Information and Inference: A Journal of the IMA (2022) **00**, 1–23

https://doi.org/DOI HERE

Statistical characterization of the chordal product determinant of Grassmannian codes

JAVIER ÁLVAREZ-VIZOSO*

Department of Communications Engineering, Universidad de Cantabria, Plaza de la Ciencia 1, 39005, Santander, Spain

CARLOS BELTRÁN

Department of Mathematics, Statistics and Computing, Universidad de Cantabria, Av. de los Castros 48, 39005, Santander, Spain

DIEGO CUEVAS AND IGNACIO SANTAMARÍA

Department of Communications Engineering, Universidad de Cantabria, Plaza de la Ciencia 1, 39005, Santander, Spain

AND

VÍT TUČEK AND GUNNAR PETERS

Department of Wireless Algorithms, Huawei Technologies, 16440, Kista, Sweden

*Corresponding author: javier.alvarezvizoso@unican.es

[Received on Date Month Year; revised on Date Month Year; accepted on Date Month Year]

We consider the chordal product determinant, a measure of the distance between two subspaces of the same dimension. In information theory, collections of elements in the complex Grassmannian are searched with the property that their pairwise chordal products are as large as possible. We characterize this function from an statistical perspective, which allows us to obtain bounds for the minimal chordal product and related energy of such collections.

Keywords: Coherence, energy, Grassmannian codes, noncoherent communication.

1. Introduction and statement of the main results

Let $T \geq 2M$ be two positive integers and consider the complex Grassmannian $\mathbb{G}r(M,\mathbb{C}^T)$, i.e. the space of M–dimensional complex vector subspaces of $\mathbb{C}^T$. Finite collections of points (also called *codes* or *packings*) in $\mathbb{G}r(M,\mathbb{C}^T)$ with different desired separation properties have been investigated by several authors (in Section 1.3 we describe some relevant references). The limitation $T \geq 2M$ is motivated by information-theoretic arguments, cf. [25], otherwise one would have a scenario in which any two points in the Grassmannian packing would generically intersect, as linear subspaces, in $2M-T$ dimensions. The most frequent criterion for "well–separated" codes is the maximization of the minimal mutual squared chordal distance, which is the sum the squared sines of the principal angles of two subspaces. However, following [11, 15, 18, 20] (see also Section 1.1 below), a more relevant measure for its application to information theory is given by the *chordal product energy*, related to the product of the squared sines of the principal angles, which justifies its name. Given a code $[\mathbf{X}_1],\ldots,[\mathbf{X}_K] \in \mathbb{G}r(M,\mathbb{C}^T)$, its chordal product energy with integer parameter N is

$$\mathscr{E}(\mathbf{X}_1,\ldots,\mathbf{X}_K) = \sum_{i\neq j} \det(\mathbf{I}_M - \mathbf{X}_i^H\mathbf{X}_j\mathbf{X}_j^H\mathbf{X}_i)^{-N}, \tag{1.1}$$

where $\mathbf{I}_M$ is the identity matrix and we have chosen $T\times M$ matrices $\mathbf{X}_i$ as representatives of each point $[\mathbf{X}_i]$ satisfying $\mathbf{X}_i^H\mathbf{X}_i = \mathbf{I}_M$, where $(\cdot)^H$ stands for conjugate transpose. Note that the energy is well-defined in the sense that it does not change if other representatives with that property are chosen. An elementary computation shows that, if $\mathbf{UDV}^H$ is the SVD decomposition of $\mathbf{X}_i^H\mathbf{X}_j$, then the diagonal elements of $\mathbf{D}$ are the cosines of the principal angles between the subspaces $[\mathbf{X}_i]$ and $[\mathbf{X}_j]$, $\cos\theta_1,\ldots,\cos\theta_M$, so each determinant in the chordal product energy reduces to a product of terms $1-\cos^2\theta_1,\ldots,1-\cos^2\theta_M$, i.e.:

$$\det\left(\mathbf{I}_M - \mathbf{X}_i^H\mathbf{X}_j\mathbf{X}_j^H\mathbf{X}_i\right) = \det\left(\mathbf{I}_M - \mathbf{D}^2\right) = \prod_{k=1}^{M}\sin^2\theta_k, \tag{1.2}$$

while the squared chordal distance between the two subspaces $[\mathbf{X}_i]$ and $[\mathbf{X}_j]$ is given by $\sum_{k=1}^M \sin^2\theta_k$. The sum in (1.1) is a pairwise interaction energy in the spirit of the well–studied Riesz or logarithmic energies of importance in Potential Theory (see [7] for a complete monograph dedicated to energy minimization in the sphere and other spaces). We refer to the function (again, choosing representatives $\mathbf{A}$ and $\mathbf{B}$ such that $\mathbf{A}^H\mathbf{A} = \mathbf{B}^H\mathbf{B} = \mathbf{I}_M$)

$$[\mathbf{A}], [\mathbf{B}] \in \mathbb{G}r(M, \mathbb{C}^T) \mapsto \det(\mathbf{I}_M - \mathbf{A}^H\mathbf{B}\mathbf{B}^H\mathbf{A}) = \det(\mathbf{I}_M - \mathbf{B}^H\mathbf{A}\mathbf{A}^H\mathbf{B}), \tag{1.3}$$

as the *chordal product determinant* or, simply, the chordal product, and note that it is *not* a metric in $\mathbb{G}r(M, \mathbb{C}^T)$, for it may happen that $[\mathbf{A}] \neq [\mathbf{B}]$ and yet $\det(\mathbf{I}_M - \mathbf{A}^H\mathbf{B}\mathbf{B}^H\mathbf{A}) = 0$, if the intersection of $[\mathbf{A}]$ and $[\mathbf{B}]$ is nontrivial.

In this paper we perform the first theoretical study of the chordal product energy (for numerical results, see [11] and references therein). We start by describing the context where the problem arises, following [15]. The organization of the paper is as follows: in Sec. 1.1 we describe the role of Grassmannian codes in information theory models, both with and without noise, relating the asymptotic pairwise error probability to the chordal product energy. In Sec. 1.2 we state the main result giving a complete statistical characterization of the chordal product as a product of beta-distributed random variables, including an upper bound on the energy and a Gilbert-Varshamov lower bound. In Sec. 1.3 we briefly put these results in a historical context. In Sec. 2 we provide the proof of the main theorem, and in Sec. 3 the probability density function of the chordal product is derived. Then in Sec. 4 we prove the Gilbert-Varshamov bound. Finally, in Appendix A, an alternative parametrization of the Grassmannian is introduced.

1.1. *The importance of Grassmannian codes in information theory*

Consider a *transmitter*, i.e. some device that is able to send a signal, which here means a collection of numbers ordered in a complex $T \times M$ matrix $\mathbf{X}$. Physically, this corresponds to the setting where the transmitter has M antennas and there is a total amount of T time slots where the communication channel is assumed to be constant (i.e. the contour conditions of the communication are considered constant during the time that these TM numbers are sent). At each time instant, a complex number is sent through each of the M antennas, i.e. a row of the transmitted matrix $\mathbf{X}$ is sent in each time slot (this corresponds to the sampling period in a practical digital processor). After T time slots or channel uses, the whole matrix $\mathbf{X}$ has been transmitted. The *receiver* is another device, that we consider equipped with N antennas, and the signal it receives is

$$\mathbf{Y} = \mathbf{X}\mathbf{H} + \sqrt{\frac{M}{T\rho}}\mathbf{W},$$

where $\mathbf{H}$ is an unknown $M \times N$ matrix (termed *the channel*), $\mathbf{W}$ is an unknown $T \times N$ matrix that describes the noise and ρ, called the signal-to-noise-ratio (SNR), measures the magnitude of the signal against the noise. Hence, the receiver acquires over the T time slots, synchronously with the transmitter, a distorted version, $\mathbf{Y}$, of the transmitted matrix, $\mathbf{X}$, and the decoding process to decide which matrix was sent starts. Notice that the pairwise error probability of this decision depends on the packing from which $\mathbf{X}$ is chosen. In particular N is always a positive integer. In this paper, we will assume that $\mathbf{H}$ and $\mathbf{W}$ are independent and have complex Gaussian $\mathscr{C}\mathscr{N}(0,1)$ entries. This assumption for $\mathbf{H}$ is reasonable to model the effect of heavily built-up urban environments on radio signals, and this assumption for $\mathbf{W}$ mimics the effect of many random natural noise sources.

