TWO-SIDED ESTIMATES FOR ORDER STATISTICS OF LOG-CONCAVE RANDOM VECTORS

RAFAŁ LATAŁA AND MARTA STRZELECKA

Abstract. We establish two-sided bounds for expectations of order statistics (k-th maxima) of moduli of coordinates of centered log-concave random vectors with uncorrelated coordinates. Our bounds are exact up to multiplicative universal constants in the unconditional case for all k and in the isotropic case for $k \leq n - cn^{5/6}$. We also derive two-sided estimates for expectations of sums of k largest moduli of coordinates for some classes of random vectors.

1. Introduction and main results

For a vector $x \in \mathbb{R}^n$ let $k\text{-}\max x_i$ (or $k\text{-}\min x_i$) denote its *k-th maximum* (respectively its *k-th minimum*), i.e. its k-th maximal (respectively k-th minimal) coordinate. For a random vector $X = (X_1, \ldots, X_n)$, $k\text{-}\min X_i$ is also called the k-th order statistic of X.

Let $X = (X_1, \ldots, X_n)$ be a random vector with finite first moment. In this note we try to estimate $\mathbb{E}k\text{-}\max_i |X_i|$ and

$$\mathbb{E} \max_{|I|=k} \sum_{i\in I} |X_i| = \mathbb{E} \sum_{l=1}^{k} l\text{-}\max_i |X_i|.$$

Order statistics play an important role in various statistical applications and there is an extensive literature on this subject (cf. [2, 5] and references therein).

We put special emphasis on the case of log-concave vectors, i.e. random vectors X satisfying the property $\mathbb{P}(X \in \lambda K + (1-\lambda)L) \geq \mathbb{P}(X \in K)^{\lambda}\mathbb{P}(X \in L)^{1-\lambda}$ for any $\lambda \in [0,1]$ and any nonempty compact sets K and L. By the result of Borell [3] a vector X with full dimensional support is log-concave if and only if it has a log-concave density, i.e. the density of a form $e^{-h(x)}$ where h is convex with values in $(-\infty, \infty]$. A typical example of a log-concave vector is a vector uniformly distributed over a convex body. In recent years the study of log-concave vectors attracted attention of many researchers, cf. monographs [1, 4].

Date: Revised version.

The research of RL was supported by the National Science Centre, Poland grant 2015/18/A/ST1/00553 and of MS by the National Science Centre, Poland grants 2015/19/N/ST1/02661 and 2018/28/T/ST1/00001.

To bound the sum of k largest coordinates of X we define

$$t(k,X) := \inf\left\{t>0\colon \frac{1}{t}\sum_{i=1}^{n}\mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\geq t\}} \leq k\right\}. \tag{1}$$

and start with an easy upper bound.

Proposition 1. *For any random vector X with finite first moment we have*

$$\mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i| \leq 2kt(k,X). \tag{2}$$

Proof. For any $t>0$ we have

$$\max_{|I|=k}\sum_{i\in I}|X_i| \leq tk + \sum_{i=1}^{n}|X_i|\mathbf{1}_{\{|X_i|\geq t\}}.$$

□

It turns out that this bound may be reversed for vectors with independent coordinates or, more generally, vectors satisfying the following condition

$$\mathbb{P}(|X_i|\geq s, |X_j|\geq t) \leq \alpha\mathbb{P}(|X_i|\geq s)\mathbb{P}(|X_j|\geq t) \qquad \text{for all } i\neq j \text{ and all } s,t>0. \tag{3}$$

If $\alpha=1$ this means that moduli of coordinates of X are negatively correlated.

Theorem 2. *Suppose that a random vector X satisfies condition* (3) *with some $\alpha\geq 1$. Then there exists a constant $c(\alpha)>0$ which depends only on α such that for any $1\leq k\leq n$,*

$$c(\alpha)kt(k,X) \leq \mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i| \leq 2kt(k,X).$$

We may take $c(\alpha)=(288(5+4\alpha)(1+2\alpha))^{-1}$.

In the case of i.i.d. coordinates two-sided bounds for $\mathbb{E}\max_{|I|=k}\sum_{i\in I}|a_iX_i|$ in terms of an Orlicz norm (related to the distribution of X_i) of a vector $(a_i)_{i\leq n}$ where known before, see [7].

Log-concave vectors with diagonal covariance matrices behave in many aspects like vectors with independent coordinates. This is true also in our case.

Theorem 3. *Let X be a log-concave random vector with uncorrelated coordinates (i.e. $\mathrm{Cov}(X_i,X_j)=0$ for $i\neq j$). Then for any $1\leq k\leq n$,*

$$ckt(k,X) \leq \mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i| \leq 2kt(k,X).$$

In the above statement and in the sequel c and C denote positive universal constants.

The next two examples show that the lower bound cannot hold if $n\gg k$ and only marginal distributions of X_i are log-concave or the coordinates of X are highly correlated.

Example 1. Let $X=(\varepsilon_1 g, \varepsilon_2 g, \ldots, \varepsilon_n g)$, where $\varepsilon_1,\ldots,\varepsilon_n, g$ are independent, $\mathbb{P}(\varepsilon_i=\pm 1)=1/2$ and g has the normal $\mathcal{N}(0,1)$ distribution. Then $\mathrm{Cov}X=\mathrm{Id}$ and it is not hard to check that $\mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i| = k\sqrt{2/\pi}$ and $t(k,X)\sim \ln^{1/2}(n/k)$ if $k\leq n/2$.

Example 2. Let $X = (g, \ldots, g)$, where $g \sim \mathcal{N}(0,1)$. Then, as in the previous example, $\mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i| = k\sqrt{2/\pi}$ and $t(k,X) \sim \ln^{1/2}(n/k)$.

Question 1. Let $X' = (X_1', X_2', \ldots, X_n')$ be a decoupled version of X, i.e. X_i' are independent and X_i' has the same distribution as X_i. Due to Theorem 2 (applied to X'), the assertion of Theorem 3 may be stated equivalently as

$$\mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i| \sim \mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i'|.$$

Is the more general fact true that for any symmetric norm and any log-concave vector X with uncorrelated coordinates

$$\mathbb{E}\|X\| \sim \mathbb{E}\|X'\|?$$

Maybe such an estimate holds at least in the case of unconditional log-concave vectors?

We turn our attention to bounding k-maxima of $|X_i|$. This was investigated in [8] (under some strong assumptions on the function $t \mapsto \mathbb{P}(|X_i| \geq t)$) and in the weighted i.i.d. setting in [7, 9, 15]. We will give different bounds valid for log-concave vectors, in which we do not have to assume independence, nor any special conditions on the growth of the distribution function of the coordinates of X. To this end we need to define another quantity:

$$t^*(p,X) := \inf\left\{t>0\colon\ \sum_{i=1}^n \mathbb{P}(|X_i|\geq t) \leq p\right\} \quad \text{for } 0<p<n.$$

Theorem 4. *Let X be a mean zero log-concave n-dimensional random vector with uncorrelated coordinates and $1 \leq k \leq n$. Then*

$$\mathbb{E}k\text{-}\max_{i\leq n}|X_i| \geq \frac{1}{2}\mathrm{Med}\Big(k\text{-}\max_{i\leq n}|X_i|\Big) \geq ct^*\left(k-\frac{1}{2},X\right).$$

Moreover, if X is additionally unconditional then

$$\mathbb{E}k\text{-}\max_{i\leq n}|X_i| \leq Ct^*\left(k-\frac{1}{2},X\right).$$

The next theorem provides an upper bound in the general log-concave case.

Theorem 5. *Let X be a mean zero log-concave n-dimensional random vector with uncorrelated coordinates and $1 \leq k \leq n$. Then*

$$\mathbb{P}\left(k\text{-}\max_{i\leq n}|X_i| \geq Ct^*\left(k-\frac{1}{2},X\right)\right) \leq 1-c \tag{4}$$

and

$$\mathbb{E}k\text{-}\max_{i\leq n}|X_i| \leq Ct^*\left(k-\frac{1}{2}k^{5/6},X\right). \tag{5}$$

In the isotropic case (i.e. $\mathbb{E}X_i = 0, \mathrm{Cov}X = \mathrm{Id}$) one may show that $t^*(k/2,X) \sim t^*(k,X) \sim t(k,X)$ for $k \leq n/2$ and $t^*(p,X) \sim \frac{n-p}{n}$ for $p \geq n/4$ (see Lemma 24 below). In particular $t^*(n-k+1-(n-k+1)^{5/6}/2, X) \sim k/n + n^{-1/6}$ for $k \leq n/2$. This together with the two previous theorems implies the following corollary.

Corollary 6. *Let X be an isotropic log-concave n-dimensional random vector and $1 \leq k \leq n/2$. Then*

$$\mathbb{E}k\text{-}\max_{i\leq n}|X_i| \sim t^*(k,X) \sim t(k,X)$$

and

$$c\frac{k}{n} \leq \mathbb{E}k\text{-}\min_{i\leq n}|X_i| = \mathbb{E}(n-k+1)\text{-}\max_{i\leq n}|X_i| \leq C\left(\frac{k}{n} + n^{-1/6}\right).$$

If X is additionally unconditional then

$$\mathbb{E}k\text{-}\min_{i\leq n}|X_i| = \mathbb{E}(n-k+1)\text{-}\max_{i\leq n}|X_i| \sim \frac{k}{n}.$$

Question 2. Does the second part of Theorem 4 hold without the unconditionality assumptions? In particular, is it true that in the isotropic log-concave case $\mathbb{E}k\text{-}\min_{i\leq n}|X_i| \sim k/n$ for $1 \leq k \leq n/2$?

Notation. Throughout this paper by letters C, c we denote universal positive constants and by $C(\alpha), c(\alpha)$ constants depending only on the parameter α. The values of constants $C, c, C(\alpha), c(\alpha)$ may differ at each occurrence. If we need to fix a value of constant, we use letters $C_0, C_1, \ldots$ or $c_0, c_1, \ldots$. We write $f \sim g$ if $cf \leq g \leq Cg$. For a random variable Z we denote $\|Z\|_p = (\mathbb{E}|Z|^p)^{1/p}$. Recall that a random vector X is called isotropic, if $\mathbb{E}X = 0$ and $\mathrm{Cov}X = \mathrm{Id}$.

This note is organised as follows. In Section 2 we provide a lower bound for the sum of k largest coordinates, which involves the Poincaré constant of a vector. In Section 3 we use this result to obtain Theorem 3. In Section 4 we prove Theorem 2 and provide its application to comparison of weak and strong moments. In Section 5 we prove the first part of Theorem 4 and in Section 6 we prove the second part of Theorem 4, Theorem 5, and Lemma 24.

