Electronic Journal of Probability

Electron. J. Probab. **24** (2019), no. 107, 1–15.
ISSN: 1083-6489 https://doi.org/10.1214/19-EJP365

Estimates of norms of log-concave random matrices with dependent entries*

Marta Strzelecka†

Abstract

We prove estimates for $\mathbb{E}\|X : \ell_{p'}^n \to \ell_q^m\|$ for $p,q \geqslant 2$ and any random matrix X having the entries of the form $a_{ij}Y_{ij}$, where $Y = (Y_{ij})_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n}$ has i.i.d. isotropic log-concave rows and p' denotes the Hölder conjugate of p. This generalises a result of Guédon, Hinrichs, Litvak, and Prochno for Gaussian matrices with independent entries. Our estimate is optimal up to logarithmic factors. As a byproduct we provide an analogous bound for $m \times n$ random matrices, whose entries form an unconditional vector in $\mathbb{R}^{mn}$. We also prove bounds for norms of matrices whose entries are certain Gaussian mixtures.

Keywords: random matrices; operator norm; log-concave vectors; unconditional vectors.
AMS MSC 2010: 60B20; 46B09; 15B52.
Submitted to EJP on April 3, 2019, final version accepted on September 18, 2019.

1 Introduction and main results

By $\|A\|_{p,q}$ we denote the operator norm of the matrix A from ℓ_p to ℓ_q. A classical result regarding spectra of random matrices is Wigner's Semicircle Law, which describes the limit of the empirical spectral measure 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 [16] that for a random matrix X with i.i.d. symmetric entries $\mathbb{E}\|X\|_{2,2}$ is of the same order as the expectation of the maximum Euclidean norm of rows and columns of X. The same holds true for structured Gaussian matrices (i.e. when $X_{ij} = a_{ij}g_{ij}$ and g_{ij} are i.i.d. standard Gaussian variables), as was shown recently by Latała, van Handel, and Youssef in [14], and up to a logarithmic factor for any random matrix X with independent centred entries, see [15]. The main novelty of these two results is that they do not require the entries of X to be equally distributed (nor to have equal variances).

*The research was supported by the National Science Centre, Poland via the grants 2015/19/N/ST1/02661 and 2018/28/T/ST1/00001.

†Institute of Mathematics, University of Warsaw, Poland. E-mail: martast@mimuw.edu.pl

In [9] another upper bound for $\mathbb{E}\|X\|_{2,2}$ was proved:

$$\mathbb{E}\|X\|_{2,2} \leqslant C\Big[\max_i\Big(\sum_j \mathbb{E}X_{ij}^2\Big)^{1/2} + \max_j\Big(\sum_i \mathbb{E}X_{ij}^2\Big)^{1/2} + \Big(\sum_{i,j}\mathbb{E}X_{ij}^4\Big)^{1/4}\Big],$$

where C is a universal constant. It also requires the independence of entries, but not the equality of their distributions. This bound is dimension free, but is in some cases worse than the one from [15].

Upper bounds for the expectation of other operator norms were investigated in [2] in the case of independent centred entries bounded by 1. For $q \geqslant 2$ and $m \times n$ matrices the authors proved that $\mathbb{E}\|X\|_{2,q} \lesssim \max\{m^{1/q}, \sqrt{n}\}$. In [6] Guédon, Hinrichs, Litvak, and Prochno proved that for a structured Gaussian matrix $X = (a_{ij}X_{ij})_{i\leqslant m, j\leqslant n}$ and $p, q \geqslant 2$,

$$\begin{aligned}\mathbb{E}\|X\|_{p',q} \leqslant C(p,q)\Big[\big(\log m\big)^{1/q} \max_{1\leqslant i\leqslant m}\Big(\sum_{j=1}^n |a_{ij}|^p\Big)^{1/p} + \max_{1\leqslant j\leqslant n}\Big(\sum_{i=1}^m |a_{ij}|^q\Big)^{1/q} \\ + \big(\log m\big)^{1/q}\mathbb{E}\max_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant n}}|X_{ij}|\Big]. \quad (1.1)\end{aligned}$$

This estimate is optimal up to logarithmic factors (see Remark 1.2 below). Estimating more general norms required a different approach than those used for estimating the spectral norm. In particular the moment method fails in estimating $\mathbb{E}\|X\|_{p',q}$ for $(p,q) \neq (2,2)$ as it gives information only about the spectrum of X.

All the results mentioned above require the independence of entries of X. We will show how to relax the independence assumption. The aim of this article is to generalise the main result of [6] to a wide class of random matrices with independent uncorrelated log-concave rows by following the scheme of proof of the original theorem from [6]. We work with a more general class of random matrices, thus the proof from [6] may not be rewritten verbatim, but it requires some extra tools: the comparison of weak and strong moments of ℓ_p-norm of X from [11] and a Sudakov minoration-type bound from [10].

Before we state our main results, let us say a few words about log-concave vectors, which are widely investigated in convex geometry and high dimensional probability. We call a random vector Z in $\mathbb{R}^n$ log-concave, if for any compact nonempty sets $K, L \subset \mathbb{R}^n$ and $\lambda \in [0,1]$,

$$\mathbb{P}\big(Z \in \lambda K + (1-\lambda)L\big) \geqslant \mathbb{P}(Z\in K)^\lambda \mathbb{P}(Z \in L)^{1-\lambda}.$$

The class of log-concave vectors is closed under linear transformations, convolutions and weak limits. By the result of Borell [3] an n-dimensional vector with a full dimensional support is log-concave if and only if it has a log-concave density, i.e. a density (with respect to the Lebesgue measure) of the form e^{-h}, where h is a convex function with values in $(-\infty, \infty]$.

Log-concave vectors are a natural generalisation of vectors distributed uniformly over convex bodies. The distribution of any log-concave vector can be obtained as a weak limit of projections of uniform measures over (higher dimensional) convex bodies (see for example [1]).

We will frequently use a basic property of log-concave vectors: the regularity of $f(Z)$ for log-concave vectors X and seminorms f, which states that

$$\big(\mathbb{E}f(Z)^p\big)^{1/p} \leqslant C_1\frac{p}{q}\big(\mathbb{E}f(Z)^q\big)^{1/q} \quad \text{for } p \geqslant q \geqslant 1, \quad (1.2)$$

where C_1 is a universal constant (see [4, Theorem 2.4.6]). Other results and conjectures concerning log-concave vectors are discussed in the monograph [4].

We say that a vector in $\mathbb{R}^n$ is isotropic if its covariance matrix is the identity. If Z is a log-concave random vector in $\mathbb{R}^n$ with a full dimensional support, then there exists a linear transformation T such that $\mathrm{Cov}(TZ) = \mathrm{Id}$, so the isotropicity is only a matter of normalisation.

If $A = (A_{ij})_{i\leqslant m, j\leqslant n}$ is an $m \times n$ matrix, we denote by $A_i \in \mathbb{R}^n$ its i-th row and by $A^{(j)} \in \mathbb{R}^m$ its j-th column. We are now ready to present the main theorem.

Theorem 1.1. *Let $m \geqslant 2$, let $Y_1, \dots, Y_m$ 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 $i \leqslant m, j \leqslant n$, where Y_{ij} is the j-th coordinate of Y_i. Then for every $p, q \geqslant 2$ we have*

$$\mathbb{E}\|X\|_{p',q} \leqslant C(p,q)\Big[\big(\log m\big)^{1/q}\max_{1\leqslant i\leqslant m}\big\|A_i\big\|_p + \max_{1\leqslant j\leqslant n}\big\|A^{(j)}\big\|_q + \big(\log m\big)^{1+\frac{1}{q}}\mathbb{E}\max_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}|X_{ij}|\Big], \tag{1.3}$$

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

Let us stress that the theorem requires the independence only between the rows and does not require the independence of the entries of X.

