# Bounds on Packings of Spheres in the Grassmann Manifold

Alexander Barg, *Senior Member, IEEE,* and Dmitry Yu. Nogin

***Abstract*—We derive the Gilbert–Varshamov and Hamming bounds for packings of spheres (codes) in the Grassmann manifolds over $\mathbb{R}$ and $\mathbb{C}$. Asymptotic expressions are obtained for the geodesic metric and projection Frobenius (chordal) metric on the manifold.**

***Index Terms*—Invariant measure, minimum distance, principal angles, volume bounds.**

## I. INTRODUCTION

### A. Distances in $G_{k,n}$

LET $X$ be a Riemannian homogeneous space with metric $\rho$ and (normalized) invariant measure $dh$. A *code* $C$ is a finite subset of $X$. Let $\delta = \delta(C)$ be the minimum distance between distinct points in $C$. One of the main problems of coding theory is establishing the maximum size of a code with a given distance $\delta$. The best known examples are the sphere $S^{n-1}(\mathbb{R})$ and, in the discrete case, the Hamming space $H_q^n$. One of the possible generalizations of the former is studying codes in the Grassmann manifold $G_{k,n}(L)$, where $L = \mathbb{R}$ or $\mathbb{C}$. It is a homogeneous space of the group $O(n)$ or $U(n)$, respectively. For instance

$$G_{k,n}(\mathbb{R}) \cong O(n)/O(k) \times O(n-k).$$

Recently, this space has been the focus of attention for a number of reasons. From a purely geometric point of view, packings in $G_{k,n}$ form a natural generalization of spherical codes. However, their study seems to have been first addressed only a few years ago [3], [13], motivated in part by a group-theoretic application that connects them with the theory of quantum codes. Apart from this, codes in $G_{k,n}(\mathbb{C})$ arise naturally in the area of multiple-antenna transmission (especially, in the case of low noise), and a few papers that underline this connection [1], [15] have been published. Therefore, it seems timely to address basic coding-theoretic questions such as the sphere-packing bounds on the size of codes. This is the aim of the present paper.

Manuscript received May 30, 2000; revised November 5, 2001. This work was supported in part by the Russian Foundation for Basic Research under Projects 99-01-01204 and 02-01-01041, and INTAS under Grant 00-738. This work was performed in part while D. Yu. Nogin was visiting DIMACS, Bell Labs, November 1999.

A. Barg is with Bell Labs, Lucent Technologies, Murray Hill, NJ 07974 USA and with IPPI RAN, Moscow, Russia (e-mail abarg@research.bell- labs.com).

D. Yu. Nogin is with IPPI RAN, Moscow 101447 Russia (e-mail nogin@iitp.ru).

Communicated by P. Solé, Associate Editor for Coding Theory.

Publisher Item Identifier 10.1109/TIT.2002.801469.

Let $B(\delta)$ be the metric ball of radius $\delta$ in $X$. Obviously, if $M$ is any number such that

$$Mh(B(\delta)) < 1 \tag{1}$$

there exists a code in $X$ of size $M+1$ and distance $\delta$. Indeed, as long as (1) holds true, it is always possible to pick a point $x \in X$ so that the centers of already chosen balls together with $x$ form a code with distance at least $\delta$. This principle is called the Gilbert–Varshamov lower bound. On the other hand, obviously for any code $C$

$$|C|h(B(\delta/2)) \le 1.$$

This is the Hamming upper bound on the size of codes. We note that if the metric on $X$ is not "strictly intrinsic," i.e., the triangle inequality is never satisfied with equality (for pairwise distinct points), the Hamming bound can be improved.

To define the distance in $G_{k,n}$, we have to introduce principal angles between planes $p$ and $q$. Let $a \in p$ and $b \in q$ be two unit vectors and $\theta = \arccos|\langle a, b\rangle|$ the angle between them. As $a$ varies over $p$ and $b$ varies over $q$, $\theta$ has $k$ stationary points

