ISRAEL JOURNAL OF MATHEMATICS **132** (2002), 337–358

PROPORTIONAL CONCENTRATION PHENOMENA ON THE SPHERE

BY

SHIRI ARTSTEIN*

School of Mathematical Sciences, Tel Aviv University
Ramat Aviv, Tel Aviv 69978, Israel
e-mail: artst@post.tau.ac.il

ABSTRACT

In this paper we establish concentration phenomena for subspaces with arbitrary dimension. Namely, we display conditions under which the Haar measure on the sphere concentrates on a neighborhood of the intersection of the sphere with a subspace of R^n of a given dimension. We display applications to a problem of projections of points on the sphere, and to the duality of entropy numbers conjecture.

1. Introduction

The classical concentration phenomenon refers to concentration of the Haar measure on the sphere S^{n-1} around the intersection with $(n-1)$-dimensional subspaces. Namely, for any fixed ε, the area in $S^{n-1}$of the ε-neighborhood of an $(n-1)$-dimensional subspace converges to 1 when $n \to \infty$. This phenomenon and its extensions have attracted much attention, and have a variety of applications in the Asymptotic Theory of Normed Spaces; see, e.g., [7], [4]. The complementary case is that of the area in S^{n-1} of the ε-neighborhood of a 1-dimensional subspace. It converges, of course, to 0 when $n \to \infty$. In both the case of dimension 1 and the case of co-dimension 1, the convergence is exponentially fast, a fact which plays a central role in the applications.

* Research was partially supported by the Israel Science Foundation.
Received August 13, 2001

The results displayed here relate to the limit as $n \to \infty$ of the area in S^{n-1} of the ε-neighborhood of a k-dimensional subspace E_k, for the full range of k between 3 and $n-3$. A particular case is when k is proportional to n, namely $k = \lambda n$. In this case it is not difficult to show (we do this in section 2) that for every $0 < \lambda < 1$, there exists a critical value $\varepsilon(\lambda)$, determined by the formula $\sin \varepsilon(\lambda) = \sqrt{1-\lambda}$, such that: If $\varepsilon > \varepsilon(\lambda)$ then $\mu((E_k)_\varepsilon) \to 1$ as $n \to \infty$, and if $\varepsilon < \varepsilon(\lambda)$ then $\mu((E_k)_\varepsilon) \to 0$ as $n \to \infty$ (as customary, $(E_k)_\varepsilon$ denotes the set $\{x \in S^{n-1} : \rho(x, E_k \cap S^{n-1}) < \varepsilon\}$ where ρ is the geodesic distance on the sphere, and μ denotes the normalized rotation invariant Haar measure on the sphere). Our main concern is finding the exact rates of convergence. Approximations for these rates turn out to be useful in several applications. We show that in both the case $\varepsilon > \varepsilon(\lambda)$ and the case $\varepsilon < \varepsilon(\lambda)$, the rate of convergence is exponential in n, namely $1 - e^{-\gamma n}$ and $e^{-\gamma n}$ respectively, where $\gamma = \gamma(\lambda, \delta)$ is a constant depending only on λ and on $\delta = \varepsilon - \varepsilon(\lambda)$. We provide precise estimates for the constant $\gamma(\lambda, \delta)$.

The established estimates are then applied to several problems in the Asymptotic Theory of Normed Spaces. The first application deals with projections of points on the sphere into lower dimensional subspaces as follows. A random point on S^{n-1}, projected on a λn-dimensional subspace, has, with high probability, euclidean norm close to $\sqrt{\lambda}$. We estimate precisely this probability. This enables us, for instance, to provide an isomorphic Johnson–Lindenstrauss lemma. The second application is concerned with the duality of entropy numbers conjecture. In the special case where one of the bodies is the euclidean ball $B(l_2^n)$, and where the covering number of K (a general convex body) by $B(l_2^n)$ is exponentially large, we show duality, namely that the covering number of $B(l_2^n)$ by the polar body K° is also exponentially large. We also use the estimates to compare, following Diaconis and Freedman, the distribution of the first k coordinates of a random vector on the $(n-1)$-dimensional sphere with a random gaussian vector with k coordinates.

The paper is organized as follows. In section 2 we display a useful representation and derive the critical value of the concentration. Section 3 contains the main estimate when λ and ε are fixed and $n \to \infty$. This estimate is generalized in section 4 to the case in which λ and ε are not necessarily fixed and may change with n. Section 5 includes several special cases, among them a case which is useful in the applications. In section 6 we give the application regarding projections of points on the sphere. The latter result is further applied in section 7 to the duality of entropy numbers question. In chapter 8 we give the estimates related to the mentioned result by Diaconis and Freedman.

ACKNOWLEDGEMENT: The research was carried out under the supervision of Professor V. D. Milman. I wish to thank Professor Milman for the many contributions to the paper and for his attentive guidance.

2. A useful representation of the spherical area

In this section we represent the area of the ε-neighborhood of a k-dimensional subspace by means of Beta-distributed random variables. Through this representation, some of the properties of the behavior of the area become transparent. The area, on the sphere, of the ε-neighborhood of a k-dimensional subspace is

$$\mu((E_k)_\varepsilon) = \frac{1}{\int_0^{\pi/2} \sin^{n-k-1} x \cos^{k-1} x dx} \int_0^\varepsilon \sin^{n-k-1} x \cos^{k-1} x dx.$$

By the change of variables $u = \sin^2 x$ one gets

$$\mu((E_k)_\varepsilon) = \frac{1}{\int_0^1 u^{\frac{n-k}{2}-1}(1-u)^{\frac{k}{2}-1} du} \int_0^{\sin^2 \varepsilon} u^{\frac{n-k}{2}-1}(1-u)^{\frac{k}{2}-1} du. \tag{2.1}$$

Recall that the integral $\int_0^1 u^{m-1}(1-u)^{l-1} du$ coincides with the Beta function

$$\mathrm{Beta}(m, l) = \frac{\Gamma(m)\Gamma(l)}{\Gamma(m+l)} = \frac{(m-1)!(l-1)!}{(m+l-1)!}.$$

This can be easily verified by induction. The expression (2.1) implies, therefore, that as a function of ε, the measure $\mu((E_k)_\varepsilon)$ is the distribution of a Beta random variable with parameters $(\frac{n-k}{2}, \frac{k}{2})$. The expectation of such a variable is $\frac{n-k}{n}$. Letting $k = \lambda n$, we can write

$$\mu((E_k)_\varepsilon) = \frac{\Gamma(n/2)}{\Gamma(\lambda n/2)\Gamma((1-\lambda)n/2)} \int_0^{\sin^2 \varepsilon} u^{\frac{(1-\lambda)n}{2}-1}(1-u)^{\frac{\lambda n}{2}-1} du. \tag{2.2}$$

For a random variable Y_n with distribution $\mathrm{Beta}((1-\lambda)\frac{n}{2}, \lambda\frac{n}{2})$, we have

$$\begin{aligned} \mu((E_k)_\varepsilon) &= \mathrm{Prob}[Y_n \le \sin^2 \varepsilon], \\ \mathrm{E}[Y_n] &= (1-\lambda), \end{aligned} \tag{2.3}$$

and

$$\mathrm{Var}[Y_n] = \frac{\lambda(1-\lambda)}{n/2+1} \to_{n\to\infty} 0.$$

Remark 2.1: From the derived representation it is possible to get, easily, the existence of, and the formula for, the critical value $\varepsilon(\lambda)$. The expectation of the distance of a random point x from the subspace E_k is determined by the expectation of Y_n which is independent of n, and it is therefore $\varepsilon(\lambda) = \arcsin(\sqrt{1-\lambda})$. Approximately half of the measure on the sphere is within a distance smaller than $\varepsilon(\lambda)$ from the subspace. The argument is that the median and the expectation are close. This already verifies the existence of the critical value of ε. Moreover, the Beta random variables Y_n concentrate, when $n \to \infty$, around their mutual expectation. This holds, for example, since the variance of Y_n tends to 0 as $n \to \infty$. By (2.3) we see that, therefore, the measure on the sphere concentrates within distance $\varepsilon(\lambda)$ of the subspace.