Remark 1.2. Note that the bound from Theorem 1.1 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 the regularity of Y_{ij} (see (1.2)) that $\mathbb{E}|Y_{ij}| \geqslant (2C_1)^{-1}\big(\mathbb{E}Y_{ij}^2\big)^{1/2} = (2C_1)^{-1}$. Hence for every $j \leqslant n$,

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

Since $\|X\|_{p',q} = \|X^T\|_{q',p}$, we also have $\mathbb{E}\|X\|_{p',q} \geqslant (2C_1)^{-1}\|A_i\|_p$ for all $i \leqslant m$. Moreover, for all $i \leqslant m$ and $j \leqslant n$,

$$\|X\|_{p',q} = \sup_{u\in\ell_{p'}^n}\sup_{v\in\ell_{q'}^n} v^T X u \geqslant e_i^T X(\operatorname{sgn} X_{ij} e_j) = |X_{ij}|.$$

Therefore

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

which yields the claim.

Note that in (1.3) the logarithmic term appears in front of the norm of rows, but not in front of the norm of columns, so our bound is not symmetric. This is not so strange, since the assumptions of the theorem are also non-symmetric: we assume that the rows are weighted i.i.d. random vectors, but no independence between the columns is required. However, the asymmetry of the bound in Theorem 1.1 is mainly a residue of the proof and the author does not know if one may skip the logarithmic factor in front of the norm of rows.

Since $\|X\|_{p',q} = \|X^T\|_{q',p}$, one can assume in (1.3) that the columns (instead of the rows) of Y are i.i.d. isotropic log-concave vectors. Then Theorem 1.1 yields

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

We mentioned previously results from [16, 14, 15], which provide the bounds of the expected value of the operator norm in terms of expected values of norms of rows and columns. Theorem 1.1 easily implies an estimate of the same kind: by (1.2), the assumption $\mathbb{E}Y_{ij}^2 = 1$, and the Jensen inequality we get

$$\begin{aligned}
&\big(\log m\big)^{1/q} \max_{1\leqslant i\leqslant m} \big\|A_i\big\|_p + \max_{1\leqslant j\leqslant n} \big\|A^{(j)}\big\|_q + \big(\log m\big)^{1+\frac{1}{q}} \mathbb{E} \max_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} |X_{ij}| \\
&\qquad \leqslant 2C_1 \bigg(\big(\log m\big)^{1/q} \max_{1\leqslant i\leqslant m} \big\|(A_{ij}\mathbb{E}|Y_{ij}|)_j\big\|_p + \max_{1\leqslant j\leqslant n} \big\|(A_{ij}\mathbb{E}|Y_{ij}|)_i\big\|_q \\
&\qquad\qquad + \big(\log m\big)^{1+\frac{1}{q}} \mathbb{E} \max_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} |X_{ij}| \bigg) \\
&\qquad \leqslant 2C_1 \bigg(\big(\log m\big)^{1/q} \mathbb{E} \max_{1\leqslant i\leqslant m} \|X_i\|_p + \mathbb{E} \max_{1\leqslant j\leqslant n} \|X^{(j)}\|_q + (\log m)^{1+\frac{1}{q}} \mathbb{E} \max_{1\leqslant i\leqslant m} \|X_i\|_p \bigg) \\
&\qquad \leqslant C \bigg((\log m)^{1+\frac{1}{q}} \mathbb{E} \max_{1\leqslant i\leqslant m} \|X_i\|_p + \mathbb{E} \max_{1\leqslant j\leqslant n} \|X^{(j)}\|_q \bigg), \qquad (1.4)
\end{aligned}$$

where C is a universal constant. Therefore (1.3) yields the following corollary.

Corollary 1.3. *Under the assumptions of Theorem 1.1 we have*

$$\mathbb{E}\|X\|_{p',q} \leqslant C(p,q) \bigg((\log m)^{1+\frac{1}{q}} \mathbb{E} \max_{1\leqslant i\leqslant m} \|X_i\|_p + \mathbb{E} \max_{1\leqslant j\leqslant n} \|X^{(j)}\|_q \bigg).$$

Remark 1.4. If the rows and columns of a random matrix Y are isotropic and log-concave (we do not require independence), and $p, q \geqslant 1$, then

$$\begin{aligned}
&\mathbb{E} \max_{1\leqslant i\leqslant m} \Big(\sum_{j=1}^n |A_{ij}Y_{ij}|^p \Big)^{1/p} + \mathbb{E} \max_{1\leqslant j\leqslant n} \Big(\sum_{i=1}^m |A_{ij}Y_{ij}|^q \Big)^{1/q} \\
&\qquad \leqslant C \Big(p^2 \max_{1\leqslant i\leqslant m} \big\|A_i\big\|_p + q^2 \max_{1\leqslant j\leqslant n} \big\|A^{(j)}\big\|_q + (p+q)\log(m\vee n) \mathbb{E} \max_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} |A_{ij}Y_{ij}| \Big). \qquad (1.5)
\end{aligned}$$

This means that inequality (1.4) may be reversed up to a logarithmic factor and constants depending only on p and q in the log-concave setting. Therefore the estimates from Theorem 1.1 and Corollary 1.3 are equivalent up to a logarithmic factor.

Inequality (1.5) follows directly from the following proposition.

Proposition 1.5. *Let Y be an $m \times n$ random matrix, with isotropic and log-concave rows, let B be a deterministic $m \times n$ matrix, and let $p \geqslant 1$. Then*

$$\mathbb{E} \max_{1\leqslant i\leqslant m} \Big(\sum_{j=1}^n |B_{ij}Y_{ij}|^p \Big)^{1/p} \leqslant C \bigg(p^2 \max_{1\leqslant i\leqslant m} \Big(\sum_{j=1}^n |B_{ij}|^p \Big)^{1/p} + p\log(m\vee n) \mathbb{E} \max_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} |B_{ij}Y_{ij}| \bigg),$$

where C is a universal constant.

Our next result concerns unconditional matrices. Recall that we say that a random vector Z in $\mathbb{R}^d$ is unconditional, if for every choice of signs $\eta \in \{-1,1\}^d$ the vectors Z and $(\eta_i Z_i)_{i\leqslant d}$ are equally distributed (or, equivalently, that Z and $(\varepsilon_i Z_i)_{i\leqslant d}$ are equally distributed, where $\varepsilon_1, \ldots, \varepsilon_d$ are i.i.d. symmetric Bernoulli variables, independent of Z). The assertion of the next corollary is expressed in the spirit of Corollary 1.3, which is more natural in the non log-concave setting (without the assumption of log-concavity inequality (1.5) may no longer be true, even up to additional logarithmic factors).

Corollary 1.6. *Assume that X is a random matrix such that the (mn)-dimensional vector $(X_{1,1}, \ldots X_{1,n}, X_{2,1}, \ldots, X_{2,n}, X_{m,1}, \ldots, X_{mn})$ is unconditional. Then for every $p, q \geqslant 2$ we have*

$$\mathbb{E}\|X\|_{p',q} \leqslant C(p,q) \bigg((\log m)^{\frac{1}{2}+\frac{1}{q}} \mathbb{E} \max_{1\leqslant i\leqslant m} \Big(\sum_{j=1}^n |X_{ij}|^p \Big)^{1/p} + \sqrt{\log n}\, \mathbb{E} \max_{1\leqslant j\leqslant n} \Big(\sum_{i=1}^m |X_{ij}|^q \Big)^{1/q} \bigg), \qquad (1.6)$$

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

The asymmetry of the bound in (1.6) is just a residue of its proof: the factor $(\log m)^{1/q}$ could be skipped if one could prove (1.1) in the Gaussian case with constants independent of the dimensions. Such an improved version of (1.1) was conjectured in [6].

