Moment estimates for convex measures

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

Abstract

Let $p \geq 1$, $\varepsilon > 0$, $r \geq (1+\varepsilon)p$, and X be a $(-1/r)$-concave random vector in $\mathbb{R}^n$ with Euclidean norm $|X|$. We prove that

$$(\mathbb{E}|X|^p)^{1/p} \leq c\,(C(\varepsilon)\mathbb{E}|X| + \sigma_p(X))\,,$$

where

$$\sigma_p(X) = \sup_{|z|\leq 1} (\mathbb{E}|\langle z, X\rangle|^p)^{1/p},$$

$C(\varepsilon)$ depends only on ε and c is a universal constant. Moreover, if in addition X is centered then

$$(\mathbb{E}|X|^{-p})^{-1/p} \geq c(\varepsilon)\,(\mathbb{E}|X| - C\sigma_p(X))\,.$$

2010 Math. Subject Classification: Primary: 46B06, 60E15, 60F10; Secondary: 52A23, 52A40

Key Words and Phrases: convex measures, κ-concave measure, tail inequalities, small ball probability estimate.

1 Introduction

Let X be a random vector with values in a finite dimensional Euclidean space E with Euclidean norm $|\cdot|$ and scalar product $\langle \cdot, \cdot \rangle$. For any $p > 0$, we define the weak p-th moment of X by

$$\sigma_p(X) = \sup_{|z|\leq 1} (\mathbb{E}|\langle z, X\rangle|^p)^{1/p}.$$

Clearly $(\mathbb{E}|X|^p)^{1/p} \geq \sigma_p(X)$ and by Hölder's inequality, $(\mathbb{E}|X|^p)^{1/p} \geq \mathbb{E}|X|$. In this paper we are interested in reversed inequalities of the form

$$(\mathbb{E}|X|^p)^{1/p} \leq C_1\, \mathbb{E}|X| + C_2\sigma_p(X) \tag{1.1}$$

for $p \geq 1$ and constants C_1 and C_2.

[*]Research partially supported by MNiSW Grant no. N N201 397437.
[†]Research partially supported by ANR GeMeCoD, ANR 2011 BS01 007 01.
[‡]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.

This is known for some classes of distributions and the question has been studied in a more general setting (see [20] and references therein) and our objective in this paper is to describe new classes for which the relationship (1.1) is satisfied.

Let us recall some known results when (1.1) holds. It clearly holds for Gaussian vectors and it is not difficult to see that (1.1) is true for subgaussian vectors (see below for definitions) for every $p \geq 1$, with C_1 and C_2 depending only on the subgaussian parameter.

Another example of such a class is the class of so-called log-concave vectors. A probability measure μ on $\mathbb{R}^m$ is called log-concave if for all $0 < \theta < 1$ and for all compact subsets $A, B \subset \mathbb{R}^m$ with positive measure one has

$$\mu((1-\theta)A + \theta B) \geq \mu(A)^{1-\theta}\mu(B)^{\theta}. \tag{1.2}$$

A random vector with a log-concave distribution is called log-concave. It is known that for every log-concave random vector X in a finite dimensional Euclidean space and any $p > 0$,

$$(\mathbb{E}|X|^p)^{1/p} \leq C\big(\mathbb{E}|X| + \sigma_p(X)\big),$$

where $C > 0$ is a universal constant. See Corollary 5.3 and references below.

In this paper we will consider the class of convex measures introduced by Borell. Let $\kappa < 0$. A probability measure μ on $\mathbb{R}^m$ is called κ-concave if for all $0 < \theta < 1$ and for all compact subsets $A, B \subset \mathbb{R}^m$ with positive measure one has

$$\mu((1-\theta)A + \theta B) \geq \left((1-\theta)\mu(A)^{\kappa} + \theta\mu(B)^{\kappa}\right)^{1/\kappa}. \tag{1.3}$$

A random vector with a κ-concave distribution is called κ-concave. Note that a log-concave vector is also κ-concave for any $\kappa < 0$.

We show in Theorem 5.2 that for $\kappa > -1$, a κ-concave random vector satisfies (1.1) for all $0 < (1+\varepsilon)p < -1/\kappa$ with C_1 and C_2 depending only on ε.

In fact, in Definition 4.1 we will introduce a general assumption on the distribution, called $H(p, \lambda)$. The main result of the first part of the paper is Theorem 4.2 in which we show that this assumption is sufficient in order to have (1.1). In Theorem 5.1 we prove that convex measures satisfy this assumption.

One of the main applications of the relationship (1.1) consists in tail inequalities for $\mathbb{P}\left(|X| \geq t\,\mathbb{E}|X|\right)$. In Corollary 5.4 we show that for $r > 2$ and for a $(-1/r)$-concave isotropic random vector $X \in \mathbb{R}^n$ the above probability is bounded by $\left(\frac{c\max\{1, r/\sqrt{n}\}}{t}\right)^{r/2}$. From this bound we deduce that the empirical covariance matrix of a sample of size proportional to n is a good approximation of the covariance matrix of X, extending results of [1, 2] from log-concave measures to convex measures. This provides thus a new class of random vectors satisfying such approximation. See Corollary 5.6 and the remark following it.

The second part of the paper deals with negative moments. We are looking for relationship of the form

$$\left(\mathbb{E}|X|^{-p}\right)^{-1/p} \geq C_1\,\mathbb{E}|X| - C_2\sigma_p(X) \tag{1.4}$$

for $p > 0$ and constants C_1 and C_2.

We show in Theorem 6.2 that for $\kappa > -1$, an n-dimensional κ-concave random vector satisfies (1.4) for all $0 < (1+\varepsilon)p < \min\{n/2, (-1/\kappa)\}$ with C_1 and C_2 depending only on ε. As an application we show a small ball probability estimate for κ-concave random vectors. In the log-concave setting it was proved in [28].

2 Preliminaries

The space $\mathbb{R}^m$ is equipped with the scalar product $\langle \cdot, \cdot \rangle$, the Euclidean norm $|\cdot|$, the unit ball B_2^m and the volume measure $\mathrm{vol}(\cdot)$. The canonical basis is denoted by $e_1, e_2, \ldots, e_m$. A gauge or Minkowski functional $\|\cdot\|$ on $\mathbb{R}^m$ is a non-negative function on $\mathbb{R}^m$ satisfying: $\|\lambda x\| = \lambda \|x\|$ and $\|x+y\| \le \|x\| + \|y\|$ for every $x, y \in \mathbb{R}^m$ and every real $\lambda \ge 0$ and such that $\|x\|$=0 if and only if $x = 0$. The dual gauge is defined for every $x \in \mathbb{R}^m$ by $\|x\|^* = \max\{\langle x, t\rangle : \|t\| \le 1\}$. A body is a compact subset of $\mathbb{R}^m$ with a non-empty interior. Any convex body $K \subset \mathbb{R}^m$ containing the origin in its interior defines the gauge by $\|x\| = \inf\{\lambda \ge 0 : x \in \lambda K\}$. It is called the Minkowski functional of K. If $K \subset \mathbb{R}^m$ is a convex body containing the origin in its interior, the polar body K° is defined by $K^\circ = \{x \in \mathbb{R}^m : \langle x, t\rangle \le 1 \text{ for all } t \in K\}$. The diameter of K in the Euclidean metric is denoted by $\mathrm{diam}(K)$.

For a linear subspace $F \subset \mathbb{R}^m$ we denote the orthogonal projection on F by P_F. Note that $P_F K^\circ := P_F(K^\circ) = (K \cap F)^\circ$, when the polar is taken in F.

For a random vector X in $\mathbb{R}^m$ with a density h and a subspace $F \subset \mathbb{R}^m$, we denote the density of $P_F X$ by h_F.

A random vector X in $\mathbb{R}^m$ will be called non-degenerate if it is not supported in a proper affine subspace of $\mathbb{R}^m$. It is called isotropic if it is centered and for all $\theta \in \mathbb{R}^m$, $\mathbb{E}|\langle X, \theta\rangle|^2 = |\theta|^2$.

Given a non-negative bounded function f on $\mathbb{R}^m$ we introduce the following associated set. For any $\alpha \ge 1$, let

$$K_\alpha(f) = \{t \in \mathbb{R}^m : f(t) \ge \alpha^{-m} \|f\|_\infty\}, \tag{2.1}$$

where $\|f\|_\infty = \sup_{t \in \mathbb{R}^m} |f(t)|$.

By g_i, $g_{i,j}$ we denote independent standard Gaussian random variables, i.e. centered and of variance one. A standard Gaussian vector in $\mathbb{R}^m$ is denoted by G, i.e. $G = (g_1, g_2, ..., g_m)$. The standard Gaussian matrix is the matrix whose entries are i.i.d. standard Gaussian variables, i.e. $\Gamma = \{g_{i,j}\}$. By γ_p we denote the L_p norm of g_1. Note that $\gamma_p/\sqrt{p} \to 1/\sqrt{e}$ as $p \to \infty$.