Remark 2.2: Note that the ε-neighborhoods we took were with respect to the geodesic distance on the sphere. If, instead, we would compute the measure with respect to the euclidean distance in R^n, the term "$\sin \varepsilon$" in formula (2.2) would become simply "ε".

Finding the critical value was straightforward. A deeper investigation of the asymptotic deviation of a Beta variable from its mean (using methods to handle Beta random variables with large parameters with a constant ratio) is needed in order to get the convergences rates. This is described in the next section.

3. The main estimate

In this section we display and prove the main result of the paper. Throughout what follows we use the following notation. For variables A and B depending on n, we write $A \simeq B$ in the following two meanings.

(1) If both A and B are close to 0, $A \simeq B$ means that $\frac{A}{B} \to 1$ as $n \to \infty$.

(2) If both A and B are close to 1, $A \simeq B$ means that $\frac{1-A}{1-B} \to 1$ as $n \to \infty$.

When A and B depend on several parameters, we specify which parameters are fixed, and we mean that the convergence is uniform with respect to all other parameters.

THEOREM 3.1: *Let E_k be a k-dimensional subspace of R^n, and denote by $\mu((E_k)_\varepsilon)$ the Haar measure on the sphere S^{n-1} of the set of points within a geodesic distance smaller than ε of E_k. We write $k = \lambda n$. Fix $0 < \varepsilon < \pi/2$ and $0 < \lambda < 1$. The following estimates hold as $n \to \infty$.*

(i) *If $\sin^2 \varepsilon > 1 - \lambda$, then*

$$\mu((E_k)_\varepsilon) \simeq 1 - \frac{1}{\sqrt{n\pi}} \frac{\sqrt{\lambda(1-\lambda)}}{\sin^2 \varepsilon - (1-\lambda)} e^{\frac{n}{2} u(\lambda, \varepsilon)}. \tag{3.1}$$

(ii) *If* $\sin^2\varepsilon < 1-\lambda$, *then*

$$\mu((E_k)_\varepsilon) \simeq \frac{1}{\sqrt{n\pi}} \frac{\sqrt{\lambda(1-\lambda)}}{(1-\lambda)-\sin^2\varepsilon} e^{\frac{n}{2}u(\lambda,\varepsilon)} \tag{3.2}$$

where $u(\lambda,\varepsilon) = (1-\lambda)\ln\frac{(1-\lambda)}{\sin^2\varepsilon} + \lambda\ln\frac{\lambda}{\cos^2\varepsilon}$.

For the proof of Theorem 3.1 we need the following preliminary observations. We employ ideas given in Alfers and Dinjes [1] for comparing a $\mathrm{Beta}(\alpha m,(1-\alpha)m)$ distributed random variable with a standard Gaussian random variable, where m is large and $0<\alpha<1$ is constant. Following [1], define the two functions:

$$A(\alpha,p) = \mathrm{sign}(\alpha-p)\sqrt{2}\Big((1-\alpha)\ln\Big(\frac{1-\alpha}{1-p}\Big) + \alpha\ln\frac{\alpha}{p}\Big)^{1/2}$$

(thus $u(\lambda,\varepsilon) = A^2((1-\lambda),\sin^2\varepsilon)/2)$, and

$$D(\alpha,a) = \frac{\sqrt{\alpha(1-\alpha)}}{\alpha-p}A(\alpha,p) \quad \text{where } a = A(\alpha,p).$$

The mapping $A(\alpha,p)$ is well defined (the term in the parentheses is always positive), and for every fixed α it maps $(0,1)$ bijectively onto $(-\infty,\infty)$. Notice, however, that if $A(\alpha,p)=0$ then $\alpha=p$, and then $D(\alpha,0)$ is not defined. We define it in the natural way as the limit of $D(\alpha,a)$ when $a\to 0$ and get that $D(\alpha,0)=1$ for every α. A straightforward calculation (using the change of variables $a=A(\alpha,t)$, the observation that $\frac{\partial}{\partial p}A(\alpha,p) = -\frac{\alpha-p}{p(1-p)A(\alpha,p)}$, the Stirling formula and the behavior of $D(\alpha,a)$) yields the following results.

Lemma 3.2 (See [1, Theorem 1.1]): *If* Y *is beta-distributed for the parameters* $(\alpha m,(1-\alpha)m)$ *then*

$$\mathrm{Prob}[Y\ge p] = e^{S^m(\alpha)}\int_{-\infty}^{A(\alpha,p)}\sqrt{\frac{m}{2\pi}}e^{-\frac{m}{2}a^2}D(\alpha,a)da \tag{3.3}$$

where $S^m(\alpha)$ *is very small for large* m, *namely* $S^m(\alpha)\simeq(\frac{1}{12m})(1-\frac{1}{\alpha}-\frac{1}{1-\alpha})_{m\to\infty}\to 0$.

Corollary 3.3 (See [1, Corollary 1]): *Let* Z *be a standard Gaussian random variable, and let* Y_m *be Beta-distributed with parameters* $(\alpha m,(1-\alpha)m)$. *For* $m\to\infty$ *and for fixed* p

(i) *If* $p>\alpha$ *then*

$$\lim_{m\to\infty}\frac{\mathrm{Prob}[Y_m\ge p]}{\mathrm{Prob}[Z\le\sqrt{m}A(\alpha,p)]} = D(\alpha,A(\alpha,p)). \tag{3.4}$$

(ii) *If $p < \alpha$ then*

$$\lim_{m\to\infty} \frac{\text{Prob}[Y_m \le p]}{\text{Prob}[Z \ge \sqrt{m}A(\alpha, p)]} = D(\alpha, A(\alpha, p)). \tag{3.5}$$

(Note that the case $p < a$ is not covered in [1, Corollary 1].)

Following [1], another result can be obtained as follows (we use it in section 4).

PROPOSITION 3.4 (See [1, Theorems 2.1' and 2.1"]): *Assume Y is a Beta-distributed random variable for the parameters $(\alpha m, (1-\alpha)m)$, and denote $\alpha' = \alpha\frac{m}{m-1}$ and $\alpha'' = \frac{\alpha m - 1}{m-1}$. Then*

$$\begin{aligned} \text{Prob}[Y \le p] &\ge \text{Prob}[Z \ge \sqrt{m-1}A(\alpha', p)], \\ \text{Prob}[Y \le p] &\le \text{Prob}[Z \ge \sqrt{m-1}A(\alpha'', p)]. \end{aligned} \tag{3.6}$$

Proof of Theorem 3.1: We use Corollary 3.3 with $m = n/2$, $\alpha = (1-\lambda)$, and $p = \sin^2\varepsilon$. Again, let Z be a standard Gaussian random variable. First examine the case $p > 1-\lambda$. Denote by Φ the Gaussian distribution function $\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}t^2}dt$. Then $\text{Prob}[Z \le \sqrt{\frac{n}{2}}A((1-\lambda), p)] = \Phi(\sqrt{\frac{n}{2}}A((1-\lambda), p))$. Notice that in this case $A((1-\lambda), p)$ is negative. Differentiation easily yields the following approximation for the Gaussian integral for y positive:

$$\frac{1}{\sqrt{2\pi}}\frac{1}{y^{-1}+y}e^{-\frac{y^2}{2}} \le \frac{1}{\sqrt{2\pi}}\int_y^{\infty} e^{-\frac{1}{2}x^2}dx \le \frac{1}{\sqrt{2\pi}}\frac{1}{y}e^{-\frac{y^2}{2}}. \tag{3.7}$$

This approximation implies that

$$\text{Prob}[Z \le \sqrt{\frac{n}{2}}A((1-\lambda), p)] \simeq \frac{1}{\sqrt{2\pi}}\frac{e^{-\frac{n}{2}(A^2((1-\lambda,p)/2))}}{\sqrt{n/2}(-A((1-\lambda), p))}$$

(here, and throughout this section, both λ and ε ($= \arcsin\sqrt{p}$) are fixed). A direct substitution yields

$$D((1-\lambda), A((1-\lambda), p)) = \frac{\sqrt{\lambda(1-\lambda)}}{(1-\lambda)-p}A((1-\lambda), p).$$

