The Annals of Probability
2003, Vol. 31, No. 1, 195–215

ON CONCENTRATION OF DISTRIBUTIONS OF RANDOM WEIGHTED SUMS[1]

BY S. G. BOBKOV

University of Minnesota

For noncorrelated random variables, we study the rate of approximation of distributions of weighted sums by "typical" distributions.

1. Introduction. Let $X = (X_1, \dots, X_n)$, $n \geq 2$, be a vector of n random variables with finite second moments such that

$$\mathbf{E} X_k X_j = \delta_{kj}, \tag{1.1}$$

where δ_{kj} is Kronecker's symbol. Consider the sums

$$S_\theta = \sum_{k=1}^{n} \theta_k X_k$$

with coefficients $\theta = (\theta_1, \dots, \theta_n)$ taken from the unit sphere S^{n-1}, that is, such that $\theta_1^2 + \cdots + \theta_n^2 = 1$.

When X_k's are independent, identically distributed and have mean zero, the standard theory indicates (cf., e.g., [15]), that the distribution function of S_θ,

$$F_\theta(t) = \mathbf{P}\{S_\theta \leq t\}, \qquad t \in \mathbf{R},$$

is close to the standard normal law Φ, as soon as $\max_k |\theta_k|$ is small. Various extensions of this fundamental fact are known under different assumptions on dependence of the coordinates X_k's and for prescribed (fixed) coefficients. Note that the condition $\max_k |\theta_k| = o(1)$ defines a set Θ on the sphere whose spherical (normalized Lebesgue) measure $\sigma_{n-1}(\Theta)$ is almost 1. One may wonder therefore whether or not, in a reasonably general situation, one can choose a similar big set $\Theta \subset S^{n-1}$ depending on X such that, for all $\theta \in \Theta$, the uniform distance $\|F_\theta - \Phi\|_\infty = \sup_{t \in \mathbf{R}} |F_\theta(t) - \Phi(t)|$ is small enough. Other metrics would also be of interest.

In essence, this question contains two different concentration problems. First, one may ask whether most of F_θ's are close to a certain distribution, say, to the average distribution

$$F(t) = \int_{S^{n-1}} F_\theta(t)\, d\sigma_{n-1}(\theta).$$

Received June 2001; revised July 2002.

[1]Supported in part by NSF Grant.

AMS 2000 subject classifications. Primary 60F, 60G; secondary 46G.

Key words and phrases. Concentration, typical distributions, central limit theorem.

The second problem would be then how to measure the difference between F and Φ. With this formulation, the first problem, interesting in itself, was first studied by Sudakov [18]. He applied the isoperimetric theorem on the sphere to obtain in particular the following qualitative result: for each $\delta > 0$, there is an integer n_δ such that, if $n \geq n_\delta$ one can choose a set $\Theta_\delta \subset S^{n-1}$ of measure $\sigma_{n-1}(\Theta_\delta) \geq 1 - \delta$ with

$$\kappa(F_\theta, F) = \int_{-\infty}^{+\infty} |F_\theta(t) - F(t)|\, dt \leq \delta \qquad \text{for all } \theta \in \Theta_\delta. \tag{1.2}$$

Thus, the concentration property of the family $\{F_\theta\}_{\theta \in S^{n-1}}$ has a very universal character since no additional requirement on the distribution of X beyond (1.1) is needed. The latter can further be relaxed to $\mathbf{E}\,\langle\theta, X\rangle^2 \leq 1$ ($\theta \in S^{n-1}$), and this is what was actually assumed in [18]. Moreover, with a similar conclusion, the spherical measure σ_{n-1} may be replaced with a suitably normalized Gaussian measure on $\mathbf{R}^n$. A different proof of this result was recently suggested by von Weizsäcker [19]. He also obtained a quantitative version for Gaussian coefficients and for a specially constructed metric κ^*. As turns out, the "Gaussian coefficients" approach allows one to reach a rather general formulation for infinite-dimensional Gaussian cylindrical measure.

The study of the problem was continued by Antilla, Ball and Perissinaki [1] who considered an important special situation where the random vector X is uniformly distributed over an arbitrary centrally symmetric convex body in $\mathbf{R}^n$. Under (1.1), they prove that, for any $\delta > 0$, except for a set of directions of measure at most $4\sqrt{n}\log n\, e^{-n\delta^2/50}$, one has $\|F_\theta - F\|_\infty \leq \delta$; that is,

$$\sigma_{n-1}\{\|F_\theta - F\|_\infty \geq \delta\} \leq 4\sqrt{n}\log n\, e^{-n\delta^2/50}. \tag{1.3}$$

In addition to the concentation phenomenon on the sphere, the proof of [1] essentially relies on some deep facts from convex geometry (such as Busemann's theorem, convexity of the floating body). It is therefore to be understood what feature the convexity assumption brings in the concentration property of the weighted sums, and how to quantify for canonical metrics Sudakov's observation in nonconvex case.

To treat the general case, it is unlikely to be possible to work with the uniform distance since the average distribution F may degenerate at zero (at least asymptotically). Together with the Kantorovich–Rubinshtein distance $\kappa(F_\theta, F)$, we consider the Lévy distance $L(F_\theta, F)$ defined as the minimum over all $\delta \geq 0$ such that $F(t - \delta) - \delta \leq F_\theta(t) \leq F(t + \delta) + \delta$ for all $t \in \mathbf{R}$. As well as κ, the metric L is responsible for the weak convergence, and since F and all F_θ's have unit second moments, there is a simple relation $\kappa(F_\theta, F) \leq C L(F_\theta, F)^{1/2}$ (with C universal). In Section 2, we prove the following:

THEOREM 1.1. *If* (1.1) *holds true, for all* $\delta > 0$,

$$\sigma_{n-1}\{L(F_\theta, F) \geq \delta\} \leq 4n^{3/8}\, e^{-n\delta^4/8}. \tag{1.4}$$

Thus, in (1.2) one can take n_δ at least of order $c_\delta \delta^{-8}$ up to a logarithmically growing factor c_δ (as $\delta \downarrow 0$).

The difference between inequalities (1.3) and (1.4) appears in particular in the strength of concentration, and we believe this is due to the additional assumption on the shape of the distribution of X. We do not know how sharp the estimate (1.4) is; nevertheless, the inequality (1.3) can further be sharpened and extended to the family of all (isotropic) log-concave probability distributions on $\mathbf{R}^n$. For short, we say that the random vector X is log-concave, if it has a density p on $\mathbf{R}^n$ such that $p(tx + (1-t)y) \geq p(x)^t p(y)^{1-t}$, whenever $x, y \in \mathbf{R}^n$ and $t \in (0,1)$. In Section 3 we prove the theorem.

THEOREM 1.2. *Assume a log-concave random vector X has mean zero and satisfies the correlation condition* (1.1). *Then, for all $\delta > 0$,*

$$\sigma_{n-1}\left\{ \sup_{t \in \mathbf{R}} e^{c|t|} \, |F_\theta(t) - F(t)| \geq \delta \right\} \leq C\sqrt{n} \log n \, e^{-cn\delta^2}, \tag{1.5}$$

where c and C are positive universal constants.

According to (1.4), in order to approximate F_θ's by the standard normal distribution function Φ, one needs to estimate the uniform distance $\|F - \Phi\|_\infty$. The average distribution F can be characterized as the distribution of the product $\theta_1 |X|$, where θ_1, the first coordinate of a point on the sphere, is regarded as a random variable independent of the Euclidean norm $|X| = (X_1^2 + \cdots + X_n^2)^{1/2}$. As is well known, the distribution function Φ_n of $\theta_1 \sqrt{n}$ under the measure σ_{n-1} satisfies $\|\Phi_n - \Phi\|_\infty = O(\frac{1}{\sqrt{n}})$, as $n \to \infty$. Therefore, the distance $\|F - \Phi\|_\infty$ depends on how strong the distribution of $\frac{|X|}{\sqrt{n}}$ is concentrated around the point $t = 1$. In particular, it depends on the smallest value $\varepsilon_n = \varepsilon_n(X)$ such that

$$\mathbf{P}\left\{ \left| \frac{|X|}{\sqrt{n}} - 1 \right| \geq \varepsilon_n \right\} \leq \varepsilon_n. \tag{1.6}$$

Combining Theorem 1.1 with (1.6), we may conclude that, for any $\delta > 0$, except for a set of directions of measure at most $4n^{3/8} e^{-n\delta^4/8}$,

$$\sup_{t \in \mathbf{R}} |F_\theta(t) - \Phi(t)| \leq C(\delta + \varepsilon_n), \tag{1.7}$$

where C is a universal constant. In particular, this leads to the following version of the central limit theorem. A similar observation was made by Diaconis and Freedman ([6], Theorem 1.1) and by von Weizsäcker ([19], Theorem 3).

