

Remarks on the Central Limit Theorem for Non-convex Bodies

Uri Grupel

Abstract In this note, we study possible extensions of the Central Limit Theorem for non-convex bodies. First, we prove a Berry-Esseen type theorem for a certain class of unconditional bodies that are not necessarily convex. Then, we consider a widely-known class of non-convex bodies, the so-called p -convex bodies, and construct a counter-example for this class.

1 Introduction

Let X_1, \dots, X_n be random variables with $\mathbb{E}X_i = 0$ and $\mathbb{E}X_i X_j = \delta_{i,j}$ for $i, j = 1, 2, \dots, n$. Let $\theta \in S^{n-1}$, where $S^{n-1} \subseteq \mathbb{R}^n$ is the unit sphere centered at 0, and let G be a standard Gaussian random variable, that is G has density function $1/\sqrt{2\pi}e^{-x^2/2}$. We denote $X = (X_1, \dots, X_n)$. In this paper we examine different conditions on X under which $X \cdot \theta$ is close to G in distribution. The classical central limit theorem states that if X_1, \dots, X_n are independent then for most $\theta \in S^{n-1}$ the marginal $X \cdot \theta$ is close to G . It was conjectured by Anttila et al. [1] and by Brehm and Voigt [6] that if X is distributed uniformly in a convex body $K \subseteq \mathbb{R}^n$, then for most $\theta \in S^{n-1}$ the marginal $X \cdot \theta$ is close to G . This is known as the central limit theorem for convex sets and was first proved by Klartag [12].

In this note we examine extensions of the above theorem to non-convex settings. Our study was motivated by the following observation on the unit balls of l_p spaces for $0 < p < 1$:

We denote by $B_p^n = \{x \in \mathbb{R}^n; |x_1|^p + \dots + |x_n|^p \leq 1\}$ the unit ball of the space l_p^n . For $X = (X_1, \dots, X_n)$ that is distributed uniformly on $c_{p,n} B_p^n$, $p > 0$, $\theta \in S^{n-1}$, and G a standard Gaussian, one can show that

$$|\mathbb{P}(\theta \cdot X \leq t) - \mathbb{P}(G \leq t)| \leq C_p \sum_{k=1}^n |\theta_k|^3$$

U. Grupel (✉)
Tel Aviv University, Tel Aviv, Israel
e-mail: urigrupe@post.tau.ac.il

where $c_{p,n}$ is chosen such that $\mathbb{E}X_i = 0$ and $\mathbb{E}X_i X_j = \delta_{i,j}$ for $i, j = 1, 2, \dots, n$, and $C_p > 0$ does not depend on n .

In order to formulate our results we use the following definitions: Let $X = (X_1, \dots, X_n)$ be a random vector in \mathbb{R}^n . A random vector X is called isotropic if $\mathbb{E}X_i = 0$ and $\mathbb{E}X_i X_j = \delta_{i,j}$ for $i, j = 1, 2, \dots, n$. A random vector X is called unconditional if the distribution of $(\varepsilon_1 X_1, \dots, \varepsilon_n X_n)$ is the same as the distribution of X for any $\varepsilon_i = \pm 1, i = 1, \dots, n$.

The first class of densities we define is based on Klartag’s recent work [14] and includes the uniform distribution over B_p^n for $0 < p < 1$.

Theorem 1. *Let X be an unconditional, isotropic random vector with density $e^{-u(x)}$, where the function $u(x_1^\kappa, \dots, x_n^\kappa)$ is convex in $\mathbb{R}_+^n = \{x \in \mathbb{R}^n; x_i \geq 0 \forall i \in \{1, 2, \dots, n\}\}$ for $\kappa > 1$. Let G be a standard Gaussian random variable and $\theta \in S^{n-1}$. Then*

$$|\mathbb{P}(\theta \cdot X \geq t) - \mathbb{P}(G \geq t)| \leq C_\kappa \sum_{k=1}^n |\theta_k|^3,$$

where $C_\kappa > 0$ depends on κ only, and does not depend on n .

In order to see that Theorem 1 includes the uniform distribution over B_p^n for $0 < p < 1$ take

$$u(x) = \begin{cases} 0, & x_1^p + \dots + x_n^p \leq 1 \\ \infty, & \text{otherwise} \end{cases},$$

and set $\kappa = 1/p$.

The error rate in Theorem 1 is the same as in the classical Central Limit Theorem. For example, by choosing $\theta = (1/\sqrt{n}, \dots, 1/\sqrt{n})$, we get an error rate of $O(1/\sqrt{n})$.

The symmetry conditions in Theorem 1 are highly restrictive. Hence, we are led to study p -convex bodies, which satisfy fewer symmetry conditions and are shown to share some of the properties of convex bodies.

We say that $K \subset \mathbb{R}^n$ is p -convex with $0 < p < 1$ if $K = -K$ and for all $x, y \in K$ and $0 < \lambda < 1$, we have

$$\lambda^{1/p}x + (1 - \lambda)^{1/p}y \in K.$$

These bodies are related to unit balls of p -norms and were studied in relation to local theory of Banach spaces by Gordon and Lewis [10], Gordon and Kalton [9], Litvak et al. [16] and others (see [2, 7, 11, 15, 17]).

The following discussion explains why the class of p -convex bodies does not give the desired result.

Theorem 2. *Set $N = n + n^{5/2} \log^2 n$. There exists a random vector X distributed uniformly in a $1/2$ -convex body $K \subseteq \mathbb{R}^N$, and a subspace E with $\dim(E) = n$,*

such that for any $\theta \in S^{N-1} \cap E$, the random variable $\theta \cdot \text{Proj}_E X$ is not close to a Gaussian random variable in any reasonable sense (Kolmogorov distance, Wasserstein distance and others).