Therefore, using Corollary 3.3, we get the estimate

$$\begin{aligned} \text{Prob}[Y \ge p] &\simeq \Phi\left(\sqrt{\frac{n}{2}}A((1-\lambda), p)\right) D((1-\lambda), A((1-\lambda), p)) \\ &\simeq \frac{1}{\sqrt{n\pi}}\frac{\sqrt{\lambda(1-\lambda)}}{p-(1-\lambda)}e^{-\frac{n}{2}(A^2((1-\lambda),p)/2)}. \end{aligned}$$

In other words, if $\sin^2\varepsilon > (1-\lambda)$ then

$$\begin{aligned}\mathrm{Prob}[Y \le \sin^2\varepsilon] &= 1 - \mathrm{Prob}[Y \ge \sin^2\varepsilon]\\ &\simeq 1 - \frac{1}{\sqrt{n\pi}}\frac{\sqrt{\lambda(1-\lambda)}}{\sin^2\varepsilon - (1-\lambda)}e^{-\frac{n}{2}(A^2((1-\lambda),\sin^2\varepsilon)/2)}.\end{aligned}$$

This completes the proof in the case $\sin^2\varepsilon > (1-\lambda)$. The case $\sin^2\varepsilon < (1-\lambda)$ follows, in fact, from the former case. Assume $\sin^2\varepsilon < (1-\lambda)$, and let Y' be a Beta-distributed random variable with parameters $(\lambda n/2, (1-\lambda)n/2)$. Then Y' has the same distribution as $1-Y$ and therefore (since now $\cos^2\varepsilon > \lambda$) we get

$$\begin{aligned}\mathrm{Prob}[Y \le \sin^2\varepsilon] &= \mathrm{Prob}[Y' \ge \cos^2\varepsilon]\\ &\simeq \frac{1}{\sqrt{n\pi}}\frac{\sqrt{\lambda(1-\lambda)}}{(\cos^2\varepsilon - \lambda)}e^{-\frac{n}{2}(A^2(\lambda,\cos^2\varepsilon)/2)}.\end{aligned}$$

Since $A(t,s) = A(1-t, 1-s)$, the proof is complete.

4. The case of λ and ε not necessarily fixed

The estimates in the previous section apply to the case $n \to \infty$ with λ and ε fixed. In many problems, however, we would like to allow λ to change with n. To cover such instances we use Proposition 3.4. This proposition gives estimates from above and from below, rather than limit estimates, for the required probabilities. The expressions we get resemble the estimates attained for λ and ε fixed.

We first give Theorem 4.1, which is in the most general form. In Theorem 4.2 we place an extra condition which simplifies the estimates. This condition can be relaxed, to form a condition which holds in all cases relevant to us in this paper. The result is given in Theorem 4.3.

THEOREM 4.1: *Let $n \ge 0$, $3 \le k \le n-3$, $\lambda = \frac{k}{n}$, and let E_k be a k-dimensional subspace of R^n. For $\varepsilon > 0$, let $\mu((E_k)_\varepsilon)$ be the Haar measure on the sphere S^{n-1} of all the points within a geodesic distance smaller than ε of E_k. Denote $l = \frac{\sin^2\varepsilon}{1-\lambda}$ and $l' = \frac{\cos^2\varepsilon}{\lambda}$. Then there exist positive constants $c_{\lambda,n}$ and $c'_{\lambda,n}$, bounded by an absolute constant M, such that:*

(i) *If $\sin^2\varepsilon < 1-\lambda$ then*

$$\begin{aligned}\frac{1}{\sqrt{2\pi}}\frac{e^{-u-c'_{\lambda,n}-\ln l'}}{\frac{1}{\sqrt{u+c'_{\lambda,n}+\ln l'}} + \sqrt{u+c'_{\lambda,n}+\ln l'}} &\le \mu((E_k)_\varepsilon)\\ &\le \frac{1}{\sqrt{2\pi}}\frac{e^{-u-c_{\lambda,n}-\ln l}}{\sqrt{u+c_{\lambda,n}+\ln l}}.\end{aligned} \tag{4.1}$$

(ii) *If* $\sin^2\varepsilon > 1-\lambda$ *then*

(4.2)

$$1 - \frac{1}{\sqrt{2\pi}} \frac{e^{-u-c'_{\lambda,n}-\ln l'}}{\frac{1}{\sqrt{u+c'_{\lambda,n}+\ln l'}} + \sqrt{u+c'_{\lambda,n}+\ln l'}} \le \mu((E_k)_\varepsilon)$$

$$\le 1 - \frac{1}{\sqrt{2\pi}} \frac{e^{-u-c_{\lambda,n}-\ln l}}{\sqrt{u+c_{\lambda,n}+\ln l}}$$

where $u = \frac{n}{2}[(1-\lambda)\ln\frac{(1-\lambda)}{\sin^2\varepsilon} + \lambda\ln\frac{\lambda}{\cos^2\varepsilon}]$.

If we restrict our discussion to the case where the terms l and l' in Theorem 4.1 are bounded from below and from above, we can give a simpler formulation of the result.

THEOREM 4.2: *Let* $n \ge 0$, $3 \le k \le n-3$, $\lambda = k/n$, *and let* E_k *be a* k*-dimensional subspace of* R^n. *Fix* $l_1 > 0$ *and* $l_2 < \infty$, *and assume that* $l_1 < \frac{\sin^2\varepsilon}{1-\lambda} < l_2$, $l_1 < \frac{\cos^2\varepsilon}{\lambda} < l_2$. *Then there exist absolute constants* $c, c', c_1, \dots, c_4$, *depending only on* l_1 *and* l_2, *such that:*

(i) *If* $\sin^2\varepsilon < 1-\lambda$, *then*

$$c_1\frac{e^{-u}}{\sqrt{u+c}} \le \mu((E_k)_\varepsilon) \le c_2\frac{e^{-u}}{\sqrt{u+c'}}. \tag{4.3}$$

(ii) *If* $\sin^2\varepsilon > 1-\lambda$, *then*

$$1 - c_3\frac{e^{-u}}{\sqrt{u+c'}} \le \mu((E_k)_\varepsilon) \le 1 - c_4\frac{e^{-u}}{\sqrt{u+c}} \tag{4.4}$$

where $u = \frac{n}{2}[(1-\lambda)\ln\frac{(1-\lambda)}{\sin^2\varepsilon} + \lambda\ln\frac{\lambda}{\cos^2\varepsilon}]$.

The restrictions of boundedness imposed on l and l' in the statement of Theorem 4.2 can be relaxed. In the next theorem we assume, instead, only that l and l' are between $1/n$ and n.

THEOREM 4.3: *Let* $n \ge 0$, $3 \le k \le n-3$, $\lambda = k/n$, *and let* E_k *be a* k*-dimensional subspace of* R^n. *Assume that*

$$\frac{1}{n} < \frac{\sin^2\varepsilon}{1-\lambda} < n, \quad \frac{1}{n} < \frac{\cos^2\varepsilon}{\lambda} < n.$$

Then there exist absolute constants $c, c', c_1, \dots, c_4$, *and a sequence* $\alpha_n \to 1$, *such that:*

(i) *If* $\sin^2\varepsilon < 1-\lambda$, *then*

$$c_1\frac{e^{-\alpha_n u}}{\sqrt{u+c}} \le \mu((E_k)_\varepsilon) \le c_2\frac{e^{-\alpha_n u}}{\sqrt{u+c'}}. \tag{4.3}$$

(ii) *If* $\sin^2\varepsilon > 1-\lambda$, *then*

$$1 - c_3\frac{e^{-\alpha_n u}}{\sqrt{u+c'}} \le \mu((E_k)_\varepsilon) \le 1 - c_4\frac{e^{-\alpha_n u}}{\sqrt{u+c}} \tag{4.4}$$

where $u = \frac{n}{2}[(1-\lambda)\ln\frac{(1-\lambda)}{\sin^2\varepsilon} + \lambda\ln\frac{\lambda}{\cos^2\varepsilon}]$.

