Bounds for Codes in Products of Spaces, Grassmann and Stiefel Manifolds

Christine Bachoc, Yael Ben-Haim, and Simon Litsyn, *Senior Member, IEEE*

Abstract

Upper bounds are derived for codes in Stiefel and Grassmann manifolds with given minimum chordal distance. They stem from upper bounds for codes in the product of unit spheres and projective spaces. The new bounds are asymptotically better than the previously known ones.

I. Introduction

Use of multiple transmit and receive antennas essentially increases the spectral efficiency of wireless systems (see [1] and references therein). Analysis of Rayleigh flat-fading multiple-input multiple-output (MIMO) scenarios with m transmit antennas and n transmitted symbols, reveals that relevant coding schemes can be designed as collections of elements (points) in the complex Grassmann manifold - the set of m-dimensional linear subspaces in $\mathbb{C}^n$, if the channel is unknown to the receiver, and in the complex Stiefel manifold - the set of m orthonormal vectors in $\mathbb{C}^n$, if the channel is known to the receiver. An appropriately defined distance measure between the points characterizes diversity of the designed scheme. Following standard coding theory considerations, we study the relation between the number of points (the size of a code) and the minimum distance between distinct code points. Our aim in this paper is to obtain new upper bounds for the size of codes in Grassmann and Stiefel manifolds.

The most powerful technique for this kind of problems is the linear programming method (called also the polynomial method), initiated by Delsarte [2]. The method is very well understood in the case of 2-point homogeneous spaces (defined in the next section), where very explicit bounds, and also good asymptotic bounds on the rate of codes have been derived. Examples are the Hamming and Johnson schemes, treated in [3], the unit sphere of $\mathbb{R}^n$ [4], and the projective spaces [4], [5].

When the underlying space is homogeneous and symmetric but not 2-point homogeneous, the situation is much more complicated, although the principles of the linear programming method remain valid. The difficulties come from the fact that the zonal functions defined for these spaces are not functions of one variable, but rather of several variables. The Grassmann spaces considered in this paper fall into this category. An attempt to overcome

C. Bachoc is with Laboratoire A2X, Université Bordeaux I, 351, cours de la Libération, 33405 Talence France. (email: `bachoc@math.u-bordeaux1.fr`)

Y. Ben-Haim and S. Litsyn are with the School of Electrical Engineering, Tel-Aviv University, Tel-Aviv 69978, Israel (email: `yael@eng.tau.ac.il; litsyn@eng.tau.ac.il`).

this problem was carried out in [6]. An asymptotic bound for the rate of Grassmannian codes was obtained, involving the asymptotics of the largest eigenvalue of some symmetric endomorphism. This bound however is not optimal since it was improved for $m > 1$ by some volume-type arguments for a large range of values of the minimum distance [7].

There is one trivial case of symmetric spaces of rank $m > 1$ for which the classical treatment of the linear programming method is easily extended: it is the direct product of 2-point homogeneous spaces, such as the direct product of m copies of the unit sphere. An example of a similarly easy case is provided by the non-binary Johnson space [8], [9], [10], that is the product of the Hamming and the binary Johnson scheme.

The approach developed in this paper is to relate Grassmann and Stiefel spaces and their associated codes to various products of 2-point homogeneous spaces, and hence to derive upper bounds for these codes in a rather easy way. The asymptotic versions of the new bounds (Theorems 2.7 and 2.8) provide the best currently known asymptotic bounds.

The paper is organized as follows. Definitions and results are given in Section II. Section III describes various relations between the spaces and their codes. The simplest one connects Grassmann and Stiefel spaces to the unit sphere of an asymptotically equal dimension; this yields, for example, to a bound for the asymptotic rate of Grassmannian codes that already improves upon the previous ones. Section IV develops the Delsarte polynomial method for the products of spaces under consideration, including the classical method that involves the Christoffel-Darboux formula, and derives upper bounds for the size of the associated codes. A bound for the asymptotic rate of these codes is obtained. Section V discusses the consequences for the Grassmannian and Stiefel codes. In particular, we show that the bound obtained on the asymptotic rate of Grassmannian codes from the product of projective spaces is sometimes better than the one obtained in Section III. We conclude in Section VI.

II. Definitions and results

We shall use the following notations and definitions. We say that $f(n) \lesssim g(n),\ f(n) \simeq g(n),\ f(n) \gtrsim g(n)$ if $\lim_{n\to\infty} \frac{f(n)}{g(n)} \leq 1,\ \lim_{n\to\infty} \frac{f(n)}{g(n)} = 1,\ \lim_{n\to\infty} \frac{f(n)}{g(n)} \geq 1$, respectively. A code in a metric space (X, d) is a finite set contained in the space, and a codeword is an element of the code. The size of a code C is its cardinality, denoted $|C|$. The rate of a code is $R(C) := \frac{1}{n} \ln |C|$, where $\ln$ denotes the natural logarithm, and n will be defined separately for each space. The minimum distance of a code is the minimum distance (induced by the relevant metric) between a pair of distinct codewords. A metric space (X, d) is called 2-point homogeneous, if X affords the transitive action of a group G, such that the orbits of the action of G on $X \times X$ are characterized by the distance d. In other words, for all $(x, y) \in X$ and $g \in G$, $d(g(x), g(y)) = d(x, y)$, and moreover, for all pairs (x, y), $(x', y') \in G$, there exists $g \in G$ such that $g(x) = x'$ and $g(y) = y'$ if and only if $d(x, y) = d(x', y')$. It is a well-known fact that the compact Riemannian manifolds that are two-point homogeneous are exactly: the unit sphere S^{n-1}, the projective spaces $\mathbb{P}^{n-1}(K)$ where $K = \mathbb{R}, \mathbb{C}, \mathbb{H}$ and the projective plane over the octonions $\mathbb{P}^2(\mathbb{O})$ (see [11], and [12] for more about the octonions and $\mathbb{P}^2(\mathbb{O})$).

A. The real compact two-point homogeneous spaces

The unit sphere of the Euclidean space $\mathbb{R}^n$ is denoted S^{n-1}, namely,

$$S^{n-1} := \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n \mid \sum_{i=1}^{n} x_i^2 = 1 \right\}. \tag{1}$$

The standard scalar product in $\mathbb{R}^n$, given by $(u \cdot v) = \sum_{i=1}^{n} u_i v_i$, defines the Euclidean distance between two points of S^{n-1}:

$$\|u - v\| = \sqrt{\sum_{i=1}^{n} (u_i - v_i)^2} = \sqrt{2}\sqrt{1 - (u \cdot v)}. \tag{2}$$

The angular distance between u and v is defined by the angle $\theta = \theta(u, v) \in [0, \pi]$ such that $\cos\theta = (u \cdot v)$. We have of course

$$\|u - v\| = \sqrt{2}\sqrt{1 - \cos\theta}. \tag{3}$$

The best known asymptotic bound on the rate of spherical codes as a function of the minimum distance is given in the following theorem. It will be extensively used in the derivation of the new results.

Theorem 2.1 ([4]): Let C be a spherical code in S^{n-1} with minimum angular distance $0 \leq \theta \leq \pi/2$ and rate $R(C) = \frac{1}{n} \ln |C|$. Then, when $n \to \infty$,

$$R(C) \lesssim R_S(\theta) := \min\{R_{LP}(\theta), R_Y(\theta)\}, \tag{4}$$

where

$$R_{LP}(\theta) := (1 + \rho)\ln(1 + \rho) - \rho \ln \rho, \tag{5}$$

$$\rho = \frac{1 - \sin\theta}{2 \sin\theta}, \tag{6}$$

and

$$R_Y(\theta) := -\ln \sqrt{1 - \cos\theta} - 0.0686 \tag{7}$$

(the numerical value 0.0686 is approximated). □

Remark 2.2:
- The asymptotic rate of spherical codes with minimum angular distance at least $\pi/2$ is known to be equal to zero. This is a consequence of the Rankin bound ([18], see also [13] or [14]).
- For small values of θ (approximately $\theta < 63°$), we have $R_S(\theta) = R_Y(\theta)$. Otherwise, $R_S(\theta) = R_{LP}(\theta)$.

□

The other real compact manifolds which are two-point homogeneous can be treated in a similar way. These are the projective spaces $\mathbb{P}^{n-1}(K)$ where $K = \mathbb{R}, \mathbb{C}, \mathbb{H}$ (the field of real quaternions) and $n \geq 3$, and the projective plane over the octonions $\mathbb{P}^2(\mathbb{O})$. In order to treat the fields of coefficients in a uniform way, we extend the definition of $(x \cdot y)$ so that, for all $x, y \in K^n$, $(x \cdot y) = \sum_{i=1}^{n} x_i \overline{y_i}$, where the conjugation $x \to \overline{x}$ is the standard one over $K = \mathbb{C}, \mathbb{H}, \mathbb{O}$ and is the identity over $\mathbb{R}$. Also we conventionally assume that $n = 3$ when $K = \mathbb{O}$. The group

G under which these spaces are two-point homogeneous is respectively the orthogonal group $O(\mathbb{R}^n)$, the unitary groups $U(K^n)$ with $K = \mathbb{C}, \mathbb{H}$, and the Lie group F_4 (see [12] for this last case).

