Tail estimates for norms of sums of log-concave random vectors

Radosław Adamczak [*] Rafał Latała[*]
Alexander E. Litvak [†] Alain Pajor [‡]
Nicole Tomczak-Jaegermann [§]

July 14, 2011

Abstract

We establish new tail estimates for order statistics and for the Euclidean norms of projections of an isotropic log-concave random vector. More generally, we prove tail estimates for the norms of projections of sums of independent log-concave random vectors, and uniform versions of these in the form of tail estimates for operator norms of matrices and their sub-matrices in the setting of a log-concave ensemble. This is used to study a quantity $A_{k,m}$ that controls uniformly the operator norm of the sub-matrices with k rows and m columns of a matrix A with independent isotropic log-concave random rows. We apply our tail estimates of $A_{k,m}$ to the study of Restricted Isometry Property that plays a major role in the Compressive Sensing theory.

AMS Classification: 46B06, 15B52, 60E15, 60B20

Key Words and Phrases: log-concave random vectors, random matrices, Compressive Sensing, Restricted Isometry Property,

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

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

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

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

1 Introduction

In the recent years a lot of work was done on the study of the empirical covariance matrix, and on understanding related random matrices with independent rows or columns. In particular, such matrices appear naturally in two important (and distinct) directions. Namely,
– approximation of covariance matrices of high-dimensional distributions by empirical covariance matrices; and
– the Restricted Isometry Property of sensing matrices defined in the Compressive Sensing theory.

To illustrate, let n, N be integers. For $1 \le m \le N$ by $U_m = U_m(\mathbb{R}^N)$ we denote the set of m-sparse vectors of norm one, that is, vectors $x \in S^{N-1}$ with at most m non-zero coordinates. For any $n \times N$ random matrix A, treating A as a linear operator $A : \mathbb{R}^N \to \mathbb{R}^n$ we define $\delta_m(A)$ by $\delta_m(A) = \sup_{x \in U_m} \Big| |Ax|^2 - \mathbb{E}|Ax|^2 \Big|$. (Here $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^n$.)

Now let $X \in \mathbb{R}^N$ be a centered random vector with the covariance matrix equal to the identity, that is, $\mathbb{E} X \otimes X = Id$; such vectors are called isotropic. Consider n independent random vectors $X_1, \dots, X_n$ distributed as X and let A be the $n \times N$ matrix whose rows are $X_1, \dots, X_n$. Then

$$\begin{aligned} \delta_m\Big(\frac{A}{\sqrt{n}}\Big) &= \sup_{x \in U_m} \Big| \frac{1}{n} \left(|Ax|^2 - \mathbb{E}|Ax|^2 \right) \Big| \\ &= \sup_{x \in U_m} \Big| \frac{1}{n} \sum_{i=1}^{n} \left(|\langle X_i, x\rangle|^2 - \mathbb{E}|\langle X_i, x\rangle|^2 \right) \Big|. \qquad (1.1) \end{aligned}$$

In the particular case of $m = N$ it is also easy to check that

$$\delta_N\Big(\frac{A}{\sqrt{n}}\Big) = \Big\| \frac{1}{n} \sum_{i=1}^{n} \left(X_i \otimes X_i - \mathbb{E} X \otimes X \right) \Big\|. \qquad (1.2)$$

We first discuss the case $n \ge N$. In this case we will work only with the parameter $\delta_N(A/\sqrt{n})$. By the law of large numbers, under some moment hypothesis, the empirical covariance matrix $\frac{1}{n}\sum_{i=1}^{n} X_i \otimes X_i$ converges to $\mathbb{E}\, X \otimes X = Id$ in the operator norm, as $n \to \infty$. A natural goal important for many classes of distributions is to get quantitative estimates of the rate of this convergence, in other words, to estimate the error term $\delta_N(A/\sqrt{n})$ with high probability, as $n \to \infty$.

This question was raised and investigated in [17] motivated by a problem of complexity in computing volume in high dimensions. In this setting it was natural to consider uniform measures on convex bodies, or more generally, log-concave measures (see below for all the definitions). Partial solutions were given in [10] and [26] soon after the question was raised, and in the intervening years further partial solutions were produced. A full and optimal answer to the Kannan-Lovász- Simonovits question was given in [5] and [7]. For recent results on similar questions for other distributions, see e.g., [29] and [27].

The answer from [5] and [7] to the K-L-S question on the rate of convergence stated that:

$$\mathbb{P}\left(\sup_{x\in S^{N-1}}\left|\frac{1}{n}\sum_{i=1}^{n}\left(|\langle X_i,x\rangle|^2-1\right)\right|\leq C\sqrt{\frac{N}{n}}\right)\geq 1-e^{-c\sqrt{N}}, \tag{1.3}$$

where C and c are absolute positive constants. The proofs are based on an approach initiated by J. Bourgain [10] where the following norm of a matrix played a central role. Let $1\leq k\leq n$, then

$$A_{k,N}=\sup_{\substack{J\subset\{1,\dots,n\}\\ |J|=k}}\ \sup_{x\in S^{N-1}}\left(\sum_{j\in J}|\langle X_j,x\rangle|^2\right)^{1/2}=\sup_{\substack{J\subset\{1,\dots,n\}\\ |J|=k}}\ \sup_{x\in S^{N-1}}|P_JAx|, \tag{1.4}$$

where for $J\subset\{1,\dots,n\}$, P_J denotes the orthogonal projection on the coordinate subspace $\mathbb{R}^J$ of $\mathbb{R}^n$.

To understand the role of $A_{k,N}$ for estimating $\delta_N(A/\sqrt{n})$, let us explain the standard approach. For each individual x on the sphere, the rate of convergence may be estimated via some probabilistic concentration inequality. The method consists of a discretization of the sphere and then the use of an approximation argument to complete the proof. This approach works perfectly as long as the trade-off between complexity and concentration allows it.

Thus when the random variables $\frac{1}{n}\sum_{i=1}^{n}|\langle X_i,x\rangle|^2$ satisfy a good concentration inequality sufficient to handle uniformly exponentially many points, the method works. This is the case for instance when the random variables $\langle X,x\rangle$ are sub-gaussian or bounded, due to Bernstein inequalities. In the general case, we decompose the function $|\langle X_i,x\rangle|^2$ as the sum of two terms, the first being its truncation at the level B^2, for some $B>0$. Now let us discuss the second term in the decomposition of $\sum_{i=1}^{n}|\langle X_i,x\rangle|^2$. Let

$$E_B=E_B(x)=\{i\leq n\,:\,|\langle X_i,x\rangle|>B\}.$$

For simplicity, let us assume that the maximum cardinality of the sets of the family $\{E_B(x) : x \in S^{N-1}\}$ is a fixed non-random number k, then clearly the second term is controlled by

$$\sum_{i \in E_B} |\langle X_i, x\rangle|^2 \leq A_{k,N}^2.$$

In order to estimate k, let x such that $k = |E_B(x)| = |E_B|$, then

$$B^2 k = B^2 |E_B| \leq \sum_{i \in E_B} |\langle X_i, x\rangle|^2.$$

Thus we get the implicit relation $A_{k,N}^2 \geq B^2 k$. From this relation and an estimate of the parameter $A_{k,N}$ we eventually deduce an upper bound for k. To conclude the argument of Bourgain the bounded part is uniformly estimated by a classical concentration inequality and the rest is controlled by the parameter $A_{k,N}$.

Notice that we only need tail inequalities to estimate $A_{k,N}$, that is to control uniformly the norms of sub-matrices of A. This is still a difficult task however because of a high complexity of the problem and the lack of matching probability estimates; and a more sophisticated argument has been developed in [5] to handle it.

We now pass to the complementary case $n < N$, which is one of central points of the present paper, and was announced in [4].

Let A be an $n \times N$ random matrix with rows $X_1, \ldots, X_n$ which are independent random centered and with covariance matrices equal to the identity, but not necessarily identically distributed. Clearly, A is then not invertible. The uniform concentration on the sphere $U_N = S^{N-1}$ (which appeared in the definition of $\delta_m(A/\sqrt{n})$ for $m = N$) does not hold and the expressions in (1.1) are not uniformly small on $U_N = S^{N-1}$. The best one can hope for is that A may be "almost norm-preserving" on some subsets of S^{N-1}. This is true for subsets U_m, for some $1 \leq m \leq N$ and is indeed measured by $\delta_m(A/\sqrt{n})$.

The parameter δ_m plays a major role in the Compressive Sensing theory and an important question is to bound it from above with high probability, for some (fixed) m. For example, it can be directly used to express the so-called Restricted Isometry Property (RIP) (introduced by E. Candes and T. Tao in [12]) which in turn ensures that every m-sparse vector x can be reconstructed from its compression Ax with $n \ll N$ by the so-called ℓ_1-minimization method.

For matrices with independent rows $X_1, \dots, X_n$, questions on the RIP were understood and solved in the case of Gaussian and sub-gaussian measurements (see [12], [23] and [8]). When $X_1, \dots, X_n$ are independent log-concave isotropic random vectors, these questions remained open and this is one of our motivation for this article.

For an $n \times N$ matrix A and $m \le N$, the definition of $\delta_m(A/\sqrt{n})$ implies a uniform control of the norms of all sub-matrices of $A/\sqrt{n}$ with n rows and m columns. Passing to transposed matrices, it implies a uniform control of $|P_I X_i|$ over all $I \subset \{1, 2, \cdots, N\}$ of cardinality m and $1 \le i \le n$. In order to verify a necessary condition that for some m, $\delta_m(A/\sqrt{n})$ is small with high probability, one needs to get an upper estimate for $\sup\{|P_I X| : |I| = m\}$ valid with high probability.

The probabilistic inequality from [24]

$$\mathbb{P}\left(|P_I X| \ge C\, t\sqrt{m}\right) \le e^{-t\sqrt{m}} \tag{1.5}$$

valid for $t \ge 1$ is optimal for each individual I, but it does not allow to get directly (by a union bound argument) a uniform estimate because the probability estimate does not match the cardinality of the family of the I's. Thus the first natural goal we address in this paper is to get uniform tail estimates for some norms of log-concave random vectors.

This heuristic analysis points out to the main objective and novelty of the present paper; namely the study of high-dimensional log-concave measures and a deeper understanding of such measures and their convolutions via new tail estimates for norms of sums of projections of log-concave random vectors.

To emphasize a uniform character of our tail estimates, for an integer $N \ge 1$, an N-dimensional random vector Z, an integer $1 \le m \le N$, and $t \ge 1$, we consider the event

$$\Omega(Z, t, m, N) = \left\{ \sup_{\substack{I \subset \{1, \dots, N\} \\ |I| = m}} |P_I Z| \ge C t \sqrt{m} \log\left(\frac{eN}{m}\right) \right\}, \tag{1.6}$$

where C is a sufficiently large absolute constant. Note that the cut-off level in this definition is of the order of the median of the supremum for the exponential random vector.

Recall that $X, X_1, \ldots, X_n$ denote N-dimensional independent log-concave isotropic random vectors, and A is the $n \times N$ matrix whose rows are $X_1, \ldots, X_n$. A chain of main results of this paper provides estimates for $\mathbb{P}(\Omega(Z, t, m, N))$ in the cases when

(i) $Z = X$; and, more generally,

(ii) $Z = Y$ is a weighted sum $Y = \sum_1^n x_i X_i$, where $x = (x_i)_1^n \in \mathbb{R}^n$, with control of the Euclidean and supremum norms of x,

(iii) a uniform version of (ii) in the form of tail estimates for operator norms of sub-matrices of A.

Our first main theorem solves the question of uniform tail estimates for projections of a log-concave random vector discussed above.

Theorem 1.1. *Let X be an N-dimensional log-concave isotropic random vector. For any $1 \le m \le N$ and $t \ge 1$,*

$$\mathbb{P}\Big(\Omega(X, t, m, N)\Big) \le \exp\Big(- t\sqrt{m} \log\Big(\frac{eN}{m}\Big) / \sqrt{\log(em)}\Big).$$

The proof of the theorem is based on tail estimates for order statistics of isotropic log-concave vectors. By $(X^*(i))_i$ we denote the non-increasing rearrangement of $(|X(i)|)_i$. Combining (1.5) with methods of [18] and the formula $\sup_{\substack{I \subset \{1,\ldots,N\} \\ |I|=m}} |P_I X| = \left(\sum_{i=1}^m X^*(i)^2\right)^{1/2}$ will complete the argument.

Let us also mention that further applications (in Section 4) of inequality of this type require a stronger probability bound that involves a natural parameter $\sigma_X(p)$ – defined in (3.3) – determined by a "ℓ_p-weak" behavior of the random vector X.

More generally, the next step provides tail estimates for Euclidean norms of weighted sums of independent isotropic log-concave random vectors. Let $x = (x_i)_1^n \in \mathbb{R}^n$ and set $Y = \sum_1^n x_i X_i = A^* x$. The key estimate used later, Theorem 4.3, provides uniform estimates for the Euclidean norm of projections of Y. Namely, for every $x \in \mathbb{R}^n$, $\mathbb{P}\Big(\Omega(Y, t, m, N)\Big)$ is exponentially small with specific estimates depending on whether the ratio $\|x\|_\infty / |x|$ is larger or smaller than $1/\sqrt{m}$. Since precise formulations of probability estimates are rather convoluted we do not state them here and we refer the reader to Section 4.

The last step of this chain of results estimating probabilities of (1.6) is connected with the family of parameters $A_{k,m}$, with $1 \le k \le n$ and $1 \le m \le N$, defined by

$$A_{k,m} = \sup_{\substack{J\subset\{1,\dots,n\}\\ |J|=k}} \sup_{x\in U_m} \Big(\sum_{j\in J} |\langle X_j, x\rangle|^2\Big)^{1/2} = \sup_{\substack{J\subset\{1,\dots,n\}\\ |J|=k}} \sup_{x\in U_m} |P_J A x|. \tag{1.7}$$

That is, $A_{k,m}$ is the maximal operator norm over all sub-matrices of A with k rows and m columns (and for $m = N$ it obviously coincides with (1.4)).

Finding bounds on deviation of $A_{k,m}$ is one of our main goals. To develop an intuition of this result we state it below in a slightly less technical form. Full details are contained in Theorem 5.1.

Theorem 1.2. *For any $t \ge 1$ and $n \le N$ we have*

$$\mathbb{P}\big(A_{k,m} \ge Ct\lambda\big) \le \exp(-t\lambda/\sqrt{\log(3m)}),$$

where $\lambda = \sqrt{\log\log(3m)}\sqrt{m}\log(eN/m) + \sqrt{k}\log(en/k)$ and C is a universal constant.

The threshold value λ is optimal, up to the factor of $\sqrt{\log\log(3m)}$. Assuming additionally unconditionality of the distributions of rows (or columns), this factor can be removed to get a sharp estimate (see [3]).

We make several comments about the proof. Set $\Gamma = A^*$. Then

$$A_{k,m} = \sup_{\substack{I\subset\{1,\dots,N\}\\ |I|=m}} \sup_{x\in U_k(\mathbb{R}^n)} |P_I \Gamma x| = \sup_{\substack{I\subset\{1,\dots,N\}\\ |I|=m}} \sup_{x\in U_k(\mathbb{R}^n)} \Big|\Big(\sum x_i P_I X_i\Big)\Big|.$$

To bound $A_{k,m}$ one has then to prove uniformity with respect to two families of different character: one coming from the cardinality of the family $\{I \subset \{1,\dots,N\} : \ |I| = m\}$; and the other, from the complexity of $U_k(\mathbb{R}^n)$. This leads us to distinguishing two cases, depending on the relation between k and the quantity

$$k' = \inf\{\ell \ge 1 \colon m\log(eN/m) \le \ell\log(en/\ell)\}.$$

First, if $k \ge k'$, we adjust the chaining argument similar to the one from [5] to reduce the problem to the case $k \le k'$. In this step we use the uniform tail estimate from Theorem 3.4 for the Euclidean norm of the family of vectors

$\{P_I X : |I| = m\}$. Next, we use a different chain decomposition of x and apply Theorem 4.3.

As already alluded to, an independent interest of this paper lies in upper bounds for $\delta_m(A/\sqrt{n})$ where A is our $n \times N$ random matrix. We presently return to this subject to explain the connections.

The family $A_{k,m}$ plays a very essential role in studies of the Restricted Isometry constant, which in fact applies even in a more general setting. Namely, for an arbitrary subset $T \subset \mathbb{R}^N$ and $1 \le k \le n$ define the parameter $A_k(T)$ by

$$A_k(T) = \sup_{\substack{I \subset \{1,\dots,n\} \\ |I|=k}} \sup_{y \in T} \Big(\sum_{i \in I} |\langle X_i, y\rangle|^2 \Big)^{1/2}. \tag{1.8}$$

Thus $A_{k,m} = A_k(U_m)$. The parameter $A_k(T)$ was studied in [22] by means of Talagrand's γ-functionals.

The following lemma reduces a concentration inequality to a deviation inequality and hence is useful in studies of the RIP. It is based on an argument of truncation similar to Bourgain's approach presented earlier.

Lemma 1.3. *Let $X_1, \dots, X_n$ be independent isotropic random vectors in $\mathbb{R}^N$. Let $T \subset S^{N-1}$ be a finite set. Let $0 < \theta < 1$ and $B \ge 1$. Then with probability at least $1 - |T| \exp(-3\theta^2 n/8B^2)$ one has*

$$\sup_{y \in T} \left| \frac{1}{n} \sum_{i=1}^n (|\langle X_i, y\rangle|^2 - \mathbb{E}|\langle X_i, y\rangle|^2) \right| \le \theta + \frac{1}{n} \left(A_k(T)^2 + \mathbb{E} A_k(T)^2 \right),$$

where $k \le n$ is the largest integer satisfying $k \le (A_k(T)/B)^2$.