2. Exponential concentration

A probability measure μ on $\mathbb{R}^n$ satisfies *exponential concentration with constant $\alpha > 0$* if for any Borel set A with $\mu(A) \geq 1/2$,

$$1 - \mu(A + uB_2^n) \leq e^{-u/\alpha} \quad \text{for all } u > 0.$$

We say that a random n-dimensional vector satisfies exponential concentration if its distribution has such a property.

It is well known that exponential concentration is implied by the Poincaré inequality

$$\mathrm{Var}_\mu f \leq \beta \int |\nabla f|^2 d\mu \quad \text{for all bounded smooth functions } f\colon \mathbb{R}^n \mapsto \mathbb{R}$$

and $\alpha \leq 3\sqrt{\beta}$ (cf. [12, Corollary 3.2]).

Obviously, the constant in the exponential concentration is not linearly invariant. Typically one assumes that the vector is isotropic. For our purposes a more natural normalization will be that all coordinates have L_1-norm equal to 1.

The next proposition states that bound (2) may be reversed under the assumption that X satisfies the exponential concentration.

Proposition 7. *Assume that $Y=(Y_1,\ldots,Y_n)$ satisfies the exponential concentration with constant $\alpha>0$ and $\mathbb{E}|Y_i|\geq 1$ for all i. Then for any sequence $a=(a_i)_{i=1}^n$ of real numbers and $X_i:=a_iY_i$ we have*

$$\mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i|\geq\Big(8+64\frac{\alpha}{\sqrt{k}}\Big)^{-1}kt(k,X),$$

where $t(k,X)$ is given by (1).

We begin the proof with a few simple observations.

Lemma 8. *For any real numbers $z_1,\ldots,z_n$ and $1\leq k\leq n$ we have*

$$\max_{|I|=k}\sum_{i\in I}|z_i|=\int_0^\infty\min\Big\{k,\sum_{i=1}^n\mathbf{1}_{\{|z_i|\geq s\}}\Big\}ds.$$

Proof. Without loss of generality we may assume that $z_1\geq z_2\geq\ldots\geq z_n\geq 0$. Then

$$\int_0^\infty\min\Big\{k,\sum_{i=1}^n\mathbf{1}_{\{|z_i|\geq s\}}\Big\}ds=\sum_{l=1}^{k-1}\int_{z_{l+1}}^{z_l}lds+\int_0^{z_k}kds=\sum_{l=1}^{k-1}l(z_l-z_{l+1})+kz_k$$
$$=z_1+\ldots+z_k=\max_{|I|=k}\sum_{i\in I}|z_i|.\qquad\square$$

Fix a sequence $(X_i)_{i\leq n}$ and define for $s\geq 0$,

$$N(s):=\sum_{i=1}^n\mathbf{1}_{\{|X_i|\geq s\}}.\tag{6}$$

Corollary 9. *For any $k=1,\ldots,n$,*

$$\mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i|=\int_0^\infty\sum_{l=1}^k\mathbb{P}(N(s)\geq l)ds,$$

and for any $t>0$,

$$\mathbb{E}\sum_{i=1}^n|X_i|\mathbf{1}_{\{|X_i|\geq t\}}=t\mathbb{E}N(t)+\int_t^\infty\sum_{l=1}^\infty\mathbb{P}(N(s)\geq l)ds.$$

In particular

$$\mathbb{E}\sum_{i=1}^n|X_i|\mathbf{1}_{\{|X_i|\geq t\}}\leq\mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i|+\sum_{l=k+1}^\infty\Big(t\mathbb{P}(N(t)\geq l)+\int_t^\infty\mathbb{P}(N(s)\geq l)ds\Big).$$

Proof. We have

$$\int_0^\infty\sum_{l=1}^k\mathbb{P}(N(s)\geq l)ds=\int_0^\infty\mathbb{E}\min\{k,N(s)\}ds=\mathbb{E}\int_0^\infty\min\{k,N(s)\}ds$$
$$=\mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i|,$$

where the last equality follows by Lemma 8.

Moreover,

$$t\mathbb{E}N(t)+\int_t^\infty\sum_{l=1}^\infty \mathbb{P}(N(s)\geq l)ds = t\mathbb{E}N(t)+\int_t^\infty \mathbb{E}N(s)ds$$
$$=\mathbb{E}\sum_{i=1}^n\left(t\mathbf{1}_{\{|X_i|\geq t\}}+\int_t^\infty \mathbf{1}_{\{|X_i|\geq s\}}ds\right)$$
$$=\mathbb{E}\sum_{i=1}^n |X_i|\mathbf{1}_{\{|X_i|\geq t\}}.$$

The last part of the assertion easily follows, since

$$t\mathbb{E}N(t)=t\sum_{l=1}^n\mathbb{P}(N(t)\geq l)\leq \int_0^t\sum_{l=1}^k\mathbb{P}(N(s)\geq l)ds+\sum_{l=k+1}^\infty t\mathbb{P}(N(t)\geq l). \qquad \square$$

Proof of Proposition 7. To shorten the notation put $t_k := t(k,X)$. Without loss of generality we may assume that $a_1\geq a_2\geq\ldots\geq a_n\geq 0$ and $a_{\lceil k/4\rceil}=1$. Observe first that

$$\mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i|\geq \sum_{i=1}^{\lceil k/4\rceil} a_i\mathbb{E}|Y_i|\geq k/4,$$

so we may assume that $t_k\geq 16\alpha/\sqrt{k}$.

Let μ be the law of Y and

$$A:=\left\{y\in\mathbb{R}^n\colon\ \sum_{i=1}^n\mathbf{1}_{\{|a_iy_i|\geq\frac{1}{2}t_k\}}<\frac{k}{2}\right\}.$$

We have

$$\mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i|\geq\frac{k}{4}t_k\mathbb{P}\left(\sum_{i=1}^k\mathbf{1}_{\{|a_iY_i|\geq\frac{1}{2}t_k\}}\geq\frac{k}{2}\right)=\frac{k}{4}t_k(1-\mu(A)),$$

so we may assume that $\mu(A)\geq 1/2$.

Observe that if $y\in A$ and $\sum_{i=1}^n\mathbf{1}_{\{|a_iz_i|\geq s\}}\geq l>k$ for some $s\geq t_k$ then

$$\sum_{i=1}^n(z_i-y_i)^2\geq\sum_{i=\lceil k/4\rceil}^n(a_iz_i-a_iy_i)^2\geq(l-3k/4)(s-t_k/2)^2>\frac{ls^2}{16}.$$

Thus we have

$$\mathbb{P}(N(s)\geq l)\leq 1-\mu\left(A+\frac{s\sqrt{l}}{4}B_2^n\right)\leq e^{-\frac{s\sqrt{l}}{4\alpha}}\quad\text{ for } l>k,\ s\geq t_k.$$

Therefore

$$\int_{t_k}^\infty\mathbb{P}(N(s)\geq l)ds\leq\int_{t_k}^\infty e^{-\frac{s\sqrt{l}}{4\alpha}}ds=\frac{4\alpha}{\sqrt{l}}e^{-\frac{t_k\sqrt{l}}{4\alpha}}\quad\text{ for } l>k,$$

and

$$\sum_{l=k+1}^{\infty}\Big(t_k\mathbb{P}(N(t_k)\geq l)+\int_{t_k}^{\infty}\mathbb{P}(N(s)\geq l)ds\Big)\leq\sum_{l=k+1}^{\infty}\Big(t_k+\frac{4\alpha}{\sqrt{l}}\Big)e^{-\frac{t_k\sqrt{l}}{4\alpha}}$$
$$\leq\Big(t_k+\frac{4\alpha}{\sqrt{k+1}}\Big)\int_k^{\infty}e^{-\frac{t_k\sqrt{u}}{4\alpha}}du\leq\Big(t_k+\frac{4\alpha}{\sqrt{k+1}}\Big)e^{-\frac{t_k\sqrt{k}}{4\sqrt{2}\alpha}}\int_k^{\infty}e^{-\frac{t_k\sqrt{u-k}}{4\sqrt{2}\alpha}}du$$
$$=\Big(t_k+\frac{4\alpha}{\sqrt{k+1}}\Big)\frac{64\alpha^2}{t_k^2}e^{-\frac{t_k\sqrt{k}}{4\sqrt{2}\alpha}}\leq\Big(t_k+\frac{1}{4}t_k\Big)\frac{k}{4}\leq\frac{1}{2}kt_k,$$

where to get the next-to-last inequality we used the fact that $t_k\geq 16\alpha/\sqrt{k}$.

Hence Corollary 9 and the definition of t_k yields

$$kt_k\leq\mathbb{E}\sum_{i=1}^n|X_i|\mathbf{1}_{\{|X_i|\geq t_k\}}$$
$$\leq\mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i|+\sum_{l=k+1}^{\infty}\Big(t_k\mathbb{P}(N(t_k)\geq l)+\int_{t_k}^{\infty}\mathbb{P}(N(s)\geq l)ds\Big)$$
$$\leq\mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i|+\frac{1}{2}kt_k,$$

so $\mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i|\geq\frac{1}{2}kt_k$. □

We finish this section with a simple fact that will be used in the sequel.

Lemma 10. *Suppose that a measure μ satisfies exponential concentration with constant α. Then for any $c\in(0,1)$ and any Borel set A with $\mu(A)>c$ we have*

$$1-\mu(A+uB_2^n)\leq\exp\Big(-\Big(\frac{u}{\alpha}+\ln c\Big)_+\Big)\qquad\textit{for } u\geq 0.$$

Proof. Let $D:=\mathbb{R}^n\setminus(A+rB_2^n)$. Observe that $D+rB_2^n$ has an empty intersection with A so if $\mu(D)\geq 1/2$ then

$$c<\mu(A)\leq 1-\mu(D+rB_2^n)\leq e^{-r/\alpha},$$

and $r<\alpha\ln(1/c)$. Hence $\mu(A+\alpha\ln(1/c)B_2^n)\geq 1/2$, therefore for $s\geq 0$,

$$1-\mu(A+(s+\alpha\ln(1/c))B_2^n)=1-\mu((A+\alpha\ln(1/c)B_2^n)+sB_2^n)\leq e^{-s/\alpha},$$

and the assertion easily follows. □

3. Sums of largest coordinates of log-concave vectors

We will use the regular growth of moments of norms of log-concave vectors multiple times. By [4, Theorem 2.4.6], if $f:\mathbb{R}^n\to\mathbb{R}$ is a seminorm, and X is log-concave, then

$$(\mathbb{E}f(X)^p)^{1/p}\leq C_1\frac{p}{q}(\mathbb{E}f(X)^q)^{1/q}\quad\text{for } p\geq q\geq 1,\tag{7}$$

where C_1 is a universal constant.