We denote by $\mu_{m,k}$ the Haar probability measure on the Grassmannian $G_{m,k}$ of k-dimensional subspaces of $\mathbb{R}^m$.

Recall that for a real number s, $\lceil s \rceil$ denotes the smallest integer which is not less than s.

By C, C_0, C_1, C_2, ..., c, c_0 c_1, c_2 we denote absolute positive constants, whose values can change from line to line.

For two functions f and h we write $f \sim h$ if there are absolute positive constants c and C such that $cf \le h \le Cf$.

3 Convex probability measures

In this section by a measure we always mean a probability measure.

Let $\kappa \le 1/m$. A Borel measure μ on $\mathbb{R}^m$ is called κ-concave if it satisfies (1.3). When $\kappa = 0$, this inequality should be read as (1.2) and it defines μ as a log-concave measure.

In this paper we will be interested in the case $\kappa \le 0$, which we consider from now on.

The class of κ-concave measures was introduced and studied by Borell. We refer to [10, 11] for a general study and to [9] for more recent development. A κ-concave measure is supported on some convex subset of an affine subspace where it has a density. When the support of κ-concave measure μ generates the whole space, a characterization of Borell ([10, 11]) states that μ is absolutely continuous with respect to the Lebesgue measure and has a density h which is log-concave when $\kappa = 0$ and when $\kappa < 0$, is of the form

$$h = f^{-\beta} \quad \text{with} \quad \beta = m - \frac{1}{\kappa},$$

where $f : \mathbb{R}^m \to (0, \infty]$ is a convex function. The class of m-dimensional κ-concave measures is increasing as κ is decreasing. In particular a log-concave measure is κ-concave for any $\kappa < 0$.

As we mentioned in the Introduction, a random vector with a κ-concave distribution is called κ-concave. Clearly, the linear image of a κ-concave measure is also κ-concave. Recall that any semi-norm of an m-dimensional vector with a κ-concave distribution has moments up to the order $p < -1/\kappa$ (see [10] and Lemmas 7.3 and 7.4 below). Since we are interested in comparison of moments with the moment of order 1, we will always assume that $-1 < \kappa \leq 0$.

4 Strong and weak moments

In this section we consider a random vector X in a finite dimensional Euclidean space E.

Definition 4.1. *Let $p > 0$, $m = \lceil p \rceil$, and $\lambda \geq 1$. We say that a random vector X in E satisfies the assumption $H(p, \lambda)$ if for every linear mapping $A : E \to \mathbb{R}^m$ such that $Y = AX$ is non-degenerate in $\mathbb{R}^m$ there exists a gauge $\| \cdot \|$ on $\mathbb{R}^m$ such that $\mathbb{E}\|Y\| < \infty$ and*

$$(\mathbb{E}\|Y\|^p)^{1/p} \leq \lambda \, \mathbb{E}\|Y\|. \tag{4.1}$$

Remark. Let us give a first example of a random vector satisfying $H(p, \lambda)$. Let X be a random vector in an n-dimensional Euclidean space E, satisfying, for some $\psi \geq 1$,

$$\forall z \in E \quad \forall\, 1 \leq p \leq n \quad (\mathbb{E}|\langle z, X \rangle|^p)^{1/p} \leq \psi \sqrt{p}\, \mathbb{E}|\langle z, X \rangle|. \tag{4.2}$$

Then X satisfies $H(p, C\psi^2)$ for every $p \geq 1$. For example, the standard Gaussian and Rademacher vectors satisfy the above condition with ψ being a numerical constant. Note that one of equivalent definitions of a subgaussian vector says that X is subgaussian if it satisfies (4.2) for every $p \geq 1$.

To prove that (4.2) implies $H(p, C\psi^2)$, let $p > 0$, $m = \lceil p \rceil$ and let $A : E \to \mathbb{R}^m$ be such that $Y = AX$ is non-degenerate. We may assume that $m \geq 2$. Clearly, because of the linear invariance of the property (4.2), we may also assume that $Y = AX$ is isotropic. Thus (4.2) yields,

$$\forall z \in \mathbb{R}^m \quad (\mathbb{E}|\langle z, Y \rangle|^p)^{1/p} \leq \psi \sqrt{p}\, \mathbb{E}|\langle z, Y \rangle| \leq \psi \sqrt{m} |z| \leq \sqrt{2} \psi^2 |z| \mathbb{E}|Y|, \tag{4.3}$$

where the last inequality follows from isotropicity of Y by applying (4.2) with $p = 2$, $z_i = A^\top e_i$, $i \leq m$, and the Cauchy-Schwarz inequality.

Now let us make the following general observation. Let $p \geq 1$ and $m = \lceil p \rceil$. Let Y be a random vector in an m-dimensional normed space with norm $\| \cdot \|$. Since any m-dimensional norm can be estimated, up to a multiplicative constant, by the supremum over an exponential (in m) number of norm one linear forms, we deduce that

$$(\mathbb{E}\|Y\|^p)^{1/p} \leq C' \sup_{\|\varphi\|^* \leq 1} (\mathbb{E}|\varphi(Y)|^p)^{1/p}, \tag{4.4}$$

where C' is a universal constant (see [21] Proposition 3.20). Combining this with (4.3) we conclude that

$$(\mathbb{E}|Y|^p)^{1/p} \le C' \sup_{|z|\le 1} (\mathbb{E}|\langle z, Y\rangle|^p)^{1/p} \le C\, C'\, \psi\, \mathbb{E}|Y|,$$

which shows that X satisfies $H(p, CC'\psi)$.

The main result of this section states a relationship between weak and strong moments under the assumption $H(p,\lambda)$.

Theorem 4.2. *Let $p > 0$ and $\lambda \ge 1$. If a random vector X in a finite dimensional Euclidean space satisfies $H(p,\lambda)$, then*

$$(\mathbb{E}|X|^p)^{1/p} \le c(\lambda\, \mathbb{E}|X| + \sigma_p(X)),$$

where c is a universal constant.

The first step of the proof of Theorem 4.2 consists of showing that there exists some z such that $(\mathbb{E}(\langle z, Y\rangle)_+^p)^{1/p}$ is small, with comparison to $\mathbb{E}|Y|$. This is the purpose of the following lemma.

Lemma 4.3. *Let Y be a random vector in $\mathbb{R}^m$. Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two gauges on $\mathbb{R}^m$ and $\|\cdot\|_1^*$ and $\|\cdot\|_2^*$ be their dual gauges. Then for all $p > 0$,*

$$\min_{\|z\|_2^*=1} (\mathbb{E}(\langle z, Y\rangle)_+^p)^{1/p} \le \frac{(\mathbb{E}\|Y\|_1^p)^{1/p}}{\mathbb{E}\|Y\|_1}\, \mathbb{E}\|Y\|_2.$$

Proof. Let r be the largest real number such that $r\|t\|_1 \le \|t\|_2$ for all $t \in \mathbb{R}^m$. By duality r is the largest number such that $r\|w\|_2^* \le \|w\|_1^*$ for $w \in \mathbb{R}^m$. Pick $z \in \mathbb{R}^m$ such that $\|z\|_2^* = 1$ and $\|z\|_1^* = r$. Then $\langle z, t\rangle \le \|z\|_1^* \|t\|_1 \le r\|t\|_1$ for all $t \in \mathbb{R}^m$. Therefore, for any $p > 0$, $(\mathbb{E}(\langle z, Y\rangle)_+^p)^{1/p} \le r\,(\mathbb{E}\|Y\|_1^p)^{1/p}$. Thus the lemma follows from the inequality $r\,\mathbb{E}\|Y\|_1 \le \mathbb{E}\|Y\|_2$. □

The second step of the proof of Theorem 4.2 is contained in the next lemma.

Lemma 4.4. *Let $n, m \ge 1$ be integers. Let $p \ge 1$. Let X be an n-dimensional random vector and Γ be an $n \times m$ standard Gaussian matrix. Then*

$$(\mathbb{E}|X|^p)^{1/p} \le 2^{1/p}\, \gamma_p^{-1} \left(\mathbb{E} \min_{|t|=1} \|\Gamma t\|_{p,+} + (C\gamma_p + \sqrt{m})\, \sigma_p(X) \right),$$

where $\|z\|_{p,+} = \left(\mathbb{E}\,(\langle z, X\rangle)_+^p\right)^{1/p}$ and C is a universal constant.