Proof of Theorems: Proposition 3.4 together with the approximation (3.7) give, for $\sin^2\varepsilon < 1-\lambda$, that

$$\frac{1}{\sqrt{2\pi}}\frac{e^{-\xi'}}{\frac{1}{\sqrt{\xi'}}+\sqrt{\xi'}} \le \mu((E_k)_\varepsilon) \le \frac{1}{\sqrt{2\pi}}\frac{e^{-\xi}}{\sqrt{\xi}},$$

and for $\sin^2\varepsilon > 1-\lambda$, that

$$1 - \frac{1}{\sqrt{2\pi}}\frac{e^{-\xi'}}{\sqrt{\xi'}} \le \mu((E_k))_\varepsilon \le 1 - \frac{1}{\sqrt{2\pi}}\frac{e^{-\xi}}{\frac{1}{\sqrt{\xi}}+\sqrt{\xi}},$$

where

$$\begin{aligned}\xi =& \frac{n}{2}\Big((1-\lambda)\Big(1-\frac{2}{(1-\lambda)n}\Big)\ln\Big(\frac{(1-\lambda)(1-\frac{2}{(1-\lambda)n})(1+\frac{2}{(n-2)})}{\sin^2\varepsilon}\Big)\\ &+\lambda\ln\Big(\frac{\lambda(1+\frac{2}{n-2})}{\cos^2\varepsilon}\Big)\Big),\\ \xi' =& \frac{n}{2}\Big((1-\lambda)\ln\Big(\frac{(1-\lambda)(1+\frac{2}{n-2})}{\sin^2\varepsilon}\Big)+\lambda\Big(1-\frac{2}{\lambda n}\Big)\ln\Big(\frac{\lambda(1-\frac{2}{\lambda n})(1+\frac{2}{n-2})}{\cos^2\varepsilon}\Big)\Big).\end{aligned}$$

Rearrangement of terms yields

$$\begin{aligned}\xi &= u + \ln\Big(1+\frac{2}{n-2}\Big)^{\frac{n-2}{2}} + \ln\Big(1-\frac{2}{(1-\lambda)n}\Big)^{\frac{(1-\lambda)n}{2}} + \ln\frac{\sin^2\varepsilon}{(1-\lambda-\frac{2}{n})},\\ \xi' &= u + \ln\Big(1+\frac{2}{n-2}\Big)^{\frac{n-2}{2}} + \ln\Big(1-\frac{2}{\lambda n}\Big)^{\frac{\lambda n}{2}} + \ln\frac{\cos^2\varepsilon}{\lambda-\frac{2}{n}}\end{aligned}$$

when u is spelled out in the statement of the Theorems. Restriction to the case $(1-\lambda)n \ge 3$ and $\lambda n \ge 3$ implies the existence of positive constants $c_{\lambda,n}$ and $c'_{\lambda,n}$, bounded by a constant M such that

$$\xi = u + c_{\lambda,n} + \ln l, \quad \xi' = u + c'_{\lambda,n} + \ln l'.$$

This proves Theorem 4.1. In the case of Theorem 4.2, for large enough n, ξ and ξ' are bigger than 1 (actually they are tending to infinity) and hence the additional $1/\sqrt{\xi'}$ or $1/\sqrt{\xi}$ in the denominator of the displayed inequalities can be ignored, resulting perhaps in a slightly larger constant. Since l and l' are assumed bounded from below and from above, we can substitute $\ln l$ and $\ln l'$ by constants as well, and the proof of Theorem 4.2 is also complete.

To prove Theorem 4.3, we use Theorem 4.1 and the boundedness assumption in a similar way.

5. Some special cases

In this section we provide useful expressions for the general estimates in several particular cases which may become handy in applications. The computations and substitutions are straightforward, and are included for completeness.

CASE 5.1: $\sin^2\varepsilon = (1-\lambda)\pm\lambda\delta$, where $\lambda \to 0$, $\lambda \geq c''/n$ for an absolute constant c'', and $0 < \delta < 1$ is fixed.

(a) If $\sin^2\varepsilon = (1-\lambda)+\lambda\delta$, then

(5.1)
$$\mu((E_k)_\varepsilon) \simeq 1 - c\frac{1}{\sqrt{\frac{n}{2}(\lambda(\ln(\frac{1}{1-\delta}) - \delta) + O(\lambda^2))}} e^{-\frac{n}{2}(\lambda(\ln(\frac{1}{1-\delta})-\delta)+O(\lambda^2))}.$$

(b) If $\sin^2\varepsilon = (1-\lambda)-\lambda\delta$, then

$$\mu((E_k)_\varepsilon) \simeq \frac{c'}{\sqrt{\frac{n}{2}(\lambda(\ln(\frac{1}{1+\delta}) + \delta) + O(\lambda^2))}} e^{-\frac{n}{2}(\lambda(\ln(\frac{1}{1+\delta})+\delta)+O(\lambda^2))} \tag{5.2}$$

where, in general, c and c' depend on δ. The constants c and c' become absolute if we further assume δ to be bounded away from 1 (say if we assume $\delta < 9/10$).

In order to verify (5.1) and (5.2) we establish the boundedness condition required from the expressions $\frac{\sin^2\varepsilon}{1-\lambda}$ and $\frac{\cos^2\varepsilon}{\lambda}$ in the statement of Theorem 4.2. Such a bound exists for a fixed δ, and it is uniform with respect to δ bounded away from 1. Now, Theorem 4.2 implies the following. When $\sin^2\varepsilon = (1-\lambda)+\lambda\delta$,

$$\mu((E_k)_\varepsilon) \simeq 1 - \frac{c}{\sqrt{-\frac{n}{2}(A^2((1-\lambda),\sin^2\varepsilon)/2)}} e^{-\frac{n}{2}(A^2((1-\lambda),\sin^2\varepsilon)/2)}.$$

We can make use of the equality

$$\begin{aligned}\frac{1}{2}A^2((1-\lambda),(1-\lambda)+\lambda\delta) &= (1-\lambda)\ln\Big(\frac{1-\lambda}{1-\lambda+\lambda\delta}\Big)+\lambda\ln\Big(\frac{1}{1-\delta}\Big)\\ &= -(1-\lambda)\ln\Big(1+\frac{\lambda\delta}{1-\lambda}\Big)-\lambda\ln(1-\delta)\\ &= -\lambda\delta+\frac{1}{2}\lambda^2\delta^2/(1-\lambda)+\cdots-\lambda\ln(1-\delta)\\ &= \lambda\Big(\ln\Big(\frac{1}{1-\delta}\Big)-\delta\Big)+O(\lambda^2),\end{aligned}$$

and get the estimate, thus verify, (5.1).

When $\sin^2\varepsilon=(1-\lambda)-\lambda\delta$ (namely $\cos^2\varepsilon=\lambda(1+\delta)$), we look at

$$\mu((E_k)_\varepsilon)=\text{Prob}[Y'>\cos^2\varepsilon]\simeq\frac{c'}{\sqrt{\frac{n}{2}(A^2(\lambda,\cos^2\varepsilon)/2)}}e^{-\frac{n}{2}(A^2(\lambda,\cos^2\varepsilon)/2)}.$$

Now we make use of the equality

$$\begin{aligned}\frac{1}{2}A^2(\lambda,\lambda(1+\delta)) &= \lambda\ln\Big(\frac{1}{1+\delta}\Big)+(1-\lambda)\ln\Big(\frac{1-\lambda}{1-\lambda-\lambda\delta}\Big)\\ &= \lambda\ln\Big(\frac{1}{1+\delta}\Big)-(1-\lambda)\ln\Big(1-\frac{\lambda}{1-\lambda}\delta\Big)\\ &= \lambda\ln\Big(\frac{1}{1+\delta}\Big)+\lambda\delta+\frac{1}{2}\lambda^2\delta^2/(1-\lambda)+\cdots\\ &= \lambda\Big(\delta+\ln\Big(\frac{1}{1+\delta}\Big)\Big)+O(\lambda^2)\end{aligned}$$

and deduce estimate (5.2).

