Available online at www.sciencedirect.com
ScienceDirect
Journal of Functional Analysis 261 (2011) 681–696
JOURNAL OF Functional Analysis
www.elsevier.com/locate/jfa

Order statistics and concentration of l_r norms for log-concave vectors ☆

Rafał Latała [a,b,*]

[a] *Institute of Mathematics, University of Warsaw, Banacha 2, 02-097 Warszawa, Poland*
[b] *Institute of Mathematics, Polish Academy of Sciences, ul. Śniadeckich 8, 00-956 Warszawa, Poland*

Received 30 November 2010; accepted 17 February 2011
Available online 2 March 2011
Communicated by K. Ball

Abstract

We establish upper bounds for tails of order statistics of isotropic log-concave vectors and apply them to derive a concentration of l_r norms of such vectors.

Keywords: Log-concave measures; Order statistics; Concentration of volume

1. Introduction and notation

An n-dimensional random vector is called log-concave if it has a log-concave distribution, i.e. for any compact nonempty sets $A, B \subset \mathbb{R}^n$ and $\lambda \in (0, 1)$,

$$\mathbb{P}\big(X \in \lambda A + (1-\lambda)B\big) \geqslant \mathbb{P}(X \in A)^{\lambda}\mathbb{P}(X \in B)^{1-\lambda},$$

where $\lambda A + (1-\lambda)B = \{\lambda x + (1-\lambda)y\colon\ x \in A,\ y \in B\}$. By the result of Borell [3] a vector X with full dimensional support is log-concave if and only if it has a density of the form e^{-f}, where $f : \mathbb{R}^n \to (-\infty, \infty]$ is a convex function. Log-concave vectors are frequently studied in convex geometry, since by the Brunn–Minkowski inequality uniform distributions on convex sets as well as their lower dimensional marginals are log-concave.

☆ Research partially supported by MNiSW Grant no. N N201 397437 and the Foundation for Polish Science.
* Address for correspondence: Institute of Mathematics, University of Warsaw, Banacha 2, 02-097 Warszawa, Poland.
E-mail address: rlatala@mimuw.edu.pl.

0022-1236/$ – see front matter
doi:10.1016/j.jfa.2011.02.013

A random vector $X = (X_1, \dots, X_n)$ is isotropic if $\mathbb{E}X_i = 0$ and $\mathrm{Cov}(X_i, X_j) = \delta_{i,j}$ for all $i, j \leqslant n$. Equivalently, an n-dimensional random vector with mean zero is isotropic if $\mathbb{E}\langle t, X\rangle^2 = |t|^2$ for any $t \in \mathbb{R}^n$. For any nondegenerate log-concave vector X there exists an affine transformation T such that TX is isotropic.

In recent years there were derived numerous important properties of log-concave vectors. One of such results is the Paouris concentration of mass [10] that states that for any isotropic log-concave vector X in $\mathbb{R}^n$,

$$\mathbb{P}\big(|X| \geqslant Ct\sqrt{n}\big) \leqslant \exp(-t\sqrt{n}) \quad \text{for } t \geqslant 1. \tag{1}$$

One of purposes of this paper is the extension of the Paouris result to l_r norms, that is deriving upper bounds for $\mathbb{P}(\|X\|_r \geqslant t)$, where $\|x\|_r = (\sum_{i=1}^n |x_i|^r)^{1/r}$. For $r \in [1, 2)$ this is an easy consequence of (1) and Hölder's inequality, however the case $r > 2$ requires in our opinion new ideas. We show that

$$\mathbb{P}\big(\|X\|_r \geqslant C(r)tn^{1/r}\big) \leqslant \exp\big(-tn^{1/r}\big) \quad \text{for } t \geqslant 1,\ r > 2,$$

where $C(r)$ is a constant depending only on r – see Theorem 8. Our method is based on suitable tail estimates for order statistics of X.

For an n-dimensional random vector X by $X_1^* \geqslant X_2^* \geqslant \dots \geqslant X_n^*$ we denote the nonincreasing rearrangement of $|X_1|, \dots, |X_n|$ (in particular $X_1^* = \max\{|X_1|, \dots, |X_n|\}$ and $X_n^* = \min\{|X_1|, \dots, |X_n|\}$). Random variables X_k^*, $1 \leqslant k \leqslant n$, are called order statistics of X.

By (1) we immediately get for isotropic log-concave vectors X,

$$\mathbb{P}\big(X_k^* \geqslant t\big) \leqslant \exp\left(-\frac{1}{C}\sqrt{k}t\right)$$

for $t \geqslant C\sqrt{n/k}$. The main result of the paper is Theorem 3 which states that the above inequality holds for $t \geqslant C\log(en/k)$ – as shows the example of exponential distribution this range of t is for $k \leqslant n/2$ optimal up to a universal constant.

Tail estimates for order statistics can be also applied to provide optimal estimates for $\sup_{\#I=m} |P_I X|$, where the supremum is taken over all subsets of $\{1, \dots, n\}$ of cardinality $m \in [1, n]$ and P_I denotes the coordinatewise projection. The details will be presented in the forthcoming paper [1].

The organization of the article is as follows. In Section 2 we discuss upper bounds for tails of order statistics and their connections with exponential concentration and Paouris' result. Section 3 is devoted to the derivation of tail estimates of l_r norms for log-concave vectors. Finally Section 4 contains a proof of Theorem 4, which is a crucial tool used to derive our main result.

Throughout the article by $C, C_1, \dots$ we denote universal constants. Values of a constant C may differ at each occurrence. For $x \in \mathbb{R}^n$ we put $|x| = \|x\|_2 = (\sum_{i=1}^n x_i^2)^{1/2}$.

2. Tail estimates for order statistics

If the coordinates of X are independent symmetric exponential random variables with variance one then it is not hard to see that $\mathrm{Med}(X_k^*) \geqslant \frac{1}{C}\log(en/k)$ for any $1 \leqslant k \leqslant n/2$. So we may obtain a reasonable bound for $\mathbb{P}(X_k^* \geqslant t)$, $k \leqslant n/2$ in the case of isotropic log-concave vectors only for $t \geqslant \frac{1}{C}\log(en/k)$. Using the idea that exponential random vectors are extremal in the class of unconditional log-concave vectors (i.e. such vectors that $(\eta_1 X_1, \dots, \eta_n X_n)$ has the same distribution as X for any choice of signs $\eta_i \in \{-1, 1\}$) one may easily derive the following fact.

