Union Bound Minimization Approach for Designing Grassmannian Constellations

Diego Cuevas, Javier Álvarez-Vizoso, Carlos Beltrán, Ignacio Santamaria, *Senior Member, IEEE*, Vít Tuček, and Gunnar Peters

***Abstract*—In this paper, we propose an algorithm for designing unstructured Grassmannian constellations for noncoherent multiple-input multiple-output (MIMO) communications over Rayleigh block-fading channels. Unlike the majority of existing unitary space-time or Grassmannian constellations, which are typically designed to maximize the minimum distance between codewords, in this work we employ the asymptotic pairwise error probability (PEP) union bound (UB) of the constellation as the design criterion. In addition, the proposed criterion allows the design of MIMO Grassmannian constellations specifically optimized for a given number of receiving antennas. A rigorous derivation of the gradient of the asymptotic UB on a Cartesian product of Grassmann manifolds, is the main technical ingredient of the proposed gradient descent algorithm. A simple modification of the proposed cost function, which weighs each pairwise error term in the UB according to the Hamming distance between the binary labels assigned to the respective codewords, allows us to jointly solve the constellation design and the bit labeling problem. Our simulation results show that the constellations designed with the proposed method outperform other structured and unstructured Grassmannian designs in terms of symbol error rate (SER) and bit error rate (BER), for a wide range of scenarios.**

***Index Terms*—Noncoherent communications, MIMO communications, Grassmannian constellations, pairwise error probability (PEP), union bound (UB), bit-labeling.**

I. Introduction

In multiple-input multiple-output (MIMO) communications systems, the channel state information (CSI) is typically estimated at the receiver side by sending a few known pilots and then used for decoding at the receiver and/or for precoding at the transmitter. These are known as coherent schemes. When the channel remains approximately constant over a long coherence time (slowly fading scenarios), the channel capacity for coherent MIMO systems is known to increase linearly with the minimum number of transmit and receive antennas at high signal-to-noise (SNR) ratios [1], [2]. In fast fading scenarios, however, to obtain an accurate channel estimate would require pilots to occupy a disproportionate fraction of communication resources. Even in slowly-varying channels, CSI acquisition by orthogonal pilot-based schemes can result in significant overheads in massive MIMO systems [3], and the performance of coherent massive MIMO systems can be degraded by channel aging [4]. These scenarios motivate the use of noncoherent MIMO communications schemes in which neither the transmitter nor the receiver have any knowledge about the instantaneous CSI (although they might have some knowledge about the statistical or long-term CSI such as its fading distribution).

We consider in this work a block-fading model in which the MIMO channel matrix with M transmit and N receive antennas remains constant during a T-symbol coherence interval, after which it changes to a new independent realization for another T symbols. Under additive Gaussian noise, it was proved by Hochwald and Marzetta [5], [6] that the $T \times M$ space-time transmit matrices that achieve the ergodic noncoherent capacity for this channel model can be factored as the product of an isotropically distributed $T \times M$ unitary matrix and a diagonal $M \times M$ matrix with real nonnegative entries. Further, when $T >> M$ the nonzero entries of the diagonal matrix take the same value, showing that in this regime it is optimal to transmit unitary space time codewords $\mathbf{X}^H\mathbf{X} = \mathbf{I}_M$. Using the same signal model, Zheng and Tse proved that at high SNR and when $T \geq 2M$, ergodic capacity can be achieved by transmitting isotropically distributed unitary matrices. The pre-log factor in the high-SNR capacity expression is $M^*(1-M^*/T)$, where $M^* = \min\{M, N\}$ is the minimum between the transmit and receive number of antennas. This result shows that despite the absence of *instantaneous* CSI at the receiver, noncoherent MIMO communication systems can achieve a significant fraction of the coherent capacity at high SNR. Moreover, the noncoherent multiplexing gain approaches the coherent multiplexing gain as $T \to \infty$. Under the assumption of equal-energy signals, it is also shown in [7] that unitary space-time matrices also minimize the asymptotic union bound on the error probability. Therefore, the use of unitary space-time constellations in noncoherent MIMO communications is well justified.

In noncoherent MIMO communication systems information is carried by the column span of the transmitted $T \times M$ matrix, $\mathbf{X}$, which is not affected by the MIMO channel $\mathbf{H}$. In other words, the column span of $\mathbf{X}$ is identical to the column space

This work was supported by Huawei Technologies, Sweden under the project GRASSCOM. The work of D. Cuevas was also partly supported under grant FPU20/03563 funded by Ministerio de Universidades (MIU), Spain. The work of Carlos Beltrán was also partly supported under grant PID2020-113887GB-I00 funded by MCIN/ AEI /10.13039/501100011033. The work of I. Santamaria was also partly supported under grant PID2019-104958RB-C43 (ADELE) funded by MCIN/ AEI /10.13039/501100011033.

D. Cuevas, J. Álvarez Vizoso and I.Santamaria are with the Department of Communications Engineering, Universidad de Cantabria, 39005 Santander, Spain (e-mail: diego.cuevas@unican.es; javier.alvarezvizoso@unican.es; i.santamaria@unican.es).

C. Beltrán is with the Department of Mathematics, Statistics and Computing, Universidad de Cantabria, 39005 Santander, Spain (e-mail: carlos.beltran@unican.es).

V. Tucek and G. Peters are with the Department of Wireless Algorithms, Huawei Technologies, 16440 Kista, Sweden (email: vit.tucek@huawei.com; gunnar.peters@huawei.com)

of $\mathbf{XH}$. The problem of designing noncoherent codebooks formed by unitary space-time matrices thus becomes equivalent to finding optimal (according to some appropriate distance or metric) packings in Grassmann manifolds [8], [9]. While at high SNR, or for codebooks of very high cardinality, the use of random codebooks (with codewords generated by uniformly sampling the Grassmann manifold) is essentially optimal, this is not the case for constellations of moderate cardinality, or at moderate SNRs. For this reason, there has been extensive research on the design of noncoherent constellations as optimal packings on the Grassmann manifold [10]–[15]. Existing constellation designs can be generically categorized into two groups: structured or unstructured. Structured designs impose some kind of structure on the constellation points through parameterized mappings such as the Exp-Map design in [12], algebraic constructions such as the Fourier-based constellation in [16], structured partitions of the Grassmannian like the recently proposed Cube-Split constellation [13], or designs based on group representations [17], [18]. Structured designs may simplify either codeword generation/storage or detection, but the packing efficiency is lower than that achieved by unstructured codes, which in turn translates into poorer performance in terms of symbol error rate (SER). Since our goal is to design quasi-optimal constellations in terms of SER/BER, in this paper we focus on unstructured constellations designed through numerical optimization methods.

Among the unstructured designs we can mention the alternating projection method [14], which enforces in each iteration both structural and spectral properties of the Gram matrix formed by the inner products between codewords, as well as the numerical methods in [10], [11], which minimize certain distance measures on the Grassmannian. For example, [10] employs as a suitable distance metric the chordal distance between subspaces, while [11] uses the spectral distance (the cosine square of the minimum principal angle). The method proposed in [10] is computationally costly because it works on a Grassmann manifold whose dimension grows with the number of constellation points. All these methods ultimately aim at maximizing the minimum distance between codewords, usually using the chordal distance as a suitable metric. As shown in [19], however, the chordal distance is related to the pairwise error probability only at low SNR. The criterion that emerges from the PEP analysis at high SNR is the so-called diversity product [19], which is the *product of the squared sines* of the principal angles between subspaces, while the squared chordal distance is the *sum of the squared sines*. In addition, as it has been shown in [20], maximizing the minimum chordal distance does not guarantee obtaining full-diversity codes when $M > 1$. That is, it is not guaranteed that the M-dimensional subspaces designed according to the chordal distance do not intersect each other. These reasons make the diversity product criterion, also called coherence criterion, in [19] more appropriate for designing Grassmannian constellations. A method for designing Grassmannian constellations according to the diversity product criterion has been recently proposed in [21].