Let us compare (1.6) with the following result of Seginer coming from [16]. Let $X = (A_{ij}\varepsilon_{ij})_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n}$ where ε_{ij} are independent random signs. Then

$$\mathbb{E}\|X\|_{2\to 2} \leqslant C\sqrt[4]{\log\min(m,n)}\left(\max_{1\leqslant i\leqslant m}\|A_i\|_2 + \max_{1\leqslant j\leqslant n}\|A^{(j)}\|_2\right), \tag{1.7}$$

where C is a universal constant. Moreover, for every m, n, there exists a matrix $(A_{ij})_{1\leqslant i\leqslant m, 1\leqslant j\leqslant n}$, for which (1.7) may be reversed (up to a universal constant). Therefore one may not skip the dependence on the dimension also in estimate (1.6).

In the case $(p,q) = (2,2)$ and $X = (A_{ij}\varepsilon_{ij})$ Corollary 1.6 provides a bound with worse dependence of the dimension than in (1.7). However, our result works also in much more general setting than the result of Seginer does (in particular, our Corollary works also when the entries of X are dependent).

The rest of this note is organised as follows. Section 2 contains results from other articles to be used in a sequel. Section 3 contains generalisations of Lemmas 3.1 and 3.2 from [6] to the log-concave setting and the proof of Theorem 1.1. In Section 4 we show how to deduce an analogue of Theorem 1.1 for Gaussian mixtures (see Corollary 4.2) and we provide a proof of Proposition 1.5. Section 5 is devoted to the proof of Corollary 1.6.

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 $C_1, C_2, \dots$ We may and do always assume that $C, C_i \geqslant 1$. For two quantities a, b we write $a \lesssim b$ if there exists a constant C, such that $a \leqslant Cb$, and $a \sim b$, if $a \lesssim b$ and $b \lesssim a$. For two numbers a and b we write $a \vee b$ instead of $\max\{a,b\}$.

For a random variable Z by $\|Z\|_p$ we denote the p-th integral norm of Z, i.e. the quantity $(\mathbb{E}|Z|^p)^{1/p}$; in the case when $Z = \|U\|$ we also call this quantity the p-th strong moment of a random vector U associated with the norm $\|\cdot\|$. For a vector $x \in \mathbb{R}^n$ (in particular for a random vector Z) and $r \geqslant 1$, 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 $\|Z\|_q$ means for a random object Z, so the double meaning of $\|\cdot\|_q$ will not lead to any misunderstanding. Recall that for an $m \times n$ matrix A by $\|A\|_{p,q}$ we denote its norm from ℓ_p^n to ℓ_q^m. For $p \in [1,\infty]$ we denote by p' the Hölder conjugate of p, i.e. the number such that $1 = \frac{1}{p} + \frac{1}{p'}$.

2 Preliminaries

In the proof of the main theorem we will need the comparison of weak and strong moments for ℓ_p-norms of log-concave vectors:

Theorem 2.1 ([11, Theorem 5]). *Let Z be a log-concave vector in $\mathbb{R}^n$ and let $p \in [1,\infty)$. Then*

$$(\mathbb{E}\|Z\|_p^q)^{1/q} \leqslant Cp\Big(\mathbb{E}\|Z\|_p + \sigma_{p,Z}(q)\Big) \quad \textit{for } q \geqslant 1,$$

where

$$\sigma_{p,Z}(q) := \sup_{t\in B_{p'}^n}\Big\|\sum_{i=1}^n t_i Z_i\Big\|_q$$

is the q-th weak moment of X associated with the ℓ_p-norm.

We will use the previous theorem also in the tail-bound version:

Corollary 2.2. *Assume Z is a log-concave vector in $\mathbb{R}^n$ and $p \in [1,\infty)$. Then for all $u > 0$,*

$$\mathbb{P}\Big(\|Z\|_p \geqslant C_2 p\big(u + \mathbb{E}\|Z\|_p\big)\Big) \leqslant C_3 \sup_{t\in B^n_{p'}} \mathbb{P}\bigg(\Big|\sum_{i=1}^n t_i Z_i\Big| \geqslant u\bigg). \tag{2.1}$$

For the Reader's convenience we give a proof of this corollary, which goes along the lines of the proof of Corollary 1.3 in [12].

Proof. Define a random variable $S := \|Z\|_p$. By the Paley–Zygmund inequality and (1.2) we have for $t \in \mathbb{R}^n$, and $q \geqslant 1$,

$$\begin{aligned}\mathbb{P}\bigg(\Big|\sum_{i=1}^n t_i Z_i\Big| \geqslant \frac{1}{2}\Big\|\sum_{i=1}^n t_i Z_i\Big\|_q\bigg) &= \mathbb{P}\bigg(\Big|\sum_{i=1}^n t_i Z_i\Big|^q \geqslant 2^{-q}\mathbb{E}\Big|\sum_{i=1}^n t_i Z_i\Big|^q\bigg)\\ &\geqslant (1-2^{-q})^2\left(\frac{\big\|\sum_{i=1}^n t_i Z_i\big\|_q}{\big\|\sum_{i=1}^n t_i Z_i\big\|_{2q}}\right)^{2q} \geqslant e^{-C_4 q}.\end{aligned} \tag{2.2}$$

In order to show (2.1) we consider 3 cases.

Case 1. $2u < \sup_{t\in B^n_{p'}} \|\sum_{i=1}^n t_i Z_i\|_2$. Then by (2.2)

$$\sup_{t\in B^n_{p'}} \mathbb{P}\bigg(\Big|\sum_{i=1}^n t_i Z_i\Big| \geqslant u\bigg) \geqslant e^{-2C_4}$$

and (2.1) obviously holds if $C_3 \geqslant \exp(2C_4)$.

Case 2. $\sup_{t\in B^n_{p'}} \|\sum_{i=1}^n t_i Z_i\|_2 \leqslant 2u < \sup_{t\in B^n_{p'}} \|\sum_{i=1}^n t_i Z_i\|_\infty$. Let us then define

$$q := \sup\bigg\{r \geqslant 2C_4 \colon \sup_{t\in B^n_{p'}} \Big\|\sum_{i=1}^n t_i Z_i\Big\|_{r/C_4} \leqslant 2u\bigg\}.$$

By (2.2) we have

$$\sup_{t\in B^n_{p'}} \mathbb{P}\bigg(\Big|\sum_{i=1}^n t_i Z_i\Big| \geqslant u\bigg) \geqslant e^{-q}.$$

By (1.2), Theorem 2.1, and Chebyshev's inequality we have

$$\mathbb{P}\big(S \geqslant C_5 p(\mathbb{E}S + u)\big) \leqslant \mathbb{P}(S \geqslant e\|S\|_q) \leqslant e^{-q}$$

for C_5 large enough. Thus (2.1) holds in this case.