Proposition 1. *If X is a log-concave and unconditional n-dimensional isotropic random vector then*

$$\mathbb{P}\big(X_k^* \geqslant t\big) \leqslant \exp\left(-\frac{1}{C}kt\right) \quad \text{for } t \geqslant C\log\left(\frac{en}{k}\right).$$

Proof. The result of Bobkov and Nazarov [2] implies that for any $i_1 < i_2 < \dots < i_k$ and $t > 0$,

$$\mathbb{P}\big(|X_{i_1}| \geqslant t, \dots, |X_{i_k}| \geqslant t\big) = 2^k \mathbb{P}(X_{i_1} \geqslant t, \dots, X_{i_k} \geqslant t) \leqslant 2^k \exp\left(-\frac{1}{C}kt\right).$$

Hence

$$\begin{aligned}\mathbb{P}\big(X_k^* \geqslant t\big) &\leqslant \sum_{1\leqslant i_1<\dots<i_k\leqslant n} \mathbb{P}\big(|X_{i_1}| \geqslant t, \dots, |X_{i_k}| \geqslant t\big) \leqslant \binom{n}{k} 2^k \exp\left(-\frac{1}{C}kt\right)\\ &\leqslant \left(\frac{2en}{k}\right)^k \exp\left(-\frac{1}{C}kt\right) \leqslant \exp\left(-\frac{1}{2C}kt\right)\end{aligned}$$

if $t \geqslant C' \log(en/k)$. $\square$

However for a general isotropic log-concave vector without unconditionality assumption we may bound $\mathbb{P}(X_{i_1} \geqslant t, \dots, X_{i_k} \geqslant t)$ only by $\exp(-\sqrt{k}t/C)$ for $t \geqslant C$. This suggests that we should rather expect bound $\exp(-\sqrt{k}t/C)$ than $\exp(-kt/C)$. If we try to apply the union bound as in the proof of Proposition 1 it will work only for $t \geqslant C\sqrt{k}\log(en/k)$.

Another approach may be based on the exponential concentration. We say that a vector X satisfies exponential concentration inequality with a constant α if for any Borel set A,

$$\mathbb{P}\big(X \in A + \alpha t B_2^n\big) \geqslant 1 - \exp(-t) \quad \text{if } \mathbb{P}(X \in A) \geqslant \frac{1}{2} \text{ and } t > 0.$$

Proposition 2. *If the coordinates of an n-dimensional vector X have mean zero and variance one and X satisfies exponential concentration inequality with a constant $\alpha \geqslant 1$ then*

$$\mathbb{P}\big(X_k^* \geqslant t\big) \leqslant \exp\left(-\frac{1}{3\alpha}\sqrt{k}t\right) \quad \text{for } t \geqslant 8\alpha\log\left(\frac{en}{k}\right).$$

Proof. Since $\mathrm{Var}(X_i) = 1$ we have $\mathbb{P}(|X_i| \leqslant 2) \geqslant 1/2$ so $\mathbb{P}(|X_i| \geqslant 2 + t) \leqslant \exp(-t/\alpha)$ for $t > 0$. Let μ be the distribution of X. Then the set

$$A(t) = \left\{x \in \mathbb{R}^n \colon \#\big\{i \colon |x_i| \geqslant t\big\} < \frac{k}{2}\right\}$$

has measure μ at least $1/2$ for $t \geqslant 4\alpha\log(en/k)$ – indeed we have for such t

$$1-\mu\big(A(t)\big)=\mathbb{P}\left(\sum_{i=1}^{n}\mathbb{1}_{\{|X_i|\geqslant t\}}\geqslant\frac{k}{2}\right)\leqslant\frac{2}{k}\mathbb{E}\left(\sum_{i=1}^{n}\mathbb{1}_{\{|X_i|\geqslant t\}}\right)$$
$$\leqslant\frac{2n}{k}\exp\left(-\frac{t}{2\alpha}\right)\leqslant\frac{2n}{k}\left(\frac{en}{k}\right)^{-2}\leqslant\frac{1}{2}.$$

Let $A=A(4\alpha\log(en/k))$. If $z=x+y\in A+\sqrt{k}sB_2^n$ then less than $k/2$ of $|x_i|$'s are bigger than $4\alpha\log(en/k)$ and less than $k/2$ of $|y_i|$'s are bigger than $\sqrt{2}s$, so

$$\mathbb{P}\left(X_k^*\geqslant 4\alpha\log\left(\frac{en}{k}\right)+\sqrt{2}s\right)\leqslant 1-\mu\big(A+\sqrt{k}sB_2^n\big)\leqslant\exp\left(-\frac{1}{\alpha}\sqrt{k}s\right). \qquad \square$$

For log-concave vectors it is known that exponential inequality is equivalent to several other functional inequalities such as Cheeger's and spectral gap – see [9] for a detailed discussion and recent results. The strong conjecture due to Kannan, Lovász and Simonovits [6] states that every isotropic log-concave vector satisfies Cheeger's (and therefore also exponential) inequality with a uniform constant. The conjecture however is wide open – a recent result of Klartag [7] shows that in the unconditional case KLS conjecture holds up to $\log n$ constant (see also [5] for examples of nonproduct distributions that satisfy spectral gap inequality with uniform constants). Best known upper bound for Cheeger's constant for general isotropic log-concave measure is n^{α} for some $\alpha\in(1/4,1/2)$ (see [9,4]).

