right)^{rl} \exp(-rl2^{j-3}).$$

To finish the proof of (2.18), note that

$$\sum_{j=1}^{\infty} \left(2^{3j}\right)^{rl} \exp(-rl2^{j-3}) \leq C^{rl} \sum_{j=1}^{\infty} \exp(-rl2^{j-4}) \leq C^{rl}. \qquad \square$$

2.3 Proofs in the case of independent coordinates

In Subsection 2.3.1 we prove Theorem 2.9 for unconditional sets T only. Using this result we generalize it to the case of an arbitrary T in Subection 2.3.2. In Subsection 2.3.3 we prove Corollaries 2.11 and 2.12. Finally, in Subection 2.3.4 we present the proof of Theorem 2.10.

Throughout this section we will frequently work with a Bernoulli sequence ε_i of i.i.d. symmetric random variables taking values ± 1. We assume that variables ε_i are independent of other random variables.

2.3.1 The case of unconditional sets

In this subsection we show that Theorem 2.9 holds under additional assumptions that the set T is unconditional and the variables X_i are symmetric. Recall that a set T in $\mathbb{R}^n$ is called *unconditional* if it is symmetric with respect to the coordinate axes, i.e. $(\eta_i t_i)_{i=1}^n \in T$ for any $t = (t_i)_{i=1}^n \in T$ and any choice of signs $\eta_1, \ldots, \eta_n \in \{-1, 1\}$.

Proposition 2.18. *Let $r \in (0,1)$ and $L \geq 1$. Assume that variables $Y_1, \ldots, Y_n$ are independent and symmetric and*

$$\left(\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iY_i\Big|^p\right)^{1/p} \leq L\left[\mathbb{E}\sup_{t\in T}\sum_{i=1}^n t_iY_i + \sup_{t\in T}\left(\mathbb{E}\Big|\sum_{i=1}^n t_iY_i\Big|^p\right)^{1/p}\right] \qquad (2.20)$$

for all $p \geq 1$ and all nonempty unconditional sets T. Then variables $X_i := |Y_i|^{1/r}\operatorname{sgn} Y_i$ satisfy

$$\left(\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big|^p\right)^{1/p} \leq (2L)^{1/r}\left[\mathbb{E}\sup_{t\in T}\sum_{i=1}^n t_iX_i + \sup_{t\in T}\left(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\right)^{1/p}\right] \qquad (2.21)$$

for all $p \geq 1$ and all nonempty unconditional sets $T \subset \mathbb{R}^n$.

Proof. Definition of X_i and unconditionality of T yield

$$\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big| = \sup_{t\in T}\Big|\sum_{i=1}^n t_i|Y_i|^{1/r}\operatorname{sgn} Y_i\Big| = \sup_{t\in T}\sum_{i=1}^n |t_i||Y_i|^{1/r}.$$

Let $s = (1-r)^{-1}$ and let B_s^n denote the unit ball of ℓ_s^n. Then $1/s + r = 1$ and by Hölder's duality we have

$$\sup_{t\in T}\Big|\sum_{i=1}^n |t_i||Y_i|^{1/r}\Big|^r = \sup_{t\in T}\sup_{u\in B_s^n}\sum_{i=1}^n u_i|t_i|^rY_i = \sup_{t\in T_r}\sum_{i=1}^n t_iY_i,$$

where

$$T_r := \{(u_i|t_i|^r)_{i=1}^n :\ t \in T, u \in B_s^n\}$$

is unconditional in $\mathbb{R}^n$. Therefore (2.20) applied with p/r and T_r instead of p and T yields

$$\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big|^p \le L^{p/r}\left[\mathbb{E}\sup_{t\in T_r} t_iY_i + \sup_{t\in T_r}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iY_i\Big|^{p/r}\Big)^{r/p}\right]^{p/r}.$$

We have

$$\mathbb{E}\sup_{t\in T_r}\sum_{i=1}^n t_iY_i = \mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big|^r \le \Big(\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big|\Big)^r.$$

Moreover,

$$\begin{aligned}\sup_{t\in T_r}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iY_i\Big|^{p/r}\Big)^{r/p} &\le \sup_{t\in T}\Big(\mathbb{E}\sup_{u\in B_s^n}\Big|\sum_{i=1}^n u_i|t_i|^rY_i\Big|^{p/r}\Big)^{r/p}\\ &= \sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n |t_i||X_i|\Big|^p\Big)^{r/p} = \sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_i|X_i|\Big|^p\Big)^{r/p}.\end{aligned}$$

Estimates above together with the inequality $(a+b)^{1/r} \le 2^{1/r-1}(a^{1/r}+b^{1/r})$ yield

$$\Big(\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p} \le \frac{1}{2}(2L)^{1/r}\left[\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big| + \sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_i|X_i|\Big|^p\Big)^{1/p}\right].$$

Hence, in order to prove (2.21) it suffices to show that

$$\sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_i|X_i|\Big|^p\Big)^{1/p} \le \mathbb{E}\sup_{t\in T}\sum_{i=1}^n t_iX_i + 2\sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p}. \tag{2.22}$$

Let $(X_1',\ldots,X_n')$ be an independent copy of $(X_1,\ldots,X_n)$. By the triangle inequality for the p-th integral norm and Jensen's inequality we get

$$\begin{aligned}\sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_i|X_i|\Big|^p\Big)^{1/p} &\le \sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_i\big(|X_i|-\mathbb{E}|X_i'|\big)\Big|^p\Big)^{1/p} + \sup_{t\in T}\Big|\mathbb{E}\sum_{i=1}^n t_i|X_i|\Big|\\ &\le \sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_i\big(|X_i|-|X_i'|\big)\Big|^p\Big)^{1/p} + \mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_i|X_i|\Big|\\ &= \sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_i\big(|X_i|-|X_i'|\big)\Big|^p\Big)^{1/p} + \mathbb{E}\sup_{t\in T}\sum_{i=1}^n t_iX_i,\end{aligned} \tag{2.23}$$

where the equation follows by the unconditionality of T.

Since the sequence $(|X_i| - |X_i'|)_{i=1}^n$ has the same distribution as $(\varepsilon_i(|X_i| - |X_i'|))_{i=1}^n$, for every $t \in \mathbb{R}^n$ we have

$$\begin{aligned}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_i(|X_i| - |X_i'|)\Big|^p\Big)^{1/p} &= \Big(\mathbb{E}\Big|\sum_{i=1}^n t_i\varepsilon_i(|X_i| - |X_i'|)\Big|^p\Big)^{1/p} \\ &\le \Big(\mathbb{E}\Big|\sum_{i=1}^n t_i\varepsilon_i|X_i|\Big|^p\Big)^{1/p} + \Big(\mathbb{E}\Big|\sum_{i=1}^n t_i\varepsilon_i|X_i'|\Big|^p\Big)^{1/p} \\ &= 2\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p}. \qquad (2.24)\end{aligned}$$

Putting (2.23) and (2.24) together we get (2.22), what completes the proof of (2.21). □

Corollary 2.19. *Let $X_1, \ldots, X_n$ be independent symmetric random variables with finite moments such that*

$$\|X_i\|_{2p} \le \alpha\|X_i\|_p \qquad \textit{for } p \ge 2 \textit{ and } i = 1, \ldots, n, \qquad (2.25)$$

where α is a finite positive constant. Then for every $p \ge 1$ and every nonempty unconditional set $T \subset \mathbb{R}^n$ we have

$$\Big(\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p} \le C(\alpha)\Big[\mathbb{E}\sup_{t\in T}\sum_{i=1}^n t_iX_i + \sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p}\Big], \quad (2.26)$$

where $C(\alpha)$ is a constant, which depends only on α.

Proof. Let us first note, that the assumption (2.25) applied k times yields that

$$\|X_i\|_{2^kp} \le \alpha^k\|X_i\|_p \qquad \text{for } p \ge 2.$$

Therefore

$$\|X_i\|_q \le \alpha^{\lceil \log_2(\frac{q}{p})\rceil}\|X_i\|_p \le \alpha\Big(\frac{q}{p}\Big)^{\log_2\alpha}\|X_i\|_p \qquad \text{for } q \ge p \ge 2.$$

Let $Y_i := |X_i|^{1/\log_2\alpha}\operatorname{sgn} X_i$. We may assume without loss of generality that $\alpha \ge 2$. Then $X_i = |Y_i|^{1/r}\operatorname{sgn} Y_i$ with $r := \frac{1}{\log_2\alpha} \in (0,1)$, and

$$\|Y_i\|_q \le 2\frac{q}{p}\|Y_i\|_p \qquad \text{for } q \ge p \ge 2\log_2\alpha. \qquad (2.27)$$

Take $2\log_2\alpha = q \ge p \ge 2$. Then by Hölder's inequality and (2.27) with exponents $\frac{p(q-1)}{p-1}$ and q we get

$$\|Y_i\|_q^q = \mathbb{E}|Y_i||Y_i|^{q-1} \le \left(\mathbb{E}|Y_i|^p\right)^{\frac{1}{p}} \left(\mathbb{E}|Y_i|^{\frac{p(q-1)}{p-1}}\right)^{\frac{p-1}{p}} \le \|Y_i\|_p \|Y_i\|_q^{q-1} \left(2\frac{p(q-1)}{q(p-1)}\right)^{q-1}.$$

Observe that

$$\left(2\frac{p(q-1)}{q(p-1)}\right)^{q-1} \le 4^{q-1} \le \frac{1}{4}\alpha^4,$$

so

$$\|Y_i\|_q \le \frac{1}{4}\alpha^4 \|Y_i\|_p \qquad \text{for } 2\log_2\alpha = q \ge p \ge 2.$$

Thus for any value of α we get

$$\|Y_i\|_q \le \max\left\{2, \frac{1}{2}\alpha^4\right\}\frac{q}{p}\|Y_i\|_p \qquad \text{for } q \ge p \ge 2.$$

Hence, by [24, Theorem 2.3] the variables $Y_1, \ldots, Y_n$ satisfy (2.20) with a constant L depending only on α (in fact for arbitrary, not only unconditional sets T) and the assertion follows by Proposition 2.18. □

2.3.2 Symmetrization argument

We will use the following proposition to prove that we may skip the unconditionality assumption in Corollary 2.19.

Proposition 2.20. *Let $(X_i)_{i=1}^n$ be a sequence of independent random variables with finite second moments and let $(\varepsilon_i)_{i=1}^n$ be a Bernoulli sequence independent of $(X_i)_{i=1}^n$. Then for any nonempty $T \subset \mathbb{R}^n$ and $p \ge 1$,*

$$\mathbb{E}_X \sup_{t\in T}\left(\mathbb{E}_\varepsilon\left|\sum_{i=1}^n t_i\varepsilon_i X_i\right|^p\right)^{1/p} \le C\left[\mathbb{E}\sup_{t\in T}\sum_{i=1}^n t_i\varepsilon_i X_i + \sup_{t\in T}\left(\mathbb{E}\left|\sum_{i=1}^n t_i\varepsilon_i X_i\right|^p\right)^{1/p}\right]. \tag{2.28}$$

Proof. Since this is only a matter of normalization we may and do assume that $\mathbb{E}X_i^2 = 1$ for all i. We will frequently use the result of Hitczenko from [15]:

$$\left\|\sum_{i=1}^n t_i\varepsilon_i\right\|_p \sim \sum_{i\le p} t_i^* + \sqrt{p}\sqrt{\sum_{i>p}|t_i^*|^2}.$$

Recall that $(t_i^*)_{i=1}^n$ denotes the non-increasing rearrangement of $(|t_i|)_{i=1}^n$. Since for every $t \in \mathbb{R}^n$ we know that $\|\sum_{i=1}^n t_i\varepsilon_i\|_1 \sim \|\sum_{i=1}^n t_i\varepsilon_i\|_2$, it is enough to consider $p \ge 2$ only.

Let m be such an integer that $2m \le p < 2(m+1)$. Then, by the symmetry of X_i, ε_i, and the independence of $X_1, \ldots, X_n, \varepsilon_1, \ldots, \varepsilon_n$ we have

$$\begin{aligned}\Big\|\sum_{i=1}^n t_i\varepsilon_i X_i\Big\|_p &\ge \Big\|\sum_{i=1}^n t_i\varepsilon_i X_i\Big\|_{2m} \\ &= \Big(\sum_{i_1+\ldots i_n=m} c_{i_1,\ldots,i_n} t_1^{2i_1}\ldots t_n^{2i_n}\mathbb{E}X_1^{2i_1}\ldots\mathbb{E}X_n^{2i_n}\Big)^{1/2m} \\ &\ge \Big(\sum_{i_1+\ldots i_n=m} c_{i_1,\ldots,i_n} t_1^{2i_1}\ldots t_n^{2i_n}\Big)^{1/2m} \\ &= \Big(\sum_{i_1+\ldots i_n=m} c_{i_1,\ldots,i_n} t_1^{2i_1}\ldots t_n^{2i_n}\mathbb{E}\varepsilon_1^{2i_1}\ldots\mathbb{E}\varepsilon_n^{2i_n}\Big)^{1/2m} = \Big\|\sum_{i=1}^n t_i\varepsilon_i\Big\|_{2m},\end{aligned}$$

where

$$c_{i_1,\ldots,i_n} = \frac{(2i_1+\ldots+2i_n)!}{(2i_1)!\ldots(2i_n)!}.$$

Therefore to establish (2.28) it is enough to show that

$$\mathbb{E}\sup_{t\in T}\Big(\mathbb{E}_\varepsilon\Big|\sum_{i=1}^n t_i\varepsilon_i X_i\Big|^p\Big)^{1/p} \le C\Big(\mathbb{E}\sup_{t\in T}\sum_{i=1}^n t_i\varepsilon_i X_i + pa\Big), \tag{2.29}$$

where

$$a := \frac{1}{p}\sup_{t\in T}\Big(\sum_{i\le p} t_i^* + \sqrt{p}\Big(\sum_{i>p}|t_i^*|^2\Big)^{1/2}\Big).$$

To this end observe that since

$$\Big\|\sum_{i=1}^n u_i\varepsilon_i\Big\|_p \le C\sqrt{p}\|u\|_2, \quad \Big\|\sum_{i=1}^n u_i\varepsilon_i\Big\|_p \le \|u\|_1,$$
$$\text{and} \quad \Big\|\sum_{i=1}^n u_i\varepsilon_i\Big\|_p = \Big\|\sum_{i=1}^n |u_i|\varepsilon_i\Big\|_p,$$

we have

$$\Big\|\sum_{i=1}^n u_i\varepsilon_i\Big\|_p \le \sum_{i=1}^n (|u_i|-a)_+ + C\sqrt{p}\Big(\sum_{i=1}^n \min\{u_i^2, a^2\}\Big)^{1/2}.$$

Thus

$$\begin{aligned}&\mathbb{E}_X\sup_{t\in T}\Big(\mathbb{E}_\varepsilon\Big|\sum_{i=1}^n t_i\varepsilon_i X_i\Big|^p\Big)^{1/p} \\ &\quad\le \mathbb{E}\sup_{t\in T}\sum_{i=1}^n\big(|t_iX_i|-a\big)_+ + C\sqrt{p}\Big(\mathbb{E}\sup_{t\in T}\sum_{i=1}^n\min\big\{(t_iX_i)^2, a^2\big\}\Big)^{1/2}.\end{aligned} \tag{2.30}$$

To estimate the first term above observe that

$$\mathbb{E}\sup_{t\in T}\sum_{i=1}^n\big(|t_iX_i|-a\big)_+$$
$$\le \sup_{t\in T}\mathbb{E}\sum_{i=1}^n\big(|t_iX_i|-a\big)_+ + \mathbb{E}\sup_{t\in T}\sum_{i=1}^n\Big(\big(|t_iX_i|-a\big)_+ - \mathbb{E}\big(|t_iX_i'|-a\big)_+\Big),$$

where $(X_i')_i$ is a copy of $(X_i)_i$, independent of $(\varepsilon_i)_i$ and $(X_i)_i$.