Case 3. $u > \sup_{t\in B^n_{p'}} \|\sum_{i=1}^n t_i Z_i\|_\infty = \|S\|_\infty$. Then $\mathbb{P}(S \geqslant u) = 0$ and (2.1) holds for any $C_2 \geqslant 1$. □

In the proof of Theorem 1.1 we will use Theorem 2.1 from [6], which is a version of the results provided by Guédon–Rudelson in [8], and by Guédon–Mendelson–Pajor–Tomczak-Jaegerman in [7]. Below we give only a particular version of the theorem; the general result is stated in [6].

Theorem 2.3 ([6, Theorem 2.1])**.** *Let $X_1, \dots X_m \in \mathbb{R}^n$ be independent random vectors, and let $p, q \geqslant 2$. Define*

$$u := \sup_{t\in B^n_{p'}} \Big(\sum_{i=1}^m \mathbb{E}\big|\langle X_i, t\rangle\big|^q\Big)^{1/q} \tag{2.3}$$

and

$$v := \Big(p^9 \log m \,\mathbb{E} \max_{1\leqslant i\leqslant m} \|X_i\|_p^q\Big)^{1/q}. \tag{2.4}$$

Then

$$\Bigg[\mathbb{E} \sup_{t\in B^n_{p'}} \Big|\sum_{i=1}^m \Big(\big|\langle X_i,t\rangle\big|^q - \mathbb{E}\big|\langle X_i,t\rangle\big|^q\Big)\Big|\Bigg]^{1/q} \leqslant C(\sqrt{uv}+v) \leqslant 2C(u+v).$$

3 Proof of Theorem 1.1

The next two lemmas provide estimates of the quantities u and v appearing in Theorem 2.3.

Lemma 3.1. *Assume that p, q, X, and Y are as in Theorem 1.1. Then*

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

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

Lemma 3.2. *Assume that p, q, X, and Y are as in Theorem 1.1. Then*

$$\sup_{t\in B^n_{p'}} \Big(\sum_{i=1}^m \mathbb{E}\big|\langle X_i,t\rangle\big|^q\Big)^{1/q} \leqslant C_1 q \max_{1\leqslant j\leqslant n} \big\|A^{(j)}\big\|_q. \tag{3.1}$$

In the proof of Lemma 3.1 we will also need the following estimate:

Lemma 3.3. *Assume that Z is an isotropic log-concave vector in $\mathbb{R}^m$. Then for all $1 \leqslant k \leqslant m$ and all $a \in \mathbb{R}^m$ we have*

$$\mathbb{E} \max_{1\leqslant i\leqslant m} |a_i Z_i| \geqslant \frac{1}{C_6} \max_{k\leqslant m}\big(a_k^* \min_{i\leqslant m} \|Z_i\|_{\log(k+1)}\big),$$

where $(a_i^*)_{i=1}^m$ denotes the non-increasing rearrangement of $(|a_i|)_{i=1}^m$.

In order to prove Theorem 1.1, we repeat the scheme of the proof from [6].

Proof of Theorem 1.1. Let u and v be given by formulas (2.3) and (2.4). The triangle inequality, Theorem 2.3, Lemma 3.1, and Lemma 3.2 yield

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

The main contribution of this article lies in the proofs of Lemmas 3.1, 3.2, and 3.3.

Proof of Lemma 3.3. We may and do assume that $a_1 \geqslant a_2 \geqslant \ldots \geqslant a_m \geqslant 0$, i.e. $a_i^* = a_i$ for $i \leqslant m$. By [10, Proposition 3.3] we have for all $k \leqslant m$,

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

Thus

$$\mathbb{E} \max_{1\leqslant i\leqslant m} |a_i Z_i| = \max_{1\leqslant k\leqslant m} \mathbb{E} \max_{1\leqslant i\leqslant k} |a_i Z_i| \geqslant C^{-1} \max_{1\leqslant k\leqslant m}\big(a_k \min_{1\leqslant i\leqslant m} \|Z_i\|_{\log(k+1)}\big). \qquad \square$$

Proof of Lemma 3.1. We may and do assume that $m \geqslant 2$.

Since we may approximate A_{ij} by nonzero numbers, we may and do assume that $A_{ij} \neq 0$ for all i, j. Let C_2, C_3 be the constants from (2.1), let C_6 be the constant from Lemma 3.3, 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,

$$\begin{aligned} &\Big(\mathbb{E} \max_{1\leqslant i\leqslant m} \|X_i\|_p^q\Big)^{1/q} \\ &\leqslant \bigg(\mathbb{E} \max_{1\leqslant i\leqslant m}\Big[\big(\|X_i\|_p - C_2 p\mathbb{E}\|X_i\|_p\big)^q \mathbf{1}_{\{\|X_i\|_p \geqslant C_2 p \mathbb{E}\|X_i\|_p\}}\Big]\bigg)^{1/q} + C_2 p \max_{1\leqslant i\leqslant m} \mathbb{E}\|X_i\|_p. \quad (3.2)\end{aligned}$$

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

$$\begin{aligned} \mathbb{E}\|X_i\|_p &\leqslant \Big(\sum_{j=1}^n \mathbb{E}|Y_{ij}|^p |A_{ij}|^p\Big)^{1/p} \leqslant \max_{j\leqslant n} \|Y_{ij}\|_p \|A_i\|_p \leqslant C_1 p \|A_i\|_p \\ &\leqslant C_1 p \max_{1\leqslant k\leqslant m} \|A_k\|_p. \quad (3.3)\end{aligned}$$

Now it suffices to estimate the first term of (3.2). Let

$$B := C_1^2 C_6 \log(m+1)\, \mathbb{E} \max_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} |X_{ij}| \qquad \text{and} \qquad \sigma := \big(\max_{1\leqslant i\leqslant m} \sigma_{p,X_i}(2)\big) \vee B.$$

By (2.1) and the integration by parts we have

$$\begin{aligned} &\mathbb{E} \max_{1\leqslant i\leqslant m}\Big[\big(\|X_i\|_p - C_2 p\mathbb{E}\|X_i\|_p\big)^q \mathbf{1}_{\{\|X_i\|_p \geqslant C_2 p \mathbb{E}\|X_i\|_p\}}\Big] \\ &\quad\leqslant (2C_2 pe\sigma)^q + \int_{2C_2 pe\sigma}^\infty qv^{q-1} \mathbb{P}\Big(\max_{1\leqslant i\leqslant m} \big(\|X_i\|_p - C_2 p\mathbb{E}\|X_i\|_p\big) \geqslant v\Big) dv \\ &\quad\leqslant (2C_2 pe\sigma)^q + \sum_{i=1}^m \int_{2C_2 pe\sigma}^\infty qv^{q-1} \mathbb{P}\big(\|X_i\|_p - C_2 p\mathbb{E}\|X_i\|_p \geqslant v\big) dv \\ &\quad= (2C_2 pe\sigma)^q + (C_2 p)^q \sum_{i=1}^m \int_{2e\sigma}^\infty qu^{q-1} \mathbb{P}\big(\|X_i\|_p - C_2 p\mathbb{E}\|X_i\|_p \geqslant C_2 pu\big) du \\ &\quad\leqslant (2C_2 pe\sigma)^q + (C_2 p)^q C_3 \sum_{i=1}^m \int_{2e\sigma}^\infty qu^{q-1} \sup_{\|t\|_{p'}\leqslant 1} \mathbb{P}\Big(\Big|\sum_{j=1}^n t_j X_{ij}\Big| \geqslant u\Big) du. \quad (3.4)\end{aligned}$$

We want to estimate the function we integrate in (3.4). Fix $u \geqslant 2e\sigma$. For i such that $u \geqslant \sup_{\|t\|_{p'}\leqslant 1} \|\sum_{j=1}^n t_j X_{ij}\|_\infty$ the function we integrate vanishes, so from now on we will consider only i's for which $u < \sup_{\|t\|_{p'}\leqslant 1} \|\sum_{j=1}^n t_j X_{ij}\|_\infty$.

