Chevet type inequality and norms of sub-matrices

Radosław Adamczak[1], Rafał Latała[1], Alexander E. Litvak[2], Alain Pajor[3], Nicole Tomczak-Jaegermann[4]

Abstract

We prove a Chevet type inequality which gives an upper bound for the norm of an isotropic log-concave unconditional random matrix in terms of expectation of the supremum of "symmetric exponential" processes compared to the Gaussian ones in the Chevet inequality. This is used to give sharp upper estimate for a quantity $\Gamma_{k,m}$ that controls uniformly the Euclidean operator norm of the sub-matrices with k rows and m columns of an isotropic log-concave unconditional random matrix. We apply these estimates to give a sharp bound for the Restricted Isometry Constant of a random matrix with independent log-concave unconditional rows. We also show that our Chevet type inequality does not extend to general isotropic log-concave random matrices.

AMS 2010 Classification: Primary 52A23, 46B06, 46B09, 60B20; Secondary 15B52, 60E15, 94B75.

Keywords: log-concave random vectors, concentration inequalities, deviation inequalities, random matrices, Compressive Sensing, Restricted Isometry Property, Chevet inequality

---

[1]Research partially supported by MNiSW Grant no. N N201 397437 and the Foundation for Polish Science.

[2]Research partially supported by the E.W.R. Steacie Memorial Fellowship.

[3]Research partially supported by the ANR project ANR-08-BLAN-0311-01.

[4]This author holds the Canada Research Chair in Geometric Analysis.

1 Introduction

Let n, N be positive integers. Let $K \subset \mathbb{R}^N$ and $L \subset \mathbb{R}^n$ be origin symmetric convex bodies, $\|\cdot\|_K$ and $\|\cdot\|_L$ be the corresponding gauges on $\mathbb{R}^N$ and $\mathbb{R}^n$, that is the norms for which K and L are the unit balls.

To shorten the notation we write $\|\Gamma : K \to L\|$ for the operator norm of a linear operator $\Gamma : (\mathbb{R}^N, \|\cdot\|_K) \to (\mathbb{R}^n, \|\cdot\|_L)$. In particular, $\|\Gamma : K \to B_2^N\|$ will denote the operator norm of Γ considered as a linear operator from $(\mathbb{R}^N, \|\cdot\|_K)$ to ℓ_2^N, where ℓ_2^N is $\mathbb{R}^N$ equipped with the canonical Euclidean norm, whose unit ball is B_2^N; similarly for $\|\Gamma : B_2^n \to L\|$. Also note that the dual normed space $(\mathbb{R}^N, \|\cdot\|_K)^*$ of $(\mathbb{R}^N, \|\cdot\|_K)$ may be identified (via the canonical inner product) with $(\mathbb{R}^N, \|\cdot\|_{K^\circ})$, where K° denotes the polar of K (see the next section for all definitions). The canonical basis on $\mathbb{R}^d$ is denoted by $\{e_i\}_{1 \le i \le d}$.

Let $(g_i)_{1 \le i \le \max(n,N)}$ be i.i.d. standard Gaussian random variables that is centered Gaussian variables with variance one, and Γ be a Gaussian matrix whose entries are i.i.d. standard Gaussian. Then one side of the Chevet inequality ([7], see also [8] for sharper constants) states that

$$\begin{aligned} \mathbb{E}\|\Gamma : K \to L\| \le C\|\mathrm{Id} : K \to B_2^N\| \cdot \mathbb{E}\left\|\sum_{i=1}^n g_i e_i\right\|_L \\ + C\|\mathrm{Id} : B_2^n \to L\| \cdot \mathbb{E}\left\|\sum_{i=1}^N g_i e_i\right\|_{K^\circ}, \end{aligned} \tag{1}$$

where Id stays for the formal identity operator and C is an absolute constant. This inequality plays an important role in Probability in Banach Spaces and in Asymptotic Geometric Analysis ([5, 16]).

We say that a random matrix $\Gamma = (\gamma_{ij})$ is *isotropic* if all entries (γ_{ij}) are uncorrelated centered with variance one. It is *log-concave* if the joint distribution of the γ_{ij}'s has a density which is log-concave on its support. Finally we say that the matrix Γ is *unconditional* if for any choice of signs (ε_{ij}) the matrices Γ and $(\varepsilon_{ij}\gamma_{ij})$ have the same distribution. There are similar definitions for random vectors.

In Theorem 3.1 we prove that an inequality similar to the Chevet inequality (1) holds for any isotropic log-concave unconditional random matrix Γ.

Namely, we show that for such a matrix one has

$$\begin{aligned}&\mathbb{E}\|\Gamma : K \to L\|\\&\leq C\left(\|\mathrm{Id} : K \to B_2^N\| \cdot \mathbb{E}\left\|\sum_{i=1}^n E_i e_i\right\|_L + \|\mathrm{Id} : B_2^n \to L\| \cdot \mathbb{E}\left\|\sum_{i=1}^N E_i e_i\right\|_{K^\circ}\right),\end{aligned}$$

where E_i's denote i.i.d. random variables with symmetric exponential distribution with variance 1. Moreover, in Corollary 3.2 we provide the corresponding probability estimates.

A result from [9] of the second named author of this article states that if $X = (X_1, \ldots, X_d)$ is an isotropic log-concave unconditional random vector in $\mathbb{R}^d$ and if $Y = (E_1, \ldots, E_d)$, where $E_1, \ldots, E_d$ are i.i.d. symmetric exponential random variables, then for any norm $\|\cdot\|$ on $\mathbb{R}^d$ one has

$$\mathbb{E}\|X\| \leq C\ \mathbb{E}\|Y\|, \tag{2}$$

where C is an absolute constant.

The proof of our Chevet type inequality consists of two steps. First, using the comparison (2), we reduce the case of a general isotropic log-concave unconditional random matrix A to the case of an exponential random matrix, i.e. the matrix whose entries are i.i.d. standard symmetric exponential random variables. The second step uses Talagrand's result ([14]) on relations between some random processes associated to the symmetric exponential distribution and so-called γ_p functionals.

We apply our inequality of Chevet type to obtain sharp uniform bounds on norms of sub-matrices of isotropic log-concave unconditional random matrices Γ. More precisely, for any subsets $J \subset \{1, \ldots, n\}$ and $I \subset \{1, \ldots, N\}$ denote by $\Gamma(J, I)$ the sub-matrix of Γ consisting of the rows indexed by elements from J and the columns indexed by elements from I. Given $k \leq n$ and $m \leq N$ define the parameter $\Gamma_{k,m}$ by

$$\Gamma_{k,m} = \sup \|\Gamma(J, I)\ :\ \ell_2^m \to \ell_2^k\|,$$

where the supremum is taken over all subsets $J \subset \{1, \ldots, n\}$ and $I \subset \{1, \ldots, N\}$ with cardinalities $|J| = k$, $|I| = m$. That is, $\Gamma_{k,m}$ is the maximal operator norm of a sub-matrix of Γ with k rows and m columns. We prove that

$$\Gamma_{k,m} \leq C\left(\sqrt{m}\ln\left(\frac{3N}{m}\right) + \sqrt{k}\ln\left(\frac{3n}{k}\right)\right),$$

with high probability. This estimate is sharp up to absolute constants.

Furthermore, we provide applications of this result to the Restricted Isometry Property (RIP) of a matrix with independent isotropic log-concave unconditional random rows. We give sharp estimate for the restricted isometry constant of such matrices.

It is well known and follows from Talagrand's majorizing measure theorem (see [15]) that if $X = (X_1, \ldots, X_d)$ is a centered sub-gaussian random vector in $\mathbb{R}^d$ with parameter $\alpha > 0$, that is, all coordinates X_i are centered and for any $x \in \mathbb{R}^d$ of Euclidean norm 1, any $t > 0$, $\mathbb{P}(|\sum x_i X_i| \geq t) \leq 2\exp(-t^2/\alpha^2)$, then for any norm $\|\cdot\|$ on $\mathbb{R}^d$, one has

$$\mathbb{E}\|X\| \leq C\alpha\, \mathbb{E}\|Y\|, \tag{3}$$

where $Y = (g_1, \ldots, g_d)$ and $C > 0$ is an absolute constant.

It is interesting to view both inequalities (2) and (3) in parallel. There are both based on majorizing measure theorems of Talagrand; inequality (3) states that the expectation of the norm of a sub-gaussian vector is up to a multiplicative constant, dominated by its Gaussian replica. So Gaussian vectors are almost maximizers. To which class of random vectors does inequality (2) correspond? Since in many geometric and probabilistic inequalities involving isotropic log-concave vectors, Gaussian and exponential vectors are the extreme cases, it was naturally conjectured that the expectation of the norm of isotropic log-concave vector is similarly dominated by the corresponding expectation of the norm of an exponential random vector. This conjecture would have many applications. For instance the estimate of $\Gamma_{k,m}$ above would extend to general log-concave random matrices, which is open (see [1]).