Proof. For every $x, y \in \mathbb{R}^n$, $|\,\|x\|_{p,+} - \|y\|_{p,+}\,| \le |x-y|\sigma_p(X)$. The classical Gaussian concentration inequality (see [13] or inequality (2.35) and Proposition 2.18 in [22]) gives that

$$\mathbb{P}\left(\big|\,\|G\|_{p,+} - \mathbb{E}\|G\|_{p,+}\,\big| \ge s\right) \le 2\exp\left(-s^2/2\sigma_p^2(X)\right),$$

and implies (cf. [23], Statement 3.1)

$$(\mathbb{E}\|G\|_{p,+}^p)^{1/p} \le \mathbb{E}\|G\|_{p,+} + C\gamma_p\sigma_p(X), \tag{4.5}$$

where C is a universal constant. Since $\langle G, X\rangle$ has the same distribution as $|X|\,g_1$, we have

$$\mathbb{E}\,(\langle G, X\rangle)_+^p = (1/2)\,\mathbb{E}|\langle G, X\rangle|^p \quad \text{and} \quad \mathbb{E}|\langle G, X\rangle|^p = \gamma_p^p\, \mathbb{E}|X|^p. \tag{4.6}$$

Therefore

$$(\mathbb{E}|X|^p)^{1/p} = 2^{1/p}\,\gamma_p^{-1}\left(\mathbb{E}\|G\|_{p,+}^p\right)^{1/p} \le 2^{1/p}\,\gamma_p^{-1}\big(\mathbb{E}\|G\|_{p,+} + C\gamma_p\sigma_p(X)\big).$$

The Gordon minimax lower bound (see [16], Theorem 2.5) states that for any norm $\|\cdot\|$

$$\mathbb{E}\|G\| \le \mathbb{E}\min_{|t|=1}\|\Gamma t\| + \mathbb{E}|H|\max_{|z|=1}\|z\|,$$

where H is a standard Gaussian vector in $\mathbb{R}^m$. It is easy to check the proof and to show that this inequality remains true when $\|\cdot\|$ is a gauge. This gives us that

$$\mathbb{E}\|G\|_{p,+} \le \mathbb{E}\min_{|t|=1}\|\Gamma t\|_{p,+} + \mathbb{E}|H|\max_{|z|=1}\|z\|_{p,+} \le \mathbb{E}\min_{|t|=1}\|\Gamma t\|_{p,+} + \sqrt{m}\max_{|z|=1}\|z\|_{p,+}$$

and it is enough to observe that $\max_{|z|=1}\|z\|_{p,+} \le \sigma_p(X)$. □

Proof of Theorem 4.2. We may assume that $p \ge 1$ since Theorem 4.2 is obviously true when $0 < p \le 1$. Let m be the integer so that $1 \le p \le m < p+1$, thus $m \le 2p$. We use the notation of Lemma 4.4. We first condition on Γ. We have

$$\|\Gamma z\|_{p,+} = (\mathbb{E}_X(\langle \Gamma z, X\rangle)_+^p)^{1/p} = (\mathbb{E}_X(\langle z, \Gamma^* X\rangle)_+^p)^{1/p}.$$

Let $Y = \Gamma^* X \in \mathbb{R}^m$. If Y is supported by a hyperplane then

$$\min_{|z|=1}(\mathbb{E}_X(\langle z, \Gamma^* X\rangle)_+^p)^{1/p} = 0.$$

Otherwise by our assumption $H(p,\lambda)$ there exists a gauge in $\mathbb{R}^m$ such that

$$(\mathbb{E}\|Y\|^p)^{1/p} \le \lambda\,\mathbb{E}\|Y\|.$$

From Lemma 4.3 we get

$$\min_{|z|=1}(\mathbb{E}_X(\langle z, \Gamma^* X\rangle)_+^p)^{1/p} \le \lambda \mathbb{E}_X|\Gamma^* X|.$$

We now take the expectation with respect to Γ and get

$$\mathbb{E}\min_{|t|=1}\|\Gamma t\|_{p,+} \le \lambda\mathbb{E}|\Gamma^* X| = \lambda\mathbb{E}|H|\mathbb{E}|X| \le \lambda\sqrt{m}\,\mathbb{E}|X|,$$

where H is a standard m-dimensional Gaussian vector. The proof is concluded using Lemma 4.4 and the fact that $\gamma_p^{-1}\sqrt{p}$ is bounded. Indeed,

$$\begin{aligned}(\mathbb{E}|X|^p)^{1/p} &\le 2^{1/p}\gamma_p^{-1}\lambda\sqrt{m}\,\mathbb{E}|X| + 2^{1/p}(C + \gamma_p^{-1}\sqrt{m})\sigma_p(X)\\ &\le c'\left(\lambda\mathbb{E}|X| + \sigma_p(X)\right).\end{aligned}$$

□

5 Tail behavior of convex measures

Theorem 5.1. *Let $n \ge 1$ and $r > 1$. Let X be a centered $(-1/r)$-concave random vector in a finite dimensional Euclidean space. Then for every $0 < p < r$, X satisfies the assumption $H(p, \lambda(p,r))$ with $\lambda(p,r) = c\left(\frac{r}{r-1}\right)^3\left(\frac{r}{r-p}\right)^4$, where c is a universal constant.*

Remark: Note that the parameter $\lambda(p,r)$ in Theorem 5.1 is bounded by a universal constant if the parameters p and r are not close, for instance if $r \geq 2\max\{1,p\}$.

Theorem 5.2. *Let $r > 1$ and let X be a $(-1/r)$-concave random vector in a finite dimensional Euclidean space. Then, for every $0 < p < r$,*

$$(\mathbb{E}|X|^p)^{1/p} \leq c(C_2(p,r)\mathbb{E}|X| + \sigma_p(X)), \tag{5.1}$$

where $C_2(p,r) = c\left(\frac{r}{r-1}\right)^3\left(\frac{r}{r-p}\right)^4$ and c is a universal constant.

Proof. Without loss of generality we assume that $p \geq 1$. The proof may be reduced to the case of a centered random vector. Indeed, let X be a $(-1/r)$-concave random vector, then so is $X - \mathbb{E}X$. Since

$$(\mathbb{E}|X|^p)^{1/p} \leq (\mathbb{E}|X-\mathbb{E}X|^p)^{1/p} + |\mathbb{E}X| \leq (\mathbb{E}|X-\mathbb{E}X|^p)^{1/p} + \mathbb{E}|X|,$$

$\mathbb{E}|X - \mathbb{E}X| \leq 2\mathbb{E}|X|$ and $\sigma_p(X - \mathbb{E}X) \leq 2\sigma_p(X)$, we may assume that X is centered. The theorem now follows immediately by Theorems 4.2 and 5.1. $\square$

Note that trivially a reverse inequality to (5.1) is valid, for $p \geq 1$:

$$2(\mathbb{E}|X|^p)^{1/p} \geq \mathbb{E}|X| + \sigma_p(X).$$

Therefore Theorem 5.2 states an equivalence

$$(\mathbb{E}|X|^p)^{1/p} \sim_{C_2(p,r)} \mathbb{E}|X| + \sigma_p(X).$$

Since a log-concave measure is κ-concave for any $\kappa < 0$, we obtain

Corollary 5.3. *For any log-concave random vector X in a finite dimensional Euclidean space and any $p > 0$,*

$$(\mathbb{E}|X|^p)^{1/p} \leq C\big(\mathbb{E}|X| + \sigma_p(X)\big),$$

where $C > 0$ is a universal constant.

As formulated here, Corollary 5.3 first appeared as Theorem 2 in [4] (see also [3]). A short proof of this result was given in [5]. It can be deduced directly from Paouris work in [27] (see [5]).

As it was mentioned above, if $X \in E$ is $(-1/r)$-concave then so is $\langle z, X\rangle$ for any $z \in E$. From Lemma 7.3, we have that for any $1 \leq p < r$,

$$(\mathbb{E}|\langle z, X\rangle|^p)^{1/p} \leq C_1(p,r)\,\mathbb{E}|\langle z, X\rangle|, \tag{5.2}$$

where $C_1(p,r)$ is defined in Lemma 7.3. Assume that $r > 2$. Let n be the dimension of E. If moreover X is centered and has the identity as the covariance matrix – such a vector is called an *isotropic random vector* – then one has for any $z \in S^{n-1}$ and any $1 \leq p < r$,

$$(\mathbb{E}|\langle z, X\rangle|^p)^{1/p} \leq C_1(p,r)\,\mathbb{E}|\langle z, X\rangle| \leq C_1(p,r)(\mathbb{E}|\langle z, X\rangle|^2)^{1/2} = C_1(p,r). \tag{5.3}$$

Since in that case, $\mathbb{E}|X| \leq (\mathbb{E}|X|^2)^{1/2} = \sqrt{n}$, it follows from Theorem 5.2 that for any $1 \leq p < r$,

$$(\mathbb{E}|X|^p)^{1/p} \leq c(C_2(p,r)\sqrt{n} + C_1(p,r)). \tag{5.4}$$

Together with Markov's inequality this gives the following Corollary.