1.1.1. The zero–noise case

Since $\mathbf{H}$ is unknown (it is common to assume that it has random complex Gaussian entries), even in the event that $\mathbf{W} = 0$ the receiver cannot recover the whole matrix $\mathbf{X}$:

- If two matrices $\mathbf{X}_1$ and $\mathbf{X}_2$ have the same column span, then one can easily find an full–rank matrix $\mathbf{H}$ such that $\mathbf{X}_1\mathbf{H} = \mathbf{X}_2\mathbf{H}$, hence the receiver just cannot distinguish which of these two matrices was the original signal.
- On the other hand, if two matrices $\mathbf{X}_1$ and $\mathbf{X}_2$ have the property that the intersection of the column span of $\mathbf{X}_1$ and $\mathbf{X}_2$ is trivial, the the receiver can easily distinguish if a given matrix $\mathbf{Y}$ has been constructed by $\mathbf{X}_1\mathbf{H}$ or by $\mathbf{X}_2\mathbf{H}$: if the column span of $\mathbf{Y}$ intersected with the column span of $\mathbf{X}_1$ (resp. $\mathbf{X}_2$) is nontrivial, then $\mathbf{X}_1$ (resp. $\mathbf{X}_2$) was sent.

Summarizing, if a previously agreed code of possible signals $[\mathbf{X}_1],\ldots,[\mathbf{X}_K]\in\mathbb{G}r(M,\mathbb{C}^T)$ is fixed with the property that the column spans of $\mathbf{X}_i$ and $\mathbf{X}_j$ have trivial intersection for $i\neq j$, the receiver will be able to recover, at least in the zero–noise scenario, *the element of the Grassmannian represented by the sent signal, but not the concrete representative of that element.* Hence, collections of points in $\mathbb{G}r(M,\mathbb{C}^T)$ are searched with that property.

1.1.2. The general case

In the more realistic context of the presence of non–zero noise, the analysis is quite more involved since there is always a non–zero probability of error in the detection procedure. The pioneer work [15] showed that, in order to recover the element $\mathbf{X}_i$ of $\mathbb{G}r(M,\mathbb{C}^T)$ just by knowing $\mathbf{Y}$, the optimal method is to use the so called maximum–likelihood decoder:

$$i=\operatorname{argmax}_{j=1,\ldots,K}\operatorname{tr}(\mathbf{Y}^H\mathbf{X}_j\mathbf{X}_j^H\mathbf{Y})=\operatorname{argmax}_{j=1,\ldots,K}\operatorname{tr}(\mathbf{X}_j^H\mathbf{Y}\mathbf{Y}^H\mathbf{X}_j).$$

Then, [8] showed that if only 2 codewords are permitted, i.e. if $K=2$, and assuming that the entries of $\mathbf{H}$ and $\mathbf{W}$ are complex Gaussian $\mathscr{CN}(0,1)$ numbers, then the probability $P_e(\mathbf{X}_1,\mathbf{X}_2,\rho)$ of erroneously decoding $\mathbf{X}_1$ if $\mathbf{X}_2$ was sent can be given by a (quite complicated) formula involving the residues of a certain rational function. Luckily, the asymptotic expansion of this *Pairwise Error Probability* (PEP) in the case $\rho\to\infty$ (which is the vanishing noise scenario, ideal for communications), called the high-SNR asymptotic analysis, admits a much more concise expression, see [11, 18]:

$$P_e(\mathbf{X}_1,\mathbf{X}_2,\rho)\approx C\rho^{-NM}\det(\mathbf{I}_M-\mathbf{X}_1^H\mathbf{X}_2\mathbf{X}_2^H\mathbf{X}_1)^{-N},\quad \rho\to\infty, \tag{1.4}$$

where $C=\frac{1}{2}\left(\frac{4M}{T}\right)^{NM}\frac{(2NM-1)!!}{(2NM)!!)}$. It is assumed that any two distinct points have trivial intersection as linear subspaces (which relates to the requirement $T\geq 2M$), and the representatives $\mathbf{X}_i$ of each $[\mathbf{X}_i]$ are such that $\mathbf{X}_i^H\mathbf{X}_i=\mathbf{I}_M$. If we have K elements $[\mathbf{X}_1],\ldots,[\mathbf{X}_K]$ in the code of possible signals and we send one of them at random, all with equal probability $1/K$, then the total probability of erroneously decoding a signal is bounded above by the so-called *Union Bound* (UB)

$$P_e \leq \frac{1}{K}\sum_{i=1}^{K}\sum_{j=i+1}^{K} P_e\left(\mathbf{X}_i,\mathbf{X}_j,\rho\right) \tag{1.5}$$

where P_e is the total probability of erroneusly decoding a signal, which is then asymptotically approximated by

$$\frac{1}{K}\sum_{i<j} P_e(\mathbf{X}_i,\mathbf{X}_j,\rho) \approx \frac{C}{K}\rho^{-NM}\sum_{i<j}\det(\mathbf{I}_M - \mathbf{X}_i^H\mathbf{X}_j\mathbf{X}_j^H\mathbf{X}_i)^{-N}. \tag{1.6}$$

The determinant in (1.4) is the chordal product (1.3) and the sum in the right–hand side in (1.6) is the energy (1.1).

1.1.3. Criteria for the design of Grassmannian codes

It follows from the previous discussion that reasonable criteria for the design of a code $[\mathbf{X}_1],\ldots,[\mathbf{X}_K]$ would be to maximize the minimum pairwise chordal product (1.3), or to minimize the chordal product energy (1.1). In [11], these approaches are considered, numerically showing that the obtained codes are very well suited for their use in non–coherent communications, in which neither the transmitter nor the receiver have any knowledge about the channel matrix $\mathbf{H}$ except from its probability distribution, with a slight advantage in the use of the chordal product energy. Yet, little to no theory exists about the behavior of the optimal pairwise chordal product or energy. The main purpose of this paper is to put the basis for the study of these questions.