We will also apply a few times the functional version of the Grünbaum inequality (see [14, Lemma 5.4]) which states that

$$\mathbb{P}(Z \geq 0) \geq \frac{1}{e} \quad \text{for any mean-zero log-concave random variable } Z. \tag{8}$$

Let us start with a few technical lemmas. The first one will be used to reduce proofs of Theorem 3 and lower bound in Theorem 4 to the symmetric case.

Lemma 11. *Let X be a log-concave n-dimensional vector and X' be an independent copy of X. Then for any $1 \leq k \leq n$,*

$$\mathbb{E} \max_{|I|=k} \sum_{i \in I} |X_i - X_i'| \leq 2\mathbb{E} \max_{|I|=k} \sum_{i \in I} |X_i|,$$

$$t(k, X) \leq et(k, X - X') + \frac{2}{k} \max_{|I|=k} \sum_{i \in I} \mathbb{E}|X_i|, \tag{9}$$

and

$$t^*(2k, X - X') \leq 2t^*(k, X). \tag{10}$$

Proof. The first estimate follows by the easy bound

$$\mathbb{E} \max_{|I|=k} \sum_{i \in I} |X_i - X_i'| \leq \mathbb{E} \max_{|I|=k} \sum_{i \in I} |X_i| + \mathbb{E} \max_{|I|=k} \sum_{i \in I} |X_i'| = 2\mathbb{E} \max_{|I|=k} \sum_{i \in I} |X_i|.$$

To get the second bound we may and will assume that $\mathbb{E}|X_1| \geq \mathbb{E}|X_2| \geq \ldots \geq \mathbb{E}|X_n|$. Let us define $Y := X - \mathbb{E}X$, $Y' := X' - \mathbb{E}X$ and $M := \frac{1}{k} \sum_{i=1}^{k} \mathbb{E}|X_i| \geq \max_{i \geq k} \mathbb{E}|X_i|$. Obviously

$$\sum_{i=1}^{k} \mathbb{E}|X_i| \mathbf{1}_{\{|X_i| \geq t\}} \leq kM \quad \text{for } t \geq 0. \tag{11}$$

We have $\mathbb{E}Y_i = 0$, thus $\mathbb{P}(Y_i \leq 0) \geq 1/e$ by (8). Hence

$$\mathbb{E}Y_i \mathbf{1}_{\{Y_i > t\}} \leq e\mathbb{E}Y_i \mathbf{1}_{\{Y_i > t, Y_i' \leq 0\}} \leq e\mathbb{E}|Y_i - Y_i'| \mathbf{1}_{\{Y_i - Y_i' > t\}} = e\mathbb{E}|X_i - X_i'| \mathbf{1}_{\{X_i - X_i' > t\}}$$

for $t \geq 0$. In the same way we show that

$$\mathbb{E}|Y_i| \mathbf{1}_{\{Y_i < -t\}} \leq e\mathbb{E}|Y_i| \mathbf{1}_{\{Y_i < -t, Y_i' \geq 0\}} \leq e\mathbb{E}|X_i - X_i'| \mathbf{1}_{\{X_i' - X_i > t\}}$$

Therefore

$$\mathbb{E}|Y_i| \mathbf{1}_{\{|Y_i| > t\}} \leq e\mathbb{E}|X_i - X_i'| \mathbf{1}_{\{|X_i - X_i'| > t\}}.$$

We have

$$\begin{aligned}\sum_{i=k+1}^n \mathbb{E}|X_i|\mathbf{1}_{\{|X_i|>et(k,X-X')+M\}} &\leq \sum_{i=k+1}^n \mathbb{E}|X_i|\mathbf{1}_{\{|Y_i|>et(k,X-X')\}}\\ &\leq \sum_{i=k+1}^n \mathbb{E}|Y_i|\mathbf{1}_{\{|Y_i|>t(k,X-X')\}} + \sum_{i=k+1}^n |\mathbb{E}X_i|\mathbb{P}(|Y_i|>et(k,X-X'))\\ &\leq e\sum_{i=1}^n \mathbb{E}|X_i-X_i'|\mathbf{1}_{\{|X_i-X_i'|>t(k,X-X')\}} + M\sum_{i=1}^n \mathbb{P}(|Y_i|>et(k,X-X'))\\ &\leq ekt(k,X-X') + M\sum_{i=1}^n \big(et(k,X-X')\big)^{-1}\mathbb{E}|Y_i|\mathbf{1}_{\{|Y_i|>et(k,X-X')\}}\\ &\leq ekt(k,X-X') + Mt(k,X-X')^{-1}\sum_{i=1}^n \mathbb{E}|X_i-X_i'|\mathbf{1}_{\{|X_i-X_i'|>t(k,X-X')\}}\\ &\leq ekt(k,X-X') + kM.\end{aligned}$$

Together with (11) we get

$$\sum_{i=1}^n \mathbb{E}|X_i|\mathbf{1}_{\{|X_i|>et(k,X-X')+M\}} \leq k(et(k,X-X')+2M)$$

and (9) easily follows.

In order to prove (10), note that for $u>0$,

$$\mathbb{P}(|X_i-X_i'|\geq 2u) \leq \mathbb{P}\big(\max\{|X_i|,|X_i'|\}\geq u\big) \leq 2\mathbb{P}\big(|X_i|\geq u\big),$$

thus the last part of the assertion follows by the definition of parameters t^*. □

Lemma 12. *Suppose that V is a real symmetric log-concave random variable. Then for any $t>0$ and $\lambda\in(0,1]$,*

$$\mathbb{E}|V|\mathbf{1}_{\{|V|\geq t\}} \leq \frac{4}{\lambda}\mathbb{P}(|V|\geq t)^{1-\lambda}\mathbb{E}|V|\mathbf{1}_{\{|V|\geq \lambda t\}}.$$

Moreover, if $\mathbb{P}(|V|\geq t)\leq 1/4$, then $\mathbb{E}|V|\mathbf{1}_{\{|V|\geq t\}}\leq 4t\mathbb{P}(|V|\geq t)$.

Proof. Without loss of generality we may assume that $\mathbb{P}(|V|\geq t)\leq 1/4$ (otherwise the first estimate is trivial).

Observe that $\mathbb{P}(|V|\geq s)=\exp(-N(s))$ where $N\colon[0,\infty)\to[0,\infty]$ is convex and $N(0)=0$. In particular

$$\mathbb{P}(|V|\geq \gamma t)\leq \mathbb{P}(|V|\geq t)^{\gamma} \quad \text{for } \gamma>1$$

and

$$\mathbb{P}(|V|\geq \gamma t)\geq \mathbb{P}(|V|\geq t)^{\gamma} \quad \text{for } \gamma\in[0,1].$$

We have

$$\mathbb{E}|V|\mathbf{1}_{\{|V|\geq t\}} \leq \sum_{k=0}^{\infty} 2^{k+1} t\mathbb{P}(|V| \geq 2^k t) \leq 2t \sum_{k=0}^{\infty} 2^k \mathbb{P}(|V| \geq t)^{2^k}$$

$$\leq 2t\mathbb{P}(|V| \geq t) \sum_{k=0}^{\infty} 2^k 4^{1-2^k} \leq 4t\mathbb{P}(|V| \geq t).$$

This implies the second part of the lemma.

To conclude the proof of the first bound it is enough to observe that

$$\mathbb{E}|V|\mathbf{1}_{\{|V|\geq \lambda t\}} \geq \lambda t\mathbb{P}(|V| \geq \lambda t) \geq \lambda t\mathbb{P}(|V| \geq t)^{\lambda}. \qquad \square$$

Proof of Theorem 3. By Proposition 1 it is enough to show the lower bound. By Lemma 11 we may assume that X is symmetric. We may also obviously assume that $\|X_i\|_2^2 = \mathbb{E}X_i^2 > 0$ for all i.

Let $Z = (Z_1, \ldots, Z_n)$, where $Z_i = X_i/\|X_i\|_2$. Then Z is log-concave, isotropic and, by (7), $\mathbb{E}|Z_i| \geq 1/(2C_1)$ for all i. Set $Y := 2C_1 Z$. Then $X_i = a_i Y_i$ and $\mathbb{E}|Y_i| \geq 1$. Moreover, since any m-dimensional projection of Z is a log-concave, isotropic m-dimensional vector, we know by the result of Lee and Vempala [13], that it satisfies the exponential concentration with a constants $Cm^{1/4}$. (In fact an easy modification of the proof below shows that for our purposes it would be enough to have exponential concentration with a constant Cm^{γ} for some $\gamma < 1/2$, so one may also use Eldan's result [6] which gives such estimates for any $\gamma > 1/3$). So any m-dimensional projection of Y satisfies exponential concentration with constant $C_2 m^{1/4}$.

Let us fix k and set $t := t(k, X)$, then (since X_i has no atoms)

$$\sum_{i=1}^{n} \mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\geq t\}} = kt. \tag{12}$$

For $l = 1, 2, \ldots$ define

$$I_l := \{i \in [n]\colon\ \beta^{l-1} \geq \mathbb{P}(|X_i| \geq t) \geq \beta^l\},$$

where $\beta = 2^{-8}$. By (12) there exists l such that

$$\sum_{i\in I_l} \mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\geq t\}} \geq kt2^{-l}.$$

Let us consider three cases.

(i) $l = 1$ and $|I_1| \leq k$. Then

$$\mathbb{E}\max_{|I|=k} \sum_{i\in I} |X_i| \geq \sum_{i\in I_1} \mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\geq t\}} \geq \frac{1}{2}kt.$$

(ii) $l = 1$ and $|I_1| > k$. Choose $J \subset I_1$ of cardinality k. Then

$$\mathbb{E}\max_{|I|=k} \sum_{i\in I} |X_i| \geq \sum_{i\in J} \mathbb{E}|X_i| \geq \sum_{i\in J} t\mathbb{P}(|X_i| \geq t) \geq \beta kt.$$

(iii) $l > 1$. By Lemma 12 (applied with $\lambda = 1/8$) we have

$$\sum_{i\in I_l} \mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\geq t/8\}} \geq \frac{1}{32}\beta^{-7(l-1)/8}\sum_{i\in I_l}\mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\geq t\}} \geq \frac{1}{32}\beta^{-7(l-1)/8}2^{-l}kt. \tag{13}$$

Moreover for $i \in I_l$, $\mathbb{P}(|X_i| \geq t) \leq \beta^{l-1} \leq 1/4$, so the second part of Lemma 12 yields

$$4t|I_l|\beta^{l-1} \geq \sum_{i\in I_l}\mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\geq t\}} \geq kt2^{-l}$$

and $|I_l| \geq \beta^{1-l}2^{-l-2}k = 2^{7l-10}k \geq k$.