Observe that for any u and i

$$\mathbb{E}\big(|uX_i|-a\big)_+ \le |u|\mathbb{E}|X_i| \le |u|\|X_i\|_2 = |u|$$

and, by the Cauchy-Schwarz inequality and the Markov inequality

$$\mathbb{E}\big(|uX_i|-a\big)_+ \le |u|\mathbb{E}|X_i|I_{\{|X_i|\ge a/|u|\}} \le |u|\|X_i\|_2\big(\mathbb{P}(|X_i|\ge a/|u|)\big)^{1/2}$$
$$\le |u|\|X_i\|_2^2\frac{|u|}{a} = \frac{u^2}{a}.$$

Hence for any $t\in T$

$$\sum_{i=1}^n\mathbb{E}\big(|t_iX_i|-a\big)_+ \le \sum_{i\le p}t_i^* + \frac{1}{a}\sum_{i>p}(t_i^*)^2 \le 2pa.$$

Moreover, by the Jensen inequality

$$\mathbb{E}\sup_{t\in T}\sum_{i=1}^n\Big(\big(|t_iX_i|-a\big)_+ - \mathbb{E}\big(|t_iX_i'|-a\big)_+\Big)$$
$$\le \mathbb{E}\sup_{t\in T}\sum_{i=1}^n\Big(\big(|t_iX_i|-a\big)_+ - \big(|t_iX_i'|-a\big)_+\Big)$$
$$= \mathbb{E}\sup_{t\in T}\sum_{i=1}^n\varepsilon_i\Big(\big(|t_iX_i|-a\big)_+ - \big(|t_iX_i'|-a\big)_+\Big)$$
$$\le \mathbb{E}\sup_{t\in T}\sum_{i=1}^n\varepsilon_i\big(|t_iX_i|-a\big)_+ + \mathbb{E}\sup_{t\in T}\sum_{i=1}^n -\varepsilon_i\big(|t_iX_i'|-a\big)_+.$$
$$= 2\mathbb{E}\sup_{t\in T}\sum_{i=1}^n\varepsilon_i\big(|t_iX_i|-a\big)_+.$$

Function $x\mapsto(|x|-a)_+$ is 1-Lipschitz, so Talagrand's comparison theorem for Bernoulli processes [36, Theorem 2.1] yields

$$\mathbb{E}\sup_{t\in T}\sum_{i=1}^n\varepsilon_i\big(|t_iX_i|-a\big)_+ \le \mathbb{E}\sup_{t\in T}\sum_{i=1}^n t_i\varepsilon_iX_i.$$

Therefore

$$\mathbb{E}\sup_{t\in T}\sum_{i=1}^{n}\big(|t_iX_i|-a\big)_+\le 2pa+2\mathbb{E}\sup_{t\in T}\sum_{i=1}^{n}t_i\varepsilon_iX_i. \tag{2.31}$$

Now we turn our attention to the other term in (2.30). We have

$$\mathbb{E}\sup_{t\in T}\sum_{i=1}^{n}\min\big\{(t_iX_i)^2,a^2\big\}$$
$$\le\sup_{t\in T}\mathbb{E}\sum_{i=1}^{n}\min\big\{(t_iX_i)^2,a^2\big\}+\mathbb{E}\sup_{t\in T}\sum_{i=1}^{n}\Big(\min\big\{(t_iX_i)^2,a^2\big\}-\mathbb{E}\min\big\{(t_iX_i')^2,a^2\big\}\Big).$$

We have

$$\sum_{i=1}^{n}\mathbb{E}\min\big\{(t_iX_i)^2,a^2\big\}\le\sum_{i=1}^{n}\min\big\{a^2,t_i^2\mathbb{E}X_i^2\big\}\le pa^2+\sum_{i>p}(t_i^*)^2\le 2pa^2.$$

Moreover, by the Jensen inequality

$$\begin{aligned}
&\mathbb{E}\sup_{t\in T}\sum_{i=1}^{n}\Big(\min\big\{(t_iX_i)^2,a^2\big\}-\mathbb{E}\min\big\{(t_iX_i')^2,a^2\big\}\Big)\\
&\qquad\le\mathbb{E}\sup_{t\in T}\sum_{i=1}^{n}\Big(\min\big\{(t_iX_i)^2,a^2\big\}-\min\big\{(t_iX_i')^2,a^2\big\}\Big)\\
&\qquad=\mathbb{E}\sup_{t\in T}\sum_{i=1}^{n}\varepsilon_i\Big(\min\big\{(t_iX_i)^2,a^2\big\}-\min\big\{(t_iX_i')^2,a^2\big\}\Big)\\
&\qquad\le 2\mathbb{E}\sup_{t\in T}\sum_{i=1}^{n}\varepsilon_i\min\big\{(t_iX_i)^2,a^2\big\}.
\end{aligned}$$

Function $x\mapsto\min\{x^2,a^2\}$ is $2a$-Lipschitz, so using the comparison theorem for Bernoulli processes again we get

$$\mathbb{E}\sup_{t\in T}\sum_{i=1}^{n}\varepsilon_i\min\big\{(t_iX_i)^2,a^2\big\}\le 2a\mathbb{E}\sup_{t\in T}\sum_{i=1}^{n}t_i\varepsilon_iX_i.$$

Thus

$$\begin{aligned}
p\mathbb{E}\sup_{t\in T}\sum_{i=1}^{n}\min\big\{(t_iX_i)^2,a^2\big\}&\le 2p^2a^2+4pa\mathbb{E}\sup_{t\in T}\sum_{i=1}^{n}t_i\varepsilon_iX_i\\
&\le\Big(2pa+\mathbb{E}\sup_{t\in T}\sum_{i=1}^{n}t_i\varepsilon_iX_i\Big)^2. \qquad (2.32)
\end{aligned}$$

Estimate (2.29) follows by (2.30)-(2.32). □

Proof of Theorem 2.9. Since it is enough to consider $T \cup (-T)$ instead of T, we may and do assume that the set T is symmetric, i.e. $T = -T$.

Assume first that the variables X_i are also symmetric. Let $\varepsilon = (\varepsilon_i)_{i=1}^n$ be a Bernoulli sequence independent of $(X_i)_{i=1}^n$. Weak and strong moments of $(\varepsilon_i)_{i=1}^n$ are comparable:

$$\Big(\mathbb{E}\sup_{s\in S}\Big|\sum_{i=1}^n s_i\varepsilon_i\Big|^p\Big)^{1/p} \le C\left[\mathbb{E}\sup_{s\in S}\Big|\sum_{i=1}^n s_i\varepsilon_i\Big| + \sup_{s\in S}\Big(\mathbb{E}\Big|\sum_{i=1}^n s_i\varepsilon_i\Big|^p\Big)^{1/p}\right]$$

(this follows for example by Corollary 3.4, since ε_i have log-concave tails) Hence the symmetry of X_i and inequalities $(a+b)^p \le 2^p(a^p+b^p)$, $(a+b)^{1/p} \le a^{1/p}+b^{1/p}$ yield

$$\begin{aligned}\Big(\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p} &= \Big(\mathbb{E}_X\mathbb{E}_\varepsilon\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\varepsilon_i\Big|^p\Big)^{1/p} \\ &\le 2C\left[\Big(\mathbb{E}_X\Big(\mathbb{E}_\varepsilon\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\varepsilon_i\Big|\Big)^p\Big)^{1/p} + \Big(\mathbb{E}_X\sup_{t\in T}\mathbb{E}_\varepsilon\Big|\sum_{i=1}^n t_iX_i\varepsilon_i\Big|^p\Big)^{1/p}\right].\end{aligned} \tag{2.33}$$

Since T is symmetric, for $x \in \mathbb{R}^n$ we have

$$\mathbb{E}_\varepsilon\sup_{t\in T}\Big|\sum_{i=1}^n t_ix_i\varepsilon_i\Big| = \sup_{t\in T_1}\sum_{i=1}^n t_ix_i,$$

where

$$T_1 := \{(\mathbb{E}_\varepsilon s_i(\varepsilon)\varepsilon_i)_{i=1}^n :\ s\colon \{-1,1\}^n \to T\}$$

is an unconditional subset of $\mathbb{R}^n$. Estimate (2.26) applied for T_1 instead of T yields

$$\begin{aligned}&\Big(\mathbb{E}_X\Big(\mathbb{E}_\varepsilon\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\varepsilon_i\Big|\Big)^p\Big)^{1/p} \\ &\qquad\le C(\alpha)\left[\mathbb{E}_X\mathbb{E}_\varepsilon\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\varepsilon_i\Big| + \sup_{t\in T_1}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p}\right].\end{aligned}$$

By the symmetry of X_i and T we have

$$\mathbb{E}_X\mathbb{E}_\varepsilon\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\varepsilon_i\Big| = \mathbb{E}\sup_{t\in T}\sum_{i=1}^n t_iX_i.$$

Moreover,

$$T_1 \subset S(T) := \operatorname{conv}\{(\eta_it_i)_{i=1}^n : \eta\in\{-1,1\}^n, t\in T\},$$

hence

$$\sup_{t\in T_1}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p} \le \sup_{t\in S(T)}\Big(\mathbb{E}\Big|\sum t_iX_i\Big|^p\Big)^{1/p} = \sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p}.$$

Thus

$$\Big(\mathbb{E}_X\Big(\mathbb{E}_\varepsilon \sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\varepsilon_i\Big|\Big)^p\Big)^{1/p} \le C(\alpha)\Big[\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big| + \sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p}\Big]. \tag{2.34}$$

Let $q = p/(p-1)$ be the Hölder dual of p. For $x\in\mathbb{R}^n$ we have

$$\Big(\sup_{t\in T}\mathbb{E}_\varepsilon\Big|\sum_{i=1}^n t_ix_i\varepsilon_i\Big|^p\Big)^{1/p} = \sup_{t\in T_2}\sum_{i=1}^n t_ix_i,$$

where

$$T_2 = \{(\mathbb{E}_\varepsilon t_i\varepsilon_i h(\varepsilon))_{i=1}^n\colon\ t\in T, h\colon\{-1,1\}^n\to\mathbb{R}, \mathbb{E}_\varepsilon|h(\varepsilon)|^q\le 1\}$$

is an unconditional subset of $\mathbb{R}^n$. Estimate (2.26) applied for T_2 instead of T yields

$$\Big(\mathbb{E}_X\sup_{t\in T}\mathbb{E}_\varepsilon\Big|\sum_{i=1}^n t_iX_i\varepsilon_i\Big|^p\Big)^{1/p}$$
$$\le C(\alpha)\Big[\mathbb{E}_X\Big(\sup_{t\in T}\mathbb{E}_\varepsilon\Big|\sum_{i=1}^n t_iX_i\varepsilon_i\Big|^p\Big)^{1/p} + \sup_{t\in T_2}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p}\Big].$$

Proposition 2.20 and the symmetry of X_i gives

$$\mathbb{E}_X\Big(\sup_{t\in T}\mathbb{E}_\varepsilon\Big|\sum_{i=1}^n t_iX_i\varepsilon_i\Big|^p\Big)^{1/p} \le C\Big[\mathbb{E}\sup_{t\in T}\sum_{i=1}^n t_iX_i + \sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p}\Big].$$

Since $T_2\subset S(T)$ we have

$$\sup_{t\in T_2}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p} \le \sup_{t\in S(T)}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p} = \sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p}.$$

Thus

$$\Big(\mathbb{E}_X\sup_{t\in T}\mathbb{E}_\varepsilon\Big|\sum_{i=1}^n t_iX_i\varepsilon_i\Big|^p\Big)^{1/p} \le C(\alpha)\Big[\mathbb{E}\sup_{t\in T}\sum_{i=1}^n t_iX_i + \sup_{t\in T}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p}\Big]. \tag{2.35}$$

Estimate (2.6) follows (for symmetric X_i's) by (2.33)-(2.35)

In the case when the variables X_i are centred, but not necessarily symmetric let $(X_1', \ldots, X_n')$ be an independent copy of $(X_1, \ldots, X_n)$. Then $X_i - X_i'$ are symmetric. The Jensen inequality and the assumption on X_i imply that for any $p \geq 2$ we have

$$\|X_i - X_i'\|_{2p} \leq 2\|X_i\|_{2p} \leq 2\alpha\|X_i - \mathbb{E}X_i\|_p \leq 2\alpha\|X_i - X_i'\|_p.$$

Therefore, Theorem 2.9 applied to $(X_1 - X_1', \ldots, X_n - X_n')$ implies

$$\begin{aligned}
\left(\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big|^p\right)^{1/p} &= \left(\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_i(X_i - \mathbb{E}X_i')\Big|^p\right)^{1/p} \leq \left(\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_i(X_i - X_i')\Big|^p\right)^{1/p} \\
&\leq C(2\alpha)\left[\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_i(X_i - X_i')\Big| + \sup_{t\in T}\left(\mathbb{E}\Big|\sum_{i=1}^n t_i(X_i - X_i')\Big|^p\right)^{1/p}\right] \\
&\leq 2C(2\alpha)\left[\mathbb{E}\sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big| + \sup_{t\in T}\left(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\right)^{1/p}\right],
\end{aligned}$$

what finishes the proof in the general case. □

Remark 2.21. It follows by the proof of [24, Theorem 2.3] that if $(X_i)_{i=1}^n$ is symmetric, independent and for any i moments of X_i grow β-regularly, then the comparison of weak and strong moments of suprema of linear combinations of variables X_i holds with a constant $C(\beta) = C\beta^{11}$. Therefore, we may follow the constants in the proofs above to obtain that Theorem 2.9 holds with $C(\alpha) = C^{\log_2^2 \alpha}$.

2.3.3 From comparison of weak and strong moments to comparison of weak and strong tails

In this subsection we prove Corollary 2.11 and Corollary 2.12. To this end we need the following lemma.

Lemma 2.22. *Assume $X_1, X_2, \ldots$ satisfy the assumptions of Theorem 2.9. Then for any $t \in \mathbb{R}^n$,*

$$\Big\|\sum_{i=1}^n t_iX_i\Big\|_p \leq C(\alpha)\left(\frac{p}{q}\right)^{\max\{1/2,\log_2\alpha\}}\Big\|\sum_{i=1}^n t_iX_i\Big\|_q \quad \textit{for } p \geq q \geq 2. \tag{2.36}$$

Proof. Let $\beta := \max\{1/2, \log_2\alpha\}$. It is enough to show that for positive integers $k \geq l$ we have

$$\Big\|\sum_{i=1}^n t_iX_i\Big\|_{2k} \leq C\alpha\left(\frac{k}{l}\right)^{\beta}\Big\|\sum_{i=1}^n t_iX_i\Big\|_{2l}.$$

A standard symmetrization argument shows that we may assume that the random variables X_i are symmetric (see the proof of Theorem 2.9 in the non-symmetric case).

Using the hypercontractivity method [16, Section 3.3], it is enough to show that for $1 \le i \le n$,

$$\left\| s + \frac{t}{2\sqrt{2}e\alpha}\left(\frac{l}{k}\right)^{\beta} X_i \right\|_{2k} \le \left\| s + tX_i \right\|_{2l} \quad \text{for all } s,t \in \mathbb{R}.$$

This reduces to the following claim.

