Projecting the surface measure of the sphere of ℓ_p^n

Assaf Naor * Dan Romik †

January 16, 2002

Abstract.- We prove that the total variation distance between the cone measure and surface measure on the sphere of ℓ_p^n is bounded by a constant times $1/\sqrt{n}$. This is used to give a new proof of the fact that the coordinates of a random vector on the ℓ_p^n sphere are approximately independent with density proportional to $\exp(-|t|^p)$, a unification and generalization of two theorems of Diaconis and Freedman. Finally, we show in contrast that a projection of the surface measure of the ℓ_p^n sphere onto a *random* k-dimensional subspace is "close" to the k-dimensional Gaussian measure.

Résumé.- Nous montrons que la distance de la variation totale entre la mesure de cône et la mesure d'aire sur la sphère de ℓ_p^n est bornée par une constante fois $1/\sqrt{n}$. Cela fournit une nouvelle démonstration du fait que les coordonées d'un vecteur alèatoire dans la sphère de ℓ_p^n sont approximativement indépendantes avec une densité proportionelle à $\exp(-|t|^p)$, ce qui constitue une unification et une généralization de deux théorèmes de Diaconis et Freedman. Nous montrons ensuite que la projection de la mesure d'aire de la sphère de ℓ_p^n sur un sous-espace aléatoire k-dimensionnel est "proche" de la mesure Gaussienne k-dimensionnelle.

1 Introduction

In this paper, we study projections of the surface measure on the ℓ_p^n sphere onto k-dimensional subspaces. For $p = 2$, all these projections are clearly equal.

*Institute of Mathematics, Hebrew University, Jerusalem 91904, Israel. E-mail: naor@math.huji.ac.il. Supported in part by the Bi-National Science Foundation Israel-USA and by the Clore Foundation. This work is part of a Ph.D. thesis being prepared under the supervision of Professor Joram Lindenstrauss.

†School of Mathematical Sciences, Tel-Aviv University, Tel-Aviv 69978. E-mail: romik@math.tau.ac.il. This work is part of a Ph.D. thesis being prepared under the supervision of Professor David Gilat.

By a result of Diaconis-Freedman [7], for k small with respect to n this measure is close, in total variation distance, to the Gaussian measure. For general $p \geq 1$, we prove in section 4, that for a *random* choice of k-dimensional subspace, the projection will be asymptotically close to the Gaussian measure. On the other hand, the concrete k-dimensional subspace spanned by the first k coordinates exhibits a different behavior: the measure will be asymptotically close to the product measure of k i.i.d. random variables with density proportional to $\exp(-|t|^p)$. This result unifies and generalizes two results of Diaconis-Freedman, namely the above statement in the cases $p = 1$ and $p = 2$. A direct attempt to prove such a statement, leads instead to a similar statement, with the surface measure replaced by the so-called "cone measure", a distinction that does not appear in the cases $p = 1, 2$. (The cone measure is a measure for which a natural polar coordinate integration formula holds - see section 2.) In fact, the above statement was proved by Rachev and Rüschendorf in [15] (see also [3] and [14]) for the cone measure, and was conjectured to be true also for the surface measure. This problem was solved positively by Mogul'skiĭ in [11]. In this paper we propose a different approach to this problem. Our main result is that the cone measure and surface measure are in fact close in total variation distance for large n. Since the cone measure has a simple probabilistic representation, this result shows how one can approximate the geometric surface measure by the more concrete probability distribution given by the cone measure. The above result is applied to give a new proof of Mogul'skiĭ's solution to the problem posed in [15]. In fact, we show that the solution follows from the results of [15] (although it was conjectured there that it requires a completely different proof).

Section 4 deals with a version of the so-called Randomized Central Limit Theorem for the ℓ_p^n sphere. This theorem was studied for product measures by the second-named author in [16], [17]. For uniform measures on convex bodies in isotropic position, a one-dimensional version of it was proved by Antilla, Ball and Perissinaki [2] (see also [9] and [6]). Our results hold for general isotropic measures on $\mathbb{R}^n$ satisfying a certain negative correlation property, which arises naturally from the proof in [16] (In fact, it is a standard technique in probability theory to generalize results known to hold for independent random variables, by assuming that the variables involved are only negatively correlated). The usefulness of this property was also noted in [2], where it was shown to imply the so-called Central Limit Property for the volume measure on the ball of ℓ_p^n.

In section 5 we discuss several related open problems.

2 Cone measure and surface measure

For every star-shaped body $K \subset \mathbb{R}^n$, one can define two natural measures on the boundary of K. One is the regular surface measure and the other is the "cone measure". The cone measure of a subset A of ∂K is the volume of $[0,1]A$, i.e. the cone with base A and cusp 0. Both these measures have appeared in various contexts in the literature. Most notably, the cone measure appears in the Gromov-Milman theorem for concentration of Lipschitz functions on uniformly convex bodies. As far as we know, the relation between these two measures has not been studied; each measure appears naturally in different contexts, and most authors have been satisfied with an ad hoc choice of the measure most suitable for their particular application.

In this section we will show that, for the case of the ℓ_p^n sphere ($p \geq 1$), these measures are asymptotically close. More precisely, the total variation distance between the two measures is at most a constant (depending on p) times $1/\sqrt{n}$.

We recall some basic facts about the total variation distance. For P, Q probability measures on a measurable space $(\Omega, \mathcal{F})$, the total variation distance between them is defined as $||P - Q|| = 2\sup\{|P(A) - Q(A)| : A \in \mathcal{F}\}$. If P, Q are absolutely continuous with respect to some reference measure λ, with respective densities f and g, then the total variation distance is known to be equal to $\int_\Omega |f - g| d\lambda$.

Fix $p > 0$ and an integer n. Recall that the ℓ_p^n norm is defined by :

$$||x||_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} .$$

The ℓ_p^n sphere is defined by: $S(\ell_p^n) = \{x \in \mathbb{R}^n ; ||x||_p = 1\}$, and the ℓ_p^n ball is defined by $B(\ell_p^n) = \{x \in \mathbb{R}^n ; ||x||_p \leq 1\}$. We denote by σ the normalized surface measure on $S(\ell_p^n)$, and by μ the normalized cone measure. In other words, for every measurable $A \subset S(\ell_p^n)$ we put :

$$\mu(A) = \frac{1}{\mathrm{vol}(B(\ell_p^n))} \mathrm{vol}([0,1]A).$$

Here "vol" refers to the Lebesgue measure on $\mathbb{R}^n$, and $[0,1]A = \{ta : a \in A, 0 \leq t \leq 1\}$.

The measure μ has a useful probabilistic description. Let g be a random variable with density $1/(2\Gamma(1+1/p))e^{-|t|^p}$ $(t \in \mathbb{R})$. If $g_1, ..., g_n$ are i.i.d. copies of g, put:

$$S = \sum_{i=1}^{n} |g_i|^p,$$

and consider the random vector:

$$X = \left(\frac{g_1}{S^{1/p}}, ..., \frac{g_n}{S^{1/p}}\right) \in \mathbb{R}^n.$$

The following result appeared in the paper of Schechtman and Zinn [19], and later independently also in [15]:

Theorem 1 *The random vector X is independent of S. Moreover, For every measurable $A \subset S(\ell_p^n)$ we have :*

$$\mu(A) = P(X \in A).$$

We will now estimate the total variation distance between the surface measure and cone measure on $S(\ell_p^n)$.

In this and in what follows, c will denote a numerical constant, which may change in each particular appearance. Likewise, c_p will denote a constant depending on the parameter $p > 0$.

Theorem 2 *For all $1 \leq p < \infty$, on $S(\ell_p^n)$,*

$$||\mu - \sigma|| \leq \frac{c_p}{\sqrt{n}}$$

The measures μ and σ are in fact equal for $p = 1, 2$, and ∞. Since we are mainly concerned here with the probabilistic applications, the discussion of the constant c_p is postponed to a later (more geometrically oriented) paper [13], where it is proved that there is an absolute constant $C > 0$ such that for all $p \geq 1$ we can take $c_p = C\left(1 - \frac{1}{p}\right)\left|\frac{1}{p} - \frac{1}{2}\right|\frac{\sqrt{np}}{n+p}$. In particular, c_p is bounded for $p \geq 1$. See Section 5 for a discussion of the case $0 < p < 1$.