Remark: If in the previous example we assume that δ is small, the terms in (5.1) and (5.2) are of order $e^{-n\delta^2}$. In the classical concentration result a term $e^{-n\varepsilon^2}$ appears. We note that although there is a square in both expressions, ε^2 and δ^2, they originate differently. In classical concentration the square appears when one computes the modulus of convexity of the euclidean ball. In our setting the square is a result of the behavior of the term u in Theorem 4.1 for small δ. This is the reason that in Case 5.4 below, when we use the formula in Theorem 4.1 for large δ, we get completely different results. Indeed, the expression $(\ln(\frac{1}{1-\delta})-\delta)$, which for small δ behaves like δ^2, explodes when $\delta\to 1$.

CASE 5.2: $\sin^2\varepsilon=(1-\lambda)\pm(1-\lambda)\delta$, where $\lambda\to 1$ and $0<\delta<1$ is fixed.

(a) If $\sin^2\varepsilon = (1-\lambda) + (1-\lambda)\delta$, then

(5.3)
$$\mu((E_k)_\varepsilon) \simeq 1- \frac{c}{\sqrt{\frac{n}{2}((1-\lambda)(\ln(\frac{1}{1+\delta})+\delta)+O((1-\lambda)^2))}} e^{-\frac{n}{2}((1-\lambda)(\ln(\frac{1}{1+\delta})+\delta)+O((1-\lambda)^2))}.$$

(b) If $\sin^2\varepsilon = (1-\lambda) - (1-\lambda)\delta$, then

(5.4)
$$\mu((E_k)_\varepsilon) \simeq \frac{c'}{\sqrt{\frac{n}{2}((1-\lambda)(\ln(\frac{1}{1-\delta})-\delta)+O((1-\lambda)^2))}} e^{-\frac{n}{2}((1-\lambda)(\ln(\frac{1}{1-\delta})-\delta)+O((1-\lambda)^2))}$$

where, again, in general c and c' depend on δ, but become absolute if δ is assumed to be bounded away from 1.

Verifying (5.3) and (5.4) follows the same lines as in example 5.1.

Case 5.3: $\sin^2\varepsilon = (1-\lambda) \pm \delta$, where $0<\lambda<1$ is fixed and $\delta \to 0$.

(a) If $\sin^2\varepsilon = (1-\lambda)+\delta$, then

$$\mu((E_k)_\varepsilon) \simeq 1 - \frac{c}{\sqrt{\frac{n}{2}(\frac{\delta^2}{2\lambda(1-\lambda)}+O(\delta^3))}} e^{-\frac{n}{2}(\frac{\delta^2}{2\lambda(1-\lambda)}+O(\delta^3))}. \tag{5.5}$$

(b) If $\sin^2\varepsilon = (1-\lambda)-\delta$, then

$$\mu((E_k)_\varepsilon) \simeq \frac{c'}{\sqrt{\frac{n}{2}(\frac{\delta^2}{2\lambda(1-\lambda)}+O(\delta^3))}} e^{-\frac{n}{2}(\frac{\delta^2}{2\lambda(1-\lambda)}+O(\delta^3))} \tag{5.6}$$

where, in general, the constants c and c' depend on λ. If λ is assumed to be bounded away both from 0 and from 1, the constants become absolute.

In order to verify (5.5) and (5.6) we establish again the boundedness conditions for the expressions $\frac{\sin^2\varepsilon}{1-\lambda}$ and $\frac{\cos^2\varepsilon}{\lambda}$ in the statement of Theorem 4.2. Such a bound exists for a fixed λ, and it is uniform with respect to λ bounded away from 0 and from 1. Now, Theorem 4.2 implies the following.

When $\sin^2\varepsilon = (1-\lambda)+\delta$,

$$\mu((E_k)_\varepsilon) \simeq 1 - \frac{c}{\sqrt{\frac{n}{2}(A^2((1-\lambda),\sin^2\varepsilon)/2)}} e^{-\frac{n}{2}(A^2((1-\lambda),\sin^2\varepsilon)/2)}.$$

We can make use of the equality

$$\begin{aligned}\frac{1}{2}A^2((1-\lambda),(1-\lambda)+\delta) &= (1-\lambda)\ln\left(\frac{1-\lambda}{1-\lambda+\delta}\right)+\lambda\ln\left(\frac{\lambda}{\lambda-\delta}\right)\\ &= \frac{1}{2}\frac{\delta^2}{\lambda(1-\lambda)}+O(\delta^3)\end{aligned}$$

and get the desired estimate (5.5). The second case, (5.6), is verified in the same way.

CASE 5.4: $\sin^2\varepsilon = (1-\lambda) - (1-\lambda)\delta$, where λ is close to 1. (Note that the assumption here is the same as in Case 5.2, except that here no conditions are imposed on δ. We make use of this specific condition in the proof of Proposition 6.2 below.) In this case we get an inequality as follows:

$$(5.7) \qquad \mu((E_k)_\varepsilon) \leq c\frac{e^{-(1-\lambda)(\ln(\frac{1}{1-\delta})-\delta)n/2+\ln(\frac{1}{1-\delta})}}{\sqrt{(1-\lambda)(\ln(\frac{1}{1-\delta})-\delta)n/2+\ln(\frac{1}{1-\delta})+c'}}.$$

In order to verify (5.7) we use Theorem 4.1. (We cannot use, as we did in the previous cases, Theorem 4.2, since the boundedness conditions in Theorem 4.2 do not hold uniformly for δ close to 1.) The terms u, l and l' in Theorem 4.1 are in this case

$$u = \frac{n}{2}\Big((1-\lambda)\Big(\ln\Big(\frac{1}{1-\delta}\Big)-\delta\Big)+\frac{1}{2}(1-\lambda)^2\delta^2/\lambda^2-\frac{1}{3}(1-\lambda)^3\delta^3/\lambda^3+\cdots\Big),$$
$$l = \frac{\sin^2\varepsilon}{1-\lambda} = 1-\delta,$$
$$l' = \frac{\cos^2\varepsilon}{\lambda} = 1+\frac{1-\lambda}{\lambda}\delta.$$

Therefore

$$u+\ln l \geq \frac{n}{2}(1-\lambda)\Big(\ln\Big(\frac{1}{1-\delta}\Big)-\delta\Big)+\ln(1-\delta).$$

Inserting this inequality in Theorem 4.1 implies (5.7).

Similar inequalities can be deduced for the case $\sin^2\varepsilon = (1-\lambda)+(1-\lambda)\delta$, and likewise for the other inequality in Theorem 4.1. We omit the details.

6. An application to a projection problem

Consider the orthogonal projection of a random (uniformly distributed) point x in the $(n-1)$-dimensional sphere onto a k-dimensional subspace. The expectation of $|Px|$, the euclidean norm of the projection Px, is close to $\sqrt{k/n}$. The reason is that the expectation of the square of the euclidean norm is exactly k/n. Moreover, the set of points x for which $|Px|$ is close to $\sqrt{k/n}$ has a large measure. This gives rise to the following general question, several aspects of which we address in this section.

QUESTION 6.1: *Given a set of cardinality N on S^{n-1}, does there exist a subspace of dimension k such that the euclidean norms of the elements in the projection of the set on the subspace are not far from $\sqrt{k/n}$?*

An answer to this question can be one of two types. First, for a given N (which depends on the dimension n), one may provide an estimate $\Delta(N)$ such that whenever a set of cardinality N on the sphere is given, a k-dimensional subspace exists such that all the projections on the subspace of the points in the set have euclidean norms which are $\Delta(N)$ close to $\sqrt{k/n}$. Second, for a given degree of closeness Δ, one may estimate the maximal cardinality $N(\Delta)$ such that whenever a set of cardinality $N(\Delta)$ is given, a k-dimensional subspace exists such that all the projections on the subspace of the points in the set have euclidean norms which are Δ-close to $\sqrt{k/n}$. In both cases we are interested in the behavior of the estimates for large dimension n. We refer to an answer of the first type as an isomorphic answer and to the second type, for a fixed Δ independent of dimension, as a Δ-isometric answer. In the first part of this section we give a general, however not transparent, answer to Question 6.1, addressing both types of estimates. In the second part of the section we provide a more concrete form of the answer in a particular case (which we then use in the next section).

Our way to establish the existence of a subspace with a certain property is to show that the measure of subspaces in $G_{n,k}$ with this property is positive. ($G_{n,k}$ is, as customary, the Grassman manifold of k-dimensional subspaces of R^n, endowed with the normalized Haar measure.)