A similar construction can be made for any fixed parameter $0 < p < 1$. Since $\dim(E)$ tends to infinity with n , a similar theorem is not true in the convex case. Hence, the central limit theorem for convex sets cannot be extended for the p -convex case. Thus, we need to look for a new class of bodies (densities) that includes the l_p^n unit balls, with a weaker condition than the unconditional one.

Remark 1. In [15] Litvak constructed an example of a p -convex body for which the volume distribution is very different from the convex case. Litvak’s work studies the large deviations regime for p -convex distributions, while our work is focused on the central limit theorem.

Throughout the text the letters c, C, c', C' will denote universal positive constants that do not depend on the dimension n . The value of the constant may change from one instance to another. We use $C_\alpha, C(\alpha)$ for constants that depend on a parameter α and nothing else. σ_{n-1} will denote the Haar probability measure on S^{n-1} . $f(n) = O(g(n))$ is the big O notation, i.e. there exists a constant $C > 0$ such that $|f(n)| \leq Cg(n), \forall n \in \mathbb{N}$.

2 A Class of Densities with Symmetries

In this section we use Klartag’s recent work [14] in order to exhibit a family of functions, which includes the indicator functions of l_p^n unit balls, for $0 < p < 1$, having almost Gaussian marginals.

A special case of Theorem 1.1 in [14] gives us the following lemma.

Lemma 1. *Let $\kappa > 1$ and let $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ be an unconditional function such that $e^{-\phi(x)}$ is a probability density function and $\phi(x_1^\kappa, \dots, x_n^\kappa)$ is convex on \mathbb{R}_+^n . Let X be a random vector with density $e^{-\phi(x)}$. Then*

$$\text{Var}|X|^2 \leq c_\kappa \sum_{j=1}^n \mathbb{E}|X_j|^4,$$

where c_κ depends only on κ .

Lemma 2. *Let $\kappa \geq 1$ and let $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ be an unconditional function such that $e^{-\phi(x)}$ is a probability density function and $\phi(x_1^\kappa, \dots, x_n^\kappa)$ is convex on \mathbb{R}_+^n . Let X be a random vector with density $e^{-\phi(x)}$. Then for any $p \geq 1$ and $i = 1, \dots, n$,*

$$\mathbb{E}|X_i|^p \leq c_{p,\kappa} (\mathbb{E}|X_i|^2)^{p/2}.$$

Proof. If $p \leq 2$ then, by Hölder's inequality, we have $c_{p,\kappa} = 1$. Assume that $p \geq 2$. Define $\pi : \mathbb{R}_+^n \rightarrow \mathbb{R}_+^n$ by $\pi(x) = (|x_1|^\kappa, \dots, |x_n|^\kappa)$. The Jacobian of π is $\prod_{j=1}^n \kappa |x_j|^{\kappa-1}$. Using the symmetry of ϕ we obtain

$$\begin{aligned} \int_{\mathbb{R}^n} |x_i|^p e^{-\phi(x)} dx &= 2^n \int_{\mathbb{R}_+^n} |x_i|^p e^{-\phi(x)} dx \\ &= 2^n \int_{\mathbb{R}_+^n} |x_i|^{p\kappa} \left(\prod_{j=1}^n \kappa |x_j|^{\kappa-1} \right) e^{-\phi(\pi(x))} dx \end{aligned}$$

Now set $u(x) = \phi(\pi(x)) - (\kappa - 1) \sum_{j=1}^n \log |x_j|$. The function $e^{-u(x)}$ is log-concave on \mathbb{R}_+^n , with $\kappa^n \int_{\mathbb{R}_+^n} e^{-u(x)} = 1/2^n$, and

$$\int_{\mathbb{R}_+^n} |x_i|^p e^{-\phi(x)} dx = \kappa^n \int_{\mathbb{R}_+^n} |x_i|^{p\kappa} e^{-u(x)} dx.$$

By Borell's Lemma (see [4, 5, 18]) we obtain

$$\begin{aligned} (2\kappa)^n \int_{\mathbb{R}_+^n} |x_i|^{p\kappa} e^{-u(x)} dx &\leq C_{\kappa,p} \left((2\kappa)^n \int_{\mathbb{R}_+^n} |x_i|^{2\kappa} e^{-u(x)} dx \right)^{p/2} \\ &= C_{\kappa,p} \left(\int_{\mathbb{R}^n} |x_i|^2 e^{-\phi(x)} dx \right)^{p/2} \end{aligned}$$

Lemma 3. Let $\kappa > 1$ and let $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ be an unconditional function such that $e^{-\phi(x)}$ is an isotropic probability density and $\phi(x_1^\kappa, \dots, x_n^\kappa)$ is convex on \mathbb{R}_+^n . Let X be a random vector with density $e^{-\phi(x)}$. Then, for any $a \in \mathbb{R}^n$

$$\text{Var}(a_1^2 X_1^2 + \dots + a_n^2 X_n^2) \leq C_\kappa \sum_{j=1}^n |a_j|^4.$$

Proof. By applying a linear transformation, Lemma 1 gives

$$\text{Var}(a_1^2 X_1^2 + \dots + a_n^2 X_n^2) \leq C'_\kappa \sum_{j=1}^n \mathbb{E} a_j^4 |X_j|^4.$$

By Lemma 2, we obtain

$$\text{Var}(a_1^2 X_1^2 + \dots + a_n^2 X_n^2) \leq C'_\kappa \sum_{j=1}^n \mathbb{E} a_j^4 |X_j|^4 \leq C_\kappa \sum_{j=1}^n a_j^4 (\mathbb{E}|X_j|^2)^2 = C_\kappa \sum_{j=1}^n |a_j|^4.$$

We are now ready to prove Theorem 1.