In this paper we focus on the compressive sensing setting where T is the set of sparse vectors. The lemma above shows that after a suitable discretisation, estimating δ_m or checking the RIP, can be reduced to estimating $A_{k,m}$. This generalizes naturally Bourgain's approach explained above for $m = N$.

Using the lemma, we can show that if $0 < \theta < 1$, $B \ge 1$, and $m \le N$ satisfy $m \log(CN/m) \le 3\theta^2 n/16B^2$, then with probability at least $1 - \exp(-3\theta^2 n/16B^2)$ one has

$$\delta_m(A/\sqrt{n}) \le \theta + \frac{1}{n} \left(A_{k,m}^2 + \mathbb{E} A_{k,m}^2 \right),$$

where $k \le n$ is the largest integer satisfying $k \le (A_{k,m}/B)^2$ (note that k is a random variable).

Combining this with tail inequalities from Theorem 1.2 allows us to prove the following result on the RIP of matrices with independent isotropic log-concave rows.

Theorem 1.4. *Let $0 < \theta < 1$, $1 \le n \le N$. Let A be an $n \times N$ random matrix with independent isotropic log-concave rows. There exists $c(\theta) > 0$ such that $\delta_m(A/\sqrt{n}) \le \theta$ with an overwhelming probability, whenever*

$$m \log^2(2N/m) \log\log 3m \le c(\theta)n.$$

The result is optimal, up to the factor $\log\log 3m$, as shown in [6]. As for Theorem 5.1, assuming unconditionality of the distributions of the rows, this factor can be removed (see [3]).

The paper is organized as follows. In the next section we collect the notation and necessary preliminary tools concerning log-concave random variables. In Section 3, given an isotropic log-concave random vector X, we present several uniform tail estimates for Euclidean norms of the whole family of projections of X on coordinate subspaces of dimension m. As already mentioned, these estimates are based on tail estimates for order statistics of X. The main result, Theorem 3.4, provides a strong probability bound in terms of the "ℓ_p-weak" parameter $\sigma_X(p)$ defined in (3.3). The proofs of the main technical results, Theorems 3.2 and 3.4, are given in Section 7. Section 4 provides tail estimates for Euclidean norms of projections of weighted sums of independent isotropic log-concave random vectors. The proof of the main Theorem 4.3 is a combination of Theorem 3.4 and one-dimensional Proposition 4.3. In Section 5 we prove the result announced above on deviation of $A_{k,m}$. Section 6 treats the Restricted Isometry Property and estimates of $\delta_m(A/\sqrt{n})$. The last Section 7 is devoted to the proofs of technical results of Section 3.

Acknowledgment: The research on this project was partially done when the authors participated in the Thematic Program on Asymptotic Geometric Analysis at the Fields Institute in Toronto in Fall 2010 and in the Discrete Analysis Programme at the Isaac Newton Institute in Cambridge in Spring 2011. The authors wish to thank these institutions for their hospitality and excellent working conditions.

2 Notation and preliminaries

Let L be an origin symmetric convex compact body in $\mathbb{R}^d$. This is the unit ball of a norm that we denote by $\|\cdot\|_L$. Let $K \subset \mathbb{R}^d$. We say that a set $\Lambda \subset K$ is an ε-net of K with respect to the metric corresponding to L if

$$K \subset \bigcup_{z\in\Lambda}(z+\varepsilon L).$$

In other words, for every $x \in K$ there exists $z \in \Lambda$ such that $\|x-z\|_L \le \varepsilon$. We will mostly use ε-nets in the case $K = L$. It is well-known (and follows by the standard volume argument) that for every symmetric convex compact body K in $\mathbb{R}^d$ and every $\varepsilon > 0$ there exists an ε-net Λ of K with respect to metric corresponding to K, of cardinality not exceeding $(1+2/\varepsilon)^d$. It is also easy to see that $\Lambda \subset K \subset (1-\varepsilon)^{-1}\,\mathrm{conv}\Lambda$. In particular, for any convex positively 1-homogenous function f one has

$$\sup_{x\in K} f(x) \le (1-\varepsilon)^{-1}\sup_{x\in\Lambda} f(x).$$

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

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

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

A random vector X in $\mathbb{R}^n$ is called log-concave if for all compact nonempty sets $A, B \subset \mathbb{R}^n$ and $\theta \in [0,1]$, $\mathbb{P}(X \in \theta A + (1-\theta)B) \ge \mathbb{P}(X \in A)^\theta \mathbb{P}(X \in B)^{1-\theta}$. By the result of Borell [9] a random vector X with full dimensional support is log-concave if and only if it admits a log-concave density f, i.e. such density for which

$$f(\theta x + (1-\theta)y) \ge f(x)^\theta f(y)^{1-\theta} \quad \text{for all } x,y\in\mathbb{R}^n,\ \theta\in[0,1].$$

It is known that any affine image, in particular any projection, of a log-concave random vector is log-concave. Moreover, if X and Y are independent log-concave random vectors then so is $X+Y$ (see [9, 14, 25]).

One important and simple model of a centered log-concave random variable with variance 1 is the symmetric exponential random variable E which has density $f(t) = 2^{-1/2}\exp(-\sqrt{2}|t|)$. In particular for every $s > 0$ we have $\mathbb{P}(|E| \ge s) = \exp(-s/\sqrt{2})$.

Every centered log-concave random variable Z, with variance 1 satisfies a sub-exponential inequality:

$$\text{for every } s > 0, \quad \mathbb{P}(|Z| \geq s) \leq C \exp(-s/C), \tag{2.1}$$

where $C > 0$ is an absolute constant (see [9]).

Definition 2.1. *For a random variable Z we define the ψ_1-norm by*

$$\|Z\|_{\psi_1} = \inf \{C > 0 \, : \, \mathbb{E} \exp(|Z|/C) \leq 2\}$$

and we say that Z is ψ_1 with constant ψ, if $\|Z\|_{\psi_1} \leq \psi$.

A consequence of (2.1) is that there exists an absolute constant $C > 0$ such that any centered log-concave random variable with variance 1 is ψ_1 with constant C.

It is well known that the ψ_1-norm of a random variable may be estimated from the growth of the moments. More precisely if a random variable Z is such that for any $p \geq 1$, $\|Z\|_p \leq pK$, for some $K > 0$, then $\|Z\|_{\psi_1} \leq cK$ where c is an absolute constant.

By $|\cdot|$ we denote the standard Euclidean norm on $\mathbb{R}^n$ as well as the cardinality of a set. By $\langle \cdot, \cdot \rangle$ we denote the standard inner product on $\mathbb{R}^n$. We denote by B_2^n and S^{n-1} the standard Euclidean unit ball and unit sphere in $\mathbb{R}^n$.

A vector $x \in \mathbb{R}^n$ is called sparse or k-sparse for some $1 \leq k \leq n$ if the cardinality of its support satisfies $|\operatorname{supp} x| \leq k$.

We let

$$U_k = U_k(\mathbb{R}^n) := \{x \in S^{n-1} \colon x \text{ is } k\text{-sparse}\}. \tag{2.2}$$

For any subset $I \subset \{1, \ldots, N\}$ let P_I denote the orthogonal projection on the coordinate subspace $\mathbb{R}^I := \{y \in \mathbb{R}^N \colon \operatorname{supp} y \subset I\}$.

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.

3 New bounds for log-concave vectors

In this section we state several new estimates for Euclidean norms of log-concave random vectors. Proofs of Theorems 3.2 and 3.4 are given in Section 7.

We start with the following theorem, which was essentially proved by Paouris in [24]. Indeed, it is a consequence of Theorem 8.2 combined with Lemma 3.9 in that paper, after checking that Lemma 3.9 holds not only for convex bodies but for log-concave measures as well.

Theorem 3.1. *For any N-dimensional log-concave random vector X and any $p \geq 1$ we have*

$$(\mathbb{E}|X|^p)^{1/p} \leq C\Big((\mathbb{E}|X|^2)^{1/2} + \sup_{t \in S^{N-1}} (\mathbb{E}|\langle t, X\rangle|^p)^{1/p}\Big), \tag{3.1}$$

where C is an absolute constant.

Remarks. 1. It is well known (cf. [9]) that if Z is a log-concave random variable then

$$(\mathbb{E}|Z|^p)^{1/p} \leq C\frac{p}{q}(\mathbb{E}|Z|^q)^{1/q} \quad \text{ for } p \geq q \geq 2.$$

If Z is symmetric one may in fact take $C = 1$ (cf. Proposition 3.8 in [20]) and if Z is centered then denoting by Z' an independent copy of Z we get for $p \geq q \geq 2$,

$$(\mathbb{E}|Z|^p)^{1/p} \leq (\mathbb{E}|Z - Z'|^p)^{1/p} \leq \frac{p}{q}(\mathbb{E}|Z - Z'|^q)^{1/q} \leq 2\frac{p}{q}(\mathbb{E}|Z|^q)^{1/q}.$$

Therefore if $X \in \mathbb{R}^N$ is isotropic log-concave then

$$\sup_{t \in S^{N-1}} (\mathbb{E}|\langle t, X\rangle|^p)^{1/p} \leq p \sup_{t \in S^{N-1}} (\mathbb{E}|\langle t, X\rangle|^2)^{1/2} = p.$$

Also note that $(\mathbb{E}|X|^2)^{1/2} = \sqrt{N}$. Combining these estimates together with inequality (3.1), we get that $(\mathbb{E}|X|^p)^{1/p} \leq C(\sqrt{N} + p)$. Using Chebyshev's inequality we conclude that there exists $C > 0$ such that for every isotropic log-concave random vector $X \in \mathbb{R}^N$ and every $s \geq 1$

$$\mathbb{P}\left(|X| \geq C\, s\sqrt{N}\right) \leq e^{-s\sqrt{N}} \tag{3.2}$$

which is Theorem 1.1 from [24].

2. It is well known and it follows from [9] that for any $p \geq 1$, $(\mathbb{E}|X|^{2p})^{1/2p} \leq C(\mathbb{E}|X|^p)^{1/p}$ where C is an absolute constant. From the comparison between the first and second moment it is clear that inequality (3.1) is an equivalence. Moreover, there exists $C > 0$ such that

$$\mathbb{P}\left(|X| \geq C\Big((\mathbb{E}|X|^2)^{1/2} + \sup_{t\in S^{N-1}} (\mathbb{E}|\langle t, X\rangle|^p)^{1/p}\Big)\right) \leq e^{-p}$$

and

$$\mathbb{P}\left(|X| \geq \frac{1}{C}\Big((\mathbb{E}|X|^2)^{1/2} + \sup_{t\in S^{N-1}} (\mathbb{E}|\langle t, X\rangle|^p)^{1/p}\Big)\right) \geq \min\Big\{\frac{1}{C}, e^{-p}\Big\}.$$

The upper bound follows trivially from Chebyshev's inequality. The lower bound is a consequence of Paley-Zygmund's inequality and comparison between the p-th and $(2p)$-th moments of $|X|$.

3. Since for any Euclidean norm $\|\cdot\|$ on $\mathbb{R}^N$ there exists a linear map T such that $\|x\| = |Tx|$ and the class of log-concave random vectors is closed under linear transformations, Theorem 3.1 implies that for any N-dimensional log-concave vector X, any Euclidean norm $\|\cdot\|$ on $\mathbb{R}^N$ and $p \geq 1$ we have

$$(\mathbb{E}\|X\|^p)^{1/p} \leq C\Big((\mathbb{E}\|X\|^2)^{1/2} + \sup_{\|t\|_*\leq 1} (\mathbb{E}|\langle t, X\rangle|^p)^{1/p}\Big),$$

where $(\mathbb{R}^N, \|\cdot\|_*)$ is the dual space to $(\mathbb{R}^N, \|\cdot\|)$. It is an open problem whether such an inequality holds for arbitrary norms – see [19] for a discussion of this question and for related results.

We now introduce our main technical notations. For a random vector $X = (X(1), \ldots, X(N))$ in $\mathbb{R}^N$, $p \geq 1$ and $t > 0$ consider the functions

$$\sigma_X(p) = \sup_{t\in S^{N-1}} (\mathbb{E}|\langle t, X\rangle|^p)^{1/p} \tag{3.3}$$

and

$$N_X(t) = \sum_{i=1}^N \mathbf{1}_{\{X(i)\geq t\}}.$$

That is, $N_X(t)$ is equal to the number of coordinates of X larger than or equal to t. By σ_X^{-1} we denote the inverse of σ_X i.e.,

$$\sigma_X^{-1}(s) = \sup\{t\colon \sigma_X(t) \leq s\}.$$

Remark 1 after Theorem 3.1 implies that for isotropic vectors X, $\sigma_X(tp) \leq 2t\sigma_X(p)$ for $p \geq 2, t \geq 1$ and $\sigma_X^{-1}(2ts) \geq t\sigma_X^{-1}(s)$ for $t, s \geq 1$.

We also denote a nonincreasing rearrangement of $|X(1)|, \ldots, |X(N)|$ by $X^*(1) \geq X^*(2) \geq \ldots \geq X^*(N)$.

One of the main technical tools of this paper says:

Theorem 3.2. *For any N-dimensional log-concave isotropic random vector X, $p \geq 2$ and $t \geq C \log\Big(Nt^2/\sigma_X^2(p)\Big)$ we have*

$$\mathbb{E}(t^2 N_X(t))^p \leq (C\sigma_X(p))^{2p},$$

where C is an absolute positive constant.

We apply Theorem 3.2 to obtain probability estimates on order statistics $X^*(i)$'s.

Theorem 3.3. *For any N-dimensional log-concave random isotropic vector X, any $1 \leq \ell \leq N$ and $t \geq C\log(eN/\ell)$,*

$$\mathbb{P}(X^*(\ell) \geq t) \leq \exp\Big(- \sigma_X^{-1}\Big(\frac{1}{C} t\sqrt{\ell}\Big)\Big),$$

where C is an absolute positive constant.

Proof. Observe that $\sigma_{-X}(p) = \sigma_X(p)$ and that $X^*(\ell) \geq t$ implies that $N_X(t) \geq \ell/2$ or $N_{-X}(t) \geq \ell/2$. So by Chebyshev's inequality and Theorem 3.2,

$$\mathbb{P}(X^*(\ell) \geq t) \leq \Big(\frac{2}{\ell}\Big)^p (\mathbb{E}N_X(t)^p + \mathbb{E}N_{-X}(t)^p) \leq \Big(\frac{C'\sigma_X(p)}{t\sqrt{\ell}}\Big)^{2p}$$

provided that $t \geq C'' \log(Nt^2/\sigma_X^2(p))$, where C', C'' are absolute positive constants. To conclude the proof it is enough to take $p = \sigma_X^{-1}(\frac{1}{C'e} t\sqrt{\ell})$ and to notice that the restriction on t follows by the condition $t \geq C\log(eN/\ell)$. □

We can now state one of the main results of this paper.

Theorem 3.4. *Let X be an isotropic log-concave random vector in $\mathbb{R}^N$ and $m \leq N$. For any $t \geq 1$,*

$$\mathbb{P}\left(\sup_{\substack{I\subset\{1,\ldots,N\}\\ |I|=m}} |P_I X| \geq Ct\sqrt{m}\log\left(\frac{eN}{m}\right)\right) \leq \exp\left(-\sigma_X^{-1}\left(\frac{t\sqrt{m}\log\left(\frac{eN}{m}\right)}{\sqrt{\log(em/m_0)}}\right)\right),$$

where C is an absolute positive constant and

$$m_0 = m_0(X,t) = \sup\left\{k \leq m\colon\ k\log\left(\frac{eN}{k}\right) \leq \sigma_X^{-1}\Big(t\sqrt{m}\log\left(\frac{eN}{m}\right)\Big)\right\}.$$

Remark. We believe that the probability estimate should not contain any logarithmic term in the denominator, but it seems that our methods fail to show it. However it is not crucial in the sequel.

Since $\sigma_X(p) \le p$, Theorem 1.1 is an immediate consequence of Theorem 3.4.

4 Tail estimates for projections of sums of log-concave random vectors

We shall now study consequences that the results of Section 3 have for tail estimates for Euclidean norms of projections of sums of log-concave random vectors. Namely, we investigate the behavior of a random vector $Y = Y_x = \sum_{i=1}^n x_i X_i$, where $X_1, \dots, X_n$ are independent isotropic log-concave random vectors in $\mathbb{R}^N$ and $x = (x_i)_1^n \in \mathbb{R}^n$ is a fixed vector. We provide uniform bounds on projections of such a vector. We start with the following proposition.

Proposition 4.1. *Let $X_1, \dots, X_n$ be independent isotropic log-concave random vectors in $\mathbb{R}^N$, $x = (x_i)_1^n \in \mathbb{R}^n$, and $Y = \sum_{i=1}^n x_i X_i$. Then for every $p \ge 1$ one has*

$$\sigma_Y(p) = \sup_{t \in S^{N-1}} (\mathbb{E}|\langle t, Y\rangle|^p)^{1/p} \le C(\sqrt{p}|x| + p\|x\|_\infty),$$

where C is an absolute positive constant.

