

---

This is an electronic reprint of the original article.  
This reprint may differ from the original in pagination and typographic detail.

Pitaval, Renaud-Alexandre; Wei, Lu; Tirkkonen, Olav; Hollanti, Camilla

## Density of Spherically-Embedded Stiefel and Grassmann Codes

*Published in:*  
IEEE Transactions on Information Theory

*DOI:*  
[10.1109/TIT.2017.2759194](https://doi.org/10.1109/TIT.2017.2759194)

Published: 01/01/2018

*Document Version*  
Peer-reviewed accepted author manuscript, also known as Final accepted manuscript or Post-print

*Please cite the original version:*  
Pitaval, R.-A., Wei, L., Tirkkonen, O., & Hollanti, C. (2018). Density of Spherically-Embedded Stiefel and Grassmann Codes. *IEEE Transactions on Information Theory*, 64(1), 225 - 248.  
<https://doi.org/10.1109/TIT.2017.2759194>

# Density of Spherically-Embedded Stiefel and Grassmann Codes

Renaud-Alexandre Pitaval, Lu Wei, Olav Tirkkonen, and Camilla Hollanti

**Abstract**—The density of a code is the fraction of the coding space covered by packing balls centered around the codewords. A high density indicates that a code performs well when used as a uniform point-wise discretization of an ambient space. This paper investigates the density of codes in the complex Stiefel and Grassmann manifolds equipped with the chordal distance arising from an Euclidean embedding, including the unitary group as a special case. The choice of distance enables the treatment of the manifolds as subspaces of Euclidean hyperspheres. In this geometry, the densest packings are not necessarily equivalent to maximum-minimum-distance codes. Computing a code's density follows from computing i) the normalized volume of a metric ball and ii) the kissing radius, the radius of the largest balls one can pack around the codewords without overlapping. First, the normalized volume of a metric ball is evaluated by asymptotic approximations. The volume of a small ball can be well-approximated by the volume of a locally-equivalent tangential ball. In order to properly normalize this approximation, the precise volumes of the manifolds induced by their spherical embedding are computed. For larger balls, a hyperspherical cap approximation is used, which is justified by a volume comparison theorem showing that the normalized volume of a ball in the Stiefel or Grassmann manifold is asymptotically equal to the normalized volume of a ball in its embedding sphere as the dimension grows to infinity. Then, bounds on the kissing radius are derived alongside corresponding bounds on the density. Unlike spherical codes or codes in flat spaces, the kissing radius of Grassmann or Stiefel codes cannot be exactly determined from its minimum distance. It is nonetheless possible to derive bounds on density as functions of the minimum distance. Stiefel and Grassmann codes have larger density than their image spherical codes when dimensions tend to infinity. Finally, the bounds on density lead to refinements of the standard Hamming bounds for Stiefel and Grassmann codes.

**Index Terms**—Unitary matrix codes, spherical codes, Grassmann manifold, Stiefel manifold, density, packing, metric ball, volume, Hamming bound.

Manuscript received July 11, 2016; revised April 6, 2017; accepted July 22, 2017. This work was supported in part by the Academy of Finland (grants #276031 and #299916) and by the Nokia Foundation. Part of this work was presented at the 2011 IEEE International Symposium on Information Theory and the 2012 Asilomar Conference of the Signals, Systems and Computers.

R.-A. Pitaval was with the Department of Communications and Networking, Aalto University, FI-00076 Espoo, Finland. He is now with Huawei Technologies Sweden AB, SE-164 94 Kista, Sweden (e-mail: renaud-alexandre.pitaval@alumni.aalto.fi).

L. Wei is with the Department of Electrical and Computer Engineering, University of Michigan-Dearborn, Michigan 48128, USA (e-mail: luwe@umich.edu).

O. Tirkkonen is with the Department of Communications and Networking, Aalto University, FI-00076 Espoo, Finland (email: olav.tirkkonen@aalto.fi).

C. Hollanti is with the Department of Mathematics and Systems Analysis, P.O. BOX 11100, FI-00076 AALTO, Finland (email: camilla.hollanti@aalto.fi).

Communicated by S. Mesnager, Associate Editor for Sequences.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier

## I. INTRODUCTION

Sphere packing is a classical problem with a long history from geometry to information theory [1]–[4]. In his 1948 seminal work [5], Shannon made the connection between the capacity of an additive white Gaussian noise channel and packing of multidimensional spheres. This interpretation was later generalized to non-coherent multi-antenna channels and packings in products of Grassmann manifolds [6].

Stiefel and Grassmann codes are matrix codes with applications to Multiple Input Multiple Output (MIMO) communications [6]–[9], Code Division Multiple Access (CDMA) wireless systems [10], and compressive sensing [11], [12]. The complex Stiefel manifold is the space of rectangular semi-unitary matrices. The Grassmann manifold is the space of eigenspaces spanned by the Stiefel matrices. An element in the Grassmann manifold is an equivalence class of Stiefel matrices, which can be alternatively represented by a unique projection matrix. When discussing codes in Grassmann manifolds of one-dimensional subspaces, the equivalent language of frame theory is often used [12].

Depending of the application and convenience, several non-equivalent distances can be defined on these spaces [13], [14]. A common distance for the unitary group and Stiefel manifold is simply the Frobenius norm of the difference of two matrices [15]. Similarly for the Grassman manifold, an often-used distance arises from the Frobenius norm of the difference between two projection matrices [8], [16]. These distances correspond to embed the manifolds into Euclidean spaces. Through these embeddings, each manifold is a subset of an hypersphere and the distance is the length of a chord.

For the application of limited-feedback MIMO precoding, unitary precoder designs have been reduced to quantization problems on Grassmann and Stiefel manifolds with the associated chordal distance [8], [9], [17]. In the context of space-time coding, the code performance in the low signal-to-noise-ratio (SNR) regime has also been related to the chordal distance [16], [18]. More recently, analogical results have been derived for linear classification: In the small model mismatch regime, the number of classes that can be distinguished by a subspace classifier is governed by the chordal distance among subspaces, and more generally by their principal angles [19].

In the last decade, basic coding-theoretic results estimating the relationship between the cardinality and the minimum distance of codes in Grassmann and Stiefel manifolds have been published [15], [17], [20]–[27]. The standard Hamming bound relates the minimum distance to the notion of code density. The density of a code is the maximum portion of the coding space covered by non-intersecting balls of equal radius.

In classical geometry, maximizing the code's density is known to be equivalent to maximizing its minimum distance. For the Stiefel and Grassmann codes measured by chordal distance, this equivalence does not always hold.

The present paper investigates this surprising fact and discusses the density of codes in Grassmann and Stiefel manifolds equipped with their chordal distance. We provide a direct connection between minimum distance and density via lower and upper bounds on the density for a given distance. This connection leads to new conjectured bounds on the minimum distance as a function of code cardinality. Related to applications, numerical evaluations suggest that the density may be a more relevant criterion for MIMO precoding than the minimum distance. While for space-time coding, the minimum distance of a code may be more important than its density.

There are two main difficulties in evaluating the density of codes in these spaces: 1) evaluating the normalized volume of a ball, and 2) estimating the kissing radius of codes.

The problem of estimating the volume of a metric ball has been addressed in [17], [20], [22], [23], [26]–[29]. In the large code cardinality regime, balls are small and can be approximated to be balls in flat space. While some exact evaluations were obtained in Grassmannian cases [17], [28], the case of the Stiefel manifold has been less addressed. A powerful and general framework is provided in [22], [26], [27]. However, it appears that the state-of-the-art volumes in the literature do not correspond to the desired metrics. Indeed, the volume element is unique up to a non-vanishing scaling factor which is often dismissed, as it can be absorbed in the overall normalization. From Nash embedding theorem [30], every Riemannian metric can be seen as induced by an appropriate Euclidean embedding. There exists an intrinsic metric locally equivalent to the chordal distance, and thus a consistent volume density that defines the notion of volume on the manifold. In this respect, we provide the exact scaling of the volume for Grassmann and Stiefel manifolds induced by the spherical embedding, leading in turn to precise small ball approximations.

For larger radius, we show that the volume of a ball in the manifold can be well approximated by the “area” of the hyperspherical cap the ball is embedded into. The approximation is supported by a volume comparison theorem showing that the normalized volume of a ball in the manifold is asymptotically equal to the normalized volume of a ball in the embedding sphere. This result generalizes and provides a structural unification of our previous results in [31]–[33]. The derivation is a by-product of the asymptotic Gaussianity of the chordal distance arising through its reduction to linear statistics. The intuition behind relates to a classical result by Borel [34] who proved that coordinates of a hypersphere are asymptotically Gaussian as the dimension tends to infinity [35], and the long history of central limit theorems showing the Gaussian behavior of linear statistics [36] in random matrix theory.

Next, the paper addresses the evaluation of the kissing radius and density of Grassmann and Stiefel codes. The kissing radius is the analog of the packing radius for linear codes [37], which has applications in, e.g., sphere-decoder optimization [38],

[39]. It is the largest possible radius of packing balls around the codewords of a code. The kissing radius also relates to rate-distortion theory as it is the smallest possible distance from a codeword to the border of its Voronoi cell. The problem reduces to finding the minimum mid-distance between two points at a given distance  $\delta$ . For a geodesic distance on a flat space, the answer is simply  $\delta/2$ , the so-called packing radius. With a strictly extrinsic distance, the triangle inequality is never satisfied with equality and the kissing radius is greater than  $\delta/2$ . While for spherical codes there is a one-to-one mapping between the kissing radius and the minimum distance, this is not always the case for the Grassmann and Stiefel codes with chordal distance. As a consequence, the density is not a single-variable function of the minimum distance of the code, and two codes with equal minimum distance could have different densities. This is in sharp contrast with classical packing problems where maximizing the density is equivalent to maximizing the minimum distance. The kissing radius and the density cannot be determined solely from the minimum distance but it is possible to derive bounds. Combining these bounds with the volume comparison theorem discussed above shows that the densities of Grassmann and Stiefel codes are asymptotically greater than or equal to the densities of their image spherical codes. The bounds are shown to be tight by simulations.

Finally, a direct application of bounds on density is to revisit the Hamming bounds. The results of this paper improve the Hamming bound for the Grassmann case in [20], [22]; and generalize the bound for the unitary group in [23] to the Stiefel manifold.

The rest of this paper is organized as follows. Section II introduces the considered spaces and their geometry. Section III states the problem and necessary definitions. Section IV addresses the problem of the volume of a ball. In Section V, bounds on kissing radius and density are derived. Section VI provides concluding remarks.

## II. THE GRASSMANN AND STIEFEL MANIFOLDS

We consider the following Riemann manifolds equipped with a chordal distance induced by their canonical spherical embedding. Throughout the paper,  $\mathcal{M}$  stands for the unitary group  $\mathcal{U}_n$ , Stiefel manifold  $\mathcal{V}_{n,p}^{\mathbb{C}}$ , or Grassmann manifold  $\mathcal{G}_{n,p}^{\mathbb{C}}$  embedded into the sphere  $\mathcal{S}^{D-1}(R)$  of radius  $R = \sqrt{n}$ ,  $\sqrt{p}$  or  $\sqrt{\frac{p(n-p)}{2n}}$ , in a Euclidean space of dimension  $D = 2n^2$ ,  $2np$  or  $n^2 - 1$ , respectively. Elements in the manifold are represented by matrices in  $\mathbb{C}^{n \times n}$  or  $\mathbb{C}^{n \times p}$  endowed by the inner product  $\langle \cdot, \cdot \rangle = \Re \text{Tr} \cdot^H \cdot$ , where  $\Re$  is the real part and  $\text{Tr}$  is the matrix trace. The geometric description below follows directly by generalizing [13] and [14] to the complex space. The chordal distances, the dimension  $\dim$  of the manifolds, and the corresponding embeddings are summarized in Table I.

### A. Hypersphere

The Euclidean  $(D-1)$ -sphere of radius  $R$  in  $\mathbb{R}^D$  is defined as

$$\mathcal{S}^{D-1}(R) = \{\mathbf{x} \in \mathbb{R}^D \mid \|\mathbf{x}\|_2 = R\}. \quad (1)$$

TABLE I  
MANIFOLDS OF DIMENSION  $\dim$  WITH THEIR SPHERICAL EMBEDDINGS:  
( $\mathcal{M}, d_c$ )  $\hookrightarrow \mathcal{S}^{D-1}(R)$ .

$\mathcal{M}$	$\dim$	$d_c(\mathbf{U}, \mathbf{V})$	$D$	$R$
$\mathcal{U}_n$	$n^2$	$\ \mathbf{U} - \mathbf{V}\ _F$	$2n^2$	$\sqrt{n}$
$\mathcal{V}_{n,p}^{\mathbb{C}}$	$2np - p^2$	$\ \mathbf{U} - \mathbf{V}\ _F$	$2np$	$\frac{\sqrt{p}}{\sqrt{2n}}$
$\mathcal{G}_{n,p}^{\mathbb{C}}$	$2np - 2p^2$	$\frac{1}{\sqrt{2}}\ \mathbf{U}\mathbf{U}^H - \mathbf{V}\mathbf{V}^H\ _F$	$n^2 - 1$	$\sqrt{\frac{p(n-p)}{2n}}$

For  $R = 1$ , we simply write  $\mathcal{S}^{D-1}$ . The *chordal distance* is the natural Euclidean distance applied to elements on the sphere. Given  $\mathbf{x}, \mathbf{y} \in \mathcal{S}^{D-1}(r)$  it is simply the norm of the difference:

$$d_c(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2. \quad (2)$$

It is an extrinsic distance as it measures the length of a chord outside of the manifold itself, which here is the surface of the sphere.

### B. Unitary Group

The *unitary group* is the set of unitary matrices,

$$\mathcal{U}_n = \{\mathbf{U} \in \mathbb{C}^{n \times n} \mid \mathbf{U}^H \mathbf{U} = \mathbf{U} \mathbf{U}^H = \mathbf{I}_n\}, \quad (3)$$

where  $\cdot^H$  denotes the Hermitian conjugate. This compact Lie group is a manifold of dimension  $\dim \mathcal{U}_n = n^2$ . By differentiating the unitary constraint, one can verify that the tangent space  $\mathcal{T}_{\mathbf{U}} \mathcal{U}_n$  at  $\mathbf{U}$  is the set of matrices  $\Delta \in \mathbb{C}^{n \times n}$  such that  $\mathbf{U}^H \Delta$  is skew-Hermitian. Specifically, at identity, the tangent space  $\mathcal{T}_{\mathbf{I}} \mathcal{U}_n$  is the Lie algebra of skew-Hermitian  $n \times n$  matrices  $\mathfrak{u}(n)$ .

One can parametrize  $\mathcal{U}_n$  with reference to a fixed  $\mathbf{U} \in \mathcal{U}_n$  via skew-Hermitian matrices as  $\mathbf{V} = \mathbf{U} \exp(\mathbf{U}^H \Delta) \in \mathcal{U}_n$  with  $\mathbf{U}^H \Delta \in \mathfrak{u}(n)$ . For a fixed  $\Delta$ , the exponential map  $\exp(\cdot)$  defines the geodesic between  $\mathbf{U}$  and  $\mathbf{V}$  by mapping the tangent space to the manifold as  $\mathbf{V}(t) = \mathbf{U} \exp(t \mathbf{U}^H \Delta)$  with  $0 \leq t \leq 1$ . A Riemannian metric may be defined from the canonical embedding of  $\mathcal{U}_n$  in the Euclidean space  $(\mathbb{C}^{n \times n}, \langle \cdot, \cdot \rangle)$ , then  $\exp(\cdot)$  is the matrix exponential.

Consider the eigenvalue decomposition  $\mathbf{U}^H \mathbf{V} = \Omega \text{diag}(e^{i\theta_1}, e^{i\theta_2}, \dots, e^{i\theta_n}) \Omega^H$  where the diagonal elements of  $\Omega \in \mathcal{U}_n$  are non-negative and real. This decomposition is unique if these *principal angles* can be strictly ordered  $\pi \geq \theta_n > \dots > \theta_2 > \theta_1 \geq -\pi$ . This leads to the corresponding eigenvalue decomposition  $\mathbf{U}^H \Delta = \Omega \text{diag}(\theta_1, \theta_2, \dots, \theta_n) \Omega^H$ , and accordingly the geodesic between  $\mathbf{U}$  and  $\mathbf{V}$  becomes

$$\mathbf{V}(t) = \mathbf{U} \Omega \text{diag}(e^{i\theta_1 t}, e^{i\theta_2 t}, \dots, e^{i\theta_n t}) \Omega^H. \quad (4)$$

The geodesic distance is the intrinsic distance between two points obtained by integrating the geodesic path along the manifold. It is given by the norm of the tangent direction according to the considered Riemannian metric

$$d_g(\mathbf{U}, \mathbf{V}) = \|\Delta\|_F = \|\mathbf{U}^H \Delta\|_F = \sqrt{\sum_{i=1}^n \theta_i^2}. \quad (5)$$

Alternatively, the canonical distance in the ambient matrix-space is

$$\begin{aligned} d_c(\mathbf{U}, \mathbf{V}) &= \|\mathbf{U} - \mathbf{V}\|_F = \sqrt{2n - 2\Re \text{Tr}(\mathbf{U}^H \mathbf{V})} \\ &= \sqrt{4 \sum_{i=1}^n \sin^2 \frac{\theta_i}{2}} \end{aligned} \quad (6)$$

where the last equality follows from the decomposition (4).

By observing that  $d_c(\mathbf{0}, \mathbf{V}) = \|\mathbf{V}\|_F = \sqrt{n}$  for all  $\mathbf{V} \in \mathcal{U}_n$ , one verifies that the unitary group equipped with  $d_c$  has an isometric embedding into the sphere  $\mathcal{S}^{2n^2-1}(\sqrt{n})$ . The concrete embedding  $(\mathcal{U}_n, d_c) \hookrightarrow \mathcal{S}^{2n^2-1}(\sqrt{n})$  is obtained from the classical mapping  $\mathbb{C}^{n \times n} \hookrightarrow \mathbb{R}^{2n^2}$  by vectorizing a complex matrix into a real vector.

### C. Stiefel Manifold

The complex *Stiefel manifold* is the space of rectangular semi-unitary matrices,

$$\mathcal{V}_{n,p}^{\mathbb{C}} = \{\mathbf{Y} \in \mathbb{C}^{n \times p} \mid \mathbf{Y}^H \mathbf{Y} = \mathbf{I}_p\}. \quad (7)$$

This provides a generalization of both the hypersphere and the unitary group. For  $p = n$ , one directly recovers  $\mathcal{V}_{n,n}^{\mathbb{C}} = \mathcal{U}_n$ , while for  $p = 1$  and by identification of  $\mathbb{C}^n$  with  $\mathbb{R}^{2n}$ , this corresponds to the unit sphere  $\mathcal{V}_{n,1}^{\mathbb{C}} = \mathcal{S}^{2n-1}$ .

As for the unitary group, the tangent space at  $\mathbf{Y}$  is the set  $\mathcal{T}_{\mathbf{Y}} \mathcal{V}_{n,p}^{\mathbb{C}}$  of matrices  $\Delta \in \mathbb{C}^{n \times p}$  such that  $\mathbf{Y}^H \Delta \in \mathfrak{u}(p)$  is skew-Hermitian. In general, tangents have the form  $\Delta = \mathbf{Y} \mathbf{A} + \mathbf{Y}_{\perp} \mathbf{B}$  where  $\mathbf{A}$  is  $p \times p$  skew-Hermitian,  $\mathbf{Y}_{\perp}$  is the orthogonal complement of  $\mathbf{Y}$  such that the concatenation  $(\mathbf{Y} \mathbf{Y}_{\perp}) \in \mathcal{U}_n$  is unitary, and  $\mathbf{B}$  is an  $(n-p) \times p$  arbitrary complex matrix. Specifically at “identity”  $\mathbf{I}_{n,p} \triangleq \begin{pmatrix} \mathbf{I}_p \\ \mathbf{0} \end{pmatrix}$ , tangents are of the form

$$\Delta = \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} \in \mathbb{C}^{n \times p} \text{ with } \mathbf{A} \in \mathfrak{u}(p). \quad (8)$$

By counting the degrees of freedom in the tangent space, one finds that the Stiefel manifold is a space with dimension  $\dim \mathcal{V}_{n,p}^{\mathbb{C}} = 2np - p^2$ .

Given a starting point  $\mathbf{Y} = \mathbf{Y}(0)$  and a fixed tangent direction  $\Delta = \mathbf{Y} \mathbf{A} + \mathbf{Y}_{\perp} \mathbf{B} \in \mathcal{T}_{\mathbf{Y}} \mathcal{V}_{n,p}^{\mathbb{C}}$ , the canonical embedding of  $\mathcal{V}_{n,p}^{\mathbb{C}}$  in the Euclidean space  $(\mathbb{C}^{n \times p}, \langle \cdot, \cdot \rangle)$  leads to the following geodesic equation [14],

$$\mathbf{Y}(t) = (\mathbf{Y} \quad \Delta) \exp t \begin{pmatrix} \mathbf{Y}^H \Delta & -\Delta^H \Delta \\ \mathbf{I}_p & \mathbf{Y}^H \Delta \end{pmatrix} \mathbf{I}_{2p,p} e^{-t \mathbf{Y}^H \Delta}. \quad (9)$$

The geodesic distance between  $\mathbf{Y}$  and  $\mathbf{Z} = \mathbf{Y}(1)$  is then

$$d_g(\mathbf{Y}, \mathbf{Z}) = \|\Delta\|_F = (\|\mathbf{A}\|_F^2 + \|\mathbf{B}\|_F^2)^{1/2}. \quad (10)$$

**Remark 1.** To the best of our knowledge, when  $p \neq n$ , unlike for the unitary group, it is not known how to compute the geodesic  $\mathbf{Y}(t)$  in closed-form, given only  $\mathbf{Y} = \mathbf{Y}(0)$  and  $\mathbf{Z} = \mathbf{Y}(1)$ .

The corresponding Euclidean/chordal distance from the ambient space is

$$d_c(\mathbf{Y}, \mathbf{Z}) = \|\mathbf{Y} - \mathbf{Z}\|_F. \quad (11)$$



Similarly as for the unitary group, one can verify that  $d_c(\mathbf{0}, \mathbf{Z}) = \sqrt{p}$  for all  $\mathbf{Z}$ , and thus this gives an isometric embedding of the Stiefel manifold  $\mathcal{V}_{n,p}^{\mathbb{C}}$  into  $\mathcal{S}^{2np-1}(\sqrt{p})$ .