Set $k' := \beta^{-7l/8}2^{-l}k = 2^{6l}k$. If $k' \geq |I_l|$ then, using (13), we estimate

$$\mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i| \geq \frac{k}{|I_l|}\sum_{i\in I_l}\mathbb{E}|X_i| \geq \beta^{7l/8}2^l\sum_{i\in I_l}\mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\geq t/8\}} \geq \frac{1}{32}\beta^{7/8}kt = 2^{-12}kt.$$

Otherwise set $X' = (X_i)_{i\in I_l}$ and $Y' = (Y_i)_{i\in I_l}$. By (12) we have

$$kt \geq \sum_{i\in I_l}\mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\geq t\}} \geq |I_l|t\beta^l,$$

so $|I_l| \leq k\beta^{-l}$ and Y' satisfies exponential concentration with constant $\alpha' = C_2k^{1/4}\beta^{-l/4}$. Estimate (13) yields

$$\sum_{i\in I_l}\mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\geq 2^{-12}t\}} \geq \sum_{i\in I_l}\mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\geq t/8\}} \geq 2^{-12}k't,$$

so $t(k', X') \geq 2^{-12}t$. Moreover, by Proposition 7 we have (since $k' \leq |I_l|$)

$$\mathbb{E}\max_{I\subset I_l,|I|=k'}\sum_{i\in I}|X_i| \geq \frac{1}{8+64\alpha'/\sqrt{k'}}k't(k',X').$$

To conclude observe that

$$\frac{\alpha'}{\sqrt{k'}} = C_2 2^{-l}k^{-1/4} \leq \frac{C_2}{4}$$

and since $k' \geq k$,

$$\mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i| \geq \frac{k}{k'}\mathbb{E}\max_{I\subset I_l,|I|=k'}\sum_{i\in I}|X_i| \geq \frac{1}{8+16C_2}2^{-12}tk.$$ □

4. Vectors satisfying condition (3)

Proof of Theorem 2. By Proposition 1 we need to show only the lower bound. Assume first that variables X_i have no atoms and $k \geq 4(1+\alpha)$.

Let $t_k = t(k, X)$. Then $\mathbb{E}\sum_{i=1}^n |X_i|\mathbf{1}_{\{|X_i|\geq t_k\}} = kt_k$. Note, that (3) implies that for all $i \neq j$ we have

$$\mathbb{E}|X_iX_j|\mathbf{1}_{\{|X_i|\geq t_k,|X_j|\geq t_k\}} \leq \alpha\mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\geq t_k\}}\mathbb{E}|X_j|\mathbf{1}_{\{|X_j|\geq t_k\}}. \tag{14}$$

We may assume that $\mathbb{E}\max_{|I|=k}\sum_{i\in I}|X_i| \leq \frac{1}{6}kt_k$, because otherwise the lower bound holds trivially.

Let us define

$$Y := \sum_{i=1}^{n} |X_i| \mathbf{1}_{\{kt_k \geq |X_i| \geq t_k\}} \quad \text{and} \quad A := (\mathbb{E}Y^2)^{1/2}.$$

Since

$$\mathbb{E} \max_{|I|=k} \sum_{i \in I} |X_i| \geq \mathbb{E}\left[\frac{1}{2} kt_k \mathbf{1}_{\{Y \geq kt_k/2\}}\right] = \frac{1}{2} kt_k \mathbb{P}\left(Y \geq \frac{kt_k}{2}\right),$$

it suffices to bound below the probability that $Y \geq kt_k/2$ by a constant depending only on α.

We have

$$\begin{aligned} A^2 = \mathbb{E}Y^2 &\leq \sum_{i=1}^{n} \mathbb{E}X_i^2 \mathbf{1}_{\{kt_k \geq |X_i| \geq t_k\}} + \sum_{i \neq j} \mathbb{E}|X_i X_j| \mathbf{1}_{\{|X_i| \geq t_k, |X_j| \geq t_k\}} \\ &\overset{(14)}{\leq} kt_k \mathbb{E}Y + \alpha \sum_{i \neq j} \mathbb{E}|X_i| \mathbf{1}_{\{|X_i| \geq t_k\}} \mathbb{E}|X_j| \mathbf{1}_{\{|X_j| \geq t_k\}} \\ &\leq kt_k A + \alpha \left(\sum_{i=1}^{n} \mathbb{E}|X_i| \mathbf{1}_{\{|X_i| \geq t_k\}} \right)^2 \leq \frac{1}{2}(k^2 t_k^2 + A^2) + \alpha k^2 t_k^2. \end{aligned}$$

Therefore $A^2 \leq (1 + 2\alpha)k^2 t_k^2$ and for any $l \geq k/2$ we have

$$\begin{aligned} \mathbb{E}Y \mathbf{1}_{\{Y \geq kt_k/2\}} &\leq lt_k \mathbb{P}(Y \geq kt_k/2) + \frac{1}{lt_k} \mathbb{E}Y^2 \\ &\leq lt_k \mathbb{P}(Y \geq kt_k/2) + (1 + 2\alpha)k^2 l^{-1} t_k. \end{aligned} \tag{15}$$

By Corollary 9 we have (recall definition (6))

$$\begin{aligned} \sum_{i=1}^{n} \mathbb{E}|X_i| \mathbf{1}_{\{|X_i| \geq kt_k\}} &\leq \mathbb{E} \max_{|I|=k} \sum_{i \in I} |X_i| + \sum_{l=k+1}^{\infty} \left(kt_k \mathbb{P}(N(kt_k) \geq l) + \int_{kt_k}^{\infty} \mathbb{P}(N(s) \geq l) ds \right) \\ &\leq \frac{1}{6} kt_k + \sum_{l=k+1}^{\infty} \left(kt_k \mathbb{E}N(kt_k)^2 l^{-2} + \int_{kt_k}^{\infty} \mathbb{E}N(s)^2 l^{-2} ds \right) \\ &\leq \frac{1}{6} kt_k + \frac{1}{k} \left(kt_k \mathbb{E}N(kt_k)^2 + \int_{kt_k}^{\infty} \mathbb{E}N(s)^2 ds \right). \end{aligned} \tag{16}$$

Assumption (3) implies that

$$\begin{aligned} \mathbb{E}N(s)^2 &= \sum_{i=1}^{n} \mathbb{P}(|X_i| \geq s) + \sum_{i \neq j} \mathbb{P}(|X_i| \geq s, |X_j| \geq s) \\ &\leq \sum_{i=1}^{n} \mathbb{P}(|X_i| \geq s) + \alpha \left(\sum_{i=1}^{n} \mathbb{P}(|X_i| \geq s) \right)^2. \end{aligned}$$

Moreover for $s \geq kt_k$ we have

$$\sum_{i=1}^{n} \mathbb{P}(|X_i| \geq s) \leq \frac{1}{s} \sum_{i=1}^{n} \mathbb{E}|X_i| \mathbf{1}_{\{|X_i| \geq s\}} \leq \frac{kt_k}{s} \leq 1,$$

so

$$\mathbb{E}N(s)^2 \leq (1+\alpha) \sum_{i=1}^{n} \mathbb{P}(|X_i| \geq s) \quad \text{ for } s \geq kt_k.$$

Thus

$$kt_k \mathbb{E}N(kt_k)^2 \leq kt_k(1+\alpha) \sum_{i=1}^{n} \mathbb{P}(|X_i| \geq kt_k) \leq (1+\alpha) \sum_{i=1}^{n} \mathbb{E}|X_i| \mathbf{1}_{\{|X_i| \geq kt_k\}},$$

and

$$\int_{kt_k}^{\infty} \mathbb{E}N(s)^2 ds \leq (1+\alpha) \sum_{i=1}^{n} \int_{kt_k}^{\infty} \mathbb{P}(|X_i| \geq s) ds \leq (1+\alpha) \sum_{i=1}^{n} \mathbb{E}|X_i| \mathbf{1}_{\{|X_i| \geq kt_k\}}.$$

This together with (16) and the assumption that $k \geq 4(1+\alpha)$ implies

$$\sum_{i=1}^{n} \mathbb{E}|X_i| \mathbf{1}_{\{|X_i| \geq kt_k\}} \leq \frac{1}{3} kt_k$$

and

$$\mathbb{E}Y = \sum_{i=1}^{n} \mathbb{E}|X_i| \mathbf{1}_{\{|X_i| \geq t_k\}} - \sum_{i=1}^{n} \mathbb{E}|X_i| \mathbf{1}_{\{|X_i| \geq kt_k\}} \geq \frac{2}{3} kt_k.$$

Therefore

$$\mathbb{E}Y \mathbf{1}_{\{Y \geq kt_k/2\}} \geq \mathbb{E}Y - \frac{1}{2} kt_k \geq \frac{1}{6} kt_k.$$

This applied to (15) with $l = (12+24\alpha)k$ gives us $\mathbb{P}(Y \geq kt_k/2) \geq (144+288\alpha)^{-1}$ and in consequence

$$\mathbb{E} \max_{|I|=k} \sum_{i \in I} |X_i| \geq \frac{1}{288(1+2\alpha)} kt(k, X).$$

Since $k \mapsto kt(k, X)$ is non-decreasing, in the case $k \leq \lceil 4(1+\alpha) \rceil =: k_0$ we have

$$\begin{aligned} \mathbb{E} \max_{|I|=k} |X_i| \geq \frac{k}{k_0} \mathbb{E} \max_{|I|=k_0} |X_i| &\geq \frac{k}{5+4\alpha} \cdot \frac{1}{288(1+2\alpha)} k_0 t(k_0, X) \\ &\geq \frac{1}{288(5+4\alpha)(1+2\alpha)} kt(k, X). \end{aligned}$$

The last step is to loose the assumption that X_i has no atoms. Note that both assumption (3) and the lower bound depend only on $(|X_i|)_{i=1}^n$, so we may assume that X_i are nonnegative almost surely. Consider $X^\varepsilon := (X_i + \varepsilon Y_i)_{i=1}^n$, where $Y_1, \ldots, Y_n$ are i.i.d.