Proof. For every $t \in S^{N-1}$ we have

$$\langle t, Y\rangle = \sum_{i=1}^n x_i \langle t, X_i\rangle.$$

Let E_i be independent symmetric exponential random variables with variance 1. Let $t \in S^{N-1}$ and $x = (x_i)_1^n \in \mathbb{R}^n$. The variables $Z_i = \langle t, X_i\rangle$ are one dimensional centered log-concave with variance 1, therefore by (2.1) for every $s > 0$ one has

$$\mathbb{P}\left(|Z_i| \ge s\right) \le C_0\, \mathbb{P}\left(|E_i| \ge s/C_0\right).$$

Let (ε_i) be independent Bernoulli ± 1 random variables, independent also from (Z_i). A classical symmetrization argument and Lemma 4.6 of [21] imply that there exists C such that

$$(\mathbb{E}|\langle t, Y\rangle|^p)^{1/p} \le 2\Big(\mathbb{E}\Big|\sum_{i=1}^n x_i\varepsilon_i Z_i\Big|^p\Big)^{1/p} \le C\Big(\mathbb{E}\Big|\sum_{i=1}^n x_i E_i\Big|^p\Big)^{1/p}.$$

The well-known estimate (which follows e.g. from Theorem 1 in [16])

$$\left(\mathbb{E}\Big|\sum_{i=1}^n x_i E_i\Big|^p\right)^{1/p} \le C(\sqrt{p}|x| + p\|x\|_\infty)$$

concludes the proof. □

Corollary 4.2. *Let $X_1, \dots, X_n$, x and Y be as in Proposition 4.1 and $1 \le \ell \le N$. Then for any $t \ge C|x|\log\left(\frac{eN}{\ell}\right)$ one has*

$$\mathbb{P}(Y^*(\ell) \ge t) \le \exp\left(-\frac{1}{C}\min\left\{\frac{t^2\ell}{|x|^2}, \frac{t\sqrt{\ell}}{\|x\|_\infty}\right\}\right),$$

where C is an absolute positive constant.

Proof. The vector $Z = Y/|x|$ is isotropic and log-concave. Moreover by Proposition 4.1 we have

$$\sigma_Z(p) = \frac{1}{|x|}\sigma_Y(p) \le C_1\Big(\sqrt{p} + p\frac{\|x\|_\infty}{|x|}\Big).$$

Therefore for every $t \ge C_1$

$$\sigma_Z^{-1}(t) \ge \frac{1}{C_2}\min\Big\{t^2, \frac{|x|}{\|x\|_\infty}t\Big\}$$

and by Theorem 3.3 we get for every $t \ge C_3|x|\log\left(\frac{eN}{\ell}\right)$

$$\begin{aligned}\mathbb{P}(Y^*(\ell) \ge t) &= \mathbb{P}\Big(Z^*(\ell) \ge \frac{t}{|x|}\Big) \le \exp\Big(-\sigma_Z^{-1}\Big(\frac{t\sqrt{\ell}}{C_4|x|}\Big)\Big)\\ &\le \exp\Big(-\frac{1}{C}\min\Big\{\frac{t^2\ell}{|x|^2}, \frac{t\sqrt{\ell}}{\|x\|_\infty}\Big\}\Big).\end{aligned}$$

□

The next theorem provides uniform estimates for the Euclidean norm of projections of sums Y_x, considered above, in terms of the Euclidean and ℓ_∞ norms of the vector $x \in \mathbb{R}^n$.

Theorem 4.3. *Let $X_1, \ldots, X_n$ be independent isotropic log-concave random vectors in $\mathbb{R}^N$, $x = (x_i)_1^n \in \mathbb{R}^n$, and $Y = \sum_{i=1}^n x_i X_i$. Assume that $|x| \leq 1$, $\|x\|_\infty \leq b \leq 1$ and let $1 \leq m \leq N$.*

i) If $b \geq \frac{1}{\sqrt{m}}$ then for any $t \geq 1$

$$\mathbb{P}\left(\sup_{\substack{I \subset \{1,\ldots,N\} \\ |I|=m}} |P_I Y| \geq C t \sqrt{m} \log\left(\frac{eN}{m}\right) \right) \leq \exp\left(- \frac{t\sqrt{m}\log\left(\frac{eN}{m}\right)}{b\sqrt{\log(e^2 b^2 m)}} \right);$$

ii) if $b \leq \frac{1}{\sqrt{m}}$ then for any $t \geq 1$

$$\mathbb{P}\left(\sup_{\substack{I \subset \{1,\ldots,N\} \\ |I|=m}} |P_I Y| \geq C t \sqrt{m} \log\left(\frac{eN}{m}\right) \right)$$
$$\leq \exp\left(- \min\left\{ t^2 m \log^2\left(\frac{eN}{m}\right), \frac{t}{b}\sqrt{m}\log\left(\frac{eN}{m}\right) \right\} \right),$$

where C is an absolute positive constant.

Remark. Basically the same proof as the one given below shows that in i) the term $\sqrt{\log(e^2 b^2 m)}$ may be replaced by $\sqrt{\log(e^2 b^2 m / t^2)}$ and the condition $b \geq \frac{1}{\sqrt{m}}$ by $b \geq \frac{t}{\sqrt{m}}$. We omit the details.

The proof of Theorem 4.3 is based on Theorem 3.4. Let us first note that we may assume that vector Y is isotropic, i.e. $|x| = 1$. Indeed, we may find vector $y = (y_1, \ldots, y_\ell)$ such that $\|y\|_\infty \leq b$ and $|x|^2 + |y|^2 = 1$ and take $Y' = \sum_{i=1}^{\ell} y_i G_i$, where G_i are i.i.d. canonical N-dimensional Gaussian vectors, independent of vectors X_i's. Then the vector $Y + Y'$ is isotropic, satisfies assumptions of the theorem and for any $u > 0$,

$$\mathbb{P}\left(\sup_{\substack{I \subset \{1,\ldots,N\} \\ |I|=m}} |P_I Y| \geq u \right) \leq 2\mathbb{P}\left(\sup_{\substack{I \subset \{1,\ldots,N\} \\ |I|=m}} |P_I (Y + Y')| \geq u \right).$$

Similarly as in the proof of Corollary 4.2, for $t \geq C$ we have

$$\sigma_Y^{-1}(t) \geq \frac{1}{C} \min\left\{ t^2, \frac{t}{\|x\|_\infty} \right\}. \tag{4.1}$$

This allows us to estimate the quantity m_0 in Theorem 3.4. For $1/\sqrt{m} \leq b \leq 1$ define $m_1 = m_1(b) > 0$ by the equation

$$m_1(b) \log\Big(\frac{eN}{m_1(b)}\Big) = \frac{\sqrt{m}}{b} \log\Big(\frac{eN}{m}\Big). \tag{4.2}$$

One may show that $m_1(b) \sim \frac{\sqrt{m}}{b} \log(\frac{eN}{m})/\log(\frac{eNb}{\sqrt{m}})$, we will however need only the following simple estimate.

Lemma 4.4. *If* $1/\sqrt{m} \leq b \leq 1$ *then* $\log(m/m_1(b)) \leq 2\log(eb\sqrt{m})$.

Proof. Let $f(z) = z\log(eN/z)$. Using $1/\sqrt{m} \leq b$ we observe

$$\begin{aligned} f\Big(\frac{1}{e^2b^2}\Big) &= \frac{1}{e^2b^2}\Big(\log\Big(\frac{eN}{m}\Big) + \log(me^2b^2)\Big) \leq \frac{\sqrt{m}}{e^2b}\log\Big(\frac{eN}{m}\Big) + \frac{1}{e^2b^2}\sqrt{m}eb \\ &\leq \frac{\sqrt{m}}{b}\log\Big(\frac{eN}{m}\Big)\Big(\frac{1}{e^2} + \frac{1}{e}\Big) < \frac{\sqrt{m}}{b}\log\Big(\frac{eN}{m}\Big) = f(m_1(b)). \end{aligned}$$

Since f increases on $(0, N]$, we obtain $m_1(b) \geq (eb)^{-2}$, which implies the result. $\square$

Proof of Theorem 4.3. As we noticed after remark following Theorem 4.3, without loss of generality we may assume that $|x| = 1$, i.e. that Y is isotropic.

i) Assume $b \geq 1/\sqrt{m}$. By (4.1) for every $t \geq C/(\sqrt{m}\log(eN/m))$ we have

$$\sigma_Y^{-1}\Big(t\sqrt{m}\log\Big(\frac{eN}{m}\Big)\Big) \geq \frac{t}{Cb}\sqrt{m}\log\Big(\frac{eN}{m}\Big). \tag{4.3}$$

By (4.2) it follows that for every $t \geq |x| = 1$

$$m_1(b)\log\Big(\frac{eN}{m_1(b)}\Big) \leq \sigma_Y^{-1}\Big(Ct\sqrt{m}\log\Big(\frac{eN}{m}\Big)\Big).$$

By the definition of m_0, given in Theorem 3.4, this implies that $m_0(Y, Ct) \geq \lfloor m_1(b) \rfloor$, and since $m_0(Y, Ct) \geq 1$ we get $m_0(Y, Ct) \geq m_1(b)/2$. By Lemma 4.4 this yields $\log(em/m_0(Y, Ct)) \leq 2 + 2\log(eb\sqrt{m}) \leq 4\log(eb\sqrt{m})$.

Writing $t = t'\sqrt{2\log(e^2b^2m)}$ and applying (4.3) we obtain

$$\sigma_Y^{-1}\left(\frac{t\sqrt{m}\log\big(\frac{eN}{m}\big)}{\sqrt{\log\frac{em}{m_0(Y,Ct)}}}\right) \geq \sigma_Y^{-1}\left(t'\sqrt{m}\log\left(\frac{eN}{m}\right)\right) \geq \frac{t'}{Cb}\sqrt{m}\log\Big(\frac{eN}{m}\Big).$$

Theorem 3.4 applied to Ct instead of t implies the result (one needs to adjust absolute constants).
ii) Assume $b \le 1/\sqrt{m}$. By (4.1) for every $t \ge C|x|$ we have

$$\sigma_Y^{-1}\Big(t\sqrt{m}\log\Big(\frac{eN}{m}\Big)\Big) \ge \frac{1}{C}\min\Big\{t^2 m\log^2\Big(\frac{eN}{m}\Big), \frac{t}{b}\sqrt{m}\log\Big(\frac{eN}{m}\Big)\Big\}$$
$$\ge \frac{t}{C} m\log\Big(\frac{eN}{m}\Big),$$

which by the definition of m_0 implies that $m_0(Y, Ct) = m$. As in part i), Theorem 3.4 implies the result. $\square$

5 Uniform bounds for norms of sub-matrices

In this section we establish uniform estimates for norms of submatrices of a random matrix, namely for the quantity $A_{k,m}$ defined below.

Fix integers n and N. Let $X_1, \dots, X_n \in \mathbb{R}^N$ be independent log-concave isotropic random vectors. Let A be the $n \times N$ random matrix with rows $X_1, \dots, X_n$.

For any subsets $J \subset \{1, \dots, n\}$ and $I \subset \{1, \dots, N\}$, by $A(J, I)$ we denote the submatrix of A consisting of the rows indexed by elements from J and the columns indexed by elements from I.

Let $k \le n$ and $m \le N$. We define the parameter $A_{k,m}$ by

$$A_{k,m} = \sup \|A(J, I)\|_{\ell_2^m \to \ell_2^k}, \tag{5.1}$$

where the supremum is taken over all subsets $J \subset \{1, \dots, n\}$ and $I \subset \{1, \dots, N\}$ with cardinalities $|J| = k, |I| = m$. That is, $A_{k,m}$ is the maximal operator norm of a submatrix of A with k rows and m columns.

It is often more convenient to work with matrices with log-concave columns rather than rows, therefore in this section we fix the notation

$$\Gamma = A^*.$$

Thus Γ is an $N \times n$ matrix with columns $X_1, \dots, X_n$. In particular, given $x \in \mathbb{R}^n$ the sum $Y = \sum_{i=1}^n x_i X_i$ considered in Section 4 satisfies $Y = \Gamma x$. Clearly,

$$\Gamma(I, J) = \Big(A(J, I)\Big)^*$$

so that, recalling that U_k was defined in (2.2), we have

$$A_{k,m} = \Gamma_{m,k} = \sup\{|P_I \Gamma x| \colon I \subset \{1, \dots, N\},\ |I| = m,\ x \in U_k\}. \tag{5.2}$$

Define $\lambda_{k,m}$ and λ_m by

$$\lambda_{k,m} = \sqrt{\log\log(3m)}\ \sqrt{m}\log\Big(\frac{e\max\{N,n\}}{m}\Big) + \sqrt{k}\log\Big(\frac{en}{k}\Big), \tag{5.3}$$

and

$$\lambda_m = \frac{\sqrt{\log\log(3m)}\ \sqrt{m}}{\sqrt{\log(3m)}}\ \log\Big(\frac{e\max\{N,n\}}{m}\Big). \tag{5.4}$$

The following theorem is our main result providing estimates for the operator norms of submatrices of A (and of Γ). Its first part in the case $n \le N$ was stated as Theorem 1.2.

Theorem 5.1. *There exists a positive absolute constant C such that for any positive integers $n, N,\ k \le n,\ m \le N$ and any $t \ge 1$ one has*

$$\mathbb{P}\left(A_{k,m} \ge Ct\lambda_{k,m}\right) \le \exp\left(-\frac{t\lambda_{k,m}}{\sqrt{\log(3m)}}\right).$$

In particular, there exists an absolute positive constant C_1 such that for every $t \ge 1$ and for every $m \le N$ one has

$$\mathbb{P}\left(\exists k\ \ A_{k,m} \ge C_1 t\lambda_{k,m}\right) \le \exp\left(-t\lambda_m\right). \tag{5.5}$$

First we show the "in particular" part, which is easy.

Proof of inequality (5.5). The main part of the theorem implies that for every $t \ge 1$

$$p_m := \mathbb{P}\left(\exists k\ \ A_{k,m} \ge Ct\lambda_{k,m}\right) \le \sum_{k=1}^{n} \exp\left(-t\lambda_{k,m}/\sqrt{\log(3m)}\right).$$

Thus if $m > \log^3 n$ then for every $t \ge 100$ one has

$$p_m \le n\exp\left(-t\lambda_m\right) \le \exp\left(-(t/2)\ \lambda_m\right).$$

If $m \le \log^3 n$ (in particular $n \ge 3$) then for every $t \ge 100$ one has

$$p_m \le \sum_{k \le (\log n)^2} \exp\left(-t\lambda_{k,m}/\sqrt{\log(3m)}\right) + \sum_{k \ge (\log n)^2} \exp\left(-t\lambda_{k,m}/\sqrt{\log(3m)}\right)$$

$$\le (\log n)^2 \exp(-t\lambda_m) + \exp(-t\lambda_m) \sum_{k \ge (\log n)^2} \exp(-t a_{k,m}),$$

where

$$a_{k,m} = \frac{\sqrt{k}}{\sqrt{\log(3m)}} \log\left(\frac{en}{k}\right).$$

Since $m \le \log^3 n$, we obtain for every $t \ge 100$

$$p_m \le \exp(-(t/2)\lambda_m) + \exp(-t\lambda_m) \le \exp(-(t/4)\lambda_m).$$

The result follows by writing $t = 100t'$ and by adjusting absolute constants. □

Now we prove the main part of the theorem. Its proof consists of two steps that depend on the relation between m and k. The Step I is applicable if $m \log\left(\frac{eN}{m}\right) < k \log\left(\frac{en}{k}\right)$, and it reduces this case to the second complementary case $m \log\left(\frac{eN}{m}\right) \ge k \log\left(\frac{en}{k}\right)$. The latter case will then be treated in Step II. To make this reduction we define k' as follows

$$k' = \inf\left\{\tilde{k} \in \mathbb{N} \ : \ \tilde{k} \le n \ \text{ and } \ \tilde{k} \log\left(\frac{en}{\tilde{k}}\right) \ge m \log\left(\frac{eN}{m}\right)\right\} \tag{5.6}$$

(of course if the set in (5.6) is empty, we immediately pass to Step II).

5.1 Step I: $k \log\left(\frac{en}{k}\right) > m \log\left(\frac{eN}{m}\right)$, in particular $k \ge k'$.

Proposition 5.2. *Assume that $k \ge k'$. Then for any $t \ge 1$ we have*

$$\sup_{\substack{I \subseteq \{1,\dots,N\} \\ |I|=m}} \sup_{x \in U_k} |P_I \Gamma x| \le C\Bigg(\sup_{\substack{I \subseteq \{1,\dots,N\} \\ |I|=m}} \sup_{x \in U_{k'}} |P_I \Gamma x| + t\sqrt{m} \log\left(\frac{eN}{m}\right) + t\sqrt{k} \log\left(\frac{en}{k}\right)\Bigg) \tag{5.7}$$

with probability at least

$$1-n\exp\left(-t\frac{\sqrt{m}\log(eN/m)+\sqrt{k}\log(en/k)}{\sqrt{\log em}}\right)-\exp\left(-tk'\log\left(\frac{en}{k'}\right)\right), \quad (5.8)$$

where C is a positive absolute constant.

The proof of Proposition 5.2 is based on the ideas from [5]. We start it with the following fact.