We start with some general facts concerning the cone and surface measures. For the sake of greater generality, and in anticipation of future developments, we state these results for a general convex body $K \subset \mathbb{R}^n$. Let σ_K be the normalized surface measure on K, and let μ_K be the normalized cone measure on ∂K, defined as before by $\mu_K(A) = \frac{\mathrm{vol}([0,1]A)}{\mathrm{vol}(K)}$. We will denote by $||\cdot||_K$ the Minkowski functional (norm) of K.

The cone measure can be thought of as the measure for which a polar coordinate integration formula holds:

Proposition 1 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be an integrable function (w.r.t. Lebesgue measure). Then*

$$\int_{\mathbb{R}^n} f(x)dx = n \cdot \mathrm{vol}(K) \int_0^\infty r^{n-1} \int_{\partial K} f(r \cdot z) d\mu_K(z) dr$$

Proof: By approximation, it is enough to verify the formula for indicator functions of sets of the form $(a,b)E$, where $a < b$ and $E \subset \partial K$. For such sets the formula is trivial. ■

Note 1 *An equivalent formulation of Proposition 1, is the statement that the mapping $x \to (x/||x||_K, ||x||_K)$ transforms Lebesgue measure on $\mathbb{R}^n$ into the product of the cone measure on ∂K and the measure $n \cdot \mathrm{vol}(K) \cdot r^{n-1}dr$ on $[0,\infty)$.*

In the next lemma we compute the density of the surface measure with respect to the cone measure.

Lemma 1 *σ_K is absolutely continuous with respect to μ_K, and its density is given for almost every $x \in \partial K$ by*

$$\frac{d\sigma_K}{d\mu_K}(x) = \frac{n \cdot \mathrm{vol}(K)}{\mathrm{area}(\partial K)} \left\| \nabla(||\cdot||_K)(x) \right\|_2.$$

Where $||\cdot||_2$ denotes the Euclidean norm on $\mathbb{R}^n$.

Proof: We will denote by A the (not normalized) surface area measure of ∂K and by C the unnormalized cone measure of ∂K (i.e. $C(F) = \mathrm{vol}([0,1]F)$). We will also denote by $B(x,t)$ the Euclidean ball with radius t and center $x \in \mathbb{R}^n$. The volume of the k-dimensional unit Euclidean ball is denoted by ω_k. Recall that the measure A is defined for any open $U \subset \mathbb{R}^n$ by:

$$A(U) = \lim_{\epsilon \to 0} \frac{\mathrm{vol}(U \cap \partial K + B(0,\epsilon))}{2\epsilon}.$$

It is a classical fact (see for example [10] theorem 16.2) that almost every $x \in \partial K$ is a *density point* of A, in the sense that

$$\lim_{\epsilon \to 0} \frac{A(B(x,\epsilon))}{\epsilon^{n-1}\omega_{n-1}} = 1.$$

Fix $x \in \partial K$ which is a density point of A, and which is a point of differentiability of $||\cdot||_K$ (almost every $x \in \partial K$ has these properties). This means that we can write:

$$||x+y||_K = 1 + \langle \nabla(||\cdot||_K)(x), y\rangle + r(y),$$

where:

$$\rho(\delta) = \sup\left\{ \frac{|r(y)|}{||y||_2} : 0 < ||y||_2 \le \delta \right\} \underset{\delta \to 0}{\longrightarrow} 0.$$

Let H be the tangent hyperplane to ∂K at x, i.e. $H = x + \{\nabla(||\cdot||_K)(x)\}^{\perp}$. For simplicity define $z = \nabla(||\cdot||_K)(x)$. It is well known that $\langle x, z\rangle = 1$ (to see this note that $1+\delta = ||x+\delta x||_K = 1 + \delta\langle x,z\rangle + r(\delta x)$. Division by δ and taking the limit $\delta \to 0$ gives the required result). Similarly, for every $y \in \partial K$, $|\langle z, y\rangle| \le 1$ (in other words, z is the norming functional of x). Now, for every $0 < \epsilon < \min\{1/(2||z||_2), ||x||_2\}$ we claim that:

$$[0,1]\left(B(x,\epsilon) \cap \partial K\right) \subseteq [0,1]\left(B\left(x, \epsilon + 4||x||_2 \epsilon\rho(\epsilon)\right) \cap H\right).$$

Indeed, take $0 \le t \le 1$ and $y \in \partial K$ with $||y-x||_2 \le \epsilon$. Then,

$$\langle y, z\rangle = 1 + \langle y-x, z\rangle \ge 1 - \epsilon||z||_2 \ge 1/2,$$

and $1 = ||y||_K = 1 + \langle y-x, z\rangle + r(y-x)$ so that $|\langle y-x,z\rangle| \le \epsilon\rho(\epsilon)$. If we put $v = \langle y,z\rangle^{-1} y$ and $s = t\langle y,z\rangle$ then $ty = sv$, $0 \le s \le 1$, $v \in H$ (since $\langle v-x, z\rangle = \langle y,z\rangle^{-1}\langle y,z\rangle - 1 = 0$) and:

$$||v-x||_2 = \left\| \frac{y}{\langle y,z\rangle} - x \right\|_2 \le ||y-x||_2 + \left| \frac{1}{\langle y,z\rangle} - 1 \right| ||y||_2 \le$$

$$\le \epsilon + 2\epsilon\rho(\epsilon)(||x||_2 + \epsilon) \le \epsilon + 4||x||_2\epsilon\rho(\epsilon),$$

and this proves our claim.

To prove a reverse inclusion, fix $\epsilon > 0$ such that $\epsilon - 2||x||_2\epsilon\rho(\epsilon) = \delta > 0$ and take $y \in H$ with $||y-x||_2 \le \delta$ and $0 < t < 1$. Now,

$$||y||_K = 1 + \langle y-x, z\rangle + r(y-x) = 1 + r(y-x),$$

so that $1 \le ||y||_K \le 1 + \delta\rho(\delta) \le 1 + \epsilon\rho(\epsilon)$. Hence, if we put $v = y/||y||_K$ and $s = t||y||_K$ then $0 < s \le 1 + \epsilon\rho(\epsilon)$, $ty = sv$, $v \in \partial K$ and as long as $\delta \le ||x||_2$:

$$||v - x||_2 \le ||y - x||_2 + \left| \frac{1}{||y||_K} - 1 \right| ||y||_2 \le \delta + 2||x||_2\epsilon\rho(\epsilon) = \epsilon.$$

We have proved that as long as ϵ is small enough:

$$(1 + \epsilon\rho(\epsilon))[0,1]\left(B(x,\epsilon) \cap \partial K\right) \supseteq [0,1]\left(B\left(x, \epsilon - 2||x||_2\epsilon\rho(\epsilon)\right) \cap H\right).$$

Note that for every $a > 0$, $[0,1](B(x,a) \cap H)$ is a cone with cusp 0 and base $B(x,a) \cap H$. The (perpendicular) height of this cone is $\langle z, x\rangle/||z||_2 = 1/||z||_2$, so that:

$$\mathrm{vol}([0,1](B(x,a) \cap H)) = \frac{a^{n-1}\omega_{n-1}}{n||z||_2}.$$

Using this observation and the previous two inclusions we get that for ϵ small enough:

$$\frac{(\epsilon - 2||x||_2\epsilon\rho(\epsilon))^{n-1}}{(1+\epsilon\rho(\epsilon))^n} \cdot \frac{\omega_{n-1}}{n||z||_2} \le C(B(x,\epsilon) \cap \partial K) \le \frac{(\epsilon + 4||x||_2\epsilon\rho(\epsilon))^{n-1}\omega_{n-1}}{n||z||_2}.$$

Hence:

$$\lim_{\epsilon \to 0} \frac{C(B(x,\epsilon) \cap \partial K)}{\epsilon^{n-1}\omega_{n-1}} = \frac{1}{n\left\|\nabla(||\cdot||_K)(x)\right\|_2}.$$

Finally using the fact that x is a density point of A:

$$\frac{d\sigma_K}{d\mu_K} = \frac{\mathrm{vol(K)}}{A(\partial K)} \cdot \lim_{\epsilon \to 0} \frac{A(B(x,\epsilon) \cap \partial K)}{C(B(x,\epsilon) \cap \partial K)} = \frac{n \cdot \mathrm{vol(K)}}{A(\partial K)} \left\|\nabla(||\cdot||_K)(x)\right\|_2.$$

■

Applying Lemma 1 to the special case of $B(\ell_p^n)$ (and reverting to our earlier notation), we have easily

Lemma 2

$$\frac{d\sigma}{d\mu} = \frac{d\sigma_{B(\ell_p^n)}}{d\mu_{B(\ell_p^n)}}(x) = C_{n,p} \cdot \left(\sum_{i=1}^n |x_i|^{2p-2}\right)^{1/2}$$

Where $C_{n,p}$ is a constant depending on n and p.