1.2. *Main results of the paper*

We will start our study by computing the moments of the chordal product when $[\mathbf{B}]$ is fixed and $[\mathbf{A}]$ is chosen at random uniformly in $\mathbb{G}r(M,\mathbb{C}^T)$, w.r.t. the unique, standard rotation–invariant probability measure. This yields a complete statistical characterization of the chordal product as a product of beta–distributed random variables:

Theorem 1 *Assume that* $T \geq 2M$. *Let* $p \in (2M-T-1,\infty)$ *(notice that* p *may be negative and/or noninteger). Let* $[\mathbf{B}] \in \mathbb{G}r(M,\mathbb{C}^T)$ *be any fixed element and let* $[\mathbf{A}] \in \mathbb{G}r(M,\mathbb{C}^T)$ *be uniformly distributed on the Grassmannian. Then, the* p*–th moment of* $\det(\mathbf{I}_M - \mathbf{B}^H\mathbf{A}\mathbf{A}^H\mathbf{B})$ *is:*

$$\mathrm{E}_{\mathbf{A}}[\det(\mathbf{I}_M - \mathbf{B}^H\mathbf{A}\mathbf{A}^H\mathbf{B})^p] = \prod_{m=1}^{M} \frac{\Gamma(T-m+1)\Gamma(T+p-m-M+1)}{\Gamma(T-m-M+1)\Gamma(T+p-m+1)}, \tag{1.7}$$

where $\Gamma(\cdot)$ *is Euler's Gamma function. Moreover,* $\det(\mathbf{I}_M - \mathbf{B}^H\mathbf{A}\mathbf{A}^H\mathbf{B})$ *is distributed as the product of* M *independent beta random variables,* z_m*, with parameters* $\alpha_m = T-M+1-m$ *and* $\beta_m = M$, $m = 1,\dots,M$*, i.e.*

$$\det(\mathbf{I}_M - \mathbf{B}^H\mathbf{A}\mathbf{A}^H\mathbf{B}) \sim \prod_{m=1}^{M} z_m, \quad z_m \sim \mathrm{Beta}(T-M+1-m, M). \tag{1.8}$$

An immediate consequence is that, at least for moderate values of N, we can upper bound the energy (1.1) and hence the probability of error (1.6) of random codes $[\mathbf{X}_1],\dots,[\mathbf{X}_K]$ when they are all independently and uniformly distributed:

Corollary 1 *Assume that* $N \leq T-2M$. *For i.i.d. chosen* $[\mathbf{X}_1],\dots,[\mathbf{X}_K]$*, the expected value of the chordal product energy* (1.1) *is*

$$K(K-1)\prod_{m=1}^{M} \frac{(T-m)!(T-N-m-M)!}{(T-m-M)!(T-N-m)!},$$

In particular, there exists a code such that the upper bound in the union bound (1.5) *is at most:*

$$C(K-1)\rho^{-NM}\prod_{m=1}^{M} \frac{(T-m)!(T-N-m-M)!}{(T-m-M)!(T-N-m)!},$$

where $C = \frac{1}{2}\left(\frac{4M}{T}\right)^{NM}\frac{(2NM-1)!!}{(2NM)!!)}$.

In Section 3 we use Theorem 1 to compute exactly the probability density function and the cumulative density function of the random variable $\det(\mathbf{I}_M - \mathbf{B}^H\mathbf{A}\mathbf{A}^H\mathbf{B})$ in the same hypotheses of the

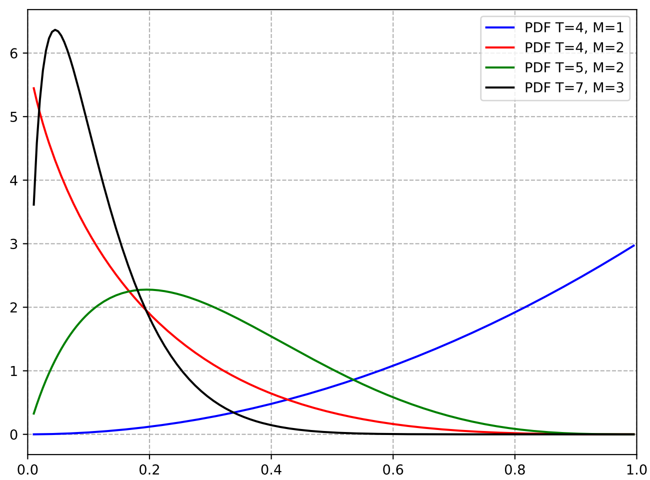

Figure 1. Probability density functions of $\det(\mathbf{I}_M - \mathbf{B}^H\mathbf{A}\mathbf{A}^H\mathbf{B})$, when $[\mathbf{B}] \in \mathbb{G}r(M, \mathbb{C}^T)$ is fixed and $[\mathbf{A}] \in \mathbb{G}r(M, \mathbb{C}^T)$ is uniformly distributed on the Grassmannian.

theorem. The expressions we get are exact and can be obtained in closed form for any fixed value of M. A reduced version of that result for $M = 2$ is now shown:

Corollary 2 *Fix any* $[\mathbf{B}] \in \mathbb{G}r(2, \mathbb{C}^T)$ *with* $T \geq 4$. *The probability that a randomly chosen* $[\mathbf{A}] \in \mathbb{G}r(2, \mathbb{C}^T)$ *satisfies* $\det(\mathbf{I}_2 - \mathbf{B}^H\mathbf{A}\mathbf{A}^H\mathbf{B}) \leq \delta \in (0, 1]$ *is exactly:*

$$F_2(\delta, T) = \frac{1}{2}(T-1)(T-2)^2(T-3)\delta^{T-3}\left(\frac{1}{T-3} - \frac{2\delta}{(T-2)^2} - \frac{\delta^2}{T-1} + \frac{2\delta\log\delta}{T-2}\right),$$

See the complete result in Corollary 4. As an illustrative example, Figs. 1 and 2 depict, respectively, the computed pdf and cdf of the chordal product for different values of T and M.

Using the statistical characterization above, we have derived a lower bound on the number of elements in any code in the Grassmannian with a given minimum value of chordal product δ. Following [4], we call this result a Gilbert-Varshamov bound since its proof mimics the argument of that classical result.

Corollary 3 (Gilbert–Varshamov lower bound) *Assume that* $T \geq 2M$. *For any fixed* $K \geq 2$, *there exists a code* $[\mathbf{X}_1], \ldots, [\mathbf{X}_K] \in \mathbb{G}r(M, \mathbb{C}^T)$ *such that* $\det(\mathbf{I}_M - \mathbf{X}_i^H\mathbf{X}_j\mathbf{X}_j^H\mathbf{X}_i) \geq \delta$ *where* δ *is the unique solution*

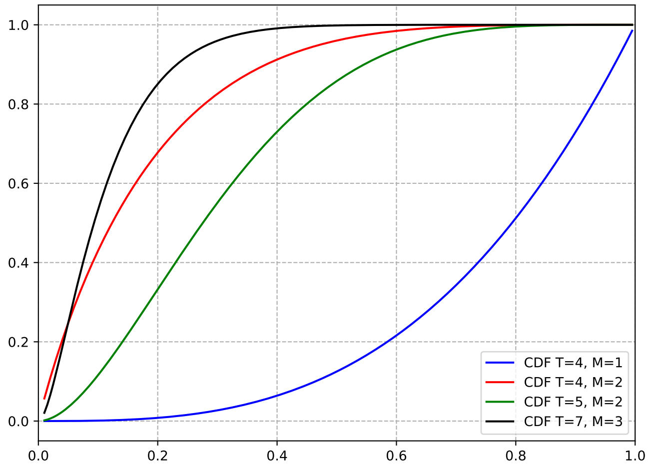

Figure 2. Cumulative distribution function of $\det(\mathbf{I}_M - \mathbf{B}^H\mathbf{A}\mathbf{A}^H\mathbf{B})$, when $[\mathbf{B}] \in \mathbb{G}r(M,\mathbb{C}^T)$ is fixed and $[\mathbf{A}] \in \mathbb{G}r(M,\mathbb{C}^T)$ is uniformly distributed on the Grassmannian.

of the equation:

$$F_M(\delta;T) = \frac{1}{K};\ \textit{that is}\ \delta = F_M^{-1}\left(\frac{1}{K};T\right)$$

Equivalently, given $\delta \in (0,1)$, there exists a code consisting of $K \geq \frac{1}{F_M(\delta;T)}$ elements and satisfying $\det(\mathbf{I}_M - \mathbf{X}_i^H\mathbf{X}_j\mathbf{X}_j^H\mathbf{X}_i) \geq \delta$ for $i \neq j$.