The angular distance between p and q in $\mathbb{P}^{n-1}(K)$ is defined by the angle $\theta = \theta(p,q) \in [0, \pi/2]$ such that $\cos\theta(p,q) = |(e \cdot f)|$ where e, f are arbitrary chosen unit vectors of the lines p, q. It is shown in [4] and [5] that the linear programming method applies to these spaces. The derived asymptotic bound for the rate can also be obtained from the bounds for spherical codes, because to a code C in $\mathbb{P}^{n-1}(K)$ one can associate a code in S^{cn-1} with the same size and a minimum angular distance at least equal to the one of C, selecting a unit vector in each element of C. One obtains:

Theorem 2.3 ([4]): Let C be a code in $\mathbb{P}^{n-1}(K)$ with minimum angular distance $0 \le \theta \le \pi/2$ and rate $R(C) := \frac{1}{n} \ln |C|$. Let $c := 1, 2, 4$ respectively when $K = \mathbb{R}, \mathbb{C}, \mathbb{H}$ (so that $c = [K : \mathbb{R}]$). Then, when $n \to \infty$,

$$R(C) \lesssim c R_S(\theta) \tag{8}$$

□

B. *The Grassmann space*

Let K be the real or the complex field. The Grassmann space $\mathcal{G}_{m,n}(K)$ is the set of all subspaces of dimension m in K^n. It is a homogeneous space under the action of either the orthogonal group $O(\mathbb{R}^n)$ or the unitary group $U(\mathbb{C}^n)$. We will denote $\mathcal{G}_{m,n}$ when K is arbitrary. It is worth noticing that when $m = 1$ we recover the projective space. Several metrics have been defined in $\mathcal{G}_{m,n}$, see [19], [20]. In this paper we consider the *chordal distance*, which was introduced in [21] and studied in [19], [7], [6], [21], [20]. The chordal distance $d_c(p,q)$ is defined as follows.

Definition 2.4: Given the planes $p, q \in \mathcal{G}_{m,n}$, apply the following procedure. Initialize the sets of unit vectors $A = \emptyset$ and $B = \emptyset$. In the ith step, choose the vectors a_i, b_i such that:

(i) a_i is contained in p and b_i is contained in q.
(ii) a_i is orthogonal to all the vectors in A and b_i is orthogonal to all the vectors in B.
(iii) Among all the vectors satisfying the conditions in (i) and (ii), the angle between a_i and b_i is minimal (i.e., their inner / Hermitian product module is maximal).

Set θ_i to be the angle between a_i and b_i, insert a_i to A and b_i to B, and proceed to the next step, until m angles $0 \le \theta_1 \le \theta_2 \le \ldots \le \theta_m \le \pi/2$, called the principal angles between p and q, have been defined. Then the chordal distance is

$$d_c(p,q) := \sqrt{\sum_{i=1}^{m} \sin^2 \theta_i} = \sqrt{m - \sum_{i=1}^{m} \cos^2 \theta_i}.$$

□

The following lemma provides an equivalent definition, which is more convenient.

Lemma 2.5 ([21]): For a plane $p \in \mathcal{G}_{m,n}$, let A_p be a $p \times n$ matrix whose rows form an orthonormal basis of p, and let $\pi_p := A_p^* A_p$ be the matrix of the orthogonal projection on p (A_p^* denotes the Hermitian conjugate of A_p). Then, the projection matrix π_p does not depend on the choice of A_p, and, given two planes $p, q \in \mathcal{G}_{m,n}$, the chordal distance is

$$d_c(p,q) := \sqrt{m - \mathrm{trace}(\pi_p \circ \pi_q)}. \tag{9}$$

□

Bounds on the size of codes in Grassmann spaces were considered in [21], [19], [7], [6], [22]. The best known asymptotic upper bounds are given in the following theorem.

Theorem 2.6 ([7], [6]): Let C be a code in $\mathcal{G}_{m,n}(K)$ with minimum chordal distance d and rate $R(C) := \frac{1}{n} \ln |C|$. Then, when $n \to \infty$,

$$R(C) \lesssim -cm \ln \left(\sqrt{1 - \sqrt{1 - d^2/m}} \right) \tag{10}$$

and

$$R(C) \lesssim cm[(1+\rho)\ln(1+\rho) - \rho \ln \rho], \tag{11}$$

where

$$\rho = \frac{1}{2} m (\sqrt{m}/d - 1). \tag{12}$$

□

The bound (10) was derived by Barg and Nogin, using Blichfeldt's density method [7]. The bound (11) is a linear programming bound due to Bachoc [6]. Both works considered only the real Grassmann space, but can be easily extended to the complex case.

The main contribution of this paper is the following theorem.

Theorem 2.7: Let C be a code in $\mathcal{G}_{m,n}(K)$ with minimum chordal distance $d = \sqrt{m-s}$, and let $\theta = \arccos(\sqrt{s/m})$. Then, when $n \to +\infty$,

$$R(C) \lesssim \min\left\{R_1(d), R_2(d)\right\}, \tag{13}$$

where

$$R_1(d) = cmR_S(\theta) \tag{14}$$

and

$$R_2(d) = \min_{\substack{(\theta_1,\ldots,\theta_m) \in [0,\pi/2]^m \\ \sum_{i=1}^m \cos^2\theta_i = m\cos^2\theta}} c\big(R_{LP}(\theta_1) + \cdots + R_{LP}(\theta_m)\big) \tag{15}$$

(R_S and R_{LP} are defined in Theorem 2.1). □

Our method will be to establish relations between the Grassmann space and other spaces, and then apply bounds for codes in these spaces. A reduction from Grassmannian to spherical codes is given in Theorem 3.2, implying the bound $R_1(d)$. The bound $R_2(d)$ follows from a reduction to codes in products of projective spaces (Theorem 3.7) and Theorem 4.9. The bound $R_1(d)$ is better than the bounds of Theorem 2.6 for all values of d and m (see Lemma 3.3). We show in Section IV-F that for some values of d and m, it is further improved by $R_2(d)$.

C. *The Stiefel manifold*

The Stiefel manifold $\mathcal{V}_{m,n}(K)$ is the set of m-tuples of orthonormal vectors in K^n, or equivalently

$$\mathcal{V}_{m,n}(K) = \{X \in M^{m\times n}(K) \mid XX^* = \mathrm{Id}_m\},$$

where Id_m is the $m \times m$ identity matrix. The orthogonal group $O(\mathbb{R}^n)$ if $K = \mathbb{R}$, respectively the unitary group $U(\mathbb{C}^n)$ if $K = \mathbb{C}$ acts transitively on $\mathcal{V}_{m,n}(K)$, and this space can be identified with the set of classes $O(\mathbb{R}^n)/O(\mathbb{R}^{n-m})$, respectively $U(\mathbb{C}^n)/U(\mathbb{C}^{n-m})$.

The distance considered in coding theory is

$$d(X,Y) := \|X - Y\| = \sqrt{\mathrm{trace}((X-Y)(X^* - Y^*))}.$$

In other words, $d(X,Y)$ is the Euclidean distance between X and Y, when X and Y are regarded as one-dimensional vectors of length mn.

In [22], estimates for the volume of balls in the Stiefel manifold are given, approximated by the geodesic distance, and Gilbert-Varshamov and sphere packing bounds are derived for small distances.

The following theorem will be proved in Section III-B. It follows from a relation between codes in the Stiefel space and spherical codes (Theorem 3.6).

Theorem 2.8: Let C be a code in $\mathcal{V}_{m,n}(K)$ with minimum distance $d = \sqrt{2}\sqrt{m-s}$, and let $\theta = \arccos(s/m)$. Then, when $n \to +\infty$,

$$R(C) \lesssim cmR_S(\theta) \tag{16}$$

where R_S is defined in (4). □

III. More spaces and their interconnections

The simplest of these connections relate Grassmann and Stiefel spaces to a single unit sphere, and allow to apply directly the known bounds for spherical codes to the Grassmannian and Stiefel codes. We start with them, then we introduce the products of spheres and projective spaces and their relations with Grassmann and Stiefel spaces.

A. $\mathcal{G}_{m,n}$ *and* S^{cmn-1}

We define a mapping

$$\beta : \mathcal{G}_{m,n}(K) \to S^{cmn-1}$$

in the following way. We select for all $p \in \mathcal{G}_{m,n}(K)$, an orthonormal basis $(e_1, \dots, e_m)$ of p whose elements belong to K^n. With the usual identification of $\mathbb{C}$ and $\mathbb{R}\times\mathbb{R}$ through the mapping $z = x+iy \to (x,y)$, we consider these elements in $\mathbb{R}^{cn}$. Then $\beta(p)$ is chosen to be the element of $\mathbb{R}^{cmn}$ obtained by the concatenation of $e_1, \ldots, e_m$, divided by $\sqrt{m}$. Obviously, $\beta(p) \in S^{cmn-1}$.