COROLLARY 1.3. *Let $(X_k)_{k \geq 1}$ be a sequence of random variables satisfying* (1.1) *and such that the following weak law of large numbers is fulfilled*:

$$\frac{X_1^2 + \cdots + X_n^2}{n} \to 1 \qquad \text{as } n \to \infty.$$

Then for some collection $\{\theta_{nk}\}_{k=1}^{n}$ *such that* $\theta_{n1}^2 + \cdots + \theta_{nn}^2 = 1$,

$$\sum_{k=1}^{n} \theta_{nk} X_k \to N(0,1) \qquad \text{as } n \to \infty.$$

The smallest value of ε_n in (1.6) is indeed very small in many interesting situations, and so the inequality (1.6) itself can be viewed as a certain general concentration hypothesis needing to be verified for wide classes of distributions on $\mathbf{R}^n$. From this point of view, an inequality similar to (1.7) was derived on the basis of (1.3) in [1], where the hypothesis (1.6) was verified for several subclasses of convex bodies. As for the general log-concave case, the property that, under (1.1), a sequence $\varepsilon_n \to 0$ in (1.6) can be chosen independent of X represents a weak form of Kannan–Lovász–Simonovits' conjecture on Cheeger-type isoperimetric constants (cf. [10]).

2. Concentration in Lévy and uniform metric. As usual, $\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathbf{R}^n$. As noticed and used in [18], for each function g on $\mathbf{R}$ with Lipschitz seminorm $\|g\|_{\mathrm{Lip}} \le C$, provided that $\mathbf{E}|\langle \theta, X\rangle| \le |\theta|$, for all $\theta \in \mathbf{R}^n$, the function

$$f(\theta) = \mathbf{E} g(\langle \theta, X \rangle) \tag{2.1}$$

has on $\mathbf{R}^n$ the Lipschitz seminorm at most C. Any such function, being considered on the sphere, is strongly concentrated around its mean $\bar{f} = \int_{S^{n-1}} f(\theta)\, d\sigma_{n-1}(\theta)$, in the sense that

$$\sigma_{n-1}\{|f - \bar{f}| \ge \delta\} \le 2e^{-(n-1)\delta^2/(2C^2)}, \qquad \delta > 0. \tag{2.2}$$

This concentration inequality is known as a consequence of Lévy's isoperimetric theorem on the sphere (cf. [13, 12]) and can also be refererred to the property that the logarithmic Sobolev constant of S^{n-1} is equal to $n-1$ [14].

PROOF OF THEOREM 1.1. We apply (2.2) to functions $f_a(\theta)$, $a \in \mathbf{R}$, defined via (2.1) with

$$g_a(t) = \begin{cases} 1, & \text{if } t \le a, \\ 1 - 2(t-a)/\delta, & \text{if } t \in [a, a+\delta/2], \\ 0, & \text{if } t \ge a + \delta/2. \end{cases}$$

Since $C = \|g_a\|_{\mathrm{Lip}} = \frac{2}{\delta}$, we thus have

$$\sigma_{n-1}\{|f_a - \bar{f}_a| \ge \delta\} \le 2e^{-(n-1)\delta^4/8}, \qquad \delta > 0, \tag{2.3}$$

where $\bar{f}_a$ is the mean of f_a over (S^{n-1}, σ_{n-1}).

Now, due to $\mathbb{1}_{(-\infty,a]} \le g_a \le \mathbb{1}_{(-\infty,a+\delta/2]}$,

$$F_\theta(a) \le f_a(\theta) \le F_\theta(a + \delta/2) \tag{2.4}$$

for all θ. Taking the average over θ's, we also have

$$F(a) \le \bar{f}_a \le F(a + \delta/2). \tag{2.5}$$

For a fixed $\delta \in (0, 1)$, put $\Omega(a) = \{\theta \in S^{n-1} : |f_a(\theta) - \bar{f}_a| < \delta\}$. We may take a sequence of numbers $a_0 < a_1 < \cdots < a_N$ such that $a_k - a_{k-1} \le \delta/2$, $k = 1, \ldots, N$, and $a_0 = -(a_N + \delta) = -1/\sqrt{\delta}$. Assume $\theta \in \bigcap_{k=1}^N \Omega(a_k)$ and take any real $t \in [a_{k-1}, a_k)$. Then, by (2.4) and (2.5),

$$F_\theta(t) \le F_\theta(a_k) \le f_{a_k}(\theta) \le \bar{f}_{a_k} + \delta \le F(a_k + \delta/2) + \delta \le F(t + \delta) + \delta.$$

Similarly, $F(t) \le F_\theta(t + \delta) + \delta$.

Analogous estimates will also hold for all $t > a_N$ and $t < a_0$, if the numbers $-a_0$ and a_N are so large that

$$\begin{aligned} 1 - F(a_N + \delta) &\le \delta, \qquad & F(a_0) &\le \delta, \\ 1 - F_\theta(a_N + \delta) &\le \delta, \qquad & F_\theta(a_0) &\le \delta. \end{aligned}$$

Here, we involve the assumption (1.1) from which it follows that $\int_{-\infty}^{+\infty} t^2\, dF_\theta(t) = \int_{-\infty}^{+\infty} t^2\, dF(t) = 1$. By Chebyshev's inequality, since $a_0 = -(a_N + \delta) = -1/\sqrt{\delta}$, we get

$$\big(1 - F(a_N + \delta)\big) + F(a_0) = \big(1 - F(-a_0)\big) + F(a_0) \le \frac{1}{a_0^2} = \delta.$$

The same argument may be repeated for F_θ. Thus, for all $t \in \mathbf{R}$, $F(t) \le F_\theta(t + \delta) + \delta$ and $F_\theta(t) \le F(t + \delta) + \delta$, which is equivalent to saying that $L(F_\theta, F) \le \delta$.

Thus, we arrived at the inclusion $\bigcap_{k=1}^N \Omega(a_k) \subset \{\theta \in S^{n-1} : L(F_\theta, F) \le \delta\}$. As a result, by (2.3), applied to $a = a_1, \ldots, a_k$, we obtain

$$\sigma_{n-1}\{L(F_\theta, F) > \delta\} \le \sum_{k=1}^N \big(1 - \sigma_{n-1}(\Omega(a_k))\big) \le 2N e^{-(n-1)\delta^4/8}.$$

It remains to estimate from above the minimal possible N. The total length of the intervals $[a_{k-1}, a_k]$, $1 \le k \le N$, is at most $N\delta/2$. On the other hand, it must be equal to $a_N - a_0 = -2a_0 - \delta = \frac{2}{\sqrt{\delta}} - \delta$. Therefore, $\frac{1}{2} N\delta \ge \frac{2}{\sqrt{\delta}} - \delta$, that is, $N \ge \frac{4}{\delta\sqrt{\delta}} - 2$. It should also be clear that the minimal possible N is at most $\frac{4}{\delta\sqrt{\delta}} - 1$. Hence,

$$\sigma_{n-1}\{L(F_\theta, F) > \delta\} \le \left(\frac{8}{\delta^{3/2}} - 2\right) e^{-(n-1)\delta^4/8},$$

and the desired inequality (1.4) will follow from $\frac{8}{\delta^{3/2}} - 2 \le 4n^{3/8} e^{-\delta^4/8}$. In view of the assumption $\delta \le 1$, the latter is implied by $\frac{8}{\delta^{3/2}} - 2 \le 4n^{3/8} e^{-1/8}$. Here, for $n\delta^4 \ge 8$, we are reduced to $\frac{8}{\delta^{3/2}} - 2 \le \frac{4 \cdot 8^{3/8}}{\delta^{3/2}} e^{-1/8}$ which in the worst case $\delta = 1$

is easily verified to hold true. In the case $n\delta^4 \le 8$, there is nothing to prove since then

$$\sigma_{n-1}\{L(F_\theta, F) > \delta\} \le 1 \le 4n^{3/8}\, e^{-n\delta^4/8}. \qquad \square$$

Theorem 1.1 is proved and, as its immediate consequence, we obtain the corollary.

COROLLARY 2.1. *Under* (1.1), $\inf_{\theta\in S^{n-1}} L(F_\theta, F) \le C\,(\frac{\log n}{n})^{1/4}$, *where* C *is a universal constant.*

REMARK 2.2. As already mentioned, given two distribution functions G and H with $\int_{-\infty}^{+\infty} t^2\, dG(t) = \int_{-\infty}^{+\infty} t^2\, dH(t) = 1$, the Kantorovich–Rubinshtein distance $\kappa(G, H)$ and the Lévy metric $L(G, H)$ are related by

$$\kappa(G, H) \le C L(G, H)^{1/2}. \tag{2.6}$$

Indeed, put $\delta = L(G, H)$, $0 \le \delta \le 1$, so that $H(t-\delta) - \delta \le G(t) \le H(t+\delta) + \delta$ for all $t \in \mathbf{R}$. Hence, for all $t \in \mathbf{R}$,

$$|G(t) - H(t)| \le \big(G(t) - G(t-\delta)\big) + \big(H(t) - H(t-\delta)\big) + \delta.$$

Integrating this inequality in $-b < t < b$, we get $\int_{-b}^{b} |G(t) - H(t)|\, dt \le 2(b+1)\delta$. From Chebyshev's inequalities $G(-t) + (1 - G(t)) \le 1/t^2$, $H(-t) + (1 - H(t)) \le 1/t^2$, $t > b$, one also derives $\int_{-\infty}^{-b} |G(t) - H(t)|\, dt + \int_{b}^{+\infty} |G(t) - H(t)|\, dt \le 2/b$. As a result, $\kappa(G, H) \le 2(b+1)\delta + 2/b$. Optimizing over $b \ge 1$ yields $\kappa(G, H) \le 2(2\sqrt{\delta} + \delta) \le 6\sqrt{\delta}$.