$$0 \le \theta_k \le \cdots \le \theta_1 \le \pi/2$$

corresponding to some pairs of vectors $(a_i, b_i)$, $1 \le i \le k$. The sets of vectors $(a_i)$ and $(b_i)$ form orthogonal bases in their respective planes, and if $k \le n/2$, then $a_i$ is orthogonal to $b_j$ for any $i \ne j$. For a proof see, for instance, [10]. Note also that if $A_p$ is a generator matrix of a plane $p$, i.e., the matrix whose rows form an orthonormal basis of $p$, and $A_q$ is the same for $q$, then for the eigenvalues $\lambda_1, \ldots, \lambda_k$ of $A_q A_p^* A_p A_q^*$ we have $\lambda_i = \cos^2\theta_i$ (here $*$ denotes the Hermitian conjugate). In other words, the singular value decomposition of the matrix $A_p A_q^*$ has the form $UCV^*$, where

$$C = \operatorname{diag}(\cos\theta_1, \cos\theta_2, \ldots, \cos\theta_k).$$

This also means that we can change the basis and apply a suitable unitary rotation so that $p$ is generated by

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ & & \cdots & & \cdots & & \cdots & \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{2}$$

and $q$ by (3), as shown at the bottom of the following page.

Let

$$\theta = (\theta_1, \theta_2, \ldots, \theta_k)$$
$$\sin\theta = (\sin\theta_1, \sin\theta_2, \ldots, \sin\theta_k).$$

0018-9448/02$17.00 © 2002 IEEE

Throughout the text, $||x||_2$ denotes the $\ell_2$-norm and $||x||_\infty$ the $\ell_\infty$-norm of a vector $x$. For a matrix $M$

$$||M||_2 = \sup_{x\neq 0} ||Mx||_2/||x||_2$$

is its operator 2-norm and

$$||M||_F = \sqrt{\sum_{i,j} |m_{i,j}|^2}$$

its Frobenius, or Euclidean norm.

Let $p, q \in G_{k,n}(L)$ be two planes. There are several possibilities for the distance between $p$ and $q$, see [4]:

*geodesic distance*

$$g(p, q) = ||\theta||_2;$$

*chordal distance*

$$\rho(p, q) = 2^{-1/2}||A_p^* A_p - A_q^* A_q||_F = ||\sin\theta||_2;$$

*Fubini–Study distance*

$$d_{\mathrm{FS}}(p, q) = \arccos|\det A_p A_q^*| = \arccos\left(\prod_i \cos\theta_i\right);$$

*chordal 2-norm*

$$d_{c2}(p, q) = ||A_p^* U - A_q^* V||_2 = \left\|2\sin\frac{1}{2}\theta\right\|_\infty;$$

*chordal Frobenius norm*

$$d_{cF}(p, q) = ||A_p^* U - A_q^* V\Big\|_F = ||2\sin\frac{1}{2}\theta\Big\|_2;$$

*projection 2-norm*

$$d_{p2}(p, q) = ||A_p^* A_p - A_q^* A_q||_2 = ||\sin\theta||_\infty.$$

Note that the metric $\rho(p, q)$ is sometimes called the projection Frobenius distance. The term "chordal distance" was introduced in [3] for the reasons discussed in the next paragraph. The projection 2-distance is the same as the chordal distance, except that the Frobenius norm is replaced by the 2-norm. The geodesic distance is the arc length in a natural geometry of the Grassmann manifold viewed as a quotient space of the orthogonal group. The Fubini–Study distance is derived via the Plücker embedding of $G_{k,n}$ in the projective space $\mathbb{P}(\bigwedge^k(\mathbb{R}^n))$. The chordal 2-norm and Frobenius-norm distances are derived by embedding the Grassmann manifold in the vector space $\mathbb{R}^{kn}$, then using the operator 2-norm and Frobenius norm, respectively.

Advantages of the chordal distance $\rho$ are discussed in [3]. One of them is that under this definition of the norm, $G_{k,n}(\mathbb{R})$ affords an isometric embedding in a sphere $S_k$ of radius $r = \sqrt{k(n-k)/n}$ in $\mathbb{R}^{(n-1)(n+2)/2}$. It is realized as follows. The matrix $\Pi_p = A_p^* A_p$ is an orthogonal projection from $\mathbb{R}^n$ on $p$. As shown in [3], for any $p$ the Euclidean norm of the matrix $\Pi_p - (k/n)I_n$ equals $r$, so $p$ is represented as a point on the sphere $S_k$. For different $k$, these spheres lie on a large sphere in $\mathbb{R}^{n(n+1)/2}$ of radius $\frac{1}{2}\sqrt{n}$. The main result of [3] is that this embedding of $G_{k,n}(\mathbb{R})$ in $S_k$ is isometric in the sense that $\rho^2(p, q) = (1/2)||\Pi_p - \Pi_q||_F^2$, i.e., the distance $\rho$ is proportional to the length of the chord that joins the projection matrices. Observe that other embeddings, such as the Plücker embedding, usually map $G_{k,n}$ into a space of a much higher dimension.