As an alternative to (7), the Stiefel manifold can be treated as the quotient space  $\mathcal{V}_{n,p}^{\mathbb{C}} \cong \mathcal{U}_n / \mathcal{U}_{n-p}$ , where a point in  $\mathcal{V}_{n,p}^{\mathbb{C}}$  is an equivalence class of unitary matrices  $[\mathbf{U}] = \left\{ \mathbf{U} \begin{pmatrix} \mathbf{I}_p & \mathbf{0} \\ \mathbf{0} & \mathbf{\Omega} \end{pmatrix} \mid \mathbf{\Omega} \in \mathcal{U}_{n-p} \right\}$ . A natural geometry in this interpretation of the Stiefel manifold is the one inherited from the geometry of the unitary group  $\mathcal{U}_n$  embedded in  $(\mathbb{C}^{n \times n}, \langle \cdot, \cdot \rangle)$ . One can split the tangent space of the unitary group at  $\mathbf{U}$  between the so-called vertical and horizontal spaces:  $\mathcal{T}_{\mathbf{U}}\mathcal{U}_n = \mathcal{H}_{\mathbf{U}} \oplus \mathcal{V}_{\mathbf{U}}$ . The vertical space  $\mathcal{V}_{\mathbf{U}}$  is the tangent space of  $[\mathbf{U}] \hookrightarrow \mathcal{U}_n$  at  $\mathbf{U}$  corresponding to “movements” inside the equivalence class. The horizontal space  $\mathcal{H}_{\mathbf{U}}$  is the orthogonal complement of the vertical space, providing a unique representation for tangents to the quotient space. For  $\mathcal{V}_{n,p}^{\mathbb{C}} \cong \mathcal{U}_n / \mathcal{U}_{n-p}$  at  $\mathbf{U} \in \mathcal{U}_n$ , it is the collection of matrices  $\mathbf{\Delta}_* = \mathbf{U} \begin{pmatrix} \mathbf{A} & -\mathbf{B}^H \\ \mathbf{B} & \mathbf{0} \end{pmatrix}$ , where  $\mathbf{A} \in \mathfrak{u}(n)$  and  $\mathbf{B} \in \mathbb{C}^{(n-p) \times p}$ . With this non-equivalent geometry, the geodesic distance between  $\mathbf{U}$  and  $\mathbf{V} = \mathbf{U} \exp(\mathbf{U}^H \mathbf{\Delta}_*)$  is  $d_{g*}(\mathbf{U}, \mathbf{V}) = \|\mathbf{\Delta}_*\|_F = (\|\mathbf{A}\|_F^2 + 2\|\mathbf{B}\|_F^2)^{1/2}$ . The metric induced by this embedding, and the resulting volume is discussed in [27]. This metric is not further considered in this paper.

#### D. Grassmann Manifold

The complex Grassmann manifold  $\mathcal{G}_{n,p}^{\mathbb{C}}$  is the quotient space of  $\mathcal{V}_{n,p}^{\mathbb{C}}$  over  $\mathcal{U}_p$ :

$$\mathcal{G}_{n,p}^{\mathbb{C}} \cong \mathcal{V}_{n,p}^{\mathbb{C}} / \mathcal{U}_p \subset \mathbb{C}^{n \times p}. \quad (12)$$

Elements in  $\mathcal{G}_{n,p}^{\mathbb{C}}$  are equivalence classes of rectangular semi-unitary matrices  $\mathbf{Y} \in \mathcal{V}_{n,p}^{\mathbb{C}}$ :

$$[\mathbf{Y}] = \{\mathbf{Y}\mathbf{Q} \mid \mathbf{Q} \in \mathcal{U}_p\}. \quad (13)$$

We can identify the equivalence class with a unique matrix representation, which is desirable for practical implementation and computation. By symmetry (c.f. the discussion in [33]) and for simplicity, we will assume  $p \leq n/2$  for the Grassmann manifold, all along this paper.

As described above, tangent vectors at  $\mathbf{Y} \in \mathcal{V}_{n,p}^{\mathbb{C}}$  take the form  $\{\mathbf{Y}\mathbf{A} + \mathbf{Y}_{\perp}\mathbf{B}\}$ . This can be again split between the vertical space  $\mathcal{V}_{\mathbf{Y}} = \{\mathbf{Y}\mathbf{A} \mid \mathbf{A} \in \mathfrak{u}_p\}$  and the horizontal space  $\mathcal{H}_{\mathbf{Y}} = \{\mathbf{Y}_{\perp}\mathbf{B} \mid \mathbf{B} \in \mathbb{C}^{n \times (n-p)}\}$ . At  $\mathbf{I}_{n,p}$ , horizontal tangents are thus of the form  $\mathbf{\Delta} = \begin{pmatrix} \mathbf{0} \\ \mathbf{B} \end{pmatrix}$ , and we have  $\dim \mathcal{G}_{n,p}^{\mathbb{C}} = 2p(n-p)$ . Consider the compact singular value decomposition  $\mathbf{B} = \mathbf{L}\mathbf{\Theta}\mathbf{R}^H$ , where  $\mathbf{L} \in \mathcal{V}_{n-p,p}^{\mathbb{C}}$ ,  $\mathbf{R} \in \mathcal{U}_p$  and  $\mathbf{\Theta} = \text{diag}(\theta_1, \dots, \theta_p)$ . By setting  $\mathbf{A} = \mathbf{0}$ , the geodesic equation (9) for the Grassmann manifold reduces to [14]

$$\mathbf{Y}(t) = \mathbf{Y}\mathbf{R}\cos(\mathbf{\Theta}t)\mathbf{R}^H + \mathbf{Y}_{\perp}\mathbf{L}\sin(\mathbf{\Theta}t)\mathbf{R}^H. \quad (14)$$

The geodesic equation could be rotated from the right by any unitary matrix as it will still belong to the same equivalence class. Without loss of generality one can thus restrict the range of the singular values of  $\mathbf{B}$  to  $0 \leq \theta_i \leq \frac{\pi}{2}$  for all  $i$ . These

values are known as the *principal angles* between the planes  $[\mathbf{Y}(0)]$  and  $[\mathbf{Y}(1)]$ . Contrary to the Stiefel manifold, here given two end points  $\mathbf{Y} = \mathbf{Y}(0)$  and  $\mathbf{Z}$ , one can compute the tangent in the geodesic (14) and satisfy  $[\mathbf{Z}] = [\mathbf{Y}(1)]$  by singular value decomposition of  $\mathbf{Z}^H\mathbf{Y}$  and  $\mathbf{Z}^H\mathbf{Y}_{\perp}$ , where the singular values of  $\mathbf{Z}^H\mathbf{Y}$  are  $\cos \theta_1, \dots, \cos \theta_p$ .

The geodesic distance is given by

$$d_g([\mathbf{Y}], [\mathbf{Z}]) = \|\mathbf{\Delta}\|_F = \|\mathbf{B}\|_F = \sqrt{\sum_{i=1}^p \theta_i^2}. \quad (15)$$

In [13], the following distance was defined

$$d_c([\mathbf{Y}], [\mathbf{Z}]) = \sqrt{\sum_{i=1}^p \sin^2 \theta_i}, \quad (16)$$

which is locally-equivalent to (15) as  $\sin^2 \theta_i \approx \theta_i^2$  when  $\theta_i \approx 0$ , and corresponds to a Euclidean embedding since

$$\begin{aligned} d_c([\mathbf{Y}], [\mathbf{Z}]) &= \sqrt{p - \sum_{i=1}^p \cos^2 \theta_i} = \sqrt{p - \|\mathbf{Z}^H\mathbf{Y}\|_F^2} \\ &= \frac{1}{\sqrt{2}} \|\mathbf{Y}\mathbf{Y}^H - \mathbf{Z}\mathbf{Z}^H\|_F. \end{aligned} \quad (17)$$

This is in fact an isometric embedding into  $\mathcal{S}^{n^2-2}(\sqrt{\frac{p(n-p)}{2n}})$  [13] which follows from mapping any  $[\mathbf{Y}] \in \mathcal{G}_{n,p}^{\mathbb{C}}$  to the space of detraced Hermitian matrices as

$$[\mathbf{Y}] \rightarrow \bar{\Pi}_{\mathbf{Y}} = \frac{1}{\sqrt{2}} \left( \mathbf{Y}\mathbf{Y}^H - \frac{p}{n} \mathbf{I} \right), \quad (18)$$

and the chordal distance is the ambient Euclidean distance

$$d_c([\mathbf{Y}], [\mathbf{Z}]) = \|\bar{\Pi}_{\mathbf{Y}} - \bar{\Pi}_{\mathbf{Z}}\|_F. \quad (19)$$

In this ambient space of dimension  $n^2 - 1$ , it can be further verified that the distance of any  $\bar{\Pi}_{\mathbf{Y}}$  from the origin is  $\|\bar{\Pi}_{\mathbf{Y}} - \mathbf{0}\|_F^2 = \frac{p(n-p)}{2n}$ , and thus  $\bar{\Pi}_{\mathbf{Y}}$  belongs to a sphere of radius  $\sqrt{\frac{p(n-p)}{2n}}$ . When there is no ambiguity on the considered space, we will simply write  $d_c(\mathbf{Y}, \mathbf{Z})$ , which is well-defined as the distance does not depend on the Stiefel representatives.

Alternatively, the Grassmann manifold can be expressed as the quotient space  $\mathcal{G}_{n,p}^{\mathbb{C}} \cong \mathcal{U}_n / (\mathcal{U}_p \times \mathcal{U}_{n-p}) \subset \mathbb{C}^{n \times n}$ . In this representation, elements in a Grassmann manifold are equivalence classes of unitary matrices. The tangents of  $\mathcal{G}_{n,p}^{\mathbb{C}}$  at the identity are of the form  $\mathbf{\Delta}_* = \begin{pmatrix} \mathbf{0} & -\mathbf{B}^H \\ \mathbf{B} & \mathbf{0} \end{pmatrix}$  for this second quotient representation. Again, we do not consider this representation, but we note here that the natural geodesic distance induced by this embedding only differs from (15) by a scaling factor,

$$\begin{aligned} d_{g*}([\mathbf{Y}], [\mathbf{Z}]) &= \|\mathbf{\Delta}_*\|_F = \sqrt{2}\|\mathbf{B}\|_F \\ &= \sqrt{2}d_g([\mathbf{Y}], [\mathbf{Z}]). \end{aligned} \quad (20)$$

### III. PACKING PROBLEM AND MAXIMUM CODE DENSITY

A packing is a maximal set of non-intersecting balls of fixed radius, covering the space so that it is not possible to fit in another ball. For a given code size, a packing thus gives the maximum density, i.e. the maximum fraction of the space that one can cover by non-intersecting balls. This problem is considered to be the dual of a coding problem: maximizing the code cardinality for a given minimum distance or reciprocally maximizing the minimum distance for a given code cardinality. Surprisingly, these two problems are not necessarily equivalent for the Grassmann and Stiefel manifolds with chordal distance. This is because there is not necessarily a one-to-one mapping between the minimum distance and the kissing radius of the code, which follows from the choice of an extrinsic distance combined with the two-point *inhomogeneity* of the spaces [15], [40].

#### A. Code, Minimum Distance, and Metric Balls

An  $(N, \delta)$ -code is a finite subset of  $N$  elements in  $\mathcal{M}$ ,

$$\mathcal{C} = \{\mathbf{C}_1, \dots, \mathbf{C}_N\} \subset \mathcal{M}, \quad (21)$$

where  $\delta$  is the minimum distance defined as

$$\delta = \min\{d_c(\mathbf{C}_i, \mathbf{C}_j) ; \mathbf{C}_i, \mathbf{C}_j \in \mathcal{C}, i \neq j\}. \quad (22)$$

A metric ball  $B_{\mathcal{C}}(\gamma) \subset \mathcal{M}$  of radius  $\gamma$  with center  $\mathbf{C} \in \mathcal{C}$  is the subset

$$B_{\mathcal{C}}(\gamma) = \{\mathbf{P} \in \mathcal{M} \mid d_c(\mathbf{P}, \mathbf{C}) \leq \gamma\} \subseteq \mathcal{M}. \quad (23)$$

#### B. Kissing Radius

Given a code, one can surround each codeword by metric balls of the same radius and enlarge them until two balls touch. This leads to the notion of *kissing radius*. The kissing radius is a generalization of the so-called *packing radius*. We choose a different terminology to emphasize here that the kissing radius of a code may not be a function of the minimum distance, and also because the packing radius is sometimes defined as  $\delta/2$  irrelevantly of the choice of distance [41] by extension from flat geometry.

**Definition 1.** *The kissing radius of a code  $\mathcal{C}$  is the maximum radius of non-overlapping metric balls centered at the codewords:*

$$\varrho = \sup_{\substack{B_{\mathbf{C}_k}(\gamma) \cap B_{\mathbf{C}_l}(\gamma) = \emptyset \\ \mathbf{C}_k \neq \mathbf{C}_l}} \gamma. \quad (24)$$

For each codeword pair  $(\mathbf{C}_k, \mathbf{C}_l)$ , there exists a mid-point  $\mathbf{M}_{k,l}$  which is the closest equidistant point satisfying  $\mathbf{M}_{k,l} = \arg \min_{\mathbf{M}} d_c(\mathbf{C}_k, \mathbf{M})$  subject to  $d_c(\mathbf{C}_k, \mathbf{M}) = d_c(\mathbf{C}_l, \mathbf{M})$ . The mid-points  $\mathbf{M}_{k,l}$  define the mid-distances  $\varrho_{k,l} = d_c(\mathbf{C}_k, \mathbf{M}_{k,l}) = d_c(\mathbf{C}_l, \mathbf{M}_{k,l})$ , and the kissing radius is the minimum of all of them  $\varrho = \min_{k \neq l} \varrho_{k,l}$ .

As the chordal distance is measured along a geodesic defined by the associated Riemann metric, the midpoints can be computed accordingly: with  $\mathbf{C}_{k,l}(t)$  the geodesic such that  $\mathbf{C}_{k,l}(0) = \mathbf{C}_k$  and  $\mathbf{C}_{k,l}(1) = \mathbf{C}_l$ , the midpoint is  $\mathbf{M}_{k,l} = \mathbf{C}_{k,l}(1/2)$ . In the case of the unitary group and the Grassmann

manifold, the geodesic equation is fully parametrized by a set of principal angles  $\{\theta_i\}$  as given in (4) and (14), respectively, and the mid-distance  $\varrho_{k,l}$  can then be directly computed from half of these angles  $\{\frac{\theta_i}{2}\}$ . However, for the case of the Stiefel manifold, the geodesic equation between two points is not known as explained in Remark 1. Alternatively, the midpoint can be computed<sup>1</sup> by an orthogonal projection of the center of mass  $\frac{\mathbf{C}_k + \mathbf{C}_l}{2}$ , which for the Stiefel manifold is given by polar decomposition [9].

For spherical codes one has only one principal angle. This results in a one-to-one mapping between the minimum distance and kissing radius. In contrast, for the Grassmann and Stiefel manifolds with chordal distance, the kissing radius and the minimum distance may not be directly expressible as functions of each other. First, the kissing radius corresponds to the mid-distance between two codewords, which may not be at the minimum distance from each other. Second, for a given pairwise distance, one may have different mid-distances. Grassmann and Stiefel manifolds are not in general two-point homogeneous spaces [15], [40]: one cannot necessarily find a unitary mapping between two pairs of equidistant points, i.e., a pair  $(\mathbf{C}_k, \mathbf{C}_l)$  cannot always be mapped to a pair  $(\mathbf{C}_i, \mathbf{C}_j)$  even if  $d_c(\mathbf{C}_k, \mathbf{C}_l) = d_c(\mathbf{C}_i, \mathbf{C}_j)$ . In fact, in the case of  $\mathcal{M} = \mathcal{U}_n$  and  $\mathcal{G}_{n,p}^{\mathbb{C}}$ , the collection of principal angles provides the complete relative position between two points which is transitive under the action of the unitary group. However, by compressing this “vector-like distance” to the scalar distance  $d_c$ , one loses transitivity.

#### C. Density

The density of a code is the fraction of  $\mathcal{M}$  covered by metric balls centered around the codewords with radius equal to the kissing radius. We consider a uniform measure  $\mu$  on  $\mathcal{M}$  inherited from the Haar measure on the unitary group. Recall that the Grassmann and Stiefel manifolds are homogeneous spaces of the unitary group. For any measurable set  $\mathcal{S} \subset \mathcal{M}$  and any  $\mathbf{U} \in \mathcal{U}_n$ , the uniform measure satisfies  $\mu(\mathbf{U}\mathcal{S}) = \mu(\mathcal{S})$ . Due to the homogeneity of  $\mathcal{M}$ , the characteristics of this ball are independent of its center which for convenience will often not be specified. The measure  $\mu$  corresponds to a normalized volume

$$\mu(B(\gamma)) = \frac{\text{vol } B(\gamma)}{\text{vol } \mathcal{M}}, \quad (25)$$

satisfying  $\mu(\mathcal{M}) = 1$ .

**Definition 2.** *The density of a code  $\mathcal{C} \in \mathcal{M}$  is defined as*

$$\begin{aligned} \Delta(\mathcal{C}) &= \mu \left( \bigcup_{\mathbf{C}_i \in \mathcal{C}} B_{\mathbf{C}_i}(\varrho) \right) = \sum_{\mathbf{C}_i \in \mathcal{C}} \mu(B_{\mathbf{C}_i}(\varrho)) \\ &= N \mu(B(\varrho)). \end{aligned} \quad (26)$$

By definition, it satisfies  $\Delta(\mathcal{C}) \leq 1$ . A maximum packing code has the maximum density for a given cardinality. The problem of maximizing the density for a given cardinality corresponds to maximizing the kissing radius of the code. This

<sup>1</sup>Except if  $\mathbf{C}_k$  and  $\mathbf{C}_l$  are antipodal then their center of mass is  $\mathbf{0}$  which does not have a unique projection.

TABLE II

AN EXAMPLE MAXIMAL-MINIMUM-DISTANCE CODE  $\mathcal{C}_1$  IN  $\mathcal{G}_{4,2}^{\mathbb{C}}$  WHICH IS NOT AN OPTIMAL PACKING SINCE IT HAS A LOWER DENSITY THAN  $\mathcal{C}_2$ .

$\mathcal{C}_1$	$\mathcal{C}_2$
density $\Delta = \frac{8}{9} (7 - 4\sqrt{3}) \approx 0.0638$ min. dist. $\delta = \frac{2}{\sqrt{3}} \approx 1.15$ kissing radius $\varrho = \sqrt{2}\alpha_- \approx 0.65$	density $\Delta = \frac{1}{8} = 0.125$ min. dist. $\delta = 1$ kissing radius $\varrho = \frac{1}{\sqrt{2}} \approx 0.71$
$\left\{ \begin{bmatrix} \alpha_+ & 0 \\ \alpha_- & 0 \\ 0 & \alpha_+ \\ 0 & \alpha_- \end{bmatrix}, \begin{bmatrix} \alpha_+ & 0 \\ -\alpha_- & 0 \\ 0 & \alpha_+ \\ 0 & -\alpha_- \end{bmatrix}, \begin{bmatrix} \alpha_- & 0 \\ i\alpha_+ & 0 \\ 0 & \alpha_- \\ 0 & i\alpha_+ \end{bmatrix}, \begin{bmatrix} \alpha_- & 0 \\ -i\alpha_+ & 0 \\ 0 & \alpha_- \\ 0 & -i\alpha_+ \end{bmatrix} \right\}$ where $\alpha_{\pm} = \sqrt{\frac{1}{6}(3 \pm \sqrt{3})}$	$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}$

TABLE III

INFINITE CODE FAMILIES  $\mathcal{C}_1^m$  AND  $\mathcal{C}_2^m$  IN  $\mathcal{G}_{2^m,2}^{\mathbb{C}}$  GENERALIZING THE CODES IN EXAMPLE 1

$\mathcal{C}_1^m$	$\mathcal{C}_2^m$
Cardinality $N = 4^{m-1}$ min. dist. $\delta = \frac{2}{\sqrt{3}} \approx 1.15$ kissing radius $\varrho = \sqrt{2}\alpha_- \approx 0.65$	Cardinality $N = \binom{2^m}{2}$ min. dist. $\delta = 1$ kissing radius $\varrho = \frac{1}{\sqrt{2}} \approx 0.71$

could be in fact reformulated as another maximum-minimum distance problem where distances among codewords would be defined by the mid-distances. In the case of the Grassmann manifold and unitary group, this would correspond to halve the principal angles in the chordal distance definition, which for the Grassmann manifold can be identified as the “chordal Frobenius-norm” [14] (up to a scaling factor).

In general, the packing problem is not necessarily equivalent to maximizing the minimum distance of the code, and codes with same minimum distance may also have different kissing radiuses and thus densities. We illustrate these statements with two concrete examples below. The first example describes two Grassmann codes where the (proven) maximum minimum-distance code has a lower density than another code.

**Example 1.** Consider the two Grassmannian codes given in Table II. The density of these codes of cardinality four in  $\mathcal{G}_{4,2}^{\mathbb{C}}$  is  $\Delta = 2\varrho^8$  according to the volume formula provided in [17].

The first code  $\mathcal{C}_1$  is an optimal max-min distance code reaching the Rankin simplex bound [1], [13], constructed by embedding the optimal tetrahedron code of  $\mathcal{G}_{2,1}^{\mathbb{C}}$  [42]. The embedding is obtained by a tensor product with the identity matrix, following [43, Prop. 12]. This is a strongly simplicial configuration in the sense that all principal angles equal  $\arccos(\frac{1}{\sqrt{3}})$ . The mid-distances between each pair of codewords are thus all the same, approximately 0.65, leading to a density of  $\approx 0.0638$ .

The second code  $\mathcal{C}_2$  is obtained by circular permutation of the rows of the truncated identity matrix. Its distance distribution corresponds to the embedding of a square. The principal angles between two codewords are either  $\{0, \frac{\pi}{2}\}$  or  $\{\frac{\pi}{2}, \frac{\pi}{2}\}$ . The mid-distances between each of the codewords are either  $1/\sqrt{2}$  or 1. The density of this code is 0.125, almost twice the density of the optimum max-min distance code  $\mathcal{C}_1$  while it has a smaller minimum distance.

The codes in Example 1 can be generalized to two infinite families of codes  $\mathcal{C}_1^m$  and  $\mathcal{C}_2^m$  in  $\mathcal{G}_{2^m,2}^{\mathbb{C}}$  where both

have constant minimum distance and kissing radius: Given  $m \geq 2$ , the codewords  $\mathcal{C}_k \in \mathcal{C}_1^m$  are constructed by the  $m$ -fold tensor product  $\mathcal{C}_k = \mathcal{I}_2 \otimes \mathbf{c}_1 \otimes \cdots \otimes \mathbf{c}_{m-1}$  for all  $\mathbf{c}_i \in \{[\alpha_+, \pm\alpha_-]^T, [\alpha_-, \pm i\alpha_+]^T\}$ , leading to a code with  $4^{m-1}$  codewords in  $\mathcal{G}_{2^m,2}^{\mathbb{C}}$  with constant minimum distance  $\delta = 2/\sqrt{3}$  and kissing radius  $\varrho = \sqrt{2}\alpha_- \approx 0.65$ . The code  $\mathcal{C}_2^m$  is obtained by row permutations of the truncated identity matrix  $\mathcal{I}_{2^m,2}$  leading to  $\binom{2^m}{2}$  codewords with constant minimum distance  $\delta = 1$  and kissing radius  $\varrho = 1/\sqrt{2} \approx 0.71$ . For all  $m$ , the first construction  $\mathcal{C}_1^m$  has a larger minimum distance than  $\mathcal{C}_2^m$ , but the latter has a larger kissing radius and density as summarized in Table III. The code  $\mathcal{C}_2^m$  is larger than  $\mathcal{C}_1^m$  but one could select only  $4^{m-1}$  codewords as in  $\mathcal{C}_1^m$  without changing the kissing radius and minimum distance.