We show that this is not the case. Namely, in Theorem 5.1 we prove that for any $d \geq 1$, there exists an isotropic log-concave random vector $X \in \mathbb{R}^d$ and a norm $\|\cdot\|$ on $\mathbb{R}^d$ such that

$$\mathbb{E}\|X\| \geq c\sqrt{\ln d}\, \mathbb{E}\|Y\|, \tag{4}$$

where Y is of "symmetric exponential" type and c is a positive universal constant. Similarly we show that our Chevet type inequality does not extend to the setting of general log-concave random matrices (non unconditional). In fact it would be interesting to find the best dependence on the dimension in the reverse inequality to (4). More precisely, to solve the following problem.

Problem. *Find tight (in terms of dimension d) estimates for the following quantity*

$$C(d) = \sup_{\|\cdot\|} \sup_{X} \frac{\mathbb{E}\|X\|}{\mathbb{E}\|Y\|},$$

where $Y = (E_1, \ldots, E_d)$ and the supremum is taken over all norms $\|\cdot\|$ on $\mathbb{R}^d$ and all isotropic log-concave random vectors $X \in \mathbb{R}^d$.

Theorem 5.1 and Remark 2 following it show that $c\sqrt{\ln d} \le C(d) \le C\sqrt{d}$ for some absolute positive constants c and C.

The results on norms of sub-matrices and applications were partially announced in [2]. For the related estimates in the non-unconditional case, see [1].

The paper is organized as follows. In the next section we introduce notation and quote known results which will be used in the sequel. In Section 3, Theorem 3.1, we prove the Chevet type inequality (and corresponding probability estimates) for unconditional log-concave matrices. In remarks following Theorem 3.1 we discuss its sharpness showing that in general one can't expect the lower bound of the same order and providing a relevant lower bound. In Section 4 we apply our Chevet type inequality to obtain sharp uniform estimates for norms of sub-matrices. Then we apply the results to the RIP. Section 5 is devoted to examples showing that one can't drop the assumption of unconditionality in the comparison theorem of the second named author and in our Chevet type inequality. Finally, in Section 6, we present a direct approach to uniform estimates of norms of sub-matrices, which does not involve Chevet type inequalities and γ_p functionals, but is based only on tail estimates for suprema of linear combinations of independent exponential variables and on a chaining argument in the spirit of [4].

2 Notation and Preliminaries

By $|\cdot|$ and $\langle \cdot, \cdot \rangle$ we denote the canonical Euclidean norm and the canonical inner product on $\mathbb{R}^d$. The canonical basis of $\mathbb{R}^d$ is denoted by $e_1, \ldots, e_d$.

As usual, $\|\cdot\|_p$, $1 \le p \le \infty$, denotes the ℓ_p-norm, i.e. for every $x = (x_i)_{i=1}^d \in \mathbb{R}^d$

$$\|x\|_p = \left(\sum_{i=1}^d |x_i|^p \right)^{1/p} \quad \text{for} \quad p < \infty \quad \text{and} \quad \|x\|_\infty = \sup_{i \le d} |x_i|$$

and $\ell_p^d = (\mathbb{R}^d, \|\cdot\|_p)$. The unit ball of ℓ_p^d is denoted by B_p^d. For a non-empty set $T \subset \mathbb{R}^d$ we write $\mathrm{diam}_p(T)$ to denote the diameter of T with respect to the ℓ_p-norm.

For an origin symmetric convex body $K \subset \mathbb{R}^d$, the Minkowski functional of K is

$$\|x\|_K = \inf\{\lambda > 0 \mid x \in \lambda K\},$$

i.e. the norm, whose unit ball is K. The polar of K is

$$K^\circ = \{x \mid \langle x, y\rangle \leq 1 \quad \text{for all} \quad y \in K\}.$$

Note that K° is the unit ball of the space dual to $(\mathbb{R}^d, \|\cdot\|_K)$.

Given an $n \times N$ matrix Γ and origin symmetric convex bodies $K \subset \mathbb{R}^N$, $L \subset \mathbb{R}^n$ we denote by

$$\|\Gamma : K \to L\|$$

the operator norm of Γ from $(\mathbb{R}^N, \|\cdot\|_K)$ to $(\mathbb{R}^n, \|\cdot\|_L)$. We also denote

$$R(K) = \|\mathrm{Id} : K \to B_2^N\|, \quad R(L^\circ) = \|\mathrm{Id} : B_2^n \to L\| = \|\mathrm{Id} : L^\circ \to B_2^n\|,$$

where Id denotes the formal identity $\mathbb{R}^N \to \mathbb{R}^N$ or $\mathbb{R}^n \to \mathbb{R}^n$.

Given a subset $K \subset \mathbb{R}^d$ the convex hull of K is denoted by $\mathrm{conv}(K)$.

A random vector $X = (X_1, \ldots, X_N)$ is called unconditional if for every sequence of signs $\varepsilon_1, \ldots, \varepsilon_N$, the law of $(\varepsilon_1 X_1, \ldots, \varepsilon_N X_N)$ is the same as the law of X.

A random vector X in $\mathbb{R}^n$ is called isotropic if

$$\mathbb{E}\langle X, y\rangle = 0, \quad \mathbb{E}\,|\langle X, y\rangle|^2 = \|y\|_2^2 \quad \text{for all } y \in \mathbb{R}^n,$$

in other words, if X is centered and its covariance matrix $\mathbb{E}\, X \otimes X$ is the identity.

A random vector X in $\mathbb{R}^n$ with full dimensional support is called log-concave if it has a log-concave density. Notice that all isotropic vectors have full dimensional support.

By E_i, E_{ij} we denote independent symmetric exponential random variables with variance 1 (i.e. with the density $2^{-1/2}\exp(-\sqrt{2}\,|x|)$). By g_i, g_{ij} we denote standard independent $\mathcal{N}(0,1)$ Gaussian random variables. The $n \times N$ random matrix with entries g_{ij} will be called the Gaussian matrix, the $n \times N$ random matrix with entries E_{ij} will be called the exponential random matrix. Similarly, the vectors $G = (g_1, \ldots, g_d)$ and $Y = (E_1, \ldots, E_d)$ are called Gaussian and exponential random vectors.

In the sequel we often consider $n \times N$ matrices as vectors in $\mathbb{R}^d$ with $d = nN$ and the inner product defined by

$$\langle A, B\rangle = \sum_{i,j} a_{ij} b_{ij}$$

for $A = (a_{ij})$, $B = (b_{ij})$. Clearly, the corresponding Euclidean structure is given by Hilbert-Schmidt norm of a matrix:

$$|A| = \|A\|_2 = \left(\sum_{i,j} |a_{ij}|^2\right)^{1/2}.$$

In this notation we have $\|A\|_\infty = \max_{i,j} |a_{ij}|$. We say that such a matrix A is isotropic/log-concave/unconditional if it is isotropic/log-concave/unconditional as a vector in $\mathbb{R}^d$, $d = nN$ (cf. the definition given in the introduction).

Given $x \in \mathbb{R}^N$ and $y \in \mathbb{R}^n$, denote by $x \otimes y = yx^\top$ the matrix $\{y_i x_j\}_{ij}$, i.e. the matrix corresponding to the linear operator defined by

$$x \otimes y\ (z) = \langle z, x\rangle y.$$

Then, for an $n \times N$ matrix $\Gamma = (\gamma_{ij})$,

$$\|\Gamma : K \to L\| = \sup_{x\in K} \sup_{y \in L^\circ} \sum_{i,j} \gamma_{ij} x_j y_i = \sup_T \langle \Gamma, x \otimes y\rangle,$$

where the latter supremum is taken over

$$T = K \otimes L^\circ = \{x \otimes y \ :\ x \in K,\ y \in L^\circ\}.$$

We will use the letters $C, C_0, C_1, \ldots$, $c, c_0, c_1, \ldots$ to denote positive absolute constants whose values may differ at each occurrence. We also use the notation $F \approx G$ if there are two positive absolute constants C and c such that $c\,G \le F \le C\,G$.

Now we state some results which will be used in the sequel. We start with the following lemma, which provides asymptotically sharp bounds on the norm of the exponential matrix considered as an operator $\ell_1^N \to \ell_1^n$. We will use it in our examples on sharpness of some estimates.