Note that if $1 \leqslant i \leqslant m$ and $\sup_{\|t\|_{p'}\leqslant 1} \|\sum_{j=1}^n t_j X_{ij}\|_\infty > u \geqslant e\sigma \geqslant e\sigma_{p,X_i}(2)$, then

$$r := r(i) := \sup\{s \geqslant 2 : \sigma_{p,X_i}(s) \leqslant u/e\} \in [2, \infty)$$

and $\sigma_{p,X_i}(r) = u/e$. Therefore

$$\sup_{\|t\|_{p'}\leqslant 1} \mathbb{P}\Big(\Big|\sum_{j=1}^n t_j X_{ij}\Big| \geqslant u\Big) \leqslant \frac{\sup_{\|t\|_{p'}\leqslant 1} \|\langle t, X_i\rangle\|_r^r}{u^r} = e^{-r}. \quad (3.5)$$

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

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

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

$$\begin{aligned}\sigma_{p,X_i}(t) \leqslant \sigma_{2,X_i}(t) \leqslant C_1 t \max_{|x|\leqslant 1}\Big(\mathbb{E}\Big(\sum_{j=1}^n A_{ij}Y_{ij}x_j\Big)^2\Big)^{1/2} &= C_1 t \max_{|x|\leqslant 1}\Big(\sum_{j=1}^n A_{ij}^2 x_j^2\Big)^{1/2} \\ = C_1 t \max_{1\leqslant j\leqslant n} |A_{ij}|\cdot\|Y_{ij}\|_2 \leqslant C_1\varphi(t) \max_{1\leqslant j\leqslant n}|A_{ij}|. \qquad (3.6)\end{aligned}$$

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

$$\max_{1\leqslant j\leqslant n}|A_{1j}| \geqslant \ldots \geqslant \max_{1\leqslant j\leqslant n}|A_{mj}|.$$

Let $j(i) \leqslant n$ be such an index that $|A_{ij(i)}| = \max_{1\leqslant j\leqslant n}|A_{ij}|$. Lemma 3.3 applied to $Z_i = Y_{ij(i)}$ and the non-increasing sequence $a_i = |A_{ij(i)}|$ implies

$$\mathbb{E}\max_{\substack{1\leqslant i\leqslant m\\1\leqslant j\leqslant n}}|X_{ij}| \geqslant \mathbb{E}\max_{1\leqslant i\leqslant m}|A_{ij(i)}Y_{ij(i)}| \geqslant C_6^{-1}\big(\log(m+1)\big)^{-1}\max_{1\leqslant i\leqslant m}\Big(\varphi\big(\log(i+1)\big)|A_{ij(i)}|\Big),$$

so for all $i \leqslant m$ we have

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

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

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

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

Therefore for all $i \leqslant m$ we have

$$\frac{B}{C_1} = \frac{1}{\lambda C_1}\sigma_{p,X_i}(r) \leqslant \sigma_{p,X_i}(r/\lambda) \overset{(3.6)}{\leqslant} C_1\varphi\Big(\frac{r}{\lambda}\Big)\max_{1\leqslant j\leqslant n}|A_{ij}| \leqslant \frac{B\varphi(\frac{r}{\lambda})}{C_1\varphi(\log(i+1))}. \qquad (3.7)$$

Since the function φ is strictly increasing, the previous inequality yields $r \geqslant \lambda\log(i+1)$. This together with (3.5) implies that (recall that $\lambda = \frac{u}{Be} \geqslant 2$)

$$\sum_{i=1}^m \sup_{\|t\|_{p'}\leqslant 1}\mathbb{P}\Big(\Big|\sum_{j=1}^n t_jX_{ij}\Big| \geqslant u\Big) \leqslant \sum_{i=1}^m (i+1)^{-\frac{u}{eB}} \leqslant 2^{-\frac{u}{eB}} + \int_2^\infty x^{-\frac{u}{eB}}dx \leqslant 3\cdot 2^{-\frac{u}{e\sigma}}. \qquad (3.8)$$

Inequalities (3.4), (3.8), and the Stirling formula yield

$$\begin{aligned}&\Big(\mathbb{E}\Big[\max_{1\leqslant i\leqslant m}\big(\|X_i\|_p - C_2\mathbb{E}\|X_i\|_p\big)^q\mathbf{1}_{\{\|X_i\|_p\geqslant C_2\mathbb{E}\|X_i\|_p\}}\Big]\Big)^{1/q} \\ &\qquad\leqslant \Big((2C_2pe\sigma)^q + 3(C_2p)^qC_3\int_0^\infty qu^{q-1}2^{-\frac{u}{e\sigma}}du\Big)^{1/q} \\ &\qquad\leqslant \Big((2C_2pe\sigma)^q + (CC_2p\sigma)^qC_3\int_0^\infty qv^{q-1}e^{-v}dv\Big)^{1/q} \\ &\qquad\leqslant CC_2C_3^{1/q}\sigma pq. \qquad (3.9)\end{aligned}$$

If $B \geqslant \max_{1\leqslant i\leqslant m} \sigma_{p,X_i}(2)$, then $\sigma = B = C\log(m+1)\, \mathbb{E}\max_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} |X_{ij}|$ and the assertion follows by (3.2), (3.3), and (3.9). Otherwise, by (1.2) we get

$$\sigma = \max_{1\leqslant i\leqslant m} \sigma_{p,X_i}(2) \leqslant 2C_1 \max_{1\leqslant i\leqslant m} \sigma_{p,X_i}(1) \leqslant 2C_1 \max_{1\leqslant i\leqslant m} \mathbb{E}\|X_i\|_p,$$

where the second inequality holds since the weak first moment is bounded above by the strong first moment. This together with (3.2), (3.3), and (3.9) gives the assertion. $\square$

Proof of Lemma 3.2. Since $p \geqslant 2$, $B_{p'} \subset B_2$. Thus we may and do assume $p = 2$. By (1.2), the isotropicity of Y, and the Jensen inequality we have

$$\begin{aligned}
\sup_{t\in B_2^n}\Big(\sum_{i=1}^m \mathbb{E}\big|\langle X_i,t\rangle\big|^q\Big)^{1/q} &\leqslant C_1 q \sup_{\|t\|_2\leqslant 1}\Big(\sum_{i=1}^m\Big(\mathbb{E}\big|\langle X_i,t\rangle\big|^2\Big)^{q/2}\Big)^{1/q}\\
&= C_1 q \sup_{\|t\|_2=1}\Big(\sum_{i=1}^m\Big(\sum_{j=1}^n A_{ij}^2 t_j^2\Big)^{q/2}\Big)^{1/q}\\
&\leqslant C_1 q \sup_{\|t\|_2=1}\Big(\sum_{i=1}^m\sum_{j=1}^n |A_{ij}|^q t_j^2\Big)^{1/q}\\
&= C_1 q \Big(\sup_{\|t\|_2=1}\sum_{j=1}^n \big\|A^{(j)}\big\|_q^q t_j^2\Big)^{1/q}\\
&= C_1 q \max_{1\leqslant j\leqslant n}\big\|A^{(j)}\big\|_q.
\end{aligned}$$
$\square$

Remark 3.4. By the same reasoning as in the log-concave case, we may prove the following result for matrices with independent heavy tailed entries. 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 $\mathbb{E}Y_{ij}^2 = 1$, and let $L > 0$. Assume that for any $r \geqslant 2$ and any $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$ we have $\frac{r^\beta}{L} \leqslant \|Y_{ij}\|_r \leqslant Lr^\beta$ with $\beta \in [\frac12, 1]$. Then for every $p, q \geqslant 2$ we have

$$\mathbb{E}\|X\|_{p',q} \leqslant C(p,q,L)\Big[\big(\log m\big)^{1/q}\max_{1\leqslant i\leqslant m}\big\|A_i\big\|_p + \max_{1\leqslant j\leqslant n}\big\|A^{(j)}\big\|_q + \big(\log m\big)^{1/q}\mathbb{E}\max_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}|X_{ij}|\Big]. \tag{3.10}$$

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

At the end of Section 4 we provide another result concerning this type of random matrices (see Corollary 4.5).