Proof of Theorem 2: By Lemma 2, we are faced with the problem of bounding the expression

$$||\mu - \sigma|| = \int_{S(\ell_p^n)} \left| C_{n,p} \left(\sum_{i=1}^{n} |x_i|^{2p-2} \right)^{1/2} - 1 \right| d\mu,$$

where $C_{n,p} = \left[\int_{S(\ell_p^n)} \left(\sum_{i=1}^{n} |x_i|^{2p-2} \right)^{1/2} d\mu \right]^{-1}$ is merely a normalizing constant. Now fix $q = 2p - 2$. Note that for any random variable Z and any $a \in \mathbb{R}$, $\mathbb{E}|Z - \mathbb{E}Z| \leq \mathbb{E}|Z - a| + \mathbb{E}|\mathbb{E}Z - a| \leq \mathbb{E}|Z - a| + \mathbb{E}(\mathbb{E}|Z - a|) = 2\mathbb{E}|Z - a|$. Using this fact, and using the fact that for $a > b > 0$, $\sqrt{a} - \sqrt{b} = \frac{a-b}{\sqrt{a}+\sqrt{b}} \leq \frac{a-b}{\sqrt{b}}$, we calculate:

$$||\mu - \sigma|| = \int_{S(\ell_p^n)} \left| \left(\int_{S(\ell_p^n)} ||y||_q^{q/2} d\mu(y) \right)^{-1} ||x||_q^{q/2} - 1 \right| d\mu(x) =$$

$$= \left(\int_{S(\ell_p^n)} ||y||_q^{q/2} d\mu(y) \right)^{-1} \int_{S(\ell_p^n)} \left| ||x||_q^{q/2} - \int_{S(\ell_p^n)} ||y||_q^{q/2} d\mu(y) \right| d\mu(x) \leq$$

$$\leq 2 \left(\int_{S(\ell_p^n)} ||y||_q^{q/2} d\mu(y) \right)^{-1} \int_{S(\ell_p^n)} \left| ||x||_q^{q/2} - \left(\int_{S(\ell_p^n)} ||y||_q^{q} d\mu(y) \right)^{1/2} \right| d\mu(x) \leq$$

$$\leq 2 \left(\int_{S(\ell_p^n)} ||y||_q^{q/2} d\mu(y) \right)^{-1} \left(\int_{S(\ell_p^n)} ||y||_q^{q} d\mu(y) \right)^{-1/2} \cdot$$

$$\cdot \int_{S(\ell_p^n)} \left| ||x||_q^{q} - \int_{S(\ell_p^n)} ||y||_q^{q} d\mu(y) \right| d\mu(x) \leq$$

$$\leq 2 \left(\int_{S(\ell_p^n)} ||y||_q^{q/2} d\mu(y) \right)^{-1} \left(\int_{S(\ell_p^n)} ||y||_q^{q} d\mu(y) \right)^{-1/2} \cdot$$

$$\cdot \left[\int_{S(\ell_p^n)} ||x||_q^{2q} d\mu(x) - \left(\int_{S(\ell_p^n)} ||y||_q^{q} d\mu(y) \right)^2 \right]^{1/2}.$$

The last inequality used the fact that for any random variable Z, $\mathbb{E}|Z - \mathbb{E}Z| \leq \sqrt{\mathbb{E}Z^2 - (\mathbb{E}(Z))^2}$.

We now go back to the probabilistic realization of the measure μ given in Theorem 1. Using the notation of Theorem 1, put $T = \sum_{i=1}^{n} |g_i|^q$. Note that the independence of S and X (defined as before) implies in particular that for any $\alpha > 0$, $\int_{S(\ell_p^n)} ||x||_q^{\alpha q} d\mu(x) = \mathbb{E}[T^\alpha / S^{\alpha q/p}] = (\mathbb{E}T^\alpha)/(\mathbb{E}S^{\alpha q/p})$. Using this observation, and Theorem 1, the above inequality translates into:

$$||\mu - \sigma|| \leq 2 \cdot \frac{(\mathbb{E}T)^{1/2}}{\mathbb{E}T^{1/2}} \cdot \frac{(\mathbb{E}S^{q/2p})(\mathbb{E}S^{q/p})^{1/2}}{(\mathbb{E}S^{2q/p})^{1/2}} \cdot \sqrt{\frac{\mathbb{E}T^2}{(\mathbb{E}T)^2} - \frac{\mathbb{E}S^{2q/p}}{(\mathbb{E}S^{q/p})^2}}$$

The first fraction in this expression is bounded by a constant (depending on p). Indeed,

$$\frac{(\mathbb{E}T)^{1/2}}{\mathbb{E}T^{1/2}} = \frac{(\mathbb{E}|g_1|^q)^{1/2}\sqrt{n}}{\mathbb{E}T^{1/2}} = \frac{(\mathbb{E}|g_1|^q)^{1/2}\sqrt{n}}{\mathbb{E}||(|g_1|^{q/2}, \ldots, |g_n|^{q/2})||_2} \leq$$
$$\frac{(\mathbb{E}|g_1|^q)^{1/2}\sqrt{n}}{\mathbb{E}||(|g_1|^{q/2}, \ldots, |g_n|^{q/2})||_1/\sqrt{n}} = \frac{(\mathbb{E}|g_1|^q)^{1/2} \cdot n}{n \cdot \mathbb{E}|g_1|^{q/2}} = c_p$$

The second fraction is trivially bounded by 1, by Jensen's inequality. To conclude our proof of Theorem 2, we thus need to bound the radical by a universal constant times $n^{-1/2}$. Note that:

$$\frac{\mathbb{E}T^2}{(\mathbb{E}T)^2} = \frac{n\mathbb{E}|g|^{2q} + n(n-1)(\mathbb{E}|g|^q)^2}{n^2(\mathbb{E}|g|^q)^2} \leq 1 + \frac{c_p}{n},$$

for some $c_p \in \mathbb{R}$, and the required inequality follows since $\frac{\mathbb{E}S^{2q/p}}{(\mathbb{E}S^{q/p})^2} \geq 1$. ■

3 The asymptotic distribution of the coordinates of a random vector on the ℓ_p^n sphere

In this section we apply the result of the previous section to prove that for a random vector on $S(\ell_p^n)$ (chosen according to either the surface or cone measures), the joint distribution of k of its coordinates will be close in total variation to the law of k i.i.d. r.v.s having density $1/(2\Gamma(1+1/p))e^{-|t|^p}$ $(t \in \mathbb{R})$, as long as $k = o(n)$. More precisely, for any k we will show that the variation distance is of the order k/n for the cone measure, and $k/n + c/\sqrt{n}$ for the surface measure. These statements are essentially the results of [15] and [11] (for the surface measure the estimate in [11] is also of order k/n so that our proof gives a somewhat worse estimate than that of [11] in the range $k \leq \sqrt{n}$. In Section 5 we suggest a method which may lead to an improvement of this estimate.). This application is a generalization of results that appeared in [7]. For the sake of a more streamlined presentation of these results, and since we wish to emphasize the simplicity of the proof, we will repeat the proof of a technical calculation that was needed also in [7] and [15]. We will then reduce the problem to a one dimensional computation. A reduction argument was also needed in [7] and [15]. We give here a direct and simple argument which achieves such a reduction, and this makes the presentation self contained.

In what follows, f_Z will denote the density of a r.v. Z.

The following Lemma is essentially contained in [7]:

Lemma 3 *Assume that $0 < \alpha < \beta$ and $\beta > 1$. Let X be a r.v. which is $(\alpha+\beta)$ times a Beta(α, β) r.v.; that is, X has density*

$$f_X(x) = \frac{\Gamma(\alpha+\beta)}{(\alpha+\beta)\Gamma(\alpha)\Gamma(\beta)} \cdot \left(\frac{x}{\alpha+\beta}\right)^{\alpha-1} \cdot \left(1 - \frac{x}{\alpha+\beta}\right)^{\beta-1} \quad (0 \le x \le \alpha+\beta);$$

and let Y be an r.v. with Gamma$(\alpha, 1)$ distribution, that is

$$f_Y(x) = \frac{1}{\Gamma(\alpha)} \cdot e^{-x} x^{\alpha-1} \quad (0 \le x < \infty).$$

Then the variation distance between the distributions of X and Y is at most $\frac{4\alpha+12}{\beta}$.