Likewise, in the complex case the Hermitian matrix $\Pi_p = A_p^* A_p$ is an orthogonal projector on $p$. We again have $\rho^2(p, q) = \frac{1}{2}||\Pi_p - \Pi_q||_F^2$, $\mathrm{tr}\Pi_p = k$, and $||\Pi_p - \frac{k}{n} I_k||_F^2 = k(n-k)/n$. A Hermitian matrix with fixed trace can be represented by a point on the sphere in $\mathbb{R}^{n^2-1}$.

The existence of these embeddings implies that upper bounds on the size of codes on the sphere apply to codes in $G_{k,n}(L)$ with distance function $\rho$. In particular, the well-known Rankin bounds in the real case imply the following [3]:

$$\delta(C) \le \begin{cases} \frac{k(n-k)}{n}\,\frac{|C|}{|C|-1}, & \text{if } |C| \le n(n+1)/2, \\ \frac{k(n-k)}{n}, & \text{if } |C| > n(n+1)/2. \end{cases}$$

This inequality is an analog of the Plotkin bound of coding theory. Interestingly, there exist sequences of codes in $G_{k,n}(\mathbb{R})$ that meet these bounds [13].

We are interested in codes in $G_{k,n}$ whose size grows exponentially with $n$. Our goal is to derive an expression for the Gilbert–Varshamov and Hamming bounds on codes. The answer can be written in a compact form only in the asymptotic setting. We assume that $n \to \infty$ and $k$ is a fixed constant. Further, the quantity $R = (1/n)\ln|C|$ is called the rate of the code. Note that the case $G_{1,n}(\mathbb{R})$ with the metric $g(p, q) = \theta$ was treated by Shannon [12]. He proved that there exist sequences of codes with distance $\theta$ and

$$R \ge -\ln\sin\theta - o(1).$$

Note that in the case $k = 1$ codes optimal for the metric $g$ will also be optimal for the metric $\rho$.

### *B. Sphere Packing in $G_{k,n}$.*

Our main result, proved in Section III, is as follows.

*Theorem 1:* Let $B(\delta)$ be the ball of radius $\delta$ in $G_{k,n}(L)$. Then

i) for the chordal distance

$$h(B(\delta)) = \left(\frac{\delta}{\sqrt{k}}\right)^{\beta nk + o(n)}; \tag{4}$$

---

$$\begin{bmatrix} \cos\theta_1 & 0 & \cdots & 0 & \sin\theta_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \cos\theta_2 & \cdots & 0 & 0 & \sin\theta_2 & \cdots & 0 & 0 & \cdots & 0 \\ & & \cdots & & & & \cdots & & & \cdots & \\ 0 & 0 & \cdots & \cos\theta_k & 0 & 0 & \cdots & \sin\theta_k & 0 & \cdots & 0 \end{bmatrix}. \tag{3}$$

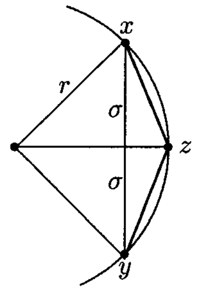


Fig. 1. Chordal distance and the Hamming bound.

ii) for the geodesic distance

$$h(B(\delta)) = \left(\sin\frac{\delta}{\sqrt{k}}\right)^{\beta nk+o(n)} \tag{5}$$

where $n\to\infty$, $k = \text{const}$, and $\beta = 1, 2$ for $L = \mathbb{R}, \mathbb{C}$, respectively.

With the above, this implies the following theorem.