nonnegative r.v's with $\mathbb{E}Y_i < \infty$ and a density g, independent of X. Then for every $s,t>0$ we have (observe that (3) holds also for $s<0$ or $t<0$).

$$\begin{aligned}\mathbb{P}(X_i^\varepsilon \geq s, X_j^\varepsilon \geq t) &= \int_0^\infty \int_0^\infty \mathbb{P}(X_i+\varepsilon y_i \geq s,\ X_j+\varepsilon y_j \geq t) g(y_i)g(y_j)dy_idy_j \\ &\overset{(3)}{\leq} \alpha \int_0^\infty \int_0^\infty \mathbb{P}(X_i \geq s-\varepsilon y_i)\mathbb{P}(X_j \geq t-\varepsilon y_j) g(y_i)g(y_j)dy_idy_j \\ &= \alpha \mathbb{P}(X_i^\varepsilon \geq s)\mathbb{P}(X_j^\varepsilon \geq t).\end{aligned}$$

Thus X^ε satisfies assumption (3) and has the density function for every $\varepsilon>0$. Therefore for all natural k we have

$$\mathbb{E}\max_{|I|=k}\sum_{i=1}^n X_i^\varepsilon \geq c(\alpha)kt(k,X^\varepsilon) \geq c(\alpha)kt(k,X).$$

Clearly, $\mathbb{E}\max_{|I|=k}\sum_{i=1}^n X_i^\varepsilon \to \mathbb{E}\max_{|I|=k}\sum_{i=1}^n X_i$ as $\varepsilon \to 0$, so the lower bound holds in the case of arbitrary X satisfying (3). □

We may use Theorem 2 to obtain a comparison of weak and strong moments for the supremum norm:

Corollary 13. *Let X be an n-dimensional centered random vector satisfying condition* (3)*. Assume that*

$$\|X_i\|_{2p} \leq \beta \|X_i\|_p \qquad \text{for every } p \geq 2 \text{ and } i=1,\ldots,n. \tag{17}$$

Then the following comparison of weak and strong moments for the supremum norm holds: for all $a \in \mathbb{R}^n$ and all $p \geq 1$,

$$\big(\mathbb{E}\max_{i\leq n}|a_iX_i|^p\big)^{1/p} \leq C(\alpha,\beta)\Big[\mathbb{E}\max_{i\leq n}|a_iX_i| + \max_{i\leq n}\big(\mathbb{E}|a_iX_i|^p\big)^{1/p}\Big],$$

where $C(\alpha,\beta)$ is a constant depending only on α and β.

Proof. Let $X'=(X_i')_{i\leq n}$ be a decoupled version of X. For any $p>0$ a random vector $(|a_iX_i|^p)_{i\leq n}$ satisfies condition (3), so by Theorem 2

$$\big(\mathbb{E}\max_{i\leq n}|a_iX_i|^p\big)^{1/p} \sim \big(\mathbb{E}\max_{i\leq n}|a_iX_i'|^p\big)^{1/p}$$

for all $p>0$, up to a constant depending only on α. The coordinates of X' are independent and satisfy condition (17), so due to [11, Theorem 1.1] the comparison of weak and strong moments of X' holds, i.e. for $p\geq 1$,

$$\big(\mathbb{E}\max_{i\leq n}|a_iX_i'|^p\big)^{1/p} \leq C(\beta)\Big[\mathbb{E}\max_{i\leq n}|a_iX_i'| + \max_{i\leq n}\big(\mathbb{E}|a_iX_i'|^p\big)^{1/p}\Big],$$

where $C(\beta)$ depends only on β. These two observations yield the assertion. □

5. Lower estimates for order statistics

The next lemma shows the relation between $t(k,X)$ and $t^*(k,X)$ for log-concave vectors X.

Lemma 14. *Let X be a symmetric log-concave random vector in $\mathbb{R}^n$. For any $1 \le k \le n$ we have*

$$\frac{1}{3}\left(t^*(k,X) + \frac{1}{k}\max_{|I|=k}\sum_{i\in I}\mathbb{E}|X_i|\right) \le t(k,X) \le 4\left(t^*(k,X) + \frac{1}{k}\max_{|I|=k}\sum_{i\in I}\mathbb{E}|X_i|\right).$$

Proof. Let $t_k := t(k,X)$ and $t_k^* := t^*(k,X)$. We may assume that any X_i is not identically equal to 0. Then $\sum_{i=1}^n \mathbb{P}(|X_i| \ge t_k^*) = k$ and $\sum_{i=1}^n \mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\ge t_k\}} = kt_k$.

Obviously $t_k^* \le t_k$. Also for any $|I| = k$ we have

$$\sum_{i\in I}\mathbb{E}|X_i| \le \sum_{i\in I}\left(t_k + \mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\ge t_k\}}\right) \le |I|t_k + kt_k = 2kt_k.$$

To prove the upper bound set

$$I_1 := \{i \in [n]\colon\ \mathbb{P}(|X_i| \ge t_k^*) \ge 1/4\}.$$

We have

$$k \ge \sum_{i\in|I_1|}\mathbb{P}(|X_i| \ge t_k^*) \ge \frac{1}{4}|I_1|,$$

so $|I_1| \le 4k$. Hence

$$\sum_{i\in I_1}\mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\ge t_k^*\}} \le \sum_{i\in I_1}\mathbb{E}|X_i| \le 4\max_{|I|=k}\sum_{i\in I}\mathbb{E}|X_i|.$$

Moreover by the second part of Lemma 12 we get

$$\mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\ge t_k^*\}} \le 4t_k^*\mathbb{P}(|X_i| \ge t_k^*) \quad \text{for } i \notin I_1,$$

so

$$\sum_{i\notin I_1}\mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\ge t_k^*\}} \le 4t_k^*\sum_{i=1}^n\mathbb{P}(|X_i| \ge t_k^*) \le 4kt_k^*.$$

Hence if $s = 4t_k^* + \frac{4}{k}\max_{|I|=k}\sum_{i\in I}\mathbb{E}|X_i|$ then

$$\sum_{i=1}^n\mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\ge s\}} \le \sum_{i=1}^n\mathbb{E}|X_i|\mathbf{1}_{\{|X_i|\ge t_k^*\}} \le 4\max_{|I|=k}\sum_{i\in I}\mathbb{E}|X_i| + 4kt_k^* = ks,$$

that is $t_k \le s$. □

To derive bounds for order statistics we will also need a few facts about log-concave vectors.

Lemma 15. *Assume that Z is an isotropic one- or two-dimensional log-concave random vector with a density g. Then $g(t) \le C$ for all t. If Z is one-dimensional, then also $g(t) \ge c$ for all $|t| \le t_0$, where $t_0 > 0$ is an absolute constant.*

Proof. We will use a classical result (see [4, Theorem 2.2.2, Proposition 3.3.1, Proposition 3.3.2, and Proposition 2.5.9]): $\|g\|_{\sup} \sim g(0) \sim 1$ (note that here we use the assumption that Z is isotropic, in particular that $\mathbb{E}Z = 0$, and that the dimension of Z is 1 or 2). This implies the upper bound on g.

In order to get the lower bound in the one-dimensional case, it suffices to prove that $g(u) \geq c$ for $|u| = \varepsilon\mathbb{E}|Z| \geq (2C_1)^{-1}\varepsilon$, where $1/4 > \varepsilon > 0$ is fixed and its value will be chosen later (then by the log-concavity we get $g(u)^s g(0)^{1-s} \leq g(su)$ for all $s \in (0,1)$). Since $-Z$ is again isotropic we may assume that $u \geq 0$.

If $g(u) \geq g(0)/e$, then we are done. Otherwise by log-concavity of g we get

$$\mathbb{P}(Z \geq u) = \int_u^\infty g(s)ds \leq \int_u^\infty g(u)^{s/u} g(0)^{-s/u+1} ds \leq g(0) \int_u^\infty e^{-s/u} ds \leq C_0 u \leq C_0 \varepsilon.$$

On the other hand, Z has mean zero, so $\mathbb{E}|Z| = 2\mathbb{E}Z_+$ and by the Paley–Zygmund inequality and (7) we have

$$\mathbb{P}(Z \geq u) = \mathbb{P}(Z_+ \geq 2\varepsilon\mathbb{E}Z_+) \geq (1-2\varepsilon)^2 \frac{(\mathbb{E}Z_+)^2}{\mathbb{E}Z_+^2} \geq \frac{1}{16} \frac{(\mathbb{E}|Z|)^2}{\mathbb{E}Z^2} \geq c_0.$$

For $\varepsilon < c_0/C_0$ we get a contradiction. □

Lemma 16. *Let Y be a mean zero log-concave random variable and let $\mathbb{P}(|Y| \geq t) \leq p$ for some $p > 0$. Then*

$$\mathbb{P}\left(|Y| \geq \frac{t}{2}\right) \geq \frac{1}{\sqrt{ep}} \mathbb{P}(|Y| \geq t).$$

Proof. By the Grünbaum inequality (8) we have $\mathbb{P}(Y \geq 0) \geq 1/e$, hence

$$\mathbb{P}\left(Y \geq \frac{t}{2}\right) \geq \sqrt{\mathbb{P}(Y \geq t)\mathbb{P}(Y \geq 0)} \geq \frac{1}{\sqrt{e}}\sqrt{\mathbb{P}(Y \geq t)} \geq \frac{1}{\sqrt{ep}} \mathbb{P}(Y \geq t).$$

Since $-Y$ satisfies the same assumptions as Y we also have

$$\mathbb{P}\left(-Y \geq \frac{t}{2}\right) \geq \frac{1}{\sqrt{ep}} \mathbb{P}(-Y \geq t).$$ □

Lemma 17. *Let Y be a mean zero log-concave random variable and let $\mathbb{P}(|Y| \geq t) \geq p$ for some $p > 0$. Then there exists a universal constant C such that*

$$\mathbb{P}(|Y| \leq \lambda t) \leq \frac{C\lambda}{\sqrt{p}} \mathbb{P}(|Y| \leq t) \quad \textit{for } \lambda \in [0,1].$$

Proof. Without loss of generality we may assume that $\mathbb{E}Y^2 = 1$. Then by Chebyshev's inequality $t \leq p^{-1/2}$. Let g be the density of Y. By Lemma 15 we know that $\|g\|_\infty \leq C$ and $g(t) \geq c$ on $[-t_0, t_0]$, where c, C and $t_0 \in (0,1)$ are universal constants. Thus

$$\mathbb{P}(|Y| \leq t) \geq \mathbb{P}(|Y| \leq t_0\sqrt{p}t) \geq 2ct_0\sqrt{p}t,$$

and

$$\mathbb{P}(|Y| \leq \lambda t) \leq 2\|g\|_\infty \lambda t \leq 2C\lambda t \leq \frac{C\lambda}{ct_0\sqrt{p}} \mathbb{P}(|Y| \leq t).$$ □

Now we are ready to give a proof of the lower bound in Theorem 4. The next proposition is a key part of it.