Lemma 2.1. *Let $\Gamma = (E_{ij})_{i\le n, j \le N}$. Then*

$$\mathbb{E}\ \|\Gamma\ :\ \ell_1^N \to \ell_1^n\| \approx n + \ln N.$$

Proof. First note

$$\|\Gamma\ :\ \ell_1^N \to \ell_1^n\| = \max_{j\le N} \sum_{i=1}^n |E_{ij}|. \tag{5}$$

By the Chebyshev inequality for every $j \le N$ we have

$$\mathbb{P}\Big(\sum_{i=1}^n |E_{ij}| \ge t\Big) \le \exp\Big(-\frac{t}{2}\Big)\ \mathbb{E}\exp\Big(\frac{1}{2}\sum_{i=1}^n |E_{ij}|\Big) \le C^n \exp\Big(-\frac{t}{2}\Big)$$

for some absolute constant $C > 0$. Hence the union bound and integration by parts gives

$$\mathbb{E}\ \|\Gamma\ :\ \ell_1^N \to \ell_1^n\| \le C\,(n + \ln N)\,.$$

On the other hand, by (5)

$$\mathbb{E}\|\Gamma\ :\ \ell_1^N \to \ell_1^n\| \ge \mathbb{E}\sum_{i=1}^n |E_{i1}| = n/\sqrt{2}$$

and

$$\mathbb{E}\|\Gamma\ :\ \ell_1^N \to \ell_1^n\| \ge \mathbb{E}\max_{j\le N} |E_{1j}| \approx 1 + \ln N$$

(the last equivalence is well-known and follows from direct computations). This completes the proof. □

The next theorem is a comparison theorem from [9].

Theorem 2.2. *Let X be an isotropic log-concave unconditional random vector in $\mathbb{R}^d$ and $Y = (E_1, \dots, E_d)$ be an exponential random vector. Let $\|\cdot\|$ be a norm on $\mathbb{R}^d$. Then*

$$\mathbb{E}\|X\| \le C\ \mathbb{E}\|Y\|,$$

where C is an absolute positive constant. Moreover, for every $t \ge 1$,

$$\mathbb{P}(\|X\| \ge t) \le C\ \mathbb{P}(\|Y\| \ge t/C).$$

Remark. The condition "X is unconditional" cannot be omitted in Theorem 2.2. We show an example proving that in Section 5.

We will also use two Talagrand's results on behavior of random processes. The first one characterizes suprema of Gaussian and exponential processes in terms of the γ_q functionals.

For a metric space (E,ρ) and $q>0$ we define the γ_q functional as

$$\gamma_q(E,\rho)=\inf_{(A_s)_{s=0}^\infty}\sup_{x\in E}\sum_{s=0}^{\infty}2^{s/q}\,\mathrm{dist}(x,A_s),$$

where the infimum is taken over all sequences $(A_s)_{s=0}^\infty$ of subsets of E, such that $|A_0|=1$ and $|A_s|\le 2^{2^s}$ for $s\ge 1$.

The following theorem combines Theorems 2.1.1 and 5.2.7 in [15].

Theorem 2.3. *Let $T\subset\mathbb{R}^d$ and ρ_q denote the ℓ_q metric. Then*

$$\mathbb{E}\sup_{z\in T}\sum_{i=1}^{d} z_i g_i\approx\gamma_2(T,\rho_2)\quad\text{and}\quad\mathbb{E}\sup_{z\in T}\sum_{i=1}^{d} z_i E_i\approx\gamma_2(T,\rho_2)+\gamma_1(T,\rho_\infty).$$

We will also use Talagrand's result on the deviation of supremum of exponential processes from their averages. It follows by Talagrand's two level concentration for product exponential measure ([13]).

Theorem 2.4. *Let T be a compact subset of $\mathbb{R}^d$. Then for any $t\ge 0$,*

$$\mathbb{P}\left(\sup_{z\in T}\left|\sum_{i=1}^{d} z_iE_i\right|\ge\mathbb{E}\sup_{z\in T}\left|\sum_{i=1}^{d} z_iE_i\right|+t\right)\le\exp\left(-c\min\left\{\frac{t^2}{a^2},\frac{t}{b}\right\}\right),$$

where $a=\sup_{z\in T}|z|$, $b=\sup_{z\in T}\|z\|_\infty$.

3 Chevet type inequality

Theorem 3.1. *Let Γ be an isotropic log-concave unconditional random $n\times N$ matrix. Let $K\subset\mathbb{R}^N$, $L\subset\mathbb{R}^n$ be origin symmetric convex bodies. Then*

$$\begin{aligned}&\mathbb{E}\|\Gamma:K\to L\|\\ &\le C\left(\|\mathrm{Id}:K\to B_2^N\|\cdot\mathbb{E}\left\|\sum_{i=1}^{n}E_ie_i\right\|_L+\|\mathrm{Id}:B_2^n\to L\|\cdot\mathbb{E}\left\|\sum_{i=1}^{N}E_ie_i\right\|_{K^\circ}\right).\end{aligned}$$

Remarks. 1. One of the most important examples of matrices satisfying the hypothesis of Theorem 3.1 are matrices whose rows (or columns) are independent isotropic log-concave unconditional random vectors. Indeed, it is easy to see that if X, Y are independent isotropic log-concave random vectors then so is (X,Y). If X, Y are in addition unconditional then clearly (X,Y) is unconditional. Therefore, if rows (or columns) of a matrix Γ are independent isotropic log-concave random vectors then Γ is isotropic log-concave. If rows (resp. columns) are in addition unconditional, then so is Γ. We will use it in Section 4.

2. A particular case of matrices from the previous remark are matrices Γ, whose entries are symmetric i.i.d. isotropic log-concave variables γ_{ij}. It is natural to ask if the Chevet type inequality holds for such matrices with γ_i's instead of E_i's, where γ_i, $i \leq \max\{n, N\}$, are independent copies of γ_{11}. The answer is NO – we indeed need to pass to the exponential variables. To show this, let variables γ_i and γ_{ij} be as above and, in addition be bounded. Let $K = B_1^N$ and $L = B_2^n$. Then for N large enough one has

$$\mathbb{E}\|\Gamma : K \to L\| = \mathbb{E} \max_{j \leq N} \left(\sum_{i \leq n} \gamma_{ij}^2 \right)^{1/2} \approx \sqrt{n} \|\gamma_1\|_\infty$$

and

$$\|\mathrm{Id} : K \to B_2^N\| \cdot \mathbb{E} \left\| \sum_{i=1}^n \gamma_i e_i \right\|_L + \|\mathrm{Id} : B_2^n \to L\| \cdot \mathbb{E} \left\| \sum_{i=1}^N \gamma_i e_i \right\|_{K^\circ}$$

$$= \mathbb{E} \left(\sum_{i \leq n} \gamma_i^2 \right)^{1/2} + \mathbb{E} \max_{i \leq N} |\gamma_i| \approx \sqrt{n} + \|\gamma_1\|_\infty.$$

Thus, if we choose variables satisfying $\|\gamma_1\|_\infty = \sqrt{n}$ then the ratio between two quantities will be of the order $\sqrt{n}$.

3. In fact in the Gaussian case the equivalence holds in the Chevet inequality. However, in the log-concave case one cannot hope for the reverse inequality even in the case of exponential matrix and unconditional convex bodies K, L. Indeed, consider the matrix $\Gamma = (E_{ij})$ as an operator $\ell_1^N \to \ell_1^n$, i.e. $K = B_1^N$, $L = B_1^n$. By Lemma 2.1

$$\mathbb{E}\|\Gamma \ : \ \ell_1^N \to \ell_1^n\| \approx n + \ln N.$$

On the other hand, the right hand side term in Theorem 3.1 is

$$C \left(\mathbb{E} \sum_{i=1}^n |E_i| + \sqrt{n}\ \mathbb{E} \max_{i \leq N} |E_i| \right) \approx n + \sqrt{n}\ \ln(2N).$$

Thus, if $N \geq e^n$ then the ratio between the right hand side and the left hand side is of the order $\sqrt{n}$.

4. The following weak form of a reverse inequality holds for the exponential matrix $\Gamma = (E_{ij})_{i\leq n, j\leq N}$:

$$\mathbb{E}\|\Gamma : K \to L\| \geq \frac{1}{2}\left(\max_{i\leq N}\|e_i\|_{K^\circ} \cdot \mathbb{E}\left\|\sum_{i=1}^{n} E_i e_i\right\|_L + \max_{i\leq n}\|e_i\|_L \cdot \mathbb{E}\left\|\sum_{i=1}^{N} E_i e_i\right\|_{K^\circ}\right).$$

Indeed, fix $1 \leq \ell \leq N$ and take $x \in K$ such that $\|e_\ell\|_{K^\circ} = |\langle e_\ell, x\rangle| = |x_\ell|$. Then

$$\begin{aligned}\mathbb{E}\|\Gamma : K \to L\| &\geq \mathbb{E}\|\Gamma x\|_L = \mathbb{E}\Big\|\sum_{i\leq n, j\leq N} E_{ij} x_j e_i\Big\|_L \geq \mathbb{E}\Big\|\sum_{i\leq n} E_{i\ell} x_\ell e_i\Big\|_L \\ &= |x_\ell|\ \mathbb{E}\Big\|\sum_{i\leq n} E_i e_i\Big\|_L = \|e_\ell\|_{K^\circ}\ \mathbb{E}\Big\|\sum_{i\leq n} E_i e_i\Big\|_L.\end{aligned}$$

This shows that

$$\mathbb{E}\|\Gamma : K \to L\| \geq \max_{i\leq N}\|e_i\|_{K^\circ}\ \mathbb{E}\left\|\sum_{i\leq n} E_i e_i\right\|_L$$

and by duality we have

$$\mathbb{E}\|\Gamma : K \to L\| = \mathbb{E}\|\Gamma^T : L^\circ \to K^\circ\| \geq \max_{i\leq n}\|e_i\|_L\ \mathbb{E}\left\|\sum_{i\leq N} E_i e_i\right\|_{K^\circ}.$$

5. As in Theorem 2.2, the condition "Γ is unconditional" cannot be omitted in Theorem 3.1. We show an example proving that in Section 5.