Proof:

$$\frac{f_X(x)}{f_Y(x)} = \frac{\Gamma(\alpha+\beta)}{(\alpha+\beta)^\alpha \Gamma(\beta)} \cdot e^x (1 - \frac{x}{\alpha+\beta})^{\beta-1} \quad (0 \le x \le \alpha+\beta),$$

or $\frac{f_X(x)}{f_Y(x)} = A \cdot h(x)$, where $A = \frac{\Gamma(\alpha+\beta)}{(\alpha+\beta)^\alpha \Gamma(\beta)}$ and $h(x) = e^x(1 - \frac{x}{\alpha+\beta})^{\beta-1}$ $(0 \le x \le \alpha+\beta)$. First, note that $\log h(x) = x + (\beta-1)\log(1 - \frac{x}{\alpha+\beta})$ attains its maximum when $x = \alpha+1$, therefore $\log h(x) \le \alpha + 1 + (\beta-1)\log(\frac{\beta-1}{\alpha+\beta})$ $(0 \le x \le \alpha+\beta)$. Next, we bound $\log A$ using the following version of Stirling's formula:

$$1 \le \Gamma(x)/(\sqrt{2\pi} x^{x-1/2} e^{-x}) \le e^{1/12x}$$

(see the monograph by Artin [1]).

$$\log A = -\alpha\log(\alpha+\beta) + \log\Gamma(\alpha+\beta) - \log\Gamma(\beta) \le$$

$$\le -\alpha\log(\alpha+\beta) + (\alpha+\beta-\frac{1}{2})\log(\alpha+\beta) - (\alpha+\beta) - (\beta-\frac{1}{2})\log\beta + \beta + \frac{1}{12(\alpha+\beta)} =$$

$$= (\beta - \frac{1}{2})\log(\frac{\alpha+\beta}{\beta}) - \alpha + \frac{1}{12(\alpha+\beta)}$$

Adding the two bounds we have:

$$\log(Ah(x)) \leq 1 + (\beta-1)\log(\frac{\beta-1}{\beta}) + \frac{1}{2}\log(\frac{\alpha+\beta}{\beta}) + \frac{1}{12(\alpha+\beta)} =$$

$$= 1 + (\beta-1)\log(1-\frac{1}{\beta}) + \frac{1}{2}\log(1+\frac{\alpha}{\beta}) + \frac{1}{12(\alpha+\beta)} \leq$$

$$\leq \frac{1}{\beta} + \frac{\alpha}{2\beta} + \frac{1}{12(\alpha+\beta)} \leq \frac{\alpha/2+1+1/12}{\beta} \leq \frac{\alpha+3}{2\beta} < 2.$$

Now exponentiating, and using the fact that for $0 \leq x \leq 2$, $e^x \leq 1+4x$:

$$\frac{f_X(x)}{f_Y(x)} - 1 \leq e^{\frac{\alpha+3}{2\beta}} - 1 \leq \frac{2\alpha+6}{\beta}$$

Hence:

$$||P_X - P_Y|| = \int_0^\infty |f_X(x) - f_Y(x)|dx = \int_0^\infty \left|\frac{f_X(x)}{f_Y(x)} - 1\right| f_Y(x)dx =$$

$$= 2\int_0^\infty \left(\frac{f_X(x)}{f_Y(x)} - 1\right)^+ f_Y(x)dx \leq \frac{4\alpha+12}{\beta}$$

where P_Z denotes the distribution of a r.v. Z, and $a^+ = \max(a,0)$ ■

Now let $k < n$. Let $X = (X_1, X_2, \ldots, X_n)$ be a random vector on $S(\ell_p^n)$, chosen according to the cone measure. Denote its first k coordinates by $X' = (X_1, X_2, \ldots, X_k)$. Let $G = (g_1, g_2, \ldots, g_n)$ be a random vector of i.i.d. r.v.s with density $1/(2\Gamma(1+1/p))e^{-|t|^p}$, and let $G' = (g_1, \ldots, g_k)$. We wish to estimate the variation distance $||P_{a_n X'} - P'_G||$, for normalization constants a_n which are of the order $n^{1/p}$. For convenience we choose $a_n = (n/p)^{1/p}$.

Lemma 4 *Define* $W = ||a_n X'||_p^p$ *and* $Z = ||G'||_p^p$. *Then:*

$$||P_{a_n X'} - P_{G'}|| = ||P_W - P_Z||$$

Proof: Put $S' = \sum_{j=1}^k |g_j|^p$, $S'' = \sum_{j=k+1}^n |g_j|^p$. We will denote by ϕ the density of S'', and by ψ_r the density of the random variable $(S'/(S'+r))^{1/p}$, where $r > 0$. Define:

$$H(u) = \frac{1}{n \cdot \mathrm{vol}(B(\ell_p^n)) \cdot u^{n-1}} \int_0^\infty \phi(r)\psi_r(u)dr.$$

We first claim that the density of X' is $H(||x||_p)$. Indeed, take a Borel subset $B \subset \mathbb{R}^k$. By Theorem 1,

$$P(X' \in B) = P\left(\frac{G'}{(S'+S'')^{1/p}} \in B\right) = \mathbb{E}P\left(\frac{G'}{(S'+S'')^{1/p}} \in B \,\middle|\, S''\right).$$

Therefore, using the independence of S'' and G', and the independence of S' and $G'/(S')^{1/p}$ (which is Theorem 1 in $\mathbb{R}^k$), we get:

$$P(X' \in B) = \int_0^\infty \phi(r) P\left(\frac{G'}{(S'+r)^{1/p}} \in B\right) dr =$$

$$= \int_0^\infty \phi(r) P\left(\left(\frac{S'}{S'+r}\right)^{1/p} \cdot \frac{G'}{(S')^{1/p}} \in B\right) dr =$$

$$= \int_0^\infty \phi(r) \mathbb{E}\left[P\left(\left(\frac{S'}{S'+r}\right)^{1/p} \cdot \frac{G'}{(S')^{1/p}} \in B \,\middle|\, S'\right)\right] dr =$$

$$= \int_0^\infty \phi(r) \int_0^\infty \psi_r(u) P\left(u \cdot \frac{G'}{(S')^{1/p}} \in B\right) du dr =$$

$$= \int_0^\infty \phi(r) \int_0^\infty \psi_r(u) \mu\left(\frac{B}{u}\right) du dr =$$

$$= \int_0^\infty n \cdot \mathrm{vol}(B(\ell_p^n)) \cdot u^{n-1} H(u) \mu\left(\frac{B}{u}\right) du = \int_B H(||x||_p) dx$$

In the last two steps we used, respectively, Fubini's theorem, and the polar coordinate integration formula of Proposition 1.

Having established the claim about the density of X', the lemma will follow by another application of Proposition 1:

$$||P_{a_n X'} - P_{G'}|| = \int_{\mathbb{R}^k} |f_{a_n X'} - f_{G'}| dx =$$

$$= k \cdot \mathrm{vol}(B(\ell_p^k)) \int_0^\infty r^{k-1} \int_{S(\ell_p^k)} \left| \frac{1}{a_n} H(r/a_n) - \frac{1}{(2\Gamma(1+1/p))^k} e^{-r^p} \right| d\mu(z) dr =$$

$$= k \cdot \mathrm{vol}(B(\ell_p^k)) \int_0^\infty r^{k-1} \left| \frac{1}{a_n} H(r/a_n) - \frac{1}{(2\Gamma(1+1/p))^k} e^{-r^p} \right| dr =$$

$$= \int_0^\infty |f_W(r) - f_Z(r)| dr = ||P_W - P_Z||.$$

Where in the last equality we used the fact that the density of W is equal to $k\mathrm{vol}(B(\ell_p^k)) r^{k-1} H(r/a_n)/a_n$ and the density of Z is proportional to e^{-r^p} (these facts also follow from Proposition 1). ■

Remark 1 *In the more abstract terminology of the Diaconis-Freedman paper, we have shown that the sigma-field in $\mathbb{R}^k$ pulled back from the Borel sets in $\mathbb{R}$ by the mapping $u \to ||u||_p$ is "sufficient", and this implies the lemma.*