Thus, the inequality (2.6) holds true with $C = 6$, and we have in particular

$$\kappa(F_\theta, F) \le 6L(F_\theta, F)^{1/2} \qquad \text{for all } \theta \in S^{n-1}.$$

Therefore, by Theorem 1.1,

$$\sigma_{n-1}\{\theta : \kappa(F_\theta, F) \ge \delta\} \le 4n^{3/8}\, e^{-cn\delta^8}, \qquad \delta > 0,$$

for some universal constant $c > 0$.

REMARK 2.3. The Lévy metric can also be related to the uniform distance by

$$\|G - H\|_\infty \le (1 + C)\, L(G, H), \tag{2.7}$$

provided, however, that H has a density H' bounded by a constant C. This estimate may be applied in particular to our case $G = F_\theta$, $H = F$. Recall that F represents the distribution function of $\theta_1|X|$, where θ_1 is the first coordinate of a random vector independent of $|X|$ and uniformly distributed over S^{n-1}. Hence, denoting by Φ_n and $\varphi_n = \Phi_n'$ the distribution function and the density of $\sqrt{n}\,\theta_1$, respectively, we see that, for all $t \in \mathbf{R}$,

$$F(t) = \mathbf{E}\Phi_n\left(\frac{\sqrt{n}}{|X|}t\right), \qquad F'(t) = \mathbf{E}\frac{\sqrt{n}}{|X|}\varphi_n\left(\frac{\sqrt{n}}{|X|}t\right). \tag{2.8}$$

The density φ_n is symmetric, is log-concave, and attains its maximum at zero. Hence, $F'(t)$ is maximized at $t = 0$ so that $C \equiv \|F'\|_\infty = \varphi_n(0)\mathbf{E}\frac{\sqrt{n}}{|X|}$. Moreover, as is well known, for all $t \in \mathbf{R}$, $\varphi_n(t)$ converges pointwise, as $n \to \infty$, to $\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$. Hence, $\varphi_n(0) \le 1$ at least for large enough n. More precisely, denoting by s_{n-1} the $(n-1)$-dimensional volume of S^{n-1}, we have $\varphi_n(0) = s_{n-1}/(s_{n-2}\sqrt{n}) = 1/(2\sqrt{n}\int_0^{\pi/2}\cos^{n-2}u\,du)$ from which it follows that $\varphi_n(0) \le \frac{1}{2}$ for all $n \ge 1$ (cf. [13], page 5). Note also that, by Jensen's inequality, since $\mathbf{E}|X|^2 = n$, necessarily $\mathbf{E}\frac{\sqrt{n}}{|X|} \ge 1$. As a result, we obtain the following corollary from Theorem 1.1.

COROLLARY 2.4. *Under* (1.1), *for all* $\delta > 0$, *there is a set* $\Theta_\delta \subset S^{n-1}$ *of measure* $\sigma_{n-1}(\Theta_\delta) \ge 1 - 4n^{3/8}e^{-n\delta^4/8}$ *such that, for all* $\theta \in \Theta_\delta$,

$$\sup_{t\in\mathbf{R}}|F_\theta(t) - F(t)| \le 2\delta\,\mathbf{E}\frac{\sqrt{n}}{|X|}.$$

Now, in order to connect the above estimates with the central limit theorem, assume inequality (1.6),

$$\mathbf{P}\left\{\left|\frac{|X|}{\sqrt{n}} - 1\right| \ge \varepsilon_n\right\} \le \varepsilon_n \tag{2.9}$$

holds true for a certain (small) number $\varepsilon_n \ge 0$. Then we have the following statement which immediately implies Corollary 1.3.

COROLLARY 2.5. *Under* (1.1) *and* (2.9), *for all* $\delta > 0$,

$$\sigma_{n-1}\left\{\sup_{t\in\mathbf{R}}|F_\theta(t) - \Phi(t)| \ge 4\varepsilon_n + \delta\right\} \le 4n^{3/8}\,e^{-cn\delta^4}, \tag{2.10}$$

where c is a universal constant.

For example, one may treat the sequence

$$X(\omega) = \big(1, \sqrt{2}\cos(\omega), \sqrt{2}\sin(\omega), \dots, \sqrt{2}\cos(n\omega), \sqrt{2}\sin(n\omega)\big),$$

$$-\pi < \omega < \pi,$$

as a random vector in $\mathbf{R}^{2n+1}$ with respect to the uniform distribution $\mathbf{P}$ on the interval $(-\pi, \pi)$. In this case $\varepsilon_n = 0$, so the most of the trigonometric polynomials

$T_\theta = \theta_0 + \sqrt{2}\sum_{k=1}^n (\theta_{2k-1}\cos(k\omega) + \theta_{2k}\sin(k\omega))$ with norm $\|T_\theta\|_{L^2(\mathbf{P})} = 1$ have distribution functions which are very close to Φ.

For the proof of Corollary 2.5, we need two simple claims.

LEMMA 2.6. *Let Ψ be a symmetric around zero, unimodal distribution function. Then $|\Psi(\alpha t) - \Psi(t)| \le |\alpha - 1|$ for all $\alpha > 0$ and $t \in \mathbf{R}$.*

PROOF. By symmetry, let $t > 0$. In addition, Ψ can be assumed to have a continuous nonincreasing density $\psi(t)$ in $t > 0$. For $\alpha \ge 1$, the desired inequality becomes $\Psi(\alpha t) - \Psi(t) \le \alpha - 1$ which turns into equality at $\alpha = 1$. Since the left-hand side is concave in α, it suffices to compare the derivatives at $\alpha = 1$: we arrive at $t\psi(t) \le 1$. The latter follows from the stronger bound $\frac{1}{2} \ge \int_0^t \psi(s)\,ds \ge t\psi(t)$. When $0 \le \alpha \le 1$, the desired inequality becomes $\Psi(t) - \Psi(\alpha t) \le 1 - \alpha$. It is true at the end points $\alpha = 0$ and $\alpha = 1$, so it holds for all α since the left-hand side is convex in α. $\square$

LEMMA 2.7. $\|\Phi_n - \Phi\|_\infty \le \frac{4}{\sqrt{n}}$, *for all $n \ge 1$.*

PROOF. Assume $n \ge 16$ and introduce an i.i.d. $N(0,1)$-sequence $Z = (Z_1, \dots, Z_n)$ so that $\Phi_n(t) = \mathbf{P}\{Z_1 \le \frac{|Z|}{\sqrt{n}} t\}$, $\Phi(t) = \mathbf{P}\{Z_1 \le t\}$, where $|Z| = (Z_1^2 + \cdots + Z_n^2)^{1/2}$. Then, for any $t > 0$,

$$|\Phi_n(t) - \Phi(t)| = \mathbf{P}\left\{t \le Z_1 \le \frac{|Z|}{\sqrt{n}} t\right\} + \mathbf{P}\left\{\frac{|Z|}{\sqrt{n}} t \le Z_1 \le t\right\}. \tag{2.11}$$

If $t^2 \ge n$, the second probability vanishes, while the first is bounded by $1 - \Phi(\sqrt{n}) < \frac{1}{\sqrt{n}}$. So, assume $t^2 < n$, in which case (2.11) becomes

$$|\Phi_n(t) - \Phi(t)| = \mathbf{E}\left|\Phi\left(\frac{\xi}{\sqrt{n - t^2}} t\right) - \Phi(t)\right|, \tag{2.12}$$

where $\xi = (Z_2^2 + \cdots + Z_n^2)^{1/2}$. Since $|\Phi(\frac{\xi}{\sqrt{n-t^2}} t) - \Phi(t)| \le \frac{1}{2}e^{-t^2/2} + \frac{1}{2}e^{-t^2\xi^2/2(n-t^2)}$, we get

$$\begin{aligned} |\Phi_n(t) - \Phi(t)| &\le \frac{1}{2}e^{-t^2/2} + \frac{1}{2}(\mathbf{E}e^{-t^2 Z_2^2/2(n-t^2)})^{n-1} \\ &= \frac{1}{2}e^{-t^2/2} + \frac{1}{2}\left(1 - \frac{t^2}{n}\right)^{(n-1)/2}. \end{aligned}$$

If $t^2 \ge \frac{\sqrt{n}}{2}$, the right-hand side is at most $\frac{1}{2}e^{-\sqrt{n}/4} + \frac{1}{2}(1 - \frac{1}{2\sqrt{n}})^{(n-1)/2} < \frac{2}{\sqrt{n}}$.