Proof of Theorem 3.1. First note that considering the matrix Γ as a vector in $\mathbb{R}^{nN}$ and applying Theorem 2.2, we obtain that it is enough to prove Theorem 3.1 for the case of the exponential matrix.

From now we assume that $\Gamma = (E_{ij})$. Denote as before $T = K \otimes L^\circ = \{x \otimes y : x \in K,\ y \in L^\circ\}$. Then by Theorem 2.3

$$\mathbb{E}\|\Gamma : K \to L\| = \mathbb{E}\sup_{x\in K}\sup_{y\in L^\circ}\sum_{i,j} E_{ij} x_j y_i = \mathbb{E}\sup_T \langle \Gamma, x\otimes y\rangle \approx \gamma_2(T,\rho_2) + \gamma_1(T,\rho_\infty)$$

and

$$\mathbb{E}\left\|\sum_{i=1}^{n} E_i e_i\right\|_L \approx \gamma_2(L^\circ,\rho_2)+\gamma_1(L^\circ,\rho_\infty),$$

$$\mathbb{E}\left\|\sum_{i=1}^{N} E_i e_i\right\|_{K^\circ} \approx \gamma_2(K,\rho_2)+\gamma_1(K,\rho_\infty).$$

Thus it is enough to show that

$$\gamma_2(T,\rho_2) \le C\left(R(K)\gamma_2(L^\circ,\rho_2)+R(L^\circ)\gamma_2(K,\rho_2)\right) \tag{6}$$

and

$$\gamma_1(T,\rho_\infty) \le C\left(R(K)\gamma_1(L^\circ,\rho_\infty)+R(L^\circ)\gamma_1(K,\rho_\infty)\right). \tag{7}$$

Inequality (6) is the Chevet inequality for the Gaussian case. Indeed by Theorem 2.3

$$\gamma_2(T,\rho_2) \approx \mathbb{E}\sup_{z\in T}\sum_{i,j} z_{ij}g_{ij} = \mathbb{E}\|(g_{ij}) \ : K \to L\|$$

and

$$R(K)\gamma_2(L^\circ,\rho_2)+R(L^\circ)\gamma_2(K,\rho_2) \approx R(K)\mathbb{E}\sup_{z\in L^\circ}\sum_{i=1}^{n} z_i g_i + R(L^\circ)\mathbb{E}\sup_{z\in K}\sum_{i=1}^{N} z_i g_i.$$

In fact we could prove (6) without the use of the Chevet inequality, but by the chaining argument similar to the one used for the proof of (7) below (cf. also [11]).

It remains to prove inequality (7).

Let $A_s \subset K$ and $B_s \subset L^\circ$, $s \ge 0$, be admissible sequences of sets (i.e., with $|A_0| = |B_0| = 1$, $|A_s|, |B_s| \le 2^{2^s}$ for $s \ge 1$). Define an admissible sequence $(C_s)_{s\ge 0}$ by $C_0 = \{0\}$ and

$$C_s = A_{s-1}\otimes B_{s-1} \subset K \otimes L^\circ, \quad s \ge 1.$$

Note that for all $x, \tilde{x} \in K$ and for all $y, \tilde{y} \in L^\circ$ one has

$$\begin{aligned}\|x\otimes y - \tilde{x}\otimes\tilde{y}\|_\infty &\le \|x\|_\infty \cdot \|y-\tilde{y}\|_\infty + \|\tilde{y}\|_\infty \cdot \|x-\tilde{x}\|_\infty \\ &\le R(K)\|y-\tilde{y}\|_\infty + R(L^\circ)\|x-\tilde{x}\|_\infty.\end{aligned}$$

Therefore

$$\begin{aligned}\gamma_1(K\otimes L^\circ,\rho_\infty) \leq &\sup_{x\otimes y\in K\otimes L^\circ}\sum_{s=0}^{\infty}2^s\mathrm{dist}\,(x\otimes y, C_s)\\ \leq &R(K)\sup_{y\in L^\circ}\left(\|y\|_\infty+\sum_{s=1}^{\infty}2^s\mathrm{dist}\,(y,B_{s-1})\right)\\ &+R(L^\circ)\sup_{x\in K}\left(\|x\|_\infty+\sum_{s=1}^{\infty}2^s\mathrm{dist}\,(x,A_{s-1})\right).\end{aligned}$$

Taking the infimum over all admissible sequences (A_s) and (B_s) we get

$$\begin{aligned}&\gamma_1(K\otimes L^\circ,\rho_\infty)\\ &\quad\leq R(K)\,(\mathrm{diam}_\infty L^\circ+2\gamma_1(L^\circ,\rho_\infty))+R(L^\circ)\,(\mathrm{diam}_\infty K+2\gamma_1(K,\rho_\infty))\\ &\quad\leq 4R(K)\gamma_1(L^\circ,\rho_\infty)+4R(L^\circ)\gamma_1(K,\rho_\infty),\end{aligned}$$

where in the last inequality we used the fact that the diameter is clearly dominated by doubled γ_1 functional. □

Corollary 3.2. *Let Γ, K, L be as in Theorem 3.1. Then for every $t>0$,*

$$\|\Gamma : K\to L\|\leq C\left(R(K)\cdot\mathbb{E}\left\|\sum_{i=1}^{n}E_ie_i\right\|_L+R(L^\circ)\cdot\mathbb{E}\left\|\sum_{i=1}^{N}E_ie_i\right\|_{K^\circ}+t\right)$$

with probability at least

$$1-\exp\left(-c\min\left\{\frac{t^2}{\sigma^2},\frac{t}{\sigma'}\right\}\right)\geq 1-\exp\left(-c\min\left\{\frac{t^2}{\sigma^2},\frac{t}{\sigma}\right\}\right),$$

where $\sigma=R(K)R(L^\circ)$ and $\sigma'=\sup_{x\in K}\|x\|_\infty\sup_{y\in L^\circ}\|y\|_\infty$.

Proof. As in the proof of Theorem 3.1, by Theorem 2.2, it is enough to consider the case $\Gamma=(E_{ij})$. Moreover it suffices to show that

$$\mathbb{P}(\|\Gamma : K\to L\|\geq\mathbb{E}\|\Gamma : K\to L\|+t)\leq\exp\left(-c\min\left\{\frac{t^2}{\sigma^2},\frac{t}{\sigma'}\right\}\right).$$

To obtain the above estimate we use Theorem 2.4. Recall that $\|\Gamma : K\to L\|=\sup_T\langle\Gamma,x\otimes y\rangle$, where $T=K\otimes L^\circ$. Thus we can easily compute parameters a and b in Theorem 2.4:

$$a=\sup_T|x\otimes y|=\sup_{x\in K,\ y\in L^\circ}|x|\cdot|y|=\sigma$$

and

$$b = \sup_T \|x \otimes y\|_\infty = \sup_{x \in K,\ y \in L^\circ} \|x\|_\infty \cdot \|y\|_\infty = \sigma'.$$

$\square$

4 Norms of sub-matrices and RIP

Here we estimate the norms of sub-matrices of an isotropic unconditional log-concave random $n \times N$ matrix Γ.

Recall that for subsets $J \subset \{1, \ldots, n\}$ and $I \subset \{1, \ldots, N\}$, $\Gamma(J, I)$ denotes the sub-matrix of Γ consisting of the rows indexed by elements from J and the columns indexed by elements from I. Recall also that for $k \leq n$ and $m \leq N$, $\Gamma_{k,m}$ is defined by

$$\Gamma_{k,m} = \sup \|\Gamma(J, I) \ : \ \ell_2^m \to \ell_2^k\|, \tag{8}$$

where the supremum is taken over all subsets $J \subset \{1, \ldots, n\}$ and $I \subset \{1, \ldots, N\}$ with cardinalities $|J| = k, |I| = m$. That is, $\Gamma_{k,m}$ is the maximal operator norm of a sub-matrix of Γ with k rows and m columns.