The main result of this paper states that despite the approach via the union bound or exponential concentration fails the natural estimate for order statistics is valid. Namely we have

Theorem 3. *Let X be an n-dimensional log-concave isotropic vector. Then*

$$\mathbb{P}\big(X_k^*\geqslant t\big)\leqslant\exp\left(-\frac{1}{C}\sqrt{k}t\right)\quad \text{for } t\geqslant C\log\left(\frac{en}{k}\right).$$

Our approach is based on the suitable estimate of moments of the process $N_X(t)$, where

$$N_X(t):=\sum_{i=1}^{n}\mathbb{1}_{\{X_i\geqslant t\}},\quad t\geqslant 0.$$

Theorem 4. *For any isotropic log-concave vector X and $p\geqslant 1$ we have*

$$\mathbb{E}\big(t^2N_X(t)\big)^p\leqslant(Cp)^{2p}\quad \text{for } t\geqslant C\log\left(\frac{nt^2}{p^2}\right).$$

We postpone a long and bit technical proof till the last section of the paper. Let us only mention at this point that it is based on two ideas. One is the Paouris large deviation inequality (1) and another is an observation that if we restrict a log-concave distribution to a convex set it is still log-concave.

Proof of Theorem 3. Observe that $X_k^*\geqslant t$ implies that $N_X(t)\geqslant k/2$ or $N_{-X}(t)\geqslant k/2$ and vector $-X$ is also isotropic and log-concave. So by Theorem 4 and Chebyshev's inequality we get

$$\mathbb{P}\big(X_k^* \geqslant t\big) \leqslant \left(\frac{2}{k}\right)^p \big(\mathbb{E}N_X(t)^p + \mathbb{E}N_{-X}(t)^p\big) \leqslant 2\left(\frac{Cp}{t\sqrt{k}}\right)^{2p}$$

provided that $t \geqslant C\log(nt^2/p^2)$. So it is enough to take $p = \frac{1}{eC}t\sqrt{k}$ and notice that the restriction on t follows by the assumption that $t \geqslant C\log(en/k)$. □

As we already noticed one of the main tools in the proof of Theorem 4 is the Paouris concentration of mass. One may however also do the opposite and derive large deviations for the Euclidean norm of X from our estimate of moments of $N_X(t)$ and the observation that the distribution of UX is again log-concave and isotropic for any rotation U. More precisely the following statement holds.

Proposition 5. *Suppose that X is a random vector in $\mathbb{R}^n$ such that for some constants $A_1, A_2 < \infty$ and any $U \in O(n)$,*

$$\mathbb{E}\big(t^2 N_{UX}(t)\big)^l \leqslant (A_1 l)^{2l} \quad \textit{for } t \geqslant A_2,\ l \geqslant \sqrt{n}.$$

Then

$$\mathbb{P}\big(|X| \geqslant t\sqrt{n}\big) \leqslant \exp\left(-\frac{1}{CA_1}t\sqrt{n}\right) \quad \textit{for } t \geqslant \max\{CA_1, A_2\}.$$

Proof. Let us fix $t \geqslant A_2$. Hölder's inequality gives that for any $U_1, \ldots, U_n \in O(n)$,

$$\mathbb{E}\prod_{i=1}^{l} N_{U_iX}(t) \leqslant \left(\prod_{i=1}^{l}\mathbb{E}N_{U_iX}(t)^l\right)^{1/l} \leqslant \left(\frac{A_1 l}{t}\right)^{2l} \quad \text{for } l \geqslant \sqrt{n}.$$

Now let $U_1, \ldots, U_l$ be independent random rotations in $O(n)$ (distributed according to the Haar measure) then for $l \geqslant \sqrt{n}$,

$$\begin{aligned}\left(\frac{A_1 l}{t}\right)^{2l} &\geqslant \mathbb{E}_X\mathbb{E}_U\prod_{i=1}^{l} N_{U_iX}(t) = \mathbb{E}_X\big(\mathbb{E}_{U_1}N_{U_1X}(t)\big)^l = \mathbb{E}_X\big(n\mathbb{P}_Y\big(\langle X, Y\rangle \geqslant t\big)\big)^l \\ &= n^l\mathbb{E}_X\big(\mathbb{P}_Y\big(|X|Y_1 \geqslant t\big)\big)^l,\end{aligned}$$

where Y is a random vector uniformly distributed on S^{n-1}. Since Y_1 is symmetric, $\mathbb{E}Y_1^2 = 1/n$ and $\mathbb{E}Y_1^4 \leqslant C/n^2$ we get by the Paley–Zygmund inequality that $\mathbb{P}(Y_1^2 \geqslant \frac{1}{4n}) \geqslant 1/C_1$ which gives

$$\mathbb{P}\big(|X| \geqslant 2t\sqrt{n}\big) \leqslant \mathbb{E}_X\big(C_1\mathbb{P}_Y\big(|X|Y_1 \geqslant t\big)\big)^l \leqslant \left(\frac{C_1A_1^2l^2}{t^2 n}\right)^l.$$

To conclude the proof it is enough to take $l = \lceil \frac{1}{\sqrt{eC_1}A_1}\sqrt{n}t\rceil$. □

3. Concentration of l_r norms

The aim of this section is to derive Paouris-type estimates for concentration of $\|X\|_r = (\sum_{i=1}^n |X_i|^r)^{1/r}$. We start with presenting two simple examples.

Example 1. Let the coordinates of X be independent symmetric exponential r.v.'s with variance one. Then

$$\left(\mathbb{E}\|X\|_r^r\right)^{1/r} = \left(n\mathbb{E}|X_1|^r\right)^{1/r} \geqslant \frac{1}{C} r n^{1/r} \quad \text{for } r \in [1,\infty),$$
$$\mathbb{E}\|X\|_\infty \geqslant \frac{1}{C} \log n$$

and

$$\left(\mathbb{E}\|X\|_r^p\right)^{1/p} \geqslant \left(\mathbb{E}|X_1|^p\right)^{1/p} \geqslant \frac{p}{C} \quad \text{for } p \geqslant 2,\ r \geqslant 1.$$