Proof. Since X is unconditional,

$$\mathbb{P}(\theta \cdot X \geq t) = \mathbb{P}\left(\sum_{k=1}^n \theta_k X_k \varepsilon_k \geq t\right),$$

where $\varepsilon_1, \dots, \varepsilon_n$ are i.i.d. random variables distributed uniformly on $\{\pm 1\}$ that are independent of X . By the triangle inequality,

$$\begin{aligned} \left| \mathbb{P}\left(\sum_{k=1}^n \theta_k X_k \varepsilon_k \geq t\right) - \mathbb{P}(G \geq t) \right| &\leq \mathbb{E}_X \left| \mathbb{P}(G \geq t) - \mathbb{P}_G\left(G \geq \frac{t}{\sqrt{\sum_{k=1}^n \theta_k^2 X_k^2}}\right) \right| \\ &+ \mathbb{E}_X \left| \mathbb{P}_\varepsilon\left(\sum_{k=1}^n \varepsilon_k \theta_k X_k \geq t\right) - \mathbb{P}_G\left(G \geq \frac{t}{\sqrt{\sum_{k=1}^n \theta_k^2 X_k^2}}\right) \right|. \end{aligned}$$

We estimate each term separately. Denote $Y_n = \sum_{k=1}^n \theta_k^2 X_k^2$. By the Berry-Esseen Theorem (see [8]),

$$\begin{aligned} &\mathbb{E}_X \left| \mathbb{P}_\varepsilon\left(\sum_{k=1}^n \varepsilon_k \theta_k X_k \geq t\right) - \mathbb{P}_G\left(G \geq \frac{t}{\sqrt{Y_n}}\right) \right| \\ &\leq C \left(\mathbb{E}_X \sum_{k=1}^n \frac{|\theta_k|^3 |X_k|^3}{(Y_n)^{3/2}} 1_{[1/2, \infty)}(Y_n) + 2\mathbb{P}\left(Y_n < \frac{1}{2}\right) \right) \\ &\leq C \left(10 \sum_{k=1}^n \mathbb{E}_X |\theta_k|^3 |X_k|^3 + 2\mathbb{P}\left(Y_n < \frac{1}{2}\right) \right) \leq C_\kappa \sum_{k=1}^n |\theta_k|^3 + C \mathbb{P}\left(Y_n < \frac{1}{2}\right) \end{aligned}$$

Here we used Lemma 2 to estimate $\mathbb{E}|X_k|^3$. Note that

$$\mathbb{E}_X Y_n = \mathbb{E}_X \sum_{j=1}^n \theta_j^2 X_j^2 = \sum_{j=1}^n \theta_j^2 \mathbb{E}_X X_j^2 = \sum_{j=1}^n \theta_j^2 = 1,$$

so by Chebyshev's inequality and Lemma 3

$$\mathbb{P}\left(|Y_n - 1| \geq \frac{1}{2}\right) \leq \frac{\text{Var}(Y_n)}{1/4} \leq 4C_\kappa \sum_{j=1}^n |\theta_j|^4 \quad (1)$$

Hence, since $|\theta_i| \leq 1$ for all $i = 1, \dots, n$,

$$\mathbb{E}_X \left| \mathbb{P}_\varepsilon \left(\sum_{k=1}^n \varepsilon_k \theta_k X_k \geq t \right) - \mathbb{P} \left(G \geq \frac{t}{\sqrt{Y_n}} \right) \right| \leq C_\kappa \sum_{k=1}^n |\theta_k|^3.$$

Now, in order to estimate $\mathbb{E}_X \left| \mathbb{P}(G \geq t) - \mathbb{P} \left(G \geq \frac{t}{\sqrt{Y_n}} \right) \right|$ we use (1) and Klartag's argument in [13] (Sect. 6, Lemma 7) and conclude that it is enough to show that

$$\mathbb{E} \left((Y_n - 1)^2 \middle| Y_n \geq \frac{1}{2} \right) \leq C \left(\sum_{j=1}^n |\theta_j|^3 \right)$$

By Lemma 3 we get

$$\mathbb{E} (Y_n - 1)^2 = \text{Var} (Y_n) \leq C_\kappa \sum_{j=1}^n |\theta_j|^4$$

Hence,

$$\begin{aligned} \mathbb{E} \left((Y_n - 1)^2 \middle| Y_n \geq \frac{1}{2} \right) &\leq \mathbb{E} (Y_n - 1)^2 \mathbb{P} \left(Y_n \geq \frac{1}{2} \right)^{-1} \\ &\leq C_\kappa \left(\sum_{j=1}^n |\theta_j|^4 \right) \mathbb{P} \left(Y_n \geq \frac{1}{2} \right)^{-1} \end{aligned}$$

From inequality (1) it follows that

$$\left(\mathbb{P} \left(Y_n \geq \frac{1}{2} \right) \right)^{-1} = \mathbb{P} \left(\sum_{j=1}^n \theta_j^2 X_j^2 \geq \frac{1}{2} \right)^{-1} \leq \frac{1}{1 - C_\kappa \sum_{j=1}^n |\theta_j|^4}.$$

We may assume that $\sum_{j=1}^n |\theta_j|^4$ is bounded by some small positive constant depending on κ , since otherwise the result is trivial, and obtain

$$\frac{1}{1 - C_\kappa \sum_{j=1}^n |\theta_j|^4} \leq 1 + C_\kappa \sum_{j=1}^n |\theta_j|^4$$

which completes our proof.