Proposition 18. *Let X be a mean zero log-concave n-dimensional random vector with uncorrelated coordinates and let $\alpha > 1/4$. Suppose that*

$$\mathbb{P}\big(|X_i| \geq t^*(\alpha, X)\big) \leq \frac{1}{C_3} \quad \text{for all } i.$$

Then

$$\mathbb{P}\Big(\lfloor 4\alpha \rfloor\text{-}\max_i |X_i| \geq \frac{1}{C_4} t^*(\alpha, X)\Big) \geq \frac{3}{4}.$$

Proof. Let $t^* = t^*(\alpha, X)$, $k := \lfloor 4\alpha \rfloor$ and $L = \lfloor \frac{\sqrt{C_3}}{4\sqrt{e}} \rfloor$. We will choose C_3 in such a way that L is large, in particular we may assume that $L \geq 2$. Observe also that $\alpha = \sum_{i=1}^n \mathbb{P}(|X_i| \geq t^*(\alpha, X)) \leq nC_3^{-1}$, thus $Lk \leq C_3^{1/2} e^{-1/2} \alpha \leq e^{-1/2} C_3^{-1/2} n \leq n$ if $C_3 \geq 1 > \frac{1}{e}$. Hence

$$k\text{-}\max_i |X_i| \geq \frac{1}{k(L-1)} \sum_{l=k+1}^{Lk} l\text{-}\max_i |X_i| = \frac{1}{k(L-1)} \Big(\max_{|I|=Lk} \sum_{i \in I} |X_i| - \max_{|I|=k} \sum_{i \in I} |X_i| \Big). \tag{18}$$

Lemma 16 and the definition of $t^*(\alpha, X)$ yield

$$\sum_{i=1}^n \mathbb{P}\left(|X_i| \geq \frac{1}{2} t^*\right) \geq \frac{\sqrt{C_3}}{\sqrt{e}} \alpha \geq Lk.$$

This yields $t(Lk, X) \geq t^*(Lk, X) \geq \frac{t^*}{2}$ and by Theorem 3 we have

$$\mathbb{E} \max_{|I|=Lk} \sum_{i \in I} |X_i| \geq c_1 Lk \frac{t^*}{2}.$$

Since for any norm $\mathbb{P}(\|X\| \leq t\mathbb{E}\|X\|) \leq Ct$ for $t > 0$ (see [10, Corollary 1]) we have

$$\mathbb{P}\left(\max_{|I|=Lk} \sum_{i \in I} |X_i| \geq c_2 Lkt^* \right) \geq \frac{7}{8}. \tag{19}$$

Let X' be an independent copy of X. By the Paley-Zygmund inequality and (7), $\mathbb{P}(|X_i| \geq \frac{1}{2}\mathbb{E}|X_i|) \geq \frac{(\mathbb{E}|X_i|)^2}{4\mathbb{E}|X_i|^2} > \frac{1}{C_3}$ if $C_3 > 16C_1^2$, so $\frac{1}{2}\mathbb{E}|X_i| \leq t^*$. Moreover it is easy to verify that $k = \lfloor 4\alpha \rfloor > \alpha$ for $\alpha > 1/4$, thus $t^*(k, X) \leq t^*(\alpha, X) = t^*$. Hence Proposition 1, Lemma 14, and inequality (10) yield

$$\begin{aligned}
\mathbb{E} \max_{|I|=k} \sum_{i \in I} |X_i| &= \mathbb{E} \max_{|I|=k} \sum_{i \in I} |X_i - \mathbb{E}X_i'| \leq \mathbb{E} \max_{|I|=k} \sum_{i \in I} |X_i - X_i'| \leq \mathbb{E} \max_{|I|=2k} \sum_{i \in I} |X_i - X_i'| \\
&\leq 4kt(2k, X - X') \leq 16k\big(t^*(2k, X - X') + \max_i \mathbb{E}|X_i - X_i'|\big) \\
&\leq 16k\big(2t^*(k, X) + 2\max_i \mathbb{E}|X_i|\big) \leq 96kt^*.
\end{aligned}$$

Therefore

$$\mathbb{P}\left(\max_{|I|=k}\sum_{i\in I}|X_i|\geq 800kt^*\right)\leq\frac{1}{8}. \tag{20}$$

Estimates (18)-(20) yield

$$\mathbb{P}\left(k\text{-}\max_i|X_i|\geq\frac{1}{L-1}(c_2L-800)t^*\right)\geq\frac{3}{4},$$

so it is enough to choose C_3 in such a way that $L\geq 1600/c_2$. □

Proof of the first part of Theorem 4. Let $t^*=t^*(k-1/2,X)$ and C_3 be as in Proposition 18. It is enough to consider the case when $t^*>0$, then $\mathbb{P}(|X_i|=t^*)=0$ for all i and $\sum_{i=1}^n\mathbb{P}(|X_i|\geq t^*)=k-1/2$. Define

$$I_1:=\left\{i\leq n\colon\ \mathbb{P}(|X_i|\geq t^*)\leq\frac{1}{C_3}\right\},\quad \alpha:=\sum_{i\in I_1}\mathbb{P}(|X_i|\geq t^*),$$

$$I_2:=\left\{i\leq n\colon\ \mathbb{P}(|X_i|\geq t^*)>\frac{1}{C_3}\right\},\quad \beta:=\sum_{i\in I_2}\mathbb{P}(|X_i|\geq t^*).$$

If $\beta=0$ then $\alpha=k-1/2$, $|I_1|=\{1,\ldots,n\}$, and the assertion immediately follows by Proposition 18 since $4\alpha\geq k$.

Otherwise define

$$\tilde{N}(t):=\sum_{i\in I_2}\mathbf{1}_{\{|X_i|\leq t\}}.$$

We have by Lemma 17 applied with $p=1/C_3$

$$\mathbb{E}\tilde{N}(\lambda t^*)=\sum_{i\in I_2}\mathbb{P}(|X_i|\leq\lambda t^*)\leq C_5\lambda\sum_{i\in I_2}\mathbb{P}(|X_i|\leq t^*)=C_5\lambda(|I_2|-\beta).$$

Thus

$$\mathbb{P}\left(\lceil\beta\rceil\text{-}\max_{i\in I_2}|X_i|\leq\lambda t^*\right)=\mathbb{P}(\tilde{N}(\lambda t^*)\geq|I_2|+1-\lceil\beta\rceil)\leq\frac{1}{|I_2|+1-\lceil\beta\rceil}\mathbb{E}\tilde{N}(\lambda t^*)\leq C_5\lambda.$$

Therefore

$$\mathbb{P}\Big(\lceil\beta\rceil\text{-}\max_{i\in I_2}|X_i|\geq\frac{1}{4C_5}t^*\Big)\geq\frac{3}{4}.$$

If $\alpha<1/2$ then $\lceil\beta\rceil=k$ and the assertion easily follows. Otherwise Proposition 18 yields

$$\mathbb{P}\Big(\lfloor 4\alpha\rfloor\text{-}\max_{i\in I_1}|X_i|\geq\frac{1}{C_4}t^*\Big)\geq\frac{3}{4}.$$

Observe that for $\alpha\geq 1/2$ we have $\lfloor 4\alpha\rfloor+\lceil\beta\rceil\geq 4\alpha-1+\beta\geq\alpha+1/2+\beta=k$, so

$$\mathbb{P}\left(k\text{-}\max_i |X_i| \geq \min\left\{\frac{t^*}{C_4}, \frac{t^*}{4C_5}\right\}\right) \geq \mathbb{P}\left(\lfloor 4\alpha\rfloor\text{-}\max_{i\in I_1}|X_i| \geq \frac{1}{C_4}t^*, \lceil\beta\rceil\text{-}\max_{i\in I_2}|X_i| \geq \frac{1}{4C_5}t^*\right) \geq \frac{1}{2}. \qquad \square$$

Remark 19. *A modification of the proof above shows that under the assumptions of Theorem 4 for any $p < 1$ there exists $c(p) > 0$ such that*

$$\mathbb{P}\left(k\text{-}\max_{i\leq n}|X_i| \geq c(p)t^*(k-1/2, X)\right) \geq p.$$

6. Upper estimates for order statistics

We will need a few more facts concerning log-concave vectors.

Lemma 20. *Suppose that X is a mean zero log-concave random vector with uncorrelated coordinates. Then for any $i \neq j$ and $s > 0$,*

$$\mathbb{P}(|X_i| \leq s, |X_j| \leq s) \leq C_6\mathbb{P}(|X_i| \leq s)\mathbb{P}(|X_j| \leq s).$$

Proof. Let C_7, c_3 and t_0 be the constants from Lemma 15. If $s > t_0\|X_i\|_2$ then, by Lemma 15, $\mathbb{P}(|X_i| \leq s) \geq 2c_3t_0$ and the assertion is obvious (with any $C_6 \geq (2c_3t_0)^{-1}$). Thus we will assume that $s \leq t_0\min\{\|X_i\|_2, \|X_j\|_2\}$.