At last, assume $t^2 < \frac{\sqrt{n}}{2}$. To estimate the right-hand side of (2.12) in this case, by Lemma 2.6, we get

$$\left|\Phi\left(\frac{\xi}{\sqrt{n-t^2}}\,t\right) - \Phi(t)\right| \le \frac{|\xi - \sqrt{n-t^2}|}{\sqrt{n-t^2}} \le \frac{|\xi^2 - (n-t^2)|}{n-t^2}$$
$$\le \frac{|\xi^2 - (n-1)|}{n-t^2} + \frac{|t^2 - 1|}{n-t^2} \le \frac{2|\xi^2 - (n-1)|}{n} + \frac{1}{\sqrt{n}},$$

where we used the assumption $t^2 < \frac{\sqrt{n}}{2} \le \frac{n}{2}$ on the last step. It remains to note that $\mathbf{E}|\xi^2 - (n-1)| \le \sqrt{\operatorname{Var}(\xi^2)} = \sqrt{2n}$. $\square$

PROOF OF COROLLARY 2.5. Assume $\varepsilon_n \in (0, \frac{1}{3}]$ [otherwise the inequality (2.10) is immediate]. In case $|\frac{|X|}{\sqrt{n}} - 1| \le \varepsilon_n$, by Lemma 2.6, for all $t \in \mathbf{R}$,

$$\left|\Phi_n\left(\frac{t}{|X|/\sqrt{n}}\right) - \Phi_n(t)\right| \le \left|\frac{1}{|X|/\sqrt{n}} - 1\right| \le \frac{\varepsilon_n}{1-\varepsilon_n} \le 1.5\varepsilon_n,$$

so $|\Phi_n(\frac{t}{|X|/\sqrt{n}}) - \Phi(t)| \le \|\Phi_n - \Phi\|_\infty + 1.5\varepsilon_n$. Using $|\Phi_n(\frac{t}{|X|/\sqrt{n}}) - \Phi(t)| \le 1$ for the remaining case $|\frac{|X|}{\sqrt{n}} - 1| > \varepsilon_n$, and recalling the definition (2.8) and the assumption (2.9), we get

$$\|F - \Phi\|_\infty \le \mathbf{E}\sup_t \left|\Phi_n\left(\frac{t}{|X|/\sqrt{n}}\right) - \Phi(t)\right| \le \|\Phi_n - \Phi\|_\infty + 2.5\varepsilon_n.$$

Thus, by Lemma 2.7, $L(F, \Phi) \le \|F - \Phi\|_\infty \le \frac{4}{\sqrt{n}} + 2.5\varepsilon_n$. Combaining this estimate with Theorem 1.1, we therefore obtain that, for any $\delta > 0$, except for a set of directions of σ_{n-1}-measure at most $4n^{3/8}e^{-n\delta^4/8}$, one has $L(F_\theta, \Phi) \le \delta + \frac{4}{\sqrt{n}} + 2.5\varepsilon_n$. Hence, by (2.7) applied to $G = F_\theta$ and $H = \Phi$,

$$\|F_\theta - \Phi\|_\infty \le \left(1 + \frac{1}{\sqrt{2\pi}}\right)\left(\delta + \frac{4}{\sqrt{n}} + 2.5\varepsilon_n\right) < 2\delta + \frac{6}{\sqrt{n}} + 4\varepsilon_n.$$

Here, since only the values $\delta \ge \text{const} \cdot n^{-1/4}$ are of interest, the term $\frac{6}{\sqrt{n}}$ can be absorbed by δ at the expense of a suitable constant c in the exponential $e^{-cn\delta^4}$ in (2.10). This proves Corollary 2.5. $\square$

One of the natural ways to bound the optimal value of the parameter ε_n in (2.9) is to use Chebyshev's inequality $\mathbf{P}\{|\frac{|X|}{\sqrt{n}} - 1| \ge \varepsilon\} \le \mathbf{P}\{|\frac{|X|^2}{n} - 1| \ge \varepsilon\} \le \frac{\operatorname{Var}(|X|^2)}{n^2\varepsilon^2}$. Consequently, one can take in Corollary 2.5,

$$\varepsilon_n = \frac{\operatorname{Var}(|X|^2)^{1/3}}{n^{2/3}}.$$

For example, when $\mathrm{Var}(|X|^2) = O(n)$, ε_n would be of order at most $\frac{1}{n^{1/3}}$ and then it will again be absorbed by δ in (2.10). In general,

$$\mathrm{Var}(|X|^2) = \sum_{k=1}^{n} \mathrm{Var}(X_k^2) + 2 \sum_{1 \le k < j \le n} \mathrm{cov}(X_k^2, X_j^2)$$

and one may get a previous estimate for ε_n if all the covariances $\mathrm{cov}(X_k^2, X_j^2)$ are not positive [in addition to $\mathrm{Var}(X_k^2) = O(1)$]. As shown by Ball and Perissinaki in [3] and then applied in [1], this is the case for all X uniformly distributed over ℓ^p-balls in $\mathbf{R}^n$.

There is another general condition leading to the property $\mathrm{Var}(|X|^2) = O(n)$. Under mild integrability assumptions, the requirement that the most of linear functionals $\langle\theta, X\rangle$ have distribution functions close to Φ implies that, for most of them, $\mathbf{E}\langle\theta, X\rangle^4 \le 3 + o(1)$. In turn, the latter can be shown to yield closeness to normality. In this connection, it is worthwile to mention the following interesting comparison principle: if

$$\mathbf{E}\langle\theta, X\rangle^4 \le \mathbf{E}\langle\theta, Z\rangle^4 = 3 \qquad \text{for all } \theta \in S^{n-1},$$

where Z is a standard normal vector in $\mathbf{R}^n$, then $\mathrm{Var}(|X|^2) \le \mathrm{Var}(|Z|^2) = 2n$. Indeed, in general $\int \langle\theta, X\rangle^4 \, d\sigma_{n-1}(\theta) = \frac{|X|^4}{n(n+2)}$, and in order to conclude, in addition to the above hypothesis, we need only the assumption $\mathbf{E}|X|^2 = n$.

3. Sharpening in the log-concave case. Here we involve the additional assumption that the distribution μ of the random vector X has a log-concave density, say, $p(x)$. Together with the basic isotropy condition (1.1),

$$\mathbf{E}\langle\theta, X\rangle^2 = \int_{\mathbf{R}^n} \langle\theta, x\rangle^2 p(x)\, dx = |\theta|^2, \tag{3.1}$$

we also assume that the vector X and thus all linear functionals $\langle\theta, X\rangle$ ($\theta \in \mathbf{R}^n$) have mean zero,

$$\mathbf{E}\langle\theta, X\rangle = \int_{\mathbf{R}^n} \langle\theta, x\rangle p(x)\, dx = 0. \tag{3.2}$$

The measure μ is supported by a convex set K in $\mathbf{R}^n$, bounded or not, on which the function p is positive and the function $\log p(x)$ is concave. In particular, when K is symmetric and bounded, the measure μ can represent the normalized Lebesgue measure on K. A striking observation made in [1] in this special situation is that, under (3.1) and (3.2), the functions

$$f(\theta) = \mathbf{E}g(\langle\theta, X\rangle),$$

being considered for non-Lipschitz $g(x) = \mathbb{1}_{\{\langle\theta, x\rangle \le t\}}$, still have on S^{n-1} finite Lipschitz seminorms bounded by a universal constant. For any such g,

$$f_t(\theta) = F_\theta(t) = \mathbf{P}\{\langle\theta, X\rangle \le t\}, \qquad \theta \in S^{n-1},$$

represents the values of the distribution function of $\langle\theta, X\rangle = \sum_{k=1}^n \theta_k X_k$ at a given fixed point $t \in \mathbf{R}$. Here we refine and extend this property to the class of log-concave distributions.

PROPOSITION 3.1. *Under* (3.1) *and* (3.2), *for every* $t \in \mathbf{R}$, *the function* f_t *has on the unit sphere a finite Lipschitz seminorm satisfying*

$$\|f_t\|_{\mathrm{Lip}} \le Ce^{-c|t|},$$

where C *and* c *are positive universal constants.*

The proof requires some preparations. We need some estimates for one-dimensional log-concave probability densities $\rho = \rho(t)$ which are not necessarily symmetric around the origin. The behavior of ρ is mainly determined by two parameters—by the median $m = m(\zeta)$ and by the value $\rho(m)$ at this point, or, equivalently, by the expectation $\mathbf{E}\zeta$ and by the variance $\mathrm{Var}(\zeta) = \mathbf{E}\zeta^2 - (\mathbf{E}\zeta)^2$ where ζ is a random variable having the density ρ. Often, the last two quantities are more convenient. In particular, we have:

LEMMA 3.2. *If the random variable* ζ *has a log-concave density* ρ *with expectation* $a = \mathbf{E}\zeta$ *and standard deviation* $\sigma = \sqrt{\mathrm{Var}(\zeta)}$, *then for all* $t \in \mathbf{R}$,

$$\rho(t) \le \frac{C}{\sigma} e^{-c|t-a|/\sigma}, \tag{3.3}$$

where C *and* c *are universal positive constants.*

PROOF. First we show that, for all $t \in \mathbf{R}$,

$$\rho(t) \le 6\rho(m)e^{-\rho(m)\,|t-m|}. \tag{3.4}$$

We may assume that $t > 0$ and $m = 0$ so that $\mathbf{P}\{\zeta \ge 0\} = \int_0^{+\infty} \rho(x)\,dx = \frac{1}{2}$.