For all $p, q \in \mathcal{G}_{m,n}(K)$, we set

$$\sigma(p,q) := \sum_{i=1}^{m} \cos^2 \theta_i = \operatorname{trace}(\pi_p \circ \pi_q).$$

Let us recall that the principal angles $\theta_1, \ldots, \theta_m$ associated to (p,q) are related to the construction of orthonormal basis $(a_1, \ldots, a_m)$ and $(b_1, \ldots, b_m)$ for p and q respectively, such that $\cos\theta_i = |(a_i \cdot b_i)|$ and $\sigma(p,q) = \sum_{i=1}^{m} \cos^2 \theta_i$. However, these orthonormal basis obviously depend on the pair (p,q) and not on p and q individually. There is no hope in the above construction of β to choose orthonormal basis that would satisfy the equality $\cos\theta_i = |(e_i \cdot e_i')|$ for all pairs (p,q). Still, and this is the main point of our construction, we do have a relation between $\sigma(p,q)$ and the *arbitrary chosen* orthonormal basis $(e_1, \ldots, e_m)$ of p and $(e_1', \ldots, e_m')$ of q, leading to the inequality

$$\sigma(p,q) \geq \sum_{i=1}^{m} |(e_i \cdot e_i')|^2.$$

In other words, if one defines an alternative set of "principal angles" $\theta_1', \ldots, \theta_m'$ by $\theta_i' = \arccos|(e_i \cdot e_i')|$, then $\sigma(p,q) \geq \sum_{i=1}^{m} \cos^2 \theta_i'$, leading to an upper bound on the chordal distance between p and q.

In the next lemma we prove the above assertions and settle the inequality we aim at in terms of the embedding β.

Lemma 3.1: For all $p, q \in \mathcal{G}_{m,n}(K)$,

$$\cos\theta(\beta(p), \beta(q)) \leq \sqrt{\frac{\sigma(p,q)}{m}}.$$

Proof: Let $\beta(p) = e$, obtained from an orthonormal basis $(e_1, e_2, \ldots, e_m)$ of p and $\beta(q) = e'$, obtained from an orthonormal basis $(e_1', \ldots, e_m')$ of q. We compute $\sigma(p,q) = \operatorname{trace}(\pi_p \circ \pi_q)$. Let A_p, A_q denote the $m \times n$ matrices whose rows are the basis elements e_i, e_i' respectively. Then

$$\sigma(p,q) = \operatorname{trace}(\pi_p \circ \pi_q) = \operatorname{trace}(A_p^* A_p A_q^* A_q) = \operatorname{trace}(A_p A_q^* A_q A_p^*).$$

The entries of the matrix $A_p A_q^*$ are the hermitian products $(e_i \cdot e_j')$. So we obtain:

$$\sigma(p,q) = \sum_{1 \leq i,j \leq m} |(e_i \cdot e_j')|^2 \tag{17}$$

and hence

$$\sigma(p,q) \geq \sum_{i=1}^{m} |(e_i \cdot e_j')|^2.$$

If $K = \mathbb{R}$, we obtain from Cauchy-Schwartz inequality

$$\cos\theta(e,e') = (e \cdot e') = \frac{\sum_{i=1}^{m} (e_i \cdot e_i')}{m} \leq \sqrt{\frac{\sum_{i=1}^{m} (e_i \cdot e_i')^2}{m}} \leq \sqrt{\frac{\sigma(p,q)}{m}}. \tag{18}$$

If $K = \mathbb{C}$, let us denote by $\Re(z)$ the real part of a complex number z. In the identification $\mathbb{C}^n = \mathbb{R}^{2n}$ recalled above, the standard scalar product on $\mathbb{R}^{2n}$ is given by $\Re((x \cdot y))$. With the obvious inequality $(\Re(x \cdot y))^2 \leq |(x \cdot y)|^2$, we obtain the same inequality $\cos\theta(e,e') \leq \sqrt{\frac{\sigma(p,q)}{m}}$ (where e and e' are considered in the unit sphere of R^{2mn}.)

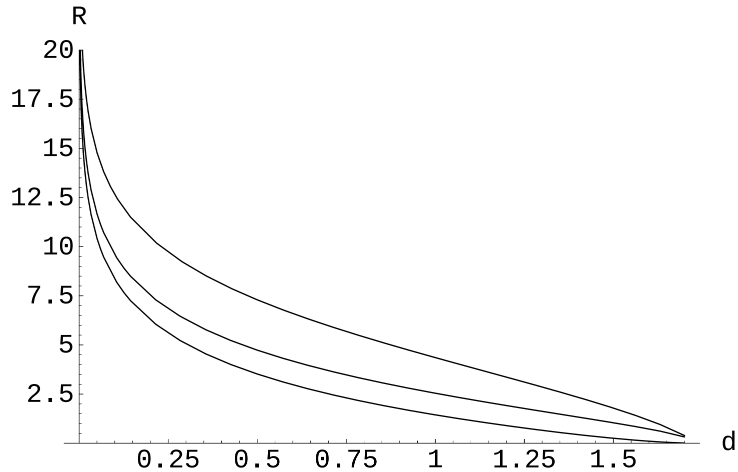

Fig. 1. Upper bound on the asymptotic rate of real Grassmannian codes with minimum chordal distance d, $m = 3$. From top to bottom: a linear programming bound (11), a Blichfeldt-type bound (10), and the new bound (19)

$\square$

The following theorem is an immediate consequence of Lemma 3.1. It states that bounds for spherical codes can be applied to codes in Grassmann spaces.

Theorem 3.2: Any upper bound on the size of codes in S^{cmn-1} with minimum angular distance $\theta = \arccos(\sqrt{s/m})$ is also an upper bound for codes in the Grassmann space $\mathcal{G}_{m,n}(K)$ with minimum distance $d = \sqrt{m-s}$. $\square$

As a corollary, we obtain the bound

$$R(C) \lesssim R_1(d) \tag{19}$$

of Theorem 2.7.

Figure 1 compares (19) with the best known asymptotic bounds, given in Theorem 2.6. It can be readily checked that for $m = 1$, (19) is equal to (11) and is smaller than (10). For $m > 1$, we have the following lemma (note that by definition, the chordal distance is always upper-bounded by $\sqrt{m}$).

Lemma 3.3: The bound (19) is smaller than the bounds of Theorem 2.6 for all values of $m > 1$ and $d \leq \sqrt{m}$.

Proof: Denote by $R_B(\theta)$ the Rankin-Blichfeldt upper bound on spherical codes with minimum angular distance θ [18]. Then the bound (10) can be expressed as

$$R(C) \lesssim cmR_B(\theta) \tag{20}$$

for $\theta = \arccos(\sqrt{s/m})$, where $s = m - d^2$. The comparison between (19) and (10) is thus reduced to a comparison between R_S and R_B. It is easy to check that R_B is larger than R_S for all values of θ.

It remains to compare (19) with (11). By definition, $R_S(\theta) \leq R_{LP}(\theta)$ (R_{LP} is defined in (5)). Therefore, it is sufficient to show that (19) with R_{LP} instead of R_S is always smaller than (11). We note that $\theta = \arccos(\sqrt{s/m}) = \arcsin(d/\sqrt{m})$. Hence, applying (19) with R_{LP} instead of R_S, we obtain:

$$R(C) \lesssim cm[(1+\rho)\ln(1+\rho) - \rho \ln \rho] \tag{21}$$

with

$$\rho = \frac{1}{2}(\sqrt{m}/d - 1). \tag{22}$$

The difference between this bound and (11) is only in the definition of ρ. The claim now follows from the fact that $f(\rho) = (1+\rho)\ln(1+\rho) - \rho\ln\rho$ is an increasing function in ρ. □

Remark 3.4: We have defined an embedding of $\mathcal{G}_{m,n}(K)$ into the unit sphere S^{cmn-1}. The dimensions of these two spaces, $cm(n-m)$ and $cmn-1$, are asymptotically equivalent. This can be compared to another embedding of $\mathcal{G}_{m,n}(\mathbb{R})$ into a unit sphere, introduced in [21]. The dimension of this sphere is $n(n+1)/2-2$, which is much larger than the one of the Grassmann space. However, unlike our embedding, the embedding from [21] is also an isometry. □

B. $\mathcal{V}_{m,n}$ and S^{cmn-1}

Lemma 3.5: Let $X, Y \in \mathcal{V}_{m,n}(K)$, $K = \mathbb{R}, \mathbb{C}$. Let $(e_1, \dots e_m)$ denote the rows of X, respectively $(e'_1, \dots e'_m)$ for the rows of Y. Then

$$d(X,Y) = \sqrt{2}\sqrt{m - \sum_{i=1}^{m} \Re((e_i \cdot e'_i))}.$$

Proof: We calculate

$$\begin{aligned} \|X-Y\|^2 &= \mathrm{trace}((X-Y)(X^*-Y^*)) = \mathrm{trace}(XX^* - XY^* - YX^* + YY^*) \\ &= 2m - 2\Re(\mathrm{trace}(XY^*)) \end{aligned}$$

since $XX^* = YY^* = \mathrm{Id}_m$ and $\overline{\mathrm{trace}(XY^*)} = \mathrm{trace}(\overline{XY^*}) = \mathrm{trace}(YX^*)$. We conclude with

$$\mathrm{trace}(XY^*) = \sum_{i=1}^{m}(e_i \cdot e'_i).$$

□

Again with the identification of $\mathbb{C}^n$ with $\mathbb{R}^{2n}$, we view $\mathcal{V}_{m,n}(\mathbb{C})$ as a submanifold of $\mathcal{V}_{m,2n}(\mathbb{R})$ endowed with the distance

$$d(X,Y) = \|X-Y\| = \sqrt{2}\sqrt{m - \sum_{i=1}^{m}(e_i \cdot e'_i)}.$$

We see that the obvious mapping:

$$\gamma : \mathcal{V}_{m,n}(K) \to S^{cmn-1}$$
$$X \mapsto \gamma(X) = \frac{1}{\sqrt{m}}(e_1, \dots, e_m)$$

is this time, up to a suitable scaling of the distances, an isometry. Hence the bounds for spherical codes also apply to $\mathcal{V}_{m,n}(K)$, probably in a quite efficient way. Still, one constraint is not encoded in it: the fact that the vectors e_i are pairwise orthogonal and of norm 1. We resume these observations in the following theorem.