Example 1 *In the case $M = 1$ we have that*

$$\det(\mathbf{I}_M - \mathbf{X}_i^H\mathbf{X}_j\mathbf{X}_j^H\mathbf{X}_i) = \sin^2\theta,$$

where $\theta \in [0,\pi/2]$ is the principal angle between the one-dimensional subspaces $[\mathbf{X}_i]$ and $[\mathbf{X}_j]$ in $\mathbb{G}(1,\mathbb{C}^T)$. That is to say, the chordal product coincides with the squared chordal distance. For uniformly distributed subspaces, the squared sine of the pairwise principal angle has cdf $F_1(\delta,T) = \delta^{T-1}$. The Gilbert-Varshamov bound shows that, for $\delta \in (0,1)$, there exist codes with cardinality K and minimum chordal product $\delta = \sin^2\theta$ such that

$$K > \delta^{-(T-1)} = (\sin\theta)^{-2(T-1)}. \tag{1.9}$$

This result provides a valid bound also for small values of T whereas, for example, in [4, Eq. 6] only asymptotic lower bounds on the so-called transmission rate, $R = \frac{\log_2 K}{T}$, are computed for a packing problem on $\mathbb{C}^T$ with $T \to \infty$.

Example 2 *Let us now take $T = 10, M = 2$. Assume that we want to allocate $K = 2^9$ points in $\mathbb{G}r(M, \mathbb{C}^T)$. Then, Corollary 3 says that there exists a code $[\mathbf{X}_1], \ldots, [\mathbf{X}_K]$ such that for all these points the chordal product is at least δ, the unique solution of:*

$$\frac{1}{2}(T-1)(T-2)^2(T-3)\delta^{T-3}\left(\frac{1}{T-3} - \frac{2\delta}{(T-2)^2} - \frac{\delta^2}{T-1} + \frac{2\delta\log\delta}{T-2}\right) = \frac{1}{K},$$

that is

$$\log_2\left(2016\delta^7\left(\frac{1}{7} - \frac{2\delta}{64} - \frac{\delta^2}{9} + \frac{\delta\log\delta}{4}\right)\right) = -9,$$

which yields $\delta \approx 0.2129$. The numerical algorithm in [11] produces in this case $[\mathbf{X}_1], \ldots, [\mathbf{X}_{2^9}]$ with minimum determinantal value $0.3958 > 0.2129$.

1.3. *Historical discussion*

There exist several results on packings on Grassmannian spaces but they are rather centered on finding codes such that the mutual chordal distance between different elements $[\mathbf{X}_i], [\mathbf{X}_j]$ is close to maximal. For example, in [21] the author finds bounds for the mutual distance of any code with a fixed number of elements (this is known as the Rankin bound). Gilbert–Varshamov bounds have also been obtained for that chordal distance by resorting to calculations of the volume of a metric ball of radius δ in $\mathbb{G}r(M, \mathbb{C}^T)$, see [4] and [12]. The case that δ is sufficiently small was analyzed in [12], [14], and the real case has also been studied, see [10] and references therein. However, to our knowledge, our results are the first theoretical bounds on codes focusing on the explicit use of the chordal product, which is the key figure of merit in non–coherent communications. The statistics of the chordal product can also be studied in the real case, for which there are known formulas describing the joint density of the principal angles between two random real subspaces, see [2, Appendix D.3] or [9]. However, the relevance of that study is not clear from the point of view of communications theory and hence we do not pursue it here.

In the case $M = 1$ the Grassmannian becomes the projective space, the chordal product equals the squared chordal distance, and the literature is much more prolific, going back to [22] (although Shannon studied the case of the geodesic, not chordal, distance), [4] for the chordal and the geodesic distance, and more recently [17], where a more complete set of references can be found. The optimal value of (1.1) in this case has been studied in [6] and [3] as a case of Riesz energy, showing that the minimum value is equal to the average computed in Corollary 1, minus a term of the form

$$O\left(K^{1+\frac{N}{T-1}}\right) = o(K^2), \quad \text{for } N \le T-2.$$

2. Proof of Theorem 1

First assume that p is an integer in the range of the hypotheses. Let $\mathrm{E}(p,T)$ be the expected value in the theorem (we omit the dependence on M in the notation). By unitary invariance, we can assume that $\mathbf{B} = \binom{\mathbf{I}_M}{\mathbf{0}}$. If we write the expected value using Proposition 1 and we pass to polar coordinates we get

$$\begin{aligned}
\mathrm{E}(p,T) =& C(T) \int_{\tilde{\mathbf{A}}\in\mathbb{C}^{(T-M)\times M}} \frac{\det\left(\mathbf{I}_M - (\mathbf{I}_M + \tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^{-1}\right)^p}{\det(\mathbf{I}_M + \tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^T}\, d\tilde{\mathbf{A}} \\
&= C(T) \int_{\tilde{\mathbf{A}}\in\mathbb{C}^{(T-M)\times M}} \frac{\det\left(\mathbf{A}^H\tilde{\mathbf{A}}\right)^p}{\det(\mathbf{I}_M + \tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^{T+p}}\, d\tilde{\mathbf{A}} \\
=& C(T) \int_0^\infty \rho^{2M(T-M)+2pM-1} \int_{\substack{\tilde{\mathbf{A}}\in\mathbb{C}^{(T-M)\times M} \\ \|\tilde{\mathbf{A}}\|_F=1}} \frac{\det(\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^p}{\det(\mathbf{I}_M + \rho^2\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^{T+p}}\, d\tilde{\mathbf{A}}\, d\rho,
\end{aligned}$$

where we omit the dependence on M in the constant:

$$C(T) = \frac{1}{Vol(\mathbb{G}r(M,\mathbb{C}^T))} \overset{Lemma\ 1}{=} \frac{(T-M)!\cdots(T-1)!}{\pi^{M(T-M)}1!\cdots(M-1)!}.$$

Since the integrand of the inner integral depends only on the singular values of $\tilde{\mathbf{A}}$, we can take it to the set $\mathbb{S}_M^+$ consisting of ordered tuples of positive numbers $\sigma_1 > \ldots > \sigma_M$ with the property that