*Theorem 2:* Let $n\to\infty$, $k=\text{const}$. There exist sequences of codes in $G_{k,n}$ with distance $\delta$ and asymptotic rate

$$R \gtrsim -\beta k \ln(\delta/\sqrt{k}) \quad \text{(chordal distance)} \tag{6}$$
$$R \gtrsim -\beta k \ln(\sin(\delta/\sqrt{k})) \quad \text{(geodesic distance).} \tag{7}$$

For any sequence of codes with distance $\delta$

$$R \lesssim -\beta k \ln\left(\sqrt{1-\sqrt{1-\frac{\delta^2}{2k}}}\right) \quad \text{(chordal distance)} \tag{8}$$
$$R \lesssim -\beta k \ln(\sin(\delta/2\sqrt{k})) \quad \text{(geodesic distance).} \tag{9}$$

*Proof:* Only the Hamming bound (8) for the chordal distance $\rho$ is not obvious. To prove it, observe that if $p$ and $q$ are two planes in $G_{k,n}$ with distance $\rho(p, q) = \delta = \sqrt{2}\sigma$, then their images $x$ and $y$ on the sphere $S_k$ are at distance $\|x-y\|_2 = 2\sigma$. Let $z$ be the "midpoint" between $x$ and $y$ (i.e., the point on $S_k$ that halves the arc). The distance between $x$ and $z$ is then easily computed: $\|x-z\|_2 = \sqrt{2}\varrho$, where

$$\varrho = r\sqrt{1-\sqrt{1-\sigma^2/r^2}}$$

and $r = \sqrt{k(n-k)/n}$ is the radius of the sphere $S_k$ (see Fig. 1). If the inverse image of $z$ is a plane in $G_{k,n}$ and the code $C$ has distance $\delta$, the spheres of radius $\varrho$ about $p$ and $q$ do not have common interior points. Thus, $|C| \le (h(B(\varrho)))^{-1}$. Since $r \sim \sqrt{k}$, we obtain (8) from (4).

### C. Distance Distribution

Apart from the minimum distance, an important parameter of codes is their distance distribution, i.e., the average number of neighbors of a code point at a given distance. For instance, consider codes in the chordal metric. Since $t$ in principle can be any number between $\delta$ and $\sqrt{k}$, it is convenient to consider the distance density of $C$ defined as follows:

$$a(t) = \frac{1}{|C|}\left|\{(p, q)\colon t-(1/n) \le \rho(p, q) \le t\}\right|.$$

Suppose $C$ is chosen in $X$ with uniform probability distribution. Then

$$\frac{\mathsf{E}a(t)}{|C|} = h(B(t)) - h(B(t-n^{-1}))$$

where $\mathsf{E}$ denotes the mathematical expectation. From (4), the right-hand side equals $h(B(t))(1-o(1))$. By the Markov inequality we conclude that among sequences of codes that meet the bound (6), i.e., for which the distance $\delta(C) \gtrsim \sqrt{k}e^{-R/\beta k}$, there exist codes whose distance density is bounded above as

$$a(t) \le p(n)|C|h(B(t))(1-o(1)) \qquad (e^{-R/\beta k} \le t/\sqrt{k} \le 1)$$

where $p(n)$ is some function of polynomial growth. In other words, the logarithm of the average number of neighbors for these codes is bounded above as

$$\ln a(t) \le n(R + \beta k \ln(t/\sqrt{k}) - o(1)).$$

Codes with similar properties in other spaces of interest to coding and information theory (the binary Hamming space $H_q^n$ and the sphere $S^{n-1}(\mathbb{R})$) have a number of interesting properties. The most important of them is related to the use of codes for transmission of information over noisy channels. In this situation, random codes account for the best known exponential upper bounds on the probability of incorrect recovery of the code vector transmitted from the noisy version of this vector received from the channel [12].

## II. Invariant Densities in $G_{k,n}$

To prove (4), we need explicit volume forms on $G_{k,n}(L)$. A general construction of invariant measures in homogeneous spaces with applications to classical groups and related manifolds is given, for instance, in [11]. A combinatorial approach is presented in [9]. Necessary background material can be looked up in any textbook on geometry, for instance, [14], [16].

We note that the metric plays no role in the construction of the measure which is unique (up to a constant factor). Let us begin with the real case. Density for the submanifold of critical angles was calculated several times in statistics (see [8]). Let $a_1, \dots, a_k$ and $b_1, \dots, b_{n-k}$ be orthonormal column vectors that span a plane $p$ and its orthogonal complement. The invariant measure on $G_{k,n}(\mathbb{R})$ is (locally) given by the $k(n-k)$ form

$$v_k^n = \bigwedge_{j=1}^{n-k} \bigwedge_{i=1}^{k} b_j^t da_i$$

where $^t$ means transposition.