A second drawback of the existing design criteria is that most of them optimize only for the worst case error. Although this may be a reasonable approximation at high SNR, multiple experiments show that at low or moderate SNRs there are several pairs of codewords that contribute to the SER, not just the two closest codewords. To achieve better SER performance a more appropriate optimization criterion is the (asymptotic) union bound (UB), which is a sum of diversity products between pairs of codewords. To the best of our knowledge, the only work that has considered a metric based on UB as a design criterion for noncoherent constellations is the method proposed by McCloud, Brehler and Varanasi in [20]. This method, however, relies on an overparametrization of the Grassmann manifold which incurs a higher computational cost and causes a loss of effectiveness due to optimizing over higher dimensional spaces than necessary.

In this paper, we present a gradient descent algorithm for minimizing the UB that operates directly on a product of Grassmann manifolds. The main contributions are summarized as follows:

- We revisit the calculation of the PEP in non-coherent communications, deriving an exact formula for the PEP that can be easily evaluated by means of an integral. The limit of the integral when the SNR tends to infinity provides an asymptotic PEP expression that coincides with the expression originally obtained by Brehler and Varanasi in [7]. Nevertheless, our proof technique is different from that of [7].
- We propose a Grassmannian constellation design criterion based on the asymptotic union bound on the error probability. The method applies a gradient descent technique on a Cartesian product of Grassmann manifolds and is particularly effective for designing Grassmannian constellations of relatively small cardinality (signal constellations with up to 2048 codewords can be designed in a few minutes).
- Unlike existing methods, the proposed UB cost function, and consequently the constellation obtained, depend not only on the coherence time T and the number of transmit antennas M, but also on the number of receive antennas N. In noncoherent massive MIMO systems, for instance, using unitary space-time codes specifically designed for a given number of receive antennas may have a significant impact.
- By weighting the pairwise UB terms according to the Hamming distance between the binary labels assigned to the respective codewords, the proposed method jointly solves the constellation design and bit labeling problem with excellent performance at no extra cost.
- We verify by simulations that the proposed UB Grassmannian constellations outperform existing unstructured and structured Grassmannian constellations in terms of symbol and bit error probability in a variety of scenarios.

The remainder of this paper is organized as follows. The system model is presented in Section II. In Section III we first revisit the PEP analysis for noncoherent multiantenna systems following a proof technique that is different from those employed in the foundational works of Hochwald and Marzetta [6], and Brehler and Varanasi [7]. Then, we describe our codebook design and provide the mathematical details for the computation of the gradient of the UB cost function on the Grassmann manifold. We next propose a joint constellation design and bit labeling scheme in Section IV. A comprehensive set of numerical simulation results to assess the performance of the proposed method in terms of symbol and bit error rates is provided in Section V. Finally, Section VI concludes the paper. In addition, the paper contains a set of appendices describing required background material on the Grassmann and Stiefel manifolds as well as the proofs of the mathematical results.

Notation: In this paper, matrices are denoted by bold-faced upper case letters, column vectors are denoted by bold-faced lower case letters, and scalars are denoted by light-faced lower case letters, e.g., a matrix $\mathbf{M}$, a vector $\mathbf{v}$, and a scalar s. The superscripts $(\cdot)^T$ and $(\cdot)^H$ denote transpose and Hermitian conjugate, respectively. The trace and determinant of a matrix $\mathbf{A}$ will be denoted, respectively, as $\mathrm{tr}(\mathbf{A})$ and $\det(\mathbf{A})$, and $\mathbf{I}_n$ denotes the identity matrix of size n. A complex Gaussian vector in $\mathbb{C}^n$ with zero mean and covariance matrix $\mathbf{R}$ is denoted as $\mathbf{x} \sim \mathcal{CN}_n(\mathbf{0}, \mathbf{R})$. The factorial of a non-negative integer n is denoted $n!$ Let us also recall that $\Gamma(n+1) = n!$, where Γ is the gamma function. The double factorial of a number n, denoted as $n!!$, is the product of all integers from 1 to n that have the same parity (odd or even) as n. For instance, for n odd, $n!! = n(n-2)(n-4)\cdots 3\cdot 1$. The Frobenius norm of a matrix $\mathbf{A}$ is $\|\mathbf{A}\|_F$, its associated Hermitian inner product is $\langle \cdot, \cdot \rangle_F$, so that for two matrices $\mathbf{A}, \mathbf{B}$ of the same size we have $\|\mathbf{A}\|_F^2 = \langle \mathbf{A}, \mathbf{A} \rangle_F = \mathrm{tr}(\mathbf{A}^H \mathbf{A})$. As usual, $\Re(z) = (z + z^*)/2$ is the real part of a complex $z \in \mathbb{C}$ whose conjugate number is denoted z^*. The complex Grassmann manifold of M-dimensional subspaces of the T-dimensional complex vector space $\mathbb{C}^T$ is denoted as $\mathbb{G}(M, \mathbb{C}^T)$ and the complex Stiefel manifold of unitary M-frames in $\mathbb{C}^T$ is denoted as $\mathbb{S}t(M, \mathbb{C}^T)$. Points in the Grassmannian are denoted as $[\mathbf{A}]$, where $\mathbf{A}$ is a unitary basis for that subspace, and $\mathbf{P}_\mathbf{A} = \mathbf{A}\mathbf{A}^H$ denotes the orthogonal projection onto $[\mathbf{A}]$. Some background material about the Stiefel and Grassmann manifolds, which is needed for the paper, is relegated to Appendix A.

II. System Model

A. System Model

We consider a noncoherent MIMO communication system in which a transmitter with M antennas transmits to a receiver equipped with N antennas over a frequency-flat block-fading channel with coherence time T symbol periods, with $T \geq 2M$. That is, the channel matrix $\mathbf{H} \in \mathbb{C}^{M \times N}$ remains constant during each coherence block of T symbols, and changes to an independent realization in the next block. The MIMO channel $\mathbf{H}$ is assumed to be Rayleigh with entries $h_{ij} \sim \mathcal{CN}(0,1)$ and unknown to both the transmitter and the receiver.

Within a coherence block, the transmitter sends a unitary matrix $\mathbf{X} \in \mathbb{C}^{T \times M}$, $\mathbf{X}^H\mathbf{X} = \mathbf{I}_M$, that is a basis for the subspace $[\mathbf{X}] \in \mathbb{G}\left(M, \mathbb{C}^T\right)$. The transmitted unitary matrices are sampled uniformly from a codebook $\mathcal{C} = \{\mathbf{X}_1, \ldots, \mathbf{X}_K\}$ of size K that here is assumed to be unstructured. For a given transmission rate R (in b/s/Hz), the codebook is composed of $K = 2^{RT}$ unitary matrices.

The signal at the receiver $\mathbf{Y} \in \mathbb{C}^{T \times N}$ is

$$\mathbf{Y} = \mathbf{X}\mathbf{H} + \sqrt{\frac{M}{T\rho}}\mathbf{W}, \tag{1}$$

where $\mathbf{W} \in \mathbb{C}^{T \times N}$ represents the additive Gaussian noise, modeled as $w_{ij} \sim \mathcal{CN}(0,1)$, and ρ represents the signal-to-noise-ratio (SNR).

Assuming equiprobable codewords, the optimal Maximum Likelihood (ML) detector that minimizes the probability of error is

$$\hat{\mathbf{X}} = \arg\max_{\mathbf{X} \in \mathcal{C}} \mathrm{tr}\left(\mathbf{Y}^H \mathbf{P}_\mathbf{X} \mathbf{Y}\right). \tag{2}$$

where $\mathbf{P}_\mathbf{X} = \mathbf{X}\mathbf{X}^H$ is the projection matrix onto the subspace $[\mathbf{X}]$. Each codeword carries $\log_2(K) = RT$ bits of information.

III. Union Bound Criterion

A. PEP Analysis

Existing criteria for the design of unstructured Grassmannian constellations arise from and are motivated by the asymptotic analysis of the pairwise error probability (PEP) $P_e(\mathbf{X}_i, \mathbf{X}_j)$ (probability of mistaking $\mathbf{X}_i$ for $\mathbf{X}_j$ when the optimal decoder is applied). The PEP analysis performed below follows a somewhat different, perhaps simpler, proof technique than that employed in the classic works [6], [7], so it may be of independent interest. The final asymptotic PEP expression and the conclusions obviously are the same.