Theorem 3.6: Any upper bound on the size of codes in S^{cmn-1} with minimum angular distance $\theta = \arccos(s/m)$ is also an upper bound for codes in the Stiefel space $\mathcal{V}_{m,n}(K)$ with minimum distance $d = \sqrt{2}\sqrt{m-s}$. □

Theorem 2.8 now easily follows as a corollary of Theorem 3.6.

C. $\mathcal{G}_{m,n}$, $\mathcal{V}_{m,n}$ and products of spaces

So far we have established a relation between codes in $\mathcal{G}_{m,n}$ and $\mathcal{V}_{m,n}$ and codes in S^{cmn-1}. We note that the ranges of the mappings β and γ contain normalized vectors from $\left(S^{cn-1}\right)^m$ (concatenations of unit vectors $e_1, \dots, e_m$ divided by $\sqrt{m}$). Hence bounds for codes in $\left(S^{cn-1}\right)^m$ will imply bounds for codes in $\mathcal{G}_{m,n}$ and $\mathcal{V}_{m,n}$. This is the motivation to the generalization of the linear programming method to the product of unit spheres, and more generally to the product of 2-point homogeneous spaces, which is proposed in the next section. As we shall see, the asymptotic bound for the rate of codes in $\left(S^{cn-1}\right)^m$ is not better than for S^{cmn-1}, hence does not improve on (19) and (16). A better result is obtained for Grassmann spaces with the product of projective spaces.

We now define more precisely the products of spaces and their associated distances that will be studied in the next section. We start with the product of m copies of the unit sphere of $\mathbb{R}^n$:

$$\left(S^{n-1}\right)^m = \{e = (e_1, \dots, e_m) \mid e_i \in S^{n-1}\}.$$

We consider on $\left(S^{n-1}\right)^m$ the distance given by

$$d(e, e') = \sqrt{\sum_{i=1}^{m} \|e_i - e'_i\|^2}$$
$$= \sqrt{2m}\sqrt{1 - \frac{\sum_{i=1}^{m} \cos\theta_i}{m}},$$

where $\cos\theta_i = (e_i \cdot e'_i)$. We attach to a pair $e, e' \in \left(S^{n-1}\right)^m$ an angle $\theta = \theta(e, e') \in [0, \pi]$ such that

$$\cos\theta = \frac{\sum_{i=1}^{m} \cos\theta_i}{m} \tag{23}$$

and we call θ the *angular distance* between e and e'. The angle θ is also the angle between the vectors $e/\sqrt{m}$ and $e'/\sqrt{m}$, viewed as elements of S^{mn-1}.

We define, for the remaining 2-point homogeneous spaces recalled above, and without specifying the field K,

$$\left(\mathbb{P}^{n-1}\right)^m = \{p = (p_1, \dots, p_m) \mid p_i \in \mathbb{P}^{n-1}\}.$$

We attach to a pair $p, p' \in \left(\mathbb{P}^{n-1}\right)^m$ an angle $\theta = \theta(p, p') \in [0, \pi/2]$ such that

$$\cos^2 \theta = \frac{\sum_{i=1}^m \cos^2 \theta_i}{m}$$

where $\theta_i = \theta(p_i, q_i)$ and we call θ the *angular distance* between p and p'. We consider on $\left(\mathbb{P}^{n-1}\right)^m$ the "chordal" distance given by

$$d(p, p') = \sqrt{m - \sum_{i=1}^m \cos^2 \theta_i} = \sqrt{m} \sin \theta.$$

In order to derive bounds for codes in Grassmann spaces $\mathcal{G}_{m,n}(K)$, we shall make use of the mapping:

$$\nu : \mathcal{G}_{m,n}(K) \to \left(\mathbb{P}^{n-1}(K)\right)^m$$

defined in the following way: for all $p \in \mathcal{G}_{m,n}(K)$, we choose a m-tuple $(p_1, \ldots, p_m)$ of pairwise orthogonal lines of p. We set $\nu(p) = (p_1, \ldots, p_m)$.

Because of the equation (17), we have similarly:

$$\cos^2 \theta(\nu(p), \nu(q)) \leq \frac{\sigma(p, q)}{m}$$

hence the bounds for codes in $\left(\mathbb{P}^{n-1}(K)\right)^m$ apply to codes in Grassmann spaces. We summarize in the following theorem:

Theorem 3.7: Any upper bound on the size of codes in $\left(\mathbb{P}^{n-1}(K)\right)^m$ with minimum angular distance $\theta = \arccos(\sqrt{s/m})$ is also an upper bound for codes in the Grassmann space $\mathcal{G}_{m,n}(K)$ with minimum distance $d = \sqrt{m-s}$.

IV. Bounds for codes in the product of 2-point homogeneous spaces

In this section, X denotes one of the spaces S^{n-1}, $\mathbb{P}^{n-1}(K)$ where $K = \mathbb{R}, \mathbb{C}, \mathbb{H}$, or the projective plane over the octonions $\mathbb{P}^2(\mathbb{O})$. We derive bounds for codes in X^m with a given minimum distance, following Delsarte's linear programming method as performed in [4]. As a reference on orthogonal polynomials, we refer to [23].

A. Review of the necessary material on the spaces X

We recall that, to each of these spaces is associated a family of orthogonal polynomials of one variable, which are the zonal polynomials relative to the action of the group G (see [4], [5], [24]). For $X = S^{n-1}$, these polynomials are the Gegenbauer polynomials with parameter $n/2 - 1$ and associated orthogonal measure $(1 - x^2)^{(n-3)/2}$ on the interval $[-1, 1]$. For $X = \mathbb{P}^{n-1}(K)$, these polynomials are Jacobi polynomials with parameters (α, β) defined by:

$$\alpha = \frac{c}{2}(n-1) - 1, \;\; \beta = \frac{c}{2} - 1.$$

More precisely, the values of (α,β) are as follows:

	c	α	β
$\mathbb{R}$	1	$(n-3)/2$	$-1/2$
$\mathbb{C}$	2	$n-2$	0
$\mathbb{H}$	4	$2n-3$	1
$\mathbb{O}$	8	7	3

The orthogonal measure associated to the parameters (α,β) is $x^\beta(1-x)^\alpha$ over the interval $[0,1]$. We generically denote by $P_k(x)$ these polynomials, with $\deg(P_k)=k$ and $P_k(1)=1$. We let $\mu(x)$ denote their normalized associated orthogonal measure and $[P,Q]$ the corresponding scalar product on $\mathbb{R}[x]$ (so that $[P,Q]=\int P(x)Q(x)\mu(x)dx$ and $[1,1]=1$). Moreover, we have $[P_k,P_k]=d_k^{-1}$ where d_k denotes the dimension of the irreducible representation of G associated to P_k (e.g. when $X=S^{n-1}$, $d_k=\dim \operatorname{Harm}_k=\binom{n+k-1}{k}-\binom{n+k-3}{k-2}$, where Harm_k is the kernel of the laplacian operator, see [13]).

The three-term relation expresses $xP_k(x)$ as a linear combination of the polynomials P_i:

$$xP_k(x)=a_kP_{k+1}(x)+b_kP_k(x)+c_kP_{k-1}(x)$$

for some sequences of rational numbers (a_k), (b_k), (c_k). It is enough for our purpose to know that (a_k) is bounded when n and k tend to $+\infty$ with n/k tending to a finite limit. For example, when $X=S^{n-1}$,

$$a_k=\frac{n-2+k}{n-2+2k}.$$

For all $(u,v)\in X$, we define

$$t(u,v)=\begin{cases}(u\cdot v)=\cos\theta(u,v) & \text{if } X=S^{n-1}\\ \cos^2\theta(u,v) & \text{if } X=\mathbb{P}^{n-1}(K).\end{cases}$$

The zonal function on X associated to P_k is given by:

$$(u,v)\to P_k(t(u,v)).$$

The so-called 'positivity property', related to these polynomials, is the basic principle underlying the linear programming method in X: for all code $C\subset X$, and for all $k\geq 0$,

$$\sum_{u\in C}\sum_{v\in C}P_k(t(u,v))\geq 0.$$

B. The linear programming method on X^m

Now we consider the product spaces X^m. The positivity property generalizes to the following:

Lemma 4.1: Let $C\subset X^m$. Let us denote elements of C by $u=(u_1,\dots,u_m)$ with $u_i\in X$. For all $(k_1,\dots,k_m)\in\mathbb{N}^m$,

$$\sum_{u\in C}\sum_{v\in C}\prod_{i=1}^m P_{k_i}(t(u_i,v_i))\geq 0.$$

Proof: This is the positivity property in the product space X^m. The group G^m acts transitively on X^m; the G^m-irreducible components of $L^2(X^m)$ are the tensor products of the G-irreducible components of each $L^2(X)$ and the associated zonal functions are given by the polynomials in the m variables $x_1, \dots, x_m$

$$\prod_{i=1}^{m} P_{k_i}(x_i), \quad (k_1, \dots, k_m) \in \mathbb{N}^m$$

in the way:

$$(u, v) \mapsto \prod_{i=1}^{m} P_{k_i}(t(u_i, v_i)).$$

□

Remark 4.2: In a sense, the polynomials $\prod_{i=1}^{m} P_{k_i}(x_i)$ are fake multivariate polynomials since the m variables are separated. As we shall see, for this reason it is much easier to deal with them, compared with the zonal polynomials for the Grassmann space (see [6]). □

The polynomials $\prod_{i=1}^{m} P_{k_i}(x_i)$ generate the polynomial algebra $\mathbb{C}[x_1, \dots, x_m]$, and are orthogonal for the product measure

$$\lambda \prod_{i=1}^{m} \mu(x_i) dx_i$$

with support $[-1, 1]^m$ when $X = S^{n-1}$, respectively $[0, 1]^m$ otherwise, and where λ is chosen so that the total measure is equal to 1. The associated scalar product on $\mathbb{R}[x_1, \dots, x_m]$ is denoted by $[,]$. We take the following notations: a multi-index in $\mathbb{N}^m$ is denoted by $\underline{k} = (k_1, \dots, k_m)$ and we define for $x = (x_1, \dots, x_m)$, $P_{\underline{k}}(x) = P_{\underline{k}}(x_1, \dots, x_m) := \prod_{i=1}^{m} P_{k_i}(x_i)$, and $d_{\underline{k}} := \prod_{i=1}^{m} d_{k_i}$. Obviously we have, for all $\underline{k}$ and $\underline{l}$,

$$\left[P_{\underline{k}}, P_{\underline{l}}\right] = \delta_{\underline{k},\underline{l}} d_{\underline{k}}^{-1}. \tag{24}$$

Moreover, we define

$$\sigma(x) := \sum_{i=1}^{m} x_i.$$

For any angle θ we denote

$$\begin{cases} t = \cos\theta & \text{if } X = S^{n-1} \\ t = \cos^2\theta & \text{if } X = \mathbb{P}^{n-1}(K) \end{cases}$$

Now we can formulate the usual associated linear programming bound:

Proposition 4.3: Assume $F \in \mathbb{R}[x_1, \dots, x_m]$ satisfies the conditions:

(i) $F = \sum_{\underline{k}} f_{\underline{k}} P_{\underline{k}}$ with $f_{\underline{k}} \geq 0$ for all $\underline{k}$, and $f_0 > 0$

$$(ii) \begin{cases} \text{If } X = S^{n-1}, \quad F(x_1, \dots, x_m) \leq 0 \text{ for all } (x_1, \dots, x_m) \in [-1, 1]^m \text{ such that } \sigma(x) \leq m\cos\theta = mt \\ \text{If } X \neq S^{n-1}, \quad F(x_1, \dots, x_m) \leq 0 \text{ for all } (x_1, \dots, x_m) \in [0, 1]^m \text{ such that } \sigma(x) \leq m\cos^2\theta = mt \end{cases}$$

Then, any code C in X^m with minimum angular distance θ satisfies

$$|C| \leq \frac{F(1,\dots,1)}{f_0}$$

Proof: We reproduce the standard argument. Let

$$S := \sum_{u\in C}\sum_{v\in C} F(t(u_1,v_1),\dots,t(u_m,v_m)).$$

The contribution in this sum of the pairs (u,v) with $u = v$ is of $|C|F(1,\dots,1)$. From condition (ii) and the assumption that for $u \neq v \in C$, $\cos\theta(u,v) \leq \cos\theta$, the other terms are non positive. Hence, $S \leq |C|F(1,\dots,1)$.

On the other hand, we have

$$S = \sum_{\underline{k}} f_{\underline{k}} \sum_{u,v\in C} P_{\underline{k}}(t(u_1,v_1),\dots,t(u_m,v_m)).$$

The term corresponding to $\underline{k} = 0 = (0,\dots,0)$ gives $f_0|C|^2$ while the other terms are non-negative from the positivity property of the polynomials $P_{\underline{k}}$ (Lemma 4.1). Hence $S \geq f_0|C|^2$. The two inequalities lead to the announced bound. $\square$

Remark 4.4: Note that $F(1,\dots,1) = \sum_{\underline{k}} f_{\underline{k}} P_{\underline{k}}(1,\dots,1) = \sum_{\underline{k}} f_{\underline{k}}$ is indeed always positive for $f_{\underline{k}} \geq 0$ and $f_0 > 0$.

C. Examples of small degree

Let us work out the case of polynomials of small degree.

(i) $X = S^{n-1}$

a) Degree 1: we take $F = (x_1+\cdots+x_m) - mt$. Since $P_1(x) = x$, F satisfies the hypothesis of Proposition 4.3 if and only if $t < 0$. We obtain:

$$\text{If } \cos\theta = t < 0, \quad |C| \leq 1 - \frac{1}{t}. \tag{25}$$

b) Degree 2: we take $F = \big((x_1+\cdots+x_m)+m\big)\big((x_1+\cdots+x_m) - mt\big)$. We have

$$\begin{aligned} F &= (x_1+\cdots+x_m)^2 + m(1-t)(x_1+\cdots+x_m) - m^2 t \\ &= \sum x_i^2 + 2\sum_{i<j} x_i x_j + m(1-t)\sum x_i - m^2 t \\ &= \sum (x_i^2 - \frac{1}{n}) + 2\sum_{i<j} x_i x_j + m(1-t)\sum x_i + \frac{m}{n} - m^2 t \end{aligned}$$

Since $P_2(x) = (x^2 - 1/n)/(1-1/n)$, F satisfies the hypothesis of Proposition 4.3 if and only if $\frac{m}{n} - m^2 t > 0$. We obtain:

$$\text{If } \cos\theta = t < \frac{1}{mn}, \quad |C| \leq \frac{2mn(1-t)}{1-mnt}. \tag{26}$$

The crossing point of the two curves corresponds to $t=-1/mn$, where the two bounds take the value $1+mn$.

(ii) $X=\mathbb{P}^{n-1}(K)$, Degree 1: we have, up to a multiplicative factor, $P_1(x)=x-\frac{\beta+1}{\alpha+\beta+2}=x-\frac{1}{n}$. We take $F=(x_1+\cdots+x_m)-mt=(x_1-\frac{1}{n})+\cdots+(x_m-\frac{1}{n})+\frac{m}{n}-mt$. F satisfies the hypothesis of Proposition 4.3 if and only if $t<1/n$. We obtain:

$$\text{If } \cos^2\theta=t<\frac{1}{n},\quad |C|\leq\frac{1-t}{1/n-t}. \tag{27}$$

D. Christoffel-Darboux formula and an explicit bound

It remains to apply the standard method with Christoffel-Darboux formula. For $\underline{k}=(k_1,\dots,k_m)$ and $\underline{l}=(l_1,\dots,l_m)$, the notation $\underline{l}\leq\underline{k}$ stands for: $l_i\leq k_i$ for all $1\leq i\leq m$.

Proposition 4.5: Let $y=(y_1,\dots,y_m)\in\mathbb{R}^m$ and $\underline{k}=(k_1,\dots,k_m)\in\mathbb{N}^m$, and define

$$K_{\underline{k}}(x,y):=\sum_{\underline{l}\leq\underline{k}}d_{\underline{l}}P_{\underline{l}}(x)P_{\underline{l}}(y)=\prod_{j=1}^{m}\big(\sum_{i=0}^{k_j}d_iP_i(x_j)P_i(y_j)\big)$$

and

$$N_{\underline{k}}(x,y):=\sum_{t=1}^{m}d_{k_t}a_{k_t}Q_{k_t}(x_t,y_t)\prod_{j\neq t}\big(\sum_{i=0}^{k_j}d_iP_i(x_j)P_i(y_j)\big)$$

where

$$Q_{k_t}(x_t,y_t):=P_{k_t+1}(x_t)P_{k_t}(y_t)-P_{k_t}(x_t)P_{k_t+1}(y_t).$$

Then we have the Christoffel-Darboux type formula:

$$K_{\underline{k}}(x,y)=\frac{N_{\underline{k}}(x,y)}{\sigma(x)-\sigma(y)}.$$

Proof: Since $\sigma(x)-\sigma(y)=\sum_{t=1}^m x_t-\sum_{t=1}^m y_t=\sum_{t=1}^m(x_t-y_t)$,

$$\begin{aligned}(\sigma(x)-\sigma(y))K_{\underline{k}}(x,y)&=\Big(\sum_{t=1}^{m}(x_t-y_t)\Big)\prod_{j=1}^{m}\big(\sum_{i=0}^{k_j}d_iP_i(x_j)P_i(y_j)\big)\\&=\sum_{t=1}^{m}\Big((x_t-y_t)\sum_{i=0}^{k_t}d_iP_i(x_t)P_i(y_t)\Big)\prod_{j\neq t}\big(\sum_{i=0}^{k_j}d_iP_i(x_j)P_i(y_j)\big)\end{aligned}$$

The Christoffel-Darboux formula for the polynomials P_k gives:

$$(x_t-y_t)\sum_{i=0}^{k_t}d_iP_i(x_t)P_i(y_t)=d_{k_t}a_{k_t}Q_{k_t}(x_t,y_t)$$

with the notations of the proposition, hence the result.