It is also known that in the independent exponential case weak and strong moments are comparable [8], hence for $r \geqslant 2$,

$$\begin{aligned}\left(\mathbb{E}\|X\|_r^r\right)^{1/r} &= \left(\mathbb{E} \sup_{\|a\|_{r'} \leqslant 1} \left|\sum_i a_i X_i\right|^r\right)^{1/r} \\ &\leqslant \left(\mathbb{E}\|X\|_r^2\right)^{1/2} + C \sup_{\|a\|_{r'} \leqslant 1} \left(\mathbb{E}\left|\sum_i a_i X_i\right|^r\right)^{1/r} \\ &\leqslant \left(\mathbb{E}\|X\|_r^2\right)^{1/2} + Cr \sup_{\|a\|_{r'} \leqslant 1} \left(\mathbb{E}\left|\sum_i a_i X_i\right|^2\right)^{1/2} \leqslant \left(\mathbb{E}\|X\|_r^2\right)^{1/r} + Cr.\end{aligned}$$

Therefore we get

$$\left(\mathbb{E}\|X\|_r^p\right)^{1/p} \geqslant \left(\mathbb{E}\|X\|_r^2\right)^{1/2} \geqslant \frac{1}{C} r n^{1/r} \quad \text{for } p \geqslant 2 \text{ and } n \geqslant C^r.$$

Example 2. For $1 \leqslant r \leqslant 2$ let X be an isotropic random vector such that $Y = (X_1 + \cdots + X_n)/\sqrt{n}$ has the exponential distribution with variance one. Then by Hölder's inequality $\|X\|_r \geqslant n^{1/r-1/2} Y$ and

$$\left(\mathbb{E}\|X\|_r^p\right)^{1/p} \geqslant n^{1/r-1/2}\|Y\|_p \geqslant \frac{1}{C} n^{1/r-1/2} p \quad \text{for } p \geqslant 2,\ 1 \leqslant r \leqslant 2.$$

The examples above show that the best we can hope is

$$\left(\mathbb{E}\|X\|_r^p\right)^{1/p} \leqslant C\left(n^{1/r} + n^{1/r-1/2} p\right) \quad \text{for } p \geqslant 2,\ 1 \leqslant r \leqslant 2, \tag{2}$$
$$\left(\mathbb{E}\|X\|_r^p\right)^{1/p} \leqslant C\left(r n^{1/r} + p\right) \quad \text{for } p \geqslant 2,\ r \in [2,\infty) \tag{3}$$

and

$$\big(\mathbb{E}\|X\|_\infty^p\big)^{1/p} \leqslant C(\log n + p) \quad \text{for } p \geqslant 2. \tag{4}$$

Or in terms of tails,

$$\mathbb{P}\big(\|X\|_r \geqslant t\big) \leqslant \exp\left(-\frac{1}{C} t n^{1/2-1/r}\right) \quad \text{for } t \geqslant C n^{1/r},\ r \in [1,2], \tag{5}$$

$$\mathbb{P}\big(\|X\|_r \geqslant t\big) \leqslant \exp\left(-\frac{1}{C} t\right) \quad \text{for } t \geqslant C r n^{1/r},\ r \in [2,\infty) \tag{6}$$

and

$$\mathbb{P}\big(\|X\|_\infty \geqslant t\big) \leqslant \exp\left(-\frac{1}{C} t\right) \quad \text{for } t \geqslant C \log n. \tag{7}$$

Case $r \in [1,2]$ is a simple consequence of the Paouris theorem.

Proposition 6. *Estimates* (2) *and* (5) *hold for all isotropic log-concave vectors* X.

Proof. We have $\|X\|_r \leqslant n^{1/r-1/2}\|X\|_2$ by Hölder's inequality, hence (2) (and therefore also (5)) immediately follows by the Paouris result. □

Case $r = \infty$ is also very simple.

Proposition 7. *Estimates* (4) *and* (7) *hold for all isotropic log-concave vectors* X.

Proof. We have

$$\mathbb{P}\big(\|X\|_\infty \geqslant t\big) \leqslant \sum_{i=1}^n \mathbb{P}\big(|X_i| \geqslant t\big) \leqslant n \exp(-t/C). \qquad \square$$

What is left is the case $2 < r < \infty$ – we would like to obtain (6) and (3). We almost get it – except that constants explode when r approaches 2.

Theorem 8. *For any* $\delta > 0$ *there exist constants* $C_1(\delta), C_2(\delta) \leqslant C(1+\delta^{-1/2})$ *such that for any* $r \geqslant 2+\delta$,

$$\mathbb{P}\big(\|X\|_r \geqslant t\big) \leqslant \exp\left(-\frac{1}{C_1(\delta)} t\right) \quad \textit{for } t \geqslant C_1(\delta) r n^{1/r}$$

and

$$\big(\mathbb{E}\|X\|_r^p\big)^{1/p} \leqslant C_2(\delta)\big(r n^{1/r} + p\big) \quad \textit{for } p \geqslant 2.$$

The proof of Theorem 8 is based on the following slightly more precise estimate.

Proposition 9. *For $r > 2$ we have*

$$\mathbb{P}\big(\|X\|_r \geqslant t\big) \leqslant \exp\biggl(-\frac{1}{C}\biggl(\frac{r-2}{r}\biggr)^{1/r} t\biggr) \quad \textit{for } t \geqslant C\biggl(rn^{1/r} + \biggl(\frac{r}{r-2}\biggr)^{1/r} \log n\biggr)$$

or in terms of moments

$$\big(\mathbb{E}\|X\|_r^p\big)^{1/p} \leqslant C\biggl(rn^{1/r} + \biggl(\frac{r}{r-2}\biggr)^{1/r} (\log n + p)\biggr) \quad \textit{for } p \geqslant 2.$$

Proof. Let $s = \lfloor \log_2 n \rfloor$. We have

$$\|X\|_r^r = \sum_{i=1}^{n} \big|X_i^*\big|^r \leqslant \sum_{k=0}^{s} 2^k \big|X_{2^k}^*\big|^r.$$

Theorem 3 yields

$$\mathbb{P}\biggl(\big|X_k^*\big|^r \geqslant C_3^r \log^r\biggl(\frac{en}{k}\biggr) + t^r\biggr) \leqslant \exp\biggl(-\frac{1}{C}\sqrt{k}t\biggr) \quad \text{for } t > 0. \tag{8}$$

Observe that

$$\sum_{k=0}^{s} 2^k \log^r\big(en2^{-k}\big) \leqslant Cn \sum_{j=1}^{\infty} j^r 2^{-j} \leqslant (Cr)^r n.$$