Remark 2 *Since S' and S'' have $Gamma(k/p, 1)$ and $Gamma((n-k)/p, 1)$ distribution, respectively (see below), the density H that appeared in the proof of Lemma 4 can be computed. We get that the density of X' at $x \in B(\ell_p^k)$ is:*

$$H(||x||_p) = \frac{p\Gamma\left(\frac{n}{p}\right)}{n\Gamma\left(\frac{k}{p}\right)\Gamma\left(\frac{n-k}{p}\right)} \cdot ||x||_p^{(k-1)} \cdot \left(1 - ||x||_p^p\right)^{\frac{n-k}{p}-1}.$$

Z has distribution $Gamma(k/p, 1)$, since it is a sum of i.i.d. $Gamma(1/p, 1)$ components. The distribution of W is (a rescaling of) a Beta distribution - this can be seen again using the Schechtman-Zinn realization of the cone measure, which implies that W has the same distribution as

$$a_n^p \cdot \frac{\sum_{i=1}^k g_i^p}{\sum_{i=1}^k g_i^p + \sum_{i=k+1}^n g_i^p},$$

an expression of the form $a_n U/(U+V)$ where U and V are independent Gamma r.v.s with the same scale parameter. Thus W has distribution a_n^p times the $Beta(k/p, (n-k)/p)$ distribution. The total variation distance can therefore be estimated by Lemma 3 above, where $\alpha = k/p$, $\beta = (n-k)/p$. Finally, we have (with our choice of $a_n = (n/p)^{1/p}$ as above):

Theorem 3 *For $1 \le k \le \min\{n/2, n-p\}$, the following estimate holds:*

$$||P_{a_n X'} - P_{G'}|| \le \frac{4k + 12p}{n-k}$$

Note that the above discussion is for a random vector on $S(\ell_p^n)$ chosen according to the cone measure. Now, let $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$ be a *uniform* vector on $S(\ell_p^n)$ - that is, a vector chosen according to the normalized surface measure. Let $\mathbf{Y}' = (Y_1, \ldots, Y_k)$. Combining theorems 2 and 3 we have:

Theorem 4 *For $1 \le k < \min\{n/2, n-p\}$, the following estimate holds:*

$$||P_{a_n Y'} - P_{G'}|| \le \frac{4k + 12p}{n-k} + \frac{c_p}{\sqrt{n}}$$

4 Random Projections

This section deals with projections of the surface measure of $S(\ell_p^n)$ onto random subspaces. This section differs from the previous sections in the techniques used and the nature of the results proved.

As we noted in the introduction, these projections will be (with high probability) approximately Gaussian for high dimensions, as long as the subspace is of dimension much smaller than n. The estimates we give will not use the total variation metric, but rather a different metric between measures:

Definition 1 *Given two probability measures P and Q on $\mathbb{R}^n$, we define the T-distance between them as:*

$$T(P,Q) = \sup_{x\in\mathbb{R}^n} |P(\{y : \langle x,y\rangle \le 1\}) - Q(\{y : \langle x,y\rangle \le 1\})| =$$

$$= \sup\{|P(H) - Q(H)| : H \textit{ affine half-space}\}$$

The use of this metric in the present context was suggested to us by B. Tsirelson.

We will begin by formulating a general principle, which states a "Randomized Central Limit Theorem" for probability measures in $\mathbb{R}^n$ satisfying certain conditions. It should be pointed out, that this result holds for more general measures than volume measures on convex bodies, and is not one-dimensional. Therefore, our result is more general than the treatment of the so-called Central Limit Problem for Convex Bodies given by e.g. Antilla, Ball and Perissinaki [2]. This generality, however, leads to a worse dependence on ϵ in Theorem 5 below (see also the remark following Theorem 5).

Let us recall some basic definitions: Given a compactly supported probability measure P on $\mathbb{R}^n$, we say that it is isotropic, if for every $\theta \in S^{n-1}$,

$$\int_{\mathbb{R}^n} \langle x,\theta\rangle^2 dP(x) = 1.$$

We also say that P has the square negative correlation property, if for every $1 \le i,j \le n$:

$$\int_{\mathbb{R}^n} x_i^2 x_j^2 dP(x) \le \left(\int_{\mathbb{R}^n} x_i^2 dP(x)\right)\left(\int_{\mathbb{R}^n} x_j^2 dP(x)\right).$$

For every $k \le n$ we denote by $G(n,k)$ the Grassmanian Manifold of all k dimensional subspaces of $\mathbb{R}^n$, and by $\lambda_{n,k}$ we denote the normalized Haar measure on $G(n,k)$. For every k dimensional subspace $E \subset \mathbb{R}^n$ we denote by $\mathrm{Proj}_E(P)$ the orthogonal projection of P onto E. In what follows, γ_k is the standard k dimensional Gaussian measure. We can now state the main theorem:

Theorem 5 *Let P be a compactly supported, non-atomic isotropic measure on $\mathbb{R}^n$ satisfying the square negative correlation property. Define:*

$$B = \left(\int_{\mathbb{R}^n} ||x||_4^4 dP(x) \right)^{1/4}.$$

Then for every $\epsilon > 0$ and $k \le c_1 \epsilon^4 n^2 / B^4$ the following inequality holds:

$$\lambda_{n,k}\left(\left\{ E \in G(n,k) : T(Proj_E(P), \gamma_k) \ge \epsilon \right\} \right) \le \frac{c_2}{\epsilon} \exp\left(-\frac{c_3 n^2 \epsilon^4}{B^4} \right),$$

where c_1, c_2, c_3 are numerical constants.

Remark 3 *When P is the normalized volume measure on a symmetric convex body $K \subset \mathbb{R}^n$, and $k = 1$, the ϵ^4-term in the above estimate can be improved to ϵ^2. This was proved in [2] using Busemann's theorem. It is unclear whether a similar estimate can be proved in the full generality of the assumptions of Theorem 5.*

In what follows, we will always assume that P satisfies the above conditions. Denote by ω the normalized surface measure on the Euclidean sphere S^{n-1}.

Lemma 5 *Let $h : \mathbb{R} \to \mathbb{R}$ be a bounded Lipschitz function with constant L. Then for every $n \ge 3$*

$$\left| \int_{S^{n-1}} \int_{\mathbb{R}^n} h(\langle x, \theta \rangle) dP(x) d\omega(\theta) - \int_{\mathbb{R}} h d\gamma_1 \right| \le \frac{B^2 L + 50 \sup_{x \in \mathbb{R}} |h(x)|}{n}.$$

Proof:

$$\left| \int_{S^{n-1}} \int_{\mathbb{R}^n} h(\langle x, \theta \rangle) dP(x) d\omega(\theta) - \int_{\mathbb{R}} h d\gamma_1 \right| =$$

$$= \left| \int_{\mathbb{R}^n} \int_{S^{n-1}} h\left(||x||_2 \langle \frac{x}{||x||_2}, \theta \rangle \right) d\omega(\theta) dP(x) - \int_{\mathbb{R}} h d\gamma_1 \right| =$$

$$= \left| \int_{\mathbb{R}^n} \int_{S^{n-1}} h\left(||x||_2 \theta_1 \right) d\omega(\theta) dP(x) - \int_{\mathbb{R}} h d\gamma_1 \right| \le$$

$$\le \left| \int_{\mathbb{R}^n} \int_{S^{n-1}} h\left(||x||_2 \theta_1 \right) d\omega(\theta) dP(x) - \int_{S^{n-1}} h\left(\sqrt{n} \theta_1 \right) d\omega(\theta) \right| +$$

$$+\left|\int_{S^{n-1}} h\left(\sqrt{n}\theta_1\right) d\omega(\theta) - \int_{\mathbb{R}} h d\gamma_1\right|,$$

Where we have used the rotational invariance of the measure ω (θ_1 denoting the first coordinate of θ).

The second summand can be estimated by the special case $p = 2$, $k = 1$ of Theorem 3 (which is in fact part of the original statement proved by Diaconis and Freedman) as follows:

$$\left|\int_{S^{n-1}} h\left(\sqrt{n}\theta_1\right) d\omega(\theta) - \int_{\mathbb{R}} h d\gamma_1\right| \leq \frac{50 \sup_{x\in\mathbb{R}} |h(x)|}{n}.$$

Note that this is where the assumption $n \geq 3$ is used.