The second example shows that two optimal maximal-minimum-distance codes with different densities can exist.

**Example 2.** In [44], an infinite family of real Grassmann codes meeting the Rankin simplex bound is described generalizing a code found in [13]. Since the simplex bound is the same for real and complex Grassmannians, these codes are also maximum-minimum-distance complex Grassmann codes. It is observed that this construction can lead to different codes with same minimum distance but a different distribution of principal angles, and thus different densities. This observation follows from the lowest dimensional examples  $\mathcal{C}_3$  and  $\mathcal{C}_4$  in Table IV. Each code is the union of four orbits under the action of the cyclic group but from different generator matrices. The density of these codes of cardinality 28 in  $\mathcal{G}_{7,3}^{\mathbb{C}}$  can be approximated [17] to  $\Delta \approx \frac{2}{33}\varrho^{24}$ . The code  $\mathcal{C}_3$  has three distinct sets of principal angles and thus three distinct mid-distances from which the kissing radius is  $\varrho = \sqrt{\frac{9-\sqrt{2}-\sqrt{3}-\sqrt{6}}{6}} \approx 0.75$ . The code  $\mathcal{C}_4$  has two distinct sets of principal angles and thus two mid-distances, among which the set of angles  $\{0, \arccos \frac{1}{3}, \arccos \frac{1}{3}\}$  with mid-distance  $\sqrt{2/3} \approx 0.817$  is also a mid-distance of  $\mathcal{C}_3$ . The kissing radius is achieved with the other set of principal angles which has a mid-distance only slightly smaller  $\approx 0.805$  (it does not seem to have a compact form). While both codes reach the optimal minimum distance of  $\frac{4}{3}$ , the code  $\mathcal{C}_4$  has a density about 5 times larger than  $\mathcal{C}_3$ .

We briefly discuss the performance of these codes when applied to MIMO communications.

MIMO precoding is well-known to be related to Grassmannian packing [8]. In this context, the mu-

TABLE IV  
EXAMPLE OF TWO MAXIMAL-MINIMUM-DISTANCE CODES  $\mathcal{C}_3$  AND  $\mathcal{C}_4$  IN  $\mathcal{G}_{7,3}^C$  WITH DIFFERENT DENSITIES [44].

$\mathcal{C}_3$	$\mathcal{C}_4$
density $\Delta \approx 10^{-4.2}$ min. dist. $\delta = \frac{4}{3}$ kissing radius $\varrho \approx 0.75$	density $\Delta \approx 10^{-3.5}$ min. dist. $\delta = \frac{4}{3}$ kissing radius $\varrho \approx 0.81$
Generator matrices: $\begin{bmatrix} 0 & 1 & 0 & \pm\sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \pm\sqrt{2} \\ 0 & 0 & 0 & 0 & 1 & \pm\sqrt{2} & 0 \end{bmatrix}^T$	Generator matrices: $\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & \pm\sqrt{2} \\ 0 & 0 & 1 & 0 & 0 & \pm\sqrt{2} & 0 \\ 0 & 0 & 0 & \pm\sqrt{2} & 1 & 0 & 0 \end{bmatrix}^T$
The signs are selected such that their product is +1. 28 codewords obtained by circular permutation of the rows of the four generator matrices	

tual information with Gaussian signaling is given by  $\mathbb{E}_{\mathbf{H}} \left[ \log_2 \det \left( \mathbf{I} + \gamma \mathbf{C}_{q(\mathbf{H})}^H \mathbf{H}^H \mathbf{H} \mathbf{C}_{q(\mathbf{H})} \right) \right]$  where  $\gamma$  is the per-stream SNR, and  $q(\cdot)$  is a quantization map that selects the codeword index in  $\mathcal{C}$  maximizing the instantaneous rate for each channel realization  $\mathbf{H}$  inside the expectation. Entries of  $\mathbf{H} \in \mathbb{C}^{p \times n}$  are assumed to be standard complex normal variables  $\mathcal{CN}(0, 1)$ . Numerical evaluations of corresponding mutual informations for the codes in Examples 1 and 2 show that the two higher density codes  $\mathcal{C}_2$  and  $\mathcal{C}_4$  slightly outperform the codes  $\mathcal{C}_1$  and  $\mathcal{C}_3$ , respectively.  $\mathcal{C}_2$  provides a SNR gain of 0.16dB over  $\mathcal{C}_1$ , and  $\mathcal{C}_4$  provides a SNR gain of 0.05dB over  $\mathcal{C}_3$ , in the SNR region  $p\gamma = 20$ dB. These examples and other numerical experiments hint that it may be in fact the density of the code that primarily governs the achievable rate of transmission rather than the maximum-minimum distance. We believe this behavior to be quite generic for the following reason. MIMO precoding reduces to a quantization problem on the Grassmann manifold with chordal distance [17], and the mid-distances of the code reflect better the rate-distortion trade-off since they represent the first border effects of the Voronoi cells.

We now look at the application of unitary codes as space-time constellations over a non-coherent MIMO transmission  $\mathbf{Y} = \sqrt{\text{snr}} \frac{n}{p} \mathbf{C} \mathbf{H} + \mathbf{N}$ . Here  $\mathbf{H} \in \mathbb{C}^{p \times p}$  and  $\mathbf{N} \in \mathbb{C}^{n \times p}$  have standard complex normal  $\mathcal{CN}(0, 1)$  entries, and  $\mathbf{C}$  is uniformly selected from the code  $\mathcal{C}$ . The standard approach in this case is to design Grassmann codes that maximize the product diversity  $\prod_i \sin^2 \theta_i$  since it is minimizing the pairwise block error rate at high-SNR [45]. Low-SNR analysis [18] shows on the other hand that it is the chordal distance (sum diversity)  $\sum_i \sin^2 \theta_i$  that dominates the block error rate, and it was additionally observed in [16] that codes maximizing the chordal distance can lead to higher mutual information  $I(\mathbf{Y}; \mathbf{C})$  than codes with high product diversity. Furthermore, subspace perturbation analysis in [46] suggests that an appropriate code metric is given by the chordal Frobenius-norm, which is exactly the kissing radius of the code with chordal distance as explained above, and thus its density. From the evaluation of the mutual informations  $I(\mathbf{Y}; \mathbf{C})$  of the codes in Examples 1 and 2 shown on Figure 1, the code density does not seem to be an ultimate performance measure when it comes to mutual information, as  $\mathcal{C}_1$  clearly outperforms  $\mathcal{C}_2$ . It should be noted that  $\mathcal{C}_1$  has both a higher minimum chordal distance and product diversity than  $\mathcal{C}_2$ , the later having zero product diversity between four codeword pairs out of six.

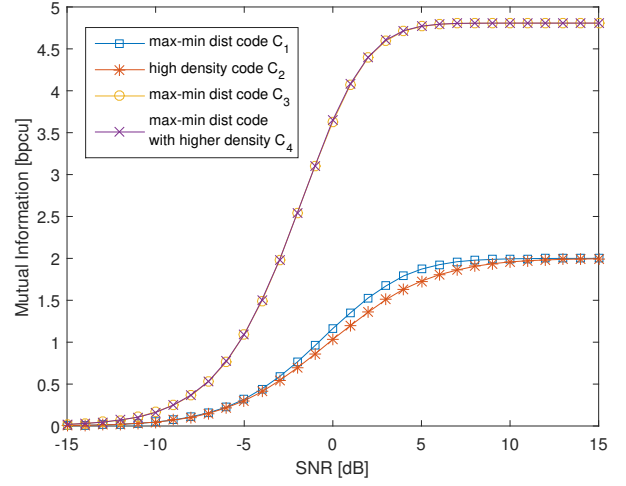


Fig. 1. Comparison of the codes in Examples 1 and 2 when used as a space-time constellation for non-coherent communication.

#### D. Hamming Bound

The notion of density directly relates to the Hamming bound: For any  $(N, \delta)$ -code, one must satisfy

$$N \leq \frac{1}{\mu(B(\frac{\delta}{2}))}. \quad (27)$$

The upper bound in (27) is a direct application of the standard Hamming bound to  $\mathcal{M}$  without taking into account the curvature of the space and the choice of distance. The chordal distance, inherited from a Euclidean embedding, is extrinsic to the considered curved space and thus never satisfies the triangle inequality with equality. Accordingly, balls of radius  $\delta/2$  around the codewords do not necessarily form a packing, as none of the balls would be touching each other and one could possibly fit in an extra ball. The kissing radius  $\varrho$  is larger than  $\delta/2$ , and the Hamming bound can be refined by any radius  $\delta/2 < r \leq \varrho$ , and ultimately for  $r = \varrho$ ,

$$N \leq \frac{1}{\mu(B(\varrho))}. \quad (28)$$

By construction  $\frac{\delta}{2} \leq \varrho$  and thus the Hamming bound (28) is always tighter than the ‘standard Hamming bound’ (27), while being asymptotically equivalent for  $\delta \rightarrow 0$ .

The difficulty in exploiting the improved bound (28) is in finding a relationship between the kissing radius and the minimum distance of the code. To obtain a bound on the minimum



distance, one needs to find a function of the minimum distance such that  $\delta/2 \leq f(\delta) \leq \varrho$ . Then, provided that both the volume expression and  $f$  are invertible, it is possible to bound the minimum distance from above.

#### IV. VOLUMES

In this section, we address the problem of volume computation in the manifolds, providing two different asymptotic approximations of the volume of balls operating in different regimes.

##### A. Spherical Volumes and Hyperspherical Caps

The  $(D-1)$ - and  $D$ -dimensional volume of  $\mathcal{S}^{D-1}(R)$  (with its natural metric  $d_c$ ) are respectively

$$A_D(R) = \frac{2\pi^{D/2}}{\Gamma(\frac{D}{2})} R^{D-1}, \quad V_D(R) = \frac{\pi^{D/2}}{\Gamma(\frac{D}{2} + 1)} R^D. \quad (29)$$

Since the manifolds of interest are submanifolds of hyperspheres, the considered balls are subsets of hyperspherical caps. A hyperspherical cap is a ball on a sphere,

$$C_{D,R}(r) = \{\mathbf{x} \in \mathcal{S}^{D-1}(R) : \|\mathbf{x} - \mathbf{y}\| \leq r\} \quad (30)$$

for some implicit center  $\mathbf{y}$ . One can define a uniform spherical measure  $\sigma$ , and the normalized volume of the spherical caps is denoted by  $\sigma(C_{D,R}(r))$ . We use a different notation for distinction with the uniform measure on the manifold  $\mathcal{M} \subset \mathcal{S}^{D-1}(R)$  but if considering  $\mathcal{M} = \mathcal{S}^{D-1}(R)$  then the two measures match  $\mu(B(r)) = \sigma(C_{D,R}(r))$ .

The  $(D-1)$ -dimensional volume of a spherical cap can be computed exactly, as given below. It is given along with two asymptotics proven in Appendix A.

**Lemma 1.** *The normalized volume (area) of a hyperspherical cap in  $\mathcal{S}^{D-1}(R)$  measured with chordal distance is given by*

$$\sigma(C_{D,R}(r)) = I_{\frac{r^2}{4R^2}} \left( \frac{D-1}{2}, \frac{D-1}{2} \right) \quad (31)$$

where  $r$  is the radius of the cap satisfying  $0 \leq r \leq 2R$ , and  $I_x(a, b)$  is the regularized incomplete beta function.

As the radius of the cap  $r$  goes to zero, the volume of the spherical cap tends to

$$\sigma(C_{D,R}(r)) \simeq \frac{1}{2\sqrt{\pi}} \frac{\Gamma(\frac{D}{2})}{\Gamma(\frac{D+1}{2})} \left( \frac{r}{R} \right)^{D-1}. \quad (32)$$

On the other hand, as the dimension of the sphere  $D$  goes to infinity with  $\sqrt{D} \left(1 - \frac{r^2}{2R^2}\right)$  fixed, the volume tends to

$$\sigma(C_{D,R}(r)) \simeq \frac{1}{2} \operatorname{erf} \left( \sqrt{\frac{D}{2}} \right) - \frac{1}{2} \operatorname{erf} \left( \sqrt{\frac{D}{2}} \left( 1 - \frac{r^2}{2R^2} \right) \right) \quad (33)$$

where  $\operatorname{erf}(x)$  is the Gauss error function defined in Eq. (75).

The volume (32) equals  $V_{D-1}(r)/A_D(R)$ . Intuitively, when a cap is small it is almost “flat” and its volume tends to be the “area of a disc”. The volume (33) corresponds to an asymptotic Gaussian behavior for large dimensions: the squared chordal distance of uniformly distributed random points is asymptotically Gaussian;  $d_c^2 \sim \mathcal{N}\left(2R^2, \frac{4R^4}{D}\right)$  where  $\mathcal{N}(m, v)$  is the

normal distribution with mean  $m$  and variance  $v$ . A similar expression formulating this high-dimensional regime can be found in [47], [48]. As  $D \rightarrow \infty$ , high-dimensional random vectors are asymptotically orthogonal and the chordal distance tends to  $\sqrt{2}R$  which is the mid-distance between two antipodal points.

##### B. Manifold Volume and Small Ball Approximation

1) *Overall Manifold Volume:* The volume of a space can be obtained from the integration of a volume element, which is unique on a Riemannian manifold up to a non-vanishing scaling factor. This scaling factor is induced by the chosen distance, and impacts the overall volume. The theorem below provides the volumes correctly scaled for the chordal distance.

**Theorem 1.** *The volumes of the manifolds induced by the chordal distance  $d_c$  (or its equivalent geodesic distance  $d_g$ ) are*

- for the unitary group [49]

$$\operatorname{vol} \mathcal{U}_n = \frac{(2\pi)^{\frac{n(n+1)}{2}}}{\prod_{i=1}^n (i-1)!}, \quad (34)$$

- for the Stiefel manifold

$$\operatorname{vol} \mathcal{V}_{n,p}^{\mathbb{C}} = \frac{2^{\frac{p(p+1)}{2}} \pi^{np - \frac{p(p-1)}{2}}}{\prod_{i=1}^p (n-i)!}, \quad (35)$$

- for the Grassmann manifold

$$\operatorname{vol} \mathcal{G}_{n,p}^{\mathbb{C}} = \pi^{p(n-p)} \prod_{i=1}^p \frac{(p-i)!}{(n-i)!} \quad (36)$$

A detailed volume computation is provided in Appendix B. While it is possible to find the exact volume for the unitary group rigorously derived for the chordal distance in [49], [50], the equivalent result for the Stiefel manifold does not seem to have been reported before. This volume differs by a constant factor from the commonly cited formula, see e.g. [6], [20], [22]. In many contexts, the overall scaling of the volume is meaningless as often the induced scaling would be absorbed or canceled out. The volumes known in the literature arise from integration of a volume element neglecting scalar prefactors. Conventional volumes in the literature are expressed as the product of the volumes of spheres:  $\operatorname{vol} \mathcal{V}_{n,p}^{\mathbb{C}} = \prod_{k=n-p+1}^n V_{2k}(1) = \prod_{k=n-p+1}^n \frac{2\pi^k}{(k-1)!}$ ,

where  $V_{2k}(1) = \operatorname{vol} \mathcal{S}^{2k-1}$  is given in (29). The two different conventions are related by  $\operatorname{vol} \mathcal{V}_{n,p}^{\mathbb{C}} = 2^{\frac{p(p-1)}{2}} \operatorname{vol} \mathcal{V}_{n,p}^{\mathbb{C}}$ . For the Grassmann manifold with the geometry induced by the spherical embedding, it appears that these normalizations cancel out so that we have  $\operatorname{vol} \mathcal{G}_{n,p}^{\mathbb{C}} = \frac{\operatorname{vol} \mathcal{V}_{n,p}^{\mathbb{C}}}{\operatorname{vol} \mathcal{U}_p} = \frac{\operatorname{vol} \mathcal{V}_{n,p}^{\mathbb{C}}}{\operatorname{vol} \mathcal{U}_p} = \frac{\operatorname{vol} \mathcal{U}_n}{\operatorname{vol} \mathcal{U}_p \operatorname{vol} \mathcal{U}_{n-p}}$ .

However, we remark that the true volume of the Grassmannian with the quotient geometry  $\mathcal{G}_{n,p}^{\mathbb{C}} \cong \mathcal{U}_n/(\mathcal{U}_p \times \mathcal{U}_{n-p})$  and distance  $d_{g*}$  is  $\operatorname{vol}_* \mathcal{G}_{n,p}^{\mathbb{C}} = \frac{\operatorname{vol} \mathcal{U}_n}{\operatorname{vol} \mathcal{U}_p \operatorname{vol} \mathcal{U}_{n-p}}$ , which differs from  $\operatorname{vol} \mathcal{G}_{n,p}^{\mathbb{C}}$  in the embedding geometry computed above

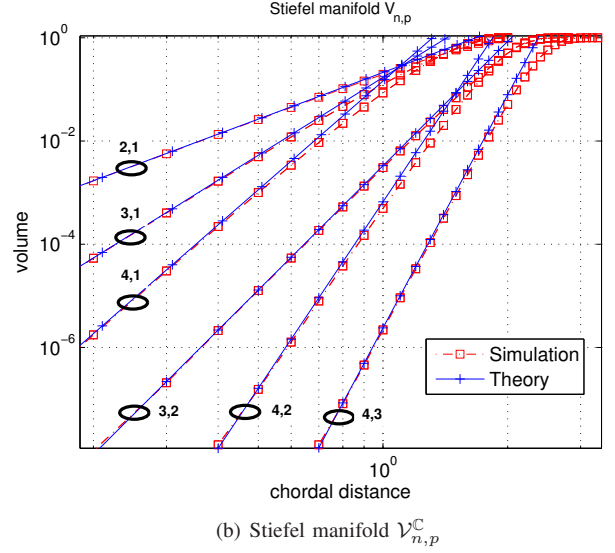
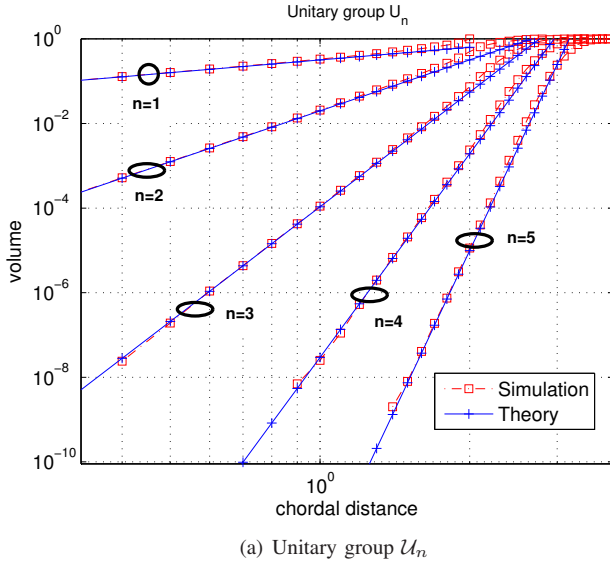


Fig. 2. The small ball volume approximation (39) of Corollary 1 compared to simulation.

by a factor of  $2^{p(n-p)}$ . Using the conventional volumes, one would get  $\text{vol}_* \mathcal{G}_{n,p}^C = \frac{\text{vol} \mathcal{U}_n}{\text{vol} \mathcal{U}_p \text{vol} \mathcal{U}_{n-p}} = \text{vol} \mathcal{G}_{n,p}^C$ , i.e. the volumes in the two geometries would erroneously be assumed to be the same. When interpreting the Haar measure as a normalized volume (25), keeping track of the scaling of the volume is necessary. We suspect that omitting the volume scaling factors is one of the reasons behind the numerical errors observed in [22], in which a standard volume formula was used irrespectively of the geometric interpretation of the Stiefel and Grassmann manifolds, and the related distances.

Considering  $d_c$ , some special manifolds of interest are isometrically isomorphic to spheres: the Stiefel manifolds  $\mathcal{V}_{n,1}^C \cong \mathcal{S}^{2n-1}$ , and the Grassmann manifolds  $\mathcal{G}_{2,1}^C \cong \mathcal{S}^2(\frac{1}{2})$ . As expected, the volumes in Theorem 1 match the spherical volume (29) in these cases.

2) *Small Ball Approximation*: The volume of a small geodesic ball can be well approximated by the volume of a ball of equal radius in the tangent space. This approximation, tight as the radius goes to zero, is actually an upper bound as discussed in [22], known as the Bishop–Gromov inequality, a volume comparison theorem valid for any Riemannian manifold. In [22], it was used to evaluate volumes of small geodesic balls in the Grassmann and Stiefel manifolds. For the case of the Stiefel manifolds, results were extended to the chordal distance in an indirect manner from the geodesic distance  $d_{g*}$  combined with local inequalities. Surprisingly, the results of [17] show that this approximation is exact for the Grassmann manifold *with chordal distance* smaller than one, in the same manner as the area of a cap of the real sphere equals the area  $\pi r^2$  of a disk. Later, the same approximation was used in [51] for volumes in simple flag manifolds. Refining the result, a power series expansion for the volume of small geodesic ball in any Riemannian manifold [52] was later leveraged in [26], [27]. Limiting this expansion to the leading term gives as  $r \rightarrow 0$

$$\text{vol } B(r) = V_{\dim}(r)(1 + O(r^2)) \quad (37)$$

where  $V_{\dim}(r)$  is according to (29) with the dimension  $\dim$  of the manifold in Table I. Intuitively, in a small neighborhood the manifold looks like a Euclidean space and can be approximated by the tangent space. Other coefficients of the series expansion are addressed in [27] requiring computation of the curvature of the manifold.

The expansion (37) given for the geodesic distance extends to the corresponding chordal distance induced by the isometric embedding in  $\mathbb{R}^D$  as  $d_g = d_c + O(d_c^3)$  [53], [54]. Therefore, the normalized volume of metric balls in  $\mathcal{M}$  with dimension  $\dim$  for *both* the geodesic distance  $d_g$  and the chordal distance  $d_c$  is given by

$$\mu(B(r)) = \frac{V_{\dim}(r)}{\text{vol } \mathcal{M}} (1 + O(r^2)) \quad (38)$$

as  $r \rightarrow 0$ . We then have the following results as a direct consequence of Theorem 1 and  $V_D(r)$  given in (29).