Proposition 5.3. *Let $(X_i)_{i\le n}$ be independent centered random vectors in $\mathbb{R}^N$ and $\psi>0$ be such that*

$$\mathbb{E}\exp\left(\frac{|\langle X_i,\theta\rangle|}{\psi}\right)\le 2 \qquad \textit{for all } i\le n, \theta\in S^{N-1}.$$

Then for $1\le p\le n$ and $t\ge 1$ with probability at least $1-\exp(-tp\log(en/p))$ the following holds:
for all $y,z\in U_p$ and all $E,F\subset\{1,\dots,n\}$ with $E\cap F=\emptyset$,

$$\Big|\Big\langle\sum_{i\in E} y_iX_i,\sum_{j\in F} z_jX_j\Big\rangle\Big|\le 20tp\log\left(\frac{en}{p}\right)\psi\max_{i\in E}|y_i|\Big(\sum_{j\in F} z_j^2\Big)^{1/2}\Gamma_p,$$

where $\Gamma_p:=\max_{x\in U_p}|\Gamma x|=\max_{x\in U_p}|\sum_{i=1}^n x_iX_i|$.

Proof. In this proof we use for simplicity the notation $[n]=\{1,\dots,n\}$. First let us fix sets $E,F\subset[n]$ with $E\cap F=\emptyset$. Since we consider $y,z\in U_p$, without loss of generality we may assume that $|E|,|F|\le p$. For $z\in U_p$ denote

$$Y_F(z)=\sum_{j\in F} z_jX_j \quad\text{and}\quad Z_F(z)=\frac{Y_F(z)}{|Y_F(z)|}$$

(if $Y_F(z)=0$ we set $Z_F(z)=0$). For any $y,z\in U_p$ we have

$$\begin{aligned}\Big|\Big\langle\sum_{i\in E} y_iX_i,\sum_{j\in F} z_jX_j\Big\rangle\Big| &\le \sum_{i\in E}|y_i|\Big|\Big\langle X_i,Y_F(z)\Big\rangle\Big|\\ &\le \max_{i\in E}|y_i|\,|Y_F(z)|\sum_{i\in E}\Big|\Big\langle X_i,Z_F(z)\Big\rangle\Big|.\end{aligned}$$

The random vector $Z_F(z)$ is independent from the vectors X_i's, $i \in E$, moreover $|Z_F(z)| \leq 1$ and $|Y_F(z)| \leq (\sum_{j\in F} z_j^2)^{1/2}\Gamma_p$. Therefore for any $z \in U_p$ and $u > 0$,

$$\begin{aligned}
\mathbb{P}\Big(\exists_{y\in U_p} \ \Big|\Big\langle \sum_{i\in E} y_i X_i, \sum_{j\in F} z_j X_j\Big\rangle\Big| > u\psi \max_{i\in E}|y_i|\Big(\sum_{j\in F} z_j^2\Big)^{1/2}\Gamma_p\Big) \\
\leq \mathbb{P}\Big(\sum_{i\in E}\Big|\Big\langle X_i, Z_F(z)\Big\rangle\Big| \geq u\psi\Big) \\
\leq e^{-u}\mathbb{E}\exp\Big(\sum_{i\in E}\frac{|\langle X_i, Z_F(z)\rangle|}{\psi}\Big) \leq 2^{|E|}e^{-u} \leq 2^p e^{-u}.
\end{aligned}$$

Let N_F denote a 1/2-net in the Euclidean metric in $B_2^n \cap \mathbb{R}^F$ of cardinality at most $5^{|F|} \leq 5^p$. We have

$$\begin{aligned}
&p_{E,F}(u) \\
&\quad := \mathbb{P}\Big(\exists_{y\in U_p}\exists_{z\in U_p} \ \Big|\Big\langle \sum_{i\in E} y_i X_i, \sum_{j\in F} z_j X_j\Big\rangle\Big| > 2u\psi \max_{i\in E}|y_i|\Big(\sum_{j\in F} z_j^2\Big)^{1/2}\Gamma_p\Big) \\
&\quad \leq \mathbb{P}\Big(\exists_{y\in U_p}\exists_{z\in N_F} \ \Big|\Big\langle \sum_{i\in E} y_i X_i, \sum_{j\in F} z_j X_j\Big\rangle\Big| > u\psi \max_{i\in E}|y_i|\Big(\sum_{j\in F} z_j^2\Big)^{1/2}\Gamma_p\Big) \\
&\quad \leq \sum_{z\in N_F}\mathbb{P}\Big(\exists_{y\in U_p} \ \Big|\Big\langle \sum_{i\in E} y_i X_i, \sum_{j\in F} z_j X_j\Big\rangle\Big| > u\psi \max_{i\in E}|y_i|\Big(\sum_{j\in F} z_j^2\Big)^{1/2}\Gamma_p\Big) \\
&\quad \leq 10^p e^{-u}.
\end{aligned}$$

Hence

$$\begin{aligned}
&\mathbb{P}\Big(\exists_{y\in U_p}\exists_{z\in U_p}\exists_{E,F\subset[n],|E|,|F|\leq p, E\cap F=\emptyset} \\
&\quad \Big|\Big\langle \sum_{i\in E} y_i X_i, \sum_{j\in F} z_j X_j\Big\rangle\Big| > 2u\psi \max_{i\in E}|y_i|\Big(\sum_{j\in F} z_j^2\Big)^{1/2}\Gamma_p\Big) \\
&\qquad \leq \sum_{E,F\subset[n],|E|,|F|\leq p, E\cap F=\emptyset} p_{E,F}(u) \leq \binom{n}{p}\binom{n}{p}10^p e^{-u} \leq e^{-u/2},
\end{aligned}$$

provided that $u \geq 10p\log(ne/p)$. This implies the desired result. ☐

Before formulating the next proposition we recall the following elementary lemma (see e.g. Lemma 3.2 in [5]).

Lemma 5.4. *Let $x_1, \ldots, x_n \in \mathbb{R}^N$, then*

$$\sum_{i \neq j} \langle x_i, x_j \rangle \le 4 \max_{E \subset \{1,\ldots,n\}} \sum_{i \in E} \sum_{j \in E^c} \langle x_i, x_j \rangle.$$

Proposition 5.5. *Let $(X_i)_{i \le n}$ be independent centered random vectors in $\mathbb{R}^N$ and $\psi > 0$ be such that*

$$\mathbb{E} \exp\Big(\frac{|\langle X_i, \theta \rangle|}{\psi} \Big) \le 2 \qquad \textit{for all } i \le n, \theta \in S^{N-1}.$$

Let $p \le n$ and $t \ge 1$. Then with probability at least $1 - \exp(-tp \ln(en/p))$ for all $x \in U_p$,

$$\Big| \sum_{i=1}^n x_i X_i \Big| \le C \Big(|x| \max_i |X_i| + tp \log\Big(\frac{en}{p}\Big) \psi \|x\|_\infty \Big),$$

where C is an absolute constant.

Proof. As in the previous proof we set $[n] = \{1, \ldots, n\}$. Fix $\alpha > 0$ and define

$$\Gamma_p(\alpha) = \sup \Big\{ \Big| \sum_{i=1}^n x_i X_i \Big| : \; x \in U_p, \|x\|_\infty \le \alpha \Big\}.$$

For $E \subset [n]$ with $|E| \le p$ let $N_E(\alpha)$ denote a $1/2$-net in $\mathbb{R}^E \cap B_2^n \cap \alpha B_\infty$ with respect to the metric defined by $B_2^n \cap \alpha B_\infty$. We may choose $N_E(\alpha)$ of cardinality $5^{|E|} \le 5^p$. Let $N(\alpha) = \bigcup_{|E|=p} N_E(\alpha)$, then

$$|N(\alpha)| \le \Big(\frac{5en}{p} \Big)^p \quad \text{and} \quad \Gamma_p(\alpha) \le 2 \sup_{x \in N(\alpha)} \Big| \sum_{i=1}^n x_i X_i \Big|. \tag{5.9}$$

Fix $E \subset [n]$ with $|E| = p$ and $x \in N_E(\alpha)$. We have

$$\Big| \sum_{i=1}^n x_i X_i \Big|^2 = \sum_{i=1}^n x_i^2 |X_i|^2 + \sum_{i \neq j} \langle x_i X_i, x_j X_j \rangle.$$

Therefore Lemma 5.4 gives

$$\Big| \sum_{i=1}^n x_i X_i \Big|^2 \le \max_i |X_i|^2 + 4 \sup_{F \subset E} \Big| \Big\langle \sum_{i \in F} x_i X_i, \sum_{j \in E \setminus F} x_j X_j \Big\rangle \Big|.$$

Notice that for any $F \subset E$, $\max_{i\in F}|x_i| \le \alpha$ and $|\sum_{j\in E\setminus F} x_j X_j| \le \Gamma_p(\alpha)$, hence as in the proof of Proposition 5.3 we can show that

$$\mathbb{P}\Big(\Big|\Big\langle \sum_{i\in F} x_i X_i, \sum_{j\in E\setminus F} x_j X_j\Big\rangle\Big| > u\psi\alpha\Gamma_p(\alpha)\Big) < 2^{|F|}e^{-u}.$$

Thus

$$\mathbb{P}\Big(\Big|\sum_{i=1}^n x_i X_i\Big|^2 > \max_i |X_i|^2 + 4u\psi\alpha\Gamma_p(\alpha)\Big) \le \sum_{F\subset E} 2^{|F|}e^{-u} \le 3^{|E|}e^{-u}.$$

This together with (5.9) and the union bound implies

$$\mathbb{P}\Big(\Gamma_p(\alpha)^2 > 4\max_i |X_i|^2 + 16u\psi\alpha\Gamma_p(\alpha)\Big) \le \sum_{x\in N(\alpha)} 3^p e^{-u} \le \Big(\frac{15en}{p}\Big)^p e^{-u}.$$

Hence

$$\mathbb{P}\Big(\Gamma_p(\alpha) > 2\sqrt{2}\max_i |X_i| + 32u\psi\alpha\Big) \le \Big(\frac{15en}{p}\Big)^p e^{-u}. \tag{5.10}$$

Using that $\Gamma_p(\alpha) \ge \Gamma_p(\beta)$ for $\alpha \ge \beta > 0$ we obtain for every $\ell \ge 1$

$$\begin{aligned}
&\mathbb{P}\Big(\exists_{x\in U_p}\ \Big|\sum_{i=1}^n x_i X_i\Big| > 2\sqrt{2}\max_i |X_i| + u\psi\max\{\|x\|_\infty, 2^{-\ell}\}\Big)\\
&= \mathbb{P}\Big(\exists_{2^{-\ell}\le\alpha\le 1}\ \Gamma_p(\alpha) > 2\sqrt{2}\max_i |X_i| + u\psi\alpha\Big)\\
&\le \mathbb{P}\Big(\exists_{0\le j\le \ell-1}\ \Gamma_p(2^{-j}) > 2\sqrt{2}\max_i |X_i| + \frac{1}{2}u\psi 2^{-j}\Big)\\
&\le \sum_{j=0}^{\ell-1}\mathbb{P}\Big(\Gamma_p(2^{-j}) > 2\sqrt{2}\max_i |X_i| + \frac{1}{2}u\psi 2^{-j}\Big) \le \ell\Big(\frac{15en}{p}\Big)^p e^{-u/64},
\end{aligned}$$

where the last inequality follows by (5.10). Taking $\ell \approx \log p$ (so that $\|x\|_\infty \ge 2^{-\ell}|x|$) and $u = Ctp\log(en/p)$, we obtain the result. $\square$

Proof of Proposition 5.2. For any $I \subset \{1,\ldots,N\}$, the vector $P_I X$ is isotropic and log-concave in $\mathbb{R}^I$, hence it satisfies the ψ_1 bound with a universal constant.

We fix $t \geq 10$. Let $s \geq 1$ be the smallest integer such that $k2^{-s} < k'$. Set $k_\mu = \lfloor k2^{1-\mu} \rfloor - \lfloor k2^{-\mu} \rfloor$ for $\mu = 1, \ldots, s$ and $k_{s+1} = \lfloor k2^{-s} \rfloor$. Then

$$\max\{1, \lfloor k2^{-\mu} \rfloor\} \leq k_\mu \leq k2^{1-\mu}, \ \frac{k'}{2} \leq k_{s+1} \leq k', \ \text{ and } \ \sum_{\mu=1}^{s+1} k_\mu = k. \quad (5.11)$$

Consider an arbitrary vector $x = (x(i))_i \in U_k$ and let $n_1, \ldots, n_k$ be pairwise distinct integers such that $|x(n_1)| \leq |x(n_2)| \leq \ldots \leq |x(n_k)|$ and $x(i) = 0$ for $i \notin \{n_1, \ldots, n_k\}$. For $\mu = 1, \ldots, s+1$ let $F_\mu = \{n_i\}_{j_\mu < i \leq j_{\mu+1}}$, where $j_\mu = \sum_{r<\mu} k_r$ ($j_1 = 0$). Let x_{F_μ} be the coordinate projection of x onto $\mathbb{R}^{F_\mu}$. Note that for each $\mu \leq s$ we have $|x_{F_\mu}| \leq 1$ and $\|x_{F_\mu}\|_\infty \leq 1/\sqrt{k - j_{\mu+1} + 1} \leq \sqrt{2^\mu / k}$.

The equality $x = \sum_{\mu=1}^{s+1} x_{F_\mu}$ yields that for every $I \subseteq \{1, \ldots, N\}$ of cardinality m,

$$|P_I \Gamma x| \leq |P_I \Gamma x_{s+1}| + \Big| \sum_{\mu=1}^{s} P_I \Gamma x_{F_\mu} \Big| \leq \sup_{\substack{I \subseteq \{1,\ldots,N\} \\ |I|=m}} \sup_{x \in U_{k'}} |P_I \Gamma x| + \Big| \sum_{\mu=1}^{s} P_I \Gamma x_{F_\mu} \Big|,$$

where in the second inequality we used that $k_{s+1} \leq k'$.

Taking the suprema over I of cardinality m and $x \in U_k$ we obtain

$$A_{k,m} = \sup_{\substack{I \subseteq \{1,\ldots,N\} \\ |I|=m}} \sup_{x \in U_k} |P_I \Gamma x| \leq \sup_{\substack{I \subseteq \{1,\ldots,N\} \\ |I|=m}} \sup_{x \in U_{k'}} |P_I \Gamma x| + \sup_{\substack{I \subseteq \{1,\ldots,N\} \\ |I|=m}} \sup_{x \in U_k} \Big| \sum_{\mu=1}^{s} P_I \Gamma x_{F_\mu} \Big|. \quad (5.12)$$

Note that

$$\Big| \sum_{\mu=1}^{s} P_I \Gamma x_{F_\mu} \Big|^2 = \sum_{\mu=1}^{s} |P_I \Gamma x_{F_\mu}|^2 + 2 \sum_{\mu=1}^{s-1} \langle P_I \Gamma x_\mu, \sum_{\nu=\mu+1}^{s} P_I \Gamma x_\nu \rangle. \quad (5.13)$$

We are going to use Proposition 5.5 to estimate the first summand and Proposition 5.3 to estimate the second one. First note that by the definition of k' and s we have

$$\binom{N}{m} \leq \left(\frac{eN}{m}\right)^m \leq \exp\left(k' \log \frac{en}{k'}\right) \quad \text{ and } \quad \frac{k}{k'} < 2^s \leq \frac{2k}{k'} \leq \frac{2n}{k'}.$$

Hence, using the definition of k_μ's, we observe that for $t \geq 10$ we have

$$\binom{N}{m} \sum_{\mu=1}^{s} \exp\Big(- tk_\mu \log\Big(\frac{en}{k_\mu}\Big)\Big) \leq s \, \exp\Big(k' \log \frac{en}{k'}\Big) \exp\Big(- tk_s \log\Big(\frac{en}{k_s}\Big)\Big)$$
$$\leq \frac{1}{2} \, \exp\Big(- (tk'/2) \log\Big(\frac{en}{k'}\Big)\Big).$$

Since $x_{F_\mu} \in U_{k_\mu}$ for every $x \in U_k$ and $\mu = 1, \ldots, s$, the union bound and Proposition 5.5 imply that with probability at least

$$1 - \sum_{\mu=1}^{s} \binom{N}{m} \exp\Big(- tk_\mu \log\Big(\frac{en}{k_\mu}\Big)\Big) \geq 1 - \frac{1}{2} \, \exp\Big(- (tk'/2) \log\Big(\frac{en}{k'}\Big)\Big),$$

for every $x \in U_k$, every I of cardinality m, and every $\mu \in \{1, \ldots, s\}$ one has

$$|P_I \Gamma x_{F_\mu}| \leq C\Big(|x_{F_\mu}| \max_i |P_I X_i| + tk_\mu \log\Big(\frac{en}{k_\mu}\Big) \sqrt{\frac{2^\mu}{k}}\Big), \tag{5.14}$$

where C is an absolute constant.

Similarly, by Proposition 5.3, with probability at least

$$1 - \frac{1}{2} \, \exp\Big(- (tk'/2) \log\Big(\frac{en}{k'}\Big)\Big),$$

for every $x \in U_k$, every I of cardinality m and every $\mu \in \{1, \ldots, s\}$ one has

$$\langle P_I \Gamma x_\mu, \sum_{\nu=\mu+1}^{s} P_I \Gamma x_\nu \rangle \leq C t k_\mu \log\Big(\frac{en}{k_\mu}\Big) \sqrt{\frac{2^\mu}{k}} A_{k,m}, \tag{5.15}$$

where we have used the facts that $\sum_{\nu=\mu+1}^{s} x_\nu \in U_{k_\mu}$ and $\sum_\nu |x_\nu|^2 \leq 1$.