We also denote the set of ℓ-sparse unit vectors on $\mathbb{R}^d$ by U_ℓ (or $U_\ell(d)$, when we want to emphasize the dimension of the underlying space) and its convex hull by $\tilde{U}_\ell$, i.e.

$$U_\ell = U_\ell(d) = \{x \in \mathbb{R}^d \colon |\operatorname{supp} x| \leq \ell \ \text{ and } \ |x| = 1\}, \quad \text{and} \quad \tilde{U}_\ell = \operatorname{conv}(U_\ell).$$

Clearly, $\|\Gamma(J, I)\| = \sup \langle \Gamma x, y \rangle$, where supremum is taken over unit vectors x and y with supports I and J correspondingly. Therefore

$$\Gamma_{k,m} = \left\| \Gamma \ : \ \tilde{U}_m(N) \to (U_k(n))^\circ \right\|.$$

Note that $(U_k(n))^\circ = (\tilde{U}_k(n))^\circ$. Below U_ℓ° means $(U_\ell)^\circ$.

Remark. For matrices with N independent log-concave columns and $k = n$ the sharp estimates for $\Gamma_{n,m}$ were obtained in [4].

To treat the general case we will need the following simple lemma.

Lemma 4.1. *For any $1 \leq \ell \leq n$ we have*

$$\mathbb{E} \left\| \sum_{i=1}^n E_i e_i \right\|_{U_\ell^\circ} \approx \sqrt{\ell} \ln \frac{3n}{\ell}.$$

Proof. By Borell's lemma ([6]) we have

$$\left(\mathbb{E}\left\|\sum_{i=1}^{n} E_i e_i\right\|_{U_\ell^\circ}\right)^2 \approx \mathbb{E}\left\|\sum_{i=1}^{n} E_i e_i\right\|_{U_\ell^\circ}^2 = \mathbb{E}\sup_{\substack{I\subset\{1,\dots,n\}\\ |I|=\ell}} \sum_{i\in I} E_i^2 = \sum_{i=1}^{\ell} \mathbb{E}|E_i^*|^2,$$

where $E_1^*,\dots,E_n^*$ denotes the nonincreasing rearrangement of $|E_1|,\dots,|E_n|$. We conclude the proof by the standard well known estimate $\mathbb{E}|E_i^*|^2 \approx (\ln(3n/i))^2$ for $i \le n/2$ (and, hence, $\mathbb{E}|E_i^*|^2 \le C$ for $i \ge n/2$). $\square$

Now observe that Γ satisfies the hypothesis of Theorem 3.1 and that $\tilde{U}_\ell \subset B_2^n$, so $R(\tilde{U}_\ell) = 1$. Thus Theorem 3.1 implies

$$\mathbb{E}\Gamma_{k,m} \le C\left(\mathbb{E}\left\|\sum_{i=1}^{N} E_i e_i\right\|_{U_m^\circ} + \mathbb{E}\left\|\sum_{i=1}^{n} E_i e_i\right\|_{U_k^\circ}\right),$$

which together with Lemma 4.1 and Corollary 3.2 implies the following theorem.

Theorem 4.2. *There are absolute positive constants C and c such that the following holds. Let $m \le N$ and $k \le n$. Let Γ be an isotropic unconditional log-concave random $n \times N$ matrix. Then*

$$\mathbb{E}\Gamma_{k,m} \le C\left(\sqrt{m}\,\ln\frac{3N}{m} + \sqrt{k}\,\ln\frac{3n}{k}\right).$$

Moreover, for every $t > 0$,

$$\Gamma_{k,m} \le C\left(\sqrt{m}\,\ln\frac{3N}{m} + \sqrt{k}\,\ln\frac{3n}{k} + t\right)$$

with probability at least

$$1 - \exp\left(-c\min\left\{t, t^2\right\}\right).$$

Remarks. 1. In the case when $\Gamma = (E_{ij})$ we have

$$\begin{aligned}\mathbb{E}\Gamma_{k,m} &\ge \max\left\{\mathbb{E}\Big\|\sum_{i=1}^{N} E_i e_i\Big\|_{U_m^\circ}, \mathbb{E}\Big\|\sum_{i=1}^{n} E_i e_i\Big\|_{U_k^\circ}\right\}\\ &\ge \frac{1}{C}\left(\sqrt{m}\,\ln\frac{3N}{m} + \sqrt{k}\,\ln\frac{3n}{k}\right).\end{aligned}$$

2. Theorem 4.2 can be proved directly (i.e. without Chevet inequality) using a chaining argument in the spirit of [4]. We provide the details in the last section. Similar estimates (with worse probability) were recently independently obtained in [10].

We now estimate the restricted isometry constant (RIC) of a random matrix Γ with independent unconditional isotropic log-concave rows. As was mentioned in the example following Theorem 3.1 such Γ is unconditional isotropic log-concave. Recall that the RIC of order m is the smallest number $\delta = \delta_m(\Gamma)$ such that

$$(1-\delta)|x|^2 \leq |\Gamma x|^2 \leq (1+\delta)|x|^2.$$

for every $x \in U_m$.

The following theorem is an "unconditional" counterpart of Theorem 6.4 from [1] (see also Theorem 7 in [2]). Its proof repeats the lines of the corresponding proof in [1] (see the remark following the proof of Theorem 6.4 from [1]). The result is sharp up to dependence on θ and absolute constants (see Proposition 5.7 in [3]).

Theorem 4.3. *Let $0 < \theta < 1$. Let Γ be an $n \times N$ random matrix, whose rows are independent unconditional isotropic log-concave random vectors in $\mathbb{R}^N$. Then $\delta_m(\Gamma/\sqrt{n}) \leq \theta$ with probability at least*

$$1 - \exp\left(-c\,\frac{\theta^2 n}{\ln^2 n}\right) - 2\exp\left(-c\,\sqrt{m}\ln\frac{3N}{m}\right),$$

provided that either
(i) $N \leq n$ *and*

$$m \approx \min\left\{N,\ \frac{\theta^2 n}{\ln^3(3/\theta)}\right\}$$

or
(ii) $N \geq n$ *and*

$$m \leq c\,\frac{\theta n}{\ln(3N/(\theta n))}\ \min\left\{\frac{1}{\ln(3N/(\theta n))}, \frac{\theta}{\ln^2(3/\theta)}\right\},$$

where $c > 0$ is an absolute constant.

Remarks. 1. The condition on m in (ii) can be written as follows

$$\text{if} \quad \theta \geq \frac{\ln^2 \ln(3N/n)}{\ln(3N/n)} \qquad \text{then} \qquad m \leq c\, \frac{\theta n}{\ln^2(3N/(\theta n))},$$

$$\text{if} \quad \theta \leq \frac{\ln^2 \ln(3N/n)}{\ln(3N/n)} \qquad \text{then} \qquad m \leq c\, \frac{\theta^2}{\ln^2(3/\theta)}\, \frac{n}{\ln(3N/(\theta n))}.$$

2. Precisely the proof of Theorem 6.4 in [1] (with estimates from our Theorem 4.2) gives that if

$$b_m := m \left(\ln \frac{3N}{m} \right)^2 \leq c\theta n$$

and

$$m \ln \frac{3N}{m} \ln^2 \frac{n}{b_m} \leq c\theta^2 n$$

then $\delta_m(\Gamma/\sqrt{n}) \leq \theta$ with probability at least

$$1 - \exp\left(-c\, \frac{\theta^2 n}{\ln^2(n/b_m)} \right) - 2 \exp\left(-c\, \sqrt{m}\, \ln \frac{3N}{m} \right).$$

5 An example

In this section we prove that the condition "X is unconditional" cannot be omitted in Theorems 2.2 and 3.1. Namely, first we construct an example of isotropic log-concave non-unconditional d-dimensional random vector X and a norm $\|\cdot\|$ on $\mathbb{R}^d$, which fails to satisfy the conclusion of Theorem 2.2. Then we consider the matrix consisting of one column X as an operator from $(\mathbb{R}, |\cdot|)$ to $(\mathbb{R}^d, \|\cdot\|)$ and show that it does not satisfy the Chevet type inequality. The idea of the construction of X is rather simple – we start with a matrix with i.i.d. exponential entries and rotate its columns by a "random" rotation. Considering the matrix as a vector with operator norm $\ell_1 \to \ell_1$ we prove the result.

Theorem 5.1. *Let $d \geq 1$ and $Y = (E_1, \ldots, E_d)$. There exists an isotropic log-concave random vector X in $\mathbb{R}^d$ and a norm $\|\cdot\|$ such that*

$$\mathbb{E}\|X\| \geq c\, \sqrt{\ln d}\; \mathbb{E}\|Y\|, \tag{9}$$

where $c > 0$ is an absolute constant. Moreover, the $d \times 1$ matrix B with the only column equal to X satisfies

$$\mathbb{E}\|B : [-1,1] \to L\| \geq c\,\sqrt{\ln d}\,\left(\mathbb{E}\left\|\sum_{i=1}^{d} E_i e_i\right\|_L + \|\mathrm{Id} : B_2^d \to L\|\right),$$

where L is the unit ball of $\|\cdot\|$.