Let $\widetilde{X}_i = X_i/\|X_i\|_2$ and let g_{ij} be the density of $(\widetilde{X}_i, \widetilde{X}_j)$. By Lemma 15 we know that $\|g_{i,j}\|_\infty \leq C_7$, so

$$\mathbb{P}(|X_i| \leq s, |X_j| \leq s) = \mathbb{P}(|\widetilde{X}_i| \leq s/\|X_i\|_2, |\tilde{X}_j| \leq s/\|X_j\|_2) \leq C_7\frac{s^2}{\|X_i\|_2\|X_j\|_2}.$$

On the other hand the second part of Lemma 15 yields

$$\mathbb{P}(|X_i| \leq s)\mathbb{P}(|X_j| \leq s) \geq \frac{4c_3^2s^2}{\|X_i\|_2\|X_j\|_2}. \qquad \square$$

Lemma 21. *Let Y be a log-concave random variable. Then*

$$\mathbb{P}(|Y| \geq ut) \leq \mathbb{P}(|Y| \geq t)^{(u-1)/2} \quad \text{for } u \geq 1, t \geq 0.$$

Proof. We may assume that Y is non-degenerate (otherwise the statement is obvious), in particular Y has no atoms. Log-concavity of Y yields

$$\mathbb{P}(Y \geq t) \geq \mathbb{P}(Y \geq -t)^{\frac{u-1}{u+1}}\mathbb{P}(Y \geq ut)^{\frac{2}{u+1}}.$$

Hence

$$\mathbb{P}(Y \geq ut) \leq \left(\frac{\mathbb{P}(Y \geq t)}{\mathbb{P}(Y \geq -t)}\right)^{\frac{u+1}{2}}\mathbb{P}(Y \geq -t) = \left(1 - \frac{\mathbb{P}(|Y| \leq t)}{\mathbb{P}(Y \geq -t)}\right)^{\frac{u+1}{2}}\mathbb{P}(Y \geq -t)$$
$$\leq (1 - \mathbb{P}(|Y| \leq t))^{\frac{u+1}{2}}\mathbb{P}(Y \geq -t) = \mathbb{P}(|Y| \geq t)^{\frac{u+1}{2}}\mathbb{P}(Y \geq -t).$$

Since $-Y$ satisfies the same assumptions as Y, we also have

$$\mathbb{P}(Y \le -ut) \le \mathbb{P}(|Y| \ge t)^{\frac{u+1}{2}} \mathbb{P}(Y \le t).$$

Adding both estimates we get

$$\mathbb{P}(|Y| \ge ut) \le \mathbb{P}(|Y| \ge t)^{\frac{u+1}{2}}(1 + \mathbb{P}(|Y| \le t)) = \mathbb{P}(|Y| \ge t)^{\frac{u-1}{2}}(1 - \mathbb{P}(|Y| \le t)^2). \qquad \square$$

Lemma 22. *Suppose that Y is a log-concave random variable and $\mathbb{P}(|Y| \le t) \le \frac{1}{10}$. Then $\mathbb{P}(|Y| \le 21t) \ge 5\mathbb{P}(|Y| \le t)$.*

Proof. Let $\mathbb{P}(|Y| \le t) = p$ then by Lemma 21

$$\mathbb{P}(|Y| \le 21t) = 1 - \mathbb{P}(|Y| > 21t) \ge 1 - \mathbb{P}(|Y| > t)^{10} = 1 - (1-p)^{10} \ge 10p - 45p^2 \ge 5p. \quad \square$$

Let us now prove (4) and see how it implies the second part of Theorem 4. Then we give a proof of (5).

Proof of (4). Fix k and set $t^* := t^*(k-1/2, X)$. Then $\sum_{i=1}^n \mathbb{P}(|X_i| \ge t^*) = k - 1/2$. Define

$$I_1 := \left\{ i \le n\colon\ \mathbb{P}(|X_i| \ge t^*) \le \frac{9}{10} \right\}, \quad \alpha := \sum_{i \in I_1} \mathbb{P}(|X_i| \ge t^*), \tag{21}$$

$$I_2 := \left\{ i \le n\colon\ \mathbb{P}(|X_i| \ge t^*) > \frac{9}{10} \right\}, \quad \beta := \sum_{i \in I_2} \mathbb{P}(|X_i| \ge t^*). \tag{22}$$

Observe that for $u > 3$ and $1 \le l \le |I_1|$ we have by Lemma 21

$$\begin{aligned}\mathbb{P}(l\text{-}\max_{i \in I_1} |X_i| \ge ut^*) &\le \mathbb{E}\frac{1}{l}\sum_{i \in I_1} \mathbf{1}_{\{|X_i| \ge ut^*\}} = \frac{1}{l}\sum_{i \in I_1} \mathbb{P}(|X_i| \ge ut^*) \\ &\le \frac{1}{l}\sum_{i \in I_1} \mathbb{P}(|X_i| \ge t^*)^{(u-1)/2} \le \frac{\alpha}{l}\left(\frac{9}{10}\right)^{(u-3)/2}.\end{aligned} \tag{23}$$

Consider two cases.

Case 1. $\beta > |I_2| - 1/2$. Then $|I_2| < \beta + 1/2 \le k$, so $k - |I_2| \ge 1$ and

$$\alpha = k - \frac{1}{2} - \beta \le k - |I_2|.$$

Therefore by (23)

$$\mathbb{P}\left(k\text{-}\max |X_i| \ge 5t^*\right) \le \mathbb{P}\left((k - |I_2|)\text{-}\max_{i \in I_1} |X_i| \ge 5t^*\right) \le \frac{9}{10}.$$

Case 2. $\beta \le |I_2| - 1/2$. Observe that for any disjoint sets J_1, J_2 and integers l, m such that $l \le |J_1|$, $m \le |J_2|$ we have

$$(l+m-1)\text{-}\max_{i \in J_1 \cup J_2} |x_i| \le \max\left\{ l\text{-}\max_{i \in J_1} |x_i|, m\text{-}\max_{i \in J_2} |x_i| \right\} \le l\text{-}\max_{i \in J_1} |x_i| + m\text{-}\max_{i \in J_2} |x_i|. \tag{24}$$

Since

$$\lceil \alpha \rceil + \lceil \beta \rceil \le \alpha + \beta + 2 < k + 2$$

we have $\lceil\alpha\rceil+\lceil\beta\rceil\le k+1$ and, by (24),

$$k\text{-}\max_i |X_i| \le \lceil\alpha\rceil\text{-}\max_{i\in I_1}|X_i| + \lceil\beta\rceil\text{-}\max_{i\in I_2}|X_i|.$$

Estimate (23) yields

$$\mathbb{P}\left(\lceil\alpha\rceil\text{-}\max_{i\in I_1}|X_i|\ge ut^*\right)\le\left(\frac{9}{10}\right)^{(u-3)/2}\qquad\text{for } u\ge 3.$$

To estimate $\lceil\beta\rceil\text{-}\max_{i\in I_2}|X_i| = (|I_2|+1-\lceil\beta\rceil)\text{-}\min_{i\in I_2}|X_i|$ observe that by Lemma 22, the definition of I_2 and assumptions on β,

$$\sum_{i\in I_2}\mathbb{P}(|X_i|\le 21t^*)\ge 5\sum_{i\in I_2}\mathbb{P}(|X_i|\le t^*) = 5(|I_2|-\beta)\ge 2(|I_2|+1-\lceil\beta\rceil).$$

Set $l := (|I_2|+1-\lceil\beta\rceil)$ and

$$\tilde{N}(t) := \sum_{i\in I_2}\mathbf{1}_{\{|X_i|\le t\}}.$$

Note that we know already that $\mathbb{E}\tilde{N}(21t^*)\ge 2l$. Thus the Paley-Zygmund inequality implies

$$\begin{aligned}\mathbb{P}\left(\lceil\beta\rceil\text{-}\max_{i\in I_2}|X_i|\le 21t^*\right) &= \mathbb{P}\left(l\text{-}\min_{i\in I_2}|X_i|\le 21t^*\right)\ge\mathbb{P}(\tilde{N}(21t^*)\ge l)\\ &\ge\mathbb{P}\left(\tilde{N}(21t^*)\ge\frac{1}{2}\mathbb{E}\tilde{N}(21t^*)\right)\ge\frac{1}{4}\frac{(\mathbb{E}\tilde{N}(21t^*))^2}{\mathbb{E}\tilde{N}(21t^*)^2}.\end{aligned}$$

However Lemma 20 yields

$$\mathbb{E}\tilde{N}(21t^*)^2\le\mathbb{E}\tilde{N}(21t^*)+C_6(\mathbb{E}\tilde{N}(21t^*)))^2\le(C_6+1)(\mathbb{E}\tilde{N}(21t^*))^2.$$

Therefore

$$\begin{aligned}\mathbb{P}\left(k\text{-}\max_i|X_i|>(21+u)t^*\right)&\le\mathbb{P}\left(\lceil\alpha\rceil\text{-}\max_{i\in I_1}|X_i|\ge ut^*\right)+\mathbb{P}\left(\lceil\beta\rceil\text{-}\max_{i\in I_2}|X_i|>21t^*\right)\\ &\le\left(\frac{9}{10}\right)^{(u-3)/2}+1-\frac{1}{4(C_6+1)}\le 1-\frac{1}{5(C_6+1)}\end{aligned}$$

for sufficiently large u. □

The unconditionality assumption plays a crucial role in the proof of the next lemma, which allows to derive the second part of Theorem 4 from estimate (4).

Lemma 23. *Let X be an unconditional log-concave n-dimensional random vector. Then for any $1\le k\le n$,*

$$\mathbb{P}\left(k\text{-}\max_{i\le n}|X_i|\ge ut\right)\le\mathbb{P}\left(k\text{-}\max_{i\le n}|X_i|\ge t\right)^u\qquad\textit{for } u>1, t>0.$$

Proof. Let ν be the law of $(|X_1|,\ldots,|X_n|)$. Then ν is log-concave on $\mathbb{R}_n^+$. Define for $t>0$,

$$A_t := \left\{x\in\mathbb{R}_n^+ :\ k\text{-}\max_{i\le n}|x_i|\ge t\right\}.$$

It is easy to check that $\frac{1}{u}A_{ut}+(1-\frac{1}{u})\mathbb{R}^n_+\subset A_t$, hence

$$\mathbb{P}\left(k\text{-}\max_{i\le n}|X_i|\ge t\right)=\nu(A_t)\ge \nu(A_{ut})^{1/u}\nu(\mathbb{R}^n_+)^{1-1/u}=\mathbb{P}\left(k\text{-}\max_{i\le n}|X_i|\ge ut\right)^{1/u}. \quad \square$$

Proof of the second part of Theorem 4. Estimate (4) together with Lemma 23 yields

$$\mathbb{P}\left(k\text{-}\max_{i\le n}|X_i|\ge Cut^*(k-1/2,X)\right)\le (1-c)^u \quad \text{for } u\ge 1,$$

and the assertion follows by integration by parts. $\square$

Proof of (5). Define I_1, I_2, α and β by (21) and (22), where this time $t^*=t^*(k-k^{5/6}/2,X)$. Estimate (23) is still valid so integration by parts yields

$$\mathbb{E}l\text{-}\max_{i\in I_1}|X_i|\le\left(3+20\frac{\alpha}{l}\right)t^*.$$

Set

$$k_\beta := \left\lceil \beta+\frac{1}{2}k^{5/6}\right\rceil.$$

Observe that

$$\lceil\alpha\rceil+k_\beta<\alpha+\beta+\frac{1}{2}k^{5/6}+2=k+2.$$

Hence $\lceil\alpha\rceil+k_\beta\le k+1$.