The first summand is estimated as follows:

$$\left|\int_{\mathbb{R}^n}\int_{S^{n-1}} h\left(||x||_2\theta_1\right) d\omega(\theta) dP(x) - \int_{S^{n-1}} h\left(\sqrt{n}\theta_1\right) d\omega(\theta)\right| \leq$$

$$\leq L\int_{R^n}\int_{S^{n-1}} \left|||x||_2 - \sqrt{n}\right| |\theta_1| d\omega(\theta) dP(x) =$$

$$= L\left(\int_{S^{n-1}} |\theta_1| d\omega(\theta)\right)\left(\int_{\mathbb{R}^n} \left|||x||_2 - \sqrt{n}\right| dP(x)\right) \leq$$

$$\leq L\left(\int_{S^{n-1}} \theta_1^2 d\omega(\theta)\right)^{1/2}\left(\int_{\mathbb{R}^n} \left|||x||_2 - \sqrt{n}\right| dP(x)\right) =$$

$$= \frac{L}{\sqrt{n}}\int_{\mathbb{R}^n} \left|||x||_2 - \sqrt{n}\right| dP(x) \leq \frac{L}{n}\int_{\mathbb{R}^n} \left|||x||_2^2 - n\right| dP(x) =$$

$$= \frac{L}{n}\int_{\mathbb{R}^n} \left|\sum_{i=1}^n (x_i^2 - 1)\right| dP(x) \leq \frac{L}{n}\left[\int_{\mathbb{R}^n}\left(\sum_{i=1}^n (x_i^2-1)\right)^2 dP(x)\right]^{1/2} =$$

$$= \frac{L}{n}\left[\int_{\mathbb{R}^n}\left(\sum_{i=1}^n (x_i^2-1)^2 + \sum_{i\neq j}(x_i^2-1)(x_j^2-1)\right) dP(x)\right]^{1/2} \leq \frac{LB^2}{n}.$$

Where in the last inequality we used the assumption that P is isotropic and the square negative correlation property. ∎

In order to prove Theorem 5 we will apply the following concentration inequality due to Gordon [8] (see also [18]). In both papers this inequality is proved in the text but is not specifically stated as a theorem . A weaker version of it (with a worse dependence on δ) is a classical interpretation of Levy's isoperimetric inequality on the sphere (see for example [12] Theorem 2.4). For the reader's convenience we will sketch the proof.

Theorem 6 *There are absolute constants $c_1, c_2, c_3 > 0$ such that for every $f : S^{n-1} \to \mathbb{R}$ which is Lipschitz with constant L (with respect to the Euclidean metric on S^{n-1}), for every $\delta > 0$ and for every $k \le c_1 \delta^2 n / L^2$:*

$$\lambda_{n,k}\left(\left\{E \in G(n,k) : \exists x \in E \cap S^{n-1} \ \left| f(x) - \int_{S^{n-1}} f(y)d\omega(y)\right| \ge \delta\right\}\right) \le$$
$$\le c_2 e^{-c_3 \delta^2 n / L^2}.$$

Denote by $e_1, \ldots, e_n$ the standard basis of $\mathbb{R}^n$ and let $\{X_{ij};\ i = 1, ..., k, \quad j = 1, ..., n\}$ be i.i.d. standard Gaussian random variables. For every $a \in S^{k-1}$ consider the random vector:

$$Y_a = \sum_{i=1}^{k} a_i \sum_{j=1}^{n} X_{ij} e_j.$$

Denote also by $G = \sum_{j=1}^{n} X_{1j} e_j$ the standard Gaussian random vector in $\mathbb{R}^n$.

In [18] the following result was proved for norms. Actually, the proof only uses the fact that a norm is Lipschitz (the parameter σ in [18] is precisely the Lipschitz constant of the norm). For general Lipschitz functions, the following proposition is the inequality obtained in the second line of page 276 of [18].

Proposition 2 *There are absolute constants $c_1, c_2, c_3 > 0$ such that if $h : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz with constant K, then for every $\eta > 0$ and $k \le c_1 \eta^2 / K^2$*

$$P\left(\sup_{a \in S^{k-1}} |h(Y_a) - \mathbb{E}h(G)| \ge \eta\right) \le c_2 e^{-c_3 \eta^2 / K^2}.$$

Sketch of the proof of Theorem 6: By translating f, we may assume without loss of generality that for every $x \in S^{n-1}$, $|f(x)| \le 2L$. Let $I = \int_{S^{n-1}} f d\omega$. Define $F_a = f(Y_a / ||Y_a||_2)$. Clearly, when $||a||_2 = 1$, $\mathbb{E}F_a = I$ and $\mathbb{E}||Y_a||_2 = \mathbb{E}||G||_2 = E_n \sim \sqrt{n}$. By standard arguments,

$$\lambda_{n,k}\left(\left\{E \in G(n,k) : \exists x \in E \cap S^{n-1} \ \left| f(x) - \int_{S^{n-1}} f(y)d\omega(y)\right| > \delta\right\}\right) =$$
$$= P\left(\sup_{a \in S^{k-1}} |F_a - I| > \delta\right).$$

Define $\tilde{f}(x) = ||x||_2 f(x/||x||_2)$. Since f is bounded by $2L$, $\tilde{f}$ is Lipschitz with constant $4L$ on $\mathbb{R}^n \setminus \{0\}$. Moreover, $\mathbb{E}\tilde{f}(Y_a) = I E_n$. Now, by Proposition 2:

$$P\left(\sup_{a \in S^{k-1}} |F_a - I| > \delta\right) \le$$

$$\leq P\left(\sup_{a\in S^{k-1}} \frac{|F_a|}{E_n}\big|||Y_a||_2 - E_n\big| + \sup_{a\in S^{k-1}} \frac{|\tilde{f}(Y_a) - IE_n|}{E_n} > \delta\right) \leq$$

$$\leq P\left(\sup_{a\in S^{k-1}} \big|||Y_a||_2 - E_n\big| \geq \frac{\delta E_n}{4L}\right) + P\left(\sup_{a\in S^{k-1}} |\tilde{f}(Y_a) - IE_n| \geq \frac{\delta E_n}{2}\right) \leq$$

$$\leq c_2 e^{-c_3\delta^2 E_n^2/(16L^2)} + c_2 e^{-c_3\delta^2 E_n^2/(64L^2)} \leq c_2' e^{-c_3' n\delta^2/L^2},$$

as long as $k \leq c_1'\delta^2 E_n^2/L^2$, which implies the required result.

■

In order to apply Lemma 5, we introduce the following functions:

$$h_{t,a}(x) = \begin{cases} 1 & x \leq t \\ \frac{(t+a)-x}{a} & t < x \leq t+a \\ 0 & t+a < x \end{cases} \qquad (t \in \mathbb{R}, a > 0)$$

It is clear that $h_{t,a}$ is Lipschitz with constant $1/a$. The following simple approximation result is the key to the application of Lemma 5:

Lemma 6 *Let $\epsilon > 0$. Then there exist $N = \lfloor 1/\epsilon \rfloor$ numbers $t_1, t_2, \ldots, t_N \in \mathbb{R}$ with the following property: If ν is a measure on $\mathbb{R}$ such that for all $i = 1, 2, \ldots, N$ we have*

$$\left|\int_{\mathbb{R}} h_{t_i,\epsilon} d\nu - \int_{\mathbb{R}} h_{t_i,\epsilon} d\gamma_1\right| \leq \epsilon$$

Then $T(\nu, \gamma_1) \leq 6\epsilon$. (Note that in dimension 1 the metric T is exactly the usual Kolmogorov metric.)