Corollary 5.4. *Let $r > 2$ and let $X \in \mathbb{R}^n$ be a $(-1/r)$-concave isotropic random vector. Then for every $t > 0$,*

$$\mathbb{P}\Big(|X| \geq t\sqrt{n}\Big) \leq \left(\frac{c\max\{1, r/\sqrt{n}\}}{t}\right)^{r/2}. \tag{5.5}$$

In particular, if $r \geq 2\sqrt{n}$, then for every $6c \leq t \leq 3cr/\sqrt{n}$,

$$\mathbb{P}\Big(|X| \geq t\sqrt{n}\Big) \leq \exp(-c_0 t\sqrt{n}), \tag{5.6}$$

where c and c_0 are universal positive constants.

Remark. A log-concave measure is $(-1/r)$-concave for every $r > 0$, thus in such a case inequality (5.6) is valid for every $t > c$, which is a result from [27].

Proof of Corollary 5.4. The inequality (5.5) follows by Markov's inequality from inequality (5.4) with $p = r/2$, since $C_2(r/2, r) \leq c$ and $C_1(r/2, r) \leq cr$ for a universal positive c.

To prove the "In particular" part denote $r' = t\sqrt{n}/(3c)$. Note that $r' \geq 2\sqrt{n}$ and that $r' \leq r$. Therefore X is $(-1/r')$-concave as well and we can apply (5.5) with r', obtaining the bound for probability $3^{-r'/2}$, which implies the result. $\square$

We now apply our results to the problem of the approximation of the covariance matrix by the empirical covariance matrix. Recall that for a random vector X the covariance matrix of X is given by $\mathbb{E}XX^\top$. It is equal to the identity operator I if X is isotropic. The empirical covariance matrix of a sample of size N is defined by $\frac{1}{N}\sum_{i=1}^N X_i X_i^\top$, where $X_1, X_2, \ldots, X_N$ are independent copies of X. The main question is how small N can be taken in order that these two matrices are close to each other in the operator norm (clearly, if X is non-degenerate then $N \geq n$ due to the dimensional restrictions and, by the law of large numbers, the empirical covariance matrix tends to the covariance matrix as N grows to infinity). See [1, 2] for references on this question and for corresponding results in the case of log-concave measures. In particular, it was proved there that for $N \geq n$ and log-concave n-dimensional vectors $X_1, \cdots, X_N$ one has

$$\Big\|\frac{1}{N}\sum_{i=1}^N X_i X_i^\top - I\Big\| \leq C\sqrt{\frac{n}{N}}$$

with probability at least $1 - 2\exp(-c\sqrt{n})$, where, as usual, I is the identity operator, $\|\cdot\|$ is the operator norm $\ell_2^n \to \ell_2^n$ and c, C are absolute positive constants.

In [30] (Theorem 1.1), the following condition was introduced: an isotropic random vector $X \in \mathbb{R}^n$ is said to satisfy the *strong regularity assumption* if for some $\eta, C > 0$ and every rank $k \leq n$ orthogonal projection P, one has for every $t \geq C$

$$\mathbb{P}\left(|PX| \geq t\sqrt{k}\right) \leq C\, t^{-2-2\eta} k^{-1-\eta}.$$

We show that an isotropic $(-1/r)$-concave random vector satisfies this assumption. For simplicity we will show this with $\eta = 1$ (one can change η by adjusting constants).

Lemma 5.5. *Let $n \geq 1$, $a > 0$ and $r = \max\{4, 2a\log n\}$. Let $X \in \mathbb{R}^n$ be an isotropic $(-1/r)$-concave random vector. Then there exists an absolute constant*

C such that for every rank k orthogonal projection P and every $t \geq C_1(a)$, one has

$$\mathbb{P}\left(|PX| \geq t\sqrt{k}\right) \leq C_2(a)\, t^{-4}k^{-2},$$

where $C_1(a) = C \exp(4/a)$ and $C_2(a) = C \max\{(a \log a)^4, \exp(32/a)\}$.

Proof. Let P be a projection of rank k. Let c be the constant from Corollary 5.4 (without loss of generality we assume $c \geq 1$) and $t > c$. If $r \leq \sqrt{k}$ then Corollary 5.4 implies that

$$\mathbb{P}\Big(|PX| \geq t\sqrt{k}\Big) \leq \left(\frac{c}{t}\right)^{r/2} \leq \left(\frac{c}{t}\right)^{a\ \log n} = n^{-a\log(t/c)}.$$

If $r > \sqrt{k}$ then X is also $(-1/\sqrt{k})$-concave, hence, assuming $k > \max\{4, 64a^2\log^2(4a)\}$ (so that $\sqrt{k} \geq 2a\log k$) and applying Corollary 5.4 again we obtain

$$\mathbb{P}\Big(|PX| \geq t\sqrt{k}\Big) \leq \left(\frac{c}{t}\right)^{\sqrt{k}/2} \leq \left(\frac{c}{t}\right)^{a\ \log k} = k^{-a\log(t/c)}.$$

Thus in both cases we have

$$\mathbb{P}\Big(|PX| \geq t\sqrt{k}\Big) \leq k^{-a\log(t/c)}.$$

One can check that for $t \geq c^2 \exp(4/a)$ and $k \geq \exp(16/a)$ this implies

$$\mathbb{P}\left(|PX| \geq t\sqrt{k}\right) \leq t^{-4}k^{-2},$$

which proves the desired result for $k > C_a := \max\{64a^2\log^2(4a), \exp(16/a)\}$ and $t \geq c^2\exp(4/a)$.

Assume now that $k \leq C_a$. Then we apply Borell's Lemma – Lemma 7.3 (note that $\mathbb{E}|PX| \leq \sqrt{k}$). We have that for every $t \geq 3$

$$\mathbb{P}\left(|PX| \geq t\sqrt{k}\right) \leq \left(1 + \frac{t}{9r}\right)^{-r}.$$

It is not difficult to see (e.g., by considering cases $t \leq 9r$, $9r < t \leq 18r$ and $t > 18r$) that for $C(a) := 54^4 C_a^2$, $t \geq 3$ and $r \geq 4$, one has

$$\mathbb{P}\left(|PX| \geq t\sqrt{k}\right) \leq C(a)t^{-4}k^{-2}.$$

This completes the proof. □

Theorem 1.1 from [30] and the above lemma immediately imply the following corollary on the approximation of the covariance matrix by the empirical covariance matrix.

Corollary 5.6. *Let $n \geq 1$, $a > 0$ and $r = \max\{4, 2a\log n\}$. Let $X_1, \ldots, X_N$ be independent $(-1/r)$-concave isotropic random vectors in $\mathbb{R}^n$. Then for every $\varepsilon \in (0,1)$ and every $N \geq C(\varepsilon, a)n$, one has*

$$\mathbb{E}\Big\|\frac{1}{N}\sum_{i=1}^{N} X_i X_i^\top - I\Big\| \leq \varepsilon,$$

where $C(\varepsilon, a)$ depends only on a and ε.

Remark. Let $r = 2a\log(2n) > 8$. Applying Corollary 5.4 for independent $(-1/r)$-concave isotropic random vectors $X_1, X_2, \ldots, X_N$ and using results of [25], it can be checked that with large probability

$$\Big\|\frac{1}{N}\sum_{i=1}^{N} X_i X_i^\top - I\Big\| \le C(a)\sqrt{\frac{n}{N}}$$

where $C(a)$ depends only on a. As we mentioned above, this extends the results of [1, 2] on the approximation of the covariance matrix from the log-concave setting to the class of convex measures.

Now we prove Theorem 5.1. We need the following lemma. Recall that K_α was defined by (2.1).

Lemma 5.7. *Let m be an integer. Let $r > 1$ and $0 < p < r$. Let $Y \in \mathbb{R}^m$ be a centered random vector with density $g = f^{-\beta}$ with $\beta = m + r$ and f convex positive. Let $F : \mathbb{R}^m \to \mathbb{R}^+$ be such that for every $t \in \mathbb{R}^m$, $F(2t) \le 2^p F(t)$ and assume that $\mathbb{E}F(Y)$ is finite. Then, there exists a universal constant $c \ge 1$ such that $0 \in K_\alpha(g)$ and*

$$\mathbb{E}F(Y) \le c(p,r)\,\mathbb{E}\left(F(Y)1_{K_\alpha(g)}(Y)\right), \tag{5.7}$$

where $c(p,r) = 1 + \frac{c}{r-p}$ and $\alpha = \left(c\frac{(m+r)^2}{(r-p)(r-1)}\right)^{\frac{m+r}{m}}$.