Suppose that two $T \times M$ unitary space-time codewords $\mathbf{X}_i$ and $\mathbf{X}_j$ are transmitted with equal probability and decoded with an ML receiver. Then, the pairwise error probability (PEP) is [6]

$$P_e(\mathbf{X}_i, \mathbf{X}_j) = \sum_{n=1}^{m} \mathrm{Res}_{w=ja_n}\left(\frac{-1}{w + j/2} \cdot \prod_{m=1}^{M}\left(\frac{1+\rho T/M}{(\rho T/M)^2(1-d_m^2)(w^2+a_m^2)}\right)^N\right), \tag{3}$$

where Res is the residue, $1 \geq d_1 \geq \ldots \geq d_M \geq 0$ are the singular values of the $M \times M$ matrix $\mathbf{C}_{ij} = \mathbf{X}_i^H\mathbf{X}_j$ (equivalently, they are the cosines of the principal angles between the subspaces), and a_m is given by

$$a_m^2 = \frac{1}{4} + \frac{1+\rho T/M}{(\rho T/M)^2(1-d_m^2)}. \tag{4}$$

We assume in (3) that all d_m are different from 1, which is needed for any full-diversity unitary space-time code. This is the expression in [6, (B.9)] (see also [7]). In [6], the authors took the real part of (3) and applied the Chernoff bound to get an upper bound for the pairwise error probability. In [7] the authors perform the integration in the complex plane and derived tight asymptotic expressions for the case $\rho \to \infty$. Here, we first present an alternative expression for the $P_e(\mathbf{X}_i, \mathbf{X}_j)$ that is more amenable to numerical optimization. Further, this alternative expression allows us to derive in a simple way a high-SNR approximation for the PEP.

Theorem 1 *The following equality holds:*

$$P_{\mathrm{e}}(\mathbf{X}_i,\mathbf{X}_j)=\frac{1}{\pi}\int_0^{\pi/2}\prod_{m=1}^{M}\left(\frac{1}{1+\frac{(\rho T/M)^2(1-d_m^2)}{4(1+\rho T/M)\cos^2\theta}}\right)^N d\theta. \quad (5)$$

Proof: The proof is given in Appendix B. □

Corollary 1 *The slope of* (5) *for large* ρ*, which is related to the diversity gain of the code, is*

$$\lim_{\rho\to\infty}\frac{d\log P_e(\mathbf{X}_i,\mathbf{X}_j)}{d\log\rho}=-ND, \quad (6)$$

where D *is the number of singular values of the* $M\times M$ *matrix* $\mathbf{X}_i^H\mathbf{X}_j$ *that are different from 1.*

Proof: The proof is given in Appendix C. □

The parameter D in Corollary 1 may well be called the *diversity* of the pair of codewords. It was shown in [8] that given a coherence time T, the degrees of freedom increase with M up to $M=T/2$ which is optimal.

Therefore we restrict our codebook constructions to the case $M\le T/2$. Further, we restrict our analysis to full diversity codes for which $D=M\le T/2$ in (6).

Corollary 2 *The pairwise error probability satisfies as* $\rho\to\infty$*:*

$$P_e(\mathbf{X}_i,\mathbf{X}_j)\sim\rho^{-NM}\frac{1}{2}\left(\frac{4M}{T}\right)^{NM}\cdot\left(\frac{1}{\prod_{m=1}^{M}(1-d_m^2)}\right)^N\frac{(2NM-1)!!}{(2NM)!!}, \quad (7)$$

where we have assumed a full diversity code with $M\le T/2$ *for which the number of singular values of* $\mathbf{X}_i^H\mathbf{X}_j$ *that are different from* 1 *is exactly* M*. The precise meaning of* $\sim$ *in* (7) *is:*

$$\begin{aligned}&\lim_{\rho\to\infty}[\rho^{NM}P_e(\mathbf{X}_i,\mathbf{X}_j)]\\&=\frac{1}{2}\left(\frac{4M}{T}\right)^{NM}\left(\frac{1}{\prod_{m=1}^{M}(1-d_m^2)}\right)^N\frac{(2NM-1)!!}{(2NM)!!}\\&=C.\end{aligned} \quad (8)$$

That is to say, $\log P_e(\mathbf{X}_i,\mathbf{X}_j)\approx\log C-NM\log\rho$*. Hence, in order to minimize the probability of error for large* ρ *one must maximize*

$$\prod_{m=1}^{M}(1-d_m^2)=\det\left(\mathbf{I}_M-\mathbf{X}_i^H\mathbf{X}_j\mathbf{X}_j^H\mathbf{X}_i\right). \quad (9)$$

Proof: The proof is given in Appendix C. □

The result in Corollary 2 is the same as the result in [7]. The design criterion that emerges from the high-SNR analysis of the PEP is the maximization of the so-called diversity product defined in [19] as

$$DP=\min_{i\neq j}\det\left(\mathbf{I}_M-\mathbf{X}_i^H\mathbf{X}_j\mathbf{X}_j^H\mathbf{X}_i\right). \quad (10)$$

B. Union Bound Minimization

We have seen in the previous subsection that the criterion for minimizing the PEP at high SNR must be to minimize the coherence between the transmitted subspaces. Since an unstructured codebook with K codewords may produce $K(K-1)/2$ distinct PEPs, it is customary to simplify the problem and consider only the minimization of the worst PEP. This simplification is supported by the fact that when the SNR grows, the worst PEP dominates the behavior of the SER curve which, ultimately, is the figure of merit to be optimized. Although this simplification is in general reasonable, some considerations can be made. For the additive white Gaussian channel (AWGN), the asymptotic behavior of the PEP with the SNR is proportional to the Gaussian Q-function, whose argument depends only on the minimum Euclidean distance between symbols and the SNR [22]. In this case, the exponential decrease of the Q-function with the SNR makes the worst case asymptotic PEP an asymptotically tight lower bound of the SER. For block-fading Rayleigh channels, such as those considered in this work, the PEP decreases polynomially (not exponentially) with the SNR (cf. (7)). This makes that considering only the worst case probability provides a looser asymptotic SNR bound for block-fading Rayleigh channels. Therefore, for Rayleigh block-fading channels and for the practical SNRs at which communications systems operate, it may be appropriate to consider other codeword pairs (not just the worst one) that contribute to the SER. One of the contributions of this work is precisely to consider the PEP union bound, which takes into account all codeword pairs, as the cost function to be optimized.

We show in this work that, for moderate cardinality constellations, it is possible to consider the PEP union bound, which takes into account all codeword pairs, as the cost function to be optimized. The symbol error probability for equiprobable symbols can be bounded by

$$P_e\le\frac{1}{K}\sum_{i=1}^{K}\sum_{j=i+1}^{K}P_e\left(\mathbf{X}_i,\mathbf{X}_j\right), \quad (11)$$

where $P_e\left(\mathbf{X}_i,\mathbf{X}_j\right)$ is the asymptotic PEP in (7). The right hand side in (11) is the union bound (UB). From Corollary 2, at high-SNR the UB is, up to a constant,

$$\mathrm{UB}(\mathbf{X}_1,\ldots,\mathbf{X}_K)=\sum_{i<j}\det\left(\mathbf{I}_M-\mathbf{X}_i^H\mathbf{X}_j\mathbf{X}_j^H\mathbf{X}_i\right)^{-N}, \quad (12)$$

which is the function we seek to minimize in this work. Specifically, the optimization problem is

$$\underset{\mathbf{X}_1,\ldots,\mathbf{X}_K}{\operatorname{argmin}}\quad\sum_{i<j}\det\left(\mathbf{I}_M-\mathbf{X}_i^H\mathbf{X}_j\mathbf{X}_j^H\mathbf{X}_i\right)^{-N}. \quad (13)$$

It is interesting to note that unlike the diversity product (DP) criterion in (10), now the number of receive antennas N appears explicitly in the UB criterion (13). It is clear that for the DP criterion maximizing $\min_{i\neq j}\left(\det(\mathbf{I}_M-\mathbf{X}_i^H\mathbf{X}_j\mathbf{X}_j^H\mathbf{X}_i)\right)^N$ is equivalent to maximizing $\min_{i\neq j}\det(\mathbf{I}_M-\mathbf{X}_i^H\mathbf{X}_j\mathbf{X}_j^H\mathbf{X}_i)$, so the number of receive antennas plays no role in the optimization. The situation is different with the UB criterion, for which there will be different optimal codebooks for each set of values (T, M, N, K).