**Corollary 1.** *The volume of metric balls as  $r \rightarrow 0$  with metrics  $d_c$  or  $d_g$  is*

$$\mu(B(r)) = c_{n,p} r^{\dim} (1 + O(r^2)), \quad (39)$$

- where for the Stiefel manifold  $\mathcal{V}_{n,p}^C$ ,

$$c_{n,p} = \frac{2^{-\frac{p(p+1)}{2}} \pi^{-\frac{p}{2}}}{\Gamma(p(n-p/2)+1)} \prod_{i=1}^p (n-i)!, \quad (40)$$

- and for the Grassmann manifold  $\mathcal{G}_{n,p}^C$  [17],

$$c_{n,p} = \frac{1}{(p(n-p))!} \prod_{i=1}^p \frac{(n-i)!}{(p-i)!}. \quad (41)$$

In the case of the geodesic distance  $d_g$ , the Bishop–Gromov inequality [22] is accordingly given by  $\mu(B(r)) \leq c_{n,p} r^{\dim}$ .

For the Stiefel manifold with  $n \neq p$ , the result above differs from the one in [27] derived for quotient geometry ( $d_{g*}$ ). The two results match for the unitary group with  $n = p$ , as expected. Comparing to the Grassmannian case in [27], there is a difference by a factor of  $2^{2p - \frac{n-p}{2}(n-p+1)}$ . The local

equivalence between the chordal distance  $d_c$  and the geodesic distance  $d_g$  makes the result identical for both metrics. For the Grassmann manifold with chordal distance, the result of [17] is stronger than Corollary 1 as it gives  $\mu(B(r)) = c_{n,p} r^{\dim}$  for  $r < 1$ .

The approximated volumes in Corollary 1 for the Stiefel manifold (with  $p \neq n$ ) and for the unitary group (with  $p = n$ ) are compared with simulation in Fig. 2. The figures are shown in logarithm scales since the region of interest is small chordal distance. We see that the approximations match almost exactly the simulations as  $r \rightarrow 0$ , which is a consequence of the exact volume normalization given in Theorem 1.

### C. Complementary Balls

As an interlude, we discuss complementary balls. Small ball approximations can be used to compute the volume of very large balls almost totally covering the space. In the same manner as a sphere can be fully covered by two complementary caps satisfying

$$\sigma(C_{D,R}(r)) = 1 - \sigma(C_{D,R}(\sqrt{4R^2 - r^2})), \quad (42)$$

we have the following analogous result for the Stiefel and Grassmann manifolds, proved in Appendix C.

**Lemma 2.** *In the Stiefel manifold  $(\mathcal{V}_{n,p}^{\mathbb{C}}, d_c)$ ,*

$$\mu(B(r)) = 1 - \mu(B(\sqrt{4p - r^2})) \quad (43)$$

*implying that the volume is symmetrical at  $\mu(B(\sqrt{2p})) = 1/2$ .*

*For the Grassmann manifold  $(\mathcal{G}_{n,p}^{\mathbb{C}}, d_c)$ ,*

$$\mu(B(r)) = 1 - \mu\left(B_{\perp}\left(\sqrt{\frac{2p(n-p)}{n} - r^2}\right)\right), \quad (44)$$

*where  $B_{\perp}(\gamma) = \{[P] \in \mathcal{G}_{n,p}^{\mathbb{C}} \mid d_c([Q], [P]) \leq \gamma\}$  for an arbitrary center  $[Q] \in \mathcal{G}_{n,n-p}^{\mathbb{C}}$  and with chordal distance<sup>2</sup> as defined in (19).*

We remark that there is a structural difference here between the Stiefel and Grassmann manifolds. The antipodal on the embedding sphere to a point in the Stiefel manifold always belongs to the same Stiefel manifold. This is not always true for the Grassmann manifold except for  $p = n/2$ , as the antipodal on the sphere to a point in  $\mathcal{G}_{n,p}^{\mathbb{C}}$  belongs to  $\mathcal{G}_{n,n-p}^{\mathbb{C}}$ . This observation will be useful in the interpretation of the high-dimensional regime discussed next.

### D. Volume Comparison Theorem and Spherical-Cap Approximation

We now present a second volume approximation: For the manifolds  $\mathcal{M}$  isometrically embedded in  $\mathcal{S}^D(R)$ , the uniform measure of a ball in  $\mathcal{M}$  can be well approximated by the spherical measure of a cap on  $\mathcal{S}^D(R)$ . The two normalized volumes are indeed asymptotically equivalent in the high-dimensional regime. Before stating the theorem, we highlight an intermediate result of independent interest.

<sup>2</sup>Note that here the definition of the chordal distance between two Grassmannian planes of non-equal dimensions differs from [17], where one would have  $\mu(B(r)) = 1 - \mu(B_{\perp}(\sqrt{p - r^2}))$ , see [33].

**Lemma 3.** *The squared chordal distance  $d_c^2 = d_c^2(\mathbf{I}_{n,p}, \mathbf{Y})$  or  $d_c^2 = d_c^2([\mathbf{I}_{n,p}], [\mathbf{Y}])$  drawn from a uniformly distributed random point  $\mathbf{Y} \in \mathcal{V}_{n,p}^{\mathbb{C}}$  and a reference point, say  $\mathbf{I}_{n,p}$ , converges in distribution to a Gaussian random variable:*

$$\frac{1}{\sqrt{\text{Var}[d_c^2]}} (d_c^2 - \mathbb{E}[d_c^2]) \xrightarrow{d} \mathcal{N}(0, 1), \quad (45)$$

- *for the Stiefel manifold  $\mathcal{V}_{n,p}^{\mathbb{C}}$ , as  $n \rightarrow \infty$  and where the finite-size regime mean and variance are given by*

$$\mathbb{E}[d_c^2] = 2p, \quad \text{Var}[d_c^2] = \frac{2p}{n}; \quad (46)$$

- *for the Grassmann manifold  $\mathcal{G}_{n,p}^{\mathbb{C}}$ , as  $n, p \rightarrow \infty$  with  $(n - 2p)$  fixed, and where*

$$\mathbb{E}[d_c^2] = \frac{p(n-p)}{n}, \quad \text{Var}[d_c^2] = \frac{p^2(n-p)^2}{n^4 - n^2}. \quad (47)$$

The proofs are given in Appendix D. They are obtained by reducing the chordal distances to the linear statistics of random matrix ensembles and studying their moment-generating functions. Such linear statistics are asymptotically Gaussian, and by computing the two first moments one finds the asymptotic forms. The case of the Stiefel manifold is closely related to the partition function of the von-Mises Fisher distribution, while for the Grassmann manifold it is related to the partition function of the Bingham distribution [55].

It follows that the chordal distance is well-approximated by a Gaussian random variable with the same finite-size mean and variance, i.e.  $d_c^2(\mathbf{I}_{n,p}, \mathbf{Y}) \sim \mathcal{N}(2p, \frac{2p}{n})$  for the Stiefel manifold  $\mathcal{V}_{n,p}^{\mathbb{C}}$ , and  $d_c^2([\mathbf{I}_{n,p}], [\mathbf{Y}]) \sim \mathcal{N}\left(\frac{p(n-p)}{n}, \frac{p^2(n-p)^2}{n^4 - n^2}\right)$  for the Grassmann manifold  $\mathcal{G}_{n,p}^{\mathbb{C}}$ .

Using (33) for the hyperspherical cap volumes, one gets a volume comparison of the ball in the embedded manifold with the embedding spherical cap:

**Theorem 2.** *The normalized volume of a ball of radius  $r$  in  $\mathcal{M}$  is asymptotically equal to the normalized volume of a cap with same radius in the embedding sphere  $(\mathcal{M}, d_c) \hookrightarrow \mathcal{S}^{D-1}(R)$ ,*

$$\mu(B(r)) \simeq \sigma(C_{D,R}(r)) \quad (48)$$

*in the high-dimensional regime as given in (33). We have*

- *for the Stiefel manifold  $\mathcal{M} = \mathcal{V}_{n,p}^{\mathbb{C}}$ ,*

$$\mu(B(r)) \simeq \frac{1}{2} \text{erf}(\sqrt{np}) - \frac{1}{2} \text{erf}\left(\sqrt{np} \left(1 - \frac{r^2}{2p}\right)\right), \quad (49)$$

*as  $n \rightarrow \infty$  with  $\sqrt{2np} \left(1 - \frac{r^2}{2p}\right)$  fixed;*

- *and for the Grassmann manifold  $\mathcal{M} = \mathcal{G}_{n,p}^{\mathbb{C}}$ ,*

$$\mu(B(r)) \simeq \frac{1}{2} \text{erf}\left(\sqrt{\frac{n^2 - 1}{2}}\right) - \frac{1}{2} \text{erf}\left(\sqrt{\frac{n^2 - 1}{2}} \left(1 - \frac{n r^2}{p(n-p)}\right)\right) \quad (50)$$

*as  $n, p \rightarrow \infty$  with  $(n - 2p)$  and  $\sqrt{n^2 - 1} \left(1 - \frac{n r^2}{p(n-p)}\right)$  fixed.*

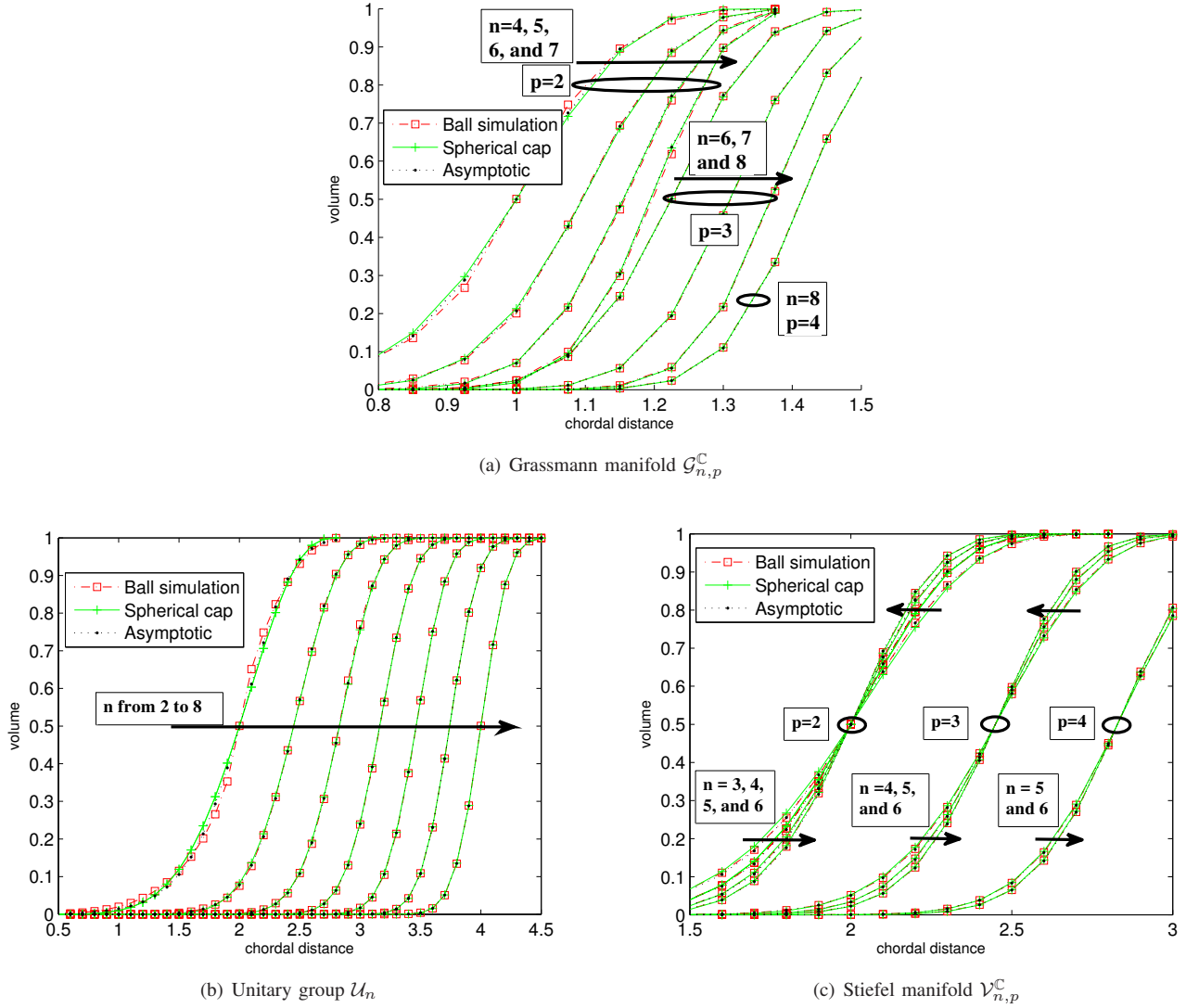


Fig. 3. Illustration of Theorem 2: volume of balls in manifolds (Monte Carlo simulation), spherical cap volume (31), and asymptotic evaluation (33) (equivalently (49) and (50)).

The proof is given in Appendix E. The volume (49) generalizes our previous result [31], [32] from the unitary group to the Stiefel manifold. The volume (50) corresponds to the special case  $p = q$  in [33]. Theorem 2 provides a geometric unification of the asymptotic expressions which are shown to be closely related due to their spherical embedding. The different asymptotic regimes for the Stiefel and Grassmann manifolds can be geometrically understood as the behaviors in Lemma 2. The Stiefel manifold fully covers all possible distances on its embedding sphere. This is not the case for the Grassmann manifold, and as a consequence  $\mu(B(r))$  and  $\sigma(C_{D,R}(r))$  are not defined on the same support, except for  $\mathcal{G}_{n,p}^C$ . In the regime  $n, p \rightarrow \infty$  with fixed  $n - 2p$ , one has  $p \rightarrow \frac{n}{2}$  and the volume expressions are then asymptotically defined on the same support.

There exists a body of literature on comparison theorems in Riemannian geometry that compares volumes among manifolds with reference to the standard sphere. However, these theorems often compare manifolds of equal dimension, e.g. the

Bishop-Gromov inequality or the Berger-Kazdan comparison theorem. To the best of our knowledge, volume comparisons of these manifolds with their embedding spheres in Theorem 2 is new.

The quality of the approximation is illustrated in Figure 3. The normalized volume of metric balls for different Grassmann and Stiefel manifolds obtained via Monte Carlo simulations is compared to the normalized volume of their respective embedding hyperspherical cap, together with their joint asymptotic expression. The exact volume of the hyperspherical cap is as given by Lemma 1. This numerical evaluation shows that the three volume expressions are very close to each other even in the low dimensional regime.

## V. KISSING RADIUS AND DENSITY

In this section, we discuss the kissing radius of codes with chordal distance, and apply volume approximations derived in the previous section to evaluate code density. Recall that for the Grassmann and Stiefel manifold with chordal distance, the



kissing radius  $\varrho$  cannot be directly expressed in term of the minimum distance of the code. In the following, upper and lower bounds on  $\varrho$  and the corresponding bounds on density are provided.

#### A. Preliminaries

1) *Hypothetical Covering Radius*: Ideally, a set of packing balls would fully cover the space, reaching a maximum density of one. This is only possible when the cardinality of the code is  $N = 2$ , otherwise one gets an upper bound. This ideal radius  $r_N$  fulfills

$$\mu(B(r_N)) = \frac{1}{N}. \quad (51)$$

Two volume approximations were discussed in the previous section. Depending of the regime, one may compute  $r_N$  accordingly.

- For  $N \geq c_{n,p}^{-1}$  the radius is less than one, and the small ball approximation leads to

$$r_N \approx (c_{n,p} N)^{\frac{-1}{\dim}}, \quad (52)$$

where  $c_{n,p}$  is given in (40), (41) and  $\dim$  is the dimension of the manifold as given in Table I. For the Grassmann manifold, (52) holds with equality with  $N \geq c_{n,p}^{-1}$  [17].

- Otherwise for a larger ball, occurring with  $n, p$  large and relatively small  $N$ , the spherical approximation of Theorem 2 leads to

$$r_N \approx \sqrt{2}R \sqrt{1 - \sqrt{\frac{2}{D}} \operatorname{erf}^{-1} \left( \operatorname{erf} \sqrt{\frac{D}{2}} - \frac{2}{N} \right)}, \quad (53)$$

where  $R, D$  are the radius and the dimension of the spherical embedding, provided in Table I.

2) *Kissing Radius for Spherical Codes*: For spherical codes with chordal distance, the kissing radius of a code is given by a one-to-one mapping from the minimum distance, directly computable by the Pythagorean theorem. Given an  $(N, \delta)$ -spherical code on  $\mathcal{S}^{D-1}(R)$ , the midpoint on the geodesic between the two codewords of distance  $\delta$  is at distance  $\varrho_s$  from the extremities:

$$\varrho_s = \sqrt{2}R \sqrt{1 - \sqrt{1 - \frac{\delta^2}{4R^2}}}, \quad (54)$$

which can be inverted as

$$\delta^2 = 4\varrho_s^2 - \frac{\varrho_s^4}{R^2}. \quad (55)$$

3) *Preliminary Bounds on Kissing Radius*: Since we are considering a manifold isometrically embedded in the Euclidean sphere  $\mathcal{S}^{D-1}(R)$ , an  $(N, \delta)$ -code in  $\mathcal{M}$  is an  $(N, \delta)$ -spherical code. As a consequence, as balls of radius  $\varrho_s$  are non-overlapping on  $\mathcal{S}^{D-1}(R)$ , their inverse image on  $\mathcal{M}$  are also non-overlapping, and we can deduce that  $\varrho_s \leq \varrho$ . On the other hand, we know that for every non-overlapping ball of radius  $r$ , we have  $r \leq r_N$ .

**Lemma 4.** *Given an  $(N, \delta)$ -code in  $\mathcal{M}$  isometrically embedded in  $\mathcal{S}^{D-1}(R)$ , the kissing radius  $\varrho$  is bounded by*

$$\varrho_s \leq \varrho \leq r_N, \quad (56)$$

where  $\varrho_s$  is given in (54) and  $r_N$  satisfies (51).

It should be noted here that the lower bound is a function of the minimum distance  $\delta$ , while the upper bound is a function of the cardinality  $N$ .

#### B. Bounds on Kissing Radius and Density

We now provide bounds on the kissing radius as a function of the minimum distance of the code only, and corresponding bounds on code density.

**Proposition 1.** *For any  $(N, \delta)$ -code  $\mathcal{C}$  in  $\mathcal{M}$ , we have*

$$\underline{\varrho} \leq \varrho \leq \bar{\varrho}, \quad (57)$$

where

$$\underline{\varrho} = \begin{cases} \sqrt{\frac{p}{2} \left( 1 - \sqrt{1 - \frac{\delta^2}{p}} \right)} & \text{for } \mathcal{M} = \mathcal{G}_{n,p}^{\mathbb{C}} \\ \sqrt{2p \left( 1 - \sqrt{1 - \frac{\delta^2}{4p}} \right)} & \text{for } \mathcal{M} = \mathcal{V}_{n,p}^{\mathbb{C}} \end{cases} \quad (58)$$

$$\bar{\varrho} = \begin{cases} \frac{1}{\sqrt{2}} \sqrt{[\delta^2] - \sqrt{[\delta^2]} - \delta^2} & \text{for } \mathcal{M} = \mathcal{G}_{n,p}^{\mathbb{C}} \\ \sqrt{2} \sqrt{[\frac{\delta^2}{4}] - \sqrt{[\frac{\delta^2}{4}]} - \frac{\delta^2}{4}} & \text{for } \mathcal{M} = \mathcal{U}_n \end{cases} \quad (59)$$

and  $\lceil x \rceil$  is the smallest integer greater than  $x$ . It follows, therefore, that the density is bounded by

$$N\mu(B(\underline{\varrho})) \leq \Delta(\mathcal{C}) \leq \min\{1, N\mu(B(\bar{\varrho}))\}. \quad (60)$$

A detailed proof can be found in Appendix F. Given two points either on the Grassmann manifold or the unitary group, their midpoint can be determined according to their principal angles. The bounds then follow by optimizing over the principal angles. For the unitary group, the obtained lower bound matches the spherical lower bound in Lemma 4, and can thus be extended to any Stiefel manifold. We conjecture that the upper bound (59) can be generalized to all Stiefel manifolds. The lack of a closed-form geodesic equation between two points, as explained in Remark 1 (or equivalently the absence of the notion of principal angles), renders a tentative proof of generalization difficult. Nevertheless, numerical experiments support this generalization, and note that the upper bounds (59) do not depend on any dimension parameter.

For the Grassmann manifold, the lower bound in Proposition 1 provides an improvement of the spherical embedding bound (see Appendix F-D).

**Corollary 2.** *For the Grassmann manifold, the lower bound in Proposition 1 is tighter than the lower bound in Lemma 4. These two bounds are equal if and only if  $p = n/2$ .*

Figure 4 illustrates the upper and lower bounds (57) on the kissing radius and shows that  $\delta/2$  is a good approximation for  $\varrho$  only when the minimum distance of the code is relatively small. In general, since the chordal distance is not strictly intrinsic, we have  $\frac{\delta}{2} < \varrho$ . The bounds are also compared to  $10^5$  simulated midpoints between a fixed center and uniformly-distributed random points. We stress that in this case the bounds are tight in the sense that it is always possible to construct two diagonal codewords fulfilling the bounds.