Using (5.12) – (5.15) we conclude that there exists an absolute constant $C_1 > 0$ such that with probability at least $1 - \exp(-(tk'/2)\log(en/k'))$,

$$A_{k,m}^2 \leq C\Big(\sup_{\substack{I \subseteq \{1,\ldots,N\} \\ |I|=m}} \sup_{x \in U_{k'}} |P_I \Gamma x|^2 + \sum_{\mu=1}^{s} |x_{F_\mu}|^2 \max_i \max_{\substack{I \subseteq \{1,\ldots,N\} \\ |I|=m}} |P_I X_i|^2$$
$$+ t^2 \sum_{\mu=1}^{s} \frac{k}{2^\mu} \log^2\Big(\frac{en2^\mu}{k}\Big) + t \sum_{\mu=1}^{s} \sqrt{\frac{k}{2^\mu}} \log\Big(\frac{en2^\mu}{k}\Big) A_{k,m}\Big)$$
$$\leq C_1\Big(\sup_{\substack{I \subseteq \{1,\ldots,N\} \\ |I|=m}} \sup_{x \in U_{k'}} |P_I \Gamma x|^2 + \max_i \max_{\substack{I \subseteq \{1,\ldots,N\} \\ |I|=m}} |P_I X_i|^2$$
$$+ t^2 k \log^2\Big(\frac{en}{k}\Big) + t\sqrt{k} \log\Big(\frac{en}{k}\Big) A_{k,m}\Big).$$

Thus, with the same probability

$$A_{k,m} \le C_2 \Big(\sup_{\substack{I \subseteq \{1,\dots,N\} \\ |I|=m}} \sup_{x \in U_{k'}} |P_I \Gamma x| + \max_i \max_{\substack{I \subseteq \{1,\dots,N\} \\ |I|=m}} |P_I X_i| + t\sqrt{k} \log\Big(\frac{en}{k}\Big)\Big),$$

where $C_2 > 0$ is an absolute constant.

But by Theorem 1.1 and the union bound we have for every $t \ge 1$,

$$\max_i \max_{\substack{I \subseteq \{1,\dots,N\} \\ |I|=m}} |P_I X_i| \le C_3 \Big(t\sqrt{m} \log(eN/m) + t\sqrt{k} \log(en/k) \Big),$$

with probability larger than or equal to

$$p_0 := 1 - n \exp\Big(-t \frac{\sqrt{m} \log(eN/m) + \sqrt{k} \log(en/k)}{\sqrt{\log em}} \Big)$$

(we added the term depending on k to get better probability, we may do it by adjusting t). This proves the result for $t \ge 10$ with probability $p_0 - \exp(-(tk'/2)\log(en/k'))$. Passing to $t_0 = t/20$ and adjusting absolute constants, we complete the proof. $\square$

5.2 Step II. $k \log\left(\frac{en}{k}\right) \le m \log\left(\frac{eN}{m}\right)$, in particular $k \le k'$.

In this case we have to be a little bit more careful than in the previous case with the choice of nets. We will need the following lemma, in which $\tilde{U}_k$ denotes the set of k-sparse vectors of the Euclidean norm at most one.

Lemma 5.6. *Suppose that $k \le n$, $k_1, k_2, \dots, k_s$ are positive integers such that $k_1 + \dots + k_s \ge k$ and $k_{s+1} = 1$. We may then find a finite subset $\mathcal{N}$ of $\frac{3}{2}\tilde{U}_k$ satisfying the following.*
i) For any $x \in U_k$ there exists $y \in \mathcal{N}$ such that $x - y \in \frac{1}{2}\tilde{U}_k$.
ii) Any $x \in \mathcal{N}$ may be represented in the form $x = \pi_1(x) + \dots + \pi_s(x)$, where vectors $\pi_1(x), \dots, \pi_s(x)$ have disjoint supports, $|\mathrm{supp}\,(\pi_i(x))| \le k_i$ for $i = 1, \dots, s$,

$$\sum_{i=1}^{s} k_{i+1} \|\pi_i(x)\|_\infty^2 \le 4$$

and

$$|\pi_i(\mathcal{N})| = |\{\pi_i(x) \colon x \in \mathcal{N}\}| \le \Big(\frac{en}{k_i}\Big)^{3k_i} \qquad \textit{for } i = 1, \dots, s.$$

Proof. First note that we can assume that $k_1 + k_2 + \ldots + k_s = k$. Indeed, otherwise denote by j the largest integer such that $k_j + k_{j+1} + \ldots + k_s \geq k$. If $j = s$ then set $\tilde{k}_j = k$, if $j < s$ then set $\tilde{k}_i = k_i$ for $j < i \leq s$, $\tilde{k}_j = k - k_{j+1} - k_{j+2} - \ldots - k_s$, $\pi_i(x) = 0$ for $i < j$ and repeat the proof below for the sequence $\tilde{k}_j, \tilde{k}_{j+1}, \ldots, \tilde{k}_s, \tilde{k}_{s+1}$, where $\tilde{k}_{s+1} = 1$ as before.

Recall that for $F \subset \{1, \ldots, n\}$, $\mathbb{R}^F$ denotes the set of all vectors in $\mathbb{R}^n$ with support contained in F.

For $i = 1, \ldots, s$ and $F \subset \{1, \ldots, n\}$ of cardinality at most k_i let $\mathcal{N}_i(F)$ denote the subset of $S_F(i) := \mathbb{R}^F \cap B_2^n \cap k_{i+1}^{-1/2} B_\infty^n$ such that

$$S_F(i) \subset \mathcal{N}_i(F) + \frac{k_i}{2n}\Big(B_2^n \cap k_{i+1}^{-1/2} B_\infty^n\Big).$$

Standard volumetric argument shows that we may choose $\mathcal{N}_i(F)$ of cardinality at most $(6n/k_i)^{|F|} \leq (6n/k_i)^{k_i}$ (additionally without loss of generality we assume that $0 \in \mathcal{N}_i(F)$). We set

$$\mathcal{N}_i := \bigcup_{|F| \leq k_i} \mathcal{N}_i(F),$$

then

$$|\mathcal{N}_i| \leq \left(\frac{en}{k_i}\right)^{k_i} \left(\frac{6n}{k_i}\right)^{k_i} \leq \left(\frac{en}{k_i}\right)^{3k_i}.$$

Fix $x \in U_k$, let F_s denote the set of indices of k_s largest coefficients of x, F_{s-1} – the set of indices of the next k_{s-1} largest coefficients, etc. Then $x = x_{F_s} + x_{F_{s-1}} + \ldots + x_{F_1}$, $\|x_{F_s}\|_\infty \leq 1$ and

$$\|x_{F_i}\|_\infty \leq \frac{1}{\sqrt{k_{i+1}}} |x_{F_{i+1}}| \leq \frac{1}{\sqrt{k_{i+1}}} \quad \text{for } i < s.$$

In particular, $x_{F_i} \in \mathbb{R}^{F_i} \cap B_2^n \cap k_{i+1}^{-1/2} B_\infty^n$ for all $i = 1, \ldots, s$. Let $\pi_i(x)$ be a vector in $\mathcal{N}_i(F_i)$ such that

$$|x_{F_i} - \pi_i(x)| \leq \frac{k_i}{2n} \quad \text{and} \quad \|x_{F_i} - \pi_i(x)\|_\infty \leq \frac{k_i}{2n\sqrt{k_{i+1}}}.$$

Define also $\pi(x) = \pi_1(x) + \ldots + \pi_s(x)$. Then

$$|x - \pi(x)| \leq \sum_{i=1}^{s} |x_{F_i} - \pi_i(x)| \leq \sum_{i=1}^{s} \frac{k_i}{2n} = \frac{k}{2n} \leq \frac{1}{2}$$

and

$$\sum_{i=1}^{s} k_{i+1}\|\pi_i(x)\|_\infty^2 \le 1 + 2\sum_{i=1}^{s-1} k_{i+1}(\|x_{F_i}\|_\infty^2 + \|x_{F_i} - \pi_i(x)\|_\infty^2)$$
$$\le 1 + 2\sum_{i=1}^{s-1}\Big(|x_{F_{i+1}}|^2 + \Big(\frac{k_i}{2n}\Big)^2\Big) \le 1 + 2|x|^2 + \frac{k^2}{2n^2} \le 4.$$

Thus we complete the proof by letting

$$\mathcal{N} = \{\pi(x) \colon x \in U_k\}.$$

□

Lemma 5.7. *Suppose that $n \le N$ and $k \le \min\{n, k'\}$. Then for some positive integer $s \le C\log\log(3m)$ we can find $s+1$ positive integers $k_1 = k$, $k_i \in [\frac{1}{6}m^{1/4}, m]$ for $2 \le i \le s$, $k_{s+1} = 1$ satisfying*

$$k_i \log\Big(\frac{en}{k_i}\Big) \le 20g(k_{i+1}), \qquad \text{for } i = 1, \dots, s, \tag{5.16}$$

where C is an absolute positive constant and

$$g(z) = \begin{cases} \frac{\sqrt{zm}}{\sqrt{\log(e^2m/z)}}\log\Big(\frac{eN}{m}\Big) & \text{if } z < m,\\ \min\Big\{\sqrt{zm}\log\Big(\frac{eN}{m}\Big), m\log^2\Big(\frac{eN}{m}\Big)\Big\} & \text{if } z \ge m.\end{cases}$$

Proof. Let us define

$$h(z) = z\log\Big(\frac{en}{z}\Big) \qquad \text{and} \qquad H(z) = z\log\Big(\frac{eN}{z}\Big).$$

Notice that $h(z) \le H(z)$, h is increasing on $(0, n]$ and H is increasing on $(0, N]$. It is also easy to see that $h(\lceil z\rceil) \le 2h(z)$ for $z \in [1, n]$.

We first establish some relations between the functions g and H. It is not hard to check that $\log^{3/2}(e^2m) \le e^2\sqrt{m}$, therefore for $z \in [1, m]$,

$$H\Big(\frac{\sqrt{zm}}{\log^{3/2}(e^2m)}\Big) = \frac{\sqrt{zm}}{\log^{3/2}(e^2m)}\Big(\log\Big(\frac{eN}{m}\Big) + \log\Big(\frac{m\log^{3/2}(e^2m)}{\sqrt{zm}}\Big)\Big)$$
$$\le \frac{\sqrt{zm}}{\log^{3/2}(e^2m)}\log\Big(\frac{eN}{m}\Big)(1 + \log(e^2m)) \le 2g(z). \tag{5.17}$$

Write $z = pm$ with $p \in (0,1)$, so $H(2z) = 2pm\left(\log\left(\frac{eN}{m}\right) + \log\left(\frac{m}{2z}\right)\right)$. Then

$$H(2z) \leq \frac{2\sqrt{p}m}{\sqrt{\log(e^2/p)}} \log\left(\frac{eN}{m}\right)\sqrt{p\log(e^2/p)}(1+\log(1/p)) \leq 10g(z), \quad (5.18)$$

where the last inequality follows since

$$\begin{aligned}\sup_{p\in(0,1)} 2\sqrt{p}\sqrt{\log(e^2/p)}(1+\log(1/p)) &= 2\sup_{u\geq 0} e^{-u}\sqrt{2+2u}(1+2u) \\ &\leq 2\sqrt{2}\sup_{u\geq 0} e^{-u/2}(1+2u) \leq 10.\end{aligned}$$

Let us define the increasing sequence $\ell_0, \ell_1, \ldots, \ell_{s-1}$ by the formula

$$\ell_0 = 1 \quad \text{ and } \quad h(\ell_i) = 10g(\ell_{i-1}), \; i = 1, 2, \ldots,$$

where s is the smallest number such that $\ell_{s-1} \geq m$ (if at some moment $10g(\ell_{j-1}) \geq n$ we set $\ell_j = m$ and $s = j+1$). First we show that such an s exists and satisfies $s \leq C\log\log(3m)$ for some absolute constant $C > 0$. We will use that $h(z) \leq H(z)$. By (5.17) if $\ell_{i-1} \leq m$ then $\ell_i \geq \sqrt{\ell_{i-1}m}/\log^{3/2}(e^2 m)$, which implies

$$\ell_1 \geq \sqrt{m}/\log^{3/2}(e^2m) \geq \frac{1}{6}m^{1/4} \quad (5.19)$$

and, by induction,

$$\ell_i \geq \left(\frac{m}{\log^3(e^2m)}\right)^{1-2^{-i}} \quad \text{for } i = 0, 1, 2, \ldots$$

In particular we have for some absolute constant $C_1 > 0$,

$$\ell_{s_1} \geq \frac{m}{2\log^3(e^2m)} \qquad \text{for some } s_1 \leq C_1 \log\log(3m).$$

By (5.18) we have $h(2z) \leq H(2z) \leq 10g(z)$ for $z \leq m$, so, if $\ell_{i-1} \leq m$ then $\ell_i \geq 2\ell_{i-1}$. It implies that for some $s \leq s_1 + C_2\log\log(3m) \leq C\log\log(3m)$ we indeed have $\ell_{s-1} \geq m$.

Finally we choose the sequence $(k_i)_{i\leq s+1}$ in the following way. Let $k_1 = k$ and $k_i := \min\{m, \lceil \ell_{s+1-i}\rceil\}$, $i = 2, \ldots, s+1$ (note that $k_2 = m$, $k_{s+1} = 1$). Using $h(\lceil z\rceil) \leq 2h(z)$ and construction of ℓ_i's, we obtain (5.16) for $i \geq 2$, while for $i = 1$ by definition of k' and since $k \leq k'$ we clearly have $h(k_1) = h(k) \leq 2g(m) = 2g(k_2)$. Since $(\ell_i)_i$ is increasing we also observe by (5.19) that $k_i \geq \frac{1}{6}m^{1/4}$ for $2 \leq i \leq s$. This completes the proof. $\square$

Proposition 5.8. *Suppose that $N \geq n$ and $k \leq \min\{n, k'\}$. Then for $t \geq 1$,*

$$\mathbb{P}\Big(\sup_{\substack{I\subset\{1,\ldots,N\}\\ |I|=m}} \sup_{x\in U_k} |P_I \Gamma x| \geq Ct\sqrt{\log\log(3m)}\sqrt{m}\log\Big(\frac{eN}{m}\Big)\Big)$$

$$\leq \exp\Big(- \frac{t\sqrt{\log\log(3m)}\sqrt{m}\log(eN/m)}{\sqrt{\log(em)}}\Big),$$

where C is a universal constant.

Proof. Let $k_1, \ldots, k_{s+1}$ be given by Lemma 5.7 and $\mathcal{N} \subset \frac{3}{2}U_k$ be as in Lemma 5.6. Notice that

$$\sup_{\substack{I\subset\{1,\ldots,N\}\\ |I|=m}} \sup_{x\in U_k} |P_I \Gamma x| \leq 2 \sup_{\substack{I\subset\{1,\ldots,N\}\\ |I|=m}} \sup_{x\in \mathcal{N}} |P_I \Gamma x|,$$

so we will estimate the latter quantity.

Let us fix $x \in \mathcal{N}$ and $1 \leq i \leq s$. We apply Theorem 4.3 to the vector $y = \pi_i(x)/(\sqrt{k_{i+1}}\|\pi_i(x)\|_\infty + |\pi_i(x)|)$ (observe that $|y| \leq 1$ and $\|y\|_\infty \leq 1/\sqrt{k_{i+1}}$ and on the other hand $1/\sqrt{k_{i+1}} \geq \frac{1}{\sqrt{m}}$) to get for $u > 0$,

$$\mathbb{P}\Big(\sup_{\substack{I\subset\{1,\ldots,N\}\\ |I|=m}} |P_I \Gamma \pi_i(x)| \geq C(\sqrt{k_{i+1}}\|\pi_i(x)\|_\infty + |\pi_i(x)|)\sqrt{m}\log\Big(\frac{eN}{m}\Big) + u\Big)$$

$$\leq \exp\Big(- 100 g(k_{i+1})\Big) \exp\Big(- \frac{\sqrt{k_{i+1}}u}{C\sqrt{\log(em)}}\Big),$$

where $g(x)$ is as in Lemma 5.7.