Thus for $t_1, \ldots, t_k \geqslant 0$ we get

$$\mathbb{P}\Biggl(\|X\|_r \geqslant C\Biggl(rn^{1/r} + \Biggl(\sum_{k=0}^{s} t_k\Biggr)^{1/r}\Biggr)\Biggr) \leqslant \mathbb{P}\Biggl(\sum_{k=0}^{s} Y_k \geqslant \sum_{k=0}^{s} t_k\Biggr),$$

where

$$Y_k := 2^k\big(\big|X_{2^k}^*\big|^r - C_3^r \log^r\big(en2^{-k}\big)\big).$$

Hence by (8)

$$\begin{aligned}\mathbb{P}\Biggl(\|X\|_r \geqslant C\Biggl(rn^{1/r} + \Biggl(\sum_{k=0}^{s} t_k\Biggr)^{1/r}\Biggr)\Biggr) &\leqslant \sum_{k=0}^{s} \mathbb{P}(Y_k \geqslant t_k) \\ &\leqslant \sum_{k=0}^{s} \exp\biggl(-\frac{1}{C} 2^{\frac{k}{2}-\frac{k}{r}} t_k^{1/r}\biggr).\end{aligned}$$

Fix $t > 0$ and choose t_k such that $t = 2^{k/2-k/r} t_k^{1/r}$. Then

$$\sum_{k=0}^{s} t_k = t^r \sum_{k=0}^{s} 2^{\frac{k(2-r)}{2}} \leqslant t^r \bigl(1 - 2^{\frac{2-r}{2}}\bigr)^{-1} \leqslant C t^r \frac{r}{r-2},$$

so we get

$$\mathbb{P}\left(\|X\|_r \geqslant C\left(r n^{1/r} + t\left(\frac{r}{r-2}\right)^{1/r}\right)\right) \leqslant (\log_2 n + 1)\exp\left(-\frac{1}{C}t\right). \qquad \square$$

Proof of Theorem 8. Observe that $(\frac{r}{r-2})^{1/r} \leqslant C(1+\delta^{-1/2})$ for $r \geqslant 2+\delta$ and $\log n \leqslant r n^{1/r}$ and apply Proposition 9. $\square$

4. Proof of Theorem 4

Our crucial tool will be the following result.

Proposition 10. *Let X be an isotropic log-concave n-dimensional random vector, $A = \{X \in K\}$, where K is a convex set in $\mathbb{R}^n$ such that $0 < \mathbb{P}(A) \leqslant 1/e$. Then*

$$\sum_{i=1}^{n} \mathbb{P}\bigl(A \cap \{X_i \geqslant t\}\bigr) \leqslant C_1 \mathbb{P}(A)\bigl(t^{-2}\log^2\bigl(\mathbb{P}(A)\bigr) + n e^{-t/C_1}\bigr) \quad \textit{for } t \geqslant C_1. \tag{9}$$

Moreover for $1 \leqslant u \leqslant \frac{t}{C_2}$,

$$\#\bigl\{i \leqslant n\colon \mathbb{P}\bigl(A \cap \{X_i \geqslant t\}\bigr) \geqslant e^{-u}\mathbb{P}(A)\bigr\} \leqslant \frac{C_2 u^2}{t^2}\log^2\bigl(\mathbb{P}(A)\bigr). \tag{10}$$

Proof. Let Y be a random vector distributed as the vector X conditioned on the set A that is

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

Notice that in particular for any set B, $\mathbb{P}(X \in B) \geqslant \mathbb{P}(A)\mathbb{P}(Y \in B)$.

The vector Y is log-concave, but no longer isotropic. Since this is only a matter of permutation of coordinates we may assume that $\mathbb{E}Y_1^2 \geqslant \mathbb{E}Y_2^2 \geqslant \cdots \geqslant \mathbb{E}Y_n^2$.

For $\alpha > 0$ let

$$m = m(\alpha) = \#\bigl\{i\colon \mathbb{E}Y_i^2 \geqslant \alpha\bigr\}.$$

We have $\mathbb{E}Y_1^2 \geqslant \cdots \geqslant \mathbb{E}Y_m^2 \geqslant \alpha$. Hence by the Paley–Zygmund inequality,

$$\mathbb{P}\left(\sum_{i=1}^{m} Y_i^2 \geqslant \frac{1}{2}\alpha m\right) \geqslant \mathbb{P}\left(\sum_{i=1}^{m} Y_i^2 \geqslant \frac{1}{2}\mathbb{E}\sum_{i=1}^{m} Y_i^2\right) \geqslant \frac{1}{4}\frac{(\mathbb{E}\sum_{i=1}^{m} Y_i^2)^2}{\mathbb{E}(\sum_{i=1}^{m} Y_i^2)^2} \geqslant \frac{1}{C}.$$

This implies that

$$\mathbb{P}\left(\sum_{i=1}^{m} X_i^2 \geqslant \frac{1}{2}\alpha m\right) \geqslant \frac{1}{C}\mathbb{P}(A).$$

However by the result of Paouris,

$$\mathbb{P}\left(\sum_{i=1}^{m} X_i^2 \geqslant \frac{1}{2}\alpha m\right) \leqslant \exp\left(-\frac{1}{C_3}\sqrt{m\alpha}\right) \quad \text{for } \alpha \geqslant C_3.$$

So for $\alpha \geqslant C_3$, $\exp(-\frac{1}{C_3}\sqrt{m\alpha}) \geqslant \mathbb{P}(A)/C$ and we get that

$$m(\alpha) = \#\{i\colon \mathbb{E}Y_i^2 \geqslant \alpha\} \leqslant \frac{C_4}{\alpha}\log^2\bigl(\mathbb{P}(A)\bigr) \quad \text{for } \alpha \geqslant C_3. \tag{11}$$

We have

$$\frac{\mathbb{P}(A \cap \{X_i \geqslant t\})}{\mathbb{P}(A)} = \mathbb{P}(Y_i \geqslant t) \leqslant \exp\left(1 - \frac{t}{C(\mathbb{E}Y_i^2)^{1/2}}\right)$$

and (10) follows by (11).