$\sigma_1^2+\cdots+\sigma_M^2=1$, see for example [5, Th. 3.3], that yields

$$\mathrm{E}(p,T)=D(T)\times \int_0^\infty \rho^{2M(T-M)+2pM-1}\int_{\mathbb{S}_M^+}\frac{(\sigma_1\cdots\sigma_M)^{2p+2T-4M+1}\prod_{j\neq k}(\sigma_k^2-\sigma_j^2)^2}{\prod_{m=1}^M(1+\rho^2\sigma_m^2)^{T+p}}\,d\sigma_1\cdots d\sigma_M\,d\rho,$$

where

$$D(T)=\frac{C(T)Vol(\mathscr{U}_{T-M})Vol(\mathscr{U}_M)}{Vol(\mathscr{U}_{T-2M})2^{M(T-M)}\pi^M},$$

and $\mathscr{U}_k$ is the unitary group of degree k. It follows immediately that $\mathrm{E}(p,T)/D(T)=\mathrm{E}(0,T+p)/D(T+p)=1/D(T+p)$, that is,

$$\begin{aligned}\mathrm{E}(p,T)=&\frac{D(T)}{D(T+p)}\\ =&\frac{C(T)Vol(\mathscr{U}_{T-M})Vol(\mathscr{U}_M)}{2^{M(T-M)}Vol(\mathscr{U}_{T-2M})}\frac{2^{M(T+p-M)}Vol(\mathscr{U}_{T+p-2M})}{C(T+p)Vol(\mathscr{U}_{T+p-M})Vol(\mathscr{U}_M)}\\ =&\frac{(2\pi)^{Mp}(T-M)!\cdots(T-1)!Vol(\mathscr{U}_{T-M})Vol(\mathscr{U}_{T+p-2M})}{Vol(\mathscr{U}_{T-2M})(T+p-M)!\cdots(T+p-1)!Vol(\mathscr{U}_{T+p-M})}.\end{aligned}$$

The volume of the unitary group is known (see [5, p. 28]):

$$Vol(\mathscr{U}_k)=\frac{(2\pi)^{k(k+1)/2}}{1!\cdots(k-1)!}.$$

The theorem (for integer p in the range) follows by substituting the known values in the constants above.

On the other hand, it is known that the pth moment of a beta distributed random variable with parameters $\alpha>0$ and $\beta>0$ denoted as $x\sim\mathrm{Beta}(\alpha,\beta)$ is [24]

$$\mathrm{E}[x^p]=\frac{\Gamma(\alpha+\beta)\Gamma(\alpha+p)}{\Gamma(\alpha+\beta+p)\Gamma(\alpha)}, \tag{2.1}$$

so the mth product term in (1.7) corresponds to the pth moment of a beta distributed random variable with parameters $\alpha_m = T - M + 1 - m$ and $\beta_m = M$, thus proving that the distribution of $\det(\mathbf{I}_M - \mathbf{B}^H\mathbf{A}\mathbf{A}^H\mathbf{B})$ is equivalent to the distribution of the product of M independent beta random variables (this is an instance of the Hausdorff moments problem, hence the distribution is uniquely determined by its moments). Notice that (2.1) is valid for $p + \alpha > 0$, and since m can get up to M this entails to $p > 2M - T - 1$. Now that we have characterized $\det(\mathbf{I}_M - \mathbf{B}^H\mathbf{A}\mathbf{A}^H\mathbf{B})$ as a product of beta distributed random variables, we can write down the formula for its moments for noninteger $p > 2M - T - 1$, finishing the proof of the theorem.

3. Probability density function of the chordal product

Can we effectively recover the pdf of the random variable $x = \det(\mathbf{I}_M - \mathbf{B}^H\mathbf{A}\mathbf{A}^H\mathbf{B})$ from its moments? If the density function is $f(x)$ and the moments are $\mathscr{M}_n$ then we have the classical formula (see, e.g., [19, Sec. 5-5]):

$$f(x) = \int_{-\infty}^{\infty} e^{2i\pi xs} \sum_{n=0}^{\infty} \frac{(-2i\pi s)^n}{n!} \mathscr{M}_n \, ds. \tag{3.1}$$

Following [23, Th. 7] a closed-form expression for the pdf can actually be written down in terms of certain special functions called Meijer G–functions. However, this expression is quite involved and requires extra work in practice for the derivation of bounds. In the following, we show that we can obtain simpler closed-form formulas for small values of $M = 1,2,3$. They represent the most practical use cases in noncoherent communications. Moreover, we also provide a general recursive procedure to obtain the pdfs for higher values:

Corollary 4 *Let $T \geq 2M$. The probability density function (pdf) of* $\det(\mathbf{I}_M - \mathbf{B}^H\mathbf{A}\mathbf{A}^H\mathbf{B})$*, when* $[\mathbf{B}] \in \mathbb{G}r(M, \mathbb{C}^T)$ *is fixed and* $[\mathbf{A}] \in \mathbb{G}r(M, \mathbb{C}^T)$ *is uniformly distributed on the Grassmannian, for* $M = 1,2,3$

is:

$$
\begin{aligned}
M=1 \rightarrow\ f_1(x;T) =& (T-1)x^{T-2},\\
M=2 \rightarrow\ f_2(x;T) =& \frac{1}{2}(T-1)(T-2)^2(T-3)x^{T-4}\left(1-x^2+2x\log x\right),\\
M=3 \rightarrow\ f_3(x;T) =& \frac{1}{288}(T-1)(T-2)^2(T-3)^3(T-4)^2(T-5)x^{T-6}\times\\
& \left(1+80x-162x^2+80x^3+x^4+24x\log x-24x^3\log x-36x^2\log^2 x\right)
\end{aligned}
$$

The cumulative distribution function (cdf) $F_M(x,T)=\int_0^x f_M(s,T)\,ds$ for these three cases is respectively:

$$
\begin{aligned}
M=1 \rightarrow & F_1(x;T) = x^{T-1}\\
M=2 \rightarrow & F_2(x;T) = \frac{1}{2}(T-1)(T-2)^2(T-3)x^{T-3}\left(\frac{1}{T-3}-\frac{2x}{(T-2)^2}-\frac{x^2}{T-1}+\frac{2x\log x}{T-2}\right)\\
M=3 \rightarrow & F_3(x;T) = \frac{1}{288}(T-1)(T-2)^2(T-3)^3(T-4)^2(T-5)x^{T-5}\times Q,
\end{aligned}
$$

with

$$
\begin{aligned}
Q = & \frac{1}{T-5}+\frac{80x}{T-4}-\frac{24x}{(T-4)^2}-\frac{162x^2}{T-3}-\frac{72x^2}{(T-3)^3}+\frac{80x^3}{T-2}\\
& +\frac{24x^3}{(T-2)^2}+\frac{x^4}{T-1}+\frac{24x\log x}{T-4}+\frac{72x^2\log x}{(T-3)^2}-\frac{24x^3\log x}{T-2}-\frac{36x^2\log^2 x}{T-3}.
\end{aligned}
$$

For arbitrary higher values of M the pdf has the form:

$$
\begin{aligned}
& f_M(x;T) = (T-M)^M\prod_{m=1}^{M-1}(T-m)^m(T-M-m)^{M-m}\cdot\\
& \left[\sum_{m=1}^{M}\sum_{l=1}^{m}\frac{A_{ml}(-1)^{l-1}}{(l-1)!}x^{T-m-1}\log^{l-1}x+\sum_{m=1}^{M-1}\sum_{l=1}^{M-m}\frac{B_{ml}(-1)^{l-1}}{(l-1)!}x^{T-m-M-1}\log^{l-1}x\right]
\end{aligned}
$$

where the M^2 coefficients A_{ml}, B_{ml} can be obtained (e.g. with the aid of symbolic computation software) by solving the linear system of M^2 equations resulting from equating coefficients on both sides for the *polynomial identity:*

$$\sum_{m=1}^{M}\sum_{l=1}^{m} A_{ml}(x-m)^{m-l}\prod_{i\neq m}^{M}(x-i)^{i}\prod_{i=1}^{M-1}(x-M-i)^{M-i}+$$
$$+\sum_{m=1}^{M-1}\sum_{l=1}^{M-m} B_{ml}(x-M-m)^{M-m-l}\prod_{i=1}^{M}(x-i)^{i}\prod_{i\neq m}^{M-1}(x-M-i)^{M-i}=1.$$

PROOF.