To provide our answer to Question 6.1, we need the following observations. First, the measure of subspaces in $G_{n,k}$ with a certain property is the same as the measure of orthogonal transformations $U \in O(n)$ such that $U(E_0)$ has this property, where E_0 is some fixed subspace of dimension k. (Here $O(n)$ is the group of orthogonal transformations on R^n endowed with the normalized Haar measure.) Second, consider a set of points $\{x_i\}_{i=1}^N$ in R^n. Placing a restriction on the norms of their projections on $U(E_0)$ is equivalent to placing the same restriction on the norms of the projections to E_0 of the set $\{U^{-1}x_i\}_{i=1}^N$. This is implied by the equalities $P_{UE_0}x = P_{UE_0}UU^{-1}x = UP_{E_0}U^{-1}x$. Third, consider again a set of points $\{x_i\}_{i=1}^N$ in R^n. If $1 - \sum_{i=1}^N \mu\{U \in O(n) : P_{E_0}U^{-1}x_i \notin I\} > 0$, for a given set $I \subset [0,1]$, then the measure of the family of orthogonal transformations satisfying $|P_{E_0}U^{-1}x_i| \in I$ for every $1 \le i \le N$, is positive.

For $k = (1-\lambda)n$ and for $\Delta > 0$ define

$$\mu_0 = \mu\{x \in S^{n-1} : ||P_{E_k}x| - \sqrt{k/n}| < \Delta\}. \tag{6.1}$$

Taking into account these three observations, we see that if $N < 1/\mu_0$, then for every set $\{x_i\}_{i=1}^N$ in S^{n-1} there exists a subspace E_k of dimension k such that for every x_i, $i = 1, \ldots, n$, $|\mathrm{P}_{E_k}x_i - \sqrt{k/n}| < \Delta$. The above enables us to establish relations $N(\Delta)$ and $\Delta(N)$ as an answer to Question 6.1. Indeed, given Δ, we can use sections 3, 4 and 5 to estimate $\mu_0 = \mu_0(\Delta)$, using the following evident equality:

$$\begin{aligned}\mu_0 &= \mu\{x : \sqrt{1-\lambda} - \Delta < |\mathrm{P}_{E_k} x| < \sqrt{1-\lambda} + \Delta)\} \\ &= \mu\{x : \mathrm{d}(x, E_{\lambda n})^2 < (\sqrt{1-\lambda}+\Delta)^2\} - \mu\{x : \mathrm{d}(x, E_{\lambda n})^2 < (\sqrt{1-\lambda}-\Delta)^2\},\end{aligned}$$

and then have the estimate $N(\Delta) = 1/\mu_0(\Delta)$. To get an estimate for $\Delta(N)$, we do the reverse. In Proposition 6.4 we use the above scheme for a special choice of N.

Remark 6.2: The preceding derivations, and in particular the estimate for $N(\Delta)$, are close in spirit to the Johnson–Lindenstrauss Lemma; see [5]. The Lemma gives an estimate for the smallest dimension $k(n)$ such that any subset of cardinality n of l_2^n can be $(1+\epsilon)$-isometrically embedded into $l_2^{k(n)}$. Such an embedding can be realized by projecting into a well chosen subspace of the appropriate dimension, and by dividing then the images by $\sqrt{\lambda}$. The reason the method works is the following. Instead of the set of points, say $\mathcal{N}$, in l_2^n, one considers the set

$$\mathcal{N}' = \left\{\frac{x_i - x_j}{|x_i - x_j|}\right\}$$

of normalized differences. To this end a restriction has to be imposed on the cardinality of $\mathcal{N}$, or equivalently on the dimension k (the exact restriction can easily be computed). Then the existence of a subspace E_k such that for every $y \in \mathcal{N}'$ the euclidean norm of its projection onto E_k is close to $\sqrt{\lambda}$ is guaranteed. This insures that the distance between $\mathrm{P}x_i$ and $\mathrm{P}x_j$ is close to $\sqrt{\lambda}$ times the distance between x_i and x_j. In other words, the relative distances do not change much. Thus, the resemblance to the derivations in the present paper is apparent. However, while in the Johnson–Lindenstrauss Lemma the embedding was required to be an ε-isometry, the present paper offers a general isomorphic version.

Remark 6.3: The following estimate was used in an article by Milman and Pajor [9],

$$\mathrm{Prob}[T \in O(n) : |P_{E_k} T x| > \xi |x|] \le \left(e(1-\xi^2)\frac{1}{1-\lambda}\right)^{(1-\lambda)n/2}. \tag{6.2}$$

In the terminology of the present paper the left hand side of (6.2) has the following form,

$$\mathrm{Prob}[x \in S^{n-1} : \mathrm{dist}^2_{\|\cdot\|_2}(x, E_k) < 1 - \xi^2].$$

Estimates for this expression better than (6.2) can be obtained from the results in sections 3–5. Such estimates would be asymptotically accurate. Moreover, (6.2) has meaning only when $\xi^2 > \lambda(1+\frac{1-\lambda}{\lambda}(1-\frac{1}{e}))$, while using the technique of the present paper yields an answer whenever $\xi^2 > \lambda$.

We turn now to the special case for which we get a more concrete answer to Question 6.1. Suppose that e^{ck} points in S^{n-1} are projected into a k-dimensional subspace E_k. If c is large, it could be that for no subspace E_k all the norms of the projections are close to $\sqrt{k/n}$. We can, however, choose E_k such that all of the projections do not enter some small neighborhood of 0, namely we can establish some kind of isomorphic result. The precise result is as follows.

PROPOSITION 6.4: *For any number c, there exists an $\varepsilon(c)$ (for instance, $\varepsilon(c) = e^{-(2c+1)}$ will work) such that the following holds for every $\lambda = k/n$ fixed. For n large enough, whenever a set $\mathcal{N}$ in S^{n-1} is of cardinality $|\mathcal{N}| = e^{ck}$, there exists a subspace E of dimension k such that*

$$\sqrt{\lambda}\varepsilon(c) < |P_E x| < \sqrt{\lambda}(1+\varepsilon(c)) \tag{6.3}$$

for every $x \in \mathcal{N}$. Moreover, by taking a small enough $\varepsilon(c)$ one can ensure that the measure of the set of subspaces in $G_{n,k}$ satisfying (6.3) is arbitrarily close to 1 (for instance, for $\varepsilon(c) = \frac{1}{2}e^{-(2c+1)}$, the latter measure is larger than $1-e^{\frac{1}{2}(1-\lambda)n}$).

Proof: We will show how to establish the left hand side inequality with $\varepsilon(c) = e^{-(c+1)}$. This is the only part we use in the application. The right hand side is attained in a similar way, and for the two inequalities to hold together we reduce $\varepsilon(c)$ to the magnitude mentioned in the statement of the theorem. To prove the left hand side, first note that the projections of N points do not enter an ε-neighborhood of 0 if and only if they all stay within a distance more than ε from $E_k^{\perp}$. Using the same reasoning as in the general answer above, we find that for the latter condition to hold it is enough that N be smaller than $1/\mu((E_{n-k})_\varepsilon)$. For $1-\lambda = k/n$ small, and when $\sin^2\varepsilon = (1-\lambda)(1-\delta)$, we use Case 5.4 to get that the inequality

$$N \le \frac{\sqrt{\frac{k}{2}(\ln(\frac{1}{1-\delta})-\delta)-\ln(\frac{1}{1-\delta})}}{c'} e^{\frac{k}{2}(\ln(\frac{1}{1-\delta})-\delta)-\ln(\frac{1}{1-\delta})} \tag{6.4}$$

ensures that $N < 1/\mu((E_{n-k})_\varepsilon)$. Since the function $(\ln(\frac{1}{1-\delta})-\delta)$ can be arbitrarily large for δ close to 1, we see that for any number N of points on the sphere, there exists a $0<\delta<1$, and a subspace E_k of dimension k such that the projections, on E_k, of all the points are out of a $\sqrt{(1-\lambda)(1-\delta)}$-neighborhood of 0. For the specific case $N = e^{ck}$ mentioned in the statement of Proposition

6.4, we identify $\delta = \delta(c)$ for which (6.4) holds. To this end, for k large enough, it is enough to ask that

$$ck \leq \frac{k}{2}\Big(\ln\Big(\frac{1}{1-\delta}\Big) - \delta\Big) - \ln\Big(\frac{1}{1-\delta}\Big),$$

or equivalently that $1-\delta \sim e^{-2c}$. This identifies the desired δ. In particular, the size of the neighborhood of 0 is

$$\varepsilon(c) = \sqrt{(1-\lambda)(1-\delta)} \sim \sqrt{(1-\lambda)}e^{-c}.$$

This verifies the first claim of Proposition 6.4. By choosing a smaller $\varepsilon(c)$ and following the same line of proof, one can verify that the inequality (6.3) is satisfied for a large measure of subspaces, and complete the proof.