Take $t \geqslant \sqrt{C_3}$ and let k_0 be a nonnegative integer such that $2^{-k_0}t \geqslant \sqrt{C_3} \geqslant 2^{-k_0-1}t$. Define

$$I_0 = \{i\colon \mathbb{E}Y_i^2 \geqslant t^2\}, \qquad I_{k_0+1} = \{i\colon \mathbb{E}Y_i^2 < 4^{-k_0}t^2\}$$

and

$$I_j = \{i\colon 4^{-j}t^2 \leqslant \mathbb{E}Y_i^2 < 4^{1-j}t^2\}, \quad j = 1, 2, \ldots, k_0.$$

By (11) we get

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

and obviously $\#I_{k_0+1} \leqslant n$. Moreover for $i \in I_j$, $j \neq 0$,

$$\mathbb{P}(Y_i \geqslant t) \leqslant \mathbb{P}\left(\frac{Y_i}{(\mathbb{E}Y_i^2)^{1/2}} \geqslant 2^{j-1}\right) \leqslant \exp\left(1 - \frac{1}{C}2^j\right).$$

Thus

$$\begin{aligned}\sum_{i=1}^{n}\mathbb{P}(Y_i \geqslant t) &= \sum_{j=0}^{k_0+1}\sum_{i\in I_j}\mathbb{P}(Y_i \geqslant t) \leqslant \#I_0 + e\sum_{j=1}^{k_0+1}\#I_j \exp\left(-\frac{1}{C}2^j\right)\\ &\leqslant C_4\left(t^{-2}\log^2\mathbb{P}(A)\left(1 + e\sum_{j=1}^{k_0}2^{2j}\exp\left(-\frac{1}{C}2^j\right)\right) + ene^{-t/C}\right)\\ &\leqslant C_1\bigl(t^{-2}\log^2\mathbb{P}(A) + ne^{-t/C_1}\bigr).\end{aligned}$$

To finish the proof of (9) it is enough to observe that

$$\sum_{i=1}^{n} \mathbb{P}\big(A \cap \{X_i \geqslant t\}\big) = \mathbb{P}(A) \sum_{i=1}^{n} \mathbb{P}(Y_i \geqslant t). \qquad \square$$

The following two examples show that estimate (9) is close to be optimal.

Example 1. Take $X_1, X_2, \ldots, X_n$ to be independent symmetric exponential random variables with variance 1 and $A = \{X_1 \geqslant \sqrt{2}\}$. Then $\mathbb{P}(A) = \frac{1}{2e}$ and

$$\sum_{i=2}^{n} \mathbb{P}\big(A \cap \{X_i \geqslant t\}\big) = \mathbb{P}(A) \sum_{i=2}^{n} \mathbb{P}(X_i \geqslant t) = (n-1)\mathbb{P}(A) \exp(-t/\sqrt{2}),$$

therefore the factor $ne^{-t/C}$ in (9) is necessary.

Example 2. Take $A = \{X_1 \geqslant t, \ldots, X_k \geqslant t\}$ then

$$\sum_{i=1}^{n} \mathbb{P}\big(A \cap \{X_i \geqslant t\}\big) \geqslant k\mathbb{P}(A).$$

So improvement of the factor $t^{-2}\mathbb{P}(A) \log^2 \mathbb{P}(A)$ in (9) would imply in particular a better estimate of $\mathbb{P}(X_1 \geqslant t, \ldots, X_k \geqslant t)$ than $\exp(-\frac{1}{C}\sqrt{k}t)$ and we do not know if such bound is possible to obtain.

Proof of Theorem 4. We have $N_X \leqslant n$, so the statement is obvious if $t\sqrt{n} \leqslant Cp$, in the sequel we will assume that $t\sqrt{n} \geqslant 10p$.

Let C_1 and C_2 be as in Proposition 10 – increasing C_i if necessary we may assume that $\mathbb{P}(X_1 \geqslant t) \leqslant e^{-t/C_i}$ for $t \geqslant C_i$ and $i = 1, 2$. Let us fix $p \geqslant 1$ and $t \geqslant C \log(\frac{nt^2}{p^2})$, then $t \geqslant \max\{C_1, 4C_2\}$ and $t^2 ne^{-t/C_1} \leqslant p^2$ if C is large enough. Let l be a positive integer such that

$$p \leqslant l \leqslant 2p \quad \text{and} \quad l = 2^k \quad \text{for some integer } k.$$

Since $(\mathbb{E}(N_X(t))^p)^{1/p} \leqslant (\mathbb{E}(N_X(t))^l)^{1/l}$ it is enough to show that

$$\mathbb{E}\big(t^2 N_X(t)\big)^l \leqslant (Cl)^{2l}.$$

Recall that by our assumption on p, we have $t\sqrt{n} \geqslant 5l$.

To shorten the notation let

$$B_{i_1,\ldots,i_s} = \{X_{i_1} \geqslant t, \ldots, X_{i_s} \geqslant t\} \quad \text{and} \quad B_\emptyset = \Omega.$$

Define

$$m(l) := \mathbb{E}N_X(t)^l = \mathbb{E}\left(\sum_{i=1}^{n} \mathbb{1}_{\{X_i \geqslant t\}}\right)^l = \sum_{i_1,\ldots,i_l=1}^{n} \mathbb{P}(B_{i_1,\ldots,i_l}),$$

we need to show that

$$m(l) \leqslant \left(\frac{Cl}{t}\right)^{2l}. \tag{12}$$

We divide the sum in $m(l)$ into several parts. Let $j_1 \geqslant 2$ be such integer that

$$2^{j_1-2} < \log\left(\frac{nt^2}{l^2}\right) \leqslant 2^{j_1-1}.$$

We set

$$I_0 = \left\{(i_1,\ldots,i_l) \in \{1,\ldots,n\}^l \colon \mathbb{P}(B_{i_1,\ldots,i_l}) > e^{-l}\right\},$$
$$I_j = \left\{(i_1,\ldots,i_l) \in \{1,\ldots,n\}^l \colon \mathbb{P}(B_{i_1,\ldots,i_l}) \in \left(e^{-2^j l}, e^{-2^{j-1} l}\right]\right\}, \quad 0 < j < j_1$$

and

$$I_{j_1} = \left\{(i_1,\ldots,i_l) \in \{1,\ldots,n\}^l \colon \mathbb{P}(B_{i_1,\ldots,i_l}) \leqslant e^{-2^{j_1-1} l}\right\}.$$