□

Following the standard method, we apply Proposition 4.3 to the function

$$\frac{N_{\underline{k}}(x,y)^2}{\sigma(x)-\sigma(y)}.$$

Proposition 4.6: Let $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$ and $\underline{k} \in \mathbb{N}^m$, and define

$$F(x) := \frac{N_{\underline{k}}(x,y)^2}{\sigma(x) - \sigma(y)} = (\sigma(x) - \sigma(y))K_{\underline{k}}(x,y)^2 = K_{\underline{k}}(x,y)N_{\underline{k}}(x,y).$$

If y satisfies the conditions:

(i) $P_i(y_t) \geq 0$ for all $0 \leq i \leq k_t$ and for all $1 \leq t \leq m$

(ii) $P_{k_t+1}(y_t) \leq 0$ for all $1 \leq t \leq m$

then F satisfies the hypothesis of Proposition 4.3 for all θ such that $mt \leq \sigma(y)$. Consequently, for any code C in X^m with minimum angular distance θ,

$$|C| \leq \frac{\Big(\sum_{t=1}^m d_{k_t} a_{k_t} \big(P_{k_t}(y_t) - P_{k_t+1}(y_t)\big) \prod_{j \neq t} \big(\sum_{i=0}^{k_j} d_i P_i(y_j)\big)\Big)^2}{-(m - \sigma(y)) \sum_{t=1}^m d_{k_t} a_{k_t} P_{k_t}(y_t) P_{k_t+1}(y_t) \prod_{j \neq t} \Big(\sum_{i=0}^{k_j} d_i (P_i(y_j))^2\Big)}. \tag{28}$$

Proof: Clearly, under the assumptions (i) and (ii), $K_{\underline{k}}(x,y)$ and $N_{\underline{k}}(x,y)$ have non-negative coefficients on the $P_{\underline{l}}$. This is enough to ensure that it is also the case for the product $K_{\underline{k}}(x,y)N_{\underline{k}}(x,y)$ (recall that the product of two polynomials with non-negative coefficients on the P_k also has non-negative coefficients on the P_k. This property transfers straightforwardly to the $P_{\underline{k}}$; it is anyway general to any family of zonal polynomials).

Obviously the sign of $F(x)$ is the sign of $\sigma(x) - \sigma(y)$ so the conditions of Proposition 4.3 are fulfilled.

It remains to compute $f_0 = [F, 1]$ and $F(1, \ldots, 1)$.

$$\begin{aligned}
&[F,1] = [K, N] \\
&= \Big[\prod_{j=1}^m \big(\sum_{i=0}^{k_j} d_i P_i(x_j)P_i(y_j)\big), \sum_{t=1}^m d_{k_t} a_{k_t} Q_{k_t}(x_t, y_t) \prod_{j \neq t} \big(\sum_{i=0}^{k_j} d_i P_i(x_j) P_i(y_j)\big)\Big] \\
&= \sum_{t=1}^m d_{k_t} a_{k_t} \Big[\prod_{j=1}^m \big(\sum_{i=0}^{k_j} d_i P_i(x_j)P_i(y_j)\big), Q_{k_t}(x_t, y_t) \prod_{j \neq t} \big(\sum_{i=0}^{k_j} d_i P_i(x_j) P_i(y_j)\big)\Big] \\
&= \sum_{t=1}^m d_{k_t} a_{k_t} \Big[\sum_{i=0}^{k_t} d_i P_i(x_t)P_i(y_t), Q_{k_t}(x_t, y_t)\Big] \prod_{j \neq t} \Big[\sum_{i=0}^{k_j} d_i P_i(x_j) P_i(y_j), \sum_{i=0}^{k_j} d_i P_i(x_j) P_i(y_j)\Big] \\
&= \sum_{t=1}^m d_{k_t} a_{k_t} \big(-P_{k_t}(y_t) P_{k_t+1}(y_t)\big) \prod_{j \neq t} \big(\sum_{i=0}^{k_j} d_i (P_i(y_j))^2\big)
\end{aligned}$$

where the last equality follows from (24).

Let us now compute $F(1, \ldots, 1)$. We have:

$$F(1, \ldots, 1) = \frac{N_{\underline{k}}(\underline{1}, y)^2}{m - \sigma(y)}$$

and

$$N_{\underline{k}}(\underline{1}, y) = \sum_{t=1}^m d_{k_t} a_{k_t} \big(P_{k_t}(y_t) - P_{k_t+1}(y_t)\big) \prod_{j \neq t} \big(\sum_{i=0}^{k_j} d_i P_i(y_j)\big).$$

Applying the resulting bound of Proposition 4.3 leads to the announced bound.

∎

We proceed to choose the parameters y and $\underline{k}$ such that the conditions of Proposition 4.6 will be satisfied. We follow the standard method. Let $a_1^{(k)} < \ldots < a_k^{(k)}$ be the zeros of the polynomial P_k. They admit the following interlacing properties [23]: $a_i^{(k)} < a_i^{(k-1)} < a_{i+1}^{(k)}$ for all $1 \leq i < k$. We choose the multi-index $\underline{k}$ such that $mt \leq \sum_{t=1}^m z_{k_t}$, where $z_{k_t} = a_{k_t}^{(k_t)}$ is the largest zero of P_{k_t}. It follows from the interlacing property that there exists y such that $z_{k_t} \leq y_t \leq z_{k_t+1}$ and

$$P_{k_t}(y_t) + P_{k_t+1}(y_t) = 0$$

for all $1 \leq t \leq m$. Thus, $P_i(y_t) > 0$ and $P_{k_t+1}(y_t) < 0$ for all $0 \leq i \leq k_t$ and $1 \leq t \leq m$.

Now we have:

$$\begin{aligned} f_0 = [F, 1] &= \sum_{t=1}^m d_{k_t} a_{k_t} \big(P_{k_t}(y_t)\big)^2 \prod_{j \neq t} \big(\sum_{i=0}^{k_j} d_i (P_i(y_i))^2\big) \\ &= \sum_{t=1}^m a_{k_t} \sum_{\substack{\underline{l} \leq \underline{k} \\ l_t = k_t}} d_{\underline{l}} \big(P_{\underline{l}}(y)\big)^2 \end{aligned}$$

and

$$\begin{aligned} N_{\underline{k}}(\underline{1}, y) &= 2 \sum_{t=1}^m d_{k_t} a_{k_t} P_{k_t}(y_t) \prod_{j \neq t} \big(\sum_{i=0}^{k_j} d_i P_i(y_j)\big) \\ &= 2 \sum_{t=1}^m a_{k_t} \sum_{\substack{\underline{l} \leq \underline{k} \\ l_t = k_t}} d_{\underline{l}} P_{\underline{l}}(y). \end{aligned}$$

With Cauchy-Schwartz inequality (applied twice),

$$\begin{aligned} F(1, \ldots, 1) &= \frac{4}{(m - \sigma(y))} \Big(\sum_{t=1}^m a_{k_t} \sum_{\substack{\underline{l} \leq \underline{k} \\ l_t = k_t}} d_{\underline{l}} P_{\underline{l}}(y) \Big)^2 \\ &\leq \frac{4}{(m - \sigma(y))} \Big(\sum_{t=1}^m a_{k_t} \Big) \Big(\sum_{t=1}^m a_{k_t} \big(\sum_{\substack{\underline{l} \leq \underline{k} \\ l_t = k_t}} d_{\underline{l}} P_{\underline{l}}(y) \big)^2 \Big) \\ &\leq \frac{4}{(m - \sigma(y))} \Big(\sum_{t=1}^m a_{k_t} \Big) \Big(\sum_{t=1}^m a_{k_t} \big(\sum_{\substack{\underline{l} \leq \underline{k} \\ l_t = k_t}} d_{\underline{l}} \big) \big(\sum_{\substack{\underline{l} \leq \underline{k} \\ l_t = k_t}} d_{\underline{l}} \big(P_{\underline{l}}(y)\big)^2 \big) \Big) \\ &\leq \frac{4}{(m - \sigma(y))} \Big(\sum_{t=1}^m a_{k_t} \Big) \Big(\sum_{\underline{l} \leq \underline{k}} d_{\underline{l}} \Big) \Big(\sum_{t=1}^m a_{k_t} \sum_{\substack{\underline{l} \leq \underline{k} \\ l_t = k_t}} d_{\underline{l}} \big(P_{\underline{l}}(y)\big)^2 \Big) \\ &= \frac{4}{(m - \sigma(y))} \Big(\sum_{t=1}^m a_{k_t} \Big) \prod_{t=1}^m \big(\sum_{i=0}^{k_t} d_i \big) f_0 \end{aligned}$$