Claim. Suppose that Y is a symmetric random variable such that $\|Y\|_{2p} \le \alpha\|Y\|_p$ for some $\alpha \ge 1$ and every $p \ge 2$. Let $k \ge l$ be positive integers. Then

$$\left\|1 + \sigma Y\right\|_{2k} \le \left\|1 + Y\right\|_{2l}, \qquad \text{where } \sigma := \frac{1}{2\sqrt{2}e\alpha}\left(\frac{l}{k}\right)^{\beta}.$$

To show the claim observe first that (to see this we proceed as in the first part of proof of Corollary 2.19)

$$\|Y\|_q \le \alpha\left(\frac{q}{p}\right)^{\log_2 \alpha} \|Y\|_p \le \alpha\left(\frac{q}{p}\right)^{\beta} \|Y\|_p \quad \text{for } q \ge p \ge 2. \tag{2.37}$$

Moreover we have

$$\begin{aligned}\mathbb{E}\left|1 + \sigma Y\right|^{2k} &= 1 + \sum_{j=1}^{k} \binom{2k}{2j} \mathbb{E}\left|\sigma Y\right|^{2j} \le 1 + \sum_{j=1}^{k}\left(\frac{ek}{j}\sigma\|Y\|_{2j}\right)^{2j} \\ &\le 1 + \sum_{i=1}^{k} 2^{-i} \sup_{1\le j\le k}\left(\frac{\sqrt{2}ek}{j}\sigma\|Y\|_{2j}\right)^{2j} \le 1 + \sup_{1\le j\le k}\left(\frac{\sqrt{2}ek}{j}\sigma\|Y\|_{2j}\right)^{2j},\end{aligned}$$

so it is enough to show that

$$1 + \left(\frac{k^{1-\beta}l^{\beta}}{2j\alpha}\|Y\|_{2j}\right)^{2j} \le \left\|1 + Y\right\|_{2l}^{2k} \quad \text{for } j = 1, 2 \ldots k. \tag{2.38}$$

To this end we will use the following deterministic inequality:

$$(1+u)^p \ge \left(1 + \frac{p}{q}u\right)^q \ge 1 + \left(\frac{p}{q}u\right)^q \quad \text{for } p \ge q \ge 1 \text{ and } u \ge 0, \tag{2.39}$$

and a simple lower bound for $\|1+Y\|_{2l}^{2l}$:

$$\mathbb{E}|1+Y|^{2l} = 1 + \sum_{r=1}^{l}\binom{2l}{2r}\mathbb{E}|Y|^{2r} \ge 1 + \sum_{r=1}^{l}\left(\frac{l}{r}\|Y\|_{2r}\right)^{2r}. \tag{2.40}$$

In order to prove the first part of (2.39) note that it is equivalent to the fact that for all $u > 0$ a function $p \mapsto (1 + \frac{u}{p})^p$ is non-decreasing. The derivative of this function is equal to $\ln(1 + \frac{u}{p}) - (\frac{p}{u} + 1)^{-1}$, what is nonnegative by the inequality $\ln(1 + x) \geq \frac{x}{1+x}$ for $x > -1$.

Assume first that $1 \leq j \leq \frac{k}{l}$. Estimate (2.37) applied with $p = 2$ and $q = 2j$ yields

$$\frac{k^{1-\beta}l^{\beta}}{2j\alpha}\|Y\|_{2j} \leq \frac{k^{1-\beta}l^{\beta}}{j^{1-\beta}}\|Y\|_2 \leq \sqrt{\frac{kl}{j}}\|Y\|_2,$$

where the last inequality holds since $\beta \geq \frac{1}{2}$ and $k \geq jl$. Inequalities (2.40) and (2.39) (applied with $p = k/l$ and $q = j$) yield

$$\big\|1 + Y\big\|_{2l}^{2k} \geq \big(1 + (l\|Y\|_2)^2\big)^{k/l} \geq 1 + \left(\sqrt{\frac{kl}{j}}\|Y\|_2\right)^{2j}$$

so (2.38) holds for $j \leq \frac{k}{l}$.

If $j \geq \frac{k}{l}$ we choose $r = \lceil jl/k \rceil$. Then $jl \leq kr \leq 2jl$. Since $1 \leq r \leq l$, the estimate (2.40) gives

$$\big\|1 + Y\big\|_{2l}^{2k} \geq \left(1 + \left(\frac{l}{r}\|Y\|_{2r}\right)^{2r}\right)^{k/l} \geq \left(1 + \left(\frac{l}{r}\|Y\|_{2r}\right)^{2r}\right)^{j/r} \geq 1 + \left(\frac{l}{r}\|Y\|_{2r}\right)^{2j},$$

where to get the last two inequalities we used $k/l \geq j/r$ and $j/r \geq 1$. Applying estimate (2.37) with $2j$ and $2r$ instead of p and q we get

$$\frac{k^{1-\beta}l^{\beta}}{2j\alpha}\|Y\|_{2j} \leq \frac{k^{1-\beta}l^{\beta}}{2j}\left(\frac{j}{r}\right)^{\beta}\|Y\|_{2r} \leq \frac{k}{2j}\|Y\|_{2r} \leq \frac{l}{r}\|Y\|_{2r},$$

which completes the proof of the claim in the remaining case. □

Remark 2.23. It will be clear from the proof below that Theorem 2.5 implies an analogue of Corollary 2.11 for log-concave random vectors:

$$\mathbb{P}\left(\sup_{t\in B^n_{r'}}\Big|\sum_{i=1}^n t_iX_i\Big| \geq rD_1\left[u + \mathbb{E}\sup_{t\in B^n_{r'}}\Big|\sum_{i=1}^n t_iX_i\Big|\right]\right) \leq D_2 \sup_{t\in B^n_{r'}}\mathbb{P}\left(\Big|\sum_{i=1}^n t_iX_i\Big| \geq u\right), \tag{2.41}$$

for all $r \in [1, \infty)$, where D_1 and D_2 are universal constants.

Proof of Corollary 2.11. Let

$$S := \sup_{t\in T}\Big|\sum_{i=1}^n t_iX_i\Big|.$$

By the Paley-Zygmund inequality and (2.36) we have for $t \in T$ and $p \geq 2$,

$$\mathbb{P}\Big(\Big|\sum_{i=1}^n t_i X_i\Big| \geq \frac{1}{2}\Big\|\sum_{i=1}^n t_i X_i\Big\|_p\Big) = \mathbb{P}\Big(\Big|\sum_{i=1}^n t_i X_i\Big|^p \geq 2^{-p}\mathbb{E}\Big|\sum_{i=1}^n t_i X_i\Big|^p\Big)$$
$$\geq (1-2^{-p})^2\left(\frac{\big\|\sum_{i=1}^n t_i X_i\big\|_p}{\big\|\sum_{i=1}^n t_i X_i\big\|_{2p}}\right)^{2p} \geq e^{-C_3(\alpha)p}. \tag{2.42}$$

In order to show (2.9) we consider 3 cases.

Case 1. $2u < \sup_{t\in T}\|\sum_{i=1}^n t_i X_i\|_2$. Then by (2.42)

$$\sup_{t\in T}\mathbb{P}\Big(\Big|\sum_{i=1}^n t_i X_i\Big| \geq u\Big) \geq e^{-2C_3(\alpha)}$$

and (2.9) obviously holds if $C_2(\alpha) \geq \exp(2C_3(\alpha))$.

Case 2. $\sup_{t\in T}\|\sum_{i=1}^n t_i X_i\|_2 \leq 2u < \sup_{t\in T}\|\sum_{i=1}^n t_i X_i\|_\infty$. Let us then define

$$p := \sup\Big\{q \geq 2C_3(\alpha)\colon\ \sup_{t\in T}\Big\|\sum_{i=1}^n t_i X_i\Big\|_{q/C_3(\alpha)} \leq 2u\Big\}.$$

By (2.42) we have

$$\sup_{t\in T}\mathbb{P}\Big(\Big|\sum_{i=1}^n t_i X_i\Big| \geq u\Big) \geq e^{-p}.$$

By (2.36) we have $\sup_{t\in T}\|\sum_{i=1}^n t_i X_i\|_p \leq C(\alpha)u$, so by Theorem 2.9 and Chebyshev's inequality we have

$$\mathbb{P}(S \geq C_1(\alpha)(\mathbb{E}S + u)) \leq \mathbb{P}(S \geq e\|S\|_p) \leq e^{-p}$$

for $C_1(\alpha)$ large enough. Thus (2.9) holds in this case.

Case 3. $2u > \sup_{t\in T}\|\sum_{i=1}^n t_i X_i\|_\infty = \|S\|_\infty$. Then $\mathbb{P}(S \geq 2u) = 0$ and (2.9) holds for any $C_1(\alpha) \geq 2$. □

Proof of Corollary 2.12. The result is an immediate consequence of Theorem 2.9 and inequality (2.36). □

2.3.4 Comparison of weak and strong moments of suprema implies comparison of moments p and $2p$

Proof of Theorem 2.10. We will use the assumption (2.7) for T containing all vectors of the standard base of $\mathbb{R}^n$ and their negatives, i.e. we will use only the inequality

$$\Big(\mathbb{E}\sup_{1\le i\le n}|X_i|^p\Big)^{1/p}\le L\Big[\mathbb{E}\sup_{1\le i\le n}|X_i|+\|X_1\|_p\Big]. \tag{2.43}$$

Fix $p\ge 2$ and let $n:=\lfloor (4L)^{2p}\rfloor+1$, $A:=n^{1/p}\|X_1\|_p$. If $A\ge \|X_1\|_{2p}$, then (2.8) holds with $\alpha=(4L)^2+1$. Hence we may and do assume $A\le\|X_1\|_{2p}$.

Obviously

$$\mathbb{P}\Big(\sup_{1\le i\le n}|X_i|\ge t\Big)\le\min\big\{1,n\mathbb{P}\big(|X_1|\ge t\big)\big\}.$$

Moreover, if $\mathbb{P}(|X_1|\ge t)\le\frac{1}{n}$,

$$\begin{aligned}\mathbb{P}\Big(\sup_{1\le i\le n}|X_i|\ge t\Big)&=1-\mathbb{P}\big(|X_1|<t\big)^n=\mathbb{P}\big(|X_1|\ge t\big)\sum_{k=0}^{n-1}\mathbb{P}\big(|X_1|<t\big)^k\\&\ge\mathbb{P}\big(|X_1|\ge t\big)\cdot n\Big(1-\frac{1}{n}\Big)^{n-1}\ge\frac{n}{3}\mathbb{P}\big(|X_1|\ge t\big).\end{aligned}$$

Since $\mathbb{P}(|X_1|\ge A)\le\frac{1}{n}$ (which follows by the Markov inequality) and $A\le\|X_1\|_{2p}$, we have

$$\begin{aligned}\mathbb{E}\sup_{1\le i\le n}|X_i|^{2p}&\ge 2p\int_A^\infty t^{2p-1}\mathbb{P}\Big(\sup_{1\le i\le n}|X_i|\ge t\Big)dt\ge 2p\int_A^\infty t^{2p-1}\frac{n}{3}\mathbb{P}\big(|X_1|\ge t\big)dt\\&=\frac{n}{3}\mathbb{E}\big(|X_1|^{2p}-A^{2p}\big)_+\ge\frac{n}{3}\big(\|X_1\|_{2p}^{2p}-A^{2p}\big)\ge\frac{n}{3}\big(\|X_1\|_{2p}-A\big)^{2p}\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}\sup_{1\le i\le n}|X_i|&\le A+\int_A^\infty\mathbb{P}\Big(\sup_{1\le i\le n}|X_i|\ge t\Big)dt\le A+n\int_A^\infty\mathbb{P}\big(|X_1|\ge t\big)dt\\&\le A+n\mathbb{E}\big(|X_1|\mathbf{1}_{\{|X_1|\ge A\}}\big)\le A+n\|X_1\|_p\mathbb{P}\big(|X_1|\ge A\big)^{1-\frac{1}{p}}\\&\le A+n^{1/p}\|X_1\|_p,\end{aligned}$$

where in the last inequality we used again the fact that $\mathbb{P}(|X_1|\ge A)\le\frac{1}{n}$.

Thus our choice of n and A, and (2.43) (applied to $2p$ instead of p) imply that

$$\begin{aligned}
2L\|X_1\|_{2p} &\le \frac{1}{2}n^{\frac{1}{2p}}\|X_1\|_{2p} \le \frac{1}{2}n^{\frac{1}{2p}}A + \Big(\mathbb{E}\sup_{1\le i\le n}|X_i|^{2p}\Big)^{1/(2p)} \\
&\le \frac{1}{2}n^{\frac{1}{2p}}A + L\Big[\mathbb{E}\sup_{1\le i\le n}|X_i| + \|X_1\|_{2p}\Big] \\
&\le \frac{1}{2}n^{\frac{1}{2p}}A + LA + Ln^{\frac{1}{p}}\|X_1\|_p + L\|X_1\|_{2p} \\
&\le \|X_1\|_p\Big(\frac{1}{2}(4L+1)n^{\frac{1}{p}} + 2Ln^{\frac{1}{p}}\Big) + L\|X_1\|_{2p} \\
&\le \left(4L + \frac{1}{2}\right)\big((4L)^2+1\big)\|X_1\|_p + L\|X_1\|_{2p}.
\end{aligned}$$

Thus

$$\|X_1\|_{2p} \le \left(4 + \frac{1}{2L}\right)\big(16L^2+1\big)\|X_1\|_p.$$

□

Remark 2.24. It is clear from the proof above that we may take $\alpha(L) = CL^2$ in Theorem 2.10.

Chapter 3

Convex infimum convolution ineqality

Functional inequalities such as the Poincaré, log-Sobolev, or Marton-Talagrand inequality to name a few, play a crucial role in studying concentration of measure, an important cornerstone of the local theory of Banach spaces. In this chapter we focus on another example of such inequalities, the infimum convolution inequality, introduced by Maurey in [27].

Let X be a random vector with values in $\mathbb{R}^n$ and let $\varphi : \mathbb{R}^n \to [0, \infty]$ be a measurable function. We say that the pair (X, φ) satisfies the *infimum convolution inequality* (ICI for short) if for every bounded measurable function $f : \mathbb{R}^n \to \mathbb{R}$,

$$\mathbb{E}\, e^{f \square \varphi(X)}\, \mathbb{E}\, e^{-f(X)} \leq 1, \tag{3.1}$$

where $f \square \varphi$ denotes the infimum convolution of f and φ defined as $f \square \varphi(x) = \inf\{f(y) + \varphi(x - y) : y \in \mathbb{R}^n\}$ for $x \in \mathbb{R}^n$. The function φ is called a *cost function* and f is called a *test function*. We also say that the pair (X, φ) satisfies the *convex infimum convolution inequality* if (3.1) holds for every convex function $f : \mathbb{R}^n \to \mathbb{R}$ bounded from below.

Maurey showed that Gaussian and exponential random variables satisfy the ICI with a quadratic and quadratic-linear cost function respectively. Thanks to the tensorisation property of the ICI, he recovered the Gaussian concentration inequality as well as the so-called Talagrand two-level concentration inequality for the exponential product measure. Moreover, Maurey proved that bounded random variables satisfy the convex ICI with a quadratic cost function (see also Lemma 3.2 in [33] for an improvement).

Later on, Maurey's idea was developed further by Latała and Wojtaszczyk who studied comprehensively the ICI in [26]. By testing with linear functions, they observed that the optimal cost function is given by the Legendre transform of the cumulant-generating function (here optimal means largest possible, up to a scaling constant, because the larger the cost function is, the better (3.1) gets). They introduced the notion of optimal infimum convolution inequalities, established them for log-concave product measures and uniform measures on ℓ_p-balls, and put forward important, challenging and far-reaching conjectures.

The recent works [10] and [9] enable to view the ICI from a different perspective. In [10] Gozlan, Roberto, Samson, and Tetali introduce weak transport-entropy inequalities and establish their dual formulations. The dual formulations are exactly the convex ICIs. In [9] Gozlan, Roberto, Samson, Shu and Tetali investigate extensively the weak transport cost inequalities on the real line, obtaining a characterisation for arbitrary cost functions which are convex and quadratic near zero, thus providing a tool for studying the convex ICI. Around the same time, the convex ICI for the quadratic-linear cost function was fully understood by Feldheim, Marsiglietti, and Nayar in [8].

In this chapter, based on [35], we go along Latała and Wojtaszczyk's line of research and study the optimal convex ICI. Using the aforementioned novel tools from [9], we show that product measures with symmetric marginals having log-concave tails satisfy the optimal convex ICI, which complements Latała and Wojtaszczyk's result about log-concave product measures. This has applications to concentration and moment comparison. We also offer an example showing that the assumption of log-concave tails cannot be weakened substantially.

3.1 Main results

For a random vector X in $\mathbb{R}^n$ we define

$$\Lambda_X^*(x) := \mathcal{L}\Lambda_X(x) := \sup_{y\in\mathbb{R}^n}\{\langle x,y\rangle - \ln \mathbb{E}\, e^{\langle y,X\rangle}\},$$

which is the Legendre transform of the cumulant-generating function

$$\Lambda_X(x) := \ln \mathbb{E}\, e^{\langle x,X\rangle}, \qquad x\in\mathbb{R}^n.$$

If X is symmetric and the pair (X,φ) satisfies the ICI, then $\varphi(x) \le \Lambda_X^*(x)$ for every $x\in\mathbb{R}^n$ (see Remark 2.12 in [26]). In other words, Λ_X^* is the optimal cost function φ for which the ICI can hold. Since this conclusion is obtained by testing (3.1) with linear functions, the same holds for the convex ICI. Following [26] we shall say that X satisfies (convex) IC(β) if the pair $(X, \Lambda_X^*(\cdot/\beta))$ satisfies the (convex) ICI.