If $k_\beta>|I_2|$, then $k-|I_2|\ge\lceil\alpha\rceil+k_\beta-1-|I_2|\ge\lceil\alpha\rceil$, so

$$\mathbb{E}k\text{-}\max_{i}|X_i|\le\mathbb{E}(k-|I_2|)\text{-}\max_{i\in I_1}|X_i|\le\mathbb{E}\lceil\alpha\rceil\text{-}\max_{i\in I_1}|X_i|\le 23t^*.$$

Therefore it suffices to consider case $k_\beta\le|I_2|$ only.

Since $\lceil\alpha\rceil+k_\beta-1\le k$ and $k_\beta\le|I_2|$, we have by (24),

$$\mathbb{E}k\text{-}\max_{i}|X_i|\le\mathbb{E}\lceil\alpha\rceil\text{-}\max_{i\in I_1}|X_i|+\mathbb{E}k_\beta\text{-}\max_{i\in I_2}|X_i|\le 23t^*+\mathbb{E}k_\beta\text{-}\max_{i\in I_2}|X_i|.$$

Since $\beta\le k-\frac{1}{2}k^{5/6}$ and $x\to x-\frac{1}{2}x^{5/6}$ is increasing for $x\ge 1/2$ we have

$$\beta\le\beta+\frac{1}{2}k^{5/6}-\frac{1}{2}\left(\beta+\frac{1}{2}k^{5/6}\right)^{5/6}\le k_\beta-\frac{1}{2}k_\beta^{5/6}.$$

Therefore, considering $(X_i)_{i\in I_2}$ instead of X and k_β instead of k it is enough to show the following claim:
Let $s>0$, $n\ge k$ and let X be an n-dimensional log-concave vector with uncorrelated coordinates. Suppose that

$$\sum_{i\le n}\mathbb{P}(|X_i|\ge s)\le k-\frac{1}{2}k^{5/6} \quad\text{and}\quad \min_{i\le n}\mathbb{P}(|X_i|\ge s)\ge 9/10$$

then

$$\mathbb{E}k\text{-}\max_{i\leq n}|X_i| \leq C_8 s.$$

We will show the claim by induction on k. For $k=1$ the statement is obvious (since the assumptions are contradictory). Suppose now that $k\geq 2$ and the assertion holds for $k-1$.

Case 1. $\mathbb{P}(|X_{i_0}|\geq s)\geq 1-\frac{5}{12}k^{-1/6}$ for some $1\leq i_0\leq n$. Then

$$\sum_{i\neq i_0}\mathbb{P}(|X_i|\geq s)\leq k-\frac{1}{2}k^{5/6}-\left(1-\frac{5}{12}k^{-1/6}\right)\leq k-1-\frac{1}{2}(k-1)^{5/6},$$

where to get the last inequality we used that $x^{5/6}$ is concave on $\mathbb{R}_+$, so $(1-t)^{5/6}\leq 1-\frac{5}{6}t$ for $t=1/k$. Therefore by the induction assumption applied to $(X_i)_{i\neq i_0}$,

$$\mathbb{E}k\text{-}\max_i|X_i|\leq \mathbb{E}(k-1)\text{-}\max_{i\neq i_0}|X_i|\leq C_8 s.$$

Case 2. $\mathbb{P}(|X_i|\leq s)\geq \frac{5}{12}k^{-1/6}$ for all i. Applying Lemma 15 we get

$$\frac{5}{12}k^{-1/6}\leq \mathbb{P}\left(\frac{|X_i|}{\|X_i\|_2}\leq \frac{s}{\|X_i\|_2}\right)\leq C\frac{s}{\|X_i\|_2},$$

so $\max_i\|X_i\|_2\leq Ck^{1/6}s$. Moreover $n\leq \frac{10}{9}k$. Therefore by the result of Lee and Vempala [13] X satisfies the exponential concentration with $\alpha\leq C_9k^{5/12}s$.

Let $l=\lceil k-\frac{1}{2}(k^{5/6}-1)\rceil$ then $s\geq t_*(l-1/2,X)$ and $k-l+1\geq \frac{1}{2}(k^{5/6}-1)\geq \frac{1}{9}k^{5/6}$. Let

$$A:=\left\{x\in\mathbb{R}^n\colon l\text{-}\max_i|x_i|\leq C_{10}s\right\}.$$

By (4) (applied with l instead of k) we have $\mathbb{P}(X\in A)\geq c_4$. Observe that

$$k\text{-}\max_i|x_i|\geq C_{10}s+u\Rightarrow \operatorname{dist}(x,A)\geq\sqrt{k-l+1}u\geq \frac{1}{3}k^{5/12}u.$$

Therefore by Lemma 10 we get

$$\mathbb{P}\left(k\text{-}\max_i|X_i|\geq C_{10}s+3C_9us\right)\leq \exp\left(-(u+\ln c_4)_+\right).$$

Integration by parts yields

$$\mathbb{E}k\text{-}\max_i|X_i|\leq (C_{10}+3C_9(1-\ln c_4))\,s$$

and the induction step is shown in this case provided that $C_8\geq C_{10}+3C_9(1-\ln c_4)$. □

To obtain Corollary 6 we used the following lemma.

Lemma 24. *Assume that X is a symmetric isotropic log-concave vector in $\mathbb{R}^n$. Then*

$$t^*(p,X)\sim\frac{n-p}{n}\quad \text{for } n>p\geq n/4. \tag{25}$$

and

$$t^*(k/2,X)\sim t^*(k,X)\sim t(k,X)\quad \text{for } k\leq n/2. \tag{26}$$

Proof. Observe that

$$\sum_{i=1}^{n} \mathbb{P}(|X_i| \le t^*(p,X)) = n - p.$$

Thus Lemma 15 implies that for $p \ge c_5 n$ (with $c_5 \in (\frac{1}{2}, 1)$) we have $t^*(p,X) \sim \frac{n-p}{n}$. Moreover, by the Markov inequality

$$\sum_{i=1}^{n} \mathbb{P}(|X_i| \ge 4) \le \frac{n}{16},$$

so $t^*(n/4, X) \le 4$. Since $p \mapsto t^*(p,X)$ is non-increasing, we know that $t^*(p,X) \sim 1$ for $n/4 \le p \le c_5 n$.

Now we will prove (26). We have

$$t^*(k,X) \le t^*(k/2,X) \le t(k/2,X) \le 2t(k,X),$$

so it suffices to show that $t^*(k,X) \ge ct(k,X)$. To this end we fix $k \le n/2$. By (25) we know that $t := C_{11} t^*(k,X) \ge C_{11} t^*(n/2,X) \ge e$, so the isotropicity of X and Markov's inequality yield $\mathbb{P}(|X_i| \ge t) \le e^{-2}$ for all i. We may also assume that $t \ge t^*(k,X)$. Integration by parts and Lemma 21 yield

$$\begin{aligned}\mathbb{E}|X_i| \mathbf{1}_{\{|X_i| \ge t\}} &\le 3t\mathbb{P}(|X_i| \ge t) + t\int_0^\infty \mathbb{P}(|X_i| \ge (s+3)t) ds \\ &\le 3t\mathbb{P}(|X_i| \ge t) + t\int_0^\infty \mathbb{P}(|X_i| \ge t) e^{-s} ds \le 4t\mathbb{P}(|X_i| \ge t).\end{aligned}$$

Therefore

$$\sum_{i=1}^{n} \mathbb{E}|X_i| \mathbf{1}_{\{|X_i| \ge t\}} \le 4t \sum_{i=1}^{n} \mathbb{P}(|X_i| \ge t) \le 4t \sum_{i=1}^{n} \mathbb{P}(|X_i| \ge t^*(k,X)) \le 4kt,$$

so $t(k,X) \le 4C_{11} t^*(k,X)$. □

References

[1] S. Artstein-Avidan, A. Giannopoulos, and V. D. Milman, *Asymptotic geometric analysis. Part I*, Mathematical Surveys and Monographs, **202**, American Mathematical Society, Providence, RI, 2015.

[2] N. Balakrishnan, and A. C. Cohen, *Order Statistics and Inference,* Academic Press, New York, 1991.

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

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

[5] H. A. David, and H. N. Nagaraja, *Order Statistics*, 3rd ed. Wiley-Interscience, Hoboken, NJ, 2003.

[6] R. Eldan, *Thin shell implies spectral gap up to polylog via a stochastic localization scheme*, Geom. Funct. Anal. **23** (2013), 532–569.

[7] Y. Gordon, A. Litvak, C. Schütt, and E. Werner, *Orlicz norms of sequences of random variables*, Ann. Probab. **30** (2002), no. 4, 1833–1853.

[8] Y. Gordon, A. Litvak, C. Schütt, and E. Werner, *On the minimum of several random variables*, Proc. Amer. Math. Soc. **134** (2006), no. 12, 3665–3675.

[9] Y. Gordon, A. Litvak, C. Schütt, and E. Werner, *Uniform estimates for order statistics and Orlicz functions*, Positivity **16** (2012), no. 1, 1–28.

[10] R. Latała, *On the equivalence between geometric and arithmetic means for log-concave measures*, Convex geometric analysis (Berkeley, CA, 1996), 123–127, Math. Sci. Res. Inst. Publ. **34**, Cambridge Univ. Press, Cambridge 1999.

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

[12] M. Ledoux, *The concentration of measure phenomenon*, American Mathematical Society, Providence, RI 2001.

[13] Y.T. Lee, and S. Vempala, *Eldan's stochastic localization and the KLS hyperplane conjecture: an improved lower bound for expansion*, 58th Annual IEEE Symposium on Foundations of Computer Science – FOCS 2017, 998–1007, IEEE Computer Soc., Los Alamitos, CA, 2017.

[14] L. Lovász, and S. Vempala, *The geometry of logconcave functions and sampling algorithms*, Proc. of the 44th IEEE Foundations of Computer Science (FOCS '03), Boston, 2003. Random Structures Algorithms **30** (2007), no. 3, 307–358.

[15] J. Prochno, and S. Riemer, *On the maximum of random variables on product spaces*, Houston J. Math. **39** (2013), no. 4, 1301–1311.

Institute of Mathematics, University of Warsaw, Banacha 2, 02–097 Warsaw, Poland.
E-mail address: `rlatala@mimuw.edu.pl, martast@mimuw.edu.pl`