Denote $D_{k_t} := \sum_{i=0}^{k_t} d_i$. We obtain

$$|C| \leq \frac{4\big(\sum_{t=1}^m a_{k_t}\big) \prod_{t=1}^m D_{k_t}}{m - \sigma(y)}.$$

We summarize the above result in the following statement:

Proposition 4.7: For any code C in X^m with minimum angular distance θ, for any multi-index $\underline{k}$ such that $mt \leq \sum_{t=1}^m z_{k_t}$, let y_t satisfy $z_{k_t} \leq y_t \leq z_{k_t+1}$ and

$$P_{k_t}(y_t) + P_{k_t+1}(y_t) = 0,$$

then

$$|C| \leq \frac{4\big(\sum_{t=1}^m a_{k_t}\big)\prod_{t=1}^m D_{k_t}}{m - \sigma(y)}. \tag{29}$$

□

Remark 4.8: Using the so-called adjacent polynomials instead of the Gegenbauer polynomials, an enhancement of (29) was derived for $m = 1$ [5], [13]. It is very likely that this can be generalized for all m. □

E. A bound for the asymptotic rate

Now we consider the limit when $n \to +\infty$ of the rate $R(C) := \frac{1}{n}\ln|C|$ (of course the space $\mathbb{P}^2(\mathbb{O})$ is not concerned anymore) of the codes C of X^m. We derive an upper bound for this limit from (29). The next theorem settles the result obtained that way only in the case $X = \mathbb{P}^{n-1}(K)$ because this bound, in the case of $X = S^{n-1}$, turns out to be the same as the one obtained from the trivial isometric embedding $\left(S^{n-1}\right)^m \to S^{mn-1}$ (see Remark 4.10).

Theorem 4.9: Let C be a code in X^m, $X = \mathbb{P}^{n-1}(K)$, with minimum angular distance θ, and let $(\theta_1, \ldots, \theta_m) \in [0, \pi/2]^m$ satisfy $\sum_{t=1}^m \cos^2\theta_t = m\cos^2\theta$. Then, when $n \to \infty$,

$$R(C) \lesssim c(R_{LP}(\theta_1) + \ldots + R_{LP}(\theta_m)), \tag{30}$$

where R_{LP} is defined in (5).

Proof: Same as in [4], involving the asymptotic estimate of z_k. We reproduce it here: Consider an infinite sequence $k(n)$ such that $2k(n)/cn$ tends to a finite limit ρ as n tends to infinity. Then [4]

$$\lim_{n\to\infty} z_{k(n)} = 4\frac{\rho^{-1}+1}{(\rho^{-1}+2)^2}$$

and, since from [5],

$$D_k \simeq \binom{\frac{c}{2}n + k - 1}{k}^2$$

we have

$$\lim_{n\to\infty}\frac{1}{n}\ln D_{k(n)} = \lim_{n\to\infty}\frac{2}{n}\ln\binom{\frac{c}{2}n + k(n) - 1}{k(n)} = c\big((1+\rho)\ln(1+\rho) - \rho\ln\rho\big).$$

Inverting the conditions

$$\cos^2\theta_t = 4\frac{\rho_t^{-1}+1}{(\rho_t^{-1}+2)^2}$$

leads to

$$\rho_t = \frac{1-\sin\theta_t}{2\sin\theta_t}$$

Let $k_t = \lfloor \rho_t n \rfloor$, and let y_t satisfy $z_{k_t} \leq y_t \leq z_{k_t+1}$ and $P_{k_t}(y_t) + P_{k_t+1}(y_t) = 0$ (the existence of y_t is guaranteed by the interlacing property of the zeros of the Jacobi polynomials). Then from (29),

$$|C| \leq \frac{4\big(\sum_{t=1}^m a_{k_t}\big)\prod_{t=1}^m D_{k_t}}{m-\sigma(y)}.$$

Since $\sigma(y) \simeq m\cos^2\theta$ and the expression $\frac{4\sum_{t=1}^m a_{k_t}}{m-\sigma(y)}$ has a finite limit when k_t/n tends to ρ_t, the rate $R(C)$ satisfies

$$R(C) \lesssim \frac{1}{n}\sum_{t=1}^m \ln D_{k_t} \simeq \sum_{t=1}^m c\big((1+\rho_t)\ln(1+\rho_t) - \rho_t\ln(\rho_t)\big).$$

□

Remark 4.10:
- It is worth noticing that the choice $\theta_t = \theta$ in (30) yields to the bound

$$R(C) \lesssim cmR_{LP}(\theta). \tag{31}$$

  This bound can be derived more easily, since every code in X^m is also a code in the cmn-th dimensional unit sphere (combining the mapping β for $m=1$ and the obvious mapping $\big(S^{n-1}\big)^m \to S^{mn-1}$). It turns out that, since the function $R_{LP}(\theta)$ *as a function of* $t=\cos^2\theta$ is not convex, the bound (30) slightly improves on (31). We discuss this in more details in the next subsection.
- The same method applied to $X = S^{n-1}$ would lead to:

$$R(C) \lesssim R_{LP}(\theta_1) + \ldots + R_{LP}(\theta_m), \text{ for all } \theta_t \text{ such that } \sum_{t=1}^m \cos\theta_t = m\cos\theta.$$

  But the function $R_{LP}(\theta)$ *as a function of* $t = \cos\theta$ is convex, therefore the choice of $(\theta_1,\ldots,\theta_m)$ that minimizes the right hand side is $\theta_1 = \cdots = \theta_m = \theta$, yielding (31).

□

F. *Analysis of* (30) *versus* (31)

Let C^2 be the set of continuous, twice differentiable functions with continuous second derivative. For a function f defined on $[0,1[$, of class C^2, we denote:

$$f^{(m)}(t) := \min_{\substack{t_1,\ldots,t_m\in[0,1[\\ \sum_{i=1}^m t_i = mt}} \frac{f(t_1)+\cdots+f(t_m)}{m}.$$

Clearly, if f is convex on $[0,1[$, we have $f^{(m)} = f$, and, if $f \leq g$, $f^{(m)} \leq g^{(m)}$. It is also easy to see that $f^{(m')} \leq f^{(m)}$ when m divides m'.

The function we are interested in is $f(t) = R_{LP}(\theta)$ where $t = \cos^2\theta$. We have

$$f(t) = (1+\rho(t))\ln(1+\rho(t)) - \rho(t)\ln(\rho(t))$$

where

$$\rho(t) = \frac{1}{2}\big(-1+(1-t)^{-1/2}\big).$$

One can check that the second derivative of f takes negative values on some interval $[0, t_0]$, $t_0 \simeq 0.208$, and then takes positive values on $[t_0, 1[$. The function f is an increasing function, with $f(0) = 0$, first concave then convex. We consider the function g on $[0, 1[$, whose graph $\mathcal{C}_g$ determines the convex hull of the portion of plane above the graph $\mathcal{C}_f$ of f. The function g is uniquely determined by the conditions:

$$\begin{cases} g \leq f \\ g \text{ is convex} \\ g \text{ is maximal with these properties} \end{cases}$$

Let us denote by t_1 the unique value for which the tangent at $(t_1, f(t_1))$ to $\mathcal{C}_f$ contains the origin $(0, 0)$. The value $t_1 \simeq 0.379$ is the unique solution to

$$f(t) = f'(t)t$$

and the slope of the tangent to $\mathcal{C}_f$ at t_1 equals $f'(t_1) \simeq 1.089$. Then the function g is defined by:

$$\begin{cases} g(t) = f'(t_1)t \simeq 1.089t & \text{for all } t \in [0, t_1] \\ g(t) = f(t) & \text{for all } t \in [t_1, 1[\end{cases}$$

Since g is convex, we have for all m and all $t \in [0, 1[$, $g^{(m)}(t) = g(t) \leq f^{(m)}(t)$. In other words, on $[0, t_1]$, $f^{(m)}$ is somewhere between g and f, and on $[t_1, 1[$, $f^{(m)} = f$. Clearly, when $m \to +\infty$, $f^{(m)} \to g$. Also, the maximum δ of $f(t) - g(t)$ is an upper bound for the maximum of $f(t) - f^{(m)}(t)$. Numerical calculation gives $\delta \simeq 0.016$. Considering our primary goal, i.e., to compare (30) and (31), this means that the improvement of (31) upon (30) is upper-bounded by $0.016m$.

It seems difficult to determine the optimal choice of $(t_1, \dots, t_m)$ that minimizes the quotient $\frac{f(t_1)+\cdots+f(t_m)}{m}$. A natural choice is $(t_1, \dots, t_m) = (0, 0, .., mt/r, \dots, mt/r)$ with r non-zero and equal coordinates. In that case, $\frac{f(t_1)+\cdots+f(t_m)}{m} = \frac{r}{m} f(\frac{mt}{r})$ and requires $t < r/m$. If $t = \frac{rt_1}{m}$, it is certainly the best choice since then the resulting point lies on $\mathcal{C}_g$. Numerical experiments seem to show that, for $m = 2, 3$, and $t < 1/m$, $r = 1$ does minimize the quotient $\frac{f(t_1)+\cdots+f(t_m)}{m}$.