We are ready to present our first main result.

Theorem 3.1. *There exists a universal constant $\beta \le 1680e$ such that every symmetric random variable with log-concave tails satisfies convex IC(β).*

The (convex) ICI tensorises and, consequently, the property (convex) IC tensorises: if independent random vectors X_i satisfy (convex) IC(β_i), $i = 1, \ldots, n$, then the vector $(X_1, \ldots, X_n)$ satisfies (convex) IC($\max \beta_i$) (see [27] and [26]). Therefore we have the following corollary.

Corollary 3.2. *Let X be a symmetric random vector with values in $\mathbb{R}^n$ and independent coordinates with log-concave tails. Then X satisfies convex IC(β) with a universal constant $\beta \leq 1680e$.*

Note that the class of distributions from Theorem 3.1 is wider than the class of symmetric log-concave product distributions considered by Latała and Wojtaszczyk in [26]. Among others, it contains measures which do not have a connected support, e.g. a symmetric Bernoulli random variable.

Recall that variables with log-concave tails are 1-regular (see Remark 1.2). However, the assumption of log-concave tails in Theorem 3.1 cannot be replaced by a weaker one of α-regularity of moments: if X is a symmetric random variable defined by

$$\mathbb{P}(|X| > t) = 1_{[0,2)}(t) + \sum_{k=1}^{\infty} e^{-2^k} 1_{[2^k, 2^{k+1})}(t), \quad t \geq 0, \tag{3.2}$$

then the moments of X grow α-regularly (for some $\alpha < \infty$), but there is no $C > 0$ such that the pair $(X, x \mapsto \max\{(Cx)^2, C|x|\})$ satisfies the convex ICI. All the more, X cannot satisfy convex IC(β) with any $\beta < \infty$ (see Section 3.4 for details). Thus it seems that the assumptions of Theorem 3.1 are not far from necessary conditions for the convex ICI to hold with an optimal cost function (random variables with moments growing regularly are akin to random variables with log-concave tails as the former can essentially be sandwiched between the latter, see (4.6) in [24]).

Our second main result is an application of Theorem 3.1 to moment comparison in a manner of the previous chapter.

Theorem 3.3. *Let X be a symmetric random vector with values in $\mathbb{R}^n$ which moments grow α-regularly. Suppose moreover that X satisfies convex IC(β). Then for every norm $\|\cdot\|$ on $\mathbb{R}^n$ and every $p \geq 2$ we have*

$$\Big(\mathbb{E}\big|\|X\| - E\|X\|\big|^p\Big)^{1/p} \leq C\alpha\beta\sigma_{\|\cdot\|,X}(p),$$

where C is a universal constant (one can take $C = 4\sqrt{2}e < 16$).

Immediately we obtain the following corollary stating that the comparison of weak and strong moments holds with a constant 1 at the first strong moment. Similar inequalities for Rademacher sums with the emphasis on exact values of constants have also been studied by Oleszkiewicz (see [28, Theorem 2.1]).

Corollary 3.4. *Let X be a symmetric random vector with values in $\mathbb{R}^n$ and with independent coordinates which have log-concave tails. Then for every norm $\|\cdot\|$ on $\mathbb{R}^n$ and every $p \geq 2$ we have*

$$\left(\mathbb{E}\,\|X\|^p\right)^{1/p} \leq \mathbb{E}\,\|X\| + D\sigma_{\|\cdot\|,X}(p), \tag{3.3}$$

where D is a universal constant (one can take $D = 6720\sqrt{2}e^2 < 70223$).

Note also that the constant standing at $\mathbb{E}\,\|X\|$ is equal to 1. If we only assume that the coordinates of X are independent and their moments grow α-regularly, then (3.3) does not always hold (the example here is a vector with independent coordinates distributed like in (3.2); see Section 3.4 for details), although by Theorem 2.9 it holds if we allow the constant at $\mathbb{E}\,\|X\|$ to be greater than 1 and to depend on α. Hence Corollary 3.4 and example (3.2) partially answer the following question raised in [23]: "For which vectors does the comparison of weak and strong moments hold with constant 1 at the first strong moment?"

The organization of the rest of this chapter is the following. In Section 3.2 and 3.3 we present the proofs of Theorem 3.1 and Proposition 3.3 respectively. In Section 3.4 we discuss example (3.2) in details.

3.2 Proof of Theorem 3.1

Our approach is based on a characterization – provided by Gozlan, Roberto, Samson, Shu, and Tetali in [9] – of measures on the real line which satisfy a weak transport-entropy inequality. We emphasize that our optimal cost functions need not be quadratic near the origin, therefore we cannot apply their characterization as is, but have to first fine-tune the cost functions a bit. We shall also need the following simple lemma.

Lemma 3.5. *If X is a symmetric random variable and $\mathbb{E}\,X^2 = \beta_1^{-2}$, then*

$$\Lambda_X^*(x/\beta_1) \leq x^2 \quad \textit{for } |x| \leq 1.$$

Proof. Since X is symmetric, we have

$$\mathbb{E}\,e^{tX} = 1 + \sum_{k=1}^{\infty} \frac{\|X\|_{2k}^{2k} t^{2k}}{(2k)!} \geq 1 + \sum_{k=1}^{\infty} \frac{\|X\|_{2}^{2k} t^{2k}}{(2k)!} = 1 + \sum_{k=1}^{\infty} \frac{\beta_1^{-2k} t^{2k}}{(2k)!} = \cosh(\beta_1^{-1}|t|).$$

Moreover, $\mathcal{L}\big(\ln\cosh(\cdot)\big)(|u|) \leq |u|^2$ for $|u| \leq 1$ (see for example the proof of [26, Proposition 3.3]). Therefore

$$\Lambda_X^*(x/\beta_1) = \mathcal{L}(\Lambda_X(\beta_1 \cdot))(x) \leq \mathcal{L}(\ln\cosh(\cdot))(x) \leq x^2 \qquad \text{for } |x| \leq 1. \qquad \square$$

Proof of Theorem 3.1. Note that $N(0) = 0$ and the function N is non-decreasing. First we tweak the assumptions and change the assertion to a more straightforward one.

Step 1 (first reduction). We claim that it suffices to prove the assertion for random variables for which the function N is strictly increasing on the set where it is finite[1] (or, in other words, $N(t) = 0$ only for $t = 0$). Indeed, suppose we have done this and let now X be any random variable satisfying the assumptions of the theorem. Let X_ε be a symmetric random variable such that $\mathbb{P}(|X_\varepsilon| \geq t) = \exp(-N_\varepsilon(t))$, where $N_\varepsilon(t) = N(t) \vee \varepsilon t$. If X and X_ε are represented in the standard way by the inverses of their CDFs on the probability space $(0, 1)$, then $|X_\varepsilon| \leq |X|$ a.s. (and also $X_\varepsilon \to X$ a.s. as $\varepsilon \to 0^+$). Hence $\Lambda_{X_\varepsilon} \leq \Lambda_X$ and therefore also $\Lambda^*_{X_\varepsilon} \geq \Lambda^*_X$.

The theorem applied to the random variable X_ε and the above inequality imply that the pair $(X_\varepsilon, \Lambda^*_X(\cdot/\beta))$ satisfies the convex ICI. Taking $\varepsilon \to 0^+$ we get the assertion for X (in the second integral we just use the fact that the test function f is bounded from below and thus e^{-f} is bounded from above; for the first integral it suffices (by the Fatou lemma) to prove the convergence of integrals on any interval $[-M, M]$, and on such an interval we have $f \square \Lambda^*_X(x/\beta) \leq f(x) + \Lambda^*_X(0) = f(x)$, and thus $\exp(\max_{[-M,M]} f)$ is a good majorant).

Step 2 (second reduction). We claim that it suffices to prove the assertion for random variables such that $\Lambda_X < \infty$. Indeed, suppose we have done this and let X be any random variable satisfying the assumptions of the theorem. Let $N_\varepsilon(t) = N(t) \vee \varepsilon^2 t^2$ and let X_ε be a symmetric random variable such that $\mathbb{P}(|X_\varepsilon| \geq t) = \exp(-N_\varepsilon(t))$. Then, similarly as in Step 1., $\Lambda_{X_\varepsilon} \leq \Lambda_Y < \infty$, where Y is symmetric and $\mathbb{P}(|Y| \geq t) = \exp(-\varepsilon^2 t^2)$. Thus we can apply the proposition to X_ε and we continue as in Step 1.

Step 3 (scaling). Due to the scaling properties of the Legendre transform, we can assume that $\mathbb{E}\, X^2 = \beta_1^{-2}$, where $\beta_1 := 2e$ (the case where $X \equiv 0$ is trivial). Note that then, by Markov's inequality, $e^{-N(1/2)} = \mathbb{P}(|X| \geq \frac{1}{2}) \leq 4\,\mathbb{E}\, X^2 = e^{-2}$, so

$$N(1/2) \geq 2. \tag{3.4}$$

Step 4 (reformulation). For $x \in \mathbb{R}$ let

$$\varphi(x) := \big(x^2 1_{\{|x|<1\}} + (2|x| - 1) 1_{\{|x|\geq 1\}}\big) \vee \Lambda^*_X(x/(2\beta_1)).$$

We claim that there exists a universal constant $\widetilde{b} \leq 1/420$, such that the pair $(X, \varphi(\widetilde{b}\cdot))$ satisfies the convex infimum convolution inequality. Of course the assertion follows immediately from that.

Note that φ is convex, increasing on $[0, \infty)$ (because $\Lambda^*_X(\cdot/(2\beta_1))$ is convex and symmetric and thus non-decreasing on $[0, \infty)$). Crucially, $\varphi(x) = x^2$ for $x \in [0, 1]$

[1] Recall that $N(t) := -\ln \mathbb{P}(|X| \geq t)$ for $t \geq 0$.

(by Lemma 3.5), so the cost function φ is quadratic on $[-1,1]$. Moreover, by Lemma 3.5, $\varphi^{-1}(3)=2$.

Let $U = F^{-1} \circ F_\nu$, where F, F_ν are the distribution functions of X and the symmetric exponential measure ν on $\mathbb{R}$, respectively. By [9, Theorem 1.1] we know that if there exists $b>0$ such that for every $x,y\in\mathbb{R}$ we have

$$\big|U(x)-U(y)\big| \le \frac{1}{b}\varphi^{-1}\big(1+|x-y|\big), \tag{3.5}$$

then the pair $(X, \varphi(\widetilde{b}\cdot))$, where $\widetilde{b} = \frac{b}{210\varphi^{-1}(2+1^2)} = \frac{b}{420}$, satisfies the convex ICI. We will show that (3.5) holds with $b=1$.

Step 5 (further reformulation). Let $a = \inf\{t>0 : N(t)=\infty\}$. We have three possibilities (recall that N is left-continuous):

- $a=\infty$. Then N is continuous, increasing, and transforms $[0,\infty]$ onto $[0,\infty]$. Also, F is increasing and therefore F^{-1} is the usual inverse of F.
- $a<\infty$ and $N(a)<\infty$. Then X has an atom at a. Moreover, $N(a) = \lim_{t\to a^-} N(t)$.
- $a<\infty$ and $N(a)=\infty=\lim_{t\to a^-} N(t)$.

Of course, in the first case one can extend N by putting $N(a)=\infty$, so that all formulas below make sense.

Note that

$$F(t) = \begin{cases} \frac{1}{2}\exp(-N(|t|)) & \text{if } t<0, \\ 1-\frac{1}{2}\exp(-N_+(t)) & \text{if } t\ge 0, \end{cases}$$

where $N_+(t)$ denotes the right-sided limit of N at t (which is different from $N(t)$ only if $t=a$ and X has an atom at a). Hence, F is continuous on the interval $(-a,a)$, the image of $(-a,a)$ under F is the interval $\big(\frac{1}{2}\exp(-N(a)), 1-\frac{1}{2}\exp(-N(a))\big)$, and we have $F(-a)=\frac{1}{2}\exp(-N(a))$ and $F(a)=1$. Since the image of $\mathbb{R}$ under U is equal to the image of $(0,1)$ under F^{-1}, we conclude that $U(\mathbb{R})=(-a,a)$ if $N(a)=\infty$ and $U(\mathbb{R})=[-a,a]$ if $N(a)<\infty$. Denote $A:=U(\mathbb{R})$.

When $N(a)<\infty$, it suffices to check condition (3.5) for $x,y\in[-a,a]$ (otherwise one can change x, y and decrease the right-hand side while not changing the value of the left-hand side of (3.5)). Take thus $x,y\in[-a,a]$ in a case $N(a)<\infty$ and arbitrary $x,y\in\mathbb{R}$ in a case $N(a)=\infty$. Then $N(|x|)\operatorname{sgn} x = U^{-1}(x)$ and $N(|y|)\operatorname{sgn} y = U^{-1}(y)$. Therefore, in order to verify (3.5) we need to check that

$$|x-y| \le \varphi^{-1}\big(1+\big|N(|x|)\operatorname{sgn}(x) - N(|y|)\operatorname{sgn}(y)\big|\big) \qquad \text{for } x,y\in A. \tag{3.6}$$

Since we consider the case when $\Lambda_X(t)$ is finite for every $t \in \mathbb{R}$, the Chernoff inequality applies, so for $t \geq \mathbb{E}\, X = 0$ we have

$$\frac{1}{2}e^{-N(t)} = \mathbb{P}(X \geq t) \leq e^{-\Lambda_X^*(t)},$$

so

$$N(t) \geq \Lambda_X^*(t) - \ln 2. \tag{3.7}$$

Note that $\varphi(|x-y|) < \infty$ for $x, y \in A$, since $\varphi(|x-y|) = \infty$ would imply $\Lambda_X^*(|x-y|/(2\beta_1)) = \infty$, and hence $\Lambda_X^*(|x-y|/2) = \infty$, and – by (3.7) – also $N(|x-y|/2) = \infty$, but for $x, y \in A$ we have $|x-y|/2 \in [0, a)$ when $N(a) = \infty$ or $|x-y|/2 \in [0, a]$ when $N(a) < \infty$ and in either case $N(|x-y|/2)$ is finite. Therefore for every $x, y \in A$ we have indeed $\varphi(|x-y|) < \infty$. Since $\varphi^{-1}(\varphi(z)) = z$ for z such that $\varphi(z) < \infty$ (because φ is then continuous and increasing on $[0, z]$), the condition (3.6) is implied by

$$\varphi\big(|x-y|\big) \leq 1 + \big|N(|x|)\operatorname{sgn} x - N(|y|)\operatorname{sgn} y\big| \qquad \text{for } x, y \in A. \tag{3.8}$$

In the next step we check that this is indeed satisfied.

Step 6 (checking the condition (3.8)*).* Let $x_0 = \inf\{x \geq 1 : 2x - 1 = \Lambda_X^*(\frac{x}{2\beta_1})\}$ (if $x_0 = \infty$ we simply do not have to consider Case 2 below). We consider three cases. We repeatedly use the fact that $uN(t) \geq N(ut)$ for $u \leq 1$, $t \geq 0$, which follows by the convexity of N and the property $N(0) = 0$.

Case 1. $|x-y| \leq 1$. Then $\varphi\big(|x-y|\big) = (x-y)^2 \leq 1$, so (3.8) is trivially satisfied.

Case 2. $|x-y| \geq x_0$. Then $\varphi\big(|x-y|\big) = \Lambda_X^*(\frac{1}{2\beta_1}|x-y|) \leq \Lambda_X^*(|x-y|/2)$. Inequality (3.7) implies that in order to prove (3.8) it suffices to show that if x, y are of the same sign, say $x, y \geq 0$, then $N\big(|x-y|/2\big) \leq |N(x) - N(y)|$ and if x, y have different signs, we have $N\big((|x| + |y|)/2\big) \leq N(|x|) + N(|y|)$.