Proof: Denote as usual $\Phi(t) = \gamma_1((-\infty, t])$. Take $t_i = \Phi^{-1}(\epsilon \cdot i)$. Note that for some $0 < \theta < 1$,

$$t_{i+1} - t_i = \epsilon \cdot (\Phi^{-1})'(i + \theta \cdot \epsilon) \geq \epsilon\sqrt{2\pi} \geq \epsilon.$$

And therefore

$$\nu((-\infty, t_i]) \leq \int_{\mathbb{R}} h_{t_i,\epsilon} d\nu \leq \int_{\mathbb{R}} h_{t_i,\epsilon} d\gamma_1 + \epsilon \leq$$

$$\leq \Phi(t_{i+1}) + \epsilon = \Phi(t_i) + 2\epsilon.$$

Similarly

$$\nu((\infty, t_i]) \geq \Phi(t_i) - 2\epsilon.$$

We have shown that $|\nu((-\infty, t_i]) - \Phi(t_i)| \leq 2\epsilon$ for $i = 1, 2, \ldots, N$. It is now easy to show that this implies $|\nu((-\infty, t]) - \Phi(t)| \leq 6\epsilon$ for all $t \in \mathbb{R}$, as required. ■

Proof of Theorem 5. To begin with, note that since P is isotropic, for every $h:\mathbb{R}\to\mathbb{R}$ which is Lipschitz with constant L, the function that maps $u\in\mathbb{R}^n$ to $\int_{\mathbb{R}^n} h(\langle x,u\rangle)dP(x)$ is also Lipschitz with the same constant. Indeed, for every distinct $u,v\in\mathbb{R}^n$,

$$\left|\int_{\mathbb{R}^n} h(\langle x,u\rangle)dP(x) - \int_{\mathbb{R}^n} h(\langle x,v\rangle)dP(x)\right| \le$$

$$\le L||u-v||_2 \int_{\mathbb{R}^n}\left|\langle x, \frac{u-v}{||u-v||_2}\rangle\right| dP(x) \le$$

$$\le L||u-v||_2 \left(\int_{\mathbb{R}^n}\langle x, \frac{u-v}{||u-v||_2}\rangle^2 dP(x)\right)^{1/2} = L||u-v||_2.$$

Fix $\epsilon > 0$. By modifying the constant c_1, we can clearly assume that $40B/\sqrt{n}\le\epsilon\le 2$ (since the metric T is bounded by 2). Note that in this case, since $B\ge n^{1/4}$, it is easy to verify that

$$\frac{\epsilon}{6} - \frac{6B^2/\epsilon+50}{n} \ge \frac{\epsilon\sqrt{n}}{12B^2}.$$

Now, applying Lemmas 5 and 6 and Theorem 6, we get:

$$\lambda_{n,k}\left[\{E\in G(n,k) : T(\mathrm{Proj}_E(P),\gamma_k)\ge\epsilon\}\right] =$$

$$= \lambda_{n,k}\left[\left\{E\in G(n,k) : \exists u\in S^{n-1}\cap E \quad T(\mathrm{Proj}_{\mathbb{R}u}(P),\gamma_1)\ge\epsilon\right\}\right] \le$$

$$\le \lambda_{n,k}\left[\bigcup_{i=1}^{\lfloor 6/\epsilon\rfloor}\left\{E\in G(n,k) : \exists u\in S^{n-1}\cap E\right.\right.$$

$$\left.\left.\left|\int_{\mathbb{R}^n} h_{t_i,\epsilon/6}(\langle x,u\rangle)dP(x) - \int_{\mathbb{R}} h_{t_i,\epsilon/6}d\gamma_1\right|\ge\epsilon/6\right\}\right] \le$$

$$\le \lambda_{n,k}\left[\bigcup_{i=1}^{\lfloor 6/\epsilon\rfloor}\left\{E\in G(n,k) : \exists u\in S^{n-1}\cap E\right.\right.$$

$$\left|\int_{\mathbb{R}^n} h_{t_i,\epsilon/6}(\langle x,u\rangle)dP(x) - \int_{S^{n-1}}\int_{\mathbb{R}^n} h_{t_i,\epsilon/6}(\langle x,u\rangle)dP(x)d\omega(u)\right| \ge$$

$$\left.\left.\ge \frac{\epsilon}{6} - \frac{6B^2/\epsilon+50}{n}\right\}\right] \le$$

$$\le \lambda_{n,k}\left[\bigcup_{i=1}^{\lfloor 6/\epsilon\rfloor}\left\{E\in G(n,k) : \exists u\in S^{n-1}\cap E\right.\right.$$

$$\left|\int_{\mathbb{R}^n} h_{t_i,\epsilon/6}(\langle x,u\rangle)dP(x) - \int_{S^{n-1}}\int_{\mathbb{R}^n} h_{t_i,\epsilon/6}(\langle x,u\rangle)dP(x)d\omega(u)\right| \geq$$
$$\left.\left.\geq \frac{\epsilon\sqrt{n}}{12B^2}\right\}\right] \leq \frac{C}{\epsilon}\exp\left(-\frac{c\epsilon^4 n^2}{B^4}\right).$$
The last estimate uses Theorem 6, which is valid as long as $k \leq c_1\epsilon^4 n/B^4$, where c_1 is a (small enough) absolute constant. ■

We will now apply the above general result to the cone measure and the surface measure on the sphere of ℓ_p^n, beginning with the case of the cone measure. The measure μ is compactly supported and non-atomic, and it is isotropic up to multiplication by a constant: that is, we need to take the cone measure on some ℓ_p^n sphere other than the unit sphere - the exact constant is calculated below. But first, we prove the square negative correlation property, which in this case is quite simple. We will give an analytic proof of a more general fact. A similar result with the cone measure replaced by the volume measure on the ball of ℓ_p^n was proved in [5] and also in [2]. See also [13] for a an even stronger result for the cone measure.

Proposition 3 *For every $\alpha_1,\ldots,\alpha_n \geq 0$:*

$$\int_{S(\ell_p^n)}\prod_{k=1}^n |x_i|^{\alpha_i}d\mu(x) \leq \prod_{k=1}^n\int_{S(\ell_p^n)}|x_1|^{\alpha_i}d\mu(x).$$

Proof: Using (again) the notation of Theorem 1,

$$\int_{S(\ell_p^n)}\prod_{k=1}^n |x_k|^{\alpha_k}d\mu(x) = \mathbb{E}\left[S^{-\frac{1}{p}\sum_{i=1}^n\alpha_i}\prod_{k=1}^n|g_k|^{\alpha_k}\right] =$$
$$= \frac{\prod_{k=1}^n\mathbb{E}|g_1|^{\alpha_k}}{\mathbb{E}\left[S^{\frac{1}{p}\sum_{i=1}^n\alpha_i}\right]} = \frac{\prod_{k=1}^n\mathbb{E}|g_1|^{\alpha_k}}{\prod_{i=1}^n\left[\mathbb{E}S^{\frac{1}{p}\sum_{i=1}^n\alpha_i}\right]^{\alpha_i/\sum_{j=1}^n\alpha_j}} \leq$$
$$\leq \frac{\prod_{k=1}^n\mathbb{E}|g_1|^{\alpha_k}}{\prod_{i=1}^n\mathbb{E}\left[\left(S^{\frac{1}{p}\sum_{i=1}^n\alpha_i}\right)^{\alpha_i/\sum_{j=1}^n\alpha_j}\right]} = \prod_{k=1}^n\frac{\mathbb{E}|g_1|^{\alpha_k}}{\mathbb{E}S^{\alpha_k/p}} =$$
$$= \prod_{k=1}^n\mathbb{E}\frac{|g_k|^{\alpha_k}}{S^{\alpha_k/p}} = \prod_{k=1}^n\int_{S(\ell_p^n)}|x_k|^{\alpha_k}d\mu(x).$$

■

Remark 4 *Note that the only property used in the above proof is the independence of S and X, rather than their specific distributions. However, this property is known to characterize these distributions - see [4].*

Because the coordinates of x are μ-uncorrelated (an obvious geometric truth) it is easily seen that for every $\theta \in S^{n-1}$,

$$\int \langle x, \theta \rangle^2 d\mu(x) = \int x_1^2 d\mu(x).$$

The right hand side can be calculated, using the methods of the previous sections:

$$\int x_1^2 d\mu(x) = \mathbb{E}\left[\frac{g_1^2}{(\sum_{i=1}^n |g_i|^p)^{2/p}}\right] = \frac{\mathbb{E}g_1^2}{\mathbb{E}(\sum_{i=1}^n |g_i|^p)^{2/p}} = \frac{\Gamma(3/p)\cdot\Gamma(n/p)}{\Gamma(1/p)\cdot\Gamma((n+2)/p)}$$

Define, therefore, $a_{n,p} = \left[\frac{\Gamma(1/p)}{\Gamma(3/p)} \cdot \frac{\Gamma((n+2)/p)}{\Gamma(n/p)}\right]^{1/2}$. Take as before a random vector $X = (X_1, \ldots, X_n)$ whose distribution law is μ. Then $\tilde{\mu}$, the distribution measure of $Y = a_{n,p} \cdot X$, is a compactly supported, nonatomic, isotropic measure. We now proceed to calculate the constant B of this measure, using again the log-convexity of the gamma function (or the case $n = 2$ of Lemma 7):

$$B^4 = \mathbb{E}||Y||_4^4 = n\mathbb{E}Y_1^4 = na_{n,p}^4\mathbb{E}X_1^4 = na_{n,p}^4 \cdot \frac{\Gamma(5/p)\Gamma(n/p)}{\Gamma(1/p)\Gamma((n+4)/p)} =$$

$$= \frac{\Gamma(1/p)\Gamma(5/p)}{\Gamma(3/p)^2} \cdot \frac{\Gamma((n+2)/p)^2}{\Gamma(n/p)\Gamma((n+4)/p)} \cdot n \leq c^4 \cdot n$$