Consider the value $M(t) = \sup_u e^{-u(t)}$ where the supremum is taken over all convex functions u on $[0, +\infty)$ such that $e^{-u(0)} = \rho(0)$ and

$$\int_0^{+\infty} e^{-u(x)}\,dx \le \tfrac{1}{2}. \tag{3.5}$$

By convexity, for any such u, all these constrains will be fulfilled for the function u_0 which is linear in $[0, t]$, is equal to $+\infty$ on $(t, +\infty)$, and which has values $u_0(0) = u(0)$, $u_0(t) = u(t)$. Hence, while computing $M(t)$, we may restrict ourselves to functions of the form $u_\alpha(x) = -\log\rho(0) + \alpha x$, $0 \le x \le t$ [and $u(x) = +\infty$ for $x > t$]. The range of α is determined by (3.5), that is, by $\rho(0)\,\frac{1-e^{-\alpha t}}{\alpha} \le \frac{1}{2}$. Since the function $\frac{1-e^{-y}}{y}$ is positive and decreases on the whole real line, we may conclude that

$$M(t) = \rho(0)e^{-y} \qquad \text{where } y \text{ is the solution of } \frac{1-e^{-y}}{y} = \frac{1}{2\rho(0)t}.$$

First consider the case $\rho(0)t \geq 1$. Then $y > 1$ since otherwise $\frac{1-e^{-y}}{y} \geq 1-e^{-1} > \frac{1}{2}$ while $\frac{1}{2\rho(0)t} \leq \frac{1}{2}$. Thus, $\frac{1}{2\rho(0)t} = \frac{1-e^{-y}}{y} \geq \frac{1}{2y}$ from which it follows that $y \geq \rho(0)t$. Therefore,

$$\rho(t) \leq M(t) \leq \rho(0)e^{-\rho(0)t}.$$

To treat the case $\rho(0)t \leq 1$, just note that, by log-concavity, if $G(x) = \int_{-\infty}^{x} \rho(z)\, dz$ is the distribution function of ζ, and if G^{-1} denotes its inverse, then $I(s) = \rho(G^{-1}(s))$ is concave on $(0, 1)$. Therefore, $\sup_x \rho(x) = \sup_s I(s) \leq 2I(1/2) = 2\rho(0)$. In particular, $\rho(t) \leq 2\rho(0) \leq 6\rho(0)\, e^{-\rho(0)t}$. In both cases, we obtain the desired estimate (3.4).

Now, starting from (3.4), we immediately obtain that $|\mathbf{E}\zeta - m| \leq 12/\rho(m)$. Hence, applying (3.4) once more, we get

$$\rho(t) \leq 6e^{12}\rho(m)e^{-\rho(m)|t-a|}. \tag{3.6}$$

As for the value $\rho(m)$, it can be related to the variance by

$$\frac{1}{12\,\mathrm{Var}(\zeta)} \leq \rho(m)^2 \leq \frac{1}{2\,\mathrm{Var}(\zeta)}. \tag{3.7}$$

These inequalities, for symmetric log-concave densities on the real line, were proved by Ball [2]. The general nonsymmetric case was considered in [4]. Combining (3.6) with (3.7), we arrive at (3.3) with $c = 1/\sqrt{12}$ and $C = 6\,e^{12}/\sqrt{2}$ (the constants are far from being optimal). Lemma 3.2 has been proved. $\square$

LEMMA 3.3. *If a random variable ξ defined on some probability space* $(\Omega, \mathbf{P})$ *has a log-concave density*, *then*

$$\frac{1}{e} \leq \mathbf{P}\{\xi \leq \mathbf{E}\xi\} \leq 1 - \frac{1}{e}. \tag{3.8}$$

Both inequalities are sharp since on the right there is equality for ξ having the standard exponential distribution $[\Omega = (0, +\infty),\ d\mathbf{P}(x) = e^{-x}\, dx,\ \xi(x) = x]$. Note that the left inequality in (3.8) is obtained from the right by applying it to $-\xi$.

The left inequality in (3.8) can be viewed as a "log-concave" version of the known fact saying that for any convex body K in $\mathbf{R}^n$ and any half-space H with ∂H passing through the centroid of K, one has $\mathrm{vol}_n(K \cap H) \geq \frac{1}{e}\mathrm{vol}_n(K)$. Moreover, in the space of the fixed dimension, the factor $1/e$ can be replaced with $(\frac{n}{n+1})^n$. This property was first observed Grünbaum [7] and Hammer [9] with similar proofs based on the Schwarz symmetrization. For completeness, we give below a simple straigthforward argument leading to (3.8).

PROOF OF LEMMA 3.3. The distribution of ξ is concentrated on an interval (a, b), finite or not, where it has a positive log-concave density, say, g. Introduce the distribution function $G(x) = \mathbf{P}\{\xi \leq x\} = \int_a^x g(z)\, dz$, $a < x < b$, and its inverse

$G^{-1}:(0,1)\to(a,b)$. Fix a point $\alpha\in(0,1)$. As in the proof of Lemma 3.2, we use the property that the function $I=g(G^{-1})$ is concave on $(0,1)$ which implies in particular that

$$\frac{I(u)}{1-u}\le\frac{I(\alpha)}{1-\alpha}\qquad\text{for } u\in(0,\alpha]\quad\text{and}\quad\frac{I(u)}{1-u}\ge\frac{I(\alpha)}{1-\alpha}\qquad\text{for } u\in[\alpha,1).$$

Since $G^{-1}(s)-G^{-1}(\alpha)=\int_\alpha^s\frac{du}{I(u)}$, $0<s<1$, we may write

$$\begin{aligned}\mathbf{E}\xi&=\int_0^1 G^{-1}(s)\,ds=\int_0^1\left[G^{-1}(\alpha)+\int_\alpha^s\frac{du}{I(u)}\right]ds\\&=G^{-1}(\alpha)+\int_\alpha^1\frac{1-u}{I(u)}\,du-\int_0^\alpha\frac{u}{I(u)}\,du\\&\le G^{-1}(\alpha)+\frac{1-\alpha}{I(\alpha)}\int_\alpha^1 du-\frac{1-\alpha}{I(\alpha)}\int_0^\alpha\frac{u}{1-u}\,du\\&=G^{-1}(\alpha)+\frac{(1-\alpha)(1+\log(1-\alpha))}{I(\alpha)}.\end{aligned}$$

Thus, $\mathbf{E}\xi\le G^{-1}(\alpha)+\frac{(1-\alpha)(1+\log(1-\alpha))}{I(\alpha)}$ for all $\alpha\in(0,1)$. Taking $\alpha=1-\frac{1}{e}$, we arrive at $\mathbf{E}\xi\le G^{-1}(1-\frac{1}{e})$ which is equivalent to $G(\mathbf{E}\xi)\le 1-\frac{1}{e}$. This is exactly the right inequality in (3.8). $\square$

LEMMA 3.4. *If a random variable ξ defined on some Lebesgue probability space $(\Omega,\mathbf{P})$ has a log-concave density, then for every measurable $A\subset\Omega$,*

$$\int_A|\xi|\,d\mathbf{P}\ge\frac{\mathbf{P}(A)^2}{4\sqrt{2}}\sqrt{\mathrm{Var}(\xi)},\qquad\int_A|\xi|^2\,d\mathbf{P}\ge\frac{\mathbf{P}(A)^3}{24}\mathrm{Var}(\xi).\tag{3.9}$$

PROOF. We use the same notation (a,b), g, G, G^{-1}, I as in the proof of Lemma 3.3. We may assume that $p=\mathbf{P}(A)>0$ and that $0<G(0)<1$ [since otherwise inequality (3.9) can be made stronger by adding a constant to ξ].

The assumption on the probability space means that $(\Omega,\mathbf{P})$ can be transformed to the interval $(0,1)$ with a Borel probability measure which in turn, since the distribution of ξ has no atom, can be assumed to be the normalized Lebesgue measure on $(0,1)$ (cf. [17]). Thus, we can start with $\Omega=(0,1)$ and $\xi(s)=G^{-1}(s)$.