By the convexity of N, for $s, t \geq 0$ we have

$$N\big((s+t)/2\big) \leq \frac{1}{2}N(s) + \frac{1}{2}N(t) \leq N(s) + N(t)$$

and

$$N(s/2) + N(t) \leq N(s) + N(t) \leq \frac{s}{s+t}N(s+t) + \frac{t}{s+t}N(s+t) = N(s+t).$$

This finishes the proof of (3.8) in Case 2.

Case 3. $1 \leq |x-y| \leq x_0$. Then $\varphi\big(|x-y|\big) = 2|x-y| - 1$. Consider two sub-cases:

(i) x, y have different signs. Without loss of generality we may assume $x \geq |y| \geq 0 \geq y$. Thus in order to obtain (3.8) it suffices to show that $N(x) \geq 2x + 2|y|$. Note that $1 \leq x + |y| \leq 2x$, so $x \geq \frac{1}{2}$. Thus

$$N(x) \geq N(1/2)2x \overset{(3.4)}{\geq} 4x \geq 2x + 2|y|,$$

which finishes the proof in case (i).

(ii) x, y have the same sign. Without loss of generality we may assume $x \geq y \geq 0$. Thus it suffices to show that $2(x - y) \leq N(x) - N(y)$. Note that due to the assumption of Case 3 we have $x \geq x - y \geq 1 \geq \frac{1}{2}$, so by the convexity of N we have

$$\frac{N(x) - N(y)}{x - y} \geq \frac{N(\frac{1}{2}) - N(0)}{\frac{1}{2} - 0} \overset{(3.4)}{\geq} 4 \geq 2$$

This ends the examination of case (ii) and the proof of the theorem. □

3.3 Comparison of weak and strong moments

The goal of this section is to establish the comparison of weak and strong moments with respect to any norm $\|\cdot\|$ for random vectors X with independent coordinates having log-concave tails (Corollary 3.4). In view of Theorem 3.1 and Remark 1.2, it is enough to show Theorem 3.3.

Our proof of Theorem 3.3 comprises three steps: first we exploit α-regularity of moments of X to control the size of its cumulant-generating function Λ_X, second we bound the infimum convolution of the optimal cost function with the convex test function being the norm $\|\cdot\|$ properly rescaled, and finally by the property convex IC(β) we obtain exponential tail bounds which integrated out give the desired moment inequality.

We start with two lemmas corresponding to the first two steps described above and then we put everything together.

Lemma 3.6. *Let $p \geq 2$ and suppose that the moments of a random vector X in $\mathbb{R}^n$ grow α-regularly. If for a vector $u \in \mathbb{R}^n$ we have $\|\langle u, X\rangle\|_p \leq 1$, then*

$$\Lambda_X((2e\alpha)^{-1}pu) \leq p.$$

Proof. Let k_0 be the smallest integer larger than p. If $\alpha e\|\langle u, X\rangle\|_p \leq 1/2$, then by

α-regularity we have

$$\begin{aligned}\Lambda_X(pu) &\le \ln\Big(\sum_{k\ge 0}\frac{\mathbb{E}\,|\langle pu, X\rangle|^k}{k!}\Big) \le \ln\Big(\sum_{0\le k\le p} p^k\frac{\|\langle u, X\rangle\|_p^k}{k!} + \sum_{k>p}(\alpha k)^k\frac{\|\langle u, X\rangle\|_p^k}{k!}\Big)\\ &\le \ln\Big(\sum_{0\le k\le p}\frac{p^k\|\langle u, X\rangle\|_p^k}{k!} + \sum_{k>p}\big(\alpha e\|\langle u, X\rangle\|_p\big)^k\Big)\\ &\le \ln\Big(\sum_{0\le k\le p}\frac{p^k\|\langle u, X\rangle\|_p^k}{k!} + 2(\alpha e\|\langle u, X\rangle\|_p)^{k_0}\Big)\\ &\le \ln\Big(\sum_{0\le k\le p}\frac{p^k\|\langle u, X\rangle\|_p^k}{k!} + \frac{(2\alpha e p\|\langle u, X\rangle\|_p)^{k_0}}{k_0!}\Big)\\ &\le \ln\Big(\sum_{0\le k\le k_0}\frac{(2\alpha e p\|\langle u, X\rangle\|_p)^k}{k!}\Big) \le 2\alpha e p\|\langle u, X\rangle\|_p \le p.\end{aligned}$$

Replace u with $(2e\alpha)^{-1}u$ to get the assertion. □

Lemma 3.7. *Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$ and let X be a random vector with values in $\mathbb{R}^n$ and moments growing α-regularly. For $\beta > 0$, $p \ge 2$, and $x \in \mathbb{R}^n$ we have*

$$\big(\Lambda_X^*\,(\cdot/\beta)\,\square a\|\cdot\|\big)(x) \ge a\|x\| - p,$$

where $a = p(2e\alpha\beta\sigma_{\|\cdot\|,X}(p))^{-1}$.

Proof. For $f(x) = a\|x\|$ with positive a being arbitrary for now we bound the infimum convolution as follows

$$\begin{aligned}\big(\Lambda_X^*(\cdot/\beta)\square f\big)(x) &= \inf_y \sup_z \big\{\beta^{-1}\langle y, z\rangle - \Lambda_X(z) + a\|x-y\|\big\}\\ &= \inf_y \sup_u \big\{(2e\alpha\beta)^{-1}p\langle y, u\rangle - \Lambda_X((2e\alpha)^{-1}pu) + a\|x-y\|\big\}\\ &\ge \inf_y \sup_{u:\|\langle u, X\rangle\|_p\le 1} \big\{(2e\alpha\beta)^{-1}p\langle y, u\rangle - p + a\|x-y\|\big\},\end{aligned}$$

where in the last inequality we have used Lemma 3.6. Choose $u = \sigma_{\|\cdot\|,X}(p)^{-1}v$ with $\|v\|_* \le 1$ such that $\langle y, v\rangle = \|y\|$. Then clearly $\|\langle u, X\rangle\|_p \le 1$ and thus

$$\Lambda_X^*(\cdot/\beta)\square f(x) \ge \inf_y \big\{(2e\alpha\beta\sigma_{\|\cdot\|,X}(p))^{-1}p\|y\| - p + a\|x-y\|\big\}.$$

If we now set $a = p(2e\alpha\beta\sigma_{\|\cdot\|,X}(p))^{-1}$, then by the triangle inequality we obtain the desired lower bound

$$\big(\Lambda_X^*\,(\cdot/\beta)\,\square a\|\cdot\|\big)(x) \ge a\|x\| - p.$$

□

Proof of Theorem 3.3. Let $f(x) = a\|x\|$ with $a = p(2e\alpha\beta\sigma_{\|\cdot\|,X}(p))^{-1}$ as in Lemma 3.7. Testing the property convex IC(β) with f and applying Lemma 3.7 yields

$$\mathbb{E}\, e^{a\|X\|}\, \mathbb{E}\, e^{-a\|X\|} \le e^p.$$

By Jensen's inequality we obtain that both $\mathbb{E}\, e^{a(\|X\|-\mathbb{E}\,\|X\|)}$ and $\mathbb{E}\, e^{a(-\|X\|+\mathbb{E}\,\|X\|)}$ are bounded above by e^p. Thus Markov's inequality implies the tail bound

$$\mathbb{P}\left(a\big|\|X\| - \mathbb{E}\,\|X\|\big| > t\right) \le 2e^{-t}e^p \le 2e^{-t/2}, \quad t \ge 2p.$$

Consequently,

$$\begin{aligned} a^p\, \mathbb{E}\big|\|X\| - \mathbb{E}\,\|X\|\big|^p &= \int_0^\infty pt^{p-1}\mathbb{P}\left(a\big|\|X\| - \mathbb{E}\,\|X\|\big| > t\right) dt \\ &\le (2p)^p + 2\int_0^\infty pt^{p-1}e^{-t/2}dt = (2p)^p + 2\cdot 2^p p\Gamma(p) \\ &\le 2(2p)^p. \end{aligned}$$

Plugging in the value of a gives the result (we can take $C = 4\sqrt{2}e < 16$). □

3.4 An example

Let X be a symmetric random variable defined by $\mathbb{P}(|X| > t) = T(t)$, where

$$T(t) := 1_{[0,2)}(t) + \sum_{k=1}^{\infty} e^{-2^k} 1_{[2^k,2^{k+1})}(t), \quad t \ge 0, \tag{3.9}$$

or, in other words, let $|X|$ have the distribution

$$(1-e^{-2})\delta_2 + \sum_{k=2}^{\infty}\left(e^{-2^{k-1}} - e^{-2^k}\right)\delta_{2^k}.$$

Let us first show that the moments of X grow 3-regularly, but X does not satisfy IC(β) for any $\beta < \infty$ (we also prove a slightly stronger statement later).

Let Y be a symmetric exponential random variable. Then Y has log-concave tails, so the moments of Y grow 1-regularly (see Remark 1.2). Moreover, if X and Y are constructed in the standard way by the inverses of their CDFs on the probability space $(0,1)$, then

$$|Y| \le |X| \le 2|Y| + 2.$$

Therefore, for $p \geq q \geq 2$,

$$\|X\|_p \leq 2\|Y\|_p + 2 \leq 2\frac{p}{q}\|Y\|_q + 2 \leq 3\frac{p}{q}\|X\|_q$$

(we used the fact that $|X| \geq 2$ in the last inequality). Thus the moments of X grow 3-regularly.

On the other hand, for every $h > 0$ there exists $t > 0$ such that

$$\mathbb{P}(|X| \geq t + h) = \mathbb{P}(|X| \geq t).$$

Therefore by [8, Theorem 1] there is no constant C such that the pair $(X, \varphi(\cdot/C))$, where $\varphi(x) = \frac{1}{2}x^2 1_{\{|x|\leq 1\}} + (|x| - 1/2)1_{\{|x|>1\}}$, satisfies the convex infimum convolution inequality. But, by symmetry and the 3-regularity of moments of X,

$$\begin{aligned}\Lambda_X(s) &\leq \ln\Big(1 + \sum_{k\geq 1} \frac{s^{2k}\,\mathbb{E}\,X^{2k}}{(2k)!}\Big) \leq \ln\Big(1 + \sum_{k\geq 1} \frac{s^{2k}(3k)^{2k}\big(\mathbb{E}\,X^2\big)^k}{(2k)!}\Big) \\ &\leq \ln\Big(1 + \sum_{k\geq 1} s^{2k}(3e/2)^{2k}\big(\mathbb{E}\,X^2\big)^k\Big) = \ln\Big(1 + \sum_{k\geq 1}\big(9e^2s^2\,\mathbb{E}\,X^2/4\big)^k\Big).\end{aligned}$$

Thus for some $A, \varepsilon > 0$ we have $\Lambda_X(s) \leq As^2$ for $|s| \leq \varepsilon$ and $2A\varepsilon^2 \geq 1$. Hence

$$\begin{aligned}\Lambda_X^*(t) &\geq \sup_{|s|\leq\varepsilon}\{st - As^2\} = \tfrac{1}{4A}t^2 1_{\{|t|\leq 2A\varepsilon\}} + (\varepsilon|t| - A\varepsilon^2)1_{\{|t|>2A\varepsilon\}} \\ &= 2A\varepsilon^2\varphi\big(t/(2A\varepsilon)\big) \geq \varphi\big(t/(2A\varepsilon)\big).\end{aligned}$$

We conclude that X cannot satisfy IC(β) for any β.

Remark 3.8. Let us also sketch an alternative approach. Take $a, c > 0$, $b \in \mathbb{R}$, and denote $\varphi(x) = \min\{x^2, |x|\}$, $f(x) = f_{a,b}(x) = a(x - b)_+$ for $x \in \mathbb{R}$. One can check that

$$\big(f\Box\varphi(c\cdot)\big)(x) = \begin{cases} 0 & \text{if } x \leq b, \\ c^2(x-b)^2 & \text{if } b < x \leq b + 1/c, \\ c(x-b) & \text{if } x > b + 1/c, \end{cases}$$

if $a > 2c$. It is rather elementary but cumbersome to show that for any $c > 0$ there exist $a > 0$ and $b \in \mathbb{R}$ such that (3.1) is violated by the test function f. We omit the details.

In fact, the above example shows that even a slightly stronger statement is true: for vectors with independent coordinates with α-regular growth of moments the comparison of weak and strong moments of norms does not hold with the constant 1 at the first strong moment. More precisely, let $X_1, X_2, \ldots$ be independent random variables with distribution given by (3.9). We claim that there does *not* exist any $K < \infty$ such that

$$\Big(\mathbb{E}\max_{i\le n}|X_i|^p\Big)^{1/p} \le \mathbb{E}\max_{i\le n}|X_i| + K\sup_{\|t\|_1\le 1}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p} \tag{3.10}$$

holds for every $p \ge 2$ and $n \in \mathbb{N}$ (note that we chose the ℓ^∞-norm as our norm). We shall estimate the three expressions appearing in (3.10).

We have

$$\sup_{\|t\|_1\le 1}\Big(\mathbb{E}\Big|\sum_{i=1}^n t_iX_i\Big|^p\Big)^{1/p} \le \sup_{\|t\|_1\le 1}\sum_{i=1}^n |t_i|\|X_i\|_p = \|X_1\|_p \tag{3.11}$$

(this inequality is in fact an equality). Since the moments of X_1 grow 3-regularly, the last term in (3.10) is bounded by $\widetilde{K}p$ for some $\widetilde{K} < \infty$.

To estimate the remaining two terms we need the following standard fact.

Lemma 3.9. *For independent events $A_1, \ldots, A_n$,*

$$(1-e^{-1})\Big(1\wedge\sum_{i=1}^n \mathbb{P}(A_i)\Big) \le \mathbb{P}\Big(\bigcup_{i=1}^n A_i\Big) \le 1\wedge\sum_{i=1}^n \mathbb{P}(A_i).$$

In particular, for i.i.d. non-negative random variables $Y_1, \ldots, Y_n$,

$$(1-e^{-1})\int_0^\infty \Big[1\wedge n\mathbb{P}(Y_1 > t)\Big]dt \le \mathbb{E}\max_{i\le n} Y_i \le \int_0^\infty \Big[1\wedge n\mathbb{P}(Y_1 > t)\Big]dt.$$

Proof. Since one of the inequalities is just a union-bound (and the second part of the assertion follows from the formula for integration by parts), it suffices to prove the left-hand side inequality of the first part of the assertion. We have

$$1-\mathbb{P}\Big(\bigcup_{i=1}^n A_i\Big) = \mathbb{P}\Big(\bigcap_{i=1}^n A_i^c\Big) = \prod_{i=1}^n \mathbb{P}(A_i^c) = \prod_{i=1}^n \big(1-\mathbb{P}(A_i)\big) \le \exp\Big(-\sum_{i=1}^n \mathbb{P}(A_i)\Big).$$

Thus we are done if $\sum_{i=1}^n \mathbb{P}(A_i) \ge 1$. If on the other hand $\sum_{i=1}^n \mathbb{P}(A_i) < 1$, then

$$\mathbb{P}\Big(\bigcup_{i=1}^n A_i\Big) \ge 1-\exp\Big(-\sum_{i=1}^n \mathbb{P}(A_i)\Big) \ge (1-e^{-1})\sum_{i=1}^n \mathbb{P}(A_i)$$

(we used the convexity of $x \mapsto e^x$). This finishes the proof. □

Fix $m \geq 2$ and let $e^{2^{m-1}} \leq n < e^{2^m}$. Then

$$1 \wedge nT(t) = \begin{cases} 1 & \text{if } 0 < t < 2^m, \\ nT(t) & \text{if } t \geq 2^m. \end{cases}$$

By the above lemma,

$$\begin{aligned} \mathbb{E} \max_{i \leq n} |X_i| &\leq \int_0^{2^m} dt + n \int_{2^m}^{\infty} T(t) dt = 2^m + n \sum_{j=m}^{\infty} e^{-2^j} (2^{j+1} - 2^j) \\ &= 2^m + n \sum_{j=m}^{\infty} e^{-2^j} 2^j \leq 2^m + ne^{-2^m} 2^m \sum_{j=0}^{\infty} (2e^{-2^m})^j = 2^m + \frac{ne^{-2^m} 2^m}{1 - 2e^{-2^m}}. \end{aligned}$$

Set $\theta = \theta(m,n) = ne^{-2^m} \in [e^{-2^{m-1}}, 1)$. Then

$$\mathbb{E} \max_{i \leq n} |X_i| \leq 2^m \Big(1 + \frac{\theta}{1 - 2e^{-2^m}} \Big). \tag{3.12}$$

Similarly,

$$\begin{aligned} \mathbb{E} \max_{i \leq n} |X_i|^p &\geq (1 - e^{-1}) \int_0^{\infty} 1 \wedge T(t^{1/p}) dt \\ &= (1 - e^{-1}) \Big[\int_0^{2^{mp}} dt + n \int_{2^{mp}}^{\infty} T(t^{1/p}) dt \Big] \\ &= (1 - e^{-1}) \Big[2^{mp} + n \sum_{j=m}^{\infty} e^{-2^j} \big(2^{(j+1)p} - 2^{jp} \big) \Big]. \end{aligned}$$

Hence

$$\mathbb{E} \max_{i \leq n} |X_i|^p > (1 - e^{-1}) ne^{-2^m} \big(2^{(m+1)p} - 2^{mp} \big) = (1 - e^{-1}) \theta 2^{mp} (2^p - 1). \tag{3.13}$$

Putting (3.11), (3.12), and (3.13) together, we see that (3.10) would imply

$$(1 - e^{-1})^{1/p} \theta^{1/p} 2^m (2^p - 1)^{1/p} \leq 2^m \Big(1 + \frac{\theta}{1 - 2e^{-2^m}} \Big) + \widetilde{K} p$$

for every $p \geq 2$, $m \geq 2$, and $\theta \in [e^{-2^{m-1}}, 1)$ of the form ne^{-2^m}, $n \in \mathbb{N}$. Take $p = 1/\theta$ and $\theta \sim 1/m$ to get

$$(1 - e^{-1})^{\theta} \theta^{\theta} (2^{1/\theta} - 1)^{\theta} \leq 1 + \frac{\theta}{1 - 2e^{-2^m}} + \frac{\widetilde{K}}{2^m \theta}.$$

Since $\theta \to 0$ and $2^m \theta \to \infty$ as $m \to \infty$ this inequality yields $2 \leq 1$, which is a contradiction. Hence inequality (3.10) cannot hold for all $p \geq 2$ and $n \in \mathbb{N}$.