By the properties of the net $\mathcal{N}$ guaranteed by Lemma 5.6 and (5.16)

$$|\pi_i(\mathcal{N})| \exp\Big(- 100 g(k_{i+1})\Big) \leq 1.$$

Therefore for all $u > 0$ and $i = 1, \ldots, s$,

$$\mathbb{P}\Big(\sup_{x\in\mathcal{N}} \sup_{\substack{I\subset\{1,\ldots,N\}\\ |I|=m}} |P_I \Gamma \pi_i(x)| \geq C(\sqrt{k_{i+1}}\|\pi_i(x)\|_\infty + |\pi_i(x)|)\sqrt{m}\log\Big(\frac{eN}{m}\Big) + u\Big)$$

$$\leq \exp\Big(- \frac{\sqrt{k_{i+1}}u}{C\sqrt{\log(em)}}\Big).$$

We have for any $x \in \mathcal{N}$,

$$\sum_{i=1}^{s}(\sqrt{k_{i+1}}\|\pi_i(x)\|_\infty + |\pi_i(x)|) \le \sqrt{s}\Big(\Big(\sum_{i=1}^{s} k_{i+1}\|\pi_i(x)\|_\infty^2\Big)^{1/2} + \Big(\sum_{i=1}^{s}|\pi_i(x)|^2\Big)^{1/2}\Big)$$
$$\le \sqrt{s}\Big(2+\frac{3}{2}\Big) \le C\sqrt{\log\log(3m)}.$$

Therefore for any $u_1, \dots, u_s > 0$,

$$\mathbb{P}\Big(\sup_{x\in\mathcal{N}}\sup_{\substack{I\subset\{1,\dots,N\}\\|I|=m}} |P_I\Gamma x| \ge C\sqrt{\log\log(3m)}\sqrt{m}\log\Big(\frac{eN}{m}\Big) + \sum_{i=1}^{s} u_i\Big)$$
$$\le \sum_{i=1}^{s}\mathbb{P}\Big(\sup_{x\in\mathcal{N}}\sup_{\substack{I\subset\{1,\dots,N\}\\|I|=m}} |P_I\Gamma\pi_i(x)| \ge C(\sqrt{k_{i+1}}\|\pi_i(x)\|_\infty + |\pi_i(x)|)\sqrt{m}\log\Big(\frac{eN}{m}\Big) + u_i\Big)$$
$$\le \sum_{i=1}^{s}\exp\Big(-\frac{\sqrt{k_{i+1}}u_i}{C\sqrt{\log(em)}}\Big).$$

Hence it is enough to choose $u_s = Ct\sqrt{\log\log(3m)}\sqrt{m}\log(eN/m)$ and $u_i = \frac{1}{s}Ct\sqrt{\log\log(3m)}\sqrt{m}\log(eN/m)$ for $i = 1, \dots, s-1$ and to use the fact that $k_i \ge \frac{1}{6}m^{1/4}$ for $i = 2, \dots, s$. $\square$

5.3 Conclusion of the proof of Theorem 5.1

Proof. First notice that it is sufficient to consider the case $n \le N$. Indeed, if $n > N$ we may find independent isotropic n-dimensional log-concave random vectors $\tilde{X}_1, \dots, \tilde{X}_n$ such that $X_i = P_{\{1,\dots,N\}}\tilde{X}_i$ for $1 \le i \le n$. Let $\tilde{A}$ be the $n \times n$ matrix with rows $\tilde{X}_1, \dots, \tilde{X}_n$ and

$$\tilde{A}_{k,m} := \sup\{|P_I(\tilde{A})^* x| \colon I \subset \{1, \dots, n\},\ |I| = m,\ x \in U_k\}.$$

Then obviously $\tilde{A}_{k,m} \ge A_{k,m}$ and this allows us to immediately deduce the case $N \le n$ from the case $N = n$.

If $\sqrt{k}\log(en/k) + \sqrt{m}\log(eN/m) \ge k'\log(en/k')$ we may apply results of [5]. Recall that $\Gamma = A^*$. Let $s \ge \sqrt{k}\log(en/k) + \sqrt{m}\log(eN/m)$. Applying "in particular" part of Theorem 3.13 of [5] and Paoris' Theorem (inequality (3.2) together with union bound) to the columns of $m \times n$ matrix $P_I\Gamma$ and adjusting corresponding constants, we obtain that

$$\mathbb{P}\Big(\sup_{x\in U_k}|P_I\Gamma x| \ge Cs\Big) \le \exp(-2s)$$

for any $I \subset \{1, \ldots, N\}$ with $|I| = m$ (cf. Theorem 3.6 of [5]). Therefore

$$\mathbb{P}(A_{k,m} \geq Cs) \leq \sum_{|I|=m} \mathbb{P}\big(\sup_{x \in U_k} |P_I \Gamma x| \geq Cs \big) \leq \binom{N}{m} \exp(-2s).$$

By the definition of k' we get

$$\binom{N}{m} \leq \exp\big(m \log(eN/m) \big) \leq \exp\big(k' \log(en/k') \big),$$

hence for s as above

$$\mathbb{P}(A_{k,m} \geq Cs) \leq \exp\big(k' \log(en/k') \big) \exp(-2s) \leq \exp(-s)$$

and Theorem 5.1 follows in this case.

Finally assume that $n \leq N$ and that $\sqrt{k} \log(en/k) + \sqrt{m} \log(eN/m) \leq k' \log(en/k')$. For simplicity put $a_k = \sqrt{k} \log(en/k)$, $b_m = \sqrt{m} \log(eN/m)$, $d_m = \sqrt{\log\log(3m)}$. If $k \leq k'$ then Theorem 5.1 follows by Proposition 5.8 applied with $t_0 = t(1 + a_k/(d_m a_m))$. If $k \geq k'$ then we apply Proposition 5.8 (with the same t_0) and Propositions 5.2 with $t_1 = t(b_m d_m + a_k)/(a_k + b_m)$ to obtain Theorem 5.1 (note that $C\sqrt{\frac{m}{\log(em)}} \log(eN/m) \geq \log N \geq \log n$, so the factor n in the probability in Propositions 5.2 can be eliminated). $\square$

6 The Restricted Isometry Property

Fix integers n and $N \geq 1$ and let A be an $n \times N$ matrix. Consider the problem of reconstructing any vector $x \in \mathbb{R}^N$ with short support (sparse vectors) from the data $Ax \in \mathbb{R}^n$, with a fast algorithm.

Compressive Sensing provides a way of reconstructing the original signal x from its compression Ax with $n \ll N$ by the so-called ℓ_1-minimization method (see [15, 11, 13]).

Let

$$\delta_m = \delta_m(A) = \sup_{x \in U_m} \big| |Ax|^2 - \mathbb{E}|Ax|^2 \big|$$

be the Restricted Isometry Constant (RIC) of order m, introduced in [12]. Its important feature is that if δ_{2m} is appropriately small then every m-sparse vector x can be reconstructed from its compression Ax by the ℓ_1-minimization method. The goal is to check this property for certain models of matrices.

The articles [1, 2, 5, 6, 7] considered random matrices with independent *columns*, and investigated the RIP for various models of matrices, including the log-concave Ensemble build with independent isotropic log-concave columns. In this setting, the quantity $A_{n,m}$ played a central role.

In this section we consider $n \times N$ random matrices A with independent rows (X_i). For $T \subset \mathbb{R}^N$ the quantity $A_k(T)$ has been defined in (1.8) and $A_{k,m} = A_k(U_m)$ was estimated in the previous section.

We start with a general Lemma 6.1 which will be used to show that after a suitable discretization, one can reduce a concentration inequality to a deviation inequality; in particular, checking the RIP is reduced to estimating $A_{k,m}$. It is a slight strengthening of Lemma 1.3 from the introduction.

Lemma 6.1. *Let $X_1, \ldots, X_n$ be independent isotropic random vectors in $\mathbb{R}^N$. Let $T \subset S^{N-1}$ be a finite set. Let $0 < \theta < 1$ and $B \geq 1$. Then with probability at least $1 - |T| \exp(-3\theta^2 n/8B^2)$ one has*

$$\begin{aligned}
&\sup_{y \in T} \left| \frac{1}{n} \sum_{i=1}^n (|\langle X_i, y\rangle|^2 - \mathbb{E}|\langle X_i, y\rangle|^2) \right| \\
&\leq \theta + \frac{1}{n} \left(A_k(T)^2 + \sup_{y \in T} \mathbb{E} \sum_{i=1}^n |\langle X_i, y\rangle|^2 \mathbf{1}_{\{|\langle X_i, y\rangle| \geq B\}} \right) \\
&\leq \theta + \frac{1}{n} \left(A_k(T)^2 + \mathbb{E} A_k(T)^2 \right),
\end{aligned}$$

where $k \leq n$ is the largest integer satisfying $k \leq (A_k(T)/B)^2$.

Remark. Note that k in Lemma 6.1 is a random variable.

To prove Lemma 6.1 we need Bernstein's inequality (see e.g., Lemma 2.2.9 in [28]).

Proposition 6.2. *Let Z_i be independent centered random variables such that $|Z_i| \leq a$ for all $1 \leq i \leq n$. Then for all $\tau \geq 0$ one has*

$$\mathbb{P}\left(\frac{1}{n} \sum_{i=1}^n Z_i \geq \tau \right) \leq \exp\left(-\frac{\tau^2 n}{2(\sigma^2 + a\tau/3)} \right),$$

where

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n \mathrm{Var}(Z_i).$$

Proof of Lemma 6.1. For $y \in T$ let

$$S(y) = \left| \frac{1}{n} \sum_{i=1}^{n} \left(|\langle X_i, y \rangle|^2 - \mathbb{E}|\langle X_i, y \rangle|^2 \right) \right|$$

and observe that

$$\begin{aligned} S(y) \quad \le \quad & \Big| \frac{1}{n} \sum_{i=1}^{n} \left((|\langle X_i, y \rangle| \wedge B)^2 - \mathbb{E}\left(|\langle X_i, y \rangle| \wedge B \right)^2 \right) \Big| \\ + \quad & \frac{1}{n} \sum_{i=1}^{n} \left(|\langle X_i, y \rangle|^2 - B^2 \right) \mathbf{1}_{\{|\langle X_i, y \rangle| \ge B\}} \\ + \quad & \frac{1}{n} \mathbb{E} \sum_{i=1}^{n} \left(|\langle X_i, y \rangle|^2 - B^2 \right) \mathbf{1}_{\{|\langle X_i, y \rangle| \ge B\}}. \end{aligned}$$

We denote the three summands by $S_1(y)$, $S_2(y)$, $S_3(y)$, respectively, and we estimate each of them separately.

Estimate for $S_1(y)$: We will use Bernstein's inequality (Proposition 6.2). Given $y \in T$ let $Z_i(y) = (|\langle X_i, y \rangle| \wedge B)^2 - \mathbb{E}\left(|\langle X_i, y \rangle| \wedge B \right)^2$, for $i \le n$. Then $|Z_i(y)| \le B^2$, so $a = B^2$. By isotropicity of X for every $i \le n$ one has

$$\mathrm{Var}(Z_i(y)) \le \mathbb{E}\left(|\langle X_i, y \rangle| \wedge B \right)^4 \le \mathbb{E}\left(|\langle X_i, y \rangle|^2 B^2 \right) = B^2,$$

which implies $\sigma^2 \le B^2$. By Proposition 6.2,

$$\mathbb{P}\left(\frac{1}{n} \sum_{i=1}^{n} Z_i(y) \ge \theta \right) \le \exp\left(-\frac{\theta^2 n}{2(B^2 + B^2\theta/3)} \right) \le \exp\left(-\frac{3\theta^2 n}{8B^2} \right).$$

Then, by the union bound,

$$\mathbb{P}\left(\sup_{y \in T} S_1(y) \ge \theta \right) = \mathbb{P}\left(\sup_{y \in T} \frac{1}{n} \sum_{i=1}^{n} Z_i(y) \ge \theta \right) \le |T| \exp\left(-\frac{3\theta^2 n}{8B^2} \right).$$

Estimates for $S_2(y)$ and $S_3(y)$: For every $y \in T$ consider

$$E_B(y) = \{ i \le n \colon |\langle X_i, y \rangle| \ge B \},$$

and let

$$k' = \sup_{y \in T} |E_B(y)|.$$

Then, by the definition of $A_{k'}(T)$,

$$B^2 k' = B^2 \sup_{y\in T} |E_B(y)| \le \sup_{y\in T} \sum_{i\in E_B(y)} |\langle X_i, y\rangle|^2 \le A_{k'}^2(T).$$

This yields

$$k' \le \frac{A_{k'}^2(T)}{B^2},$$

and therefore $k' \le k$, where $k \le n$ is the biggest integer satisfying $k \le (A_k(T)/B)^2$.

Using the definition of $A_k(T)$ again we observe

$$\begin{aligned}\sup_{y\in T} S_2(y) &\le \frac{1}{n} \sup_{y\in T} \sum_{i=1}^n |\langle X_i, y\rangle|^2 \mathbf{1}_{\{|\langle X_i, y\rangle| \ge B\}} = \frac{1}{n} \sup_{y\in T} \sum_{i\in E_B(y)} |\langle X_i, y\rangle|^2 \\ &\le \frac{1}{n} \sup_{y\in T} \sup_{|E|\le k} \sum_{i\in E} |\langle X_i, y\rangle|^2 \le \frac{1}{n} A_k^2(T).\end{aligned}$$

Similarly,

$$\sup_{y\in T} S_3(y) \le \frac{1}{n} \sup_{y\in T} \mathbb{E} \sum_{i=1}^n |\langle X_i, y\rangle|^2 \mathbf{1}_{\{|\langle X_i, y\rangle| \ge B\}} \le \frac{1}{n} \mathbb{E} A_k^2(T).$$

Combining estimates for $S_1(y)$, $S_2(y)$, $S_3(y)$ we obtain the desired result. $\square$

By an approximation argument Lemma 6.1 has the following immediate consequence (cf., [7]).

Corollary 6.3. *Let $0 < \theta < 1$ and $B \ge 1$. Let n, N be positive integers and A be an $n \times N$ matrix, whose rows are independent isotropic random vectors X_i, for $i \le n$. Assume that $m \le N$ satisfies*

$$m \log \frac{11eN}{m} \le \frac{3\theta^2 n}{16B^2}.$$

Then with probability at least

$$1 - \exp\left(-\frac{3\theta^2 n}{16B^2}\right)$$

one has

$$\delta_m\left(\frac{A}{\sqrt{n}}\right) = \sup_{y\in U_m}\left|\frac{1}{n}\sum_{i=1}^n(|\langle X_i,y\rangle|^2 - \mathbb{E}|\langle X_i,y\rangle|^2)\right|$$
$$\le 2\theta + \frac{2}{n}\left(A_{k,m}^2 + \sup_{y\in U_m}\mathbb{E}\sum_{i=1}^n |\langle X_i,y\rangle|^2 \mathbf{1}_{\{|\langle X_i,y\rangle|\ge B\}}\right)$$
$$\le 2\theta + \frac{2}{n}\left(A_{k,m}^2 + \mathbb{E}A_{k,m}^2\right),$$

where $k\le n$ is the largest integer satisfying $k\le (A_{k,m}/B)^2$.

Remarks. 1. Note that as in Lemma 6.1, k in Corollary 6.3 is a random variable.
2. In all our applications we would like to have $A_{k,m}^2$ and $\mathbb{E}A_{k,m}^2$ of order θn. To obtain this, we choose the parameter B appropriately.
3. Note that $A_{k,m}$ is increasing in k, therefore we immediately have that if $m\le N$ satisfies

$$m\log\frac{11eN}{\theta m} \le \frac{3\theta^2 n}{16B^2} \quad \text{and} \quad \mathbb{E}A_{n,m}^2\le\theta n$$

then with probability at least

$$1-\exp\left(-\frac{3\theta^2 n}{16B^2}\right) - \mathbb{P}\left(A_{n,m}^2 > \theta n\right)$$

one has

$$\delta_m\left(\frac{A}{\sqrt{n}}\right)\le 6\theta. \tag{6.1}$$

Proof of Corollary 6.3. Let $\mathcal{N}$ be a 1/5-net in U_m of cardinality $\binom{N}{m}11^m \le (11eN/m)^m$ (we can construct $\mathcal{N}$ in such a way that for every $y\in U_m$ there exists $z_y\in\mathcal{N}$ with such that $z_y/|z_y|\in U_m$ and $|y-z_y|\le 1/5$). By the assumption on m,

$$m\log\frac{11eN}{m}\le\frac{3\theta^2 n}{16B^2},$$

and thus

$$|\mathcal{N}|\exp\left(-\frac{3\theta^2 n}{8B^2}\right)\le\exp\left(-\frac{3\theta^2 n}{16B^2}\right).$$

Using this and an obvious fact that $A_k(\mathcal{N}) \le A_k(U_m)$ for all k, we get by Lemma 6.1 that

$$\sup_{z\in\mathcal{N}} \left| \frac{1}{n}\sum_{i=1}^n (|\langle X_i, z\rangle|^2 - \mathbb{E}|\langle X_i, z\rangle|^2) \right| \le \theta + \frac{1}{n}\left(A_{k,m}^2 + \mathbb{E}A_{k,m}^2\right),$$

with probability larger than or equal to $1 - \exp\left(-\frac{3\theta^2 n}{16B^2}\right)$.

The proof is now finished by an approximation argument. Note that there exists a self-adjoint operator S acting on the Euclidean space $\mathbb{R}^N$ such that

$$\frac{1}{n}\sum_{i=1}^n (|\langle X_i, z\rangle|^2 - \mathbb{E}|\langle X_i, z\rangle|^2) = \langle Sz, z\rangle$$

for all $z \in \mathbb{R}^N$. Now pick $w \in U_m$ such that

$$|\langle Sw, w\rangle| = \sup_{y\in U_m} |\langle Sy, y\rangle|,$$

and let I with $|I| = m$ contain the support of w. Write $w = x + z$ where $x \in (1/5)B_2^N$ and $z \in \mathcal{N}$ and x and z are supported by I. Then

$$\begin{aligned} |\langle Sw, w\rangle| &= |\langle S(x+z), (x+z)\rangle| \\ &\le |\langle Sx, x\rangle| + |\langle Sx, z\rangle| + |\langle Sz, x\rangle| + |\langle Sz, z\rangle| \\ &\le (1/25)\sup_{x\in B_2^I} |\langle Sx, x\rangle| + (2/5)\sup_{x\in B_2^I} |\langle Sx, x\rangle| \sup_{z\in\mathcal{N}} |z| + \sup_{z\in\mathcal{N}} |\langle Sz, z\rangle|. \end{aligned}$$

Thus

$$\sup_{y\in U_m} |\langle Sy, y\rangle| \le (25/14)\sup_{z\in\mathcal{N}} |\langle Sz, z\rangle|.$$

completing the proof. □

The following theorem is a more general version of Theorem 1.4 stated in the introduction.