By $0<G(0)<1$, and since the function G^{-1} is increasing, we have $G^{-1}(\alpha)=0$, for some $\alpha\in(0,1)$. Hence, $|G^{-1}(s)|=|\int_\alpha^s\frac{du}{I(u)}|\ge\frac{1}{\sup_s I(s)}|s-\alpha|$. Therefore, for each measurable set $A\subset(0,1)$,

$$\int_A|\xi|\,d\mathbf{P}=\int_A|G^{-1}(s)|\,ds\ge\frac{1}{\sup_s I(s)}\int_A|s-\alpha|\,ds.\tag{3.10}$$

It should also be clear that within the class of all measurable sets $A\subset(0,1)$ of Lebesgue measure $\mathrm{mes}(A)=p$, the integral $\int_A|s-\alpha|\,ds$ attains its minimum at some interval (c, d) of length p containing the point α. Moreover, in the worst situation, one can take $\alpha = 1/2$ and $c = \alpha - \frac{p}{2}$, $d = \alpha + \frac{p}{2}$, which yields $\int_A |s - \alpha|\, ds = \frac{p^2}{4}$.

As already noted in the proof of Lemma 3.2, $\sup_s I(s) \leq 2I(1/2) = 2g(m)$ where m is the median of ξ. Therefore, by (3.7), $\sup_s I(s) \leq \sqrt{\frac{2}{\mathrm{Var}(\xi)}}$. Together with (3.10) this gives the first inequality of Lemma 3.4. A similar argument gives the second inequality (although it can also be derived, with a worse constant, from the first one by applying Cauchy's inequality). $\square$

Now, we turn to the original isotropic density $p(x)$ of the measure μ on $\mathbf{R}^n$ appearing in Proposition 3.1. We write the points in the form $x = (y, t)$ where $y \in \mathbf{R}^{n-1}$ and $t \in \mathbf{R}$.

LEMMA 3.5. *Under* (3.1) *and* (3.2), *for every $t \in \mathbf{R}$ and every unit vector l in $\mathbf{R}^{n-1}$,*

$$\int_{\mathbf{R}^{n-1}} |\langle l, y\rangle| p(y, t)\, dy \leq Ce^{-c|t|},$$

where C and c are positive universal constants.

PROOF. Set $l^+(y) = \max\{\langle l, y\rangle, 0\}$. This function is log-concave on $\mathbf{R}^{n-1}$, so, by Prékopa's theorem [16], the function

$$u(t) = \int_{\mathbf{R}^{n-1}} l^+(y)\, p(y, t)\, dy$$

is log-concave on $\mathbf{R}$ (although it might be nonsymmetric even if the function p is symmetric around the origin). We consider the linear functions $\xi(y, t) = \langle l, y\rangle$ and $\eta(y, t) = t$ as log-concave random variables with respect to μ such that $\mathbf{E}\xi = \mathbf{E}\eta = 0$ and $\mathbf{E}\xi^2 = \mathbf{E}\eta^2 = 1$ according to (3.1) and (3.2). Thus,

$$\begin{aligned} &\int_{-\infty}^{+\infty} u(t)\, dt = \mathbf{E}\xi^+, \qquad \int_{-\infty}^{+\infty} tu(t)\, dt = \mathbf{E}\xi^+\eta, \\ &\int_{-\infty}^{+\infty} t^2 u(t)\, dt = \mathbf{E}\xi^+\eta^2, \end{aligned} \tag{3.11}$$

where $\xi^+ = \max\{\xi, 0\}$. Now, introduce the log-concave probability density on the line

$$\rho(t) = \frac{u(t)}{\mathbf{E}\xi^+}, \qquad t \in \mathbf{R}. \tag{3.12}$$

Since we need to estimate the function $u(t)$ from above, we apply Lemma 3.2 to ρ. Let ζ be a random variable with this density. Then, by (3.11) and (3.12),

$$\mathbf{E}\zeta = \frac{\mathbf{E}\xi^+\eta}{\mathbf{E}\xi^+}, \qquad \mathrm{Var}(\zeta) = \frac{\mathbf{E}\xi^+\eta^2\mathbf{E}\xi^+ - (\mathbf{E}\xi^+\eta)^2}{(\mathbf{E}\xi^+)^2}.$$

Introduce the half-space $A = \{\xi > 0\}$ and the normalized restriction ν of measure μ to A, that is, $\nu(B) = \mu(B \cap A)/\mu(A)$. Then, the above variance can also be written as

$$\mathrm{Var}(\zeta) = \frac{\mu(A)}{\mathbf{E}\xi^+} \int (\eta - a)^2 \xi \, d\nu = \frac{\mu(A)}{\mathbf{E}\xi^+} \|(\eta - a)^2 \xi\|_1, \tag{3.13}$$

where $a = \mathbf{E}\zeta$ and where we used notation $\|\psi\|_\alpha = (\int |\psi|^\alpha \, d\nu)^{1/\alpha}$ with respect to measure ν and for $\alpha = 1$. To bound $\|(\eta - a)^2 \xi\|_1$ from below, we consider the quantity $\|\psi\|_\alpha$ also for $\alpha = 0$ in which case $\|\psi\|_0 = \exp \int \log |\psi| \, d\nu$ and thus

$$\|(\eta - a)^2 \xi\|_1 \geq \|(\eta - a)^2 \xi\|_0 = \| \, |\eta - a| \, \|_0^2 \, \|\xi\|_0. \tag{3.14}$$

Now, we apply a theorem (due to Latala [12]; cf. also [8, 4] for different proofs) asserting that, for every norm ψ on a linear space equipped with a log-concave measure, $\|\psi\|_\alpha$-norms are equivalent to $\|\psi\|_0$ in the sense that $\|\psi\|_0 \geq c_\alpha \|\psi\|_\alpha$ with some positive constants c_α depending on $\alpha \geq 0$, only. Since the measure ν is log-concave, we may continue (3.14) to get

$$\|(\eta - a)^2 \xi\|_1 \geq c_2^3 \|\eta - a\|_2^2 \|\xi\|_2. \tag{3.15}$$

Recall that $\|\eta - a\|_2^2 = \frac{1}{\mu(A)} \int_A (\eta - a)^2 \, d\mu$ and $\|\xi\|_2^2 = \frac{1}{\mu(A)} \int_A \xi^2 \, d\mu$. Applying Lemma 3.4 to random variables $\eta - a$ and ξ on the space $(\Omega, \mathbf{P}) = (\mathbf{R}^n, \mu)$ and recalling that $\mathrm{Var}(\eta) = \mathrm{Var}(\xi) = 1$, we obtain that

$$\|\eta - a\|_2^2 \geq \frac{(\mu(A))^2}{24}, \qquad \|\xi\|_2^2 \geq \frac{(\mu(A))^2}{24}.$$

Together with (3.13) and (3.15), this gives $\mathrm{Var}(\zeta) \geq \frac{c_2^3 \mu(A)^4}{24\sqrt{24}\mathbf{E}\xi^+}$. To continue, we have $\mathbf{E}\xi^+ \leq \sqrt{\mathbf{E}\xi^2} = 1$, and by Lemma 3.3, $\mu(A) \geq \frac{1}{e}$. Hence,

$$\sigma^2 \equiv \mathrm{Var}(\zeta) \geq \frac{c_2^3}{24\sqrt{24}e^4}.$$

Now, in order to bound the expectation $a = \mathbf{E}\zeta = \frac{\mathbf{E}\xi^+ \eta}{\mathbf{E}\xi^+}$, just note that $|\mathbf{E}\xi^+ \eta|^2 \leq \mathbf{E}\xi^2 \mathbf{E}\eta^2 = 1$, and once more by Lemmas 3.4 and 3.3, we have

$$\mathbf{E}\xi^+ = \int_A \xi \geq \frac{\mu(A)^2}{4\sqrt{2}} \geq \frac{1}{4\sqrt{2}e^2}.$$

Thus, $|a| \leq 4\sqrt{2}e^2$. Therefore, by Lemma 3.2,

$$\rho(t) \leq \frac{C}{\sigma} e^{-c|t-a|/\sigma} \leq \frac{C e^{c|a|/\sigma}}{\sigma} e^{-c|t|/\sigma} \leq C' e^{-c'|t|}$$

since we have universal bounds for σ and a. It remains to note that $u(t) = \rho(t)\mathbf{E}\xi^+ \leq \rho(t)$.

At last, replacing l with $-l$, we get the same estimate for $\int_{\mathbf{R}^{n-1}} l^-(y)p(y,t)\,dy$, $l^-(y) = \max\{-l(y), 0\}$, so Lemma 3.5 follows. $\square$

PROOF OF PROPOSITION 3.1. The statement is equivalent to saying that, for every $\theta_0 \in S^{n-1}$, the modulus of the (inner) gradient of f_t at the point θ_0,

$$|\nabla_{S^{n-1}} f_t(\theta_0)| = \limsup_{\theta \to \theta_0,\, \theta \in S^{n-1}} \frac{|f_t(\theta) - f_t(\theta_0)|}{|\theta - \theta_0|},$$

satisfies $|\nabla_{S^{n-1}} f_t(\theta_0)| \leq Ce^{-c|t|}$. Due to condition (3.1), one may assume that θ_0 is the last vector $e_n = (0, \ldots, 0, 1)$ in the canonical orthonormal basis $(e_1, \ldots, e_n)$ of $\mathbf{R}^n$. Then

$$|\nabla_{S^{n-1}} f_t(\theta_0)| = \left(\sum_{k=1}^{n-1} \left|\frac{\partial f_t(\theta_0)}{\partial \theta_k}\right|^2\right)^{1/2} = \sup_{|l|=1} |\langle \nabla f_t(\theta_0), l\rangle|,$$

where $\frac{\partial f_t(\theta_0)}{\partial \theta_k}$ are usual partial derivatives of f_t, and the supremum is taken over all unit vectors $l = (l_1, \ldots, l_n)$ in $\mathbf{R}^n$ with $l_n = 0$ (it is readily verified that f_t is differentiable on the whole space $\mathbf{R}^n$ except the origin point).