Clearly, the complexity of the UB criterion grows with the number of codewords K since the number of terms in (13) is $K(K-1)/2$. Despite this difficulty, it is both of theoretical and practical interest to use the design criterion that best characterizes the SER, which is the actual figure of merit for communications. Notice also that the constellations can be designed offline with sufficient computational resources and, once they have been designed, the encoding and decoding have the same complexity as for any other unstructured design.

The UB criterion was already considered in [20], where an algorithm is suggested to generate codebooks attaining low values of UB. However, the method in [20] relies on an overparametrization of the Grassmannian, which implies high computational cost and a loss of efficiency. Moreover, the iterative method in [20] is a greedy suboptimal approach, which at each step adds a new codeword to a previously optimized constellation of smaller size. This suboptimal approach is termed in [20] *successive updates*. Different from this proposal, we propose a gradient descent approach that operates directly on the manifold, thus eliminating the need of extra parameters associated to the overparametrization used in [20].

The essential technical aspect of the algorithm is the calculation of the gradient of

$$\begin{aligned}\mathrm{UB}(\mathbf{X}_1,\ldots,\mathbf{X}_K) &= \sum_{i<j} \det\left(\mathbf{I}_M - \mathbf{X}_i^H\mathbf{X}_j\mathbf{X}_j^H\mathbf{X}_i\right)^{-N} \\ &= \sum_{i<j} \det\left(\mathbf{I}_M - \mathbf{X}_j^H\mathbf{X}_i\mathbf{X}_i^H\mathbf{X}_j\right)^{-N},\end{aligned} \tag{14}$$

in the tangent space. We have the following lemma:

Lemma 3 *The gradient of the function* UB *defined in* (12) *(seen as a function defined in $\mathbb{G}(M,\mathbb{C}^T)\times\cdots\times\mathbb{G}(M,\mathbb{C}^T)$, K-times product), at any fixed $[\mathbf{X}_1],\ldots,[\mathbf{X}_K]$, is a vector $\nabla\mathrm{UB}([\mathbf{X}_1],\ldots,[\mathbf{X}_K]) = (\dot{\mathbf{X}}_1,\ldots,\dot{\mathbf{X}}_K)$, where each $\dot{\mathbf{X}}_i \in T_{[\mathbf{X}_i]}\mathbb{G}(M,\mathbb{C}^T)$ is given by:*

$$\begin{aligned}\dot{\mathbf{X}}_i = (\mathbf{I}_T - \mathbf{X}_i\mathbf{X}_i^H)\sum_{j\neq i} 2N\det(\mathbf{I}_M - \mathbf{X}_j^H\mathbf{X}_i\mathbf{X}_i^H\mathbf{X}_j)^{-N} \\ \cdot\,\mathbf{X}_j(\mathbf{I}_M - \mathbf{X}_j^H\mathbf{X}_i\mathbf{X}_i^H\mathbf{X}_j)^{-1}\mathbf{X}_j^H\mathbf{X}_i.\end{aligned} \tag{15}$$

Proof: The proof is given in Appendix D. □

C. Algorithm Implementation

We propose a gradient descent algorithm to minimize the UB on the Grassmann manifold that, at each iteration, performs the following steps:

1) Compute the gradient $\nabla\mathrm{UB}(\mathbf{X}_1,\ldots,\mathbf{X}_K) = (\dot{\mathbf{X}}_1,\ldots,\dot{\mathbf{X}}_K)$ using (15), and its norm $||\nabla\mathrm{UB}(\mathbf{X}_1,\ldots,\mathbf{X}_K)||_1 = \sum_i ||\dot{\mathbf{X}}_i||_F$.
2) Move each $\mathbf{X}_k$ a certain amount, defined by the step-size μ, in the direction of maximally decreasing UB defined by the gradient
$$\tilde{\mathbf{X}}_k = \mathbf{X}_k - \mu\frac{\dot{\mathbf{X}}_k}{\sum_i ||\dot{\mathbf{X}}_i||_F}. \tag{16}$$
3) Retract $\tilde{\mathbf{X}}_k$ to the manifold by computing the $\mathbf{Q}$ factor in its reduced QR decomposition, which will be the new $\mathbf{X}_k$.

An important aspect of the proposed algorithm is that the value of μ is adapted using a line-search procedure to speed up convergence. The rate at which we change μ is controlled by a parameter $\alpha \in [1, 1.1]$. After every iteration, if the algorithm does not improve the UB of the codebook, μ is decreased as $\mu = \mu/\alpha$. Otherwise, if there is an improvement, μ is increased as $\mu = \mu \cdot \alpha$. The proposed algorithm uses three stopping criteria: a maximum number of iterations, a minimum value of the stepsize μ and a minimum improvement of the value of the UB.

D. Complexity Analysis

In this subsection, we provide an in-depth complexity analysis of the proposed algorithm and a comparison with other existing techniques.

Each gradient descent step of the proposed method requires the computation of: the objective function given in 12, the Riemannian gradient defined in Lemma 3, its norm, and the retraction to the manifold. The operations of computing $\mathrm{UB}(\mathbf{X}_1,\ldots,\mathbf{X}_K)$ are three matrix multiplications, one determinant, and the summation over all pairs of codewords. This yields a complexity order of $\mathcal{O}\left(K^2\left(M^3+3TM^2\right)\right)$. The operations for computing $\nabla\mathrm{UB}(\mathbf{X}_1,\ldots,\mathbf{X}_K)$ are six matrix multiplications, one determinant, one matrix inversion, and the summation over all pairs of codewords, which results in a complexity order of $\mathcal{O}\left(K^2\left(3M^3+5TM^2\right)\right)$. The order of computing $||\nabla\mathrm{UB}||_1$ is $\mathcal{O}\left(K\left(T^2M^2\right)\right)$. Finally, the complexity of the QR decomposition for the retraction to the manifold is $\mathcal{O}\left(KT^2M\right)$, which can be neglected. In summary, the total complexity order of each iteration of the proposed UB-Opt algorithm is $\mathcal{O}\left(K^2\left(4M^3+8TM^2\right)+K\left(T^2M^2\right)\right)$.

On the other hand, the computational complexity of the alternating projection method [14] requires in each iteration an SVD of a $KM \times KM$ matrix, which results in an order of complexity of $\mathcal{O}\left(K^3M^3\right)$. We can see that the order of complexity for this algorithm is at least a monomial of degree six in the constellation parameters, being greater than the leading order of our algorithm, which is five. The GMO-Chordal algorithm needs to compute in each iteration the objective function (chordal distance) for all codeword pairs, its gradient for the closest pair of codewords, and the retraction to the Grassmann manifold. The objective function for all pairs of codewords yields a leading complexity order of $\mathcal{O}\left(K^2TM^2\right)$, the gradient results in an order of $\mathcal{O}\left(4KT^2M\right)$, and the retraction requires $\mathcal{O}\left(KT^2M\right)$. The total leading order of complexity of GMO-Chordal is then $\mathcal{O}\left(K^2TM^2\right)$. Thus, the leading order of complexity of this algorithm is the same as the one for the proposed UB-Opt method. However, the latter performs a higher number of operations of complexity order five, which shall be justified by its performance advantage shown in Section V.

IV. Bit-labeling

An important practical aspect for the use of non-coherent communications based on Grassmannian constellations is the design of the bit-to-symbol mapping method. That is, it is necessary to assign a binary label to each point in the constellation. The idea is that constellation points with small distance, or large PEP, are given binary labels with small Hamming distance. In coherent communications with standard constellations carved on a regular lattice (e.g., 16-QAM), the optimal -in the sense of minimizing the average bit error rate (BER) probability- bit-to-symbol mapping or labeling scheme is the well-known Gray mapping. This optimal labeling ensures that adjacent constellation points will only differ in one bit, so in the most likely case of mistaking a symbol with its closest one only one bit will be wrong. However, binary labeling of points in a Grassmannian constellation is a notoriously more difficult problem, especially in the case of unstructured constellations. For structured constellations, it is sometimes possible to find an optimal Gray-like labeling scheme, see for instance [23] [15]. For other structured codes, like the Reed-Muller Grassmannian constellations [24], quasi-Gray labeling schemes have been recently proposed [25]. For unstructured Grassmannian constellations, the number of symbol neighbors is usually larger than the number of binary labels, so Gray labeling is not possible. When Gray labeling is not possible, finding the optimal labeling for a given constellation $\mathcal{C}$ of cardinality K would require defining a cost function (such as the average bit error rate) and performing an exhaustive search over $K!$ different labelings. Therefore, one typically resorts to suboptimal labeling schemes [26] [27] [28]. All these suboptimal schemes decouple the problem of designing a good constellation from the binary labeling problem. An important advantage of the proposed method is that the UB cost function can be modified in a simple manner to jointly solve the constellation design and the bit-to-symbol optimal mapping problems.