Since $\{1,\ldots,n\}^l = \bigcup_{j=0}^{j_1} I_j$ we get $m(l) = \sum_{j=0}^{j_1} m_j(l)$, where

$$m_j(l) := \sum_{(i_1,\ldots,i_l)\in I_j} \mathbb{P}(B_{i_1,\ldots,i_l}) \quad \text{for } 0 \leqslant j \leqslant j_1.$$

It is easy to bound $m_{j_1}(l)$ – namely since $\#I_{j_1} \leqslant n^l$ we have

$$\sum_{(i_1,\ldots,i_l)\in I_{j_1}} \mathbb{P}(B_{i_1,\ldots,i_l}) \leqslant n^l e^{-2^{j_1-1} l} \leqslant \left(\frac{l}{t}\right)^{2l}.$$

To estimate $m_0(l)$ we define first for $I \subset \{1,\ldots,n\}^l$ and $1 \leqslant s \leqslant l$,

$$P_s I = \left\{(i_1,\ldots,i_s) \colon (i_1,\ldots,i_l) \in I \text{ for some } i_{s+1},\ldots,i_l\right\}.$$

By Proposition 10 we get for $s = 1,\ldots,l-1$

$$\begin{aligned}
&\sum_{(i_1,\ldots,i_{s+1})\in P_{s+1} I_0} \mathbb{P}(B_{i_1,\ldots,i_{s+1}}) \\
&\quad \leqslant \sum_{(i_1,\ldots,i_s)\in P_s I_0} \sum_{i_{s+1}=1}^{n} \mathbb{P}\left(B_{i_1,\ldots,i_s} \cap \{X_{i_{s+1}} \geqslant t\}\right) \\
&\quad \leqslant C_1 \sum_{(i_1,\ldots,i_s)\in P_s I_0} \mathbb{P}(B_{i_1,\ldots,i_s})\left(t^{-2}\log^2 \mathbb{P}(B_{i_1,\ldots,i_s}) + ne^{-t/C_1}\right).
\end{aligned}$$

Observe that we have $\mathbb{P}(B_{i_1,\dots,i_s}) > e^{-l}$ for $(i_1,\dots,i_s) \in P_s I_0$ and recall that $t^2 n e^{-t/C_1} \leqslant p^2 \leqslant 4l^2$, hence

$$\sum_{(i_1,\dots,i_{s+1})\in P_{s+1}I_0} \mathbb{P}(B_{i_1,\dots,i_{s+1}}) \leqslant 5C_1 t^{-2} l^2 \sum_{(i_1,\dots,i_s)\in P_s I_0} \mathbb{P}(B_{i_1,\dots,i_s}).$$

So, by easy induction we obtain

$$m_0(l) = \sum_{(i_1,\dots,i_l)\in I_0} \mathbb{P}(B_{i_1,\dots,i_l}) \leqslant \left(5C_1 t^{-2} l^2\right)^{l-1} \sum_{i_1\in P_1 I_0} \mathbb{P}(B_{i_1})$$
$$\leqslant \left(5C_1 t^{-2} l^2\right)^{l-1} n e^{-t/C_1} \leqslant \left(\frac{Cl}{t}\right)^{2l}.$$

Now comes the most involved part of the proof – estimating $m_j(l)$ for $0 < j < j_1$. It is based on suitable bounds for $\#I_j$. We will need the following simple combinatorial lemma.

Lemma 11. *Let $l_0 \geqslant l_1 \geqslant \cdots \geqslant l_s$ be a fixed sequence of positive integers and*

$$\mathcal{F} = \left\{ f : \{1,2,\dots,l_0\} \to \{0,1,2,\dots,s\} \colon \forall_{1\leqslant i\leqslant s}\ \#\left\{r \colon f(r) \geqslant i\right\} \leqslant l_i \right\}.$$

Then

$$\#\mathcal{F} \leqslant \prod_{i=1}^{s} \left(\frac{e l_{i-1}}{l_i}\right)^{l_i}.$$

Proof of Lemma 11. Notice that any function $f : \{1,2,\dots,l_0\} \to \{0,1,2,\dots,s\}$ is determined by the sets $A_i = \{r \colon f(r) \geqslant i\}$ for $i = 0,1,\dots,s$. Take $f \in \mathcal{F}$, obviously $A_0 = \{1,\dots,l_0\}$. If the set A_{i-1} of cardinality $a_{i-1} \leqslant l_{i-1}$ is already chosen then the set $A_i \subset A_{i-1}$ of cardinality at most l_i may be chosen in

$$\binom{a_{i-1}}{0} + \binom{a_{i-1}}{1} + \cdots + \binom{a_{i-1}}{l_i} \leqslant \binom{l_{i-1}}{0} + \binom{l_{i-1}}{1} + \cdots + \binom{l_{i-1}}{l_i} \leqslant \left(\frac{e l_{i-1}}{l_i}\right)^{l_i}$$

ways. □

We come back to the proof of Theorem 4. Fix $0 < j < j_1$, let r_1 be a positive integer such that

$$2^{r_1} < \frac{t}{C_2} \leqslant 2^{r_1+1}.$$

For $(i_1,\dots,i_l) \in I_j$ we define a function $f_{i_1,\dots,i_l} : \{1,\dots,l\} \to \{j, j+1,\dots,r_1\}$ by the formula

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

Notice that for all i_1, $\mathbb{P}(X_{i_1} \geqslant t) \leqslant e^{-t/C_2} < \exp(-2^{r_1})\mathbb{P}(B_\emptyset)$, so $f_{i_1,\dots,i_l}(1) = r_1$ for all $i_1, \dots, i_l$.