Fix $k = 1, \ldots, n-1$. The two-dimensional random vector (X_k, X_n) has a log-concave density on $\mathbf{R}^2$, say, $p_k = p_k(x_k, x_n)$. Therefore, for every $\varepsilon > 0$,

$$\begin{aligned}
&f_t(\theta_0 + \varepsilon e_k) - f_t(\theta_0) \\
&\quad = \mathbf{P}\{\varepsilon X_k + X_n \leq t\} - \mathbf{P}\{X_n \leq t\} \\
&\quad = \int_{-\infty}^{0} \left[\int_{t}^{t-\varepsilon x_k} p_k(x_k, x_n)\,dx_n\right] dx_k - \int_{0}^{+\infty} \left[\int_{t-\varepsilon x_k}^{t} p_k(x_k, x_n)\,dx_n\right] dx_k \\
&\quad = \varepsilon \int_{-\infty}^{0} \left[\int_{0}^{-x_k} p_k(x_k, t+\varepsilon u)\,du\right] dx_k \\
&\qquad - \varepsilon \int_{0}^{+\infty} \left[\int_{-x_k}^{0} p_k(x_k, t+\varepsilon u)\,du\right] dx_k \\
&\quad = \big(\varepsilon + o(\varepsilon)\big) \int_{-\infty}^{0} -x_k p_k(x_k, t)\,dx_k - \big(\varepsilon + o(\varepsilon)\big) \int_{0}^{+\infty} x_k p_k(x_k, t)\,dx_k.
\end{aligned}$$

Note that all the integrals are well defined since log-concave densities decrease exponentially at infinity (and more precisely, they admit exponential bounds such as $Ce^{-c|x|}$). Thus,

$$\frac{\partial f_t(\theta_0)}{\partial \theta_k} = \lim_{\varepsilon \to 0} \frac{f_t(\theta_0 + \varepsilon e_k) - f_t(\theta_0)}{\varepsilon} = -\int_{-\infty}^{+\infty} x_k p_k(x_k, t)\,dx_k.$$

Since $p_k(x_k, x_n) = \int_{\mathbf{R}^{n-2}} p(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_{n-1}, x_n) \frac{dx}{x_k\,dx_n}$, we arrive at

$$\frac{\partial f_t(\theta_0)}{\partial \theta_k} = -\int_{\mathbf{R}^{n-1}} x_k p(x_1, \ldots, x_{n-1}, t)\,dx_1 \cdots dx_{n-1}.$$

Hence, given a unit vector l in $\mathbf{R}^{n-1}$, we have

$$\langle \nabla f_t(\theta_0), l\rangle = -\int_{\mathbf{R}^{n-1}} \langle l, y\rangle\, p(y,t)\, dy,$$

where we write $y = (x_1, \dots, x_{n-1})$. It remains to apply Lemma 3.5. $\square$

PROOF OF THEOREM 1.2. As in the proof of Theorem 1.1, fix an arbitrary $\delta \in (0,1)$ and put

$$\Omega(t) = \{\theta \in S^{n-1} : e^{c|t|}\,|F_\theta(t) - F(t)| \le C\delta\}, \qquad t \in \mathbf{R},$$

where c and C are constants from Proposition 2.1. By the latter, and the concentration inequality (2.2) on the sphere, for all $t \in \mathbf{R}$,

$$\sigma_{n-1}(\Omega(t)) \ge 1 - 2e^{-(n-1)\delta^2/2}.$$

Now, apply this inequality to a sequence $t = a_1, \dots, a_N$ increasing with the step δ and such that $a_1 = -a_N$ [the number of points $N = N_\delta$ will grow as $\frac{c' \log(1/\delta)}{\delta}$ when $\delta \downarrow 0$ where the constant c' has to be later specified]. For $\Omega = \bigcap_{k=1}^N \Omega(a_k)$, we have

$$\sigma_{n-1}(\Omega) \ge 1 - 2Ne^{-(n-1)\delta^2/2}. \tag{3.16}$$

Take $\theta \in \Omega$ so that, for all a_k, $1 \le k \le N$,

$$e^{c|a_k|}|F_\theta(a_k) - F(a_k)| \le C\delta. \tag{3.17}$$

By Lemma 3.2 with $a = 0$ and $\sigma = 1$, the probability density $\rho_\theta(t) = F_\theta'(t)$ of S_θ satisfies, for all $t \in \mathbf{R}$,

$$\rho_\theta(t) \le C_1 e^{-c_1|t|}, \tag{3.18}$$

where c_1 and C_1 are some positive numerical constants. Integrating this inequality over θ we have a similar inequality

$$\rho(t) \le C_1 e^{-c_1|t|} \tag{3.19}$$

for the probability density $\rho(t) = F'(t)$ (although ρ does not need to be log-concave). Since inequalities (3.17)–(3.19) hold true also with constants $\max\{C, C_1\}$, $\min\{c, c_1\}$, we may assume in the sequel that $C_1 = C$ and $c_1 = c$. In particular,

$$\max\{F_\theta(-t), 1 - F_\theta(t)\} \le \frac{C}{c} e^{-ct}, \qquad t \ge 0,$$

and a similar inequality holds for F. Equivalently, $e^{ct/2}\max\{F_\theta(-t), 1 - F_\theta(t)\} \le \frac{C}{c} e^{-ct/2}$ (and the same for F), so

$$e^{c|t|/2}|F_\theta(t) - F(t)| \le \frac{C}{c}\delta \qquad \text{for all } |t| \ge \frac{2}{c}\log\frac{1}{\delta}. \tag{3.20}$$

To get an analogous estimate for "small" $|t|$, we use (3.17). Let $a_k < t < a_{k+1}$ for some $k = 1, \dots, N-1$. Since $a_{k+1} - a_k = \delta$, by (3.17)–(3.19),

$$e^{c|t|}|F_\theta(t) - F(t)| \le e^{c(|a_k|+\delta)}\Big(|F_\theta(a_k) - F(a_k)| + \delta \sup_{a_k<s<a_{k+1}} \max\{\rho_\theta(s), \rho(s)\}\Big)$$
$$\le C\delta e^{c\delta} + C\delta e^{2c\delta} \le C(e^c + e^{2c})\delta,$$

where we used the assumption $\delta \le 1$ on the last step. Together with (3.20), we may thus conclude that, in case $a_N \ge \frac{2}{c}\log\frac{1}{\delta}$, the inequality

$$e^{c|t|/2}|F_\theta(t) - F(t)| \le C\Big(e^c + e^{2c} + \frac{1}{c}\Big)\delta \tag{3.21}$$

holds true for all $t \in \mathbf{R}$. By the construction, $(N-1)\delta = a_N - a_1 = 2a_N$, so, the condition $N - 1 \ge \frac{2}{c\delta}\log\frac{1}{\delta}$ has to be fulfilled. The least such number satisfies $N \le 2 + \frac{2}{c\delta}\log\frac{1}{\delta}$, so by (3.16) and (3.21),

$$\sigma_{n-1}\Big\{\sup_{t\in\mathbf{R}} e^{c|t|/2}|F_\theta(t) - F(t)| \ge C\Big(e^c + e^{2c} + \frac{1}{c}\Big)\delta\Big\}$$
$$\le \Big(4 + \frac{4}{c\delta}\log\frac{1}{\delta}\Big)e^{-(n-1)\delta^2/2}.$$

Appropriately introducing new positive constants, say, c and C, we obtain that

$$\sigma_{n-1}\Big\{\sup_{t\in\mathbf{R}} e^{c|t|}|F_\theta(t) - F(t)| \ge \delta\Big\} \le \frac{C}{\delta}\log\frac{1}{\delta}e^{-n\delta^2/C}, \qquad 0 < \delta < \frac{1}{2}.$$

If $n\delta^2 \ge C$, the right-hand side does not exceed $C'\sqrt{n}\log n e^{-n\delta^2/C}$ for a certain big value of C'. The latter quantity is larger than 1 when $n\delta^2 \le C$, so that in both cases, $\sigma_{n-1}\{\sup_{t\in\mathbf{R}} e^{c|t|}|F_\theta(t) - F(t)| \ge \delta\} \le C'\sqrt{n}\log n e^{-n\delta^2/C}$. This gives inequality (1.5) for the range $0 < \delta < \frac{1}{2}$. It remains to note that large values of δ in (1.5) are not interesting since $\sup_{t\in\mathbf{R}} e^{c|t|}|F_\theta(t) - F(t)| \le C''$, as already used before. Theorem 1.2 has been proved. □

Now one can compare the log-concave case with Corollaries 2.1 and 2.4.