Theorem 6.4. *Let n, N be integers and $0 < \theta < 1$. Let A be an $n \times N$ matrix, whose rows are independent isotropic log-concave random vectors X_i, $i \le n$. There exists an absolute constant $c > 0$, such that if $m \le N$ satisfies*

$$m\ \log\log 3m \left(\log \frac{3\max\{N, n\}}{m}\right)^2 \le \frac{c\ \theta^2\ n}{\log(3/\theta)}$$

then

$$\delta_m(A/\sqrt{n}) \le \theta$$

with overwhelming probability.

Remark. In fact our proof gives that there is an absolute constant $c > 0$ such that if

$$b_m := m \log\log(3m) \left(\log \frac{3\max\{N,n\}}{m}\right)^2 \le c\theta n \tag{6.2}$$

and

$$m \log \frac{3N}{m} \log^2 \frac{n}{b_m} \le c\theta^2 n \tag{6.3}$$

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

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

In particular, denoting $\alpha_n = n/\log\log(3n)$ and $C_\theta = (\theta/\log(3/\theta))^2$ one can take

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

if $N \ge C_\theta n$ and

$$m \approx \frac{\theta\alpha_n}{\log^2(\log\log(3n)/\theta)}$$

if $N \le C_\theta n$.

Proof. Clearly it is enough to prove the estimate from the remark. So set b_m as in (6.2) and assume that $b_m \le c_1\theta n$ for small enough $c_1 > 0$. Choose $B = C_1 \log \frac{n}{b_m}$, where C_1 is a sufficiently large absolute constant.

Let k be as in Corollary 6.3, i.e. $k \le n$ is the biggest integer satisfying $k \le (A_{k,m}/B)^2$. As in Theorem 5.1 denote

$$\lambda_m = \frac{\sqrt{\log\log(3m)}\sqrt{m}}{\sqrt{\log(3m)}}\, \log(e\max\{N,n\}/m)$$

and

$$\lambda_{k,m} = \sqrt{\log\log(3m)}\sqrt{m}\log(e\max\{N,n\}/m) + \sqrt{k}\log(3n/k).$$

Applying (5.5) we obtain that there are absolute constants $C_0 > 0$ and $c_0 > 0$ such that

$$A_{k,m} \le C_0 \lambda_{k,m}, \tag{6.4}$$

with probability at least $1 - \exp(-c_0 \lambda_m)$. By Hölder's inequality and the log-concavity assumption we also obtain that there exists an absolute constant $C_2 > 0$ such that for every $y \in U_m$ one has

$$\mathbb{E} \sum_{i=1}^{n} |\langle X_i, y\rangle|^2 \mathbf{1}_{\{|\langle X_i, y\rangle| \ge B\}} \le \sum_{i=1}^{n} \sup_{x \in S^{n-1}} \|\langle X_i, x\rangle\|_4^2 \, \mathbb{P}\left(|\langle X_i, x\rangle| \ge B\right)^{1/2}$$

$$\le nC_2 \exp(-B/C_2) \le nC_2 (b_m/n)^{C_1/C_2} \le c_1 \theta n$$

for large enough C_1.

Below we show that for our choice of B, (6.4) implies

$$\sqrt{k} \log(3n/k) \le \sqrt{\log\log(3m)} \sqrt{m} \log(3 \max\{N, n\}/m) = \sqrt{b_m}, \tag{6.5}$$

which means $A_{k,m} \le 2C_0 \sqrt{b_m}$.

Note that if m satisfies (6.3) then we can apply Corollary 6.3 with our choice of B. It gives that there exists a positive constant C such that

$$\delta_m \left(\frac{A}{\sqrt{n}} \right) \le C\theta$$

with probability at least

$$1 - \exp\left(-\frac{3\theta^2 n}{16 B^2} \right) - 2 \exp(-c_0 \lambda_m),$$

which proves the desired result.

Now we prove that (6.4) implies (6.5). Assume it does not hold, i.e. assume that $k \log^2 \frac{3n}{k} > b_m$. Then, by the definition of k and (6.4) we observe that

$$k \le \frac{A_{k,m}^2}{B^2} \le \frac{C_0^2}{B^2} \lambda_{k,m}^2 \le 4C_0^2 \frac{k}{B^2} \log^2 \frac{3n}{k}.$$

This implies that $B \le 2C_0 \log \frac{3n}{k}$, which yields

$$k \le \frac{3n}{\exp(B/(2C_0))}.$$

Thus we obtain

$$\begin{aligned}b_m < k\log^2\frac{3n}{k} &\le \frac{3n}{\exp(B/(2C_0))}\,\frac{B^2}{4C_0^2}\\ &= \frac{3n}{\exp(C_1\log(n/b_m)/(2C_0))}\,\frac{C_1^2\log^2(n/b_m)}{4C_0^2},\end{aligned}$$

which is impossible for large enough C_1. This proves (6.5) and hence completes the proof. $\square$

7 Proofs of results from Section 3

7.1 Proof of Theorem 3.2

Theorem 3.2 is a strengthening of the first technical result in [18]. The proof given here is a modification of the argument from [18] and we include the details for the sake of completeness.

First we show the following proposition (an analogue of Proposition 10 from [18]).

Proposition 7.1. *There exist an absolute positive constant C_0 such that the following holds. Let X be an isotropic log-concave N-dimensional random vector, $A = \{X \in K\}$, where K is a convex set in $\mathbb{R}^N$ satisfying $0 < \mathbb{P}(A) \le 1/e$. Then for every $t \ge C_0$*

$$\sum_{i=1}^{N}\mathbb{P}(A\cap\{X(i)\ge t\}) \le C_0\mathbb{P}(A)\Big(t^{-2}\sigma_X^2(-\log(\mathbb{P}(A))) + Ne^{-t/C_0}\Big) \quad (7.1)$$

and for every $1 \le u \le \frac{t}{C_0}$

$$\big|\{i\le N\colon \mathbb{P}(A\cap\{X(i)\ge t\}) \ge e^{-u}\mathbb{P}(A)\}\big| \le \frac{C_0u^2}{t^2}\sigma_X^2(-\log(\mathbb{P}(A))). \quad (7.2)$$

Proof. Let Y be a random vector defined by

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

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

It is easy to see that Y is log-concave, but not necessarily isotropic. Without loss of generality we assume that $\mathbb{E}Y(1)^2 \geq \mathbb{E}Y(2)^2 \geq \ldots \geq \mathbb{E}Y(N)^2$ (otherwise we renumerate coordinates).

Given $\alpha > 0$ denote

$$m = m(\alpha) = \left|\{i\colon \mathbb{E}Y(i)^2 \geq \alpha\}\right|.$$

Then $\mathbb{E}Y(1)^2 \geq \ldots \geq \mathbb{E}Y(m)^2 \geq \alpha$. Using the Paley-Zygmund inequality and log-concavity of Y, we get

$$\begin{aligned}\mathbb{P}\Big(\sum_{i=1}^m Y(i)^2 \geq \frac{1}{2}\alpha m\Big) &\geq \mathbb{P}\Big(\sum_{i=1}^m Y(i)^2 \geq \frac{1}{2}\mathbb{E}\sum_{i=1}^m Y(i)^2\Big) \\ &\geq \frac{1}{4}\frac{(\mathbb{E}\sum_{i=1}^m Y(i)^2)^2}{\mathbb{E}(\sum_{i=1}^m Y(i)^2)^2} \geq \frac{1}{C_1}.\end{aligned}$$

Therefore

$$\mathbb{P}\Big(\sum_{i=1}^m X(i)^2 \geq \frac{1}{2}\alpha m\Big) \geq \mathbb{P}(A)\,\mathbb{P}\Big(\sum_{i=1}^m Y(i)^2 \geq \frac{1}{2}\alpha m\Big) \geq \frac{1}{C_1}\,\mathbb{P}(A).$$

Applying Theorem 3.1 (and Chebyshev's inequality, cf. (3.2)) to the m-dimensional vector $\bar{X} = (X_1, \ldots, X_m)$ we observe

$$\mathbb{P}\Big(\sum_{i=1}^m X(i)^2 \geq \frac{1}{2}\alpha m\Big) \leq \exp\Big(-\sigma_X^{-1}\Big(\frac{1}{C_3}\sqrt{m\alpha}\Big)\Big) \quad \text{for } \alpha \geq C_3.$$

Thus $\exp(-\sigma_X^{-1}(\frac{1}{C_3}\sqrt{m\alpha})) \geq \mathbb{P}(A)/C_1$ for $\alpha \geq C_3$, so, using the fact that $\sigma_X(tp) \leq 2t\sigma_X(p)$ for $t \geq 1$, we obtain that

$$m(\alpha) = \left|\{i\colon \mathbb{E}Y(i)^2 \geq \alpha\}\right| \leq \frac{C_4}{\alpha}\sigma_X^2(-\log(\mathbb{P}(A))) \quad \text{for } \alpha \geq C_3. \tag{7.3}$$

Note that for every i the random variable $Y(i)$ is log-concave, hence

$$\frac{\mathbb{P}(A \cap \{X(i) \geq t\})}{\mathbb{P}(A)} = \mathbb{P}(Y(i) \geq t) \leq \exp\Big(1 - \frac{t}{C_5(\mathbb{E}Y(i)^2)^{1/2}}\Big).$$

Thus, if $\mathbb{P}(Y(i) \geq t) \geq e^{-u}$ then $(\mathbb{E}Y(i)^2)^{1/2} \geq t/(C_5(u+1))$. Applying (7.3) with $\alpha = t^2/(C_5(u+1))^2$ we obtain that (7.2) holds with constant C_6 provided that $1 \leq u \leq t/C_7$.

Now assume that $t \geq \sqrt{C_3}$ and define a nonnegative integer k_0 by $2^{-k_0}t \geq \sqrt{C_3} > 2^{-k_0-1}t$. Let

$$I_0 = \{i \colon \mathbb{E}Y(i)^2 \geq t^2\}, \quad I_{k_0+1} = \{i \colon \mathbb{E}Y(i)^2 < 4^{-k_0}t^2\}$$

and

$$I_j = \{i \colon 4^{-j}t^2 \leq \mathbb{E}Y(i)^2 < 4^{1-j}t^2\} \quad j = 1, 2, \ldots, k_0.$$

Clearly $|I_{k_0+1}| \leq N$ and, by (7.3),

$$|I_j| \leq C_4 4^j t^{-2} \sigma_X^2(-\log \mathbb{P}(A)) \quad \text{for } j = 0, 1, \ldots, k_0.$$

Observe also that for $j > 0$ and $i \in I_j$ one has

$$\mathbb{P}(Y(i) \geq t) \leq \mathbb{P}\Big(\frac{Y(i)}{(\mathbb{E}Y(i)^2)^{1/2}} \geq 2^{j-1}\Big) \leq \exp\Big(1 - \frac{1}{C_8}2^j\Big).$$

Therefore

$$\begin{aligned}
\sum_{i=1}^{N} \mathbb{P}(Y(i) \geq t) &= \sum_{j=0}^{k_0+1} \sum_{i \in I_j} \mathbb{P}(Y(i) \geq t) \leq |I_0| + e\sum_{j=1}^{k_0+1} |I_j| \exp\Big(-\frac{2^j}{C_8}\Big) \\
&\leq C_4\bigg(t^{-2}\sigma_X^2(-\log \mathbb{P}(A))\bigg(1 + e\sum_{j=1}^{k_0} 4^j \exp\Big(-\frac{2^j}{C_8}\Big)\bigg) + eNe^{-t/(C\sqrt{C_3})}\bigg) \\
&\leq C_1\Big(t^{-2}\sigma_X^2(-\log \mathbb{P}(A)) + Ne^{-t/C_1}\Big).
\end{aligned}$$

By the definition of Y, this proves (7.1) with constant C_1 for $t \geq \sqrt{C_3}$. Taking $C_0 = \max\{C_1, \sqrt{C_3}, C_6, C_7\}$ completes the proof. □

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

Lemma 7.2. *Let $\ell_0 \geq \ell_1 \geq \ldots \geq \ell_s$ be a fixed sequence of positive integers and*

$$\mathcal{F} = \Big\{f \colon \{1, 2, \ldots, \ell_0\} \to \{0, 1, 2, \ldots, s\} \colon \ \forall_{1 \leq i \leq s} \ \ |\{r \colon f(r) \geq i\}| \leq \ell_i\Big\}.$$

Then

$$|\mathcal{F}| \leq \prod_{i=1}^{s} \Big(\frac{e\ell_{i-1}}{\ell_i}\Big)^{\ell_i}.$$

Proof of Theorem 3.2. Since $N_X \leq N$, the statement is trivial if $t\sqrt{N} \leq C\sigma_X(p)$. Without loss of generality we assume that $t\sqrt{N} \geq C\sigma_X(p)$ for large enough absolute constant $C > 0$.

Let C_0 be the constant from Proposition 7.1. Since X is isotropic and log-concave we may assume that $\mathbb{P}(X(j) \geq t) \leq e^{-t/C_0}$ for $t \geq C_0$ and $1 \leq j \leq N$ (we increase the actual value of C_0 if needed). Fix $p \geq 1$ and

$$t \geq C \log\left(\frac{Nt^2}{\sigma_X^2(p)}\right). \tag{7.4}$$

Then, for large enough C, $t \geq 4C_0$ and $t^2 N e^{-t/C_0} \leq \sigma_X^2(p)$.

Define a positive integer ℓ by

$$p \leq \ell < 2p \quad \text{and} \quad \ell = 2^k \text{ for some integer } k.$$

Then $\sigma_X(p) \leq \sigma_X(\ell) \leq \sigma_X(2p) \leq 4\sigma_X(p)$. Since $(\mathbb{E}(N_X(t))^p)^{1/p} \leq (\mathbb{E}(N_X(t))^\ell)^{1/\ell}$, it is enough to show that

$$\mathbb{E}(t^2 N_X(t))^\ell \leq (C_1\sigma_X(\ell))^{2\ell}.$$

Define sets

$$B_{i_1,\dots,i_s} = \{X(i_1) \geq t, \dots, X(i_s) \geq t\} \quad \text{and} \quad B_\emptyset = \Omega$$

and denote

$$m(\ell) := \mathbb{E}N_X(t)^\ell = \mathbb{E}\Big(\sum_{i=1}^N \mathbf{1}_{\{X(i)\geq t\}}\Big)^\ell = \sum_{i_1,\dots,i_\ell=1}^N \mathbb{P}(B_{i_1,\dots,i_\ell}).$$

It is enough to prove

$$m(\ell) \leq \left(\frac{C\sigma_X(\ell)}{t}\right)^{2\ell}. \tag{7.5}$$

We divide the sum in $m(\ell)$ into several parts. Let $j_1 \geq 2$ be an integer satisfying

$$2^{j_1-2} < \log\Big(\frac{Nt^2}{\sigma_X^2(\ell)}\Big) \leq 2^{j_1-1}.$$

Define sets

$$I_0 = \big\{(i_1,\dots,i_\ell) \in \{1,\dots,N\}^\ell \colon \mathbb{P}(B_{i_1,\dots,i_\ell}) > e^{-\ell}\big\},$$

$$I_j = \big\{(i_1,\dots,i_\ell) \in \{1,\dots,N\}^\ell \colon \mathbb{P}(B_{i_1,\dots,i_\ell}) \in (e^{-2^j\ell}, e^{-2^{j-1}\ell}]\big\}, \quad 0 < j < j_1,$$

and

$$I_{j_1} = \big\{(i_1,\dots,i_\ell) \in \{1,\dots,N\}^\ell \colon \mathbb{P}(B_{i_1,\dots,i_\ell}) \le e^{-2^{j_1-1}\ell}\big\}.$$

Note $\{1,\dots,N\}^\ell = \bigcup_{j=0}^{j_1} I_j$, hence $m(\ell) = \sum_{j=0}^{j_1} m_j(\ell)$, where

$$m_j(\ell) := \sum_{(i_1,\dots,i_\ell)\in I_j} \mathbb{P}(B_{i_1,\dots,i_\ell}) \quad \text{for } 0 \le j \le j_1.$$

First we estimate $m_{j_1}(\ell)$ and $m_0(\ell)$. Since $|I_{j_1}| \le N^\ell$

$$m_{j_1}(\ell) = \sum_{(i_1,\dots,i_\ell)\in I_{j_1}} \mathbb{P}(B_{i_1,\dots,i_\ell}) \le N^\ell e^{-2^{j_1-1}\ell} \le \Big(\frac{\sigma_X(\ell)}{t}\Big)^{2\ell}.$$

To estimate $m_0(\ell)$, given $I \subset \{1,\dots,N\}^\ell$ and $1 \le s \le \ell$, define

$$P_s I = \{(i_1,\dots,i_s) \colon (i_1,\dots,i_\ell) \in I \text{ for some } i_{s+1},\dots,i_\ell\}.$$

By Proposition 7.1 for $s = 1,\dots,\ell-1$ one has

$$\begin{aligned}\sum_{(i_1,\dots,i_{s+1})\in P_{s+1}I_0} \mathbb{P}(B_{i_1,\dots,i_{s+1}}) &\le \sum_{(i_1,\dots,i_s)\in P_sI_0} \sum_{i_{s+1}=1}^{N} \mathbb{P}(B_{i_1,\dots,i_s} \cap \{X(i_{s+1}) \ge t\}) \\ &\le C_0 \sum_{(i_1,\dots,i_s)\in P_sI_0} \mathbb{P}(B_{i_1,\dots,i_s})(t^{-2}\sigma_X^2(-\log \mathbb{P}(B_{i_1,\dots,i_s})) + Ne^{-t/C_0}).\end{aligned}$$