Proof. Let n, N be integers such that $d = nN$. Consider an $n \times N$ matrix $\Gamma = (E_{ij})$. Denote its columns by $X_1, \ldots, X_N$, so that $\Gamma = [X_1, \ldots, X_N]$. As before, we consider Γ as a d-dimensional vector. Given $U \in O(n)$ rotate the columns of Γ by U:

$$A = A(U) = U\Gamma = [UX_1, \ldots, UX_N].$$

Then A is a log-concave isotropic vector in $\mathbb{R}^d$. Below we show that if $N = \lfloor e^{cn} \rfloor$ for some absolute constant $c > 0$ then there exists $U_0 \in O(n)$ such that

$$\mathbb{E}_\Gamma\,\|A(U_0)\ :\ \ell_1^N \to \ell_1^n\| \geq c_1\sqrt{\ln d}\ \ \mathbb{E}_\Gamma\,\|\Gamma\ :\ \ell_1^N \to \ell_1^n\|. \tag{10}$$

This will prove the first part of the theorem, since it is clearly enough to consider only such n, N, d by adjusting the constant in the main statement.

To prove (10) we estimate the average of $\|A(U)\|$ over $U \in O(n)$. For every x in $\mathbb{R}^n$ we have

$$\mathbb{P}_{O(n)}\left(\left\{\|Ux\|_1 \geq c_2\sqrt{n}\ \|x\|_2\right\}\right) = \sigma_{n-1}(\{y\colon \|y\|_1 \geq c_2\sqrt{n}\}) \geq 1 - \exp(-2cn),$$

where σ_{n-1} denotes the uniform distribution on S^{n-1} and the last inequality follows by a simple volumetric argument (or by concentration, see e.g. 2.3, 5.1 and 5.3 in [12]). Thus, if $N \leq e^{cn}$,

$$\mathbb{P}_{O(n)}\left(\left\{\forall i \leq N\ :\|UX_i\|_1 \geq c_2\sqrt{n}\ \|X_i\|_2\right\}\right) \geq 1 - \exp(-cn) \geq \frac{1}{2}.$$

Hence

$$\mathbb{E}_{O(n)}\ \max_{i \leq N} \|UX_i\|_1 \geq c_2\sqrt{n}\ \max_{i \leq N} \|X_i\|_2,$$

which implies

$$\begin{aligned}\mathbb{E}_\Gamma\ \mathbb{E}_{O(n)}\ \|A(U)\ :\ \ell_1^N \to \ell_1^n\| &\geq c_2\sqrt{n}\ \ \mathbb{E}_\Gamma\ \max_{i \leq N} \|X_i\|_2 \\ &\geq c_2\sqrt{n}\ \ \mathbb{E}_\Gamma\ \max_{i \leq N} |E_{1,i}| \geq c_3\sqrt{n}\ \ln N.\end{aligned}$$

By Lemma 2.1
$$\mathbb{E}_\Gamma\ \|\Gamma\ :\ \ell_1^N \to \ell_1^n\| \approx n + \ln N.$$
Thus, taking $N = \lfloor e^{cn} \rfloor$,
$$\frac{\mathbb{E}_{O(n)}\ \mathbb{E}_\Gamma\ \|A(U)\ :\ \ell_1^N \to \ell_1^n\|}{\mathbb{E}_\Gamma\ \|\Gamma\ :\ \ell_1^N \to \ell_1^n\|} \geq c_4 \frac{\sqrt{n}\ \ln N}{n + \ln N} \geq c_5\sqrt{\ln N} \geq c_6\sqrt{\ln d}.$$
Hence there exists $U_0 \in O(n)$ satisfying (10).

Now we will prove the "moreover" part of the theorem. Recall that L is the unit ball of the norm $\|\cdot\|$ constructed above. The log-concave vector under consideration is $X = A(U_0)$ and the matrix which provides the counterexample to the Chevet type inequality is $B = [X]$. By the above calculations we have
$$\mathbb{E}\|B : [-1,1] \to L\| = \mathbb{E}\|X\|_L = \mathbb{E}\|A(U_0)\ :\ \ell_1^N \to \ell_1^n\| \geq c\ (\ln d)^{3/2}$$
and
$$\mathbb{E}\left\|\sum_{i=1}^d E_i e_i\right\|_L = \mathbb{E}\|\Gamma\ :\ \ell_1^N \to \ell_1^n\| \approx n + \ln N \approx \ln d.$$
It is easy to check that for every $n \times N$ matrix $T = (t_{ij})$ one has
$$\|T\ :\ \ell_1^N \to \ell_1^n\| = \max_{j \leq N} \sum_{i=1}^n |t_{ij}| \leq \sqrt{n} \left(\sum_{j=1}^N \sum_{i=1}^n |t_{ij}|^2\right)^{1/2} = \sqrt{n}\ |T|,$$
where $\sqrt{n}$ is the best possible constant in the inequality. This shows that
$$\|\mathrm{Id} : B_2^d \to L\| = \sqrt{n} \approx \sqrt{\ln d}.$$
Thus
$$\mathbb{E}\left\|\sum_{i=1}^d E_i e_i\right\|_L + \|\mathrm{Id} : B_2^d \to L\| \approx \ln d,$$
which completes the proof. $\square$

Concluding remarks. 1. The above example is optimal in the sense that one can't expect better than $\sqrt{\ln d}$ dependence on dimension in (9). Indeed, let $Y = (E_1, \ldots, E_d)$. We show that for any $U \in O(d)$ and any norm $\|\cdot\|$ on $\mathbb{R}^d$ one has
$$\mathbb{E}\|UY\| \leq C\sqrt{\ln(ed)}\ \mathbb{E}\|Y\|. \tag{11}$$

First it is known that $\mathbb{E}\|Y\| \le C\sqrt{\ln(ed)}\ \mathbb{E}\|G\|$, where $G = (g_1, \ldots, g_d)$. Now note that if K is a unit ball of $\|\cdot\|_K$ then for every $U \in O(d)$ one has $\|Ux\|_K = \|x\|_{U^{-1}K}$ for every $x \in \mathbb{R}^d$. Therefore, for any $U \in O(d)$ we have

$$\mathbb{E}\|UY\| \le C\sqrt{\ln(ed)}\ \mathbb{E}\|UG\| = C\sqrt{\ln(ed)}\ \mathbb{E}\|G\|$$

(in the last equality we used that the distribution of G is invariant under rotations). Finally note that by either Theorem 2.3 or Theorem 2.2 the norm of an exponential random vector dominates the norm of the Gaussian one, i.e. $\mathbb{E}\|G\| \le C_1\,\mathbb{E}\|Y\|$, which implies (11).

2. For any isotropic vector X in $\mathbb{R}^d$ (not necessarily log-concave) and any origin symmetric convex body $K \subset \mathbb{R}^d$ we show that

$$\mathbb{E}\|X\|_K \le Cd(K, B_2^d)\ \mathbb{E}\|Y\|_K, \tag{12}$$

where $Y = (E_1, \ldots, E_d)$ and $d(K, B_2^d)$ denotes the Banach-Mazur distance between K and B_2^d. Since for every origin symmetric K one has $d(K, B_2^d) \le \sqrt{d}$ (see e.g. [16]), the inequality (12) implies that for any norm $\|\cdot\|$ on $\mathbb{R}^d$

$$\mathbb{E}\|X\| \le C\sqrt{d}\ \mathbb{E}\|Y\|.$$

Now we prove (12). First, as in Remark 1, note that the norm of an exponential random vector dominates the norm of the Gaussian one. Thus it is enough to show that $\mathbb{E}\|X\|_K \le Cd(K, B_2^d)\ \mathbb{E}\|G\|_K$, where G is as in Remark 1. Let $\alpha = d(K, B_2^d)$ and $\mathcal{E}$ be an ellipsoid such that $\mathcal{E} \subset K \subset \alpha\mathcal{E}$. Since this is only a matter of rotation of a coordinate system we may assume that $\mathcal{E} = \{x \in \mathbb{R}^d \colon \sum_{i=1}^d a_i^2 x_i^2 \le 1\}$. Then by the isotropicity of X,

$$\mathbb{E}\|X\|_K \le \mathbb{E}\|X\|_{\mathcal{E}} = \mathbb{E}\left(\sum_{i=1}^d a_i^2 X_i^2\right)^{1/2} \le \left(\sum_{i=1}^d a_i^2\right)^{1/2} \le C\mathbb{E}\|G\|_{\mathcal{E}} \le C\alpha\mathbb{E}\|G\|_K,$$

where we used comparison of the first and second moments of the norm $\|G\|_{\mathcal{E}}$ of the Gaussian vector.