Chapter 4

Estimates of norms of log-concave matrices

A classical result regarding spectra of random matrices is Wigner's Semicircle Law, which describes the limit of empirical spectral measures of a random matrix with independent centred entries with equal variance. Theorems of this type say nothing about the largest eigenvalue (i.e. the operator norm). However, Seginer proved in [32] that for a random matrix X with i.i.d. symmetric entries $\mathbb{E}\|X\|_{2,2}$[1] is of the same order as the expectation of the maximum Euclidean norm of rows and columns of X. The same holds true for the structured Gaussian matrices (i.e. when $X_{ij} = a_{ij}g_{ij}$ and g_{ij} are i.i.d. standard Gaussian variables), as was recently shown in [25], and up to a logarithmic factor for any X with independent centred entries, see [31]. The advance of the two latest results is that they do not require that the entries of X are equally distributed.

Another upper bound for $\mathbb{E}\|X\|_{2,2}$ also does not require equal distributions but only the independence of entries: by [17] we know that

$$\mathbb{E}\|X\|_{2,2} \lesssim \max_i \sqrt{\sum_j \mathbb{E}X_{ij}^2} + \max_j \sqrt{\sum_i \mathbb{E}X_{ij}^2} + \sqrt[4]{\sum_{i,j} \mathbb{E}X_{ij}^4}.$$

This bound is dimension free, but in some cases is worse than the one from [31].

Upper bounds for the expectation of other operator norms were investigated in [4] in the case of independent centred entries bounded by 1. For $q \geq 2$ and $m \times n$ matrices the authors proved that $\mathbb{E}\|X\|_{2,q} \lesssim \max\{m^{1/q}, \sqrt{n}\}$. In [11] Guédon, Hinrichs, Litvak, and Prochno proved that for a structured Gaussian

[1] Recall that $\|\cdot\|_{p,q}$ stands for the operator norm from ℓ_p to ℓ_q.

matrix $X = (a_{ij}X_{ij})_{i\le m, j\le n}$ and $p,q\ge 2$,

$$\mathbb{E}\|X\|_{p',q} \le C(p,q)\Big[\big(\log m\big)^{1/q}\max_{1\le i\le m}\Big(\sum_{j=1}^{n}|a_{ij}|^p\Big)^{1/p} + \max_{1\le j\le n}\Big(\sum_{i=1}^{m}|a_{ij}|^q\Big)^{1/q} + \big(\log m\big)^{1/q}\mathbb{E}\max_{\substack{1\le i\le m\\ 1\le j\le n}}|X_{ij}|\Big].$$

This estimate is optimal up to logarithmic terms (see Remark 4.2 below). Note that in the case $(p,q)\neq(2,2)$ moment methods fails in estimating $\mathbb{E}\|X\|_{p',q}$ (as they give information only on the spectrum of X).

All the mentioned results require the independence of entries of X. In this chapter we will see how to generalise the main result of [11] to a wide class of random matrices with independent log-concave rows, following the scheme of proof of the original theorem from [11]. Our estimate is optimal (for fixed $p,q\ge 2$) up to a factor depending logarithmically on the dimension. Let us stress that we do not require the rows of X to have independent, but only uncorrelated coordinates (and to be log-concave). We will use results described in the previous chapters of this dissertation.

To make the notation more clear, if $A=(A_{ij})_{i\le m, j\le n}$ is an $m\times n$ matrix, we denote by $A_i\in\mathbb{R}^n$ its i-th row and by $A^{(j)}\in\mathbb{R}^m$ we denote its j-th column.

Theorem 4.1. *Let $m\ge 2$, let $Y_1,\ldots,Y_m$ be i.i.d. isotropic log-concave vectors in $\mathbb{R}^n$, and let $A=(A_{ij})$ be an $m\times n$ (deterministic) matrix. Consider a random matrix X with entries $X_{ij}=A_{ij}Y_{ij}$ for $i\le m, j\le n$, where Y_{ij} is the j'th coordinate of Y_i. Then for every $p,q\ge 2$ we have*

$$\mathbb{E}\|X\|_{p',q} \le C(p,q)\Big[\big(\log m\big)^{1/q}\max_{1\le i\le m}\big\|A_i\big\|_p + \max_{1\le j\le n}\big\|A^{(j)}\big\|_q + \big(\log m\big)^{1/q+1}\mathbb{E}\max_{\substack{1\le i\le m\\ 1\le j\le n}}|X_{ij}|\Big], \tag{4.1}$$

where $C(p,q)$ depends only on p and q.

Remark 4.2. Note that the bound from Theorem 4.1 is optimal up to a constant depending on p,q and logarithmically on the dimension. Indeed, since Y_{ij} is log-concave we have by (1.2) that $\mathbb{E}|Y_{ij}|\ge (2C_1)^{-1}\big(\mathbb{E}Y_{ij}^2\big)^{1/2}=(2C_1)^{-1}$. Hence for every $j\le n$, (we take $u=e_j$, use the unconditionality of $\|\cdot\|_q$ and the Jensen inequality)

$$\mathbb{E}\|X\|_{p',q}=\mathbb{E}\sup_{u\in\ell_{p'}^n}\|Xu\|_q\ge\mathbb{E}\|X^{(j)}\|_q=\mathbb{E}\big\|\big(|Y_{ij}|A_{ij}\big)_i\big\|_q\ge(2C_1)^{-1}\|A^{(j)}\|_q.$$

Since $\|X\|_{p',q} = \|X^T\|_{q',p}$, we also get $\mathbb{E}\|X\|_{p',q} \geq (2C_1)^{-1}\|A_i\|_p$ for all $i \leq m$. Moreover, for all $i \leq m$ and $j \leq n$, (we take $v = e_i$ and $u = e_j \operatorname{sgn} X_{ij}$)

$$\|X\|_{p',q} = \sup_{u\in\ell_{p'}^n}\sup_{v\in\ell_{q'}^n} v^T X u \geq |X_{ij}|.$$

Therefore

$$\mathbb{E}\|X\|_{p',q} \geq (4C_1+1)^{-1}\Big[\max_{1\leq i\leq m}\big\|A_i\big\|_p + \max_{1\leq j\leq n}\big\|A^{(j)}\big\|_q + \mathbb{E}\max_{\substack{1\leq i\leq m\\1\leq j\leq n}}|X_{ij}|\Big],$$

what yields the claim.

The next corollary is a version of Theorem 4.1 in the spirit of the aforementioned results from [32, 25, 31]. It follows directly from (4.1), (1.2), and the Jensen inequality.

Corollary 4.3. *Under the assumptions of Theorem 4.1 we have*

$$\mathbb{E}\|X\|_{p',q} \leq C(p,q)(\log m)^{1+1/q}\bigg(\mathbb{E}\max_{1\leq i\leq m}\Big(\sum_{j=1}^n |A_{ij}Y_{ij}|^p\Big)^{1/p} + \mathbb{E}\max_{1\leq j\leq n}\Big(\sum_{i=1}^m |A_{ij}Y_{ij}|^q\Big)^{1/q}\bigg)$$

Remark 4.4. If the rows and columns of Y are isotropic and log-concave (we do not require independence), and $p, q \geq 1$, then

$$\begin{aligned}\mathbb{E}\max_{1\leq i\leq m}\Big(\sum_{j=1}^n |A_{ij}Y_{ij}|^p\Big)^{1/p} + \mathbb{E}\max_{1\leq j\leq n}\Big(\sum_{i=1}^m |A_{ij}Y_{ij}|^q\Big)^{1/q}\\ \leq C\Big(p^2\max_{1\leq i\leq m}\big\|A_i\big\|_p + q^2\max_{1\leq j\leq n}\big\|A^{(j)}\big\|_q + (p+q)\log(m\vee n)\mathbb{E}\max_{\substack{1\leq i\leq m\\1\leq j\leq n}}\big|A_{ij}Y_{ij}\big|\Big),\end{aligned} \tag{4.2}$$

so the bound we used in the proof of Corollary 4.3 (the one which uses the Jensen inequality) may be reversed up to a logarithmic factor and constants depending on p, q. Inequality (4.2) follows directly from the following proposition.

Proposition 4.5. *Let Z be an $m\times n$ random matrix with isotropic and log-concave rows, let B be a deterministic $m\times n$ matrix, and let $p\geq 1$. Then*

$$\mathbb{E}\max_{1\leq i\leq m}\Big(\sum_{j=1}^n |B_{ij}Z_{ij}|^p\Big)^{1/p} \lesssim p^2\max_{1\leq i\leq m}\Big(\sum_{j=1}^n |B_{ij}|^p\Big)^{1/p} + p\log(m\vee n)\mathbb{E}\max_{\substack{1\leq i\leq m\\1\leq j\leq n}}|B_{ij}Z_{ij}|.$$

The rest of this chapter is organised as follows. Section 4.1 contains generalisations of Lemmas 3.1 and 3.2 from [11] to the log-concave setting and the proof of Theorem 4.1. In Section 4.2 we will show how to deduce an analogue of Theorem 4.1 for Gaussian mixtures (see Corollary 4.12) and provide a proof of Proposition 4.5.

4.1 Proof of Theorem 4.1

In the proof of Theorem 4.1 we will use Theorem 2.1 from [11], which is another version of results provided before by Guédon–Rudelson in [14], and by Guédon–Mendelson–Pajor–Tomczak-Jaegerman in [12]. Below we use a slightly different notation than in [11].

Theorem 4.6 ([11, Theorem 2.1])**.** *Let E be a Banach space with modulus of convexity of power type 2 with constant λ. Let $X_1, \dots X_m \in E^*$ be independent random vectors, $q \geq 2$. Define*

$$u := \sup_{t \in B_E} \Big(\sum_{i=1}^{m} \mathbb{E} \big| X_i(t) \big|^q \Big)^{1/q}, \tag{4.3}$$

and

$$v := \Big(\lambda^8 \big(T_2(E^*) \big)^2 \log m \, \mathbb{E} \max_{1 \leq i \leq m} \|X_i\|_{E^*}^q \Big)^{1/q}, \tag{4.4}$$

where $T_2(E^*)$ is the Rademacher type 2 constant of E^*. Then

$$\Big[\mathbb{E} \sup_{t \in B_E} \Big| \sum_{i=1}^{m} \Big(\big| X_i(t) \big|^q - \mathbb{E} \big| X_i(t) \big|^q \Big) \Big| \Big]^{1/q} \leq C(\sqrt{uv} + v) \leq 2C(u+v).$$

The next two lemmas provide estimates on quantities u and v appearing in Theorem 4.6 in the case $E = B_{p'}^n$.

Lemma 4.7. *Assume p, q, X, and Y are as in Theorem 4.1. Then*

$$\Big(\mathbb{E} \max_{1 \leq i \leq m} \|X_i\|_p^q \Big)^{1/q} \leq C(p,q) \Big[\max_{1 \leq i \leq m} \|A_i\|_p + \log m \; \mathbb{E} \max_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} |X_{ij}| \Big],$$

where $C(p,q)$ depends only on p and q.

Lemma 4.8. *Assume p, q, X, and Y are as in Theorem 4.1. Then*

$$\sup_{t \in B_{p'}^n} \Big(\sum_{i=1}^{m} \mathbb{E} \big| \langle X_i, t \rangle \big|^q \Big)^{1/q} \leq C_1 q \max_{1 \leq j \leq n} \big\| A^{(j)} \big\|_q. \tag{4.5}$$

In the proof of Lemma 4.7 we will also need the following estimate:

Lemma 4.9. *Assume that Z is an isotropic log-concave vector in $\mathbb{R}^m$. Then for all $1 \leq k \leq m$ and all $a \in \mathbb{R}^m$ we have*

$$\mathbb{E} \max_{1 \leq i \leq m} |a_i Z_i| \geq D_3^{-1} \max_{k \leq m} \big(a_k^* \min_{i \leq m} \|Z_i\|_{\log(k+1)} \big),$$

where D_3 is an absolute constant.

In order to prove Theorem 4.1, we repeat the proof scheme from [11].

Proof of Theorem 4.1. We use Theorem 4.6 for $E = \ell_{p'}^n$. Then $\lambda \sim p$ (see [30, Theorem 5.3]) and $T_2(E^*) \sim \sqrt{p}$. Let u and v be given by formulas (4.3) and (4.4). The triangle inequality, Theorem 4.6, Lemma 4.8, and Lemma 4.7 yield

$$\begin{aligned}
\mathbb{E}\|X\|_{p',q} &\le \big(\mathbb{E}\|X\|_{p',q}^q\big)^{1/q} = \Big[\mathbb{E}\sup_{t\in B_{p'}^n}\sum_{i=1}^m \big|\langle t, X_i\rangle\big|^q\Big]^{1/q} \\
&\le \Big[\mathbb{E}\sup_{t\in B_{p'}^n}\Big|\sum_{i=1}^m\Big(\big|\langle X_i,t\rangle\big|^q - \mathbb{E}\big|\langle X_i,t\rangle\big|^q\Big)\Big|\Big]^{1/q} + \sup_{t\in B_{p'}^n}\Big(\mathbb{E}\sum_{i=1}^m\big|\langle t, X_i\rangle\big|^q\Big)^{1/q} \\
&\le C\cdot(u+v) \\
&\le C(p,q)\Big[\big(\log m\big)^{1/q}\max_{1\le i\le m}\big\|A_i\big\|_p + \max_{1\le j\le n}\big\|A^{(j)}\big\|_q + \big(\log m\big)^{1/q+1}\mathbb{E}\max_{\substack{1\le i\le m\\1\le j\le n}}|X_{ij}|\Big].
\end{aligned}$$

□

The main contribution of this chapter lies in the proofs of Lemmas 4.7, 4.8, and 4.9.