Note that for $(i_1,\dots,i_s) \in P_sI_0$ one has $\mathbb{P}(B_{i_1,\dots,i_s}) \ge e^{-\ell}$ and, by (7.4), $t^2Ne^{-t/C_0} \le \sigma_X^2(p) \le \sigma_X^2(\ell)$. Therefore

$$\sum_{(i_1,\dots,i_{s+1})\in P_{s+1}I_0} \mathbb{P}(B_{i_1,\dots,i_{s+1}}) \le C_4 t^{-2}\sigma_X^2(\ell) \sum_{(i_1,\dots,i_s)\in P_sI_0} \mathbb{P}(B_{i_1,\dots,i_s}).$$

By induction and since $\mathbb{P}(X(j) \ge t) \le e^{-t/C_0}$ we obtain

$$\begin{aligned}m_0(\ell) = \sum_{(i_1,\dots,i_\ell)\in I_0} \mathbb{P}(B_{i_1,\dots,i_\ell}) &\le (C_4t^{-2}\sigma_X^2(\ell))^{\ell-1} \sum_{i_1\in P_1I_0} \mathbb{P}(B_{i_1}) \\ &\le (C_4t^{-2}\sigma_X^2(\ell))^{\ell-1} Ne^{-t/C_0} \le \Big(\frac{C_5\sigma_X(\ell)}{t}\Big)^{2\ell}.\end{aligned}$$

Now we estimate $m_j(\ell)$ for $0 < j < j_1$. The upper bound is based on suitable estimates for $|I_j|$. Fix $0 < j < j_1$ and define a positive integer r_1 by

$$2^{r_1} < \frac{t}{C_0} \le 2^{r_1+1}.$$

For all $(i_1, \dots, i_\ell) \in I_j$ define a function $f_{i_1,\dots,i_\ell} \colon \{1, \dots, \ell\} \to \{j, j+1, \dots, r_1\}$ by

$$f_{i_1,\dots,i_\ell}(s) = \begin{cases} j & \text{if } \mathbb{P}(B_{i_1,\dots,i_s}) \ge \exp(-2^{j+1})\mathbb{P}(B_{i_1,\dots,i_{s-1}}), \\ r & \text{if } \exp(-2^{r+1}) \le \mathbb{P}(B_{i_1,\dots,i_s})/\mathbb{P}(B_{i_1,\dots,i_{s-1}}) < \exp(-2^r), \ j < r < r_1, \\ r_1 & \text{if } \mathbb{P}(B_{i_1,\dots,i_s}) < \exp(-2^{r_1})\mathbb{P}(B_{i_1,\dots,i_{s-1}}). \end{cases}$$

Note that for every $(i_1, \dots, i_\ell) \in I_j$ one has

$$1 = \mathbb{P}(B_\emptyset) \ge \mathbb{P}(B_{i_1}) \ge \mathbb{P}(B_{i_1,i_2}) \ge \ldots \ge \mathbb{P}(B_{i_1,\dots,i_l}) > \exp(-2^j\ell)$$

and $f_{i_1,\dots,i_\ell}(1) = r_1$, because $\mathbb{P}(X(i_1) \ge t) \le \exp(-t/C_0) < \exp(-2^{r_1})\mathbb{P}(B_\emptyset)$.

Denote

$$\mathcal{F}_j := \left\{f_{i_1,\dots,i_\ell} \colon \ (i_1, \dots, i_\ell) \in I_j\right\}.$$

Then for $f = f_{i_1,\dots,i_\ell} \in \mathcal{F}_j$ and every $r > j$ one has

$$\exp(-2^j\ell) < \mathbb{P}(B_{i_1,\dots,i_\ell}) = \prod_{s=1}^{\ell} \frac{\mathbb{P}(B_{i_1,\dots,i_s})}{\mathbb{P}(B_{i_1,\dots,i_{s-1}})} < \exp(-2^r \left|\{s \colon f(s) \ge r\}\right|).$$

Hence for every $r \ge j$ (the case $r = j$ is trivial) one has

$$|\{s \colon f(s) \ge r\}| \le 2^{j-r}\ell =: \ell_r. \tag{7.6}$$

Clearly, $\sum_{r=j+1}^{r_1} \ell_r \le \ell$ and $\ell_{r-1}/\ell_r = 2$, so by Lemma 7.2

$$|\mathcal{F}_j| \le \prod_{r=j+1}^{r_1} \left(\frac{e\ell_{r-1}}{\ell_r}\right)^{\ell_r} \le e^{2\ell}.$$

Now for every $f \in \mathcal{F}_j$ we estimate the cardinality of the set

$$I_j(f) := \{(i_1, \dots, i_\ell) \in I_j \colon \ f_{i_1,\dots,i_\ell} = f\}.$$

Fix f and for $r = j, j+1, \dots, r_1$ set

$$A_r := \{s \in \{1, \dots, \ell\} \colon f(s) = r\} \quad \text{and} \quad n_r := |A_r|\,.$$

Then $1 \in A_{r_1}$ and

$$n_j + n_{j+1} + \ldots + n_{r_1} = \ell.$$

Fixing $r < r_1$, $i_1, \ldots, i_{s-1}$, $s \in A_r$ (then $s \geq 2$ and $\mathbb{P}(B_{i_1,\ldots,i_{s-1}}) \leq \mathbb{P}(B_{i_1}) \leq \exp(-t/C_0) \leq 1/e$), applying (7.2) with $u = 2^{r+1} \leq t/C_0$, and using the definition of I_j, we observe that i_s may take at most

$$\begin{aligned}\frac{4C_0 2^{2r}}{t^2}\sigma_X^2(-\log \mathbb{P}(B_{i_1,\ldots,i_{s-1}})) &\leq \frac{4C_0 2^{2r}}{t^2}\sigma_X^2(2^j\ell) \leq \frac{16C_0 2^{2(r+j)}\sigma_X^2(\ell)}{t^2} \\ &\leq \frac{16C_0\sigma_X^2(\ell)}{t^2}\exp(2(r+j)) =: m_r\end{aligned}$$

values in order to satisfy $f_{i_1,\ldots,i_\ell} = f$. Thus

$$|I_j(f)| \leq N^{n_{r_1}} \prod_{r=j}^{r_1-1} m_r^{n_r} = N^{n_{r_1}} \Big(\frac{16C_0\sigma_X^2(\ell)}{t^2}\Big)^{\ell - n_{r_1}} \exp\Big(\sum_{r=j}^{r_1-1} 2(r+j)n_r\Big).$$

Note that (7.6) implies that $n_r \leq \ell_r = 2^{j-r}\ell$, so

$$\sum_{r=j}^{r_1-1} 2(r+j)n_r \leq 2^{j+2}\ell \sum_{r=j}^{\infty} r2^{-r} = 8(j+1)\ell \leq (50 + 2^{j-2})\ell.$$

By the definition of r_1 we also have

$$n_{r_1} \leq 2^{j-r_1}\ell \leq \frac{2C_0}{t}2^j\ell \leq \frac{2^{j-3}\ell}{\log(Nt^2/(4\sigma_X^2(\ell)))},$$

where in the last inequality we used (7.4) with large enough C. Thus we obtain that for every $f \in \mathcal{F}_j$

$$|I_j(f)| \leq \Big(\frac{C_6\sigma_X^2(\ell)}{t^2}\Big)^\ell \Big(\frac{Nt^2}{4\sigma_X^2(\ell)}\Big)^{n_{r_1}} \exp\big(2^{j-2}\ell\big) \leq \Big(\frac{C_6\sigma_X^2(\ell)}{t^2}\Big)^\ell \exp\Big(\frac{3}{8}2^j\ell\Big).$$

This implies that

$$|I_j| \leq |\mathcal{F}_j| \cdot \Big(\frac{C_6\sigma_X^2(\ell)}{t^2}\Big)^\ell \exp\Big(\frac{3}{8}2^j\ell\Big) \leq \Big(\frac{C_6\sigma_X^2(\ell)}{t^2}\Big)^\ell \exp\Big(\Big(2 + \frac{3}{8}2^j\Big)\ell\Big).$$

Hence

$$m_j(\ell) = \sum_{(i_1,\ldots,i_\ell)\in I_j} \mathbb{P}(B_{i_1,\ldots,i_\ell}) \leq |I_j| \exp(-2^{j-1}\ell) \leq \Big(\frac{C_7\sigma_X^2(\ell)}{t^2}\Big)^\ell \exp\big(-2^{j-3}\ell\big).$$

Combining estimates for $m_j(\ell)$'s we obtain

$$\begin{aligned} m(\ell) &= m_0(\ell) + m_{j_1}(\ell) + \sum_{j=1}^{j_1-1} m_j(\ell) \\ &\le \Big(\frac{\sigma_X(\ell)}{t}\Big)^{2\ell}\Big(C_5^\ell + 1 + \sum_{j=1}^{\infty} C_7^\ell \exp\big(-2^{j-3}\ell\big)\Big) \le \Big(\frac{C_8\sigma_X(\ell)}{t}\Big)^{2\ell}, \end{aligned}$$

which proves (7.5). □

7.2 Proof of Theorem 3.4

Fix $t \ge 1$ and let $m_0 = m_0(X,t)$.

Since $\sigma_{P_J X} \le \sigma_X$ for every $J \subset \{1,\dots,N\}$, Theorem 3.1 gives for any $J \subset \{1,\dots,N\}$ of cardinality m_0,

$$(\mathbb{E}|P_J X|^p)^{1/p} \le C_1(\sqrt{m_0} + \sigma_{P_J X}(p)) \le C_1(\sqrt{m} + \sigma_X(p)).$$

Using the Chebyshev inequality and $\sigma_X(up) \le 2u\sigma_X(p)$ we observe for such J,

$$\begin{aligned} \mathbb{P}\Big(|P_J X| \ge 36C_1 t\sqrt{m}\log\Big(\frac{eN}{m}\Big)\Big) &\le \exp\Big(-\sigma_X^{-1}\Big(6t\sqrt{m}\log\Big(\frac{eN}{m}\Big)\Big)\Big) \\ &\le \exp\Big(-3\sigma_X^{-1}\Big(t\sqrt{m}\log\Big(\frac{eN}{m}\Big)\Big)\Big). \end{aligned}$$

Thus, using the definition of m_0 and that $\sigma_X^{-1}(1) = 2$, we obtain

$$\begin{aligned} \mathbb{P}\Big(\sup_{|J|=m_0} |P_J X| \ge 36C_1 t\sqrt{m}\log\Big(\frac{eN}{m}\Big)\Big) & \\ &\le \binom{N}{m_0}\exp\Big(-3\sigma_X^{-1}\Big(t\sqrt{m}\log\Big(\frac{eN}{m}\Big)\Big)\Big) \\ &\le \frac{1}{2}\exp\Big(-\sigma_X^{-1}\Big(t\sqrt{m}\log\Big(\frac{eN}{m}\Big)\Big)\Big). \qquad (7.7) \end{aligned}$$

Now notice that

$$\sup_{|I|=m}|P_I X| = \Big(\sum_{i=1}^{m}|X^*(i)|^2\Big)^{1/2} \le \sup_{|J|=m_0}|P_J X| + \Big(\sum_{i=0}^{s-1} 2^i m_0 |X^*(2^i m_0)|^2\Big)^{1/2} \qquad (7.8)$$

with $s = \lceil \log_2(m/m_0) \rceil \leq 2\log(em/m_0)$. By Theorem 3.3 we get for $u \geq 0$,

$$\mathbb{P}\Big(|X^*(2^i m_0)|^2 \geq C_2 \log^2\Big(\frac{eN}{2^i m_0}\Big) + u^2\Big) \leq \exp\Big(-\sigma_X^{-1}\Big(\frac{1}{C_3} u 2^{i/2}\sqrt{m_0}\Big)\Big).$$

We have

$$C_2 \sum_{i=0}^{s-1} 2^i m_0 \log^2\Big(\frac{eN}{2^i m_0}\Big) \leq C_4 m \log^2\Big(\frac{eN}{m}\Big).$$

Therefore for any $u_0, \ldots, u_{s-1} \geq 0$,

$$\mathbb{P}\Big(\sum_{i=0}^{s-1} 2^i m_0 |X^*(2^i m_0)|^2 \geq C_4 m \log^2\Big(\frac{eN}{m}\Big) + \sum_{i=0}^{s-1} u_i^2\Big) \leq \sum_{i=0}^{s-1} \exp\Big(-\sigma_X^{-1}\Big(\frac{1}{C_3} u_i\Big)\Big).$$

Take $u_i^2 = \frac{2C_3^2}{s} t^2 m \log^2(\frac{eN}{m})$. Since $s \leq 2\log(em/m_0)$, we obtain

$$\begin{aligned}\mathbb{P}\Big(\sum_{i=0}^{s-1} 2^i m_0 |X^*(2^i m_0)|^2 &\geq (C_4 + 2C_3^2 t^2) m \log^2\Big(\frac{eN}{m}\Big)\Big)\\ &\leq s \exp\Big(-\sigma_X^{-1}\Big(\frac{\sqrt{2}}{\sqrt{s}} t\sqrt{m}\log\Big(\frac{eN}{m}\Big)\Big)\Big)\\ &\leq \frac{1}{2}\exp\Big(-\sigma_X^{-1}\Big(\frac{1}{\sqrt{\log(em/m_0)}} t\sqrt{m}\log\Big(\frac{eN}{m}\Big)\Big)\Big).\end{aligned}$$

This together with (7.7) and (7.8) completes the proof. □

References

[1] R. Adamczak, O. Guédon, A.E. Litvak, A. Pajor, and N. Tomczak-Jaegermann, *Condition number of a square matrix with i.i.d. columns drawn from a convex body*, Proc. Amer. Math. Soc., to appear.

[2] R. Adamczak, O. Guédon, A.E. Litvak, A. Pajor, and N. Tomczak-Jaegermann, *Smallest singular value of random matrices with independent columns*, C. R., Math., Acad. Sci. Paris, **346** (2008), 853–856.

[3] R. Adamczak, R. Latała, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, *Chevet type inequality and norms of submatrices*, preprint.

[4] 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, to appear.

[5] 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*, J. Amer. Math. Soc. **23** (2010), 535–561.

[6] 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.

[7] R. Adamczak, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, *Sharp bounds on the rate of convergence of empirical covariance matrix*, C.R. Math. Acad. Sci. Paris, **349** (2011), 195–200.

[8] R. Baraniuk, M. Davenport, R. DeVore, M. Wakin, *A Simple Proof of the Restricted Isometry Property for Random Matrices*, Constructive Approximation, **28** (2008), 253-263.

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

[10] J. Bourgain, *Random points in isotropic convex sets*, in: Convex geometric analysis, Berkeley, CA, 1996, Math. Sci. Res. Inst. Publ., Vol. 34, 53–58, Cambridge Univ. Press, Cambridge (1999).

[11] E.J. Candes, J. Romberg and T. Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure App. Math. **59**, (2006), 1207–1223.

[12] E.J. Candés and T. Tao, *Decoding by linear programming*, IEEE Trans. Inform. Theory **51** (2005), 4203–4215.

[13] E. J. Candes and T. Tao, *Near-optimal signal recovery from random projections: universal encoding strategies*, IEEE Trans. Inform. Theory, **52** (2006), 5406–5425.

[14] Ju.S. Davidovic, B.I. Korenbljum and B.I. Hacet, *A certain property of logarithmically concave functions*, Soviet Math. Dokl. **10** (1969), 447–480; translation from Dokl. Akad. Nauk SSSR **185** (1969), 1215–1218.

[15] D.L. Donoho, *Neighborly Polytopes and Sparse solutions of underdetermined linear equations*, Department of Statistics, Stanford University, 2005.

[16] E.D. Gluskin and S. Kwapień, *Tail and moment estimates for sums of independent random variables with logarithmically concave tails*, Studia Math. **114** (1995) 303–309.

[17] R. Kannan, L. Lovász and M. Simonovits, *Random walks and $O^*(n^5)$ volume algorithm for convex bodies*, Random structures and algorithms, **2** (1997), 1–50.

[18] R. Latała, *Order statistics and concentration of l_r norms for log-concave vectors*, J. Funct. Anal. **261** (2011), 681–696.

[19] R. Latała, *Weak and strong moments of random vectors*, preprint, http://arxiv.org/abs/1012.2703.

[20] R. Latała and J.O. Wojtaszczyk, *On the infimum convolution inequality*, Studia Math. **189** (2008), 147–187.

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

[22] S. Mendelson, *Empirical Processes with a bounded ψ_1 diameter*, Geom. Funct. Anal., **20** (2010), 988–1027.

[23] S. Mendelson, A. Pajor and N. Tomczak-Jaegermann, *Reconstruction and subgaussian operators in asymptotic geometric analysis*, Geom. Funct. Anal. **17** (2007), 1248–1282.

[24] G. Paouris, *Concentration of mass on convex bodies*, Geom. Funct. Anal. **16** (2006), 1021–1049.

[25] A. Prékopa, *Logarithmic concave measures with application to stochastic programming*, Acta Sci. Math. **32** (1971), 301–316.

[26] M. Rudelson, *Random vectors in the isotropic position*, J. Funct. Anal. **164** (1999), 60–72.

[27] N. Srivastava and R. Vershynin, *Covariance Estimation for distributions with $2+\varepsilon$ moments*, preprint.

[28] A.W. van der Vaart and J.A. Wellner, *Weak convergence and empirical processes. With applications to statistics,* Springer Series in Statistics, Springer-Verlag, New York, 1996.

[29] R. Vershynin, *How close is the sample covariance matrix to the actual covariance matrix?*, preprint.

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