V. Bounds for codes in the Grassmann and Stiefel manifolds

In this section, we summarize the consequences of the above results for Grassmann and Stiefel codes. Following a standard notation in coding theory, we denote by $A(X, d)$, the maximal number of elements of a code C of the space X with minimum distance d.

We have proved in Theorem 3.7 that the size of Grassmannian codes with minimum chordal distance $d = \sqrt{m - s}$ is upper-bounded by the size of codes in $\mathbb{P}^{n-1}(K)^m$ with minimum angular distance θ, where $\cos^2\theta = s/m$. Thus we have proved that:

$$A(\mathcal{G}_{m,n}(K), d) \leq A(\mathbb{P}^{n-1}(K)^m, \theta) \text{ with } \theta = \arccos(\sqrt{1 - d^2/m}). \tag{32}$$

Linear programming bounds on $A(\mathbb{P}^{n-1}(K)^m, \theta)$ were derived in Section IV. Applying them, we obtain the bound $R_2(d)$ from Theorem 2.7. The bounds $R_1(d)$ and $R_2(d)$ are depicted in Figure 2.

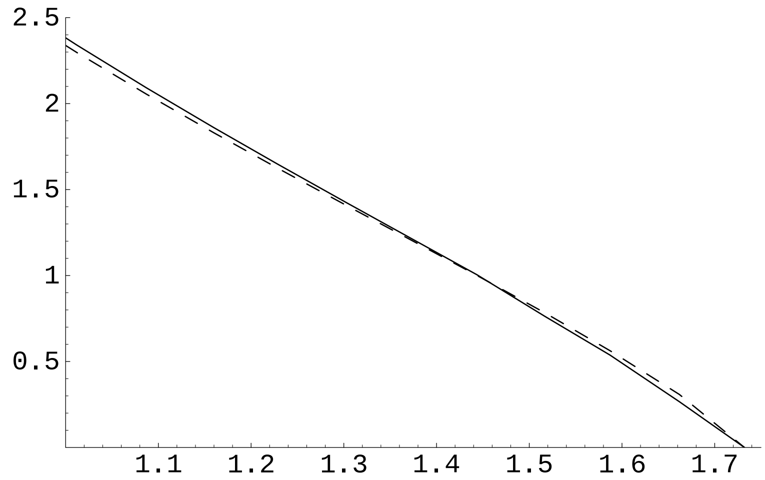

Fig. 2. Upper bounds on the asymptotic rate of real Grassmannian codes with minimum chordal distance d, $m = 3$. The dashed line is $R_1(d)$, and the solid line is $R_2(d)$ (see Theorem 2.7)

We believe that these bounds are not good in general for finite values of the parameters, because we use only a rough estimate of $\sigma(p,q) = \mathrm{trace}(\pi_p \circ \pi_q)$ in the inequality (18) (we replace $\sum_{1\leq i,j\leq m}(e_i \cdot e'_j)^2$ with $\sum_{1\leq i\leq m}(e_i \cdot e'_i)^2$). If we compare the bounds obtained with the zonal polynomials of small degree, (27) is worse than the simplex bound, obtained from the zonal polynomial of degree 1 of $\mathcal{G}_{m,n}(\mathbb{R})$. Moreover, numerical experiments for small parameters m and n (with the package LRS, by David Avis, http://cgm.cs.mcgill.ca/˜avis/C/lrs.html), confirms that the bounds obtained from the zonal polynomials of $\mathcal{G}_{m,n}(\mathbb{R})$ are sharper than the ones obtained from Proposition 4.3 for $X = \mathbb{P}^{n-1}(\mathbb{R})^m$.

Surprisingly, the consideration of $\mathbb{P}^{n-1}(K)^m$ allows us to obtain better bounds for the asymptotic rate than the ones obtained previously by either the isometric embedding given in [21] of $\mathcal{G}_{m,n}$ into a unit sphere of the dimension $n(n+1)/2-2$ (see also Remark 3.4), or the spectral method developed in [6] with the zonal polynomials of $\mathcal{G}_{m,n}$.

We have proved in Theorem 3.6 that the size of Stiefel codes with minimum chordal distance $d = \sqrt{2}\sqrt{m-s}$ is upper-bounded by the size of codes in $\left(S^{cn-1}\right)^m$ with minimum angular distance θ, where $\cos\theta = s/m$. Thus we have proved that:

$$A(\mathcal{V}_{m,n}(K), d) \leq A(\left(S^{cn-1}\right)^m, \theta) \text{ with } \theta = \arccos(1 - d^2/2m). \tag{33}$$

In Section IV, we derived linear programming bounds for codes in $\left(S^{cn-1}\right)^m$, thus implying bounds for Stiefel codes. These bounds are, up to our knowledge, the first general bounds for Stiefel codes, and we believe that they are rather sharp. For the asymptotic rate, the best result is given in Theorem 2.8.

VI. Conclusions

Using relations between Grassmann and Stiefel manifolds and other spaces, we derive new bounds on the size of Grassmannian codes (Theorem 2.7) and Stiefel codes (Theorem 2.8). These are the best known asymptotic bounds on the rate of Grassmannian and Stiefel codes.

Acknowledgement

We thank Jean Creignou for helpful discussions about spherical and Grassmannian codes.

References

[1] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: a geometric approach to the noncoherent multiple-antenna channel," vol. 48, no. 2, pp. 359–383, Feb. 2002.

[2] P. Delsarte, "An algebraic approach to the association schemes of coding theory," *Philips Res. Repts. Suppl.*, vol. 10, pp. 1–97, 1973.

[3] R. J. McEliece, E. R. Rodemich, H. Rumsey, and L. R. Welch, "New upper bounds on the rate of a code via the Delsarte-Macwilliams inequalities," *IEEE Trans. Inform. Theory*, vol. 23, pp. 157–166, Feb. 1977.

[4] G. A. Kabatyanskii and V. I. Levenshtein, "Bounds for packings on a sphere and in space," *Problemy Peredachi Informatsii*, vol. 14, no. 1, pp. 3–25, 1978.

[5] V. I. Levenshtein, "Universal bounds for codes and designs," in *Handbook of Coding Theory*, V. Pless and W. C. Huffmann, Eds. Amsterdam: Elsevier, 1998, pp. 499–648.

[6] C. Bachoc, "Linear programming bounds for codes in Grassmannian spaces," *IEEE Trans. Inform. Theory*, vol. 52, no. 5, pp. 2111–2125, May 2006.

[7] A. Barg and D. Nogin, "A bound on Grassmannian codes," *J. of Comb. Theory Ser. A*, vol. 113, no. 8, pp. 1629–1635, Nov. 2006.

[8] M. J. Aaltonen, "A new upper bound on non binary block codes," *Discrete Math.*, vol. 83, pp. 139–160, 1990.

[9] H. Tarnanen, M. J. Aaltonen, and J. M. Goethals, "On the nonbinary Johnson scheme," *Europ. J. Combin.*, vol. 6, pp. 279–285, 1985.

[10] Y. Ben-Haim and S. Litsyn, "A new upper bound on the rate of non-binary codes," *Advances in Math. of Comm.*, vol 1, no. 1, pp. 83–92, Feb. 2007.

[11] S. Helgason, *Differential Geometry, Lie Groups and Symmetric Spaces*. NY: Ac. Press, 1978.

[12] J. H. Conway and D. A. Smith, *On Quaternion and Octanions, Their Geometry, Arithmetic, and Symmetry*. A K Peters, 2003.

[13] T. Ericson and V. Zinoviev, *Codes on Euclidean Spheres*. Amsterdam, Holland: Norh-Holland.

[14] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*. New York: Springer-Verlag, 1988.

[15] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell Syst. Tech. J.*, vol. 38, no. 3, pp. 611–656, 1959.

[16] C. Chabauty, "Résultat sur l'empilement de calottes égales sur une périsphère de $\mathbb{R}^n$ et correction à un travail antérieur," *C.R.A.S.*, vol. 236, pp. 1462–1464, 1953.

[17] A. D. Wyner, "Capabilities of bounded discrepancy decoding," Bell Sys. Tech. J., vol. 43, pp. 1061–1122, July-August 1965.

[18] R. A. Rankin, "The closest packing of spherical caps in n dimensions," in *Proc. Glasgow Math. Assoc.*, vol. 2, 1955, pp. 139–144.

[19] A. Barg and D. Nogin, "Bounds on packings of spheres in the Grassmann manifold," *IEEE Trans. Inform. Theory*, vol. 48, no. 9, pp. 2450–2454, Sept. 2002.

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

[21] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc., packings in grassmannian spaces," *Experimental Mathematics*, vol. 5, no. 2, pp. 139–159, 1996.

[22] O. Henkel, "Sphere-packing bounds in the Grassmann and Stiefel manifolds," *IEEE Trans. Inform. Theory*, vol. 51, no. 10, pp. 3445–3456, Oct. 2005.

[23] G. Szegö, *Orthogonal Polynomials*. AMS Colloquim Publications, 1939, vol. 23.

[24] S. G. Hoggar, "t-designs with general angle set," *Europ. J. Combinatorics*, vol. 13, pp. 257–271, 1992.