Without loss of generality, a random constellation $\mathbf{X}_0, \ldots, \mathbf{X}_{K-1}$ has a natural labeling given by the $\log_2(K)$ bits specifying the indices $0, \ldots, K-1$ in binary code. The idea is to fix this natural labeling and then introduce weights in the UB function proportional to the Hamming distance between two given codewords. In this way, two codewords that differ by more bits will be more weighted in the cost function and will therefore end up being farther apart after the optimization process than two codewords with smaller Hamming distance.

It is therefore natural to consider that every pairwise term in the union bound error should be weighed by n_{ij}, the difference in bits of the labels of the pair $\mathbf{X}_i$, $\mathbf{X}_j$. Since these numbers are small, i.e. $1 \leq n_{ij} \leq \log_2(K)$, and each term of the union bound sum can be orders of magnitude larger, we propose using the squared bit-differences as weights:

$$\mathrm{UB}_w(\mathbf{X}_1, \ldots, \mathbf{X}_K) = \sum_{i<j} n_{ij}^2 \cdot \det\left(\mathbf{I}_M - \mathbf{X}_i^H \mathbf{X}_j \mathbf{X}_j^H \mathbf{X}_i\right)^{-N}. \tag{17}$$

Since the bit labeling is fixed, these numbers are constants and the BER optimization algorithm simply uses the gradient of Eq. (15) weighed by these coefficients. The modifications that need to be made to jointly solve the constellation design problem and the labeling problem are therefore minor. We will confirm in the results section that this modification indeed makes a significant impact in the BER performance of constellations optimized using the so-called *bit-weighed union bound criterion* UB_w.

V. Performance Evaluation

We evaluate numerically the performance of our proposed UB-Opt designs in comparison with other Grassmannian constellations and a coherent pilot-based scheme. The Riemannian optimization of the UB performed on the manifold is directly implemented in Matlab without resorting to any existing third-party optimization toolbox.

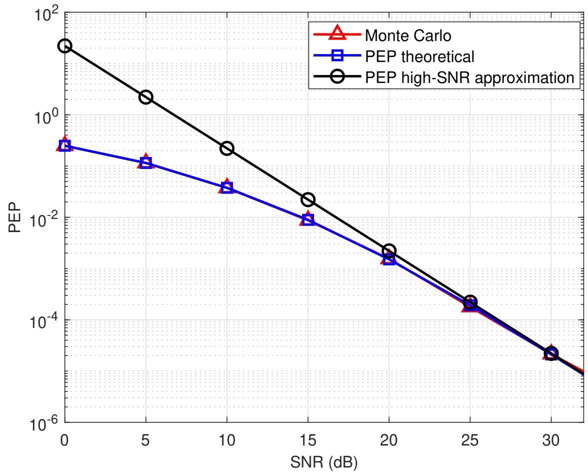
A. SER performance of the UB minimization criterion

In this subsection, we analyze the SER performance of the UB-based constellation design (labeled as UB-Opt in the figures), and we compare it to other unstructured and structured Grassmannian constellations that appear in the literature.

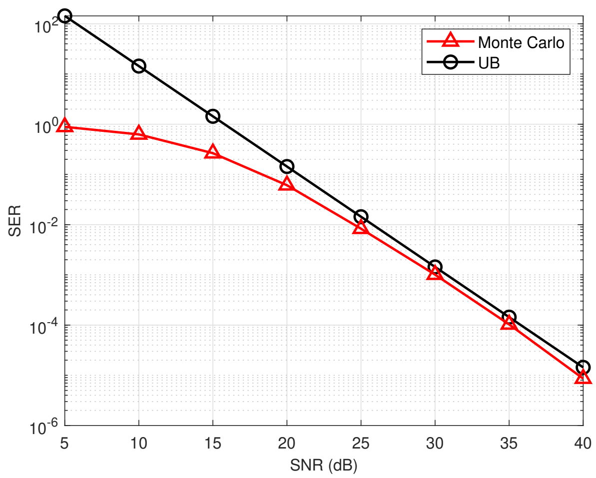
First, we check that the asymptotic PEP UB is an accurate approximation of the SER at high SNR. Fig. 1a shows the analytical expression of the PEP in (5), the high-SNR asymptotic expression in (7), and the simulated PEP estimated via Monte Carlo for two randomly selected codewords from a constellation of 256 codewords for $T = 4$, $M = 2$ and $N = 1$. Fig. 1a shows that the high-SNR asymptote perfectly fits the theoretical and simulated curves (which are almost indistinguishable) for SNRs higher than 25 dB. The high-SNR pairwise error probability expression can be used to obtain the UB (see Fig. 1b). This is an easy way to have an upper bound on the high-SNR behavior of any code or constellation, avoiding the need of time-consuming Monte Carlo simulation.

Secondly, we study the influence of the number of receive antennas N on the SER performance of the UB criterion. Recall that unlike other design criteria proposed in the literature, the number of receive antennas appears explicitly in the UB design criterion (cf. (13)). Therefore, there will be optimal codebooks for each set of values (T, M, N, K). Fig. 2 shows the SER union bound (analytical result) for a MIMO system with $M = 2$ transmit antennas and different number of receive antennas $N = \{3, 4, 5\}$. The coherence time is $T = 4$ and the constellation size is $K = 64$. For each pair of curves in the figure, the solid line represents the union bound obtained by a codebook optimized for the correct value of N, while the dashed line shows the union bound obtained by a codebook optimized for $N = 1$. As we can observe, the SER performance is better when the codebook is optimized for the correct number of receive antennas. Interestingly, this difference in performance increases as N grows, which suggests that using noncoherent constellations specifically designed for a given number of receive antennas might have a significant impact in massive MIMO systems. This is a distinctive feature of the proposed UB criterion.

We also conduct some Monte Carlo simulations to assess the SER performance of the proposed UB-Opt designs in different scenarios. Fig. 3 shows the performance of the UB-Opt designs compared to the method proposed in [29] (which maximizes the minimum chordal distance using a



(a) PEP



(b) SER

Fig. 1: Analytical vs. Simulated PEP and UB for $T = 4$, $M = 2$, $N = 1$ and $K = 256$.

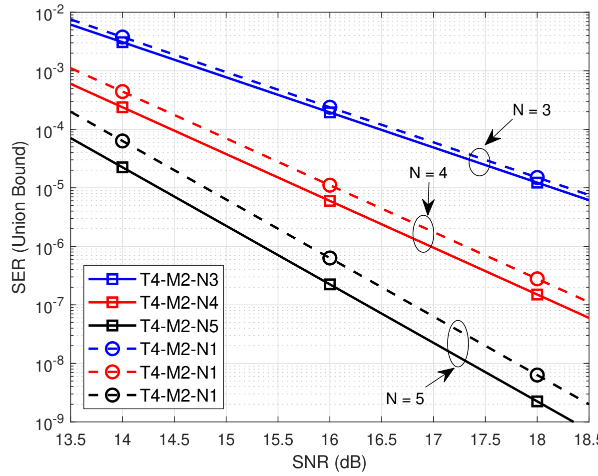


Fig. 2: Union bound for $K = 64$ codewords, $T = 4$, $M = 2$ and $N = \{3, 4, 5\}$ for different codebooks.