Proof of Lemma 4.9. We may and do assume that $a_1 \ge a_2 \ge \ldots \ge a_m \ge 0$, i.e. $a_i^* = a_i$ for $i \le m$. By [21, Proposition 3.3] we have for all $k \le m$,

$$\mathbb{E}\max_{1\le i\le k}|a_iZ_i| \ge C^{-1}\min_{1\le i\le k}\|a_iZ_i\|_{\log(k+1)} \ge C^{-1}a_k\min_{1\le i\le m}\|Z_i\|_{\log(k+1)}.$$

Thus

$$\mathbb{E}\max_{1\le i\le m}|a_iZ_i| = \max_{1\le k\le m}\mathbb{E}\max_{1\le i\le k}|a_iZ_i| \ge C^{-1}\max_{1\le k\le m}\big(a_k\min_{1\le i\le m}\|Z_i\|_{\log(k+1)}\big).$$ □

Proof of Lemma 4.7. If $m = 1$ then the assertion follows by (1.2). From now on we assume that $m \ge 2$.

Since we may approximate A_{ij} by nonzero numbers, we may and do assume that $a_{ij} \ne 0$ for all i, j. Let D_1, D_2 be the constants from (2.41) applied with $r = p$, let D_3 be the constant from Lemma 4.9, and recall that C_1 is the constant from (1.2). We may assume that all these constants are greater than 1.

Note that for any $a, b \in \mathbb{R}$ we have $a = (a-b)_+ + a\wedge b$. Thus, by the triangle inequality,

$$\begin{aligned}
&\Big(\mathbb{E}\max_{1\le i\le m}\|X_i\|_p^q\Big)^{1/q} \\
&\le \Big(\mathbb{E}\max_{1\le i\le m}\Big[\big(\|X_i\|_p - D_1p\mathbb{E}\|X_i\|_p\big)^q\mathbf{1}_{\{\|X_i\|_p\ge D_1p\mathbb{E}\|X_i\|_p\}}\Big]\Big)^{1/q} + D_1p\max_{1\le i\le m}\mathbb{E}\|X_i\|_p.
\end{aligned} \tag{4.6}$$

Moreover, for every $1 \le i \le m$ we have by (1.2) and the isotropicity of Y_i, that

$$\begin{aligned}\mathbb{E}\|X_i\|_p &\le \Big(\sum_{j=1}^n \mathbb{E}|Y_{ij}|^p|A_{ij}|^p\Big)^{1/p} \le \max_{j\le n}\|Y_{ij}\|_p\|A_i\|_p \le C_1 p\|A_i\|_p \\ &\le C_1 p \max_{1\le k\le m}\|A_k\|_p. \end{aligned} \tag{4.7}$$

Now we pass to the estimation of the fist term of (4.6). Let

$$B := C_1^2 D_3 \log(m+1)\ \mathbb{E}\max_{\substack{1\le i\le m\\ 1\le j\le n}}|X_{ij}| \qquad \text{and} \qquad \sigma := (\max_{1\le i\le m}\sigma_{\|\cdot\|_p,X_i}(2))\vee B.$$

By (2.41) we have

$$\begin{aligned}&\mathbb{E}\max_{1\le i\le m}\Big[\big(\|X_i\|_p - D_1 p\mathbb{E}\|X_i\|_p\big)^q\mathbf{1}_{\{\|X_i\|_p\ge D_1p\mathbb{E}\|X_i\|_p\}}\Big] \\ &\quad\le (2D_1pe\sigma)^q + \int_{2D_1pe\sigma}^\infty qu^{q-1}\mathbb{P}\Big(\max_{1\le i\le m}\big(\|X_i\|_p - D_1p\mathbb{E}\|X_i\|_p\big)\ge u\Big)du \\ &\quad\le (2D_1pe\sigma)^q + (D_1p)^q\sum_{i=1}^m\int_{2e\sigma}^\infty qu^{q-1}\mathbb{P}\big(\|X_i\|_p - D_1p\mathbb{E}\|X_i\|_p\ge D_1pu\big)du \\ &\quad\le (2D_1pe\sigma)^q + (D_1p)^qD_2\sum_{i=1}^m\int_{2e\sigma}^\infty qu^{q-1}\sup_{\|t\|_{p'}\le 1}\mathbb{P}\Big(\Big|\sum_{j=1}^n t_jX_{ij}\Big|\ge u\Big)du.\end{aligned} \tag{4.8}$$

For $u \ge \sup_{\|t\|_{p'}\le 1}\|\sum_{j=1}^n t_jX_{ij}\|_\infty$ the function we integrate vanishes, so from now on we will consider only i's for which $u < \sup_{\|t\|_{p'}\le 1}\|\sum_{j=1}^n t_jX_{ij}\|_\infty$.

Note that if $1\le i\le m$ and $\sup_{\|t\|_{p'}\le 1}\|\sum_{j=1}^n t_jX_{ij}\|_\infty > u \ge e\sigma \ge e\sigma_{\|\cdot\|_p,X_i}(2)$, then Remark 1.3 implies that $r := r(i) := \sup\{s\ge 2 : \sigma_{\|\cdot\|_p,X_i}(s)\le u/e\}\in[2,\infty)$ and $\sigma_{\|\cdot\|_p,X_i}(r) = u/e$. Therefore

$$\sup_{\|t\|_{p'}\le 1}\mathbb{P}\Big(\Big|\sum_{j=1}^n t_jX_{ij}\Big|\ge u\Big) \le \frac{\sup_{\|t\|_{p'}\le 1}\|\langle t, X_i\rangle\|_r^r}{u^r} = e^{-r}. \tag{4.9}$$

Now we will estimate r from below. For $t\ge 2$ let

$$\varphi(t) = t\min_{1\le j\le n}\|Y_{ij}\|_t.$$

Since $Y_i's$ are identically distributed, φ does not depend on i. By (1.2), and the isotropicity of Y we have

$$\sigma_{\|\cdot\|_p,X_i}(t) \le \sigma_{X_i}(t) \le C_1 t \max_{|x|\le 1}\Big(\mathbb{E}\Big(\sum_{j=1}^n A_{ij}Y_{ij}x_j\Big)^2\Big)^{1/2}$$
$$= C_1 t \max_{|x|\le 1}\Big(\mathbb{E}\Big(\sum_{j=1}^n A_{ij}^2 x_j^2\Big)^2\Big)^{1/2}$$
$$= C_1 t \max_{1\le j\le n}|A_{ij}|\cdot\|Y_{ij}\|_2 \le C_1\varphi(t)\max_{1\le j\le n}|A_{ij}|. \tag{4.10}$$

Since we can permute the rows of A, we may and do assume that

$$\max_{1\le j\le n}|A_{1j}| \ge \ldots \ge \max_{1\le j\le n}|A_{mj}|.$$

Let $j(i)\le n$ be such an index that $|A_{ij(i)}| = \max_{1\le j\le n}|A_{ij}|$. Lemma 4.9 applied to $Z_i = Y_{ij(i)}$ and a non-increasing sequence $a_i = |A_{ij(i)}|$ implies

$$\mathbb{E}\max_{\substack{1\le i\le m\\1\le j\le n}}|X_{ij}| \ge \mathbb{E}\max_{1\le i\le m}|A_{ij(i)}Y_{ij(i)}| \ge D_3^{-1}\big(\log(m+1)\big)^{-1}\max_{1\le i\le m}\Big(\varphi\big(\log(i+1)\big)|A_{ij(i)}|\Big),$$

so for all $i\le m$ we have

$$B \ge C_1^2\varphi(\log(i+1))|A_{ij(i)}| = C_1^2\varphi(\log(i+1))\max_{1\le j\le n}|A_{ij}|.$$

Note that by (1.2) for all $r\ge\lambda\ge 2$ we have $\sigma_{\|\cdot\|_p,X_i}(r/\lambda)\ge\sigma_{\|\cdot\|_p,X_i}(r)/(C_1\lambda)$. Take $\lambda = \sigma_{\|\cdot\|_p,X_i}(r)/B = u/(Be)\ge 2$. Then by a calculation similar to the one above we get

$$\frac{u}{e} = \sigma_{\|\cdot\|_p,X_i}(r) \le \frac{C_1 r}{2}\max_{1\le j\le n}|A_{ij}| \le C_1^2 r\max_{\substack{1\le i\le m\\1\le j\le n}}|A_{ij}|\mathbb{E}|Y_{ij}| \le C_1^2 r\mathbb{E}\max_{\substack{1\le i\le m\\1\le j\le n}}|X_{ij}| \le Br,$$

so indeed $r\ge\lambda\ge 2$.

Therefore for all $i\le m$ we have

$$\frac{B}{C_1} = \frac{1}{\lambda C_1}\sigma_{\|\cdot\|_p,X_i}(r) \le \sigma_{\|\cdot\|_p,X_i}(r/\lambda) \overset{(4.10)}{\le} C_1\varphi\Big(\frac{r}{\lambda}\Big)\max_{1\le j\le n}|A_{ij}| \le \frac{B\varphi(\frac{r}{\lambda})}{C_1\varphi(\log(i+1))}. \tag{4.11}$$

Since the function φ is strictly increasing, the previous inequality yields $r\ge\lambda\log(i+1)$. This together with (4.9) implies that (recall that $\lambda = \frac{u}{Be}\ge 2$)

$$\sum_{i=1}^m \sup_{\|t\|_{p'}\le 1}\mathbb{P}\Big(\Big|\sum_{j=1}^n t_j X_{ij}\Big| \ge u\Big) \le \sum_{i=1}^m (i+1)^{-\frac{u}{eB}} \le 2^{-\frac{u}{eB}} + \int_2^\infty x^{-\frac{u}{eB}}dx \le 3\cdot 2^{-\frac{u}{e\sigma}}. \tag{4.12}$$

Inequalities (4.8), (4.12), and the Stirling formula yield that

$$\left(\mathbb{E}\Big[\max_{1\le i\le m}\big(\|X_i\|_p - D_1\mathbb{E}\|X_i\|_p\big)^q \mathbf{1}_{\{\|X_i\|_p\ge D_1\mathbb{E}\|X_i\|_p\}}\Big]\right)^{1/q} \le CD_1D_2^{1/q}\sigma pq. \quad (4.13)$$

Moreover, by (1.2)

$$\max_{1\le i\le m}\sigma_{\|\cdot\|_p,X_i}(2) \le 2C_1\max_{1\le i\le m}\sigma_{\|\cdot\|_p,X_i}(1) \le 2C_1\max_{1\le i\le m}\mathbb{E}\|X_i\|_p,$$

where the second inequality holds since the weak first moment is bounded above by the strong first moment. This together with (4.6), (4.7), and (4.13) gives the assertion. □

Proof of Lemma 4.8. Note that if $0\le r\le s$, then for every $x\in\mathbb{R}^n$ we have $\|x\|_s\le\|x\|_r$, so we may and do assume $p=2$. By (1.2), the isotropicity of Y, and the Jensen inequality we have

$$\begin{aligned}\sup_{t\in B_2^n}\Big(\sum_{i=1}^m\mathbb{E}\big|\langle X_i,t\rangle\big|^q\Big)^{1/q} &\le C_1q\sup_{\|t\|_2\le 1}\Big(\sum_{i=1}^m\Big(\mathbb{E}\big|\langle X_i,t\rangle\big|^2\Big)^{q/2}\Big)^{1/q}\\ &= C_1q\sup_{\|t\|_2=1}\Big(\sum_{i=1}^m\Big(\sum_{j=1}^n A_{ij}^2t_j^2\Big)^{q/2}\Big)^{1/q}\\ &\le C_1q\sup_{\|t\|_2=1}\Big(\sum_{i=1}^m\sum_{j=1}^n|A_{ij}|^qt_j^2\Big)^{1/q}\\ &= C_1q\Big(\sup_{\|t\|_2=1}\sum_{j=1}^n\big\|A^{(j)}\big\|_q^qt_j^2\Big)^{1/q}\\ &= C_1q\max_{1\le j\le n}\big\|A^{(j)}\big\|_q.\end{aligned}$$

□

Remark 4.10. By the same reasoning as in the log-concave case, we may prove (using Corollary 2.11, [24, Theorem 2.1], and the claim below instead of (2.41), Lemma 4.9 and the previous estimates on $\sigma_{\|\cdot\|_{p'},X_i}(s)$, respectively) the following.

Let X be an $m\times n$ random matrix with entries $X_{ij}=a_{ij}Y_{ij}$, where Y_{ij} are independent symmetric random variables such that $\mathbb{E}Y_{ij}^2=1$. Assume that for any $r\ge 2$ and any $1\le i\le m$, $1\le j\le n$ we have $\frac{r^\gamma}{\beta}\le\|Y_{ij}\|_r\le\beta r^\gamma$ with $\gamma\in[\frac12,1]$. Then for every $p,q\ge 2$ we have

$$\begin{aligned}&\mathbb{E}\|X\|_{p',q}\\ &\quad\le C(p,q,\beta)\Big[\big(\log m\big)^{1/q}\max_{1\le i\le m}\big\|A_i\big\|_p+\max_{1\le j\le n}\big\|A^{(j)}\big\|_q+\big(\log m\big)^{1/q}\mathbb{E}\max_{\substack{1\le i\le m\\1\le j\le n}}|X_{ij}|\Big].\end{aligned}$$

where $C(p,q,\beta)$ depends only on p, q, and β.

As we mentioned, it suffices to prove the claim:

$$\left\|\sum_{j=1}^n t_jY_{ij}\right\|_r \le C\beta r^\gamma\left\|\sum_{j=1}^n t_jY_{ij}\right\|_2 = C\beta r^\gamma\|t\|_2, \tag{4.14}$$

where C is an absolute constant, and repeat the proof of Theorem 4.1.

Proof of the claim. It suffices to consider $r=2k$, where k is an integer. Let us denote

$$c_{i_1,\dots i_n} := \binom{i_1+\ldots+i_n}{i_1}\binom{i_2+\ldots+i_n}{i_2}\cdots\binom{i_n}{i_n}.$$

Let $G=(G_j)_{j=1}^n$ be the standard n-dimensional Gaussian vector. Recall that for any $t\in\mathbb{R}^n$ and $r\ge 1$ we have $\|\sum_{j=1}^n t_jG_j\|_r = \|t\|_2\|G_1\|_r \sim \|t\|_2\sqrt{r} = \sqrt{r}\|\sum_{j=1}^n t_jY_{ij}\|_2$.

By the assumptions on Y_i and by the fact that $\gamma\ge\frac12$ we get

$$\begin{aligned}
\left\|\sum_{j=1}^n t_jY_{ij}\right\|_{2k}^{2k} &= \sum_{j_1+\ldots+j_n=k} c_{2j_1,\dots 2j_n}\mathbb{E}Y_{i1}^{2j_1}\cdots\mathbb{E}Y_{in}^{2j_n}t_1^{2j_1}\cdots t_n^{2j_n}\\
&\le \beta^{2k}\sum_{j_1+\ldots+j_n=k} c_{2j_1,\dots 2j_n}(2j_1)^{2j_1\gamma}\cdots(2j_n)^{2j_n\gamma}t_1^{2j_1}\cdots t_n^{2j_n}\\
&\le (2k)^{2k\gamma-k}\beta^{2k}\sum_{j_1+\ldots+j_n=k} c_{2j_1,\dots 2j_n}(2j_1)^{j_1}\cdots(2j_n)^{j_n}t_1^{2j_1}\cdots t_n^{2j_n}\\
&\le (2k)^{2k\gamma-k}(C\beta)^{2k}\sum_{j_1+\ldots+j_n=k} c_{2j_1,\dots 2j_n}\mathbb{E}G_1^{2j_1}\cdots\mathbb{E}G_n^{2j_n}t_1^{2j_1}\cdots t_n^{2j_n}\\
&= (2k)^{2k\gamma-k}(C\beta)^{2k}\left\|\sum_{j=1}^n t_jG_j\right\|_{2k}^{2k} \le (2k)^{2k\gamma}(C\beta)^{2k}\left\|\sum_{j=1}^n t_jY_{ij}\right\|_2^{2k},
\end{aligned}$$

what finishes the proof of (4.14). □

By the claim we get

$$\sigma_{\|\cdot\|_p,cY_i}(q) \le C\beta q^\gamma \sup_{s\in B_{p*}^n}\sqrt{\sum_{j=1}^n s_j^2c_j^2} = C\beta q^\gamma\max_{1\le j\le n}|c_j| \le C\beta^2\min_{j\le n}\|Y_{ij}\|_q\max_{1\le j\le n}|c_j|,$$

what allows us to obtain a version of (4.11) for $\varphi(t) := \min_{\substack{1\le i\le m,\\ 1\le j\le n}}\|Y_{ij}\|_t$.