3 The p -Convex Case

In this section we construct a random vector X , distributed uniformly in a $1/2$ -convex body K , such that for a large subspace $E \subseteq \mathbb{R}^n$ the random vector $\text{Proj}_E X$ has no single approximately Gaussian marginal. We define a function $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that the radial density $r^{n-1}e^{-f(r)}$ is spread across an interval of length proportional to \sqrt{n} ; that is, we want $r^{n-1}e^{-f(r)}$ to be constant (or close to constant) on such an interval. Such densities have marginals that are far from Gaussian. We use the density function introduced above and an approximation argument to construct the desired body K .

In order to construct a p -convex body from a function f , we restrict ourselves to p -convex functions.

Definition 1. A function $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ is called p -convex if for any $x, y \in \mathbb{R}^n$ and $t \in [0, 1]$,

$$f(t^{1/p}x + (1-t)^{1/p}y) \leq tf(x) + (1-t)f(y). \tag{2}$$

The following proposition allows us to construct a p -convex body with $0 < p < 1$ from a p -convex function.

Proposition 1. For $\psi : \mathbb{R}^n \rightarrow \mathbb{R}_+$ p -convex function with $0 < p < 1$ and fixed $N > 0$, define $f_N(x) = (1 - \psi(x)/N)_+^N$. Then the set

$$K_N(\psi) = \left\{ (x, y); x \in \mathbb{R}^n, y \in \mathbb{R}^N, |y| < f_N^{1/N}(x) \right\}$$

is p -convex.

Proof. Let $(x_1, y_1), (x_2, y_2) \in K_N(\psi)$. Since $(x_i, y_i) \in K_N(\psi)$ we have $f_N(x_i) > 0$. Therefore,

$$f_N^{1/N}(x_i) = 1 - \frac{\psi(x_i)}{N}.$$

Let $0 \leq t \leq 1$ we get

$$\begin{aligned} f_N^{1/N}(t^{1/p}x_1 + (1-t)^{1/p}x_2) &\geq 1 - \frac{1}{N}\psi(t^{1/p}x_1 + (1-t)^{1/p}x_2) \\ &\geq 1 - \frac{1}{N}(t\psi(x_1) + (1-t)\psi(x_2)) \\ &= tf_N^{1/N}(x_1) + (1-t)f_N^{1/N}(x_2) > t|y_1| + (1-t)|y_2| \\ &\geq |t^{1/p}y_1| + |(1-t)^{1/p}y_2| \geq |t^{1/p}y_1 + (1-t)^{1/p}y_2|. \end{aligned}$$

Hence, $t^{1/p}(x_1, y_1) + (1-t)^{1/p}(x_2, y_2) \in K_N(\psi)$, as needed.

Proposition 2. *There exists a universal constant $C > 0$ such that, for $a \geq C$ the function*

$$f(x) = \begin{cases} \log a, & \text{if } 0 \leq x \leq a \\ \log x, & \text{if } a \leq x \leq 2a \\ \sqrt{x} - \sqrt{2a} + \log 2a, & \text{if } 2a \leq x \end{cases}$$

is $1/2$ -convex.

Proof. We begin by verifying that the function f is $1/2$ -convex for each interval $[0, a]$, $[a, 2a]$, $[2a, \infty)$. Then we need to check that condition (2) holds when x and y are from different intervals. By symmetry, we may assume that $x < y$. The cases $x, y \in [0, a]$ and $x, y \in [2a, \infty)$ are straightforward. In order for condition (2) to hold for the function $\log x$ on an interval $[a, b]$ we must show that for any $x, y \in [a, b]$

$$\log((1-t)^2x + t^2y) \leq (1-t)\log(x) + t\log(y) = \log(x^{1-t}y^t). \quad (3)$$

This is equivalent to

$$(1-t)^2x + t^2y - x^{1-t}y^t \leq 0.$$

Setting here $y = cx$, we obtain

$$(1-t)^2 + t^2c - c^t \leq 0.$$

This inequality holds for every $1 \leq c \leq 4$ and $0 \leq t \leq 1$. To see that note that $g(t, c) = (1-t)^2 + t^2c - c^t$ is a convex function in c (as a sum of convex functions). Hence, it is enough to verify that $g(t, 1) \leq 0$ and $g(t, 4) \leq 0$ for any $0 \leq t \leq 1$. Indeed,

$$g(t, 1) = (1-t)^2 + t^2 - 1 = 2t(t-1) \leq 0$$

and

$$g(t, 4) = (1-t)^2 + t^2 \cdot 4 - 4^t \Rightarrow \frac{\partial^2 g(t, 4)}{\partial t^2} = 2 + 8 - (\log 4)^2 4^t \geq 2.$$

Hence, $g(t, 4)$ is convex in t . Since $g(0, 4) = g(1, 4) = 0$, we obtain, $g(t, 4) \leq 0$ for all $0 \leq t \leq 1$.

Consequently (3) holds for any interval of the form $[a, b] \subseteq [a, 4a]$.

Next, we verify condition (2) for f when $x \in [a, 2a]$, $y \in [2a, \infty)$, and $t^2x + (1-t)^2y \in [a, 2a]$. We consider two cases

1. $y \in [2a, 4a]$. By inequality (3),

$$\begin{aligned} f(t^2x + (1-t)^2y) &= \log(t^2x + (1-t)^2y) \leq t \log(x) + (1-t) \log(y) \\ &\leq \log(x) + (1-t)(\log(2a) + \sqrt{y} - \sqrt{2a}) \\ &= tf(x) + (1-t)f(y). \end{aligned}$$

The second inequality holds thanks to the elementary inequality $\log(y) - \log(2a) \leq \sqrt{y} - \sqrt{2a}$. Since for $y = 2a$ we have equality, and $(\sqrt{y})' = 1/\sqrt{4y} \geq 1/y = (\log(y))'$ for $y \geq 4$, the inequality holds if $2a \geq 4$.