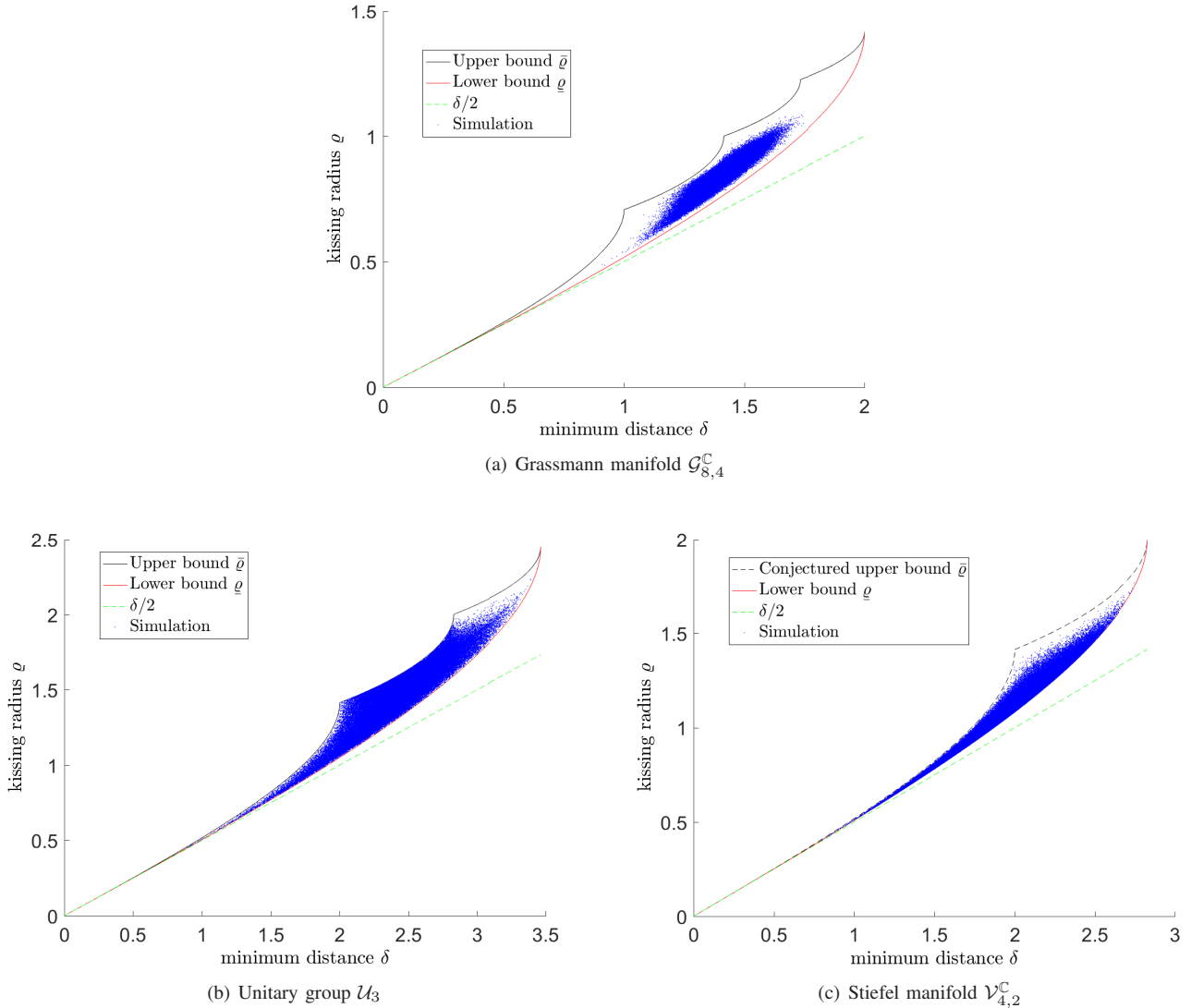


Fig. 4. Illustration of the kissing radius bounds (57) from Proposition 1. Bounds are compared to simulated midpoints between two randomly generated codewords. It is also compared with the estimate  $\delta/2$ , corresponding to the classical packing radius in flat geometry.

### C. On High-Density Codes

Since the lower bound (58) is an increasing function of  $\delta$ , a large minimum distance always guarantees a good code density but not necessarily the highest one.

1) *Numerical Experiments:* Figure 5 compares the density of codes with sizes from  $N = 2$  to  $N = 15$  in  $\mathcal{G}_{4,2}^C$  to the density bounds (60) in Proposition 1. The Grassmannian  $\mathcal{G}_{4,2}^C$  is embedded in  $\mathcal{S}^{14}(\frac{1}{\sqrt{2}})$ , and the corresponding spherical Rankin bound<sup>3</sup> provides an upper bound on the maximum possible minimum distance of Grassmannian codes. We have generated Grassmannian codes according to two criteria: a maximum-minimum-distance criterion, and a low-distortion<sup>4</sup> criterion. For visibility, the densities from the latter are displayed at  $(N + 0.1)$  while for the former at  $(N - 0.1)$ . For

<sup>3</sup>In this dimensionality, spherical codes achieve the simplex bound for  $N \leq 15$ , then the orthoplex bound for  $16 \leq N \leq 30$ .

<sup>4</sup>The code distortion refers here to the average squared quantization error of a uniform random source quantized to the code.

each code, two densities are displayed on Fig. 5: i) the density as a Grassmann code, and ii) the density as a spherical code. The bounds on Grassmann densities from Proposition 1 are shown as range bars. The higher is the minimum distance of the code, the higher is the bar.

Maximum-minimum-distance codes were obtained using the Alternating Projection algorithm from [41]. The obtained codes match the Rankin bound on squared minimum distance with  $10^{-4}$  numerical precision.

The Lloyd algorithm is used to create low-distortion codes. It is expected to provide high density codes since the kissing radius corresponds to the first border effect of Voronoi cells. In [17], it was shown that distortions are bounded by idealized codes that would have only one border effect, corresponding to a kissing radius equal to  $r_N$ , i.e. the upper bound (56) of Lemma 4.

While the Alternating Projection algorithm generates numerically optimal simplicial codes with maximum minimum distance, their density appears to be always close to the corre-

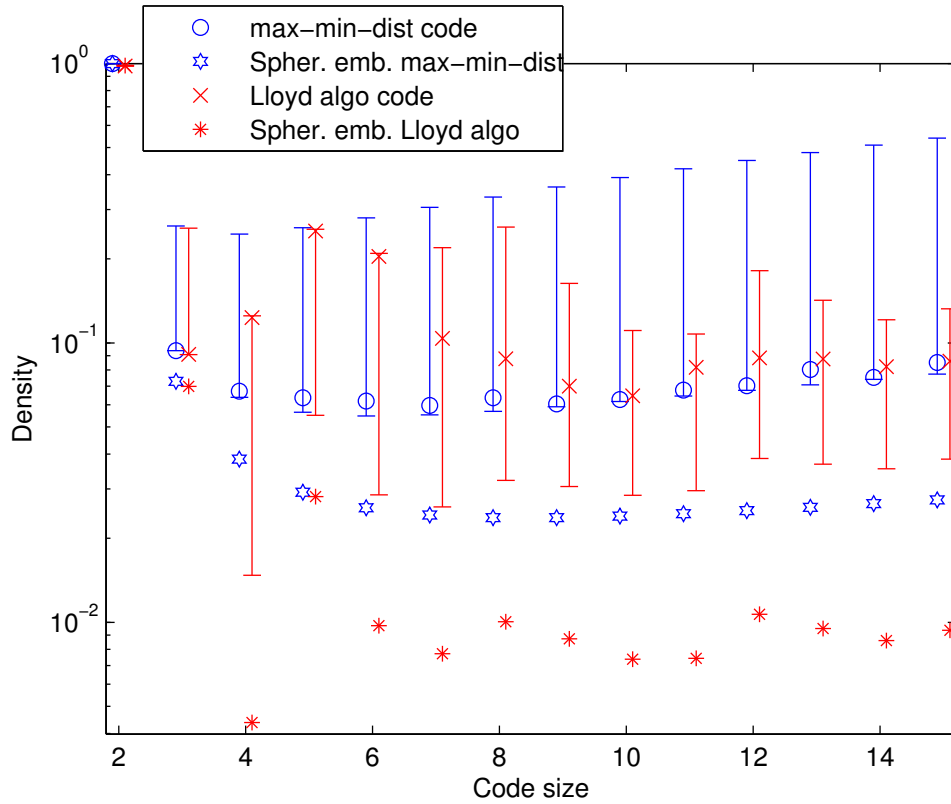


Fig. 5. Density of codes in  $\mathcal{G}_{4,2}^C$  and their image code in  $S^{14}\left(\frac{1}{\sqrt{2}}\right)$  as a function of the number of points. The error bars represent the density bounds (60) from Proposition 1 according to the minimum distance of a code.

sponding lower bound. This means that the codes are strongly simplicial configurations where all principal angles are almost equal. Comparing to the codes generated by Lloyd algorithm, the results for cardinalities 3, 4 and 5 clearly show the weak relationship between maximizing the minimum distance and maximizing the density. For  $N = 3$ , the Lloyd algorithm produced an equivalent max-min-distance code, which is at the density lower bound. For  $N = 4$ , it generated a code with smaller minimum distance but larger density, reaching the density upper bound. The configurations obtained for  $N = 4$  correspond to codes numerically equivalent to the closed form codes in Table II. For  $N = 5$ , the Lloyd algorithm produced a code with minimum distance close to the optimum, as for  $N = 3$ , but this time with a density reaching the upper bound, while the code from the Alternating Projection algorithm is close to the lower bound. For other cardinalities, the Lloyd algorithm produced codes with greater or equal density but smaller minimum distance than Alternating-Projection codes.

Lastly, an interesting characteristic of Grassmann codes is that their density for  $N = 2$  is not  $\Delta = 1$ , except for  $p = \frac{n}{2}$ . This comes from the fact that when  $p \neq \frac{n}{2}$  one cannot have two spherically antipodal points, c.f. the symmetry of Lemma 2. For example, the density of two orthogonal lines in  $\mathcal{G}_{3,1}^C$  is equal to 0.5, and one can actually pack a third orthogonal line to reach the density of 0.75.

2) *Comments on Algebraic Constructions:* For a given minimum distance  $\delta$ , the proof of Proposition 1 shows that to achieve the largest density in  $\mathcal{G}_{n,p}^C$  or  $\mathcal{U}_n$ , a highly concentrated

distribution of principal angles should be targeted where almost all non-zero principal angles should be at the maximum value which is  $\frac{\pi}{2}$  for  $\mathcal{G}_{n,p}^C$  and  $\pi$  for  $\mathcal{U}_n$ . Conversely, the lower bound is achieved when all principal angles are equal.

When constructing group codes from orbits of a symmetry group, the distribution of principal angles can be controlled. The two infinite families of codes in Table III are orbit codes.  $\mathcal{C}_1^m$  is an orbit under the action of the  $(m-1)$ -fold tensor product of the projective unitary representation of the Klein 4-group  $V_4$  [56], i.e. a direct product of the group of symmetries of a tetrahedron. As a result, the principal angles between two points at minimum distance are  $\left\{\arccos \frac{1}{\sqrt{3}}, \arccos \frac{1}{\sqrt{3}}\right\}$  and the kissing radius is at the lower bound (58) of Proposition 1. The code  $\mathcal{C}_2^m$  is an orbit of the fundamental representation of the symmetric group  $S_{2^m}$  via permutation matrices in the corresponding dimension. The resulting principal angles are either  $\left\{\frac{\pi}{2}, \frac{\pi}{2}\right\}$  or  $\left\{0, \frac{\pi}{2}\right\}$  if the two codewords have zero or one column in common, respectively. Contrary to  $\mathcal{C}_1^m$ , the codes  $\mathcal{C}_2^m$  reach the upper bound (59) of Proposition 1.

In [44], a family of optimal max-min distance Grassmann codes is presented from which the code  $\mathcal{C}_2$  in Example 1 is a subset of. These codes are orbits of a large Clifford group generalizing the symmetries of the square. This results in highly symmetric codes where many principal angles are  $\frac{\pi}{2}$ . Nevertheless, the kissing radius of these codes typically does not reach the upper bound (59). Consider the half-subspaces construction in  $\mathcal{G}_{n,\frac{n}{2}}^C$  from [44], [57]. Its complex extension [16] leads to a code of  $2(n^2 - 1)$  codewords. Due

to its orbit structure, the collection of pairwise distances or principal angles can be computed independently of the reference codeword as described in [44]. For each codeword, there is always an antipodal codeword at pairwise distance  $\delta_{i,j}^2 = n$ , and  $2(n^2 - 2)$  others at distance  $\delta_{i,j}^2 = \frac{n}{4}$ . The codes thus meet the Rankin orthoplex bound. Among the codewords at minimum distance from a given codeword, there are  $(n^2 - 4)$  that have principal angles  $\{\frac{\pi}{2}, \dots, \frac{\pi}{2}, 0, \dots, 0\}$  (each value with multiplicity  $\frac{n}{4}$ ), whereas the remaining  $n^2$  codewords have principal angles  $\{\frac{\pi}{4}, \dots, \frac{\pi}{4}\}$ . Thus roughly half of the squared mid-distances are  $\varrho_{i,j}^2 = \frac{n}{4}$ , and roughly the other half are  $\varrho_{i,j}^2 = \frac{(2-\sqrt{2})n}{8}$ . These mid-distances thus meet the upper bound (59) and lower-bound (58), respectively. The kissing radius of this code family is hence the lower bound  $\varrho = \sqrt{\frac{(2-\sqrt{2})n}{8}}$ . It should be remarked that a mid-distance at the lower bound (58), where principal angles between two codewords are equal, corresponds to *maximizing* the diversity product for the given chordal distance. Therefore, about half of the codeword pairs have a good product diversity of  $2^{-\frac{n}{2}}$ , while the others have a product diversity 0. The codes in [16], [44] provide thus a mixture of high chordal mid-distance (as advocated in [46] by the chordal Frobenius-norm) and high product diversity (as advocated in [45]), in addition to a high chordal distance (as advocated in [18]). This observation may provide further support to their good performance observed in [16] when applied to space-time constellations.

#### D. Relation to Density of Spherical Codes.

As considered in this work, for manifolds isometrically embedded in Euclidean spheres, an  $(N, \delta)$ -code in  $\mathcal{M}$  is also an  $(N, \delta)$ -spherical code. The density of a code in the manifold differs from the density of its image code in the embedding sphere. Combining the asymptotic equality from Theorem 2 between volume of balls in the manifold and hyperspherical cap with the kissing radius bounds from Proposition 1, we have the following comparison.

**Proposition 2.** *The density of codes in the Stiefel and Grassmann manifolds are asymptotically greater than the density of their image spherical codes in the high-dimensional regime of Theorem 2. The inequality is strict for the Grassmann manifold with  $N > 2$ .*

This proposition is illustrated in Figure 5 where densities of the image spherical codes of the considered Grassmann codes are also depicted. One observes that the densities of the image spherical codes are always less than the densities of the original Grassmannian codes for  $N > 2$ .

We remark also few special cases. In the case  $p = 1$ , the lower and upper bounds on the kissing radius in Proposition 1 are matching  $\varrho = \bar{\varrho}$ . Moreover with  $p = 1$ , the Stiefel manifold is isometric to the embedding sphere, while for the Grassmann manifold the volume of a ball has been calculated exactly in [17]. In these cases the density can be computed exactly as a function of the minimum distance.

**Corollary 3.** *For any packing with  $p = 1$ , on manifolds  $\mathcal{G}_{n,1}^{\mathbb{C}}$  or  $\mathcal{V}_{n,1}^{\mathbb{C}}$ ,*

$$\varrho = \underline{\varrho} = \bar{\varrho} \quad (61)$$

and

$$\Delta(C) = \begin{cases} N \left( \frac{1-\sqrt{1-\delta^2}}{2} \right)^{n-1} & \text{for } \mathcal{M} = \mathcal{G}_{n,1}^{\mathbb{C}} \\ NI_{\frac{1}{2}(1-\sqrt{1-\delta^2/4})} \left( \frac{2n-1}{2}, \frac{2n-1}{2} \right) & \text{for } \mathcal{M} = \mathcal{V}_{n,1}^{\mathbb{C}} \end{cases} \quad (62)$$

Finally, for the specific case of  $n = 2, p = 1$ , the Grassmann manifold is also isometric to the real sphere:  $\mathcal{G}_{2,1}^{\mathbb{C}} \cong S^2$  [58, Ex. 17.23] [42], and the discussed density of a code is consistent with the definition of the density for a sphere packing from the literature [59].

#### E. Hamming-type Bounds

Bounds on density directly translate to Hamming-type bounds on code cardinality and minimum distance.

1) *Hamming Bound on Cardinality:* According to Proposition 1, we have:

**Corollary 4.** *For any  $(N, \delta)$ -code in  $\mathcal{M}$ , and given  $\underline{\varrho}$  defined in (58),*

$$N \leq \frac{1}{\mu(B(\underline{\varrho}))}. \quad (63)$$

For the unitary group, this bound is equivalent to the one derived in [23]. For the Grassmann manifold, a bound based on spherical embedding and asymptotic analysis was provided in [20]. Equation (63) provides a tighter bound.

2) *Hamming-type Bound on Minimum Distance:* From the spherical embedding of the manifolds, we have the following Hamming-type bound on the minimum distance:

**Lemma 5.** *Given an  $(N, \delta)$ -code in  $\mathcal{M}$  isometrically embedded in  $S^{D-1}(R)$ , and a  $r_N$  satisfying (51), we have*

$$\delta^2 \leq 4r_N^2 - \frac{r_N^4}{R^2}. \quad (64)$$

This follows from a direct combination of (55) and Lemma 4. For the unitary group, a similar bound was derived in [23, Theorem 2.4]. For the Stiefel manifold, the result is new. For the Grassmann manifold, a tighter bound is provided as a by-product of Proposition 1, which is a generalization of a bound for line packing in [60] to any value of  $p$ :

**Lemma 6.** *Given a  $(N, \delta)$ -code in  $\mathcal{G}_{n,p}^{\mathbb{C}}$ , we have*

$$\delta^2 \leq 4r_N^2 - \frac{4}{p}r_N^4. \quad (65)$$

**Remark 2.** *For the case  $p = 1$  the bound of Lemma 6 reduces to the following bound derived in [60, (32)]:*

$$\delta^2 \leq 4N^{\frac{-1}{n-1}} - 4N^{\frac{-2}{n-1}}. \quad (66)$$

3) *Illustration and Conjectured Improvement:* The standard Hamming bound (27) and the Hamming bound from the kissing radius analysis (65) are displayed in Fig. 6 for the Grassmannian  $\mathcal{G}_{4,2}^{\mathbb{C}}$  as a function of the code rate. One verifies that (65) improves (27) and they are getting equivalent as  $N \rightarrow \infty$ . These Hamming bounds are also compared to the spherical Rankin bounds. Rankin provided three consecutive upper bounds in [1] on the number of spherical caps that can be packed for a given angular radius. The two first bounds,



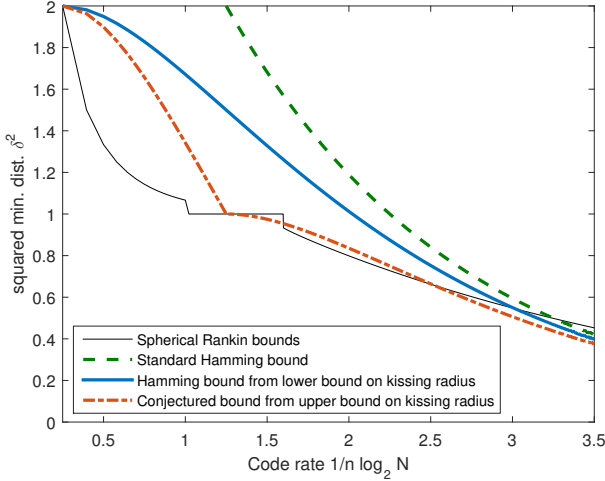


Fig. 6. The standard Hamming bound (27), the Hamming bound from the kissing radius analysis (65) and a conjectured improvement (67), compared to the Rankin bound for codes in  $\mathcal{G}_{4,2}^C$ .

known as the simplex and orthoplex bounds, can be easily inverted to bound the minimum distance. However, the third bound is expressed through an integral which we numerically inverted for Fig. 6. For  $\mathcal{G}_{4,2}^C$ , the Rankin bounds are tight at the saturating value  $N = 30$  [61], i.e. at rate 1.22. For higher cardinality the Rankin bounds are not reachable. As expected, the Hamming bounds become tighter than the Rankin bounds for large code sizes.

By construction the Hamming bound is loose even if one would be able to exactly compute the kissing radius  $\varrho$  as a function of minimum distance. The manifold can never be totally covered by packing balls (except for  $N = 2$ ), and one always has a gap between  $\varrho$  and  $r_N$ . In order to sharpen the Hamming bound one could use  $\bar{\varrho}$  rather than  $\underline{\varrho}$  to get a better estimate of the minimum distance. It is very likely that in general  $\bar{\varrho} \leq r_N$ , which can be verified at the two ends of the code size spectrum  $N = 2$  and  $N \rightarrow \infty$ . Inverting (59) gives the following approximation,

$$\delta^2 \lesssim \begin{cases} \lceil 2r_N^2 \rceil - \left( \lceil 2r_N^2 \rceil - 2r_N^2 \right)^2 & \text{for } \mathcal{M} = \mathcal{G}_{n,p}^C \\ 4\lceil \frac{r_N^2}{2} \rceil - 4\left( \lceil \frac{r_N^2}{2} \rceil - \frac{r_N^2}{2} \right)^2 & \text{for } \mathcal{M} = \mathcal{V}_{n,p}^C \end{cases}, \quad (67)$$

which is guaranteed to be an upper bound as  $N \rightarrow \infty$ . This conjectured improvement is displayed for  $\mathcal{G}_{4,2}^C$  in Fig. 6. Here,  $r_N$  is computed from (52) from [17] as in this case it is exact for any  $N \geq 2$ . Interestingly, it meets the Rankin orthoplex bound at  $N = 32$ , i.e. close to  $N = 30$  where this bound is saturating.

## VI. CONCLUSION

The density of Grassmann and Stiefel codes with chordal distance has been investigated. The analysis pertains to treat the codes as subclasses of spherical codes since the chordal distance induces embeddings in Euclidean hyperspheres. The investigation is motivated by an exotic behavior: in this context

maximizing the density of a code is not equivalent to maximizing its minimum distance. We addressed both critical steps to compute a code's density, which are computing the volume of balls and the kissing radius. For the volume of balls, our main results included a proper scaling of the manifold volume consistent with the chordal distance, which is needed in small ball approximation. Moreover, an asymptotic Gaussian behavior of chordal distances in the high-dimensional regime leading to an asymptotic equivalence between the volume of balls in the manifolds and the volume of caps in the embedding spheres was found. Then the kissing radius and density of codes were bounded as a function of the minimum distance of the codes. It was concluded that Stiefel and Grassmann codes have larger density than their image spherical codes in high dimensions. Finally, as a by-product of the analysis, refinements of the standard Hamming bounds for Stiefel and Grassmann codes were provided.

## APPENDIX A

### PROOF OF LEMMA 1—HYPERSPHERICAL CAP AREA

#### A. Proof of (31)

The area of a hyperspherical cap is given in [62] in which the radius of the cap is measured by an angle  $0 \leq \phi \leq \frac{\pi}{2}$ . This translates in our notation to a radius of  $r^2 = 2R^2(1 - \cos \phi)$ , that gives  $\sin^2 \phi = \frac{r^2}{R^2}(1 - \frac{r^2}{4R^2})$  and directly leads to

$$\text{vol}(C_{D,R}(r)) = \frac{A_D(R)}{2} I_{\frac{r^2}{R^2}(1 - \frac{r^2}{4R^2})} \left( \frac{D-1}{2}, \frac{1}{2} \right) \quad (68)$$

for  $r \leq \sqrt{2}R$ .

The result further simplifies by normalizing by the overall area of the sphere  $A_D(R)$  and using the identity  $I_x(a, a) = \frac{1}{2} I_{4x(1-x)}(a, \frac{1}{2})$  valid for  $0 \leq x \leq \frac{1}{2}$  [63, (8.17.6)] [64]. The volume for larger radius  $r \geq \sqrt{2}R$  can be computed by the complementary cap, i.e.

$$\sigma(C_{D,R}(r)) = 1 - \sigma(C_{D,R}(\sqrt{4R^2 - r^2})), \quad (69)$$

and using the identity  $I_x(a, b) = 1 - I_{1-x}(b, a)$  [63, (8.17.4)] [64]. Altogether, one obtains (31) for the whole range of the radius  $0 \leq r \leq 2R$ .