6 A direct proof of Theorem 4.2

We present here a proof of Theorem 4.2 not involving the Chevet type inequality and not relying on Theorem 2.3, but only on tail estimates for suprema of linear combinations of independent exponential variables given in Theorem 2.4 (as above, by Theorem 2.2, it is enough to consider only such variables).

We need the following lemma, which is an immediate consequence of Theorem 2.4 (recall here that for a matrix $A = (a_{ij})$, $\|A\|_\infty$ denotes $\max_{i,j} |a_{ij}|$).

Lemma 6.1. *For every $n \times N$ matrix $A = (a_{ij})$ and every $t \geq 0$ we have*

$$\mathbb{P}\left(\left|\sum_{ij} E_{ij}a_{ij}\right| \geq t\right) \leq 2\exp\left(-c\min\left(\frac{t^2}{|A|^2}, \frac{t}{\|A\|_\infty}\right)\right),$$

where $c > 0$ is an absolute constant.

Indeed, since $\mathbb{E}|\sum_{ij} E_{ij}a_{ij}| \leq (\mathbb{E}|\sum_{ij} E_{ij}x_{ij}|^2)^{1/2} = |A|$, the above Lemma follows from Theorem 2.4 for $t \geq 2|A|$. For $t \leq 2|A|$ we can make the right hand side larger than 1 by the choice of c.

Direct proof of Theorem 4.2. As in the proof of Theorem 3.1, using Theorem 2.2, we may assume that Γ is the exponential matrix, i.e. $\Gamma = (E_{ij})$. Without loss of generality we assume that $k \geq m$ and that $k = 2^r - 1$, $m = 2^s - 1$ for some positive integers $r \geq s$. It is known (and easy to see by volumetric argument) that for any origin symmetric convex body $V \subset \mathbb{R}^d$ and any $\varepsilon \leq 1$ there exists an ε-net (with respect to the metric defined by V) in V of cardinality at most $(3/\varepsilon)^d$. For $i = 0, 1, \ldots, r-1$ let $\mathcal{M}_i$ be a $(2^i/(4k))$-net (with respect to the metric defined by $B_2^n \cap (2^{-i/2}B_\infty^n)$) in the set

$$\bigcup_{\substack{I \subseteq \{1,\ldots,n\} \\ |I| \leq 2^i}} \mathbb{R}^I \cap B_2^n \cap (2^{-i/2}B_\infty^n)$$

of cardinality not greater than

$$\binom{n}{2^i}\left(\frac{12k}{2^i}\right)^{2^i} \leq \exp\left(C2^i \ln\left(\frac{2n}{2^i}\right)\right),$$

where $\mathbb{R}^I$ denotes the span of $\{e_i\}_{i \in I}$. Similarly for $i = 0, 1, \ldots, s-1$ let $\mathcal{N}_i$ be a $(2^i/(4m))$-net in the set

$$\bigcup_{\substack{I \subseteq \{1,\ldots,N\} \\ |I| \leq 2^i}} \mathbb{R}^I \cap B_2^N \cap (2^{-i/2}B_\infty^N)$$

of cardinality at most

$$\binom{N}{2^i}\Big(\frac{12m}{2^i}\Big)^{2^i} \leq \exp\Big(C2^i \ln\Big(\frac{2N}{2^i}\Big)\Big).$$

Let now $\mathcal{M}$ be the set of vectors in $2B_2^n$ that can be represented in the form $x = \sum_{i=0}^{r-1} x_i$, where $x_i \in \mathcal{M}_i$ and have pairwise disjoint supports. Analogously define $\mathcal{N}$ as the set of vectors $y = \sum_{i=0}^{s-1} y_i \in 2B_2^N$, with $y_i \in \mathcal{N}_i$ and pairwise disjoint supports. For $x \in \mathcal{M}$ and $i = 0, 1, \ldots, r-1$ let $S_i x = x_0 + \ldots + x_i$, where x_i is the appropriate vector from the above representation (this representation needs not be unique, so for each vector x we choose one of them). Similarly, for $i = 0, 1, \ldots, s-1$ and $y \in \mathcal{N}$ let $T_i y = y_0 + \ldots + y_i$. For $i = s, \ldots, r-1$ let $T_i y = y$. Additionally set $S_{-1}x = 0$, $T_{-1}y = 0$. We thus have

$$y \otimes x = \sum_{i=0}^{r-1} (T_i y \otimes S_i x - T_{i-1} y \otimes S_{i-1} x)$$

for $x \in \mathcal{M}, y \in \mathcal{N}$.

Recall that for vectors $v = (v(i))_i$ and $w = (w(i))_i$ the tensor $v \otimes w$ can be identified with the matrix $\{w(i)v(j)\}_{ij}$. Since x_i's and y_i's have pairwise disjoint supports, for every $j \geq i$ we can view $T_j y \otimes S_j x$ and $T_j y \otimes S_j x - T_{i-1} y \otimes S_{i-1} x$ as sub-matrices of $y \otimes x$. Then it is easy to see that for every $j \geq i$

$$|T_j y \otimes S_j x - T_{i-1} y \otimes S_{i-1} x| \leq |x \otimes y| = |x||y| \leq 4 \tag{13}$$

and

$$\begin{aligned} &\|T_j y \otimes S_j x - T_{i-1} y \otimes S_{i-1} x\|_\infty \\ &\leq \max\{\|(T_j - T_{i-1})y\|_\infty \|x\|_\infty, \|y\|_\infty \|(S_j - S_{i-1})x\|_\infty\} \leq 2^{-i/2}. \end{aligned} \tag{14}$$

Thus, by Lemma 6.1, for any $x \in \mathcal{M}$, $y \in \mathcal{N}$ and $t \geq 1$,

$$\mathbb{P}(|\langle \Gamma T_i y, S_i x\rangle - \langle \Gamma T_{i-1} y, S_{i-1} x\rangle| \geq t) \leq 2\exp(-c\min(t^2, 2^{i/2}t)). \tag{15}$$

Moreover, for any $i \leq s-1$, the cardinality of the set of vectors of the form $T_i y \otimes S_i x - T_{i-1} y \otimes S_{i-1} x$, $x \in \mathcal{M}, y \in \mathcal{N}$ is at most

$$\exp\Big(\sum_{j=0}^{i} \Big(C2^j \ln\Big(\frac{2n}{2^j}\Big) + C2^j \ln\Big(\frac{2N}{2^j}\Big)\Big)\Big) \leq \exp\Big(\tilde{C}2^i \ln\Big(\frac{2n}{2^i}\Big) + \tilde{C}2^i \ln\Big(\frac{2N}{2^i}\Big)\Big).$$

By (15) and the union bound we get that for $i \le s-1$ and any $t \ge 1$, with probability at least

$$1 - 2\exp\Big(-ct\Big(2^i \ln(2n/2^i) + 2^i \ln(2N/2^i)\Big)\Big),$$

one has

$$\max_{x\in\mathcal{M}, y\in\mathcal{N}} |\langle \Gamma T_i y, S_i x\rangle - \langle \Gamma T_{i-1} y, S_{i-1} x\rangle| \le Ct\Big(2^{i/2}\ln(2n/2^i) + 2^{i/2}\ln(2N/2^i)\Big).$$

By integration this yields

$$\mathbb{E}\max_{x\in\mathcal{M}, y\in\mathcal{N}} |\langle \Gamma T_i y, S_i x\rangle - \langle \Gamma T_{i-1} y, S_{i-1} x\rangle| \le C\Big(2^{i/2}\ln(2n/2^i) + 2^{i/2}\ln(2N/2^i)\Big).$$

Therefore

$$\begin{aligned}
\mathbb{E}\sup_{x\in\mathcal{M}, y\in\mathcal{N}} |\langle \Gamma T_{s-1} y, S_{s-1} x\rangle| &\le \sum_{i=0}^{s-1} \mathbb{E}\sup_{x\in\mathcal{M}, y\in\mathcal{N}} |\langle \Gamma T_i y, S_i x\rangle - \langle \Gamma T_{i-1} y, S_{i-1} x\rangle| \\
&\le \sum_{i=0}^{s-1} C\Big(2^{i/2}\ln(2n/2^i) + 2^{i/2}\ln(2N/2^i)\Big) \\
&\le C_1\Big(\sqrt{k}\ln(2n/k) + \sqrt{m}\ln(2N/m)\Big). \qquad (16)
\end{aligned}$$

On the other hand, for any $y \in \mathcal{N}$ and $i \ge s$, we have $T_{i-1}y = T_i y = y$. Thus by (15) and the fact that there are at most $\exp(C2^i \ln(2n/2^i))$ vectors of the form $S_i x - S_{i-1}x$ with $x \in \mathcal{M}$, we get for $t \ge 1$,

$$\sup_{x\in\mathcal{M}} |\langle \Gamma T_i y, (S_i x - S_{i-1} x)\rangle| \le Ct2^{i/2}\ln(2n/2^i),$$

with probability at least $1 - \exp(-ct2^i \ln(2n/2^i))$.