Proof. Let $\alpha \ge 1$ be specified later. Let $\gamma = \frac{\beta-m-1}{\beta-1}$. From Lemma 7.2 we have $\gamma f(0) \le \min f = \|g\|_\infty^{-1/\beta}$ and by definition, $\min f < \alpha^{-m/(r+m)} f(t)$ when $t \notin K_\alpha(g)$. Using the convexity of f and the last two inequalities we get

$$\forall t \notin K_\alpha(g) \quad g(t/2) \ge g(t)\left(\frac{1}{2} + \frac{1}{2}\gamma^{-1}\alpha^{-m/(r+m)}\right)^{-(r+m)}. \tag{5.8}$$

Let $\delta = \delta(\alpha) := \left(1 + \gamma^{-1}\alpha^{-m/(r+m)}\right)^{r+m}$. The inequality (5.8) can be written

$$\forall t \notin K_\alpha(g) \quad g(t) \le 2^{-r-m}\delta g(t/2).$$

Therefore

$$\mathbb{E}F(Y)1_{K_\alpha(g)^c}(Y) \le 2^{-r-m}\int_{K_\alpha(g)^c} F(t)g\left(\frac{t}{2}\right)\delta\, dt \le 2^{-r}\int_{\mathbb{R}^m} F(2t)g(t)\delta\, dt$$

and from the assumption on F, we get

$$\mathbb{E}F(Y)1_{K_\alpha(g)^c}(Y) \le 2^{p-r}\delta\mathbb{E}F(Y).$$

We conclude that if $2^{p-r}\delta < 1$ then

$$\mathbb{E}F(Y) \le (1 - 2^{p-r}\delta)^{-1}\mathbb{E}\left(F(Y)1_{K_\alpha(g)}(Y)\right).$$

Let

$$\alpha_0 = \left(\left(2^{\frac{r-p}{2(r+m)}} - 1\right)\gamma\right)^{-\frac{r+m}{m}},$$

so that $\delta_0 := \delta(\alpha_0) = 2^{\frac{r-p}{2}}$, then $(1 - 2^{p-r}\delta_0)^{-1} = (1 - 2^{\frac{p-r}{2}})^{-1} \le 1 + \frac{c}{r-p}$ and

$$\alpha_0 \le \alpha = \left(c\frac{(r+m)^2}{(r-p)(r-1)}\right)^{\frac{r+m}{m}},$$

where $c \geq 1$ is a universal constant. This concludes the proof of (5.7).

Clearly $\gamma^{-1}\alpha^{-m/(r+m)} \leq \gamma^{-1}\alpha_0^{-m/(r+m)} < 1$ and recall that $\gamma f(0) \leq \min f$. We deduce that $f(0) \leq \alpha^{\frac{m}{r+m}} \min f$ and thus $0 \in K_\alpha(g)$. $\square$

Remark. An interesting setting for the previous lemma is when r is away from 1, for instance $r \geq 2$, r and m are comparable, and p is proportional to r. In this case γ is bounded by a constant, $c(p,r)$ explodes only when $p \to r$, and α depends only on the ratio r/p.

Proof of Theorem 5.1. Let $1 \leq p < r$ and $m = \lceil p \rceil$. Let $A : E \to \mathbb{R}^m$ be a linear mapping and $Y = AX$ be a centered non-degenerate $(-1/r)$-concave random vector. By Borell's result [10, 11], there exists a positive convex function f such that the distribution of Y has a density of the form $g = f^{-(r+m)}$.

We apply Lemma 5.7 and use the notation of that lemma. Because the class of $(-1/r)$-concave measures increases as the parameter r decreases, we may assume that $r \leq 2p$ (note that $\lambda(p, 2p) \sim \lambda(p,r)$ for $r > 2p$, so we do not loose control of the constant assuming that $r \leq 2p$). Thus $1 \leq p \leq m$ and $r \leq 2m$. We deduce that the parameter α from Lemma 5.7 satisfies

$$\alpha \leq c\left(\frac{r}{r-1}\cdot\frac{r}{r-p}\right)^3,$$

where c is a numerical constant.

Now note that because $g^{-1/(r+m)}$ is convex, $K = K_\alpha(g)$ is a convex body and from Lemma 5.7, it contains 0. Let $\|\cdot\|$ be its Minkowski functional.

We have

$$1 \geq \mathbb{P}(Y \in K) = \int_K g \geq \alpha^{-m}\|g\|_\infty \mathrm{vol}(K),$$

so that

$$\mathbb{P}(\|Y\| \leq 1/(2\alpha)) = \int_{K/2\alpha} g \leq \|g\|_\infty (2\alpha)^{-m}\mathrm{vol}(K) \leq 2^{-m} \leq 1/2,$$

and therefore

$$\mathbb{E}\|Y\| \geq \frac{1}{2\alpha}\mathbb{P}(\|Y\| > 1/(2\alpha)) \geq 1/(4\alpha).$$

Let $F(t) = \|t\|^p$ for $t \in \mathbb{R}^m$. Thus $F(2t) = 2^p F(t)$ and, since $p < r$, $\mathbb{E}F(Y)$ is finite. Hence F satisfies the assumption of Lemma 5.7. Therefore for $c(p,r) = 1 + c/(r-p)$

$$\mathbb{E}\|Y\|^p \leq c(p,r)\ \mathbb{E}\left(\|Y\|^p 1_K(Y)\right) \leq c(p,r). \tag{5.9}$$

We conclude that

$$(\mathbb{E}\|Y\|^p)^{1/p}/\mathbb{E}\|Y\| \leq 4\alpha\, c(p,r)^{1/p} \leq c\left(\frac{r}{r-1}\right)^3\left(\frac{r}{r-p}\right)^4$$

for some numerical constant c. $\square$

Another application of Lemma 5.7, which will be used later, is the following lemma.

Lemma 5.8. *Let $1 \leq p < r$ and $m = \lceil p \rceil$. Let $Y \in \mathbb{R}^m$ be a centered $(-1/r)$-concave random vector with density g. There exists a universal constant c, such that $0 \in K_\alpha(g)$ and*

$$(\mathbb{E}|\langle Y, t\rangle|^p)^{1/p} = \left(\int_{\mathbb{R}^m} |\langle x, t\rangle|^p g(x)dx\right)^{1/p} \leq C_3(p,r)\max_{x\in K_\alpha(g)} |\langle x, t\rangle|, \tag{5.10}$$

where $\alpha = c\left(\frac{r^2}{(r-p)(r-1)}\right)^3$, $C_3(p,r) = \left(1+\frac{c}{r-p}\right)^{1/p}$, *and* $c>0$ *is a universal constant.*

Proof. Repeating the argument leading to (5.9) with the function $F(t) = |\langle x,t\rangle|^p$ we obtain that $0 \in K_\alpha(g)$ and

$$\left(\int_{\mathbb{R}^m} |\langle x,t\rangle|^p g(x)dx\right)^{1/p} \leq \left(1+\frac{c}{r-p}\right)^{1/p} \left(\int_{K_\alpha(g)} |\langle x,t\rangle|^p g(x)dx\right)^{1/p}.$$

Clearly

$$\left(\int_{K_\alpha(g)} |\langle x,t\rangle|^p g(x)dx\right)^{1/p} \leq \max_{x\in K_\alpha(g)} |\langle x,t\rangle| \left(\int_{\mathbb{R}^m} g(x)dx\right)^{1/p} = \max_{x\in K_\alpha(g)} |\langle x,t\rangle|,$$

which implies the result. $\square$

6 Small ball probability estimates

The following result was proved in [28].

Theorem 6.1. *Let X be a centered log-concave random vector in a finite dimensional Euclidean space. For every $\varepsilon \in (0,c')$ one has*

$$\mathbb{P}\left(|X| \leq \varepsilon (\mathbb{E}|X|^2)^{1/2}\right) \leq \varepsilon^{c(\mathbb{E}|X|^2)^{1/2}/\sigma_2(X)},$$

where $c, c' > 0$ are universal positive constants.

In this section we generalize this result to the setting of convex distributions. We first establish a lower bound for the negative moment of the Euclidean norm of a convex random vector.

Theorem 6.2. *Let $r>1$ and let X be a centered n-dimensional $(-1/r)$-concave random vector. Assume $1 \leq p < \min\{r, n/2\}$. Then*

$$\left(\mathbb{E}|X|^{-p}\right)^{-1/p} \geq C_4(p,r)\left(\mathbb{E}|X| - C\sigma_p(X)\right),$$

where

$$C_4(p,r) = c\left(\frac{r^2}{(r-p)(r-1)}\right)^{-3}\left(1+\frac{c}{r-p}\right)^{-1/p}$$

and c, C are absolute positive constants. Moreover, if $0<p<1$ then

$$\left(\mathbb{E}|X|^{-p}\right)^{-1/p} \geq c_0\,(1-p)\,\frac{r-1}{r}\,\mathbb{E}|X|,$$

where c_0 is an absolute positive constant.

From Markov's inequality we deduce a small ball probability estimates for convex measures.