#### B. Proof of (32) – Small Cap

We start from (68) which after normalization gives

$$\sigma(C_{D,R}(r)) = \frac{1}{2} I_{\frac{r^2}{R^2}(1 - \frac{r^2}{4R^2})} \left( \frac{D-1}{2}, \frac{1}{2} \right). \quad (70)$$

Expressing the regularized incomplete beta function  $I_x(a, b) = B_x(a, b)/B(a, b)$  with the Beta function  $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$  and the incomplete Beta function  $B_x(a, b)$  gives

$$\sigma(C_{D,R}(r)) = \frac{1}{2\sqrt{\pi}} \frac{\Gamma(\frac{D}{2})}{\Gamma(\frac{D-1}{2})} B_{\frac{r^2}{R^2}(1 - \frac{r^2}{4R^2})} \left( \frac{D-1}{2}, \frac{1}{2} \right). \quad (71)$$

From the identity  $B_x(a, b) = \frac{x^a}{a} F(a, 1-b; a+1; x)$  [63, (8.17.7)] where  $F(a, b; c; z)$  is the hypergeometric function

given by [63, (15.2.1)], it can be verified that  $B_x(a, b) = \frac{x^a}{a}(1 + O(x))$  as  $x \rightarrow 0$ , which in turn gives

$$\sigma(C_{D,R}(r)) = \frac{1}{2\sqrt{\pi}} \frac{\Gamma(\frac{D}{2})}{\Gamma(\frac{D+1}{2})} \frac{r^{D-1}}{R^{D-1}} (1 + O(r^2)) \quad (72)$$

as  $r \rightarrow 0$ .

### C. Proof of (33) – Large Dimension

The normalized volume of a cap  $C_{D,R}(r)$  can be interpreted as the probability that a uniformly distributed vector  $\mathbf{x} \in \mathcal{S}^{D-1}(R)$  is at distance less than  $r$  from a fixed point  $\mathbf{e}$ . Without loss of generality, we choose the center of the cap to be  $\mathbf{e} = [R, 0, \dots, 0]^T$  and express the chordal distance by  $d_c^2(\mathbf{e}, \mathbf{x}) = 2R^2 \left(1 - x_{1,D}/\sqrt{D}\right)$  where  $x_{1,D} = \frac{\sqrt{D}}{R^2} \langle \mathbf{e}, \mathbf{x} \rangle$  is the first coordinate of  $\mathbf{x}$  normalized by the sphere radius and scaled. Noting that

$$0 \leq d_c(\mathbf{e}, \mathbf{x}) \leq r \Leftrightarrow \sqrt{D} \geq x_{1,D} \geq \sqrt{D} \left(1 - \frac{r^2}{2R^2}\right),$$

the volume is

$$\begin{aligned} \sigma(C_{D,R}(r)) &= \Pr\{\mathbf{x} \in \mathcal{S}^{D-1}(R) \mid 0 \leq d_c(\mathbf{e}, \mathbf{x}) \leq r\} \\ &= F_{x_{1,D}}(\sqrt{D}) - F_{x_{1,D}}\left(\sqrt{D} \left(1 - \frac{r^2}{2R^2}\right)\right) \end{aligned} \quad (73)$$

where  $F_{x_{1,D}}$  is the cumulative distribution function of the random variable  $x_{1,D}$ . Note that since the constraint  $0 \leq d_c(\mathbf{e}, \mathbf{x})$  is always satisfied, one has  $F_{x_{1,D}}(\sqrt{D}) = 1$ .

With  $D \rightarrow \infty$ ,  $x_{1,D}$  converges in distribution to a standard Gaussian random variable  $\sim \mathcal{N}(0, 1)$  [34], [35], i.e.

$$\lim_{D \rightarrow \infty} F_{x_{1,D}}(z) = \frac{1}{2} \left(1 + \operatorname{erf}\left(\frac{z}{\sqrt{2}}\right)\right) \quad (74)$$

where

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (75)$$

is the Gauss error function. Therefore, with  $z = \sqrt{D} \left(1 - \frac{r^2}{2R^2}\right)$  fixed, we have

$$\begin{aligned} \lim_{D \rightarrow \infty} \sigma(C_{D,R}(r)) &= 1 - \lim_{D \rightarrow \infty} F_{x_{1,D}}(z) \\ &= \frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\sqrt{D} \left(1 - \frac{r^2}{2R^2}\right)\right). \end{aligned} \quad (76)$$

Finally, in order to provide a finite-size approximation that satisfies the basic property of a measure,  $\sigma(C_{D,R}(0)) = 0$ , in addition to converging to the asymptotic form above, we use in (73) the finite-size correction  $F_{x_{1,D}}(\sqrt{D}) \approx \frac{1}{2} \left(1 + \operatorname{erf}\left(\sqrt{D}\right)\right) \rightarrow 1$  which corresponds to approximating  $x_{1,D}$  to be Gaussian also in the finite-size regime. This leads to the given expression.

## APPENDIX B PROOFS OF OVERALL VOLUMES

$\mathcal{M}$  is an  $m$ -dimensional Riemann manifold with infinitesimal metric

$$ds^2 = \sum_{j,k=1}^m g_{jk} dx_j dx_k \quad (77)$$

in local coordinates  $\{x_j\}_{j=1}^m$ . The volume element of  $\mathcal{M}$  is [65]

$$d\omega = \sqrt{\det\{g_{jk}\}_{j,k=1}^m} dx_1 \dots dx_m. \quad (78)$$

Now, consider the Euclidean space of  $m \times m$  complex matrices with its canonical inner product  $(\mathbb{C}^{m \times m}, \langle \cdot, \cdot \rangle = \Re \operatorname{Tr} \cdot^H \cdot)$  from which the Riemann metrics considered are constructed. Given  $\mathbf{M} \in \mathbb{C}^{m \times m}$  we have

$$\begin{aligned} ds^2 &= \Re \operatorname{Tr}(d\mathbf{M}^H d\mathbf{M}) = \|d\mathbf{M}\|_F^2 \\ &= \sum_{j,k=1}^m \Re(d\mathbf{M}_{jk})^2 + \Im(d\mathbf{M}_{jk})^2. \end{aligned} \quad (79)$$

Restricting the metric to the space of skew-Hermitian matrices, with  $\mathbf{A} \in \mathfrak{u}(m)$  we get

$$ds_{\mathbf{A}}^2 = -\operatorname{Tr}(d\mathbf{A}^2) = \sum_{j=1}^m |d\mathbf{A}_{jj}|^2 + 2 \sum_{j < k} |d\mathbf{A}_{jk}|^2 \quad (80)$$

and the corresponding volume element

$$d\omega_{\mathbf{A}} = 2^{\frac{m(m-1)}{2}} \prod_{j=1}^m |d\mathbf{A}_{jj}| \prod_{j < k} \Re(d\mathbf{A}_{jk}) \Im(d\mathbf{A}_{jk}). \quad (81)$$

Note here that the off-diagonal elements are counted twice, leading to an overall scaling factor of  $2^{\frac{m(m-1)}{2}}$ . The overall volumes of the manifolds need to be scaled accordingly to be consistent with the chosen metric.

### A. Volume of Unitary Group

Given a unitary matrix  $\mathbf{U} \in \mathcal{U}_n$ , by differentiating  $\mathbf{U}^H \mathbf{U} = \mathbf{I}$ , we obtain

$$\mathbf{U}^H d\mathbf{U} + d\mathbf{U}^H \mathbf{U} = 0 \quad (82)$$

showing that the differential form  $\mathbf{U}^H d\mathbf{U}$  is skew-Hermitian. Due to the unitary invariance of the metric  $ds^2$ , its restriction to  $\mathcal{U}_n$  can be expressed in terms of the global form  $\mathbf{U}^H d\mathbf{U}$ . Then, the infinitesimal metric is

$$\begin{aligned} ds_{\mathbf{U}}^2 &= -\operatorname{Tr}(\mathbf{U}^H d\mathbf{U})^2 \\ &= \sum_{j=1}^n |(\mathbf{U}^H d\mathbf{U})_{jj}|^2 + 2 \sum_{j < k} |(\mathbf{U}^H d\mathbf{U})_{jk}|^2 \end{aligned} \quad (83)$$

and the volume form in local coordinates is

$$d\omega_{\mathbf{U}} = 2^{\frac{n(n-1)}{2}} d\nu_{\mathbf{U}}, \quad (84)$$

where

$$d\nu_{\mathbf{U}} = \prod_{j=1}^n |(\mathbf{U}^H d\mathbf{U})_{jj}| \prod_{j < k} \Re((\mathbf{U}^H d\mathbf{U})_{jk}) \Im((\mathbf{U}^H d\mathbf{U})_{jk}) \quad (85)$$

is a common volume element normalized so that  $\int_{\mathcal{U}_n} d\nu_{\mathbf{U}} = \frac{2^n \pi^{n^2}}{\tilde{\Gamma}_n(n)}$  [66], [67]. Finally with the metric considered,

$$\operatorname{vol} \mathcal{U}_n = \frac{2^{\frac{n(n+1)}{2}} \pi^{n^2}}{\tilde{\Gamma}_n(n)}, \quad (86)$$

where the complex multivariate gamma function is

$$\tilde{\Gamma}_p(n) = \pi^{\frac{p(p-1)}{2}} \prod_{i=1}^p \Gamma(n - i + 1). \quad (87)$$

### B. Volume for Stiefel Manifold in Theorem 1

Now let  $\mathbf{Y} \in \mathcal{V}_{n,p}^{\mathbb{C}}$  and  $\mathbf{U} \in \mathcal{U}_n$  such that  $\mathbf{U}^H \mathbf{Y} = \mathbf{I}_{n,p}$ , i.e. the first  $p$  columns of  $\mathbf{U} = (\mathbf{Y} \ \mathbf{Y}^\perp)$  are the columns of  $\mathbf{Y}$ . The differential form  $\mathbf{U}^H d\mathbf{Y}$  is “rectangular skew-Hermitian”, i.e.  $\mathbf{U}^H d\mathbf{Y} = \begin{pmatrix} \mathbf{Y}^H d\mathbf{Y} \\ \mathbf{Y}^{\perp H} d\mathbf{Y} \end{pmatrix}$  where  $\mathbf{Y}^H d\mathbf{Y}$  is  $p$ -by- $p$  skew-Hermitian. Similarly, due to unitary invariance of the metric, the volume element for the Stiefel manifold can be expressed in terms of the global form  $\mathbf{U}^H d\mathbf{Y}$  and is given in local coordinates as

$$d\omega_{\mathbf{Y}} = 2^{\frac{p(p-1)}{2}} d\nu_{\mathbf{Y}}, \quad (88)$$

where

$$d\nu_{\mathbf{Y}} = \prod_{i=1}^p |(U^H d\mathbf{Y})_{ii}| \times \prod_{k=1}^p \prod_{j=k+1}^n \Re((U^H d\mathbf{Y})_{jk}) \Im((U^H d\mathbf{Y})_{jk}) \quad (89)$$

is a common volume element normalized so that  $\int_{\mathcal{V}_{n,p}^{\mathbb{C}}} d\nu_{\mathbf{Y}} = \frac{2^p \pi^{np}}{\Gamma_p(n)}$  [66], [67]. Finally,

$$\text{vol } \mathcal{V}_{n,p}^{\mathbb{C}} = \frac{2^{\frac{p(p+1)}{2}} \pi^{np}}{\tilde{\Gamma}_p(n)}. \quad (90)$$

### C. Volume for Grassmann Manifold

The volume of the Grassmann manifold directly follows from the quotient geometry over the Stiefel manifold associated with  $d_g$ :

$$\text{vol } \mathcal{G}_{n,p}^{\mathbb{C}} = \frac{\text{vol } \mathcal{V}_{n,p}^{\mathbb{C}}}{\text{vol } \mathcal{U}_p}. \quad (91)$$

Alternatively it can be computed from the quotient geometry over the unitary group associated with  $d_{g^*}$

$$\text{vol } \mathcal{G}_{n,p}^{\mathbb{C}} = 2^{-\frac{\dim}{2}} \frac{\text{vol } \mathcal{U}_n}{\text{vol } \mathcal{U}_p \text{vol } \mathcal{U}_{n-p}}, \quad (92)$$

where the scaling coefficient comes from the  $\sqrt{2}$  in the definition of  $d_{g^*}$  compared to  $d_g$ .

### APPENDIX C

#### PROOF OF LEMMA 2—COMPLEMENTARY BALLS

a) *Stiefel Manifold*: Consider the center to be  $\mathbf{I}_{n,p} \in \mathcal{V}_{n,p}^{\mathbb{C}}$ . It has a unique antipodal point (i.e. a farthest possible point from  $\mathbf{I}_{n,p}$ ) which is  $-\mathbf{I}_{n,p}$  and we have

$$d_c(\mathbf{I}_{n,p}, -\mathbf{I}_{n,p}) = 2\sqrt{p} \triangleq d_{\max}. \quad (93)$$

Given a point  $\mathbf{Y} \in \mathcal{V}_{n,p}^{\mathbb{C}}$  such that  $d_c(\mathbf{I}, \mathbf{Y}) \geq r$ , it is a direct verification that  $d_c(-\mathbf{I}, \mathbf{Y}) \leq \sqrt{d_{\max}^2 - r^2}$ , thus  $\mathbf{Y} \notin B_{\mathbf{I}}(r)$  implies  $\mathbf{Y} \in B_{-\mathbf{I}}(\sqrt{d_{\max}^2 - r^2})$  and finally

$$\mu(B(r)) + \mu(B(\sqrt{d_{\max}^2 - r^2})) = 1. \quad (94)$$

b) *Grassmann Manifold*: From the mapping (18), the Grassmann manifolds  $\mathcal{G}_{n,p}^{\mathbb{C}}$  and  $\mathcal{G}_{n,n-p}^{\mathbb{C}}$  are both embedded in the same sphere  $\mathcal{S}^{n^2-2}(\sqrt{\frac{p(n-p)}{2n}})$ . The definition of chordal distance thus directly extends between any  $[\mathbf{Y}] \in \mathcal{G}_{n,p}^{\mathbb{C}}$  and  $[\mathbf{Z}] \in \mathcal{G}_{n,n-p}^{\mathbb{C}}$  as

$$\begin{aligned} d_c^2([\mathbf{Y}], [\mathbf{Z}]) &= \|\bar{\Pi}_{\mathbf{Y}} - \bar{\Pi}_{\mathbf{Z}}\|_F^2 \\ &= \frac{2p(n-p)}{n} - \|\mathbf{Y}^H \mathbf{Z}\|_F^2 \end{aligned} \quad (95)$$

The distance reaches its maximum at  $d_{\max} = \sqrt{\frac{2p(n-p)}{n}}$  with  $[\mathbf{Z}] = [\mathbf{Y}_\perp]$ , where the  $\mathbf{Y}_\perp$  is the orthogonal complement of  $\mathbf{Y}$  such that  $(\mathbf{Y} \ \mathbf{Y}_\perp)$  is a unitary matrix. The maximum distance is exactly twice the embedding radius,  $d_{\max} = 2R$ , i.e.  $[\mathbf{Y}_\perp]$  is the antipodal of  $[\mathbf{Y}]$  on the embedding sphere. The final result follows by the same argumentation as for the Stiefel manifold. Here the Pythagorean theorem gives  $d_c^2([\mathbf{Y}], [\mathbf{Z}]) + d_c^2([\mathbf{Y}_\perp], [\mathbf{Z}]) = \frac{2p(n-p)}{n}$ .

### APPENDIX D

#### PROOF OF THE ASYMPTOTIC GAUSSIANITY OF THE CHORDAL DISTANCES IN LEMMA 3

Before proceeding the proof, we first state some useful definitions and intermediary results.

#### A. The Hypergeometric Function of Complex Matrix Argument

For an  $n \times n$  Hermitian matrix  $\mathbf{A}$ , the hypergeometric function of a complex matrix argument is defined as [68], [69]

$${}_p\tilde{F}_q(a_1, \dots, a_p; b_1, \dots, b_q; \mathbf{A}) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_{\kappa} \cdots (a_p)_{\kappa}}{(b_1)_{\kappa} \cdots (b_q)_{\kappa}} \frac{C_{\kappa}(\mathbf{A})}{k!}, \quad (96)$$

where  $\kappa$  denotes a partition of integer  $k$  into no more than  $n$  parts, i.e.  $k = \kappa_1 + \kappa_2 + \cdots + \kappa_n$  with  $\kappa_1 \geq \kappa_2 \geq \cdots \geq \kappa_n \geq 0$ , the sum is over all partitions, and

$$(a)_{\kappa} = \prod_{j=1}^n (a - j + 1)_{\kappa_j} = \prod_{j=1}^n \frac{(\kappa_j + a - j)!}{(a - j)!} \quad (97)$$

is the multivariate hypergeometric coefficient [68, Eq. (84)]. In (96),  $C_{\kappa}(\mathbf{A})$  denotes a zonal polynomial [68], [69], which is a homogenous symmetric polynomial of degree  $k$  in the  $n$  eigenvalues of  $\mathbf{A}$ . Denoting the  $j$ -th eigenvalue of  $\mathbf{A}$  by  $a_j$ , the zonal polynomial can be represented as [68, Eq. (85)],

$$C_{\kappa}(\mathbf{A}) = \chi_{\kappa}(1) \chi_{\kappa}(\mathbf{A}), \quad (98)$$

where

$$\chi_{\kappa}(1) = \frac{k! \prod_{1 \leq i < j \leq n} (\kappa_i - \kappa_j - i + j)}{\prod_{j=1}^n (\kappa_j + n - j)!} \quad (99)$$

and

$$\chi_{\kappa}(\mathbf{A}) = \frac{\det(a_i^{\kappa_j + n - j})}{\det(a_i^{n - j})} \quad (100)$$

is a Schur polynomial<sup>5</sup>. Schur polynomials form a basis in the space of homogeneous symmetric polynomials in  $n$  variables of degree  $k$  for all  $k \leq n$ . In particular, we have [68, Eq. (17)]

$$(\text{Tr} \mathbf{A})^k = \sum_{\kappa} C_{\kappa}(\mathbf{A}). \quad (101)$$

We will need the following identity<sup>6</sup> [69, Eq. (6.2.3)]

$$\begin{aligned} & \int_{\mathbf{X}} \text{etr}(-\mathbf{XZ}) (\det(\mathbf{X}))^{n-p} \\ & \quad \times {}_r\tilde{F}_s(a_1, \dots, a_r; b_1, \dots, b_s; \mathbf{X}) d\mathbf{X} = \\ & {}_{r+1}\tilde{F}_s(a_1, \dots, a_r, n; b_1, \dots, b_s; \mathbf{Z}^{-1}) (\det(\mathbf{Z}))^{-n} \tilde{\Gamma}_p(n), \end{aligned} \quad (102)$$

where  $\mathbf{X}, \mathbf{Z}$  are  $p \times p$  Hermitian matrices and  $\tilde{\Gamma}_p(n)$  is defined in (87).

We will prove the following lemma.

**Lemma 7.** *For any  $p \times n$  complex matrix  $\mathbf{S}$ , we have*

$$\int_{\mathbf{Y} \in \mathcal{V}_{n,p}^c} \text{etr}(2\Re(\mathbf{SY})) d\mu(\mathbf{Y}) = {}_0\tilde{F}_1(n; \mathbf{SS}^H). \quad (103)$$

where  $d\mu(\mathbf{Y})$  is the uniform measure.

*Proof:* First, note that it is equivalent to show that

$$\begin{aligned} & (\det \mathbf{SS}^H)^{n-p} \int_{\mathbf{Y}} \text{etr}(\mathbf{SY} + \mathbf{Y}^H \mathbf{S}^H) d\mu(\mathbf{Y}) = \\ & (\det \mathbf{SS}^H)^{n-p} {}_0\tilde{F}_1(n; \mathbf{SS}^H). \end{aligned} \quad (104)$$

We will show that the matrix-variate Laplace transforms of both sides of the above equation are the same. The Laplace transform of the left-hand-side (LHS) is

$$\begin{aligned} T_{\text{LHS}}(\mathbf{Z}) &= \int_{\mathbf{SS}^H} \text{etr}(-\mathbf{SS}^H \mathbf{Z}) (\det \mathbf{SS}^H)^{n-p} \\ & \quad \times \int_{\mathbf{Y}} \text{etr}(\mathbf{SY} + \mathbf{Y}^H \mathbf{S}^H) d\mu(\mathbf{Y}) d(\mathbf{SS}^H) \\ &= \frac{\tilde{\Gamma}_p(n)}{\pi^{np}} \int_{\mathbf{Y}} \int_{\mathbf{S}} \text{etr}(-\mathbf{SS}^H \mathbf{Z} \\ & \quad + \mathbf{SY} + \mathbf{Y}^H \mathbf{S}^H) d\mathbf{S} d\mu(\mathbf{Y}). \end{aligned} \quad (105)$$

The equality above is established by utilizing the decomposition [69, Th. 4.5]  $\mathbf{S} = \mathbf{LU}$  where  $\mathbf{L}$  is a  $p \times p$  lower triangular matrix with positive diagonal elements and  $\mathbf{U}^H \in \mathcal{V}_{n,p}^c$ , which leads to the fact that [69, Coroll. 4.5.3]

$$d\mathbf{S} = 2^{-p} (\det \mathbf{SS}^H)^{n-p} d(\mathbf{SS}^H) d\nu_{\mathbf{U}}, \quad (106)$$

where here [69, Coroll. 4.5.2] the measure on the Stiefel manifold is normalized so that  $\int_{\mathcal{V}_{n,p}^c} d\nu_{\mathbf{U}} = \frac{2^p \pi^{np}}{\tilde{\Gamma}_p(n)}$ . Applying the transform  $\mathbf{S} = \mathbf{Z}^{-1/2} \mathbf{T}$  with  $d\mathbf{S} = (\det \mathbf{Z})^{-n} d\mathbf{T}$

in (105), we have

$$\begin{aligned} T_{\text{LHS}}(\mathbf{Z}) &= \frac{\tilde{\Gamma}_p(n) (\det \mathbf{Z})^{-n}}{\pi^{np}} \int_{\mathbf{Y}} \int_{\mathbf{T}} \text{etr} \left( -\mathbf{T} \mathbf{T}^H \right. \\ & \quad \left. + \mathbf{Z}^{-1/2} \mathbf{T} \mathbf{Y} + \mathbf{Y}^H \mathbf{T}^H (\mathbf{Z}^{-1/2})^H \right) d\mathbf{T} d\mu(\mathbf{Y}) \\ &= \frac{\tilde{\Gamma}_p(n) (\det \mathbf{Z})^{-n} \text{etr}(\mathbf{Z}^{-1})}{\pi^{np}} \times \\ & \quad \int_{\mathbf{Y}} \int_{\mathbf{T}} \text{etr} \left( -\left( \mathbf{T} - \mathbf{Y}^H (\mathbf{Z}^{-1/2})^H \right) \right. \\ & \quad \left. \left( \mathbf{T} - \mathbf{Y}^H (\mathbf{Z}^{-1/2})^H \right)^H \right) d\mathbf{T} d\mu(\mathbf{Y}) \\ &= \tilde{\Gamma}_p(n) (\det \mathbf{Z})^{-n} \text{etr}(\mathbf{Z}^{-1}), \end{aligned} \quad (107)$$

where the last step is established by the fact that  $p(\mathbf{X}) = \frac{1}{\pi^{np}} \text{etr} \left( -(\mathbf{X} - \mathbf{M})(\mathbf{X} - \mathbf{M})^H \right)$  is a matrix-variate Gaussian density function.