For $M=1$ and integer $p\geq 0$ note that

$$\int_0^1 x^p \underbrace{(T-1)x^{T-2}}_{f_1(x;T)}\, dx = \frac{T-1}{T+p-1},$$

and hence the claimed pdf satisfies Theorem 1 and must be the searched distribution. With the help of some integral formulas for the log function it is easy to check that

$$\int_0^1 x^p f_2(x;T)\, dx = \frac{(T-1)(T-2)^2(T-3)}{(T+p-1)(T+p-2)^2(T+p-3)},$$

which again satisfies Theorem 1 and we are done. A more lengthy but trivial computation gives the case $M=3$.

These formulas and the general case for higher values of M can be derived from the following procedure. The moments of the chordal product determinant from Theorem 1 are

$$\mathscr{M}_p(T,M)=\prod_{m=1}^{M}\frac{(T-m)!(T+p-m-M)!}{(T-m-M)!(T+p-m)!}.$$

By expanding the factorials and collecting terms in the product, this can be rewritten as

$$\mathscr{M}_p(T,M)=\prod_{m=1}^{M}\left(\frac{T-m}{T+p-m}\right)^{m}\cdot\prod_{m=1}^{M-1}\left(\frac{T-M-m}{T+p-M-m}\right)^{M-m}$$

so the denominator D is the value at $x = T + p$ of the polynomial

$$D(x) = \prod_{m=1}^{M} (x-m)^m \prod_{m=1}^{M-1} (x-M-m)^{M-m}.$$

Notice that this is a product of all-different real root factors $(x-\alpha)$ with varying multiplicities, so its inverse has a partial fraction decomposition

$$\frac{1}{D} = \sum_{m=1}^{M} \sum_{l=1}^{m} \frac{A_{ml}}{(x-m)^l} + \sum_{m=1}^{M-1} \sum_{l=1}^{M-m} \frac{B_{ml}}{(x-M-m)^l},$$

for some coefficients $A_{ml}, B_{ml} \in \mathbb{R}$. Following one of the usual procedures to solve for these coefficients, for general x, multiplying by $D(x)$ on both sides yields the polynomial equation of order $M^2 - 1$:

$$\begin{aligned} &\sum_{m=1}^{M} \sum_{l=1}^{m} A_{ml} (x-m)^{m-l} \prod_{i \neq m}^{M} (x-i)^i \prod_{i=1}^{M-1} (x-M-i)^{M-i} + \\ &+ \sum_{m=1}^{M-1} \sum_{l=1}^{M-m} B_{ml} (x-M-m)^{M-m-l} \prod_{i=1}^{M} (x-i)^i \prod_{i \neq m}^{M-1} (x-M-i)^{M-i} = 1. \end{aligned}$$

Expanding and gathering terms by powers of x, one can equate the coefficient of x^0 to 1 and the coefficients of x^n, for $n = 1, \ldots, M^2 - 1$, to 0 to obtain a linear system of M^2 equations in the M^2 coefficients A_{ml}, B_{ml}, and solve for them.

Now, notice that by the Laplace transform properties for a, b nonnegative integers

$$\int_0^1 x^a \log^b x \, dx = (-1)^b \mathscr{L}[t^b](a+1) = (-1)^b \frac{b!}{(a+1)^{b+1}},$$

and thus

$$\frac{1}{(T+p-m)^l} = \frac{(-1)^{l-1}}{(l-1)!} \int_0^1 x^p x^{T-m-1} \log^{l-1} x \, dx.$$

Hence, by expressing every term of the partial fraction decomposition in this integral form, the function $f_M(x;T)$ can be identified inside the moment function written as an integral:

$$\mathscr{M}_p(T,M) = (T-M)^M \prod_{m=1}^{M-1} (T-m)^m (T-M-m)^{M-m} \cdot$$
$$\cdot \int_0^1 x^p \left[\sum_{m=1}^{M} \sum_{l=1}^{m} \frac{A_{ml}(-1)^{l-1}}{(l-1)!} x^{T-m-1} \log^{l-1} x + \right.$$
$$\left. \sum_{m=1}^{M-1} \sum_{l=1}^{M-m} \frac{B_{ml}(-1)^{l-1}}{(l-1)!} x^{T-m-M-1} \log^{l-1} x \right] dx,$$

where the factors in front of the integral all come from the numerator over D in $\mathscr{M}_p(T,M)$. This finishes the proof of the corollary. □

4. Proof of Corollary 3

Let δ be the unique solution of the equation $F_M(\delta;T) = K^{-1}$. Let G be the maximum number of points in $\mathbb{G}r(M,\mathbb{C}^T)$ that can be allocated with the claimed property. We must prove that $G \geq K$. Indeed, assume that $G < K$ an let $\mathbf{X}_1,\ldots,\mathbf{X}_G$ be a code with $\det(\mathbf{I}_M - \mathbf{X}_i^H \mathbf{X}_j \mathbf{X}_j^H \mathbf{X}_i) \geq \delta$ for all $1 \leq i, j \leq G$. We note that

$$\frac{1}{Vol(\mathbb{G}r(M,\mathbb{C}^T))} Vol\left(\cup_{i=1}^{G} \{[\mathbf{A}] \in \mathbb{G}r(M,\mathbb{C}^T) : \det(\mathbf{I}_M - \mathbf{X}_i^H \mathbf{A}\mathbf{A}^H \mathbf{X}_i) \leq \delta\}\right) \leq$$
$$\frac{1}{Vol(\mathbb{G}r(M,\mathbb{C}^T))} \sum_{i=1}^{G} Vol\left(\{[\mathbf{A}] \in \mathbb{G}r(M,\mathbb{C}^T) : \det(\mathbf{I}_M - \mathbf{X}_i^H \mathbf{A}\mathbf{A}^H \mathbf{X}_i) \leq \delta\}\right) =$$
$$G F_M(\delta;T) = \frac{G}{K} < 1,$$

and we thus deduce that there exists $[\mathbf{A}] \in \mathbb{G}r(M,\mathbb{C}^T)$ such that

$$\det(\mathbf{I}_M - \mathbf{X}_i^H \mathbf{A}\mathbf{A}^H \mathbf{X}_i) > \delta \quad \forall\, 1 \leq i \leq G.$$

But then the code $\mathbf{X}_1,\ldots,\mathbf{X}_G,\mathbf{X}_{G+1}$ with $\mathbf{X}_{G+1}=\mathbf{A}$ also satisfies the claimed property and has $G+1$ points, which contradicts the definition of G.

A. Alternative parameterization of the Grassmannian and the density function of $\tilde{\mathbf{A}}$ in $\binom{\mathbf{I}_M}{\tilde{\mathbf{A}}} \in \mathbb{G}r(M,\mathbb{C}^T)$

We recall the volume of the Grassmannian for completeness.