Theorem 6.3. *Let $n \geq 1$ and $r>1$. Let X be a centered n-dimensional $(-1/r)$-concave random vector. Assume $1 \leq p < \min\{r, n/2\}$. Then, for every $\varepsilon \in (0,1)$,*

$$\mathbb{P}\left(|X| \leq \varepsilon \mathbb{E}|X|\right) \leq \left(2C_4^{-1}(p,r)\varepsilon\right)^p,$$

whenever $\mathbb{E}|X| \geq 2C\sigma_p(X)$, where c, C and $C_4(p,r)$ are the constants from Theorem 6.2.

Remark. Theorem 6.3 implies Theorem 6.1 proved in [28]. Indeed, let $p \geq 1$, $r \geq \max\{3, 2p\}$ and $A := (\mathbb{E}|X|^2)^{1/2}/\sigma_2(X)$ (note that $A \leq \sqrt{n}$). By Lemma 7.3,

$$\sigma_p(X) \leq C_1(p,r)\sigma_1(X) \leq c_0 p \sigma_2(X) \quad \text{and} \quad (\mathbb{E}|X|^2)^{1/2} \leq c_1 \mathbb{E}|X|.$$

Thus $\mathbb{E}|X|/\sigma_p(X) \geq c_2 A/p$. If $c_2 A/(2C) \geq 1$, we chose $p = c_2 A/(2C)$ and apply Theorem 6.3. Since

$$\mathbb{E}|X|/\sigma_p(X) \geq c_2 A/p \geq 2C,$$

Theorem 6.1 follows. Now assume that $A \leq 2C/c_2$. Then Theorem 6.1 follows from Lemma 7.5 and Lemma 7.4 (with $q = 2$), which for a log-concave random vector X states that $\left(\mathbb{E}|X|^2\right)^{1/2} \leq c\,\mathrm{Med}(|X|)$, where c is a numerical constant and $\mathrm{Med}(|X|)$ is a median of $|X|$.

We need the following result from [19] (Theorem 1.3 there).

Theorem 6.4. *Let $n \geq 1$ be an integer, $\|\cdot\|$ be a norm in $\mathbb{R}^n$, K be its unit ball and $\sigma := \max_{|t|=1} \|t\|$. Assume that $0 < p \leq c_0 \left(\mathbb{E}\|G\|/\sigma\right)^2$ and $m = \lceil p \rceil$. Then*

$$\frac{c\,\mathbb{E}\|G\|}{\sqrt{n}} \leq \left(\int_{G_{n,m}} (\mathrm{diam}(K \cap F))^m \, d\mu(F)\right)^{-1/m} \leq \frac{\mathbb{E}\|G\|}{c\sqrt{n}},$$

where $\mu = \mu_{n,m}$ and c is an absolute positive constant.

The proof of Theorem 6.2 is based on the following two lemmas.

Lemma 6.5. *Let $m \leq n$, $\alpha > 0$ and X be a random vector in $\mathbb{R}^n$ with density g. Then,*

$$(\mathbb{E}|X|^{-m})^{-1/m} \geq \frac{1}{\sqrt{2\pi}\alpha} (\mathbb{E}|G|^{-m})^{-1/m} \left(\int_{G_{n,m}} (\mathrm{vol}(K_\alpha(g_F)))^{-1} \, d\mu(F)\right)^{-1/m}.$$

Proof. Integrating in polar coordinates (see [28], Proposition 4.6), we obtain the following key formula

$$(\mathbb{E}|X|^{-m})^{-1/m} = (2\pi)^{-1/2} (\mathbb{E}|G|^{-m})^{-1/m} \left(\int_{G_{n,m}} g_F(0) d\mu(F)\right)^{-1/m}.$$

Note that

$$1 = \int_F g_F(x)dx \geq \int_{K_\alpha(g_F)} g_F(x)dx \geq \alpha^{-m} \|g_F\|_\infty \mathrm{vol}(K_\alpha(g_F)).$$

This implies the result, since $g_F(0) \leq \|g_F\|_\infty$. □

Below we will use the following notation. For a random vector X in $\mathbb{R}^n$, $p > 0$, and $t \in \mathbb{R}^n$ we denote

$$\|t\|_p = (\mathbb{E}|\langle X, t\rangle|^p)^{1/p}$$

(note that it is the dual gauge of the so-called centroid body, which is rather an L_p-norm than the ℓ_p-norm).

Lemma 6.6. *Let $1 \le p < r$ and $m = \lceil p \rceil$. Let X be a centered $(-1/r)$-concave random vector in $\mathbb{R}^n$ with density g. Let K denote the unit ball of $\| \cdot \|_p$. Then for every m-dimensional subspace $F \subset \mathbb{R}^n$ one has*

$$(\mathrm{vol}(P_F K^\circ))^{1/m} \le 4C_3(p,r)(\mathrm{vol}(K_\alpha(g_F)))^{1/m},$$

where $\alpha = c\left(\frac{r^2}{(r-p)(r-1)}\right)^3$, $C_3(p,r) = \left(1 + \frac{c}{r-p}\right)^{1/p}$, and $c > 0$ is a universal constant.

Proof. Applying Lemma 5.8 to $Y = P_F X$, we obtain that for every $t \in F$

$$||t||_p \le C_3(p,r) \max_{x \in K_\alpha(g_F)} |\langle x, t\rangle|$$

with α and $C_3(p,r)$ given in Lemma 5.8. Since for $t \in F$, $||t||_p = \max\langle x, t\rangle$, where the supremum is taken over $x \in (K \cap F)^\circ = P_F K^\circ$, this is equivalent to

$$P_F K^\circ \subset C_3(p,r)\mathrm{conv}(K_\alpha(g_F) \cup -K_\alpha(g_F)).$$

Lemma 5.8 also claims that $0 \in K_\alpha(g_F)$, thus

$$\mathrm{conv}(K_\alpha(g_F) \cup -K_\alpha(g_F)) \subset K_\alpha(g_F) - K_\alpha(g_F).$$

By Rogers-Sheppard inequality [29] we observe

$$(\mathrm{vol}(P_F K^\circ))^{1/m} \le \binom{2m}{m}^{1/m} C_3(p,r)(\mathrm{vol}(K_\alpha(g_F)))^{1/m}.$$

This implies the result. □

Proof of Theorem 6.2. Recall that c_1, c_2, ... denote absolute positive constants. Recall also that for a random vector X in $\mathbb{R}^n$, $p > 0$, and $t \in \mathbb{R}^n$

$$\|t\|_p = (\mathbb{E}|\langle X, t\rangle|^p)^{1/p},$$

and, given a norm $\| \cdot \|$ on $\mathbb{R}^n$, $\sigma = \sigma(\| \cdot \|) = \max_{|t|=1} \|t\|$. In particular,

$$\sigma(\| \cdot \|_p) = \sigma_p(X).$$

Finally, let K denote the unit ball of $\| \cdot \|_p$.

We assume that $p \ge 1$, X is non-degenerate in $\mathbb{R}^n$ and let $m = \lceil p \rceil$. Without loss of generality we assume that

$$\mathbb{E}|X| \ge C\sigma_p(X),$$

where C is a large enough absolute constant.

As in (4.5), since $p \le m \le 2p$, we have

$$\mathbb{E}|X| \le (\mathbb{E}|X|^p)^{1/p} = \gamma_p^{-1}(\mathbb{E}\|G\|_p^p)^{1/p} \le \gamma_p^{-1}(\mathbb{E}\|G\|_p + c_1\gamma_p\sigma_p(X))$$
$$\le c_2(\mathbb{E}\|G\|_p/\sqrt{m} + \sigma_p(X)).$$

Hence

$$\mathbb{E}\|G\|_p \ge \sqrt{p}c_2^{-1}\left(\mathbb{E}|X| - c_2\sigma_p(X)\right) \ge \sqrt{p}(c_2)^{-1}(C - c_2)\sigma_p(X). \tag{6.1}$$

This implies that for sufficiently large C we have $m \le 2p \le c_0(\mathbb{E}\|G\|_p/\sigma_p(X))^2$, where c_0 is the constant from Theorem 6.4.

Note that $(\mathbb{E}|G|^{-p})^{-1/p} \geq (\mathbb{E}|G|^{-m})^{-1/m} \geq c_3\sqrt{n}$ (the second inequality is well known for $m \leq n/2$ and can be directly computed). Combining Lemmas 6.5 and 6.6, we obtain

$$(\mathbb{E}|X|^{-m})^{-1/m} \geq \frac{c_4\sqrt{n}}{\alpha\, C_3(p,r)} \left(\int_{G_{n,m}} (\mathrm{vol}(P_F K^\circ))^{-1} d\mu(F)\right)^{-1/m},$$

with α and $C_3(p,r)$ as in Lemma 6.6.