COROLLARY 3.6. *For every log-concave random vector X satisfying* (3.1) *and* (3.2),

$$\inf_{\theta\in S^{n-1}} \|F_\theta - F\|_\infty \le C\Big(\frac{\log n}{n}\Big)^{1/2}$$

where C is a universal constant.

As in the general case, in order to reach an analogous statement with F replaced by the normal distribution function Φ, one needs an additional information about concentration of $|X|$ around the point $t = 1$. For example, in terms of the smallest value $\varepsilon_n \geq 0$ in

$$\mathbf{P}\left\{\left|\frac{|X|}{\sqrt{n}} - 1\right| \geq \varepsilon_n\right\} \leq \varepsilon_n,$$

we have $\|F - \Phi\|_\infty \leq \frac{4}{\sqrt{n}} + 2.5\varepsilon_n$. As already mentioned and used in [1], the property $\mathrm{Var}(|X|^2) = O(n)$ leads to $\varepsilon_n = O(n^{-1/3})$. For log-concave X, this estimate is, however, not optimal and can be sharpened with the help of the following.

LEMMA 3.7. *Given a log-concave vector X in $\mathbf{R}^n$, for all $h > 0$,*

$$\mathbf{P}\{\big||X| - \sqrt{\mathbf{E}|X|^2}\big| \geq h\} \leq 2\exp\left\{-\frac{c\mathbf{E}^{1/4}|X|^2}{\mathrm{Var}^{1/4}(|X|^2)}h^{1/2}\right\},$$

where $c > 0$ is some numerical constant. Under (1.1), *for all $\varepsilon > 0$, we thus have*

$$\mathbf{P}\left\{\left|\frac{|X|}{\sqrt{n}} - 1\right| \geq \varepsilon\right\} \leq 2\exp\left\{-\frac{cn^{1/2}\varepsilon^{1/2}}{\mathrm{Var}^{1/4}(|X|^2)}\right\}.$$

As an immediate consequence, we get the following corollary.

COROLLARY 3.8. *If the log-concave random vector X satisfies* (3.1) *and* (3.2), *and* $\mathrm{Var}(|X|^2) = O(n)$, *then* $\inf_{\theta \in S^{n-1}} \|F_\theta - \Phi\|_\infty = O(\frac{\log^2 n}{\sqrt{n}})$.

PROOF OF LEMMA 3.7. For any polynomial f on $\mathbf{R}^n$ of degree d and for all $p \geq 1$,

$$\mathbf{E}^{1/p}|f(X)|^p \leq (Cp)^d \mathbf{E}|f(X)|,$$

where $C > 0$ is universal. We apply this fact proved in [5] to $f(x) = |x|^2 - a^2$, $a = \sqrt{\mathbf{E}|X|^2}$, to get

$$\mathbf{E}^{1/p}\big||X|^2 - a^2\big|^p \leq (Cp)^2 \mathbf{E}\big||X|^2 - a^2\big|.$$

Hence,

$$\begin{aligned}\mathbf{E}^{1/p}\big||X| - a\big|^p &= \mathbf{E}^{1/p}\frac{||X|^2 - a^2|^p}{||X| + a|^p} \leq \frac{1}{a}\mathbf{E}^{1/p}\big||X|^2 - a^2\big|^p \\ &\leq \frac{(Cp)^2}{a}\mathbf{E}\big||X|^2 - a^2\big| \leq \frac{(Cp)^2}{a}\mathrm{Var}^{1/2}(|X|^2).\end{aligned}$$

Consider the random variable $\xi = ||X| - a|^{1/2}$. Applying the above estimate to $p = k/2$ with integer $k \geq 2$, we obtain that

$$\mathbf{E}\xi^k \leq \left(\frac{Ck}{2}\right)^k \left(\frac{\mathrm{Var}^{1/2}(|X|^2)}{a}\right)^{k/2} = C_X^k k^k,$$

where $C_X = \frac{C\,\mathrm{Var}^{1/4}(|X|^2)}{2E^{1/4}|X|^2}$. In order to involve the value $k = 1$, we may write $\mathbf{E}\xi^k \leq (2C_X)^k k^k$, and using $k^k \leq e^k k!$, we conclude that

$$\mathbf{E}e^{t\xi} = 1 + \sum_{k=1}^{\infty} \frac{t^k}{k!}\mathbf{E}\xi^k \leq 1 + \sum_{k=1}^{\infty} (2eC_X t)^k = 2$$

for $t = 1/(4eC_X)$. At last, by Chebyshev's inequality,

$$\mathbf{P}\{\big||X| - a\big| \geq h\} = \mathbf{P}\{\xi \geq h^{1/2}\} \leq 2e^{-th^{1/2}}.$$

Lemma 3.7 follows. □

Acknowledgment. We are grateful to the referee for a careful reading of the manuscript and valuable remarks.

REFERENCES

[1] ANTILLA, M., BALL, K. and PERISSINAKI, I. (1998). The central limit problem for convex bodies. Preprint.

[2] BALL, K. (1988). Logarithmically concave functions and sections of convex sets. *Stud. Math.* **88** 69–84.

[3] BALL, K. and PERISSINAKI, I. (1998). Subindependence of coordinate slabs in ℓ_p^n balls. *Israel J. Math.* **107** 289–299.

[4] BOBKOV, S. G. (1999). Isoperimetric and analytic inequalities for log-concave probability distributions. *Ann. Probab.* **27** 1903–1921.

[5] BOBKOV, S. G. (2000). Some generalizations of results of Yu. V. Prokhorov on Khinchine-type inequalities for polynomials. *Theory Probab. Appl.* **45** 745–748.

[6] DIACONIS, P. and FREEDMAN, D. (1984). Asymptotics of graphical projection pursuit. *Ann. Statist.* **12** 793–815.

[7] GRÜNBAUM, B. (1960). Partitions of mass-distributions and of convex bodies by hyperplanes. *Pacific J. Math.* **10** 1257–1261.

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

[9] HAMMER, P. C. (1995). Volumes cut from convex bodies by planes. Preprint. (This reference is given in [7].)

[10] KANNAN, R., LOVÁSZ, L. and SIMONOVITS, M. (1995). Isoperimetric problems for convex bodies and a localization lemma. *Discrete and Comput. Geom.* **13** 541–559.

[11] LATALA, R. (1999). On the equivalence between geometric and arithmetic means for log-concave measures. In *Convex Geometric Analysis* 123–127. Cambridge Univ. Press.

[12] LEDOUX, M. (1999). Concentration of measure and logarithmic Sobolev inequalities. *Séminaire de Probabilités XXXIII. Lecture Notes in Math.* **1799** 120–216. Springer, Berlin.

[13] MILMAN, V. D. and SCHECHTMAN, G. (1986). Asymptotic theory of finite-dimensional normed spaces. *Asymptotic Theory of Finite Dimensional Normed Spaces*. *Isoperimetric Inequalities in Riemannian Manifolds*. *Lecture Notes in Math.* **1200**. Springer, Berlin.

[14] MUELLER, C. E. and WEISSLER, F. B. (1992). Hypercontractivity for the heat semigroup for ultraspherical polynomials and on the n-sphere. *J. Funct. Anal.* **48** 252–283.

[15] PETROV, V. V. (1987). *Limit Theorems for Sums of Independent Random Variables.* Nauka, Moscow (in Russian).

[16] PRÉKOPA, A. (1973). On logarithmic concave measures and functions. *Acta Sci. Math.* (*Szeged*) **34** 335–343.

[17] ROKHLIN, V. A. (1949). On the fundamental ideas of measure theory. *Math. Sbornik* **25** 107–150 (in Russian). (Also in *Selected Works.* Reminiscences of V. A. Rokhlin by V. I. Arnold, A. M. Vershik, S. P. Novikov and Ya. G. Sinai, with commentaries on Rokhlin's work by N. Yu. Netsvetaev and Vershik. Edited with preface by A. M. Vershik. Moskovskii Tsentr Nepreryvnogo Mat. Obrazovaniya, Moscow, 1999.)

[18] SUDAKOV, V. N. (1978). Typical distributions of linear functionals in finite-dimensional spaces of higher dimension. *Soviet Math. Dokl.* **19** 1578–1582. [Translated from *Dokl. Akad. Nauk SSSR* **243** (1978).]

[19] VON WEIZSÄCKER, H. (1997). Sudakov's typical marginals, random linear functionals and a conditional central limit theorem. *Probab. Theory Related Fields* **107** 313–324.

SCHOOL OF MATHEMATICS
UNIVERSITY OF MINNESOTA
127 VINCENT HALL
206 CHURCH ST. S.E.
MINNEAPOLIS, MINNESOTA 55455
E-MAIL: bobkov@math.umn.edu