2. $y \geq 4a$. Define

$$g(t) = \log(t^2x + (1-t)^2y) - t \log(x) - (1-t)(\sqrt{y} - \sqrt{2a} + \log(2a)).$$

We need to show that $g(t) \leq 0$ for all $t \in [0, 1]$. Since $g(1) = 0$, it is enough to show that $g'(t) \geq 0$ for all $0 \leq t \leq 1$. We have,

$$\begin{aligned} g'(t) &= \frac{2tx - 2(1-t)y}{t^2x + (1-t)^2y} - \log(x) + \sqrt{y} - \sqrt{2a} + \log(2a) \\ &\geq \frac{2tx - 2(1-t)y}{t^2x + (1-t)^2y} + \left(1 - \frac{1}{\sqrt{2}}\right) \sqrt{y} \end{aligned}$$

Hence, if $2tx - 2(1-t)y + \left(1 - 1/\sqrt{2}\right) \sqrt{y}(t^2x + (1-t)^2y) \geq 0$, then $g'(t) \geq 0$.

Recalling that $t^2x + (1-t)^2y \geq a$, it suffices to prove that

$$2tx - 2(1-t)y + \left(1 - \frac{1}{\sqrt{2}}\right) \sqrt{y}a \geq 0.$$

Using the fact that $(1-t)^2y \leq t^2x + (1-t)^2y \leq 2a$, we obtain $(1-t)\sqrt{y} \leq \sqrt{2a}$. Hence,

$$\begin{aligned} 2tx - 2(1-t)y + \left(1 - \frac{1}{\sqrt{2}}\right) a\sqrt{y} &\geq 2ta - 2\sqrt{2a}\sqrt{y} + \left(1 - \frac{1}{\sqrt{2}}\right) a\sqrt{y} \\ &\geq \sqrt{y} \left(\left(1 - \frac{1}{\sqrt{2}}\right) a - 2\sqrt{2a} \right). \end{aligned}$$

This gives the condition

$$\left(1 - \frac{1}{\sqrt{2}}\right) a - 2\sqrt{2a} \geq 0,$$

Which is satisfied for $a \geq 100$.

When $x \in [a, 2a]$ and $y \in [2a, \infty)$ and $t^2x + (1 - t)^2y \geq 2a$, we have

$$\begin{aligned} f(t^2x + (1 - t)^2y) &= \sqrt{t^2x + (1 - t)^2y} - \sqrt{2a} + \log 2a \\ &\leq t\sqrt{x} + (1 - t)\sqrt{y} - \sqrt{2a} + \log 2a \end{aligned}$$

and

$$tf(x) + (1 - t)f(y) = t \log x + (1 - t)(\sqrt{y} - \sqrt{2a} + \log 2a).$$

Hence, (2) holds thanks to the elementary inequality $\log 2a - \log x + \sqrt{x} - \sqrt{2a} \leq 0$, which holds for $a \geq 4$.

If $x \in [0, a]$, then $f(x) = f(a)$ and $f(t^2x + (1 - t)^2y) \leq f(t^2a + (1 - t)^2y)$. Hence, for $x \in [0, a]$ and $y \in [a, \infty)$ we have

$$f(t^2x + (1 - t)^2y) \leq f(t^2a + (1 - t)^2y) \leq tf(a) + (1 - t)f(y) = tf(x) + (1 - t)f(y).$$

Proposition 3. Let $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be a p -convex function with parameter $0 < p < 1$. Then $x \mapsto f(|x|)$ is a p -convex function on \mathbb{R}^n .

Proof. First, we prove that f is non-decreasing. Let $0 < x < y$. There exists some $k \geq 1$ such that $2^{-k(1/p-1)}y \leq x$. We proceed by induction on k . For $k = 1$, note that $h(t) = t^{1/p}y + (1 - t)^{1/p}y$ is continuous, $h(0) = y$, and $h(1/2) = 2^{-(1/p-1)}y$. Hence, there exists some $0 \leq t_0 \leq 1$ for which $h(t_0) = x$, and so

$$f(x) = f(t_0^{1/p}y + (1 - t_0)^{1/p}y) \leq t_0f(y) + (1 - t_0)f(y) = f(y)$$

For $k \geq 2$, $f(2^{-(k-1)(1/p-1)}y) \leq f(y)$ by the induction hypothesis, and by the same argument as above

$$f(x) \leq f(2^{-(k-1)(1/p-1)}y) \leq f(y).$$

We thus showed that f is monotone non-decreasing. Now, by the triangle inequality, for any $x, y \in \mathbb{R}^n$ and $0 < t < 1$ we have

$$f(|t^{1/p}x + (1 - t)^{1/p}y|) \leq f(t^{1/p}|x| + (1 - t)^{1/p}|y|) \leq tf(|x|) + (1 - t)f(|y|)$$

Using the function from Proposition 2, we are ready to construct the $1/2$ -convex body K and prove Theorem 2.

Definition 2. A sequence of probability measures $\{\mu_n\}$ on \mathbb{R}^n is called essentially isotropic if $\int x d\mu_n(x) = 0$ and $\int x_i x_j d\mu_n(x) = (1 + \varepsilon_n)\delta_{ij}$ for all $i, j = 1, \dots, n$, when $\varepsilon_n \xrightarrow{n \rightarrow \infty} 0$.

Proposition 4. The probability measure $d\mu = C_n e^{-(n-1)f(|x|)} dx$, where f is defined as in Proposition 2, with $a = \sqrt{3n/7}$, is essentially isotropic. That is,

$$\int x_i x_j d\mu(x) = (1 + \varepsilon_n)\delta_{ij}$$

for all $i, j = 1, 2, \dots, n$, when $|\varepsilon_n| \leq C/n$.