Now note that $P_F K^\circ = (K \cap F)^\circ \supset (\mathrm{diam}(K \cap F))^{-1} B_2^n \cap F$. Therefore $1/\mathrm{vol}(P_F K^\circ) \leq (c_5\sqrt{m}\,\mathrm{diam}(K \cap F))^m$. Applying Theorem 6.4, we obtain

$$(\mathbb{E}|X|^{-m})^{-1/m} \geq \frac{c_6}{\alpha\, C_3(p,r)\sqrt{m}}\, \mathbb{E}\|G\|_p.$$

Applying the first inequality from (6.1), we obtain the desired result.

The "Moreover" part is an immediate corollary of Lemmas 7.4 (with $q = 1$) and 7.5. □

Conjecture 6.7. *We conjecture that for convex distributions a similar thin shell property holds as for log-concave distribution: if X is an isotropic $(-1/r)$-concave random vector in $\mathbb{R}^n$ with $r > 2$, then*

$$\forall t \in (0,1) \quad \mathbb{P}\left(\big||X| - \mathbb{E}|X|\big| \geq t\sqrt{n}\right) \to 0.$$

as n tends to ∞. See [18] for recent work in the log-concave setting.

7 Appendix

There is a vast literature on inequalities of integrals related to concave functions. Some of the following lemmas may be known but we did not find any reference. Their proofs use classical methods for demonstrating integral inequalities involving concave functions (see [12] and [26]). The first lemma is a mirror image for negative moments of a result from [24] valid for positive moments.

Lemma 7.1. *Let $s, m, \beta \in \mathbb{R}$ such that $\beta > m + 1 > 0$ and $s > 0$. Let φ be a non-negative concave function on $[s, +\infty)$. Then*

$$G(\beta) = \frac{\int_s^\infty \varphi^m(x) x^{-\beta}\, dx}{s^{m-\beta+1} B(m+1, \beta - m - 1)}$$

is an increasing function of β on $(m+1, \infty)$. Here $B(u,v) = \int_0^1 (1-t)^{u-1} t^{v-1}\, dt$ denotes the Beta function.

Proof. Let $\beta > m + 1$. Consider the function

$$H(t) = \int_s^t \varphi^m(x) x^{-\beta}\, dx - \int_s^t a^m (x-s)^m x^{-\beta}\, dx$$

for $t \geq s$, where a is chosen so that $H(\infty) = 0$. Note that H', the derivative of H, has the same sign as $(\varphi(x)/(x-s))^m - a^m$. Since $\varphi(x)/(x-s)$ is decreasing on $(s, +\infty)$, we deduce that H is first increasing and then decreasing. Since $H(s) = H(\infty) = 0$ we conclude that H is non-negative. This means that for every $t \geq s$,

$$\int_s^t \varphi^m(x) x^{-\beta}\, dx \geq \int_s^t a^m (x-s)^m x^{-\beta}\, dx. \tag{7.1}$$

Now, note that for any $\beta' > \beta$ and any non-negative function F, we have by Fubini's theorem,

$$\int_s^\infty F(x)x^{-\beta'}\,dx = \int_s^\infty (\beta'-\beta)t^{-\beta'+\beta-1}\left(\int_s^t F(x)x^{-\beta}\,dx\right)dt.$$

Using (7.1) and applying this relation to $F = \varphi^m$ and then to $F(x) = a^m(x-s)^m$, we get that

$$\int_s^\infty \varphi^m(x)x^{-\beta'}\,dx \geq a^m \int_s^\infty (x-s)^m x^{-\beta'}\,dx.$$

From the definition of a, we conclude that

$$\int_s^\infty \varphi^m(x)x^{-\beta}\,dx \Big/ \int_s^\infty (x-s)^m x^{-\beta}\,dx$$

is an increasing function of β on $(m+1,\infty)$. The conclusion follows from the computation of $\int_s^\infty (x-s)^m x^{-\beta}\,dx = s^{m-\beta+1}B(m+1,\beta-m-1)$. □

Lemma 7.2. *Let $m \geq 1$ be an integer. Let g be the density of a probability on $\mathbb{R}^m$ of the form $g = f^{-\beta}$ with f positive convex on $\mathbb{R}^m$ and $\beta > m+1$. If $\int xg(x)\,dx = 0$, then*

$$g(0) \geq \left(\frac{\beta-m-1}{\beta-1}\right)^\beta \|g\|_\infty.$$

Proof. Since f is convex it follows from Jensen's inequality that

$$f(0) = f\left(\int xg(x)\,dx\right) \leq \int f^{-\beta+1}(x)\,dx = \int_s^\infty (\beta-1)h(t)t^{-\beta}\,dt,$$

where $s = \min f = \|g\|_\infty^{-1/\beta}$ and $h(t) = \mathrm{vol}\{f \leq t\}$ denotes the Lebesgue measure of $\{f \leq t\}$. From the convexity of f and from the Brunn-Minkowski inequality, $\varphi = h^{1/m}$ is concave. Thus, using the notation of Lemma 7.1,

$$f(0) \leq (\beta-1)s^{m-\beta+1}B(m+1,\beta-m-1)G(\beta).$$

Now observe that $\int f^{-\beta} = \int_s^\infty \beta\varphi^m(x)x^{-\beta-1}\,dx = 1$ and therefore, by Lemma 7.1,

$$G(\beta) \leq G(\beta+1) = \left(\beta s^{m-\beta}B(m+1,\beta-m)\right)^{-1}.$$

The conclusion follows from combining the last two inequalities. □

Remark. When $\beta \to \infty$, which corresponds to a log-concave density, we recover the inequality from [14] saying that $g(0) \geq e^{-m}\|g\|_\infty$.

The next lemma is a well known result of Borell ([10]) stated in a way that fits our needs and stresses the dependence on the parameter of concavity.

Lemma 7.3. *Let $r > 1$ and X be a $(-1/r)$-concave random vector in $\mathbb{R}^m$. Then for any semi-norm $\|\cdot\|$ and any $t \geq 1$, one has*

$$\mathbb{P}(\|X\| \geq 3t\mathbb{E}\|X\|) \leq \left(1+\frac{t}{3r}\right)^{-r}.$$

As a consequence, for every $1 \leq p < r$,

$$(\mathbb{E}\|X\|^p)^{1/p} \leq C_1(p,r)\,\mathbb{E}\|X\|,$$

where $C_1(p,r) = cp$ for $r > p+1$, $C_1(p,r) = \frac{c\,r}{(r-p)^{1/p}}$ otherwise and c is a universal constant.

Proof. Denote $\theta := \mathbb{P}(\|X\| \le 3\mathbb{E}\|X\|)$. Assume that $\theta < 1$ (otherwise we are done). From Markov's inequality,

$$\theta = 1 - \mathbb{P}(\|X\| > 3\mathbb{E}\|X\|) \ge 2/3.$$

The subset $B = \{x \in \mathbb{R}^m : \|x\| \le 3\mathbb{E}\|X\|\}$ is symmetric and convex. From Lemma 3.1 in [10], for every $t \ge 1$, one has

$$\mathbb{P}(\|X\| \ge 3t\mathbb{E}\|X\|) \le \left(\frac{t+1}{2}\left((1-\theta)^{-1/r} - \theta^{-1/r}\right) + \theta^{-1/r}\right)^{-r}.$$

Thus,

$$\mathbb{P}(\|X\| \ge 3t\mathbb{E}\|X\|) \le \theta\left(1 + \frac{1}{2r}\log\frac{\theta}{1-\theta} + \frac{t}{2r}\log\frac{\theta}{1-\theta}\right)^{-r}.$$

We deduce that for every $t \ge 1$,

$$\mathbb{P}(\|X\| \ge 3t\mathbb{E}\|X\|) \le \theta\left(1 + \frac{t}{2r}\log\frac{\theta}{1-\theta}\right)^{-r} \le \left(1 + \frac{t}{3r}\right)^{-r}.$$

Integrating, we get

$$\begin{aligned}
\mathbb{E}\|X\|^p/(3\mathbb{E}\|X\|)^p &= \int_0^\infty pt^{p-1}\mathbb{P}(\|X\| \ge 3t\mathbb{E}\|X\|)\,dt \\
&\le 1 + \int_1^\infty pt^{p-1}\left(1 + \frac{t}{3r}\right)^{-r} dt \\
&\le 1 + (3r)^p pB(p, r-p) \\
&= 1 + (3r)^p\Gamma(p+1)\Gamma(r-p)/\Gamma(r).
\end{aligned}$$

Now, if $r > p+1$ then, by Stirling's formula,

$$(\Gamma(p+1)\Gamma(r-p)/\Gamma(r))^{1/p} \sim \frac{p}{r},$$

and if $r \le p+1$ then

$$(\Gamma(p+1)\Gamma(r-p)/\Gamma(r))^{1/p} \sim (\Gamma(r-p))^{1/p} \sim \left(\frac{1}{r-p}\right)^{1/p}.$$

This completes the proof. $\square$

The following stronger variant of Borell's lemma allows us to compare the expectation of a random variable $\|X\|$ and a median $\mathrm{Med}(\|X\|)$. The first part was proved for general functions in [7] (Theorem 1.1 and the discussion following Theorem 5.2, see also Corollary 11 in [15]). It was also implicitly proved in [17] (see inequality (4) in [15]). The second part of the lemma follows by integration (we provide its proof for the sake of completeness).