To isolate the part of this form that corresponds to the density on principal angles, we also introduce the Stiefel manifold $V_{k,n}$, i.e., the manifold of orthonormal $k$-frames in $\mathbb{R}^n$. In particular, $V_{k,k}$ is the orthogonal group $O(k)$. It is proved in [8] that the open part of $G_{k,n}$ decomposes into a direct product of the simplex

$$\Theta = \{(\theta_1, \dots, \theta_k)\colon \pi/2 > \theta_1 > \cdots > \theta_k > 0\}$$

and two manifolds, $\tilde{V}_{k,k}$ and $V_{k,n-k}$, where $\tilde{V}$ is the submanifold of the Stiefel manifold specified by those frames in which in each vector the first coordinate is positive. Based on this, it is possible to write $v_k^n$ as a product of three independent densities and compute the marginal distribution on $\Theta$ [8]. This gives the answer in the real case (see $\omega_k$ below). In the complex case, it is easier, though not so intuitive, to rely upon the distribution of eigenvalues of random Gaussian unitary matrices [6]. We can assume that $q$ is a fixed plane with generator matrix $J = [I_k\ 0]$ and $p$ is uniformly distributed on $G_{k,n}(L)$. Then we are interested in the distribution of the eigenvalues $\lambda_i = \cos^2\theta_i$ of the matrix $JA_p^*A_pJ^*$. It can be shown that this distribution is related to the distribution of eigenvalues of Wishart matrices, i.e., matrices of the form $G^*G$ where $G$ is a $k\times n$ matrix with Gaussian $\mathcal{N}(0,1)$ elements [6, p. 202]. The final answer has the form

$$\omega_k = K(k,n)\prod_{i=1}^{k}(\sin\theta_i)^{\beta(n-2k)} \cdot \prod_{1\le i<j\le k}(\sin^2\theta_i - \sin^2\theta_j)^{\beta} d\theta_1\wedge\cdots\wedge d\theta_k$$

where the constant $K(k,n)$ is chosen to normalize the measure of $\Theta$. In the real case its value is obvious from the preceding geometric considerations

$$K(k,n) = \frac{\mathrm{vol}\left(\tilde{V}_{k,k}\right)\mathrm{vol}(V_{k,n-k})}{\mathrm{vol}(G_{k,n})}.$$

Volumes of the manifolds involved are well known [11], and we get

$$K(k,n) = \prod_{i=1}^{k}\frac{O(k-i+1)^2 O(n-k-i+1)}{2O(n-i+1)}$$

where $O(k) = 2\pi^{k/2}/\Gamma(k/2)$ is the area of the unit sphere in $\mathbb{R}^k$. Note that $K(k,n)$ grows polynomially in $n$ for fixed $k$, so in our context its exact form is not essential. This also holds true for $\mathbb{C}$, namely, we have [7]

$$K(k,n) = \prod_{i=1}^{k}\frac{(n-i)!}{(i!)^2(n-k-i)!}.$$

Hence the volume of the ball of radius $\delta$ is given by $h(B(\delta)) = K(k,n)J_k(\delta)$, where

$$J_k(\delta) = \int \frac{\omega_k}{K(k,n)} \tag{10}$$

where the integration is carried over the region inside $\Theta$ given by

$$\|\sin\theta\|_2 \le \delta \qquad \text{(chordal distance)}$$
$$\|\theta\|_2 \le \delta \qquad \text{(geodesic distance).}$$

## III. Asymptotics: Proof of Theorem 1

We would like to compute the logarithmic asymptotics of $h(B(\delta))$. Both cases considered turn out to be quite similar, so let us compute the behavior of $J_k$ for the $\rho$-metric. We have

$$J_k = \int\cdots\int_{\substack{0<x_k<\cdots<x_1<1\\ \|x\|_2\le\delta}} \frac{(x_1x_2\cdots x_k)^{\beta(n-2k)}\prod_{i<j}(x_i^2-x_j^2)^{\beta}}{\sqrt{(1-x_1^2)(1-x_2^2)\cdots(1-x_k^2)}} \cdot dx_1\cdots dx_k. \tag{11}$$

Note that for $\beta = 2$ (but not for $\beta = 1$), $J_k$ is symmetric in $x_1,\ldots,x_k$, so we can divide out $k!$ and remove the ordering condition. To treat both cases simultaneously, we compute the asymptotics of the integral by the Laplace method [2], [5]. In the multidimensional case, the corresponding theorem has the following form.