Proof. The density μ is spherically symmetric, hence

$$\int_{\mathbb{R}^n} x_i x_j d\mu(x) = 0,$$

for $i \neq j$, and

$$\int_{\mathbb{R}^n} x_i^2 d\mu(x) = \frac{1}{n} \int_{\mathbb{R}^n} |x|^2 d\mu(x),$$

for $i = 1, 2, \dots, n$. Integration in spherical coordinates and using Laplace asymptotic method yields

$$\begin{aligned} \int |x|^2 d\mu(x) &= \frac{\left[\int_0^{\sqrt{3n/7}} \frac{r^{n+1}}{(\sqrt{3n/7})^{n-1}} dr + \int_{\sqrt{3n/7}}^{2\sqrt{3n/7}} r^2 dr \right. \\ &\quad \left. + \left(\frac{e^{\sqrt{2\sqrt{3n/7}}}}{2\sqrt{3n/7}} \right)^{n-1} \int_{2\sqrt{3n/7}}^{\infty} r^{n+1} e^{-(n-1)\sqrt{r}} dr \right]}{\left[\int_0^{\sqrt{3n/7}} \left(\frac{r}{\sqrt{3n/7}} \right)^{n-1} dr + \int_{\sqrt{3n/7}}^{2\sqrt{3n/7}} dr \right. \\ &\quad \left. + \left(\frac{e^{\sqrt{2\sqrt{3n/7}}}}{2\sqrt{3n/7}} \right)^{n-1} \int_{2\sqrt{3n/7}}^{\infty} r^{n-1} e^{-(n-1)\sqrt{r}} dr \right]} \\ &= \frac{\sqrt{3/7}n^{3/2} + O(\sqrt{n})}{\sqrt{\frac{3}{7}}n + O(1/\sqrt{n})} = n + O(1). \end{aligned}$$

Proposition 5. Let X be a random vector in \mathbb{R}^n distributed according to μ from Proposition 4. Then,

$$\mathbb{P}\left(\sqrt{\frac{3}{7}}n \leq |X| \leq 2\sqrt{\frac{3}{7}}n\right) \geq 1 - \frac{C}{n}.$$

Proof. By the same arguments as in Proposition 4

$$\mathbb{P}\left(\sqrt{3n/7} \leq |X| \leq 2\sqrt{3n/7}\right) = \frac{\int_{\sqrt{3n/7}}^{2\sqrt{3n/7}} dr}{\sqrt{3n/7} + O(1/\sqrt{n})} = 1 + O(1/n).$$

Proposition 6. *Let X be a random vector in \mathbb{R}^n distributed according to μ from Proposition 4, and let \tilde{X} be a random variable distributed according to $d\tilde{\mu} = \widetilde{C}_n (1 - (n - 1)f(|x|)/N)_+^N$. Then for $N \geq n^{5/2} \log^2 n$, \tilde{X} is essentially isotropic, namely*

$$\int x_i x_j d\tilde{\mu}(x) = (1 + \varepsilon'_n) \delta_{ij}$$

for all $i, j = 1, 2, \dots, n$, when $|\varepsilon'_n| \leq C/\sqrt{n}$. Also

$$\forall t, \quad |\mathbb{P}(|X| \leq t) - \mathbb{P}(|\tilde{X}| \leq t)| \leq \frac{C}{\sqrt{n}}.$$

Proof. The random vector \tilde{X} is spherically symmetric. Hence

$$\int_{\mathbb{R}^n} x_i x_j d\tilde{\mu}(x) = 0,$$

for $i \neq j$, and

$$\int_{\mathbb{R}^n} x_i^2 d\tilde{\mu}(x) = \frac{1}{n} \int_{\mathbb{R}^n} |x|^2 d\tilde{\mu}(x),$$

for $i = 1, 2, \dots, n$. Since both densities are spherically symmetric, we need to estimate the one-dimensional integrals

$$I_k = \int_0^\infty r^k \left(e^{-(n-1)f(r)} - \left(1 - \frac{(n-1)f(r)}{N} \right)_+^N \right) dr$$

for $k = n - 1, n + 1$. Define α by the equation

$$\left(\sqrt{\alpha} - \sqrt{2\sqrt{\frac{3}{7}n} + \log\left(2\sqrt{\frac{3}{7}n}\right)} \right) (n - 1) = \frac{N}{2}.$$

That is, for any $r \leq \alpha$ we have $(n - 1)f(r)/N \leq 1/2$. By Taylor's Theorem, for any $r \leq \alpha$,

$$\left| \log\left(1 - \frac{(n-1)f(r)}{N}\right)_+^N - (-(n-1)f(r)) \right| \leq C \frac{(n-1)^2}{N} f^2(r).$$

Hence, for any $r \leq \alpha$

$$\begin{aligned} & \left| e^{-(n-1)f(r)} - \left(1 - \frac{(n-1)}{N} f(r) \right)_+^N \right| \\ &= e^{-(n-1)f(r)} \left| 1 - \exp \left((n-1)f(r) - \log \left(1 - \frac{(n-1)}{N} f(r) \right)_+^N \right) \right| \\ &\leq C \frac{n^2}{N} e^{-(n-1)f(r)} f^2(r). \end{aligned}$$

Note that

$$\left| \int_\alpha^\infty \left(e^{-(n-1)f(r)} - \left(1 - \frac{(n-1)}{N} f(r) \right)_+^N \right) dr \right| \leq C \int_\alpha^\infty e^{-(n-1)f(r)} dr \leq C e^{-n}.$$

Combining the above inequalities, we obtain

$$|I_k| \leq C_1 \frac{n^2}{N} \int_0^\alpha r^k e^{-(n-1)f(r)} f^2(r) dr + C_2 e^{-n} \leq C \frac{n^2}{N} \int_0^\infty r^k e^{-(n-1)f(r)} f^2(r) dr.$$