4.2 Estimates of norms of matrices in the case of Gaussian mixtures

Let us recall the definition from [7], where the significance of Gaussian mixtures is also described.

Definition 4.11. A random variable X is called a (centred) Gaussian mixture if there exists a positive random variable r and a standard Gaussian random variable g, independent of r , such that X has the same distribution as the product rg.

We will work with a matrix $(R_{ij}A_{ij}G_{ij})_{i\le m, j\le n}$ which entries are Gaussian mixtures. We additionally assume that a random vector in $\mathbb{R}_+^{nm}$ which coordinates are the entries of $R = (R_{ij})$ is log-concave and isotropic. It will be clear from the proof, that the corollary below is true also for another type of matrices: $(r_i A_{ij} G_{ij})_{i\le m, j\le n}$, where $(r_1, \ldots, r_m)$ is an isotropic log-concave random vector.

Corollary 4.12. *Let $m, n \ge 2$, and let $G = (G_{ij})_{i\le m, j\le n}$ be a matrix which entries are i.i.d. standard Gaussian variables. Let $X_{ij} = R_{ij}B_{ij}G_{ij}$, where R is a log-concave and isotropic random matrix independent of G. Then for every $p, q \ge 2$ we have*

$$\mathbb{E}\|X\|_{p',q} \le C(p,q)\bigg(\big(\log m\big)^{1/q+1} \Big[\max_{1\le i\le m} \big\|B_i\big\|_p + \mathbb{E} \max_{\substack{1\le i\le m\\1\le j\le n}} |X_{ij}| \Big] + \log n \max_{1\le j\le n} \big\|B^{(j)}\big\|_q \bigg).$$

Proof. Theorem 4.1 applied to $Y = G$ and $A_{ij} = R_{ij}B_{ij}$ yields

$$\mathbb{E}\|X\|_{p',q} \le C(p,q)\Big[\big(\log m\big)^{1/q} \mathbb{E} \max_{1\le i\le m} \big\|(B_{ij}R_{ij})_j\big\|_p + \mathbb{E} \max_{1\le j\le n} \big\|(B_{ij}R_{ij})_i\big\|_q + \big(\log m\big)^{1/q+1} \mathbb{E} \max_{\substack{1\le i\le m\\1\le j\le n}} |X_{ij}| \Big],$$

so it suffices to prove that

$$\mathbb{E} \max_{1\le i\le m} \big\|(B_{ij}R_{ij})_j\big\|_p \le C(p) \log m \max_{1\le i\le m} \big\|B_i\big\|_p \tag{4.15}$$

and

$$\mathbb{E} \max_{1\le j\le n} \big\|(B_{ij}R_{ij})_i\big\|_q \le C(q) \log n \max_{1\le j\le n} \big\|B^{(j)}\big\|_q.$$

By the symmetry of assumptions we need only to show (4.15).

Note that for any $u \geq 1$ we have

$$
\begin{aligned}
\mathbb{E} \max_{1\leq i\leq m} \Big(\sum_{j=1}^n |B_{ij}|^p R_{ij}^p\Big)^{1/p} &\leq \bigg(\mathbb{E} \max_{1\leq i\leq m} \Big(\sum_{j=1}^n |B_{ij}|^p R_{ij}^p\Big)^{u/p}\bigg)^{1/u} \\
&\leq \bigg(\mathbb{E} \sum_{i=1}^m \Big(\sum_{j=1}^n |B_{ij}|^p R_{ij}^p\Big)^{u/p}\bigg)^{1/u} \\
&\leq m^{1/u} \max_{1\leq i\leq m} \bigg(\mathbb{E}\Big(\sum_{j=1}^n |B_{ij}|^p R_{ij}^p\Big)^{u/p}\bigg)^{1/u}. \qquad (4.16)
\end{aligned}
$$

Fix $i \leq m$. By Theorem 2.5 applied to $p = u$, $r = p$, and $X_j = B_{ij}R_{ij}$ we have

$$
(Cp)^{-1}\bigg(\mathbb{E}\Big(\sum_{j=1}^n |B_{ij}|^p R_{ij}^p\Big)^{u/p}\bigg)^{1/u} \leq \mathbb{E}\Big(\sum_{j=1}^n |B_{ij}|^p R_{ij}^p\Big)^{1/p} + \sup_{t\in B_{p'}^n} \Big\| \sum_{j=1}^n B_{ij}R_{ij}t_j\Big\|_u. \qquad (4.17)
$$

Let us use (1.2) and the isotropicity of R_i to estimate the first term in (4.17):

$$
\mathbb{E}\Big(\sum_{j=1}^n |B_{ij}|^p R_{ij}^p\Big)^{1/p} \leq \Big(\sum_{j=1}^n |B_{ij}|^p \mathbb{E} R_{ij}^p\Big)^{1/p} \leq C_1 p \|B_i\|_p. \qquad (4.18)
$$

Recall that $B_{p'}^n \subset B_2^n$. We use again (1.2) and the isotropicity of r_i to estimate the second term in (4.17):

$$
\begin{aligned}
\sup_{t\in B_{p'}^n} \Big\| \sum_{j=1}^n B_{ij}R_{ij}t_j\Big\|_u &\leq C_1 u \sup_{t\in B_2^n} \Big\| \sum_{j=1}^n B_{ij}R_{ij}t_j\Big\|_2 = C_1 u \sup_{t\in B_2^n} \Big(\sum_{j=1}^n B_{ij}^2 t_j^2\Big)^{1/2} \\
&= C_1 u \max_{1\leq j\leq n} |B_{ij}| \leq C_1 u \|B_i\|_p. \qquad (4.19)
\end{aligned}
$$

Take $u = \log m$ and put together (4.16)-(4.18) to get the assertion. □

Remark 4.13. Using [11, Theorem 1.1] instead of Theorem 4.1 in the proof above yields a slightly better estimate:

$$
\begin{aligned}
\mathbb{E}\|X\|_{p',q} \leq C(p,q)\bigg[\big(\log m\big)^{1/q+1} \max_{1\leq i\leq m} \big\|B_i\big\|_p + (\log m)^{1/q} \mathbb{E} \max_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} |X_{ij}| \\
+ \log n \max_{1\leq j\leq n} \big\|B^{(j)}\big\|_q\bigg].
\end{aligned}
$$

Proof of Proposition 4.5. We begin similarly as in the proof of (4.15), but we only estimate the second term on the right hand-side of (4.17) in a slightly different way, using (1.2). Let us repeat the whole proof for the Reader's convenience.

For any $u \geq 1$ we have

$$\mathbb{E}\max_{1\le i\le m}\Big(\sum_{j=1}^n |B_{ij}|^p Z_{ij}^p\Big)^{1/p} \le \Big(\mathbb{E}\max_{1\le i\le m}\Big(\sum_{j=1}^n |B_{ij}|^p R_{ij}^p\Big)^{u/p}\Big)^{1/u}$$
$$\le \Big(\mathbb{E}\sum_{i=1}^m\Big(\sum_{j=1}^n |B_{ij}|^p Z_{ij}^p\Big)^{u/p}\Big)^{1/u}$$
$$\le m^{1/u}\max_{1\le i\le m}\Big(\mathbb{E}\Big(\sum_{j=1}^n |B_{ij}|^p Z_{ij}^p\Big)^{u/p}\Big)^{1/u}. \tag{4.20}$$

Fix $i \le m$. By Theorem 2.5 applied to $p = u$, $r = p$, and $X_j = B_{ij}Z_{ij}$ we have

$$(Cp)^{-1}\Big(\mathbb{E}\Big(\sum_{j=1}^n |B_{ij}|^p Z_{ij}^p\Big)^{u/p}\Big)^{1/u} \le \mathbb{E}\Big(\sum_{j=1}^n |B_{ij}|^p Z_{ij}^p\Big)^{1/p} + \sup_{t\in B_{p'}^n}\Big\|\sum_{j=1}^n B_{ij}Z_{ij}t_j\Big\|_u. \tag{4.21}$$

Let us use (1.2) and the isotropicity of Z_i to estimate the first term in (4.21):

$$\mathbb{E}\Big(\sum_{j=1}^n |B_{ij}|^p Z_{ij}^p\Big)^{1/p} \le \Big(\sum_{j=1}^n |B_{ij}|^p \mathbb{E} Z_{ij}^p\Big)^{1/p} \le C_1 p\|B_i\|_p. \tag{4.22}$$

Recall that $B_{p'}^n \subset B_2^n$. We use again (1.2) and the isotropicity of r_i to estimate the second term in (4.21):

$$\sup_{t\in B_{p'}^n}\Big\|\sum_{j=1}^n B_{ij}Z_{ij}t_j\Big\|_u \le n^{1/u}\sup_{t\in B_{p'}^n}\big(\mathbb{E}\max_{1\le j\le n}|t_j B_{ij}Z_{ij}|^u\big)^{1/u} \le n^{1/u}C_1 u\mathbb{E}\max_{1\le j\le n}|B_{ij}Z_{ij}|.$$

We take $u = \log(m \vee n)$ to get the assertion. □

Bibliography

[1] R. Adamczak, R. Latała, A. E. Litvak, K. Oleszkiewicz, A. Pajor, and N. Tomczak-Jaegermann, *A short proof of Paouris' inequality*, Canad. Math. Bull. **57** (2014), no. 1, 3–8. MR 3150710

[2] S. Artstein-Avidan, B. Klartag, and V. Milman, *The Santaló point of a function, and a functional form of the Santaló inequality*, Mathematika **51** (2004), no. 1-2, 33–48 (2005). MR 2220210

[3] K. Ball, *Logarithmically concave functions and sections of convex sets in* $\mathbf{R}^n$, Studia Math. **88** (1988), no. 1, 69–84. MR 932007

[4] G. Bennett, V. Goodman, and C. M. Newman, *Norms of random matrices*, Pacific J. Math. **59** (1975), no. 2, 359–365. MR 0393085

[5] C. Borell, *Convex measures on locally convex spaces*, Ark. Mat. **12** (1974), 239–252. MR 0388475

[6] S. Brazitikos, A. Giannopoulos, P. Valettas, and B.H. Vritsiou, *Geometry of isotropic convex bodies*, Mathematical Surveys and Monographs, vol. 196, American Mathematical Society, Providence, RI, 2014. MR 3185453

[7] A. Eskenazis, P. Nayar, and T. Tkocz, *Gaussian mixtures: entropy and geometric inequalities*, Ann. Probab. **46** (2018), no. 5, 2908–2945. MR 3846841

[8] N. Feldheim, A. Marsiglietti, P. Nayar, and J. Wang, *A note on the convex infimum convolution inequality*, Bernoulli **24** (2018), no. 1, 257–270. MR 3706756

[9] N. Gozlan, C. Roberto, P.-M. Samson, Yan S., and P. Tetali, *Characterization of a class of weak transport-entropy inequalities on the line*, Ann. Inst. Henri Poincaré Probab. Stat. **54** (2018), no. 3, 1667–1693. MR 3825894

[10] N. Gozlan, C. Roberto, P.-M. Samson, and P. Tetali, *Kantorovich duality for general transport costs and applications*, J. Funct. Anal. **273** (2017), no. 11, 3327–3405. MR 3706606

[11] O. Guédon, A. Hinrichs, A.E. Litvak, and J. Prochno, *On the expectation of operator norms of random matrices*, Geometric aspects of functional analysis, Lecture Notes in Math., vol. 2169, Springer, Cham, 2017, pp. 151–162. MR 3645120

[12] O. Guédon, S. Mendelson, A. Pajor, and N. Tomczak-Jaegermann, *Majorizing measures and proportional subsets of bounded orthonormal systems*, Rev. Mat. Iberoam. **24** (2008), no. 3, 1075–1095. MR 2490210

[13] O. Guédon, P. Nayar, and T. Tkocz, *Concentration inequalities and geometry of convex bodies*, Analytical and probabilistic methods in the geometry of convex bodies, IMPAN Lect. Notes, vol. 2, Polish Acad. Sci. Inst. Math., Warsaw, 2014, pp. 9–86. MR 3329056

[14] O. Guédon and M. Rudelson, *L_p-moments of random vectors via majorizing measures*, Adv. Math. **208** (2007), no. 2, 798–823. MR 2304336

[15] P. Hitczenko, *Domination inequality for martingale transforms of a Rademacher sequence*, Israel J. Math. **84** (1993), no. 1-2, 161–178. MR 1244666

[16] S. Kwapień and W.A. Woyczyński, *Random series and stochastic integrals: single and multiple*, Probability and its Applications, Birkhäuser Boston, Inc., Boston, MA, 1992. MR 1167198

[17] R. Latała, *Some estimates of norms of random matrices*, Proc. Amer. Math. Soc. **133** (2005), no. 5, 1273–1282. MR 2111932

[18] ______, *Order statistics and concentration of l_r norms for log-concave vectors*, J. Funct. Anal. **261** (2011), no. 3, 681–696. MR 2799576

[19] ______, *Weak and strong moments of random vectors*, Marcinkiewicz centenary volume, Banach Center Publ., vol. 95, Polish Acad. Sci. Inst. Math., Warsaw, 2011, pp. 115–121. MR 2918093

[20] ______, *Modified Paouris inequality*, Geometric aspects of functional analysis, Lecture Notes in Math., vol. 2116, Springer, Cham, 2014, pp. 293–307. MR 3364693

[21] ______, *Sudakov-type minoration for log-concave vectors*, Studia Math. **223** (2014), no. 3, 251–274. MR 3274967

[22] R. Latała and M. Strzelecka, *Weak and strong moments of ℓ_r-norms of log-concave vectors*, Proc. Amer. Math. Soc. **144** (2016), no. 8, 3597–3608. MR 3503729

[23] ———, *Comparison of weak and strong moments for vectors with independent coordinates*, Mathematika **64** (2018), no. 1, 211–229. MR 3778221

[24] R. Latała and T. Tkocz, *A note on suprema of canonical processes based on random variables with regular moments*, Electron. J. Probab. **20** (2015), no. 36, 17. MR 3335827

[25] R. Latała, R. van Handel, and P. Youssef, *The dimension-free structure of nonhomogeneous random matrices*, Invent. Math. **214** (2018), no. 3, 1031–1080. MR 3878726

[26] R. Latała and J.O. Wojtaszczyk, *On the infimum convolution inequality*, Studia Math. **189** (2008), no. 2, 147–187. MR 2449135

[27] B. Maurey, *Some deviation inequalities*, Geom. Funct. Anal. **1** (1991), no. 2, 188–197. MR 1097258

[28] K. Oleszkiewicz, *Precise moment and tail bounds for Rademacher sums in terms of weak parameters*, Israel J. Math. **203** (2014), no. 1, 429–443. MR 3273447

[29] G. Paouris, *Concentration of mass on convex bodies*, Geom. Funct. Anal. **16** (2006), no. 5, 1021–1049. MR 2276533

[30] G. Pisier and Q. Xu, *Non-commutative L^p-spaces*, Handbook of the geometry of Banach spaces, Vol. 2, North-Holland, Amsterdam, 2003, pp. 1459–1517. MR 1999201

[31] S. Riemer and C. Schütt, *On the expectation of the norm of random matrices with non-identically distributed entries*, Electron. J. Probab. **18** (2013), no. 29, 13. MR 3035757

[32] Y. Seginer, *The expected norm of random matrices*, Combin. Probab. Comput. **9** (2000), no. 2, 149–166. MR 1762786

[33] Y. Shu and M. Strzelecki, *A characterization of a class of convex log-Sobolev inequalities on the real line*, Ann. Inst. Henri Poincaré Probab. Stat. **54** (2018), no. 4, 2075–2091. MR 3865667

[34] M. Strzelecka, *Estimates for norms of log-concave matrices*, work in progress (2019+).

[35] M. Strzelecka, M. Strzelecki, and T. Tkocz, *On the convex infimum convolution inequality with optimal cost function*, ALEA Lat. Am. J. Probab. Math. Stat. **14** (2017), no. 2, 903–915. MR 3731797

[36] M. Talagrand, *Regularity of infinitely divisible processes*, Ann. Probab. **21** (1993), no. 1, 362–432. MR 1207231