This implies that for $s \le i \le r-1$,

$$\mathbb{E}\max_{x\in\mathcal{M}} |\langle \Gamma T_i y, S_i x\rangle - \langle \Gamma T_{i-1} y, S_{i-1} x\rangle| \le C2^{i/2}\ln(2n/2^i)$$

and thus

$$\begin{aligned}
&\mathbb{E}\max_{x\in\mathcal{M}} |\langle \Gamma T_{r-1} y, S_{r-1} x\rangle - \langle \Gamma T_{s-1} y, S_{s-1} x\rangle| \\
&\quad\le \sum_{i=s}^{r-1} \mathbb{E}\max_{x\in\mathcal{M}} |\langle \Gamma T_i y, S_i x\rangle - \langle \Gamma T_{i-1} y, S_{i-1} x\rangle| \\
&\quad\le C\sum_{i=s}^{r-1} 2^{i/2}\ln(2n/2^i) \le \tilde{C}\sqrt{k}\ln(2n/k).
\end{aligned}$$

Applying Theorem 2.4 together with (13) and (14) (with $j = r-1$ and $i = s$) we obtain that for any $y \in \mathcal{N}$ and $t \geq 1$,

$$\max_{x\in\mathcal{M}} |\langle \Gamma T_{r-1} y, S_{r-1}x\rangle - \langle \Gamma T_{s-1}y, S_{s-1}x\rangle| \leq C\sqrt{k}\ln(2n/k) + Ct2^{s/2}\ln(2N/2^s),$$

with probability at least

$$1 - 2\exp(-\tilde{C}t2^s \ln(2N/2^s)),$$

which by the union bound and integration by parts gives

$$\begin{aligned}&\mathbb{E}\max_{x\in\mathcal{M},y\in\mathcal{N}} |\langle \Gamma T_{r-1} y, S_{r-1}x\rangle - \langle \Gamma T_{s-1}y, S_{s-1}x\rangle| \\ &\quad \leq C\sqrt{k}\ln(2n/k) + C2^{s/2}\ln(2N/2^s) \leq \tilde{C}\Big(\sqrt{k}\ln(2n/k) + \sqrt{m}\ln(2N/m)\Big).\end{aligned}$$

Combining this inequality with (16) we get

$$\mathbb{E}\max_{x\in\mathcal{M},y\in\mathcal{N}} |\langle \Gamma y, x\rangle| \leq C\Big(\sqrt{k}\ln(2n/k) + \sqrt{m}\ln(2N/m)\Big).$$

Let us now notice that for arbitrary $x \in S^{n-1}$, $y \in S^{n-1}$, with $|\operatorname{supp} x| \leq k$, $|\operatorname{supp} y| \leq m$, there exist $\tilde{x} \in \mathcal{M}, \tilde{y} \in \mathcal{N}$, such that $\operatorname{supp}\tilde{x} \subset \operatorname{supp} x$, $\operatorname{supp}\tilde{y} \subset \operatorname{supp} y$ and

$$|x-\tilde{x}|^2 \leq \sum_{i=0}^{r-1} 2^{2i}/(16k^2) \leq 1/8, \quad |y-\tilde{y}|^2 \leq \sum_{i=0}^{s-1} 2^{2i}/(16m^2) \leq 1/8.$$

We have

$$\langle \Gamma y, x\rangle = \langle \Gamma\tilde{y}, \tilde{x}\rangle + \langle \Gamma(y-\tilde{y}), x\rangle + \langle \Gamma\tilde{y}, x-\tilde{x}\rangle.$$

Taking into account that $\tilde{y} \in 2B_2^N$ and passing to suprema, we get

$$\Gamma_{k,m} \leq \max_{\tilde{x}\in\mathcal{M},\tilde{y}\in\mathcal{N}} \langle \Gamma\tilde{y}, \tilde{x}\rangle + 3\Gamma_{k,m}/8$$

and thus

$$\mathbb{E}\Gamma_{k,m} \leq 2\mathbb{E}\max_{\tilde{x}\in\mathcal{M},\tilde{y}\in\mathcal{N}} \langle \Gamma\tilde{y}, \tilde{x}\rangle \leq C(\sqrt{k}\ln(2n/k) + \sqrt{m}\ln(2N/m)),$$

which completes the proof of the first part of Theorem 4.2. The proof of the "moreover" part is obtained using Theorem 2.4 in the same way as it was used to obtain Corollary 3.2 from Theorem 3.1. □

Remark. Adjusting the chaining argument presented above one can eliminate the use of the full strength of Theorem 2.4 and obtain a proof relying only on tail inequalities for linear combinations of independent exponential random variables (which follow from classical Bernstein inequalities). The modification involves splitting the proof into two cases depending on the comparison between $m \ln(2N/m)$ and $k \ln(2n/k)$.

References

[1] R. Adamczak, R. Latała, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, *Tail estimates for norms of sums of log-concave random vectors*, preprint, http://arxiv.org/pdf/1107.4070v1.pdf

[2] R. Adamczak, R. Latała, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, *Geometry of log-concave Ensembles of random matrices and approximate reconstruction*, C.R. Math. Acad. Sci. Paris, **349** (2011), 783–786.

[3] R. Adamczak, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, *Restricted isometry property of matrices with independent columns and neighborly polytopes by random sampling*, Constructive Approximation, **34** (2011), 61–88.

[4] R. Adamczak, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, *Quantitative estimates of the convergence of the empirical covariance matrix in log-concave Ensembles*, Journal of AMS, **234** (2010), 535–561.

[5] Y. Benyamini and Y. Gordon, *Random factorization of operators between Banach spaces*, J. Analyse Math. **39** (1981), 45–74.

[6] C. Borell, *Convex measures on locally convex spaces*, Ark. Math. **12** (1974), 239–252.

[7] S. Chevet, *Séries de variables aléatoires gaussiennes à valeurs dans $E\hat{\otimes}_\varepsilon F$. Application aux produits d'espaces de Wiener abstraits*, Séminaire sur la Géométrie des Espaces de Banach (1977-1978), Exp. No. 19, École Polytech., Palaiseau, 1978.

[8] Y. Gordon, *Some inequalities for Gaussian processes and applications*, Israel J. Math. **50** (1985), 265–289.

[9] R. Latała, *On weak tail domination of random vectors*, Bull. Polish Acad. Sci. Math. **57** (2009), 75–80.

[10] S. Mendelson and G. Paouris, *On generic chaining and the smallest singular value of random matrices with heavy tails*, J. Funct. Anal. **262** (2012), 3775–3811.

[11] S. Mendelson and N. Tomczak-Jaegermann. *A subgaussian embedding theorem*, Israel J. Math. **164** (2008), 349–364.

[12] V.D. Milman and G. Schechtman, *Asymptotic theory of finite-dimensional normed spaces. With an appendix by M. Gromov*, Lecture Notes in Mathematics, 1200. Springer-Verlag, Berlin, 1986.

[13] M. Talagrand, *A new isoperimetric inequality and the concentration of measure phenomenon*, Geometric aspects of functional analysis (1989–90), 94–124, Lecture Notes in Math., 1469, Springer, Berlin, 1991.

[14] M. Talagrand, *The supremum of some canonical processes.* Amer. J. Math. **116** (1994), 283–325.

[15] M. Talagrand, *The generic chaining. Upper and lower bounds of stochastic processes*, Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2005.

[16] N. Tomczak-Jaegermann, *Banach-Mazur distances and finite-dimensional operator ideals*, Pitman Monographs and Surveys in Pure and Applied Mathematics, 38. Longman Scientific & Technical, Harlow; John Wiley & Sons, Inc., New York, 1989.

Radosław Adamczak,
Institute of Mathematics,
University of Warsaw,
Banacha 2, 02-097 Warszawa, Poland
`e-mail: radamcz@mimuw.edu.pl`

Rafał Latała,
Institute of Mathematics,
University of Warsaw,

Banacha 2, 02-097 Warszawa, Poland
and
Institute of Mathematics,
Polish Academy of Sciences,
Śniadeckich 8, 00-956 Warszawa, Poland
e-mail: rlatala@mimuw.edu.pl

Alexander E. Litvak,
Dept. of Math. and Stat. Sciences,
University of Alberta,
Edmonton, Alberta, Canada, T6G 2G1.
e-mail: alexandr@math.ualberta.ca

Alain Pajor,
Université Paris-Est
Équipe d'Analyse et Mathématiques Appliquées,
5, boulevard Descartes, Champs sur Marne,
77454 Marne-la-Vallée, Cedex 2, France
e-mail: Alain.Pajor@univ-mlv.fr

Nicole Tomczak-Jaegermann,
Dept. of Math. and Stat. Sciences,
University of Alberta,
Edmonton, Alberta, Canada, T6G 2G1.
e-mail: nicole.tomczak@ualberta.ca