Hence,

$$\begin{aligned} |I_{n-1}| &\leq C \frac{n^2}{N} \left(\sqrt{n} \log^2 n + O \left(\frac{\log^2 n}{\sqrt{n}} \right) \right) \leq C_1, \\ |I_{n+1}| &\leq C \frac{n^2}{N} \left(n^{\frac{3}{2}} \log^2 n + O(\log^2 n \sqrt{n}) \right) \leq C_2 n. \end{aligned}$$

By the estimation on I_{n-1} , and the calculations in Proposition 4 we obtain

$$\begin{aligned} & \left| \int_0^\infty \left(1 - \frac{(n-1)}{N} f(r) \right)_+^N dr - \sqrt{\frac{3}{7}n} \right| \\ &\leq \left| \int_0^\infty \left(1 - \frac{(n-1)}{N} f(r) \right)_+^N dr - \int_0^\infty e^{-(n-1)f(r)} dr \right| + O \left(\frac{1}{\sqrt{n}} \right) \\ &= |I_{n-1}| + O \left(\frac{1}{\sqrt{n}} \right) \leq C_1. \end{aligned}$$

Hence,

- $\left(\int_0^\infty \left(1 - \frac{(n-1)}{N} f(r) \right)_+^N dr \right)^{-1} = \sqrt{\frac{1}{\frac{3}{7}n}} \left(1 + O \left(\frac{1}{\sqrt{n}} \right) \right);$
- $\forall t, \quad \left| \mathbb{P}(|X| \leq t) - \mathbb{P}(|\tilde{X}| \leq t) \right| \leq \frac{C}{\sqrt{n}}.$

By the estimation of I_{n+1} we obtain,

$$|\mathbb{E}X_i^2 - \mathbb{E}\tilde{X}_i^2| = \frac{1}{n} |\mathbb{E}|X|^2 - \mathbb{E}|\tilde{X}|^2| \leq C \frac{1}{\sqrt{n}} \frac{1}{n} |I_{n+1}| \leq \frac{C}{\sqrt{n}}.$$

Remark 2. It is possible to take $a \approx \sqrt{3n/7}$ in the definition of f , such that \tilde{X} is isotropic.

We use the following estimation in our proof of Theorem 2.

Proposition 7. *Let Z_1, \dots, Z_n be independent standard Gaussian random variables, and let $0 < \delta < 1/2$. Then,*

$$\mathbb{P} \left(\left| \sqrt{Z_1^2 + \dots + Z_n^2} - \sqrt{n} \right| \leq n^\delta \right) \geq 1 - Ce^{-cn^{2\delta}},$$

where $c, C > 0$ are constants.

Proof. Note that

$$\begin{aligned} |Z_1^2 + \dots + Z_n^2 - n| &= \left| \sqrt{Z_1^2 + \dots + Z_n^2} - \sqrt{n} \right| \left| \sqrt{Z_1^2 + \dots + Z_n^2} + \sqrt{n} \right| \\ &\geq \left| \sqrt{Z_1^2 + \dots + Z_n^2} - \sqrt{n} \right| \sqrt{n}. \end{aligned}$$

Therefore it is enough to show that

$$\mathbb{P} \left(\left| Z_1^2 + \dots + Z_n^2 - n \right| \leq n^{\delta + \frac{1}{2}} \right) \geq 1 - Ce^{-cn^{2\delta}}.$$

Note that for all $m \geq 1$ and for all $i = 1, \dots, n$, we have

$$\mathbb{E}|Z_i^2 - 1|^m \leq \sum_{k=1}^m \binom{m}{k} \mathbb{E}Z_i^{2k} \leq 2^m (2m)!! \leq 4^m m!$$

where $(2m)!! = 1 \cdot 3 \cdot 5 \cdots (2m - 1)$. Hence, by Bernstein's inequality [3] we obtain

$$\mathbb{P} \left(\left| (Z_1^2 - 1) + \dots + (Z_n^2 - 1) \right| > n^{\frac{1}{2} + \delta} \right) \leq Ce^{-cn^{2\delta}}.$$

We are now ready to prove Theorem 2.

Proof. By Proposition 2, the function $(n - 1)f(|x|)$ is 1/2-convex. Proposition 1 with $N = n^{5/2} \log^2 n$ yields a 1/2-convex body K . Let X be a random vector distributed uniformly in K . By the definition of K the marginal of X with respect to the first n coordinates has density proportional to $(1 - (n - 1)f(|x|)/N)_+^N$. Denote this subspace by E . By Proposition 6, $\text{Proj}_E X$ is essentially isotropic. Let G be

a standard Gaussian random variable. In order to show that $Y = \text{Proj}_E X$ has no approximately Gaussian marginals, we examine $\mathbb{P}(|\theta_0 \cdot Y| \leq t)$, for any $\theta_0 \in S^{n-1}$. Using the symmetry of Y and the rotation invariance of σ_{n-1} , we obtain,

$$\begin{aligned} \mathbb{P}(|\theta_0 \cdot Y| \leq t) &= \mathbb{E}1_{[0,t]}(|\theta_0 \cdot Y|) = \int_{S^{n-1}} \mathbb{E}1_{[0,t]}(|\theta \cdot Y|) d\sigma_{n-1}(\theta) \\ &= \mathbb{E} \int_{S^{n-1}} 1_{[0,t]}(\theta_1 |Y|) d\sigma_{n-1}(\theta), \end{aligned}$$

where $\theta = (\theta_1, \dots, \theta_n)$. Let $Z = (Z_1, \dots, Z_n)$, where Z_i are independent standard Gaussian random variable. Since Z is invariant under rotations, $Z/|Z|$ is distributed uniformly on S^{n-1} . Hence,

$$\mathbb{P}(|\theta_0 \cdot Y| \leq t) = \mathbb{P}\left(|Z_1||Y| \leq t\sqrt{Z_1^2 + \dots + Z_n^2}\right).$$