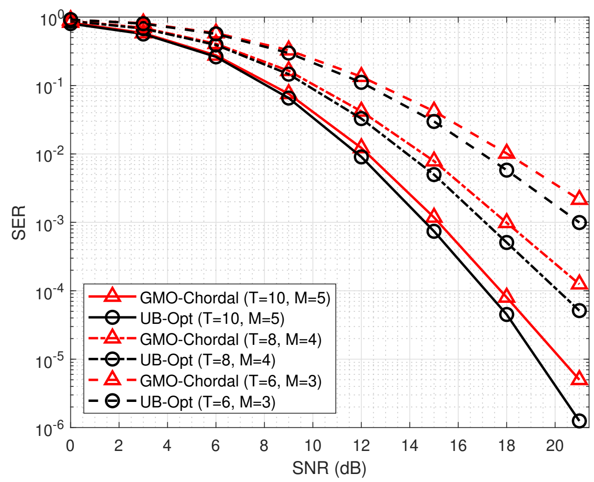


Fig. 3: Union bound vs. chordal distance optimization for $K = 128$, fixed $N = 1$ and varying $(T = 6, M = 3)$, $(T = 8, M = 4)$, $(T = 10, M = 5)$.

gradient ascent method on the Grassmannian, and hence it is labeled as GrassmannOpt-Chordal or GMO-Chordal). In this example, we design constellations with $K = 128$ codewords for varying coherence times and number of transmit antennas: $(T = 6, M = 3)$, $(T = 8, M = 4)$ and $(T = 10, M = 5)$. The number of receive antennas is $N = 1$. The union bound method provides a significant and consistent improvement over the chordal optimization for all (T, M) pairs. This suggests that the minimization of the asymptotic pairwise error probability union bound provides better signal sets than the minimum chordal distance.

Fig. 4 shows the SER performance at 20 dB of SNR for codebooks with $T = 4, M = 2, N = 2$ and varying constellation size $K \in \{16, 32, 64, 128, 256, 512, 1024\}$, designed with the union bound minimization and the minimal chordal maximization. We see that the union bound designs perform consistently better than the chordal designs for any constellation size in the given range.

Fig. 5 shows the SER for the proposed UB-Opt constellation compared to the GMO-Chordal [29] and the alternating projection (AP) method [14] for $T = 4$, $M = 2$, $N = \{1, 2\}$, and $K = 64$ codewords. We can observe that the UB constellations clearly outperform the other two methods.

B. Joint bit-labeling + constellation optimization

In this subsection we assess the BER performance of the joint binary labeling and constellation design algorithm via Monte Carlo simulations. In the first experiment we consider a scenario with $M = 2$ transmit antennas, $N = 1$ receive antenna and $T = 4$ coherence time. We compare the results of a random labeling of the codewords for a constellation designed with the chordal distance and the UB criteria, with the joint optimization of the binary labeling and the UB constellation design. The results are depicted in Figs. 6a and 6b for $K = 64$ and $K = 128$ codewords, respectively. Even with a random labeling, one can already notice a very significant BER improvement by using the union bound criterion with

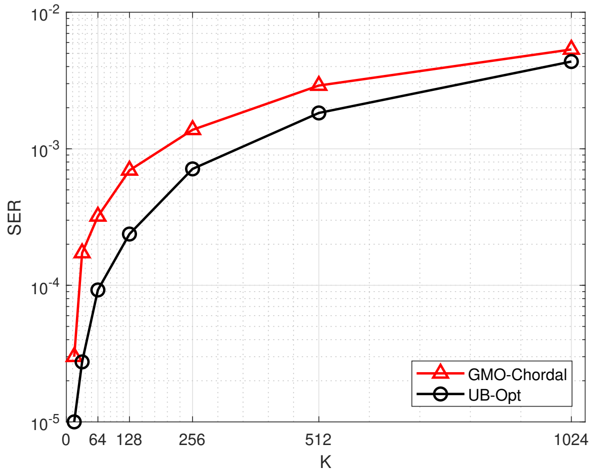


Fig. 4: SER vs. varying constellation size K of the union bound and chordal distance optimizations for $T = 4, M = 2, N = 2$ and SNR = 20 dB.

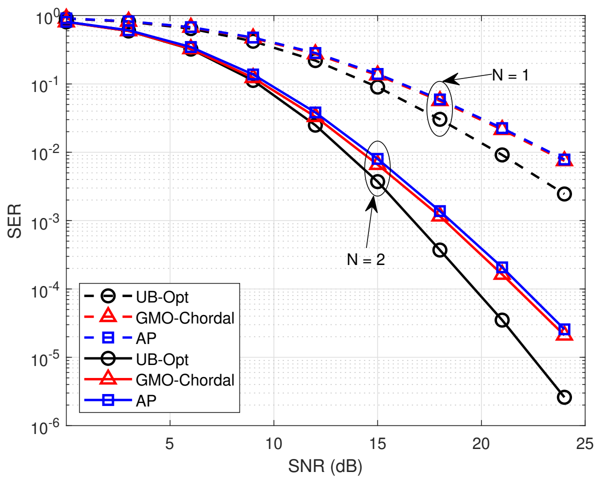
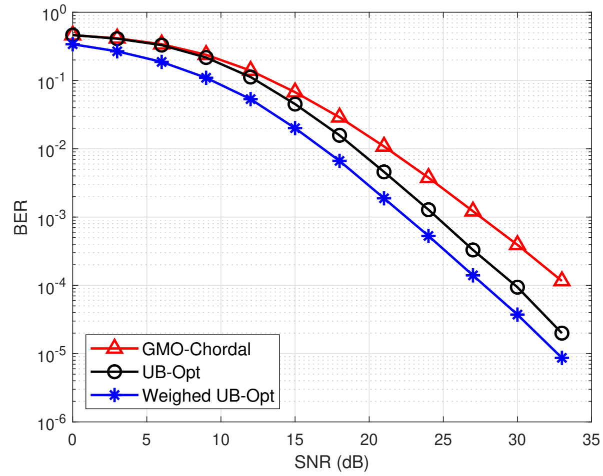
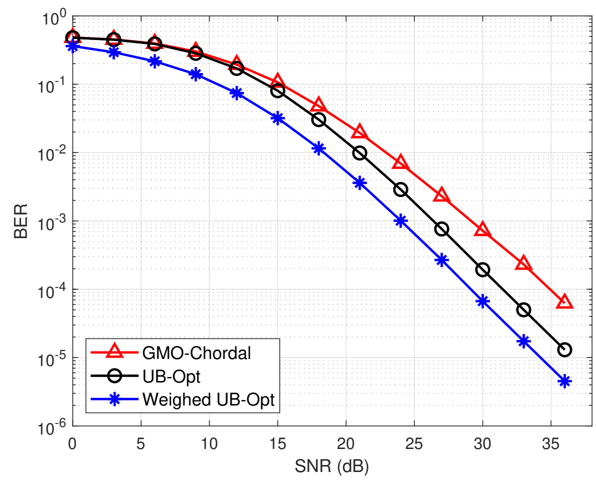


Fig. 5: SER curves for $K = 64$ codewords, $T = 4$, $M = 2$ and $N = \{1, 2\}$.



(a) $K = 64$



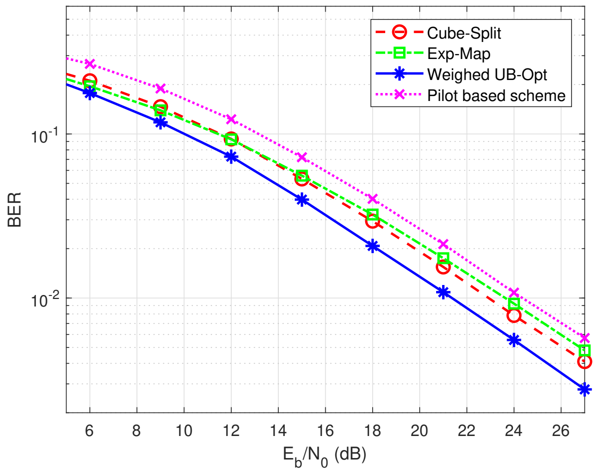
(b) $K = 128$

Fig. 6: BER performance of a random labeling for the chordal and union bound optimizations vs. the joint binary labeling and constellation design, for $T = 4$, $M = 2$, $N = 1$ and $K \in \{64, 128\}$.