Lemma 1 *The volume of the Grassmannian* $\mathbb{G}r(M,\mathbb{C}^T)$ *is:*

$$Vol(\mathbb{G}r(M,\mathbb{C}^T)) = \frac{\pi^{M(T-M)}1!\cdot 2!\cdots(M-1)!}{(T-M)!\cdot(T-M+1)!\cdots(T-1)!}.$$

Moreover, the mapping sending $[\mathbf{A}]$ *to the projection matrix* $\mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ *defines a homothety with factor* $\sqrt{2}$*.*

PROOF. This is a classical fact: since the Grassmannian is formally defined as a quotient of the Stiefel manifold $\mathbb{S}t(M,\mathbb{C}^T)$ (i.e. the set of $T\times M$ complex matrices $\mathbf{X}$ such that $\mathbf{X}^H\mathbf{X}=\mathbf{I}_M$) by the unitary group $\mathscr{U}_M$, the volume of $\mathbb{G}r(M,\mathbb{C}^T)$ is the quotient of the volumes of the Stiefel and unitary matrices which is well–known, see for example [16] (note that there exist several normalizations for the Riemannian structure of the classical groups, leading to different volume formulas. We use the standard that considers $\mathbb{S}t(M,\mathbb{C}^T)$ and $\mathscr{U}_M$ as submanifolds of their ambient affine spaces, with the inherited structure). The same identification gives the tangent space, see for example [13], since the tangent space to $\mathbb{G}r(M,\mathbb{C}^T)$ is then identified with the tangent vectors to $\mathbb{S}t(M,\mathbb{C}^T)$ which are Hermitian orthogonal to the nullspace of the quotient, that is, to the set of $\dot{\mathbf{B}}\in T_{[\mathbf{A}]}\mathbb{S}t(M,\mathbb{C}^T)$ such that $\dot{\mathbf{B}}=\mathbf{A}\dot{\mathbf{U}}$ with $\dot{\mathbf{U}}$ an $M\times M$ skew-Hermitian matrix. In other words,

$$T_{[\mathbf{A}]}\mathbb{G}r(M,\mathbb{C}^T)\equiv\{\dot{\mathbf{A}}:\dot{\mathbf{A}}^H\mathbf{A}+\mathbf{A}^H\dot{\mathbf{A}}=0, \langle\dot{\mathbf{A}},\mathbf{A}\dot{\mathbf{U}}\rangle_F=0,\ \forall\dot{\mathbf{U}}\in\mathbb{C}^{M\times M}\text{ s.t. }\dot{\mathbf{U}}+\dot{\mathbf{U}}^H=0\},$$

where $\langle \mathbf{A}, \mathbf{B} \rangle_F$ denotes the Frobenius inner product between matrices $\mathbf{A}$ and $\mathbf{B}$. It is immediate to see that $\{\dot{\mathbf{A}} \in \mathbb{C}^{T\times M} : \mathbf{A}^H \dot{\mathbf{A}} = 0\}$ is contained in this tangent space, and a dimensional argument shows that the two spaces are actually equal. For the last claim of the lemma, note that the derivative of the mapping $\mathbf{A} \to \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ applied to a tangent vector $\dot{\mathbf{A}}$ such that $\mathbf{A}^H\dot{\mathbf{A}} = 0$ equals $\dot{\mathbf{A}}\mathbf{A}^H + \mathbf{A}\dot{\mathbf{A}}^H$. By unitary invariance we can assume that $\mathbf{A} = \binom{\mathbf{I}_M}{0}$ which shows that the derivative, applied at a tangent vector $\dot{\mathbf{A}} = \binom{0}{\mathbf{B}}$, gives

$$\begin{pmatrix} 0 & \mathbf{B}^H \\ \mathbf{B} & 0 \end{pmatrix},$$

thus defining a homothety as claimed.

□

Recall that a $n \times p$ complex matrix $\mathbf{X}$ is distributed as a complex matrix-variate t distribution with ν degrees of freedom when its density is given by

$$p(\mathbf{X}) = C^{-1} \det(\mathbf{I}_p + \mathbf{X}^H\mathbf{X})^{-(\nu+p+n-1)}. \tag{A.1}$$

where C is a constant to make X distributed with respect to a probability measure.

Proposition 1 *Let $T \geq 2M$. If $[\mathbf{A}]$ is uniformly distributed in $\mathbb{G}r(M, \mathbb{C}^T)$ and we write $[\mathbf{A}] = \begin{bmatrix} \mathbf{I}_M \\ \tilde{\mathbf{A}} \end{bmatrix}$ (note that there exists a unique representative of that form), then $\tilde{\mathbf{A}}$ has density*

$$\frac{1}{Vol(\mathbb{G}r(M, \mathbb{C}^T)) \det(\mathbf{I}_M + \tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^T}. \tag{A.2}$$

Hence, $\tilde{\mathbf{A}} \in \mathbb{C}^{(T-M)\times M}$ follows a matrix–variate t distribution with $\nu = 1$ degrees of freedom. In other words, for any measurable non–negative or integrable function $f : \mathbb{G}r(M, \mathbb{C}^T) \to \mathbb{C}$,

$$\int_{[\mathbf{A}]\in\mathbb{G}r(M,\mathbb{C}^T)} f([\mathbf{A}])\,d[\mathbf{A}] = \int_{\tilde{\mathbf{A}}\in\mathbb{C}^{(T-M)\times M}} \frac{f\binom{\mathbf{I}_M}{\tilde{\mathbf{A}}}}{\det(\mathbf{I}_M+\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^T}\,d\tilde{\mathbf{A}}$$
$$= \int_{\tilde{\mathbf{A}}\in\mathbb{C}^{(T-M)\times M}} \frac{f\binom{(\mathbf{I}_M+\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^{-1/2}}{\tilde{\mathbf{A}}(\mathbf{I}_M+\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^{-1/2}}}{\det(\mathbf{I}_M+\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^T}\,d\tilde{\mathbf{A}}.$$

PROOF. This result has been proved in [1, Prop. 1 and Cor. 1], by showing that both sides of the equality are equal to

$$\int_{\mathbf{W}\in\mathbb{C}^{(T-M)\times M},\|\mathbf{W}\|_{op}<1} f\left(\begin{bmatrix}\sqrt{\mathbf{I}_M-\mathbf{W}^H\mathbf{W}}\\ \mathbf{W}\end{bmatrix}\right)d\mathbf{W},$$

with $\|\cdot\|_{op}$ the operator norm. However, for completion of this paper, we sketch an elementary proof: from the second part of Lemma 1, it suffices to see that the Jacobian of

$$\tilde{\mathbf{A}} \to \phi(\tilde{\mathbf{A}}) = \begin{pmatrix}\mathbf{I}_M\\ \tilde{\mathbf{A}}\end{pmatrix}\left(\mathbf{I}_M+\tilde{\mathbf{A}}^H\tilde{\mathbf{A}}\right)^{-1}(\mathbf{I}_M\ \tilde{\mathbf{A}}^H)$$

is equal to $2^{M(T-M)}/\det(\mathbf{I}_M+\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^T$. The derivative of ϕ is easy to compute:

$$D\phi(\tilde{\mathbf{A}})\dot{\tilde{\mathbf{A}}} = \begin{pmatrix}0\\ \dot{\tilde{\mathbf{A}}}\end{pmatrix}\left(\mathbf{I}_M+\tilde{\mathbf{A}}^H\tilde{\mathbf{A}}\right)^{-1}(\mathbf{I}_M\ \tilde{\mathbf{A}}^H) + \begin{pmatrix}\mathbf{I}_M\\ \tilde{\mathbf{A}}\end{pmatrix}\left(\mathbf{I}_M+\tilde{\mathbf{A}}^H\tilde{\mathbf{A}}\right)^{-1}(0\ \dot{\tilde{\mathbf{A}}}^H)$$
$$- \begin{pmatrix}\mathbf{I}_M\\ \tilde{\mathbf{A}}\end{pmatrix}\left(\mathbf{I}_M+\tilde{\mathbf{A}}^H\tilde{\mathbf{A}}\right)^{-1}(\dot{\tilde{\mathbf{A}}}^H\tilde{\mathbf{A}}+\tilde{\mathbf{A}}^H\dot{\tilde{\mathbf{A}}})\left(\mathbf{I}_M+\tilde{\mathbf{A}}^H\tilde{\mathbf{A}}\right)^{-1}(\mathbf{I}_M\ \tilde{\mathbf{A}}^H).$$