Lemma 7.4. *Let $r > 1$ and X be a $(-1/r)$-concave random vector in $\mathbb{R}^m$. Then for any semi-norm $\|\cdot\|$ and any $t \ge 1$, one has*

$$\mathbb{P}(\|X\| \ge t\mathrm{Med}(\|X\|)) \le (C_0 r)^r t^{-r},$$

where C_0 is an absolute positive constant. As a consequence, for every $q \in [1, r)$ one has

$$(\mathbb{E}\|X\|^q)^{1/q} \le Cq\left(\frac{r}{r-q}\right)^{1/q}\mathrm{Med}(\|X\|),$$

where C is an absolute positive constant.

Proof. As we mentioned before the lemma, the first part is known. Using it and the distribution formula (and denoting $\mathrm{Med}(\|X\|)$ by M) we observe

$$\begin{aligned}\mathbb{E}\|X\|^q &= \int_0^{(C_0rM)^q} \mathbb{P}(|X|^q > s)\,ds + \int_{(C_0rM)^q}^{\infty} \mathbb{P}(|X|^q > s)\,ds \\ &\le (C_0rM)^q + \int_{C_0rM}^{\infty} qt^{q-1}(C_0rM/t)^r\ dt \\ &= (C_0rM)^q + q(C_0rM)^q\frac{1}{r-q} = (C_0rM)^q\frac{r}{r-q},\end{aligned}$$

which implies

$$(\mathbb{E}\|X\|^q)^{1/q} \le C_0 r\left(\frac{r}{r-q}\right)^{1/q}\mathrm{Med}(\|X\|).$$

Now, if $q \ge r/2$ then this bound is equivalent to the desired one. If $1 \le q < r/2$ we denote $r' := 2q < r$. Then X is also $(-1/r')$-concave and we apply the bound with r' instead of r, obtaining

$$(\mathbb{E}\|X\|^q)^{1/q} \le C_0 r'\left(\frac{r'}{r'-q}\right)^{1/q}\mathrm{Med}(\|X\|) \le 4C_0 q\,\mathrm{Med}(\|X\|),$$

which completes the proof. □

Remarks. 1. In fact we have slightly better tail bound from inequality (17) in [15] (or from Corollary 11 there, note that parameters d_f and A_f from this Corollary in our setting are $d_f = 1$ and $A_f = 2$). Namely, for all $t \ge 1$ one has

$$\mathbb{P}\left(\|X\| \ge 2t\mathrm{Med}(\|X\|)\right) \le \left(1 + \frac{t\log 2}{r}\right)^{-r}.$$

This can be also used to obtain the upper bound without considering the cases $q < r/2$ and $q \ge r/2$.

2. Slightly changing the proof one can also obtain the bound $(1/(r-q))^{1/r}$, which is equivalent to $(r/(r-q))^{1/q}$.

3. One can show that estimate of the L_q-norm of X is sharp by considering one-dimensional functions

$$h(x) = e^{|x|/(r+1)}\,1_{|x|\le r+1} + \frac{e|x|}{r+1}\,1_{|x|>r+1} \qquad \text{and } f = Ah^{-r-1},$$

where the constant A is chosen so that f is a density. Then a random variable X with density f is $(-1/r)$-concave since h is convex, $\mathrm{Med}(|X|)$ is uniformly (over $r > 1$) bounded away from 0 and infinity, while the L_q-norm of X is of the order $q(r/(r-q))^{1/q}$.

The last lemma follows from Corollary 7.3 in [8] or Corollary 9 in [15] (as before, the second part follows by integration).

Lemma 7.5. *Let $r > 1$ and X be a $(-1/r)$-concave random vector in $\mathbb{R}^m$. Then for any semi-norm $\|\cdot\|$ and any $\varepsilon \in (0,1)$, one has*

$$\mathbb{P}\left(\|X\| \ \le \varepsilon\mathrm{Med}(\|X\|)\right) \le C_0\varepsilon,$$

where C_0 is an absolute positive constant. As a consequence, for every $p \in (0,1)$,

$$\left(\mathbb{E}\|X\|^{-p}\right)^{-1/p} \ge c(1-p)\mathrm{Med}(\|X\|),$$

where c is an absolute positive constant.

References

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

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

[3] R. Adamczak, R. Latała, A.E. Litvak, A. Pajor and N. Tomczak-Jaegermann, *Tail estimates for norms of sums of log-concave random vectors*, preprint.

[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, **349** (2011), 783–786.

[5] R. Adamczak, R. Latała, A.E. Litvak, K. Oleszkiewicz, A. Pajor and N. Tomczak-Jaegermann, *A short proof of Paouris' inequality*, Can. Math. Bull., to appear.

[6] S.G. Bobkov, *Convex bodies and norms associated to convex measures*, Probab. Theory Related Fields, **147** (2010), 303–332.

[7] S.G. Bobkov, *Large deviations and isoperimetry over convex probability measures with heavy tails*, Electr. J. Probab. **12** (2007), 1072–1100.

[8] S.G. Bobkov and F.L. Nazarov, *Sharp dilation-type inequalities with fixed parameter of convexity*, J. Math. Sci. (N. Y.) 152 (2008), 826–839.

[9] S.G. Bobkov and M. Madiman, *Reverse Brunn-Minkowski and reverse entropy power inequalities for convex measures*, J. Funct. Anal., **262** (2012), 3309–3339.

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

[11] C. Borell, *Convex set functions in d-space*, Period. Math. Hungar., **6** (1975), 111–136.

[12] C. Borell, *Inverse Hölder inequalities in one and several dimensions*, J. Math. Anal. Appl. **41** (1973), 300–312.

[13] B.S. Cirel'son, I.A. Ibragimov, V.N. Sudakov, *Norms of Gaussian sample functions*, Proc. 3rd Japan-USSR Symp. Probab. Theory, Taschkent 1975, Lect. Notes Math. **550** (1976), 20–41.

[14] M. Fradelizi, *Sections of convex bodies through their centroid*, Arch. Math. **69** (1997), 515–522.

[15] M. Fradelizi, *Concentration inequalities for s-concave measures of dilations of Borel sets and applications*, Electron. J. Probab. **14** (2009), no. 71, 2068–2090.

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

[17] O. Guédon, *Kahane-Khinchine type inequalities for negative exponent*, Mathematika **46** (1999), 165–173.

[18] O. Guédon and E. Milman, *Interpolating thin-shell and sharp large-deviation estimates for isotropic log-concave measures*, Geom. Funct. Anal. **21** (2011), 1043–1068.

[19] B. Klartag, R. Vershynin, *Small ball probability and Dvoretzky theorem*, Israel J. Math. **157** (2007), 193–207.

[20] R. Latała, *Weak and strong moments of random vectors*, Marcinkiewicz Centenary Volume, Banach Center Publ. 95 (2011), 115–121.

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

[22] M. Ledoux, *The concentration of measure phenomenon*, Mathematical Surveys and Monographs, 89. American Mathematical Society, Providence, RI, 2001.

[23] A.E. Litvak, V.D. Milman and G. Schechtman, *Averages of norms and quasi-norms*, Math. Ann. 312 (1998), 95–124.

[24] A.W. Marshall, I. Olkin and F. Proschan, *Monotonicity of ratios of means and other applications of majorization*, Inequalities edited by O. Shisha, Acad. Press, New York, London (1967), 177–190.

[25] S. Mendelson and G. Paouris, *On the singular values of random matrices*, Preprint.

[26] V.D. Milman and A. Pajor, *Isotropic position and inertia ellipsoids and zonoids of the unit ball of a normed n-dimensional space*, Lectures Notes in Mathematics **1376**, Springer, Berlin, 1989, 107–131.

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

[28] G. Paouris, *Small ball probability estimates for log-concave measures*, Trans. Amer. Math. Soc. (2012), 287–308.

[29] C.A. Rogers and G.C. Shephard, *The difference body of a convex body*, Arch. Math. **8** (1957), 220–233.

[30] N. Srivastava and R. Vershynin, *Covariance estimation for distributions with 2+epsilon moments*, Ann. of Prob., to appear.

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

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

Rafał Latała,
Institute of Mathematics,
University of Warsaw,
Banacha 2, 02-097 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: aelitvak@gmail.com

Krzysztof Oleszkiewicz,
Institute of Mathematics,
University of Warsaw,
Banacha 2, 02-097 Warszawa, Poland
e-mail: koles@mimuw.edu.pl

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