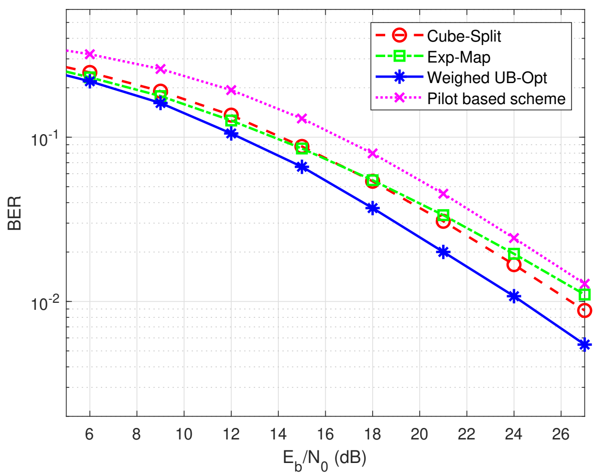
respect to the chordal distance. But more importantly, the joint optimization of the binary labeling altogether with the constellation further improves the BER performance in a very significant amount. It is worth noting the binary labeling optimization essentially comes at no additional cost in comparison to the conventional UB optimization, since we only have to weight differently the UB terms.

In the examples of Figs. 7a and 7b we consider a scenario with $T = 2$, $M = N = 1$ and compare the weighed UB-Opt design with some structured designs that allow Gray-labeling. In particular, we consider the Cube-Split [13] and the Exp-Map [12]. We include as a baseline the performance of a coherent pilot-based scheme. The transmitted signal for the pilot-based scheme is $\mathbf{x}_{coh} = [1, x_d]^T/\sqrt{2}$, where the first symbol is the constant pilot, which is known at the receiver, and the second symbol x_d is taken from a QAM constellation with cardinality 2^B. That is, when $B = 4$ we use a 16-QAM constellation and the spectral efficiency is $R = 2$ b/s/Hz and when $B = 6$ we use a 64-QAM constellation and the spectral efficiency is $R = 3$ b/s/Hz (the Cube-Split and the Exp-Map have similar spectral efficiencies). The QAM constellations are normalized such that $E[|x_d|^2] = 1$. Therefore, $E[\mathbf{x}_{coh}^H \mathbf{x}_{coh}] = 1$ and hence the average transmit power of the pilot-based scheme is the same as that of the noncoherent schemes. Notice also that the power devoted to the data transmission is the same as the power devoted to training. This is the optimal power allocation for $T = 2, M = 1$ as shown in [30].

These experiments support our proposal that a simple modification of the union bound cost function, so that it takes into account the binary code labels of the codewords, yields improved unstructured and even structured Grassmannian constellation in terms of their BER performance. This is a very encouraging result to further explore for a simultaneous joint solution of the binary labeling along with the constellation



(a) $B = 4$ ($R = 2$ b/s/Hz)



(b) $B = 6$ ($R = 3$ b/s/Hz)

Fig. 7: BER vs. E_b/N_0 for the Weighed UB-Opt constellation in comparison with Cube-Split, Exp-Map, and a pilot based scheme for $T = 2$, $M = 1$, $N = 1$ and $B \in \{4, 6\}$ ($R \in \{2, 3\}$ b/s/Hz).

design.

C. Achievable rate

In this subsection, we study the achievable rate obtained by the UB designs. For comparison, we use the asymptotic (high-SNR) approximations of the channel capacity derived in [8].

The achievable rate of our system (in bits/s/Hz) is given by the input-output mutual information

$$R = I(\mathbf{X}; \mathbf{Y}) = h(\mathbf{Y}) - h(\mathbf{Y}|\mathbf{X}) = \mathrm{E}\left[\log_2 \frac{p(\mathbf{Y}|\mathbf{X})}{p(\mathbf{Y})}\right], \tag{18}$$

where $h(\mathbf{A})$ is the differential entropy of the random matrix $\mathbf{A}$ and $p(\mathbf{A})$ is the probability distribution of $\mathbf{A}$. From the law of total probability and assuming that all constellation symbols are equally likely to be transmitted, we can write

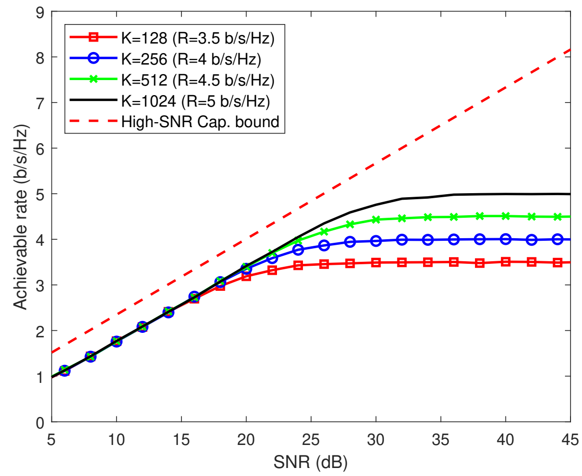


Fig. 8: Achievable rate of UB constellation for $T = 2$, $M = 1$, $N = 1$ and different constellation size K in comparison with the high-SNR approximation of the channel capacity given in (21).

$$p(\mathbf{Y}) = \sum_{k=1}^{K} p(\mathbf{X}_k)\, p(\mathbf{Y}|\mathbf{X}_k) = \frac{1}{K} \sum_{k=1}^{K} p(\mathbf{Y}|\mathbf{X}_k). \tag{19}$$

Conditioned on $\mathbf{X} = \mathbf{X}_k$, $\mathbf{Y}$ is a complex Gaussian matrix with independent columns having the same covariance matrix $\mathbf{R}_{\mathbf{Y}\mathbf{X}_k} = \mathbf{X}_k \mathbf{X}_k^H + \sigma^2 \mathbf{I}_T$, hence

$$p(\mathbf{Y}|\mathbf{X}_k) = \frac{\exp\left(-\operatorname{tr}\{\mathbf{Y}^H \mathbf{R}_{\mathbf{Y}\mathbf{X}_k} \mathbf{Y}\}\right)}{\left(\pi^T \det \mathbf{R}_{\mathbf{Y}\mathbf{X}_k}\right)^N} \tag{20}$$

The expectation in (18) does not have a closed form, so we resort to Monte Carlo simulations to estimate R.

Fig. 8 shows the achievable rate for two different UB constellations ($K = \{16, 64\}$) for $T = 4$ and $M = N = 1$. For comparison, the figure also depicts the asymptotic high-SNR noncoherent capacity given in [8]

$$C_{SNR\to\infty} = M\left(1 - \frac{M}{T}\right)\log_2 SNR + c + o(1), \tag{21}$$

where c is a constant that is independent of the SNR and its value is given in [8, Th.9].

VI. Conclusion

In this paper, we have presented an approach for designing unstructured Grassmannian constellations for noncoherent MIMO communications based on the the asymptotic pairwise error probability union bound. The optimization of the UB cost function is based on a gradient descent approach with adaptive step-size that operates directly on a Cartesian product of Grassmann manifolds. For block-fading channels and SNRs encountered in practice, our results suggest that minimizing the UB substantially improves the SER in comparison to most alternatives in the literature, which usually maximize a distance measure (usually the chordal distance) for the two closest codewords. The codebooks designed according to the proposed criterion are guaranteed to have full-diversity, also in contrast to what happens with the maximization of the minimum chordal distance. Furthermore, the proposed algorithm also allows to jointly design the constellation and solve the bit labeling problem. Numerical results show that the designed constellations outperform other unitary space-time constellations in the literature, in terms of SER and BER.

Using the proposed UB criterion constellations of cardinality $K = 2048$ codewords have been readily designed. As a future line of work, we will investigate the design of Grassmannian constellations of higher cardinality, probably with some structured design to allow for more efficient detection schemes. Another interesting line of work is to explore potential applications of noncoherent communications with other technologies of recent interest such as Reconfigurable Intelligent Surfaces (RIS). Similar to [31], [32], noncoherent schemes could be used during the beam training process in RIS-assisted networks. The best phase configuration for the RIS could be selected by measuring the average energy received in the subspaces defined by the noncoherent codebook.