We may apply Proposition 6.4 to the set of normalized differences

$$\mathcal{N}' = \Big\{\frac{x-y}{|x-y|} : x, y \in \mathcal{N}\Big\},$$

and get the following.

COROLLARY 6.5: *For any number c, there exists $\varepsilon'(c)$ (for instance, $\varepsilon'(c) = e^{-(2c+1)}$) such that the following holds for every $\lambda = k/n$ fixed. For n large enough, whenever a set $\mathcal{N}$ in R^n is of cardinality $|\mathcal{N}| = e^{ck}$, there exists a subspace E of dimension k such that*

$$\sqrt{\lambda}\varepsilon'(c)|x-y| < |P_E x - P_E y| < \sqrt{\lambda}(1+\varepsilon'(c))|x-y| \tag{6.5}$$

for every $x, y \in \mathcal{N}$.

Remark 6.6: Consider the projection of the sphere S^{n-1} into R^k, with $k = \lambda n$. The proportional concentration phenomenon displayed in section 2 implies that for every $\varepsilon > 0$ fixed, the measure of the set of points in S^{n-1} which are projected into the ε-neighborhood of the sphere of radius $\sqrt{\lambda}$ in R^k tends to 1 as $n \to \infty$. Using Theorem 3.1 we get the following result, replacing the fixed ε by a neighborhood of width $1/\sqrt{n}$. There exist constants $c_1 > 0$ and $c_2 < 1$ such that

$$c_1 < \mu\Big\{x : |\mathrm{P}_{E_k} x| \in \Big(\sqrt{\lambda} - \frac{\lambda}{\sqrt{n}}, \sqrt{\lambda} + \frac{1-\lambda}{\sqrt{n}}\Big)\Big\} < c_2. \tag{6.7}$$

Notice that (6.7) estimates the measure of the projection of the given strip with constants $c_1 > 0$ and $c_2 < 1$ independent of n.

7. An application to the duality of entropy numbers conjecture

In this section we use Corollary 6.5 to answer a question regarding entropy numbers. Recall the covering number, $N(K,T)$, of K by T, where K and T are two convex bodies in R^n, given by

$$N(K,T) = \min\{N : \exists\{x_i\}_{i=1}^N \subset K, K \subset \bigcup_{i=1}^N(\{x_i\}+T)\}.$$

The duality conjecture, due to B. Carl and A. Piestch, suggests the existence of absolute constants α and β such that for every pair of centrally symmetric convex bodies T and K the inequality

$$N(K,T) \le (N(T^\circ, \alpha K^\circ))^\beta \tag{7.1}$$

holds, where T° and K° denote the polar bodies of T and K, respectively.

In Theorem 7.2 below we verify the duality conjecture in the special case where $T = B(l_2^n)$ and $\log N(K, B(l_2^n)) = \gamma n$. We need the following result from [6].

THEOREM 7.1 (König and Milman): *There exists a universal constant C such that for any n and any two convex bodies $K, T \subset R^n$,*

$$\frac{1}{C} \le \Big(\frac{N(K,T)}{N(T^\circ,K^\circ)}\Big)^{\frac{1}{n}} \le C. \tag{7.2}$$

THEOREM 7.2: *Assume that $N(K, B(l_2^n)) = 2^{c_1 n}$.*

(a) *If $c_1 > 2\log_2 C$, then*

$$N(B(l_2^n), K^\circ) \ge 2^{c_1 n/2}. \tag{7.3}$$

(b) *If $c_1 < 2\log_2 C$, then*

$$N(B(l_2^n), (\frac{1}{4\sqrt{\log_2 C}})\sqrt{c_1}2^{-c_1}K^\circ) \ge 2^{c_1 n/2}, \tag{7.4}$$

where C is the absolute constant guaranteed in Theorem 7.1.

Proof of Theorem 7.2: Case (a) follows immediately from (7.2). Indeed, since we assume that $N(K, B(l_2^n)) = 2^{c_1 n}$, it follows from Theorem 7.1 that $N(B(l_2^n), K^\circ) \ge (2^{c_1}/C)^n$. This together with $c_1 > 2\log_2 C$ implies (7.3). It is not possible to use the same argument in case (b) since in this case c_1 is too small. Our strategy to handle case (b) is to use Corollary 6.5 in order to reduce the dimension, and only then use (7.2). The original dimension can then be recovered with the help of purely geometric arguments. Here are the details. Take a 1-separated set $\{x_1, \ldots, x_{2^{c_1 n}}\}$ in K. Such a large separated set exists. For example, a maximal 1-separated set is always a 1-covering, and since assuming that $N(K, B(l_2^n)) = 2^{c_1 n}$, it follows that a 1-covering collection has at least $2^{c_1 n}$ elements. Chose a dimension $k < n$ with $2^{c_1 n} > C^k$. By Corollary 6.5, for any given set of $2^{c_1 n}$ points, we can find a subspace E of dimension k such that after projecting the set of points into E, the mutual distances do not shrink by more than $\sqrt{\frac{k}{n}}\varepsilon'(c_1)$. This means that the new set, $\{\mathrm{P}_E x_i : 1 \leq i \leq 2^{c_1 n}\} \subset P_E K$, is ε_0-separated, for $\varepsilon_0 = \sqrt{\frac{k}{n}} e^{-2c_1+1}$. Since every $z_j + \frac{\varepsilon_0}{2} B(l_2^n)$ is of diameter ε_0 and therefore can include only one point of this new set, we get

$$N(P_E K, \frac{\varepsilon_0}{2}(B(l_2^n) \cap E)) \geq 2^{c_1 n} = 2^{(c_1 \frac{n}{k})k}.$$

Now, since in case (b), $2^{c_1 \frac{n}{k}} > C$, we can apply (7.2) to the latter formula and extract the following meaningful estimate:

$$N((B(l_2^n) \cap E)^\circ, \frac{\varepsilon_0}{2}(\mathrm{P}_E K)^\circ) \geq 2^{(c_1 \frac{n}{k} - \log_2 C)k}.$$

Using the facts that $(\mathrm{P}_E K)^\circ = K^\circ \cap E$ and that $(B(l_2^n) \cap E)^\circ = B(l_2^n) \cap E$, and the property that for every T symmetric and convex $N(B(l_2^n) \cap E, T \cap E) \leq N(D, \frac{1}{2}T)$, we get

$$N(B(l_2^n), \frac{\varepsilon_0}{4} K^\circ) \geq 2^{(c_1 \frac{n}{k} - \log_2 C)k}.$$

Applying $k = \frac{c_1 n}{2 \log_2 C}$ to the latter formula verifies (7.4). This completes the proof.

We offer now some extensions and comparisons.

Remark 7.3: The reasoning in the preceding proof can be applied without an assumption on the cardinality of $N(K, B(l_2^n))$. The estimate that would emerge is

$$N(B(l_2^n), c\sqrt{\frac{\ln N(K, (B(l_2^n)))}{n}} K^\circ) \geq (N(K, B(l_2^n)))^2. \tag{7.5}$$

In the language of entropy numbers inequality (7.5) has a more appealing form (recall the definition of entropy numbers for $u \colon X \to Y$, namely $e_k(u) = \inf\{\varepsilon : N(B(Y), \varepsilon B(X)) \leq 2^k\}$), as follows. For $u \colon l_2{}^n \to X$,

$$e_k(u^*) \geq c\sqrt{\frac{k}{n}} e_{2k}(u). \tag{7.6}$$

Note that the duality conjecture in this case suggests that the term $\sqrt{k/n}$ can be eliminated. This is still an open problem. It should be mentioned that results stronger than (7.6) regarding the duality conjecture in this case were established by V. D. Milman and S. J. Szarek in [8]. These results, however, do not improve Theorem 7.2.