In the proof of (3.10) one uses [12, Corollary 1.3], [13, Theorem 2.1], and (3.11) (see below) instead of (2.1), Lemma 3.3, and (3.6), respectively. The only non-trivial part is proving the claim:

$$\Big\|\sum_{j=1}^n t_jY_{ij}\Big\|_r \leqslant CLr^\beta\Big\|\sum_{j=1}^n t_jY_{ij}\Big\|_2 = CLr^\beta\|t\|_2, \tag{3.11}$$

where C is an absolute constant, and repeating the proof of Theorem 1.1. By (3.11) we get

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

which allows us to obtain a version of (3.6) for $\varphi(t) := \min_{\substack{1\leqslant i\leqslant m,\\ 1\leqslant j\leqslant n}}\|Y_{ij}\|_t$.

Proof of (3.11). It suffices to consider $r = 2k$, where k is an integer. Let $G = (G_j)_{j=1}^n$ be the standard n-dimensional Gaussian vector. Recall that for any $t \in \mathbb{R}^n$ and $r \geqslant 1$ we have $\| \sum_{j=1}^n t_j G_j \|_r = \|t\|_2 \|G_1\|_r \sim \|t\|_2 \sqrt{r} = \sqrt{r} \| \sum_{j=1}^n t_j Y_{ij} \|_2$.

By the assumptions on Y_i and by the fact that $\beta \geqslant \frac{1}{2}$ we get

$$\begin{aligned}
\Big\| \sum_{j=1}^n t_j Y_{ij} \Big\|_{2k}^{2k} &= \sum_{j_1 + \ldots + j_n = k} \binom{2k}{2j_1, \ldots, 2j_n} \mathbb{E} Y_{i1}^{2j_1} \cdots \mathbb{E} Y_{in}^{2j_n} t_1^{2j_1} \cdots t_n^{2j_n} \\
&\leqslant L^{2k} \sum_{j_1 + \ldots + j_n = k} \binom{2k}{2j_1, \ldots, 2j_n} (2j_1)^{2j_1\beta} \cdots (2j_n)^{2j_n\beta} t_1^{2j_1} \cdots t_n^{2j_n} \\
&\leqslant (2k)^{2k\beta - k} L^{2k} \sum_{j_1 + \ldots + j_n = k} \binom{2k}{2j_1, \ldots, 2j_n} (2j_1)^{j_1} \cdots (2j_n)^{j_n} t_1^{2j_1} \cdots t_n^{2j_n} \\
&\leqslant (2k)^{2k\beta - k} (CL)^{2k} \sum_{j_1 + \ldots + j_n = k} \binom{2k}{2j_1, \ldots, 2j_n} \mathbb{E} G_1^{2j_1} \cdots \mathbb{E} G_n^{2j_n} t_1^{2j_1} \cdots t_n^{2j_n} \\
&= (2k)^{2k\beta - k} (CL)^{2k} \Big\| \sum_{j=1}^n t_j G_j \Big\|_{2k}^{2k} \leqslant (2k)^{2k\beta} (CL)^{2k} \Big\| \sum_{j=1}^n t_j Y_{ij} \Big\|_2^{2k},
\end{aligned}$$

which finishes the proof of (3.11). □

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

Let us recall the definition of Gaussian mixtures from [5], where their significance is also described.

Definition 4.1. *A random variable X is called a (centred) Gaussian mixture if there exist a positive random variable r and a standard Gaussian random variable g, independent of r, such that X has the same distribution as rg.*

We will work with matrices of the form $(R_{ij} B_{ij} G_{ij})_{i \leqslant m, j \leqslant n}$ whose entries are Gaussian mixtures. We additionally assume that $R_{ij} = |Z_{ij}|^\gamma$, where $\gamma \geqslant 0$, and that the matrix Z is log-concave and isotropic (considered as a random vector in $\mathbb{R}^{mn}$). It will be clear from the proof, that the corollary below is true also for another type of matrices: $(R_i B_{ij} G_{ij})_{i \leqslant m, j \leqslant n}$, where $R_i = |Z_i|^\gamma$, and $(Z_1, \ldots, Z_m)$ is an arbitrary isotropic log-concave random vector.

Corollary 4.2. *Let $m, n \geqslant 2$, let $\gamma \geqslant 0$, let $B = (B_{ij})$ be a deterministic $m \times n$ matrix, and let $G = (G_{ij})_{i \leqslant m, j \leqslant n}$ be a random matrix whose entries are i.i.d. standard Gaussian variables. Let $X_{ij} = |Z_{ij}|^\gamma B_{ij} G_{ij}$, where $Z = (Z_{ij})_{i \leqslant m, j \leqslant n}$ is a log-concave and isotropic random matrix independent of G. Then for every $p, q \geqslant 2 \vee \frac{1}{\gamma}$ we have*

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

Proof. Inequality (1.1) applied to $a_{ij} = |Z_{ij}|^\gamma B_{ij}$ yields

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

so it suffices to prove that

$$\mathbb{E}\max_{1\leqslant i\leqslant m}\big\|(B_{ij}|Z_{ij}|^{\gamma})_j\big\|_p \leqslant C(p,\gamma)(\log m)^{\gamma}\max_{1\leqslant i\leqslant m}\big\|B_i\big\|_p \tag{4.1}$$

and

$$\mathbb{E}\max_{1\leqslant j\leqslant n}\big\|(B_{ij}|Z_{ij}|^{\gamma})_i\big\|_q \leqslant C(q,\gamma)(\log n)^{\gamma}\max_{1\leqslant j\leqslant n}\big\|B^{(j)}\big\|_q$$

for $p\geqslant 1\vee\frac{1}{\gamma}$. By the symmetry of assumptions we need only to show (4.1).

If $\gamma<1$, then

$$\mathbb{E}\max_{1\leqslant i\leqslant m}\big\|(B_{ij}|Z_{ij}|^{\gamma})_j\big\|_p = \mathbb{E}\max_{1\leqslant i\leqslant m}\big\|(|B_{ij}|^{1/\gamma}|Z_{ij}|)_j\big\|_{p\gamma}^{\gamma} \leqslant \Big(\mathbb{E}\max_{1\leqslant i\leqslant m}\big\|(|B_{ij}|^{1/\gamma}|Z_{ij}|)_j\big\|_{p\gamma}\Big)^{\gamma},$$

and

$$\big\||B_i|^{1/\gamma}\big\|_{p\gamma}^{\gamma} = \big\|B_i\big\|_p,$$

so it suffices to consider only $\gamma\geqslant 1$ (we used here the assumption that $p\geqslant\frac{1}{\gamma}$).