Put

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

For $f = f_{i_1,\dots,i_l} \in \mathcal{F}_j$ and $r > j$, we have

$$\exp(-2^j l) < \mathbb{P}(B_{i_1,\dots,i_l}) < \exp(-2^r \#\{s\colon f(s) \geqslant r\}),$$

so

$$\#\{s\colon f(s) \geqslant r\} \leqslant 2^{j-r} l =: l_r. \tag{13}$$

Observe that the above inequality holds also for $r = j$. We have $l_{r-1}/l_r = 2$ and $\sum_{r=j+1}^{r_1} l_r \leqslant l$ so by Lemma 11 we get

$$\#\mathcal{F}_j \leqslant \prod_{r=j+1}^{r_1} \left(\frac{e l_{r-1}}{l_r}\right)^{l_r} \leqslant e^{2l}.$$

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

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

Put

$$n_r := \#\{s \in \{1,\dots,l\}\colon f(s) = r\}, \quad r = j, j+1, \dots, r_1.$$

We have

$$n_j + n_{j+1} + \dots + n_{r_1} = l,$$

moreover if $i_1, \dots, i_{s-1}$ are fixed and $f(s) = r < r_1$ then $s \geqslant 2$ and by the second part of Proposition 10 (with $u = 2^{r+1} \leqslant t/C_2$) i_s may take at most

$$\frac{4C_2 2^{2r}}{t^2} \log^2 \mathbb{P}(B_{i_1,\dots,i_{s-1}}) \leqslant \frac{4C_2 2^{2(r+j)} l^2}{t^2} \leqslant \frac{4C_2 l^2}{t^2} \exp(2(r+j)) =: m_r$$

values. Thus

$$\#I_j(f) \leqslant n^{n_{r_1}} \prod_{r=j}^{r_1-1} m_r^{n_r} = n^{n_{r_1}} \left(\frac{4C_2 l^2}{t^2}\right)^{l-n_{r_1}} \exp\left(\sum_{r=j}^{r_1-1} 2(r+j) n_r\right).$$

Observe that by previously derived estimate (13) we get

$$n_r \leqslant l_r = 2^{j-r} l,$$

hence

$$\sum_{r=j}^{r_1-1} 2(r+j)n_r \leqslant 2^{j+2} l \sum_{r=j}^{\infty} r 2^{-r} \leqslant \bigl(C + 2^{j-2}\bigr) l.$$

We also have

$$n_{r_1} \leqslant 2^{j-r_1} l \leqslant \frac{2C_2}{t} 2^j l \leqslant \frac{1}{\log(nt^2/(4l^2))} 2^{j-3} l,$$

where the last inequality holds since $t \geqslant C \log(nt^2/(4l^2))$ and C may be taken arbitrarily large. So we get that for any $f \in \mathcal{F}_j$,

$$\#I_j(f) \leqslant \left(\frac{Cl^2}{t^2}\right)^l \left(\frac{nt^2}{4l^2}\right)^{n_{r_1}} \exp\bigl(2^{j-2} l\bigr) \leqslant \left(\frac{Cl^2}{t^2}\right)^l \exp\left(\frac{3}{8} 2^j l\right).$$

This shows that

$$\#I_j \leqslant \#\mathcal{F}_j \cdot \left(\frac{Cl^2}{t^2}\right)^l \exp\left(\frac{3}{8} 2^j l\right) \leqslant \left(\frac{Cl^2}{t^2}\right)^l \exp\left(\left(2 + \frac{3}{8} 2^j\right) l\right).$$

Hence

$$m_j(l) = \sum_{(i_1,\dots,i_l) \in I_j} \mathbb{P}(B_{i_1,\dots,i_l}) \leqslant \#I_j \exp\bigl(-2^{j-1} l\bigr) \leqslant \left(\frac{Cl^2}{t^2}\right)^l \exp\bigl(-2^{j-3} l\bigr).$$

Therefore

$$m(l) = m_0(l) + m_{j_1}(l) + \sum_{j=1}^{j_1-1} m_j(l) \leqslant \left(\frac{l}{t}\right)^{2l} \left(C^l + 1 + \sum_{j=1}^{\infty} C^l \exp\bigl(-2^{j-3} l\bigr)\right)$$
$$\leqslant \left(\frac{Cl}{t}\right)^{2l}$$

and (12) holds. □

Acknowledgments

This work was done while the author was taking part in the Thematic Program on Asymptotic Geometric Analysis at the Fields Institute in Toronto. The author would like to thank Radosław Adamczak and Nicole Tomczak-Jaegermann for their helpful comments.

References

[1] R. Adamczak, R. Latała, A. Litvak, A. Pajor, N. Tomczak-Jaegermann, Geometry of log-concave ensembles of random matrices and approximate reconstruction, preprint.
[2] S.G. Bobkov, F.L. Nazarov, On convex bodies and log-concave probability measures with unconditional basis, in: Geometric Aspects of Functional Analysis, in: Lecture Notes in Math., vol. 1807, Springer, Berlin, 2003, pp. 53–69.
[3] C. Borell, Convex measures on locally convex spaces, Ark. Mat. 12 (1974) 239–252.
[4] O. Guedon, E. Milman, Interpolating thin-shell and sharp large-deviation estimates for isotropic log-concave measures, preprint, http://arxiv.org/abs/1011.0943.
[5] N. Huet, Spectral gap for some invariant log-concave probability measures, preprint, http://arxiv.org/abs/1003.4839.
[6] R. Kannan, L. Lovász, M. Simonovits, Isoperimetric problems for convex bodies and a localization lemma, Discrete Comput. Geom. 13 (1995) 541–559.
[7] B. Klartag, A Berry–Esseen type inequality for convex bodies with an unconditional basis, Probab. Theory Related Fields 145 (2009) 1–33.
[8] R. Latała, Tail and moment estimates for sums of independent random vectors with logarithmically concave tails, Studia Math. 118 (1996) 301–304.
[9] E. Milman, On the role of convexity in isoperimetry, spectral gap and concentration, Invent. Math. 177 (2009) 1–43.
[10] G. Paouris, Concentration of mass on convex bodies, Geom. Funct. Anal. 16 (2006) 1021–1049.