The Laplace transform of the right-hand side (RHS) of (104) is

$$\begin{aligned} T_{\text{RHS}}(\mathbf{Z}) &= \int_{\mathbf{SS}^H} \text{etr}(-\mathbf{SS}^H \mathbf{Z}) (\det(\mathbf{SS}^H))^{n-p} \\ & \quad \times {}_0\tilde{F}_1(n; \mathbf{SS}^H) d(\mathbf{SS}^H) \\ &= \tilde{\Gamma}_p(n) (\det(\mathbf{Z}))^{-n} {}_1\tilde{F}_1(n; n; \mathbf{Z}^{-1}) \\ &= \tilde{\Gamma}_p(n) (\det(\mathbf{Z}))^{-n} {}_0\tilde{F}_0(\mathbf{Z}^{-1}) \\ &= \tilde{\Gamma}_p(n) (\det(\mathbf{Z}))^{-n} \text{etr}(\mathbf{Z}^{-1}) \\ &= T_{\text{LHS}}(\mathbf{Z}), \end{aligned} \quad (108)$$

where the second equality is obtained by the identity (102). By the uniqueness of Laplace transforms, we complete the proof of the lemma.

### B. Stiefel Manifold

We now prove Lemma 3 in the case of the Stiefel manifold. The expansion of the chordal distance in terms of an inner product in the ambient space gives

$$d_c^2(\mathbf{I}_{n,p}, \mathbf{Y}) = \|\mathbf{I}_{n,p} - \mathbf{Y}\|_F^2 = 2p - 2\Re \text{Tr}(\mathbf{I}_{n,p}^H \mathbf{Y}). \quad (109)$$

Accordingly, define the linear statistic

$$Y_n = \sqrt{\frac{2n}{p}} \Re \text{Tr}(\mathbf{I}_{n,p}^H \mathbf{Y}), \quad (110)$$

so that the distance is  $d_c^2(\mathbf{I}_{n,p}, \mathbf{Y}) = 2p \left( 1 - \frac{1}{\sqrt{2pn}} Y_n \right)$ .

This type of linear statistic converges in distribution to a standard Gaussian random variable as  $n$  approaches infinity [35]. To see this, consider the moment generating function of  $Y_n$  which can be represented as a hypergeometric function of matrix argument using Lemma 7 with  $\mathbf{S} = \sqrt{\frac{n}{2p}} \nu \mathbf{I}_{n,p}^H$ ,

$$\begin{aligned} \mathbb{E} [e^{\nu Y_n}] &= {}_0\tilde{F}_1 \left( n; \frac{n\nu^2}{2p} \mathbf{I}_{n,p}^H \mathbf{I}_{n,p} \right) \\ &= {}_0\tilde{F}_1 \left( n; \frac{n\nu^2}{2p} \mathbf{I}_p \right). \end{aligned} \quad (111)$$

Note that it corresponds to the partition function of the von-Mises Fisher distribution [55] with parameter matrix  $\sqrt{\frac{2n}{p}} \nu \mathbf{I}_{n,p}$ .

<sup>5</sup>When some eigenvalues of  $\mathbf{A}$  are equal, the corresponding Schur polynomials (100) are obtained by using l'Hospital's rule.

<sup>6</sup> $\text{etr}(\cdot) = e^{\text{Tr}(\cdot)}$  denotes exponential of trace.



By the definition of the hypergeometric function (96), we can further write

$$\begin{aligned}\mathbb{E}[e^{\nu Y_n}] &= \sum_{k=0}^{\infty} \sum_{\kappa} \frac{1}{(n)_{\kappa}} \frac{C_{\kappa} \left( \frac{n\nu^2}{2p} \mathbf{I}_p \right)}{k!} \\ &= \sum_{k=0}^{\infty} \frac{\left( \frac{n\nu^2}{2p} \right)^k}{k!} \sum_{\kappa} \frac{C_{\kappa}(\mathbf{I}_p)}{(n)_{\kappa}},\end{aligned}\quad (112)$$

where the last equality is established by (100). Since the leading order term in  $(n-j+1)_{\kappa_j}$  equals  $n^{\kappa_j}$ , by (97), for large  $n$  we have

$$(n)_{\kappa} = \prod_{j=1}^n (n-j+1)_{\kappa_j} \sim n^{\kappa_1 + \dots + \kappa_n} = n^k. \quad (113)$$

Using (101), the sum in (112) for large  $n$  becomes

$$\sum_{\kappa} \frac{C_{\kappa}(\mathbf{I}_p)}{(n)_{\kappa}} \xrightarrow{n \rightarrow \infty} \frac{1}{n^k} \sum_{\kappa} C_{\kappa}(\mathbf{I}_p) = \frac{1}{n^k} \text{Tr}^k(\mathbf{I}_p) = \frac{p^k}{n^k}, \quad (114)$$

and we arrive at the result,

$$\lim_{n \rightarrow \infty} \mathbb{E}[e^{\nu Y_n}] = \sum_{k=0}^{\infty} \frac{\left( \frac{\nu^2}{2} \right)^k}{k!} = e^{\frac{\nu^2}{2}}, \quad (115)$$

which is the moment generating function of a zero mean and unit variance Gaussian distribution.

Furthermore, we can directly identify in (112) the finite-size moments of  $Y_n$  from the series expansion of the moment generating function  $\mathbb{E}[e^{\nu Y_n}] = \sum_{l=0}^{\infty} \frac{\nu^l}{l!} \mathbb{E}[Y_n^l]$ . For any  $n$ , the mean is  $\mathbb{E}[Y_n] = 0$  since there are no odd powers of  $\nu$  in (112). The variance follows from the  $(k=1)$ -term in (112) for which there is only one partition  $\kappa = \{1, 0, \dots, 0\}$  such that  $(n)_{\kappa} = n$ ,  $\chi_{\kappa}(1) = 1$ ,  $C_{\kappa}(\mathbf{I}_p) = \chi_{\kappa}(\mathbf{I}_p) = p$ , and thus  $\text{Var}[Y_n] = 1$  also for any finite  $n$ .

As a byproduct, this convergence can be written in terms of the chordal distance as

$$\sqrt{\frac{n}{2p}} (d_c^2(\mathbf{I}_{n,p}, \mathbf{Y}) - 2p) \xrightarrow{d} \mathcal{N}(0, 1) \text{ as } n \rightarrow \infty, \quad (116)$$

and where the finite-size regime mean and variance are exactly  $\mathbb{E}[d_c^2(\mathbf{I}_{n,p}, \mathbf{Y})] = 2p$  and  $\text{Var}[d_c^2(\mathbf{I}_{n,p}, \mathbf{Y})] = \frac{2p}{n}$ .

### C. Grassmann Manifold

The case of the Grassmann manifold can be deduced as a by product of the volume computation in [33] by setting  $q = p$  (i.e. the dimension of the center of the ball and the elements in the ball have the same dimension). We provide below alternative lines of derivation from the hypergeometric function interpretation consistent with the Stiefel case discussed above.

If  $\mathbf{Y}$  is uniformly distributed on the Stiefel manifold, then  $[\mathbf{Y}]$  is uniformly distributed on the Grassmann manifold. A point  $[\mathbf{Y}]$  on the Grassmann manifold can be uniquely defined by its projection matrix  $\Pi_{\mathbf{Y}} = \mathbf{Y}\mathbf{Y}^H$ . Looking at the uniform measure as a probability measure, the mapping  $\mathbf{Y} \rightarrow \Pi_{\mathbf{Y}}$  maps the uniform distribution on  $\mathcal{V}_{n,p}^{\mathbb{C}}$  to the uniform distribution on  $\mathcal{G}_{n,p}^{\mathbb{C}}$ . Namely, if  $\mathbf{Y}$  has the same distribution as  $\mathbf{U}\mathbf{Y}$  for all  $\mathbf{U} \in \mathcal{U}_n$ , then  $\Pi_{\mathbf{U}\mathbf{Y}}$  has the

same distribution as  $\mathbf{U}\Pi_{\mathbf{Y}}\mathbf{U}^H$  for all  $\mathbf{U} \in \mathcal{U}_n$ . Accordingly, consider

$$\begin{aligned}d_c^2([\mathbf{I}_{n,p}], [\mathbf{Y}]) &= \frac{1}{2} \|\mathbf{E} - \Pi_{\mathbf{Y}}\|_F^2 \\ &= p - \text{Tr}(\mathbf{Y}^H \mathbf{E} \mathbf{Y})\end{aligned}\quad (117)$$

with  $\mathbf{E} = \mathbf{I}_{n,p} \mathbf{I}_{n,p}^H$  and such that  $\mathbf{Y}$  is uniformly distributed on the Stiefel manifold  $\mathcal{V}_{n,p}^{\mathbb{C}}$ .

The chordal distance can thus be expressed as a function of a linear statistic  $Z_{n,p} = \text{Tr}(\mathbf{Y}^H \mathbf{E} \mathbf{Y})$ . This type of linear statistic is asymptotically Gaussian for  $n, p \rightarrow \infty$  and  $n - 2p$  constant as shown in [33]<sup>7</sup> through its interpretation as the sum of squared principal cosines distributed according to the Jacobi ensemble [70, Sec. 2.1.2.] and using [71, Th. 2].

The moment-generating function of  $Z_{n,p}$  corresponds to the partition function of the Bingham distribution [55] with parameter matrix  $\nu \mathbf{E}$ , which can be expressed as a confluent hypergeometric function of complex matrix argument. This can be verified by matrix-variate Laplace transforms, as in the proof of Lemma 7. It follows that the moments of  $Z_{n,p}$  can also be computed by identification from the definition of the hypergeometric function (96):

$$\begin{aligned}\mathbb{E}[e^{\nu Z_{n,p}}] &= \sum_{k=0}^{\infty} \frac{\nu^k}{k!} \mathbb{E}[Z_{n,p}^k] \\ &= {}_1\tilde{F}_1(p, n; \nu \mathbf{E}) \\ &= \sum_{k=0}^{\infty} \frac{\nu^k}{k!} \sum_{\kappa} \frac{(p)_{\kappa}}{(n)_{\kappa}} C_{\kappa}(\mathbf{E}).\end{aligned}\quad (118)$$

The mean of  $Z_{n,p}$  is given for  $k=1$  for which there is only one partition such that  $\chi_{\kappa}(\mathbf{E}) = p$ , and  $\mathbb{E}[Z_{n,p}] = \frac{p^2}{n}$ . For the second moments with  $k=2$ , the possible partitions are  $\kappa = \{2, 0, \dots, 0\}$  and  $\kappa = \{1, 1, 0, \dots, 0\}$ . The corresponding Schur polynomials for each partition are  $\chi_{\kappa}(\mathbf{E}) = \frac{p}{2}(p+1)$  and  $\chi_{\kappa}(\mathbf{E}) = \frac{p}{2}(p-1)$ , respectively; while for both partitions one has  $\frac{(p)_{\kappa}}{(n)_{\kappa}} = \frac{p(p+1)}{n(n+1)}$  and  $\chi_{\kappa}(1) = 1$ . One obtains after some manipulations  $\mathbb{E}[Z_{n,p}^2] = \frac{p^2(np^2 - 2p + n)}{n(n^2 - 1)}$  and thus the variance  $\text{Var}[Z_{n,p}] = \frac{p^2(n-p)^2}{n^2(n^2 - 1)}$ . Therefore, we have

$$\frac{n\sqrt{n^2 - 1}}{p(n-p)} \left( Z_{n,p} - \frac{p^2}{n} \right) \xrightarrow{d} \mathcal{N}(0, 1).$$

Equivalently, the squared chordal distance for the Grassmann manifold converges in distribution to a Gaussian random variable as

$$\frac{n\sqrt{n^2 - 1}}{p(n-p)} \left( d_c^2([\mathbf{I}_{n,p}], [\mathbf{Y}]) - \frac{p(n-p)}{n} \right) \xrightarrow{d} \mathcal{N}(0, 1)$$

with  $n, p \rightarrow \infty$  and  $n - 2p$  constant, and its finite-regime mean and variance are exactly  $\mathbb{E}[d_c^2([\mathbf{I}_{n,p}], [\mathbf{Y}])] = \frac{p(n-p)}{n}$  and  $\text{Var}[d_c^2([\mathbf{I}_{n,p}], [\mathbf{Y}])] = \frac{p^2(n-p)^2}{n^4 - n^2}$ .

<sup>7</sup>One has  $Z_{n,p} = \sum_{i=1}^p \cos^2 \theta_i$  from the definition of the chordal distance (17). Computing the moment-generating function, one gets  $\mathbb{E}[e^{\nu Z_{n,p}}] = e^{\nu p} D_p(i\nu)$  where  $D_p(\nu)$  is defined in [33, Eq. (22)].

## APPENDIX E

## PROOF OF THE VOLUME COMPARISON IN THEOREM 2

Now with Lemma 3, we are in position to prove Theorem 2. The result follows by identification of the asymptotic forms between the manifold and the embedding sphere.

## A. Stiefel Manifold

By choosing without loss of generality  $\mathbf{I}_{n,p}$  as the center of the ball and noting that

$$0 \leq d_c(\mathbf{I}_{n,p}, \mathbf{Y}) \leq r \Leftrightarrow \sqrt{2np} \geq Y_n \geq \sqrt{\frac{n}{2p}}(2p - r^2)$$

where  $Y_n = \sqrt{\frac{n}{2p}}(2p - d_c^2(\mathbf{I}_{n,p}, \mathbf{Y}))$ ; identification of the normalized volume of the ball as a probability measure gives

$$\begin{aligned} \mu(B(r)) &= \Pr\{\mathbf{Y} \in \mathcal{V}_{n,p}^{\mathbb{C}} \mid 0 \leq d_c(\mathbf{I}_{n,p}, \mathbf{Y}) \leq r\} \\ &= F_{Y_n}(\sqrt{2np}) - F_{Y_n}\left(\sqrt{2np}\left(1 - \frac{r^2}{2p}\right)\right) \end{aligned} \quad (119)$$

where  $F_{Y_n}(z)$  is the cumulative distribution of the random variable  $Y_n$  with  $\mathbf{Y}$  being uniformly distributed on  $\mathcal{V}_{n,p}^{\mathbb{C}}$ . Note that  $F_{Y_n}(\sqrt{2np}) = 1$  by definition of the chordal distance.

As shown via Lemma 3,  $Y_n$  converges in distribution to a standard Gaussian random variable, and thus with  $n \rightarrow \infty$  and  $z = \sqrt{2np}\left(1 - \frac{r^2}{2p}\right)$  fixed, one has

$$\begin{aligned} \lim_{n \rightarrow \infty} \mu(B(r)) &= 1 - \lim_{n \rightarrow \infty} F_{Y_n}(z) \\ &= \frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\sqrt{np}\left(1 - \frac{r^2}{2p}\right)\right). \end{aligned} \quad (120)$$

The final asymptotic expression is obtained by using the finite-size correction  $F_{Y_n}(\sqrt{2np}) \approx \frac{1}{2}(1 + \operatorname{erf}(\sqrt{2np})) \rightarrow 1$  in (119) in order to provide an approximation that simultaneously satisfies  $\mu(B(0)) = 0$  and the asymptotic limit (120). This leads to (49).

Finally, one verifies that the asymptotic volume of a ball in (49) is exactly the asymptotic volume of a cap in the embedding sphere  $\mathcal{V}_{n,p}^{\mathbb{C}} \hookrightarrow \mathcal{S}^{2np-1}(\sqrt{p})$  by identification with  $D = 2np$  and  $R = \sqrt{p}$  in (33).

## B. Grassmann Manifold

Again from the results of [33], one may directly obtain (50) by setting  $q = p$ . Alternatively, the volume of a ball of radius  $r$  centered around  $[\mathbf{I}_{n,p}]$  can be expressed as

$$\begin{aligned} \mu(B(r)) &= \Pr\{\mathbf{Y} \in \mathcal{G}_{n,p}^{\mathbb{C}} \mid 0 \leq d_c([\mathbf{I}_{n,p}], [\mathbf{Y}]) \leq r\} \\ &= \Pr\{\mathbf{Y} \in \mathcal{V}_{n,p}^{\mathbb{C}} \mid 0 \leq d_c([\mathbf{I}_{n,p}], [\mathbf{Y}]) \leq r\} \\ &= F_{\tilde{Z}_{n,p}}(\sqrt{n^2 - 1}) \\ &\quad - F_{\tilde{Z}_{n,p}}\left(\sqrt{\frac{n^2 - 1}{2}}\left(1 - \frac{n r^2}{p(n-p)}\right)\right), \end{aligned} \quad (121)$$

where  $F_{\tilde{Z}_{n,p}}(z)$  is the cumulative distribution of the random variable  $\tilde{Z}_{n,p} = \frac{n\sqrt{n^2-1}}{p(n-p)}\left(\frac{p(n-p)}{n} - d_c^2([\mathbf{I}_{n,p}], [\mathbf{Y}])\right)$  such that  $\mathbf{Y}$  is uniformly distributed on the Stiefel manifold  $\mathcal{V}_{n,p}^{\mathbb{C}}$ .

As shown in Lemma 3,  $\tilde{Z}_{n,p}$  converges in distribution to a standard Gaussian random variable as  $n, p \rightarrow \infty$  with  $(n-2p)$  fixed. Thus with  $z = \sqrt{\frac{n^2-1}{2}}\left(1 - \frac{n r^2}{p(n-p)}\right)$  also fixed, one has

$$\begin{aligned} \lim_{n,p \rightarrow \infty} \mu(B(r)) &= 1 - \lim_{n,p \rightarrow \infty} F_{\tilde{Z}_{n,p}}(z) \\ &= \frac{1}{2} - \frac{1}{2} \operatorname{erf}\left(\sqrt{\frac{n^2-1}{2}}\left(1 - \frac{n r^2}{p(n-p)}\right)\right). \end{aligned} \quad (122)$$

Again, the final asymptotic expression (50) is obtained by using the finite-size correction  $F_{\tilde{Z}_{n,p}}(\sqrt{n^2-1}) \approx \frac{1}{2}(1 + \operatorname{erf}(\sqrt{n^2-1})) \rightarrow 1$  in (121) in order to provide a volume approximation that simultaneously satisfies  $\mu(B(0)) = 0$  and the asymptotic condition (122).

Finally, one verifies also here that the asymptotic volume of a ball in (50) equals the asymptotic volume of a cap in the embedding sphere  $\mathcal{G}_{n,p}^{\mathbb{C}} \hookrightarrow \mathcal{S}^{n^2-2}\left(\sqrt{\frac{p(n-p)}{2n}}\right)$  by identification with  $D = n^2 - 1$  and  $R = \sqrt{\frac{p(n-p)}{2n}}$  in (33).

## APPENDIX F

## PROOF OF KISSING RADIUS BOUNDS IN PROPOSITION 1

Consider a code  $\mathcal{C} = \{\mathbf{C}_1, \dots, \mathbf{C}_N\}$  with pairwise distances among the codewords  $\{\delta_{i,j}\}_{i \neq j}$  such that  $\delta = \min\{\delta_{i,j}\}$ . Between each codeword there is a different mid-distance  $\{\varrho_{i,j}\}_{i \neq j}$  and the kissing radius  $\varrho = \min \varrho_{i,j}$ . The detailed derivations below provide a lower and upper bound on a mid-distance  $\underline{\varrho}(\delta_{i,j}) \leq \varrho_{i,j} \leq \bar{\varrho}(\delta_{i,j})$  as function of the distance  $\delta_{i,j}$ . It can be verified that the obtained bounds  $\underline{\varrho}$  and  $\bar{\varrho}$  are increasing functions. It then follows that  $\min\{\underline{\varrho}(\delta_{i,j})\} = \underline{\varrho}(\min\{\delta_{i,j}\})$  and  $\min\{\bar{\varrho}(\delta_{i,j})\} = \bar{\varrho}(\min\{\delta_{i,j}\})$ , and one has  $\underline{\varrho}(\delta) \leq \varrho \leq \bar{\varrho}(\delta)$ .

## A. Grassmann Manifold

The principal angles  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)$  between two codewords  $\mathbf{C}_i, \mathbf{C}_j$  separated by  $\delta$  satisfies  $\sum_{i=1}^p \sin^2(\theta_i) = \delta^2$ . Without loss of generality, the code may be rotated so that the Stiefel representatives of these codewords are of the form  $\mathbf{C}_i = \mathbf{I}_{n,p}$  and  $\mathbf{C}_j = [(\operatorname{diag}(\cos \boldsymbol{\theta}) \operatorname{diag}(\sin \boldsymbol{\theta}))^T]$  [13].

The chordal distance is measured along the geodesic. The principal angles between the midpoint  $\mathbf{M}_{i,j}$  (on the geodesic joining  $\mathbf{C}_i$  and  $\mathbf{C}_j$ ) and a codeword  $\mathbf{C}_i$  or  $\mathbf{C}_j$  are  $(\frac{\theta_1}{2}, \dots, \frac{\theta_p}{2})$  [14]. It follows that the squared chordal distance between the midpoint on the geodesic and an extremity of the geodesic is

$$\varrho^2 = \left\| \sin \frac{\boldsymbol{\theta}}{2} \right\|_2^2 = \sum_{i=1}^p \sin^2 \frac{\theta_i}{2}. \quad (123)$$

Finding lower and upper bounds reduces to solving the following optimization problems:

$$\begin{aligned} &\text{minimize/maximize} && \left\| \sin \frac{\boldsymbol{\theta}}{2} \right\|_2^2 \\ &\boldsymbol{\theta} \in [0, \frac{\pi}{2}]^p && \\ &\text{subject to} && \left\| \sin \boldsymbol{\theta} \right\|_2^2 = \delta^2. \end{aligned} \quad (124)$$

First, to find the minimum, consider the corresponding Lagrange function

$$\Lambda(\theta_1, \dots, \theta_p, \lambda) = \|\sin \frac{\theta}{2}\|_2^2 + \lambda(\|\sin \theta\|_2^2 - \delta^2). \quad (125)$$

Solving:

$$\frac{\partial \Lambda}{\partial \theta_i} = \sin \theta_i(1/2 + 2\lambda \cos \theta_i) = 0 \quad \text{for } i = 1 \dots p \quad (126)$$

$$\frac{\partial \Lambda}{\partial \lambda} = \sum_{i=1}^p \sin^2 \theta_i - \delta^2 = 0 \quad (127)$$

yields a set of stationary points where at least  $x$  angles are nonzero such that  $x \geq \lceil \delta^2 \rceil$  and equal to  $\theta^* = \arcsin \frac{\delta}{\sqrt{x}}$ . It is then easy to verify that the objective function  $f(x) = \sum_{i=1}^p \sin^2 \frac{\theta_i}{2} = x/2(1 - \sqrt{1 - \delta^2/x})$  is a strictly decreasing function on  $[\lceil \delta^2 \rceil, p]$  and thus is minimized for  $x = p$ . The result follows.