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

$$\begin{aligned}\mathbb{E}\max_{1\leqslant i\leqslant m}\big\|(B_{ij}|Z_{ij}|^{\gamma})_j\big\|_p &= \mathbb{E}\max_{1\leqslant i\leqslant m}\big\|(|B_{ij}|^{1/\gamma}Z_{ij})_j\big\|_{p\gamma}^{\gamma}\\ &\leqslant \Big(\mathbb{E}\max_{1\leqslant i\leqslant m}\big\|(|B_{ij}|^{1/\gamma}Z_{ij})_j\big\|_{p\gamma}^{u\gamma}\Big)^{1/u}\\ &\leqslant \Big(\mathbb{E}\sum_{i=1}^{m}\big\|(|B_{ij}|^{1/\gamma}Z_{ij})_j\big\|_{p\gamma}^{u\gamma}\Big)^{1/u}\\ &\leqslant m^{1/u}\max_{1\leqslant i\leqslant m}\Big(\mathbb{E}\big\|(|B_{ij}|^{1/\gamma}Z_{ij})_j\big\|_{p\gamma}^{u\gamma}\Big)^{1/u}. \end{aligned}\tag{4.2}$$

Fix $i\leqslant m$. By Theorem 2.1 applied to $p=p\gamma$, $q=u\gamma$ (recall that $\gamma\geqslant 1$, so $u\gamma\geqslant 1$, $p\gamma\geqslant p\geqslant 2$), and $Z_j=|B_{ij}|^{1/\gamma}Z_{ij}$ we have

$$\begin{aligned}(Cp\gamma)^{-\gamma}\Big(\mathbb{E}\big\|(|B_{ij}|^{1/\gamma}Z_{ij})_j\big\|_{p\gamma}^{u\gamma}\Big)^{1/u} &\leqslant \left[\mathbb{E}\big\|(|B_{ij}|^{1/\gamma}Z_{ij})_j\big\|_{p\gamma} + \sup_{t\in B_{p'}^n}\Big\|\sum_{j=1}^{n}|B_{ij}|^{1/\gamma}Z_{ij}t_j\Big\|_{u\gamma}\right]^{\gamma}\\ &\leqslant 2^{\gamma-1}\left[\mathbb{E}\big\|(|B_{ij}|^{1/\gamma}Z_{ij})_j\big\|_{p\gamma}^{\gamma} + \sup_{t\in B_{p'}^n}\Big\|\sum_{j=1}^{n}|B_{ij}|^{1/\gamma}Z_{ij}t_j\Big\|_{u\gamma}^{\gamma}\right]. \end{aligned}\tag{4.3}$$

We use (1.2) and the assumption $\mathbb{E}Z_{ij}^2=1$ to estimate the first term in (4.3):

$$\mathbb{E}\Big(\sum_{j=1}^{n}|B_{ij}|^p|Z_{ij}|^{p\gamma}\Big)^{1/p} \leqslant \Big(\sum_{j=1}^{n}|B_{ij}|^p\mathbb{E}|Z_{ij}|^{p\gamma}\Big)^{1/p} \leqslant (C_1p\gamma)^{\gamma}\|B_i\|_p. \tag{4.4}$$

Recall that $B_{p'}^n\subset B_2^n$. We again use (1.2) and the isotropicity of Z_i to estimate the second term in (4.3):

$$\begin{aligned}\sup_{t\in B_{p'}^n}\Big\|\sum_{j=1}^{n}|B_{ij}|^{1/\gamma}Z_{ij}t_j\Big\|_{u\gamma}^{\gamma} &\leqslant (C_1u\gamma)^{\gamma}\sup_{t\in B_2^n}\Big\|\sum_{j=1}^{n}|B_{ij}|^{1/\gamma}Z_{ij}t_j\Big\|_2^{\gamma}\\ &= (C_1u\gamma)^{\gamma}\sup_{t\in B_2^n}\Big(\sum_{j=1}^{n}|B_{ij}|^{2/\gamma}t_j^2\Big)^{\gamma/2}\\ &= (C_1u\gamma)^{\gamma}\max_{1\leqslant j\leqslant n}|B_{ij}| \leqslant (C_1u\gamma)^{\gamma}\|B_i\|_p. \end{aligned}\tag{4.5}$$

We take $u=\log m$ and put together (4.2), (4.3), and (4.4) to get the assertion. □

Remark 4.3. Using Theorem 1.1 instead of (1.1) in the proof above yields a slightly worse estimate:

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

Remark 4.4. It is clear from the proof of Corollary 4.2 that in the case $Z_{ij} = G'_{ij}$, where G'_{ij} are i.i.d. standard Gaussian variables, inequality (4.1) may be slightly improved:

$$\mathbb{E}\max_{1\leqslant i\leqslant m}\big\|(B_{ij}|G'_{ij}|^{\gamma})_j\big\|_p \leqslant C(p,\gamma)(\log m)^{\gamma/2}\max_{1\leqslant i\leqslant m}\big\|B_i\big\|_p \tag{4.6}$$

In order to obtain this improvement one should use $\|\langle t, G_i\rangle\|_{u\gamma} \lesssim \sqrt{u\gamma}\|\langle t, G_i\rangle\|_2$ instead of $\|\langle t, Z_i\rangle\|_{u\gamma} \lesssim u\gamma\|\langle t, Z_i\rangle\|_2$. Thus the assertion of Corollary 4.2 in the case $Z_{ij} = G'_{ij}$ (where G' is independent of G) states that

$$\mathbb{E}\|X\|_{p',q} \leqslant C(p,q,\gamma)\Big[\big(\log m\big)^{\frac{1}{q}+\frac{\gamma}{2}} \max_{1\leqslant i\leqslant m}\big\|B_i\big\|_p + (\log n)^{\gamma/2} \max_{1\leqslant j\leqslant n}\big\|B^{(j)}\big\|_q \\ + (\log m)^{1/q}\mathbb{E}\max_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}|X_{ij}|\Big]. \tag{4.7}$$

Proof of Proposition 1.5. We begin similarly as in the proof of (4.1) (in the case $\gamma = 1$), but we estimate the second term on the right-hand side of (4.3) in a slightly different way, using (1.2):

$$\sup_{t\in B^n_{p'}}\Big\|\sum_{j=1}^n B_{ij}Y_{ij}t_j\Big\|_u \leqslant n^{1/u}\sup_{t\in B^n_\infty}\big(\mathbb{E}\max_{1\leqslant j\leqslant n}|t_jB_{ij}Y_{ij}|^u\big)^{1/u} \leqslant n^{1/u}C_1u\mathbb{E}\max_{1\leqslant j\leqslant n}|B_{ij}Y_{ij}|.$$

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

We may use the result concerning Gaussian mixtures to obtain an estimate similar to the one from Remark 3.4, valid for all $\beta \geqslant \frac{1}{2}$ (not only for $\beta \in [\frac{1}{2}, 1]$), but with a slightly worse constants than in Remark 3.4. The proof is based on the fact, that the variables Y_{ij} satisfying the moment assumption from Remark 3.4 are comparable with certain Gaussian mixtures.