By Proposition 7, $\mathbb{P}(|\sqrt{Z_1^2 + \dots + Z_n^2} - \sqrt{n}| \leq n^{1/100}) \geq 1 - Ce^{-cn^{1/50}}$. Hence,

$$\mathbb{P}(|\theta_0 \cdot Y| \leq t) = \mathbb{P}\left(|Z_1||Y| \leq t\sqrt{n}(1 + O(n^{-1/2+1/100}))\right) + O\left(e^{-cn^{1/50}}\right). \quad (4)$$

By Propositions 6 and 5, there exists a random vector Y' such that

$$\forall t \quad |\mathbb{P}(|Y'| \leq t) - \mathbb{P}(|Y| \leq t)| \leq \frac{C}{\sqrt{n}}, \quad \mathbb{P}\left(\sqrt{\frac{3}{7}}n \leq |Y'| \leq 2\sqrt{\frac{3}{7}}n\right) \geq 1 - \frac{C}{n},$$

and $|Y'|$ has constant density function on $\left[\sqrt{3n/7}, 2\sqrt{3n/7}\right]$. By the triangle inequality, for W distributed uniformly on $\left[\sqrt{3/7}, 2\sqrt{3/7}\right]$ and any $\sqrt{3/7} \leq \alpha \leq \beta \leq 2\sqrt{3/7}$ we have

$$|\mathbb{P}(\sqrt{n}\alpha \leq |Y| \leq \sqrt{n}\beta) - \mathbb{P}(\alpha \leq W \leq \beta)| \leq \frac{C}{\sqrt{n}}.$$

Combining with (4),

$$\mathbb{P}(|\theta_0 \cdot Y| \leq t) = \mathbb{P}\left(|G|W \leq t(1 + O(n^{-1/2+1/100}))\right) + O\left(\frac{1}{\sqrt{n}}\right).$$

We conclude that $|Y \cdot \theta_0|$ is very close to a distribution which is the product of a Gaussian with a uniform random variable, and the latter distribution is far from Gaussian.

Acknowledgements This paper is part of the authors M.Sc. thesis written under the supervision of Professor Bo'az Klartag whose guidance, support and patience were invaluable. In addition, I would like to thank Andrei Iacob for his helpful editorial comments. Supported by the European Research Council (ERC).

References

1. M. Anttila, K. Ball, I. Perissinaki, The central limit problem for convex bodies. *Trans. Am. Math. Soc.* **355**, 4723–4735 (2003)
2. J. Bastero, J. Bernues, A. Pena, An extension of Milman's reverse Brunn-Minkowski inequality. *Geom. Funct. Anal.* **5**(3), 572–581 (1995)
3. G. Bennet, Probability inequalities for the sum of independent random variables. *J. Am. Stat. Assoc.* **57**, 33–45 (1962)
4. L. Berwald, Verallgemeinerung eines Mittelwertsatzes von J. Favard für positive konkave Funktionen. *Acta Math.* **79**, 17–37 (1947)
5. C. Borell, Convex measures on locally convex spaces. *Arkiv för Matematik* **12**(1–2), 239–252 (1974)
6. U. Brehm, J. Voigt, Asymptotics of cross sections for convex bodies. *Beiträge Algebra Geom.* **41**, 437–454 (2000)
7. S.J. Dilworth, The dimension of Euclidean subspaces of quasinormed spaces. *Math. Proc. Camb. Philos. Soc.* **97**(2), 311–320 (1985)
8. W. Feller, *An Introduction to Probability Theory and its Applications*, vol. II, Sect. XVI.5 (Wiley, New York, 1971)
9. Y. Gordon, N.J. Kalton, Local structure theory for quasi-normed spaces. *Bull. Sci. Math.* **118**, 441–453 (1994)
10. Y. Gordon, D.R. Lewis, Dvoretzky's theorem for quasi-normed space. *Illinois J. Math.* **35**(2), 250–259 (1991)
11. N.J. Kalton, Convexity, type and the three space problem. *Stud. Math.* **69**, 247–287 (1981)
12. B. Klartag, A central limit theorem for convex sets. *Invent. Math.* **168**, 91–131 (2007)
13. B. Klartag, A Berry-Esseen type inequality for convex bodies with an unconditional basis. *Probab. Theory Relat. Fields* **45**(1), 1–33 (2009)
14. B. Klartag, Poincaré inequalities and moment maps. *Ann. Fac. Sci. Toulouse Math.* **22**(1), 1–41 (2013)
15. A.E. Litvak, Kahane-Khinchin's inequality for the quasi-norms. *Can. Math. Bull.* **43**, 368–379 (2000)
16. A.E. Litvak, V.D. Milman, N. Tomczak-Jaegermann, Isomorphic random subspaces and quotients of convex and quasi-convex bodies, in *GAFa*. *Lecture Notes in Mathematics*, vol. 1850 (Springer, Berlin, 2004), pp. 159–178
17. V. Milman, Isomorphic Euclidean regularization of quasi-norms in R^n . *C. R. Acad. Sci. Paris* **321**(7), 879–884 (1995)
18. V.D. Milman, A. Pajor, Isotropic position and inertia ellipsoids and zonoids of the unit ball of a normed n -dimensional space, in *Geometric Aspects of Functional Analysis - Israel Seminar*. *Lecture Notes in Mathematics*, vol. 1376 (Springer, Berlin, 1989), pp. 64–104