The unitary invariance of the determinant is then clear: $\det(D\phi(\tilde{\mathbf{A}})) = \det(D\phi(\mathbf{U}\tilde{\mathbf{A}}\mathbf{V}))$ for any unitary matrices $\mathbf{U}, \mathbf{V}$ of respective sizes $T-M$ and M. Hence, we can assume that $\tilde{\mathbf{A}} = \binom{\mathbf{D}}{0}$ with $\mathbf{D}$ the diagonal matrix containing the ordered singular values of $\tilde{\mathbf{A}}$. The rest of the proof is a long computation: consider the basis given by δ_{ij} (i.e. the $(T-M)\times M$ zero–matrix with its (i,j) entry equal to 1) and $\mathbf{j}\delta_{ij}$, consider the set of vectors $D\phi(\tilde{\mathbf{A}})\delta_{ij}$ and $D\phi(\tilde{\mathbf{A}})\mathbf{j}\delta_{ij}$ and compute the volume of the parallelepiped they span in

$\mathbb{R}^{2T^2} \equiv \mathbb{C}^{T\times T}$. This task is tedious but easy since they are all pairwise orthogonal. That volume, which is equal by definition to the determinant of $D\phi(\mathbf{A})$, is $2^{M(T-M)}/\det(\mathbf{I}_M+\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^T$ as claimed.

Note that the two integrals on $\mathbb{C}^{(T-M)\times M}$ in the last part of the lemma are equal since f is a function defined in the Grassmannian and hence its value is independent of the choice of representatives. Moreover, the advantage of the last expression in the proposition is that

$$\mathbf{X} = \begin{pmatrix} (\mathbf{I}_M+\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^{-1/2} \\ \tilde{\mathbf{A}}(\mathbf{I}_M+\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^{-1/2} \end{pmatrix}$$

is a Stiefel matrix, i. e. it satisfies $\mathbf{X}^H\mathbf{X} = \mathbf{I}_M$. ∎

Acknowledgments

The authors thank the anonymous reviewers for their valuable suggestions.

Funding

This work was supported by Huawei Technologies, Sweden AB; Ministerio de Universidades (MIU), Spain, [FPU20/03563 to D.C.]; and the Ministerio de Ciencia e Innovación/Agencia Estatal de Investigación MCIN/AEI/10.13039/501100011033, Spain, [PID2020-113887GB-I00 to C.B., PID2019-104958RB-C43 to I.S.].

Data Availability Statement

No new data were generated or analysed in support of this review.

References

1. J. Álvarez-Vizoso, D. Cuevas, C. Beltrán, I. Santamaria, V. Tucek, and G. Peters. Constellations on the sphere with efficient encoding-decoding for noncoherent communications. *arXiv:2209.04172 [cs.IT].*
2. D. Amelunxen. *Geometric analysis of the condition of the convex feasibility problem.* PhD thesis. Institute of Mathematics, University of Paderborn, 2011.

3. A. Anderson, M. Dostert, P. J. Grabner, R. W. Matzke, and T. A. Stepaniuk. Riesz and green energy on projective spaces. *arXiv:2204.04015*, 2022.
4. A. Barg and D. Y. Nogin. Bounds on packings of spheres in the Grassmann manifold. *IEEE Trans. Inf. Theory*, 48(9):2450–2454, Sept. 2002.
5. C. Beltrán. Estimates on the condition number of random rank-deficient matrices. *IMA J. Numer. Anal.*, 31 (1):25–39, 2011. ISSN 0272-4979.
6. C. Beltrán and U. Etayo. The projective ensemble and distribution of points in odd-dimensional spheres. *Constr. Approx.*, 48(1):163–182, 2018. ISSN 0176-4276.
7. S. V. Borodachov, D. P. Hardin, and E. B. Saff. *Discrete energy on rectifiable sets*. Springer Monographs in Mathematics. Springer, New York, 2019. ISBN 978-0-387-84807-5; 978-0-387-84808-2.
8. M. Brehler and M. Varanasi. Asymptotic error probability analysis of quadratic receivers in Rayleigh fading channels with application to a unified analysis of coherent and noncoherent space-time receivers. *IEEE Trans. Inf. Theory*, 47(6):2383–2399, 2001.
9. P. Bürgisser and A. Lerario. Probabilistic Schubert calculus. *J. Reine Angew. Math.*, 760:1–58, 2020. ISSN 0075-4102. doi: 10.1515/crelle-2018-0009. URL https://doi.org/10.1515/crelle-2018-0009.
10. J. H. Conway, R. H. Hardin, and N. J. A. Sloane. Packing lines, planes, etc.: Packings in Grassmannian spaces. *Experimental Mathematics*, 5(2):139–159, 1996.
11. D. Cuevas, J. Álvarez-Vizoso, C. Beltrán, I. Santamaria, V. Tuvcek, and G. Peters. Union bound minimization approach for designing Grassmannian constellations. *IEEE Transactions on Communications*, 71(4):1940–1952, 2023. doi: 10.1109/TCOMM.2023.3244965.
12. W. Dai, Y. Liu, and B. Rider. Quantization bounds on Grassmann manifolds and applications to MIMO communications. *IEEE Trans. Inf. Theory*, 54(3):1108–1123, March 2008.
13. A. Edelman, T. A. Arias, and S. T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.*, 20(2):303–353, 1999. ISSN 0895-4798.
14. O. Henkel. Sphere-packing bounds in the grassmann and stiefel manifolds. *IEEE Transactions on Information Theory*, 51:3445–3456, 2005.
15. B. Hochwald and T. Marzetta. Unitary space-time modulation for multiple-antenna communication in Rayleigh flat-fading. *IEEE Trans. Inf. Theory*, 46(6):543–564, 2000.
16. L. K. Hua. *Harmonic Analysis of Functions of Several Complex Variables in the Classical Domains*. American Mathematical Society, 1963.

17. J. Jasper, E. J. King, and D. G. Mixon. Game of Sloanes: Best known packings in complex projective space, 2019.
18. M. L. McCloud, M. Brehler, and M. Varanasi. Signal design and convolutional coding for noncoherent space-time communication on the block-Rayleigh-fading channel. *IEEE Trans. Inf. Theory*, 48(5):1186–1194, 2002.
19. A. Papoulis and U. Pillai. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill Education, 4th ed. edition, 2001.
20. D. Ramírez, I. Santamaría, and L. Scharf. *Coherence: In Signal Processing and Machine Learning*. Springer Nature Switzerland AG, 2022.
21. R. A. Rankin. The closest packing of spherical caps in n dimensions. *Proc Glasgow Math Assoc*, 2:139–144, 1955.
22. C. E. Shannon. Probability of error for optimal codes in a Gaussian channel. *Bell Syst. Tech. J.*, 38(3): 611–656, May 1959.
23. M. D. Springer and W. E. Thompson. The distribution of products of Beta, Gamma and Gaussian random variables. *SIAM Journal on Applied Mathematics*, 18(4):721–737, 1970.
24. M. S. Srivastava and C. G. Khatri. *An Introduction to Multivariate Statistics*. North Holland, 1979.
25. L. Zheng and D. Tse. Communication on the Grassmann manifold: a geometric approach to the noncoherent multiple-antenna channel. *IEEE Trans. Inf. Theory*, 48(2):359–383, Feb. 2002.