Maximization in (124) is obtained when a minimum number of angles is maximized, i.e., with  $(\theta_1^*, \dots, \theta_p^*) \in [0, \frac{\pi}{2}]^p$  such that  $\theta_1^* = \dots = \theta_{\lceil \delta^2 \rceil}^* = \frac{\pi}{2}$ ,  $\theta_{\lceil \delta^2 \rceil}^* = \arcsin(\sqrt{\delta^2 - \lceil \delta^2 \rceil})$  and  $\theta_{\lceil \delta^2 \rceil + 1}^* = \dots = \theta_p^* = 0$ . This can be verified by contradiction as in [23]: Defining  $s_i = \sin^2 \theta_i$  and  $t(s_i) = (1 - \sqrt{1 - s_i})/2$ , consider the equivalent problem of maximizing  $\sum t(s_i)$  such that  $\sum s_i = \delta^2$  and without loss of generality  $1 \geq s_1 \geq s_2 \dots \geq s_p \geq 0$ . By contradiction, assume that  $\sum t(s_i)$  is maximum at  $\mathbf{a}$  with  $a_i > 0 \forall i$ . It is possible to find a  $\mathbf{b}$  with  $b_i \geq a_i$  for  $1 \leq i \leq p-1$  and  $b_p = 0$ . Since  $t(\cdot)$  is strictly increasing, it follows from the mean value theorem that there exist  $c \in (a_{p-1}, b_{p-1})$  and  $d \in (0, a_p)$  such that

$$\begin{aligned} \sum t(b_i) - t(a_i) &\geq t'(c) \sum_{i=1}^{p-1} (b_i - a_i) + t'(d)(a_p - b_p) \\ &= (t'(d) + t'(c))a_p > 0, \end{aligned} \quad (128)$$

where the last equality is due to the constraint  $\sum b_i = \sum a_i = \delta^2$ , which is in contradiction with the fact that  $\sum t(a_i)$  is a maximum. Repeating the procedure from  $s_p$  to  $s_{\lceil \delta^2 \rceil}$  leads to the results. Lastly, the maximum is

$$\begin{aligned} \sum_{i=1}^p \sin^2 \frac{\theta_i^*}{2} &= \frac{\lceil \delta^2 \rceil}{2} + \frac{1 - \sqrt{1 - (\delta^2 - \lceil \delta^2 \rceil)}}{2} \\ &= \frac{1}{2} (\lceil \delta^2 \rceil - \sqrt{\lceil \delta^2 \rceil - \delta^2}). \end{aligned} \quad (129)$$

### B. Unitary Group

A simple adaptation of the proof for the unitary group can be done, see also [23] where similar optimization problems are considered. Consider a unitary code  $\mathcal{C}$  with minimum distance  $\delta$ . The angles  $(\theta_1, \dots, \theta_n)$  between two codewords  $C_i, C_j$  separated by  $\delta$  now satisfies  $\sum_{i=1}^n \sin^2(\frac{\theta_i}{2}) = \frac{\delta^2}{4}$ . Again, the chordal distance is measured along the geodesic and the principal angles between the midpoint and a codeword are  $(\frac{\theta_1}{2}, \dots, \frac{\theta_n}{2})$ . It follows that the squared chordal distance between the midpoint on the geodesic and an extremity of the geodesic is

$$\varrho^2 = 4 \sum_{i=1}^n \sin^2 \frac{\theta_i}{4}. \quad (130)$$

Finding lower and upper bounds reduces to solving the following optimization problems:

$$\begin{aligned} &\text{minimize/maximize} && 4\|\sin \frac{\theta}{4}\|_2^2 \\ &\theta \in [-\pi, \pi]^n && \\ &\text{subject to} && \|\sin \frac{\theta}{2}\|_2^2 = \frac{\delta^2}{4}. \end{aligned} \quad (131)$$

By using the change of variables  $\phi = (\theta + \pi)/2$ ,  $\gamma = \frac{\delta}{2}$  and  $\rho = \frac{\varrho}{2}$ , we recover the optimization problem (124). The result follows.

### C. Stiefel Manifold

For a generic Stiefel manifold, the notion of principal angles does not exist. The obtained lower bound for the unitary group is actually matching the kissing radius bound from the spherical embedding. This result thus directly extends to all Stiefel manifolds. However, we were not able to generalize the upper bound. We provide some discussion below.

Without loss of generality, we can assume that the first point is  $\mathbf{I}_{n,p}$  and the second point is  $\mathbf{Y}$ . If  $\mathbf{Y}$  is block unitary as  $\mathbf{Y} = [\mathbf{U} \mathbf{0}]^T$ , where  $\mathbf{U} \in \mathcal{U}_p$  then the geodesic between this two points stays in  $\mathcal{U}_p$  embedded in  $\mathcal{V}_{n,p}^{\mathbb{C}}$ , and the upper bound from the unitary group would apply.

For a generic  $\mathbf{Y}$ , the distance only depends on the diagonal element of  $\mathbf{Y} = \{Y_{ij}\}$ :

$$\begin{aligned} d_c^2(\mathbf{I}_{n,p}, \mathbf{Y}) &= 2(p - \Re(\text{Tr}(\mathbf{I}_{n,p}^H \mathbf{Y}))) = 2\left(p - \sum_{i=1}^p \Re(Y_{ii})\right) \\ &= 2\left(p - \sum_{i=1}^p \cos \theta_i\right) = 4 \sum_{i=1}^p \sin^2 \frac{\theta_i}{2} \\ &= 4p \sin^2 \frac{\theta}{2}, \end{aligned} \quad (132)$$

where without loss of generality, we have written  $\cos \theta_i = \Re(Y_{ii})$  and  $\cos \theta = \frac{1}{p} \sum_{i=1}^p \cos \theta_i$ . The angles  $\{\theta_i\}$  then correspond the canonical embedding of the Stiefel manifold in  $(S^{2n-1})^p$ , the angle  $\theta$  to the embedding in  $S^{2np-1}(\sqrt{p})$ .

The geodesic is not along these angles, and the midpoint, say  $\mathbf{M}_{\mathbf{Y}}$ , on the geodesic is not at  $\{\frac{\theta_i}{2}\}$  nor at  $\frac{\theta}{2}$ . However since we have an isometry for these embeddings,

$$\begin{aligned} d_c^2(\mathbf{I}_{n,p}, \mathbf{M}_{\mathbf{Y}}) &\geq 4 \sum_{i=1}^p \sin^2 \frac{\theta_i}{4} \\ &\geq 4p \sin^2 \frac{\theta}{4}. \end{aligned} \quad (133)$$

From this, by maximizing the right hand-side of the first inequality, we can deduce that the maximum is greater or equal to the upper bound of the unitary group, i.e.

$$\max_{\mathbf{Y}} d^2(\mathbf{I}_{n,p}, \mathbf{M}_{\mathbf{Y}}) \geq \bar{\varrho}. \quad (134)$$

A generalization would imply that  $\max_{\mathbf{Y}} d^2(\mathbf{I}_{n,p}, \mathbf{M}_{\mathbf{Y}}) = \bar{\varrho}$ .

### D. Proof of Corollary 2

For the Grassmann manifold with  $R^2 = \frac{p(n-p)}{2n}$ , it can be easily verified that  $p \leq 4R^2$  with equality if and only if  $p = n/2$ . Then, since  $x/2(1 - \sqrt{1 - \delta^2/x})$  is a strictly decreasing function, it follows that  $\varrho_s \leq \underline{\varrho}$  with equality if and only if  $p = n/2$ .

## REFERENCES

- [1] R. A. Rankin, "The closest packing of spherical caps in  $n$  dimensions," *Proc. Glasgow Math. Assoc.*, vol. 2, pp. 139–144, 1955.
- [2] C. A. Rogers, "The packing of equal spheres," *Proc. London Math. Soc.*, vol. 8, pp. 609–620, 1958.
- [3] L. Tóth, "Distribution of points in the elliptic plane," *Acta Math. Hungar.*, vol. 16, no. 3–4, pp. 437–440, 1965.
- [4] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*. Springer-Verlag, New York, 2nd edition, 1993.
- [5] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 4, pp. 623–656, Oct. 1948.
- [6] L. Zheng and D. Tse, "Communication on the Grassmann manifold: a geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002.
- [7] D. Agrawal, T. Richardson, and R. Urbanke, "Multiple-antenna signal constellations for fading channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2618–2626, Sep. 2001.
- [8] D. Love, R. Heath, V. Lau, D. Gesbert, B. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.
- [9] R.-A. Pitaval and O. Tirkkonen, "Joint Grassmann-Stiefel quantization for MIMO product codebook," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 210–222, Feb. 2014.
- [10] R. Heath, T. Strohmer, and A. Paulraj, "On quasi-orthogonal signatures for CDMA systems," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1217–1226, Mar. 2006.
- [11] J. A. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.
- [12] R. Calderbank, A. Thompson, and Y. Xie, "On block coherence of frames," *Appl. Comput. Harmon. Anal.*, vol. 38, no. 1, pp. 50–71, 2015.
- [13] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian space," *Exper. Math.*, vol. 5, pp. 139–159, 1996.
- [14] A. Edelman, T. 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.
- [15] C. Bachoc, Y. Ben-Haim, and S. Litsyn, "Bounds for codes in products of spaces, Grassmann, and Stiefel manifolds," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1024–1035, Mar. 2008.
- [16] A. Ashikhmin and A. Calderbank, "Grassmannian packings from operator Reed-Müller codes," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5689–5714, Nov. 2010.
- [17] W. Dai, Y. Liu, and B. Rider, "Quantization bounds on Grassmann manifolds and applications to MIMO communications," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1108–1123, Mar. 2008.
- [18] G. Han and J. Rosenthal, "Geometrical and numerical design of structured unitary space-time constellations," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3722–3735, Aug. 2006.
- [19] J. Huang, Q. Qiu, and R. Calderbank, "The role of principal angles in subspace classification," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 1933–1945, Apr. 2016.
- [20] A. Barg and D. Nogin, "Bounds on packings of spheres in the Grassmann manifold," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2450–2454, Sep. 2002.
- [21] —, "A bound on Grassmannian codes," in *Proc. IEEE Int. Symp. Inf. Theory*, July 2006, pp. 997–1000.
- [22] O. Henkel, "Sphere packing bounds in the Grassmann and Stiefel manifolds," *IEEE Trans. Inf. Theory*, vol. 51, pp. 3445–3456, Oct. 2005.
- [23] G. Han and J. Rosenthal, "Unitary space-time constellation analysis: An upper bound for the diversity," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4713–4721, Oct. 2006.
- [24] C. Bachoc, "Linear programming bounds for codes in Grassmannian spaces," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2111–2125, May 2006.
- [25] J. Creignou and H. Diet, "Linear programming bounds for unitary space time codes," in *Proc. IEEE Int. Symp. Inf. Theory*, July 2008, pp. 1073–1077.
- [26] R. Krishnamachari and M. Varanasi, "Volume of geodesic balls in the real Stiefel manifold," in *Proc. Conf. Inf. Sci. Syst.*, Mar. 2008, pp. 402–406.
- [27] —, "Volume of geodesic balls in the complex Stiefel manifold," in *Proc. Allerton Conf. on Comm. Cont. and Comp.*, Sep. 2008, pp. 902–909.
- [28] K. K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple-antenna systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2562–2579, Oct. 2003.
- [29] W. Dai, B. Rider, and Y. Liu, "Volume growth and general rate quantization on Grassmann manifolds," in *Proc. IEEE Global Telecom. Conf.*, Nov. 2007, pp. 1441–1445.
- [30] J. Nash, "C1 isometric imbeddings," *Ann. of Math. (2)*, vol. 60, no. 3, pp. 383–396, 1954.
- [31] L. Wei, R.-A. Pitaval, J. Corander, and O. Tirkkonen, "On the volume of a metric ball in unitary group," in *IEEE Int. Symp. Inf. Theory*, June 2015, pp. 191–195.
- [32] —, "From random matrix theory to coding theory: Volume of a metric ball in unitary group," *IEEE Trans. Inf. Theory*, vol. 63, no. 5, pp. 2814–2821, May 2017.
- [33] R.-A. Pitaval, L. Wei, O. Tirkkonen, and J. Corander, "Volume of metric balls in high-dimensional complex Grassmann manifolds," *IEEE Trans. Inf. Theory*, vol. 62, no. 9, pp. 5105–5116, Sep. 2016.
- [34] E. Borel, "Sur les principes de la theorie cinétique des gaz," *Ann. Sci. l'École Norm. Supér.*, vol. 23, pp. 9–32, 1906.
- [35] D. S. P. Richards, "High-dimensional random matrices from the classical matrix groups, and generalized hypergeometric functions of matrix argument," *Symmetry*, vol. 3, pp. 600–610, 2011.
- [36] J. Breuer and M. Duits, "Central limit theorems for biorthogonal ensembles and asymptotics of recurrence coefficients," *J. Amer. Math. Soc.*, vol. 30, no. 1, pp. 27–66, Jan. 2017.
- [37] P. Sole, "Packing radius, covering radius, and dual distance," *IEEE Trans. on Inf. Theory*, vol. 41, no. 1, pp. 268–272, Jan. 1995.
- [38] A. Schenk, R. Fischer, and L. Lampe, "A stopping radius for the sphere decoder and its application to MSDD of DPSK," *IEEE Comm. Letters*, vol. 13, no. 7, pp. 465–467, July 2009.
- [39] A. Schenk and R. Fischer, "A stopping radius for the sphere decoder: Complexity reduction in multiple-symbol differential detection," in *Proc. Int. ITG Conf. Source and Channel Coding*, Jan. 2010, pp. 1–6.
- [40] H.-C. Wang, "Two-point homogeneous spaces," *Ann. Math.*, vol. 55, no. 1, pp. 177–191, Jan. 1952.
- [41] I. S. Dhillon, R. W. Heath Jr, T. Strohmer, and J. A. Tropp, "Constructing packings in Grassmannian manifolds via alternating projection," *Exper. Math.*, vol. 17, no. 1, pp. 9–35, 2008.
- [42] R.-A. Pitaval, H.-L. Maattanen, K. Schober, O. Tirkkonen, and R. Wichman, "Beamforming codebooks for two transmit antenna systems based on optimum Grassmannian packings," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6591–6602, Oct. 2011.
- [43] J. Creignou, "Constructions of Grassmannian simplices," *eprint arXiv:cs/0703036*, Mar. 2007.
- [44] A. R. Calderbank, R. H. Hardin, E. M. Rains, P. W. Shor, and N. J. A. Sloane, "A group-theoretic framework for the construction of packings in Grassmannian spaces," *J. Algebraic Combin.*, vol. 9, pp. 129–140, 1999.
- [45] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 543–564, Mar. 2000.
- [46] R. H. Gohary and T. N. Davidson, "Noncoherent MIMO communication: Grassmannian constellations and efficient detection," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1176–1205, Mar. 2009.
- [47] A. M. Chudnov, "On minimax signal generation and reception algorithms (rus.)," *Probl. Inform. Trans.*, vol. 22, no. 4, p. 4954, 1986.
- [48] —, "Game-theoretical problems of synthesis of signal generation and reception algorithms (rus.)," *Probl. Inform. Trans.*, vol. 27, no. 3, p. 5765, 1991.
- [49] L. K. Hua, *Harmonic Analysis of Functions of Several Variables in the Classical Domains*. American Mathematical Society: Providence, Chinese original 1958, Russian translation, Moskva 1959.
- [50] K. Zyczkowski and H.-J. Sommers, "Hilbert-Schmidt volume of the set of mixed quantum states," *J. Phys. A: Math. Gen.*, vol. 36, no. 39, pp. 10 115–10 130, 2003.
- [51] I. Kim, S. Park, D. Love, and S. Kim, "Improved multiuser MIMO unitary precoding using partial channel state information and insights from the Riemannian manifold," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4014–4023, Aug. 2009.
- [52] A. Gray, "The volume of a small geodesic ball of a Riemannian manifold," *Michigan Math. J.*, vol. 20, no. 4, pp. 329–344, 1974.
- [53] M. Belkin and P. Niyogi, "Towards a theoretical foundation for Laplacian-based manifold methods," *J. Comput. Syst. Sci.*, vol. 74, no. 8, pp. 1289–1308, 2008.
- [54] M. Belkin, "Problems of learning on manifolds," Ph.D. dissertation, University of Chicago, 2003.



- [55] Y. Chikuse, *Statistics on special manifolds*. Lecture Notes in Statistics 174, Springer, New York, 2003.
- [56] R. A. Pitaval and O. Tirkkonen, "Grassmannian packings from orbits of projective group representations," in *Proc. Asilomar Conf. on Sig., Syst. and Comp.*, Nov. 2012, pp. 478–482.
- [57] P. W. Shor and N. J. A. Sloane, "A family of optimal packings in Grassmannian manifolds," *J. Algebraic Combin.*, vol. 7, no. 2, pp. 157–164, 1998.
- [58] R. Bott and L. W. Tu, *Differential forms in algebraic topology*. Springer – Verlag, 1982.
- [59] B. W. Clare and D. L. Kepert, "The closest packing of equal circles on a sphere," *Proc. Roy. Soc. London Ser. A*, vol. 405, no. 1829, pp. 329–344, 1986.
- [60] P. Xia, S. Zhou, and G. Giannakis, "Achieving the Welch bound with difference sets," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1900–1907, May 2005.
- [61] R.-A. Pitaval, O. Tirkkonen, and S. Blostein, "Low complexity MIMO precoding codebooks from orthoplex packings," in *Proc. IEEE Int. Conf. Commun.*, June 2011, pp. 1–5.
- [62] S. Li, "Concise formulas for the area and volume of a hyperspherical cap," *Asian J. Math. Stat.*, vol. 4, no. 1, pp. 66–70, 2011.
- [63] "NIST Digital Library of Mathematical Functions," <http://dlmf.nist.gov/>, Release 1.0.10 of 2015-08-07, online companion to [64]. [Online]. Available: <http://dlmf.nist.gov/>
- [64] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, Eds., *NIST Handbook of Mathematical Functions*. New York, NY: Cambridge University Press, 2010, print companion to [63].
- [65] L. Pastur and M. Shcherbina, *Eigenvalue Distribution of Large Random Matrices*. American Mathematical Society, 2011.
- [66] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. John Wiley & Sons, New York, 1982.
- [67] A. Edelman, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, Massachusetts Institute of Technology, 1989.
- [68] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Statist.*, vol. 35, no. 2, pp. 475–501, May 1964.
- [69] A. M. Mathai, *Jacobians of Matrix Transformations and Functions of Matrix Arguments*. Singapore: World Scientific, 1997.
- [70] I. M. Johnstone, "Multivariate analysis and jacobi ensembles: Largest eigenvalue, Tracy–Widom limits and rates of convergence," *Ann. Stat.*, vol. 36, no. 6, pp. 2638–2716, 2008.
- [71] K. Johansson, "On random matrices from the compact classical groups," *Ann. Math.*, vol. 145, no. 3, pp. 519–545, May 1997.

**Renaud–Alexandre Pitaval** received a 'Diplôme d'Ingénieur' in 2008 from the Grenoble Institute of Technology, France, a M.Sc. in 2009 and a D.Sc. in 2014 from Aalto University, all in electrical engineering. From 2008 to 2014, he was principally with the Department of Communications and Networking at Aalto University. He has also been working with the Department of Signal Processing and Acoustics at Aalto University (2007–2009); Queen's University, Canada (2009–2010); Imperial College London, UK (Spring 2013); Nokia Networks, Finland (2014); and the Department of Mathematics and Systems Analysis at Aalto University (Spring 2015). Since 2015, he is a Senior Researcher at Huawei Technologies Sweden, Stockholm.

**Lu Wei** (S'08–M'13) received the B.Eng. degree from Xi'an Jiaotong University, China, the M.Sc. degree (with distinction) from Helsinki University of Technology, Finland, and the D.Sc. degree (with distinction) from Aalto University, Finland, in 2006, 2008, and 2013, respectively. He held postdoctoral positions at the Department of Mathematics and Statistics, University of Helsinki, Finland, from 2013 to 2015 and at the School of Engineering and Applied Sciences, Harvard University, USA, from 2015 to 2016. Since fall 2016, he is an Assistant Professor with the Department of Electrical and Computer Engineering, University of Michigan–Dearborn, USA. His research interests lie in random matrix theory with applications to communications theory and information theory of both classical and quantum.

**Olav Tirkkonen** is associate professor in communication theory at the Department of Communications and Networking in Aalto University, Finland, where he has held a faculty position since August 2006. He received his M.Sc. and Ph.D. degrees in theoretical physics from Helsinki University of Technology in 1990 and 1994, respectively. Between 1994 and 1999 he held post-doctoral positions at the University of British Columbia, Vancouver, Canada, and the Nordic Institute for Theoretical Physics, Copenhagen, Denmark. From 1999 to 2010 he was with Nokia Research Center (NRC), Helsinki, Finland. In 2016–2017 he was Visiting Associate Professor at Cornell University, Ithaca, NY, USA. He has published some 200 papers, and is coauthor of the book "Multiantenna transceiver techniques for 3G and beyond". His current research interests are in coding theory, multiantenna techniques, and cognitive management of 5G cellular networks.

**Camilla Hollanti** (M'09) (born 1979, Finland) received her M.Sc. and Ph.D. degrees in mathematics from the University of Turku in 2003 and 2009, respectively. Her research interests include applications of algebra and number theory to wireless communications, physical layer security, physical layer network coding, and distributed storage systems.

For 2004–2011 Hollanti was with the University of Turku. She joined the University of Tampere as Lecturer for the academic year 2009–2010. Since Oct. 2011, she has been with the Department of Mathematics and Systems Analysis at Aalto University, where she currently works as Associate Professor and leads a research group in Algebra, Number Theory, and Applications.

Hollanti is an editor of the AIMS Journal on *Advances in Mathematics of Communications*. She is a coauthor of 80 scientific peer-reviewed publications and a recipient of several grants, including five Academy of Finland grants in 2010–2016. In 2014, she received the World Cultural Council Special Recognition Award for young researchers.