Corollary 4.5. *Let $m, n \geqslant 2$, $\beta \geqslant \frac{1}{2}$, $L > 0$, and 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 $r \geqslant 2$ and any $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$ we have $\frac{r^\beta}{L} \leqslant \|Y_{ij}\|_r \leqslant Lr^\beta$. Then for all $p, q \geqslant 2$,*

$$\mathbb{E}\|X\|_{p',q} \leqslant C(p,q,L,\beta)\Big[\big(\log m\big)^{\beta+\frac{1}{q}} \max_{1\leqslant i\leqslant m}\big\|A_i\big\|_p + (\log n)^{\beta} \max_{1\leqslant j\leqslant n}\big\|A^{(j)}\big\|_q \\ + (\log m)^{1/q}\sqrt{\log(mn)}\mathbb{E}\max_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}|X_{ij}|\Big].$$

Proof. Let G_{ij}, G'_{ij}, $i \leqslant m$, $j \leqslant n$, be i.i.d. standard Gaussian variables. Let (ε_{ij}) be i.i.d. symmetric Bernoulli random variables, independent of G and G'. Note that $Y'_{ij} := |G_{ij}|^{2\beta}\varepsilon_{ij}$ satisfies $\frac{r^\beta}{L'} \leqslant \|Y'_{ij}\|_r \leqslant L'r^\beta$ for all $r \geqslant 2$, with a universal constant L', since $\|G_{ij}\|_s \sim \sqrt{s}$ for $s \geqslant 1$. Let $X' = (X_{ij})$ be the $m \times n$ random matrix with entries $X'_{ij} = A_{ij} Y'_{ij}$. By [14, Lemma 4.7] we know that

$$\frac{1}{C(L,L',\beta)} \mathbb{E}|\!|\!|X'|\!|\!| \leqslant \mathbb{E}|\!|\!|X|\!|\!| \leqslant C(L,L',\beta)\mathbb{E}|\!|\!|X'|\!|\!|$$

for any norm $|\!|\!|\cdot|\!|\!|$ on $m \times n$ real matrices. In particular

$$\mathbb{E}\|X\|_{p',q} \leqslant C(L,\beta)\mathbb{E}\|X'\|_{p',q}, \quad \text{and} \quad \mathbb{E} \max_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} |X'_{ij}| \leqslant C(L,\beta)\mathbb{E} \max_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} |X_{ij}|.$$

Moreover, by $\mathbb{E}|G_{ij}| = \sqrt{2/\pi}$, the Jensen inequality, and (4.7) applied with $\gamma = 2\beta$, we have

$$\begin{aligned}
\mathbb{E}\big\|(X'_{ij})\big\|_{p',q} &= \mathbb{E}\big\|(\varepsilon_{ij}A_{ij}|G'_{ij}|^{2\beta})\big\|_{p',q} = \sqrt{\frac{\pi}{2}}\,\mathbb{E}\big\|\big(\mathbb{E}|G_{ij}|\varepsilon_{ij}A_{ij}|G'_{ij}|^{2\beta}\big)\big\|_{p',q} \\
&\leqslant \sqrt{\frac{\pi}{2}}\,\mathbb{E}\big\|\big(|G_{ij}|\varepsilon_{ij}A_{ij}|G'_{ij}|^{2\beta}\big)\big\|_{p',q} = \sqrt{\frac{\pi}{2}}\,\mathbb{E}_X\mathbb{E}_G\big\|(A_{ij}G_{ij}|G'_{ij}|^{2\beta})\big\|_{p',q} \\
&\leqslant C(p,q)\bigg((\log m)^{\beta+\frac{1}{q}} \max_{1\leqslant i\leqslant m}\big\|A_i\big\|_p + (\log n)^{\beta} \max_{1\leqslant j\leqslant n}\big\|A^{(j)}\big\|_q \\
&\qquad\qquad + (\log m)^{1/q}\mathbb{E}\max_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} |A_{ij}G_{ij}|\cdot|G'_{ij}|^{2\beta}\bigg) \\
&\leqslant C(p,q)\bigg((\log m)^{\beta+\frac{1}{q}} \max_{1\leqslant i\leqslant m}\big\|A_i\big\|_p + (\log n)^{\beta} \max_{1\leqslant j\leqslant n}\big\|A^{(j)}\big\|_q \\
&\qquad\qquad + (\log m)^{1/q}\mathbb{E}\max_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} |G_{ij}|\mathbb{E}\max_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} |X'_{ij}|\bigg),
\end{aligned}$$

which yields the assertion, since $\mathbb{E}\max_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}} |G_{ij}| \sim \sqrt{\log(mn)}$. □

5 The case of unconditional entries

Proof of Corollary 1.6. Proceeding like in the proof of (1.4), we prove using (1.1) that

$$\mathbb{E}\|(a_{ij}G_{ij})\|_{p',q} \leqslant C(p,q)\bigg((\log m)^{1/q}\mathbb{E}\max_{1\leqslant i\leqslant m}\|(a_{ij}G_{ij})_j\|_p + \mathbb{E}\max_{1\leqslant j\leqslant n}\|(a_{ij}G_{ij})_i\|_q\bigg). \tag{5.1}$$

where G_{ij} are i.i.d. standard Gaussian variables.

Since X is unconditional, it has the same distribution as the matrix $(\varepsilon_{ij}X_{ij})_{i\leqslant m, j\leqslant n}$, where ε_{ij} are i.i.d. symmetric Bernoulli variables independent of X. Let G_{ij} be i.i.d. standard Gaussian variables independent of X and $(\varepsilon_{ij})_{i\leqslant m, j\leqslant n}$. Using $\mathbb{E}|G_{ij}| = \sqrt{2/\pi}$, the Jensen inequality, and (5.1) (to estimate the mean with respect to G) we get

$$\begin{aligned}
\mathbb{E}\big\|(X_{ij})\big\|_{p',q} &= \mathbb{E}\big\|(\varepsilon_{ij}X_{ij})\big\|_{p',q} = \sqrt{\frac{\pi}{2}}\,\mathbb{E}\big\|\big(\varepsilon_{ij}X_{ij}\mathbb{E}|G_{ij}|\big)\big\|_{p',q} \\
&\leqslant \sqrt{\frac{\pi}{2}}\,\mathbb{E}\big\|\big(\varepsilon_{ij}X_{ij}|G_{ij}|\big)\big\|_{p',q} = \sqrt{\frac{\pi}{2}}\,\mathbb{E}_X\mathbb{E}_G\big\|(X_{ij}G_{ij})\big\|_{p',q} \\
&\leqslant C(p,q)\bigg((\log m)^{1/q}\mathbb{E}_X\mathbb{E}_G\max_{1\leqslant i\leqslant m}\|(X_{ij}G_{ij})_j\|_p + \mathbb{E}_X\mathbb{E}_G\max_{1\leqslant j\leqslant n}\|(X_{ij}G_{ij})_i\|_q\bigg),
\end{aligned}$$

We use (4.6) with $\gamma = 1$ to get the assertion. □

Remark 5.1. Using Theorem 1.1 instead of (1.1) in the proof above yields a slightly weaker estimate in Corollary 1.6:

$$\mathbb{E}\|X\|_{p',q} \leqslant C(p,q)\Bigg((\log m)^{\frac{3}{2}+\frac{1}{q}}\mathbb{E}\max_{1\leqslant i\leqslant m}\Big(\sum_{j=1}^{n}|X_{ij}|^p\Big)^{1/p} + \sqrt{\log n}\,\mathbb{E}\max_{1\leqslant j\leqslant n}\Big(\sum_{i=1}^{m}|X_{ij}|^q\Big)^{1/q}\Bigg). \quad (5.2)$$

References

[1] 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

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

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

[4] 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

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

[6] 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

[7] 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

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

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

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

[11] 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

[12] R. Latała and M. Strzelecka, *Comparison of weak and strong moments for vectors with independent coordinates*, Mathematika **64** (2018), no. 1, 211–229. MR-3778221

[13] 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

[14] 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

[15] 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

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

Acknowledgments. I would like to thank Rafał Latała for suggestions which helped me to make the presentation clearer and more reader-friendly.