Remark 7.4: Theorem 7.2 can be compared with a result of G. Pisier, written originally for operators of rank $\leq n$ (see [11, Corollary 2.4]). In the language of covering numbers it reads as follows. Under the condition $N(K, B(l_2^n)) = 2^{c_1 n}$ as in Theorem 7.2,

$$N(B(l_2^n), c' \frac{c_1^2}{(1+\log(2/c_1))} K^\circ) \geq 2^{c_1 n/2},$$

where c' is an absolute constant. Note that the inequalities (7.3) and (7.4) provide better estimates.

Remark 7.5: Theorem 7.2 can be complemented as follows.

Assume that $N(B(l_2^n), K^\circ) = 2^{c_1 n}$.

(a) If $c_1 > 2\log_2 C$ then

$$N(K, B(l_2^n)) \geq 2^{c_1 n/2}. \tag{7.7}$$

(b) If $c_1 < 2\log_2 C$ then

$$N(K, \frac{1}{18M^*(K)}\sqrt{\frac{c_1}{2C'}} B(l_2^n)) \geq 2^{c_1 n/2}, \tag{7.8}$$

where C' is an absolute constant, and $M^*(K) = \int_{S^{n-1}} ||u||_{K^\circ} du$.

Note that, unlike (7.3) and (7.4), the pair (7.7) and (7.8) does not constitute a full duality. The reason is the appearance of the term $M^*(K)$ in (7.8), which does not stay bounded when K varies. The proof uses available methods from the Asymptotic Theory of Normed Spaces, which are unrelated to the present paper, and hence is omitted.

8. A comparison with a result of Diaconis and Freedman

Diaconis and Freedman, [3], compared the limit behavior of the following two distributions. One is the distribution of the first k coordinates of a uniformly distributed random point $(x_1, \ldots, x_n)$ on the $(n-1)$-dimensional sphere of radius $\sqrt{n}$. The other is the distribution of k independent standard Gaussian random variables in R^k. Denote the two distributions by $Q_{n,\sqrt{n},k}$ and P_1^k, respectively.

Diaconis and Freedman found that when $n \to \infty$ and $k = o(n)$, the variational distance between the two distributions tends to zero. A precise estimate for the convergence is

$$\|P_1^k - Q_{n,\sqrt{n},k}\| \le \frac{2(k+3)}{n-k-3},$$

where the variational distance is defined by $\|P - Q\| = 2\sup_A |P(A) - Q(A)|$, where A runs on all measurable subsets of the probability space.

In this section we use our previous results to establish similarity of the two distributions in a different sense, and in the case $k = \lambda n$, when $0 < \lambda < 1$ is a small but fixed number. We examine a different similarity notion since we are interested in the limit behavior of the tails of the distributions. Note that both distributions concentrate strongly, as $n \to \infty$, around their expectations, with tail distributions exponentially small. Therefore, similarity of the tails would not be captured by establishing that the variational distance is small. We compare the first order terms in the exponents describing the tails. Here are the details.

The symmetry of the k coordinates in both distributions reduces the question to a comparison between the following two one-dimensional distributions (as functions of t'):

$$P_1'(t') = \mathrm{Prob}[\sum_{i=1}^k y_i^2 \le t' : y = (y_1, \dots, y_n) \text{ uniform on } \sqrt{n}S^{n-1}],$$

$$P_2'(t') = \mathrm{Prob}[\sum_{i=1}^k g_i^2 \le t' : g_i \text{ i.i.d. gaussians}, 1 \le i \le k].$$

Denote $t = t'/k$, $x = y/\sqrt{n}$ and $P_i(t) = P_i'(tk)$. The displayed distributions get the form

$$\begin{aligned} P_1(t) &= \mathrm{Prob}\left[\sum_{i=1}^k x_i^2 \le t : x = (x_1, \dots, x_n) \text{ uniform on } S^{n-1}\right] \\ &= \mathrm{Prob}[\mathrm{d}^2(x, E_{n-k}) \le \lambda t], \\ P_2(t) &= \mathrm{Prob}\left[\frac{1}{k}\sum_{i=1}^k g_i^2 \le t : g_i \text{ i.i.d. gaussians}, 1 \le i \le k\right]. \end{aligned}$$

We assume now that k/n is small and estimate $P_1(t)$ using (5.3) and (5.4) in section 5. We can also estimate $P_2(t)$ using classical results such as Cramer's theorem (see, e.g., [12]). The results of such estimates yield:

(a) If $t = 1 + \delta$ then

$$P_1(t) = 1 - e^{-\frac{k}{2}(\ln(\frac{1}{1+\delta})+\delta)+o(k)}, \quad P_2(t) = 1 - e^{-\frac{k}{2}(\ln(\frac{1}{1+\delta})+\delta)+o(k)}.$$

(b) If $t = 1 - \delta$ then

$$P_1(t) = e^{-\frac{k}{2}(\ln(\frac{1}{1-\delta})-\delta)+o(k)}, \quad P_2(t) = e^{-\frac{k}{2}(\ln(\frac{1}{1-\delta})-\delta)+o(k)}.$$

We see that for each fixed $\delta > 0$, as $k \to \infty$ the two respective tails differ by at most $o(k)$ multiplying an exponentially (in k) small number.

References

[1] D. Alfers and H. Dinges, *A normal approximation for beta and gamma tail probabilities*, Wahrscheinlichkeitstheorie und Verwandte Gebeite **65** (1984), 399–420.

[2] J. Bourgain, A. Pajor, S. J. Szarek and N. Tomczak-Jaegermann, *On the duality problem for entropy numbers of operators*, Lecture Notes in Mathematics **1376**, Springer-Verlag, Berlin, 1987–88, pp. 50–63.

[3] P. Diaconis and D. Freedman, *A dozen de Finetti-style results in search of a theory*, Annales de l'Institut Henri Poincaré **23** (1987), 397–423.

[4] A. Giannopoulos and V. D. Milman, *Euclidean structures in finite dimensional spaces*, in *Handbook on the Geometry of Banach Spaces* (W. B. Johnson and J. Lindenstrauss, eds.), North-Holland, Amsterdam, 2001, pp. 707–779.

[5] W. B. Johnson and J. Lindenstrauss, *Extensions of Lipschitz mappings into a Hilbert space*, in *Conference in Modern Analysis and Probability (New Haven, Conn., 1982)*, Contemporary Mathematics **26** (1984), 189–206.

[6] H. König and V. D. Milman, *On covering numbers of convex bodies*, Lecture Notes in Mathematics **1267**, Springer-Verlag, Berlin, 1985–86, pp. 85–95.

[7] V. D. Milman and G. Schechtmann, *Asymptotic Theory of Finite Dimensional Normed Spaces*, Lecture Notes in Mathematics **1200**, Springer-Verlag, Berlin, 1986.

[8] V. D. Milman and S. J. Szarek, *A geometric lemma and duality of entropy numbers*, in *Geometric Aspects of Functional Analysis*, Proceedings of the Israel Seminar (GAFA) 1996–2000, Lecture Notes in Mathematics **1745**, Springer-Verlag, Berlin, 2000, pp. 191–222.

[9] V. D. Milman and A. Pajor, *Entropy and asymptotic geometry of non-symmetric convex bodies*, Advances in Mathematics **152** (2000), 314–335.

[10] A. Pajor and N. Tomczak-Jaegermann, *Subspaces of small codimension of finite dimensional Banach spaces*, Proceedings of the American Mathematical Society **97** (1986), 637–642.

[11] G. Pisier, *A new approach to several results of V. Milman*, Journal für die reine und angewandte Mathematik **393** (1989), 115–131.

[12] S. R. S. Varadhan, *Large Deviations and Applications*, SIAM Publications, Philadelphia, 1984.