*Theorem 3 [5, p. 131]:* Let $\Omega$ be a connected open domain in $\mathbb{R}^k$, $\partial\Omega$ be its boundary, $[\Omega]=\Omega\cup\partial\Omega$, and let $f(x), S(x)\in C([\Omega])$ be real functions. Further, suppose that the maximum $\max_{x\in[\Omega]} S(x)$ is attained only at the point $x^0\in\partial\Omega$, and at this point

a) $\partial S(x^0)/\partial\boldsymbol{n}\ne 0$, where $\partial/\partial\boldsymbol{n}$ is differentiation along the interior normal to $\partial\Omega$ at $x^0$;

b) the matrix

$$H = \left(\frac{\partial^2 S(x^0)}{\partial\xi_i\partial\xi_j}\right)_{i,j=1}^{k-=1}$$

is negative definite, where $(\xi_1,\ldots,\xi_{k-1})$ is an orthonormal basis in the tangent space $T_{x^0}\partial\Omega$ to $\partial\Omega$ at $x^0$;

c) $f, S, \partial\Omega\in C^\infty$ in the neighborhood of $x^0$.

Then

$$\int_\Omega f(x)e^{\lambda S(x)}\,dx = \lambda^{-\frac{k+1}{2}}e^{\lambda S(x^0)}\sum_{m=0}^{\infty}a_m\lambda^{-m} \qquad (\lambda\to\infty)$$

for some constants $a_m$. Moreover, $a_0$ is proportional to $f(x^0)$.

Let us use this result in our problem. We have $x\in\mathbb{R}^k$, $\lambda = n$

$$S(x) = \beta\ln(x_1\cdots x_k) = \beta\sum_i \ln x_i$$

$$f(x) = \left(\prod_{i=1}^{k}x_i^{-2\beta k}(1-x_i^2)^{-1/2}\right)\prod_{i<j}(x_i^2-x_j^2)$$

$$[\Omega] = [(x_1,\ldots,x_k)\colon 0\le x_k\le\cdots\le x_1\le 1,\ \|x\|_2\le\delta].$$

The maximum of $S(x)$ over $[\Omega]$ is attained at

$$x^0 = k^{-1/2}(\delta,\delta,\ldots,\delta).$$

This is because $S$ is a convex function and $\Omega$ is a convex domain, so we can use Lagrange multipliers to compute the maximum. To satisfy the conditions of the theorem, we have to adjust the integration domain in several ways.

i) Observe that $S(x)$ has discontinuities at the hyperplanes $x_i = 0$ and $f(x)$ has discontinuities at the hyperplanes $x_i = 1$. Therefore, let us shift the domain $[\Omega]$ from these hyperplanes.

ii) At the point $x^0$, the boundary $\partial\Omega$ is not differentiable. Therefore, let us extend the domain by including small sectors.

For instance, for $k = 2$, the domain $[\Omega]$ is formed by the intersection of the sector of radius $\delta$ and angle $\pi/4$ in the first quadrant and the strip $0\le x_1\le 1$. The discussed extension amounts to increasing the angle to $\pi/4+\varphi$ for some small finite $\varphi$. Note that the maximum of $S(x)$ over the extended domain does not shift from $x^0$. In the general case, consider the adjusted domain

$$\Omega_\varepsilon = [(x_1,\ldots,x_k)\colon \varepsilon\le x_i\le 1-\varepsilon,\ x_{i+1}\le x_i(1+\varepsilon),\ \|x\|_2\le\delta].$$

Note that to $\Omega_\varepsilon$ corresponds an adjusted domain $\Theta_\varepsilon$ in the coordinates $\theta_1, \ldots, \theta_k$, where

$$\mathrm{vol}(\Theta \setminus \Theta_\varepsilon) + \mathrm{vol}(\Theta_\varepsilon \setminus \Theta) = O(\varepsilon).$$

Since the integrand in (10) is bounded by $1$, this implies that $J_k$ differs from the corresponding integral over $\Omega_\varepsilon$ by $O(\varepsilon)$ uniformly in $n$.

iii) Finally, the case $\delta = \sqrt{k}$ also has to be excluded since then $f(x)$ has a singularity at $x^0$. In this case, the ball exhausts the entire space $X$ except for a set of measure $0$, and the bounds are extended by continuity.