And so $B \leq c \cdot n^{1/4}$ where c is an absolute constant (although the the bound may seem to depend on p, it is easy to check that it is bounded by a numerical constant as long as $p \geq 1$). Putting all the pieces together, using Theorem 5, we have finally:

Theorem 7 *Let μ be the cone measure on $S(\ell_p^n)$, and define $\tilde{\mu}(A) = \mu(A/a_{n,p})$. Then for every $\epsilon > 0$ and $k \leq c\epsilon^4 n$:*

$$\lambda_{n,k}\left[\{E \in G(n,k) : T(Proj_E(\tilde{\mu}), \gamma_k) \geq \epsilon\}\right] \leq C\exp(-cn\epsilon^4).$$

Here $c, C > 0$ are absolute constants.

The estimate of the total variation distance between μ and σ allows us to transfer the above result immediately to a result for σ.

Theorem 8 *Let σ be the surface measure on $S(\ell_p^n)$, $p \geq 1$, and define $\tilde{\sigma}(A) = \sigma(A/a_{n,p})$. Then for every $\epsilon > 0$ and $k \leq c\epsilon^4 n$:*

$$\lambda_{n,k}\left[\{E \in G(n,k) : T(Proj_E(\tilde{\sigma}), \gamma_k) \geq \epsilon\}\right] \leq C_p \exp(-c_p n \epsilon^4),$$

where $c_p, C_p > 0$ are constants (which may depend on p).

Proof: By Theorem 2, for every $E \in G(n,k)$,

$$T(Proj_E(\tilde{\mu}), Proj_E(\tilde{\sigma})) \leq ||\mu - \sigma|| \leq \frac{c_p}{\sqrt{n}}.$$

Hence, Theorem 7 implies Theorem 8 as long as $\epsilon \geq 2c_p/\sqrt{n}$. By modifying the constant C_p, the theorem follows for every ϵ. ■

Remark 5 *As was remarked in the discussion following the statement of Theorem 2, the results in [13] imply in particular that for $p \geq 1$ the constants C_p, c_p in Theorem 8 may be taken to be independent of p.*

5 Concluding remarks

Several natural questions arise from our results:

1. We conjecture that any convex body in $\mathbb{R}^n$ has a linear image for which the surface measure and cone measure are close in total variation distance. It seems reasonable that we can estimate the above distance by a constant multiple of $1/\sqrt{n}$. Our estimates of the total variation distance between the cone measure and surface measure on $S(\ell_p^n)$ are tight and it is possible to calculate the exact dependence of the constants on p. The interested reader is referred to the paper [13], in which the first named author studies in greater depth the precise relation between the surface and cone measures on $S(\ell_p^n)$.

2. For the purpose of improving the estimate that appears in Theorem 4, it would be natural to bound the total variation distance between the projections of the surface and cone measures onto the first k coordinates, although we have not attempted to do this. The concrete density that was computed in Remark 2 may prove to be useful for such an estimate.

3. The metric T that appears in Section 4 is just one of many possible metrics on probability measures that could be used. We can in fact state similar results for the Kolmogorov distance and other natural metrics. We chose to deal with the metric T since this is a natural rotation invariant metric for which the proofs are the simplest. One might think of Theorem 5 as a measure-theoretic version of Dvoretzky's Theorem. Our result is not in complete analogy with Dvoretzky's theorem, since it does not reflect the dependence of the dimension k on p that appears there. The total variation distance is an example of a very natural metric which our methods seem insufficient to handle.

4. The Central Limit Problem, is the problem of proving that almost all projections of the volume measure of a convex body K onto 1-dimensional subspaces, are approximately Gaussian in high dimension. The results of Section 4 show that in order to prove this Central Limit property for a body K in isotropic position, it is enough to show that the cone measure on K has the square negative correlation property (As was noted in the introduction, the above fact is also proved in [2].) The square negative correlation property seems geometrically plausible when K is unconditional (i.e. when the norm of K is invariant with respect to sign changes and permutations of the coordinates) - this can be verified by elementary calculations in dimension 2, but we are presently unable to prove it for arbitrary dimension.

5. From the proof of Theorem 2 it follows that its conclusion holds as long as $\mathbb{E}|g|^{2q} < \infty$. This is true when $2q = 4p - 4 > -1$, or $p > 3/4$. We believe that this restriction is unnecessary, i.e. the statement of Theorem 2 holds for any $p > 0$. A proof of this would involve proving Theorem 2 without passing to the second moment. Preliminary calculations show that it may be possible to avoid the second moment, but the calculations quickly become tedious and beyond the scope of the present paper. We have therefore chosen to focus on the convex range $p \geq 1$ (i.e. the setting of the Central Limit Problem).

Acknowledgment: The authors would like to express their sincere gratitude to the referee, whose careful reading of the article, and many helpful suggestions, considerably improved the presentation of this paper.

References

[1] E. Artin, *The Gamma Function*, Holt, Rinehart and Winston 1964.

[2] M. Antilla, K. Ball, I. Perissinaki, *The central limit problem for convex bodies*, Trans. Amer. Math. Soc., to appear.

[3] K.A. Borovkov, *Convergence of projections of uniform distributions on balls. (Russian)*, Teor. Veroyatnost. i Primenen. 35 (1990), no. 3, 547-551. Translation in Theory Probab. Appl. 35 (1991), no. 3, 546-550.

[4] F. Barthe, M. Csörnyei, A. Naor, *A note on simultaneous polar and Cartesian decomposition*. Preprint.

[5] K. Ball, I. Perissinaki, *The subindependence of coordinate slabs in ℓ_p^n balls.* Israel J. Math. 107 (1998), 289-299.

[6] U. Brehm, J. Voigt, *Asymptotics of cross sections of convex bodies.* Beiträge Algebra Geom. 41 (2000), 437-454.

[7] P. Diaconis, D. Freedman, *A dozen de Finetti-style results in search of a theory.* Ann. Inst. Henri Poincaré **23** (1987) 397-423.

[8] Y. Gordon, *On Milman's inequality and random subspaces which escape through a mesh in $\mathbb{R}^n$*. Geometric aspects of functional analysis (1986/87), 84-106, Lecture Notes in Math., 1317, Springer, 1988.

[9] A. Koldobosky, M. Lifshits, *Average volume of sections of star bodies.* Geometric aspects of functional analysis (1996-2000), 119-146, Lecture Notes in Math., 1745, Springer, 2000.

[10] P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces*, Cambridge University Press 1995.

[11] A.A. Mogul'skiĭ, *de Finetti-type results for ℓ_p. (Russian)*, Sibirsk. Mat. Zh. 32 (1991), no. 4, 88-95. Translation in Siberian Math. J. 32 (1992), no. 4, 609-616.

[12] V. Milman, G. Schechtman, *Asymptotic Theory of Finite Dimensional Normed Spaces*, Springer, 1986.

[13] A. Naor, *The cone measure and surface measure on the sphere of ℓ_p^n.* Preprint.

[14] S.T. Rachev, *Probability Metrics and the Stability of Stochastic Models*, Wiley 1991.

[15] S.T. Rachev, L. Rüschendorf, *Approximate independence of distributions on spheres and their stability properties*, Ann. Probab. 19 (1991), no. 3, 1311-1337.

[16] D. Romik, *Measure concentration techniques for randomized central limit theorems.* Preprint (1998).

[17] D. Romik, *Projections of product measures.* Preprint (1999).

[18] G. Schechtman, *A remark concerning the dependence on ϵ in Dvoretzky's theorem.* Geometric aspects of functional analysis (1987/88), 274-277, Lecture Notes in Math., 1376, Springer, 1989.

[19] G. Schechtman, J. Zinn, *On the volume of the intersection of two L_p^n balls.* Proc. Amer. Math. Soc. **110** (1990) , 1, 217-224.

AMS Mathematics subject classification (2000): 60F05, 52A20.