We need to verify conditions a) and b) of Theorem 3. Note that in the basis $(x_1, \ldots, x_k)$ the Hessian $H(x) = (\frac{\partial^2 S(x)}{\partial x_i \partial x_j})$ has the form $-\mathrm{diag}(x_1^{-2}, \ldots, x_k^{-2})$, and so

$$H(x^0) = -\frac{\delta^2}{k} I_k.$$

This quadratic form is negative definite. Clearly, $H$ remains negative definite under the restriction to the subspace spanned by $(\xi_1 \ldots, \xi_{k-1})$. To verify a), note that the interior normal to the sphere $\partial\Omega$ has the form $\boldsymbol{n} = (-1, \ldots, -1)$. The level surfaces of $S(x)$ are given by $x_1 \cdots x_k = \mathrm{const}$, which are convex hypersurfaces. Moreover, they are homothetic to each other with respect to the origin; therefore, $S(x)$ strictly increases along the diagonal $x_1 = x_2 = \cdots = x_k$.

So, let us apply the theorem to $J_k$. We have $f(x^0) = 0$; thus,

$$J_k = n^{-\frac{k+1}{2}} \left(\frac{\delta}{\sqrt{k}}\right)^{\beta k n} O(n^{-1}).$$

This concludes the proof of (4).

To prove (5), we compute the integral in (11) over the region

$$\Omega' = [0 < x_k < \cdots < x_1 < 1, \;\; \|\arcsin x\|_2 \le \delta].$$

Since $\Omega$ is concave, so is $\Omega'$. Therefore, the maximum of the function $S(x)$ over $[\Omega']$ is attained for

$$x^0 = k^{-1/2}(\sin\delta, \ldots, \sin\delta)$$

otherwise, the argument is the same. This gives (5).

## References


[1] D. Agrawal, T. Richardson, and R. Urbanke, "Multiple-antenna signal constellations for fading channels," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2618–2626, Sept. 2001.

[2] N. G. de Bruijn, *Asymptotic Methods in Analysis*. Amsterdam, The Netherlands: North-Holland, 1958.

[3] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian spaces," *Exper. Math.*, vol. 5, no. 2, pp. 139–159, 1996.

[4] A. Edelman, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthonormality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 20, no. 2, pp. 303–353, 1998.

[5] M. V. Fedoryuk, *Asimptotika: Integraly i Ryady [Asymptotics: Integrals and Series]* (in Russian). Moscow, U.S.S.R.: Nauka, 1987.

[6] V. L. Girko, *An Introductiont to Statistical Analysis of Random Arrays*. Utrecht, The Netherlands: VSP Press, 1998.

[7] B. Hassibi, B. Hochwald, and T. Marzetta, "Space-time autocoding," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2761–2781, Nov. 2001.

[8] A. T. James, "Normal multivariate analysis and the orthogonal group," *Ann. Math. Statist.*, vol. 25, pp. 40–75, 1954.

[9] D. A. Klain and G.-C. Rota, *Introduction to Geometric Probability*. Cambridge, U.K.: Cambridge Univ. Press, 1997.

[10] S. N. Roy, "A note on critical angles between two flats in hyperspace with certain statistical applications," *Sankhyā*, vol. 8, pp. 177–194, 1947.

[11] L. A. Santaló, *Integral Geometry and Geometric Probability*. Reading, MA: Addison-Wesley, 1976.

[12] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell Syst. Tech. J.*, vol. 38, pp. 611–656, 1959.

[13] P. W. Shor and N. J. A. Sloane, "A family of optimal packings in Grassmannian manifolds," *J. Algebraic Combin.*, vol. 7, pp. 157–163, 1998.

[14] M. Spivak, *Calculus on Manifolds*. New York: Benjamin, 1965.

[15] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inform. Theory*, vol. 48, pp. 359–383, Feb. 2002.

[16] F. W. Warner, *Foundations of Differentiable Manifolds and Lie Groups*. New York: Springer-Verlag, 1983.