Appendix

A. Preliminaries

The material in this section is standard and can be found for example in [33] and [34] (although the main focus of these classic works is in the real case, the formulas carry over directly to the complex Grassmannian). We recall some basic notions. The complex Grassmannian $\mathbb{G}(M, \mathbb{C}^T)$ is the set of M–dimensional complex subspaces of $\mathbb{C}^T$, with $T > M$, that is a complex manifold of dimension $M(T-M)$. Elements in $\mathbb{G}(M, \mathbb{C}^T)$ are represented by matrices in the Stiefel manifold $\mathbf{A} \in \mathbb{S}_t(M, \mathbb{C}^T)$, that is $\mathbf{A} \in \mathbb{C}^{T\times M}$, $\mathbf{A}^H\mathbf{A} = \mathbf{I}_M$. This representation is not unique, since $\mathbf{A}$ and $\mathbf{AU}$ with $\mathbf{U}$ a unitary $M \times M$ matrix represent the same element in $\mathbb{G}(M, \mathbb{C}^T)$, so formally we should denote elements of the Grassmannian as $[\mathbf{A}]$ where $\mathbf{A} \in \mathbb{S}_t(M, \mathbb{C}^T)$ is a unitary basis for that subspace, $\mathbf{P}_{\mathbf{A}} = \mathbf{AA}^H$ denotes the orthogonal projection onto $[\mathbf{A}]$ and $[\mathbf{A}]$ is the class of $\mathbf{A}$ under the quotient by the set of $M \times M$ unitary matrices $\mathcal{U}_M$.

The tangent space to the Grassmann manifold at $[\mathbf{A}]$ is

$$T_{[\mathbf{A}]}\mathbb{G}(M, \mathbb{C}^T) \equiv= \{(\mathbf{I}_T - \mathbf{AA}^H)\mathbf{B} : \mathbf{B} \in \mathbb{C}^{T\times M}\}.$$

The following lemma is fundamental for the computation of gradients in the Grassmannian that will be used in the UB optimization algorithm.

Lemma 4 *Let $\varphi : \mathbb{C}^{T\times M} \to \mathbb{R}$ be a C^1 mapping, defined at least in some open neighborhood of the Stiefel manifold $\mathbb{S}_t(M, \mathbb{C}^T) \subseteq \mathbb{C}^{T\times M}$, and assume that φ can be defined as a function on $\mathbb{G}(M, \mathbb{C}^T)$, that is, we have:*

$$\varphi(\mathbf{A}) = \varphi(\mathbf{AU}) \quad \text{for } \mathbf{A} \in \mathbb{S}_t(M, \mathbb{C}^T), \mathbf{U} \in \mathcal{U}_M.$$

Then, the gradient of φ at $\mathbf{A} \in \mathbb{S}_t(M, \mathbb{C}^T)$ as a Grassmannian mapping is:

$$\nabla\varphi(\mathbf{A}) = (\mathbf{I}_T - \mathbf{AA}^H)\nabla^{unc}\varphi(\mathbf{A}),$$

where $\nabla^{unc}\varphi$ is the unconstrained gradient of φ as a function on the ambient space $\mathbb{C}^{T\times M}$.

Proof: A proof of this result can be found in [34]. □

B. Proof of Theorem 1

We will depart from (3) and write down this sum of residuals as a complex integral. First, since $a_m > 0$ for all m, we note that all the residuals of the form $w = ja_n$ lie in the upper complex half–plane. Hence, for large enough $R > 0$ the expresion in (3) equals the contour integral

$$P_e(\mathbf{X}_i, \mathbf{X}_j) = -\frac{1}{2\pi j}\int_{w\in\mathcal{C}_R^+}\frac{1}{w+j/2} \cdot \prod_{m=1}^{M}\left(\frac{1+\rho T/M}{(\rho T/M)^2(1-d_m^2)(w^2+1/4)+1+\rho T/M}\right)^N dw, \quad (22)$$

where $\mathcal{C}_R^+$ is the oriented curve given by the segment $[-R, R]$ followed by the half circunference $Re^{j\theta}$, $\theta \in [0, \pi]$. Since the function inside the integral in (22) is a rational function whose denominator is a polynomial of degree at least 3 and whose numerator is a constant, we can bound its modulus above by $\frac{A}{|z|^3}$ for some constant $A \in \mathbb{R}$. For large R, the integral along the semicircumference $z = Re^{j\theta}$, $0 \leq \theta \leq \pi$, is bounded above by

$$\int_0^\pi \frac{RA}{R^3}\,d\theta = \frac{\pi A}{R^2} \underset{R\to\infty}{\to} 0.$$

Hence, taking limit $R \to \infty$ in (23) we conclude that

$$P_e(\mathbf{X}_i, \mathbf{X}_j) = -\frac{1}{2\pi j}\int_{-\infty}^{\infty}\frac{1}{w+j/2} \cdot \prod_{m=1}^{M}\left(\frac{1+\rho T/M}{(\rho T/M)^2(1-d_m^2)(w^2+1/4)+1+\rho T/M}\right)^N dw. \quad (23)$$

We further simplify algebraically by noting that

$$\frac{1}{w+j/2} = \frac{w}{w^2+1/4} - \frac{j/2}{w^2+1/4},$$

which gives two integrals. The first of them equals 0 because we integrate an odd function, and up to a factor of 2 the second one can be considered only in $[0, \infty)$ because it is an even function. We thus conclude that

$$\begin{aligned} P_e(\mathbf{X}_i, \mathbf{X}_j) &= \frac{1}{2\pi}\int_0^\infty \frac{1}{w^2+1/4} \cdot \prod_{m=1}^{M}\left(\frac{1}{1+\frac{(\rho T/M)^2(1-d_m^2)(w^2+1/4)}{1+\rho T/M}}\right)^N dw \\ &\overset{\theta=\arctan(2w)}{=} \frac{1}{\pi}\int_0^{\pi/2}\prod_{m=1}^{M}\left(\frac{1}{1+\frac{(\rho T/M)^2(1-d_m^2)}{4(1+\rho T/M)\cos^2\theta}}\right)^N d\theta. \end{aligned}$$

To some extent, this proof has been obtained by reversing the ideas in the proof of (3) appearing in [6].

C. Proof of Corollary 1 and Corollary 2

We will use Lebesgue's Dominated Convergence Theorem, a classical result which gives sufficient conditions to guarantee the interchange of integral and limit, see for example [35, Th. 16.5]. It is easier to prove first Corollary 2.

1) Proof of Corollary 2: From Theorem 1, we have

$$\lim_{\rho\to\infty}[\rho^{NM}P_e(\mathbf{X}_i,\mathbf{X}_j)] = \lim_{\rho\to\infty}\frac{1}{\pi}\int_0^{\pi/2}\prod_{m=1}^{M}\left(\frac{\rho}{1+\frac{(\rho T/M)^2(1-d_m^2)}{4(1+\rho T/M)\cos^2\theta}}\right)^N d\theta. \quad (24)$$

where we have assumed that all M singular values d_m of $\mathbf{X}_i^H\mathbf{X}_j$ are distinct from 1. That is, we have assumed a full diversity unitary space-time code.

The integrand is bounded above by some constant independent of ρ and θ. This constant plays the role of the dominating function $g(x)$ in Lebesgue's Dominated Convergence Theorem and hence we can interchange limit and integral which yields:

$$\begin{aligned}&\lim_{\rho\to\infty}[\rho^{NM}P_e(\mathbf{X}_i,\mathbf{X}_j)]\\ &=\frac{1}{\pi}\int_0^{\pi/2}\lim_{\rho\to\infty}\prod_{m=1}^{M}\left(\frac{\rho}{1+\frac{(\rho T/M)^2(1-d_m^2)}{4(1+\rho T/M)\cos^2\theta}}\right)^N d\theta\\ &=\frac{1}{\pi}\int_0^{\pi/2}\prod_{m=1}^{M}\left(\frac{4(T/M)\cos^2\theta}{(T/M)^2(1-d_m^2)}\right)^N d\theta\\ &=\left(\frac{4M}{T}\right)^{NM}\frac{1}{\prod_{m=1}^{M}(1-d_m^2)^N}\frac{1}{\pi}\int_0^{\pi/2}\cos^{2NM}\theta\,d\theta.\end{aligned}$$

This last integral has a known value $\pi/2\cdot(2NM-1)!!/(2NM)!!$, which yields the result.

2) Proof of Corollary 1: The limit to be computed is

$$\lim_{\rho\to\infty}\frac{d\log P_e(\mathbf{X}_i,\mathbf{X}_j)}{d\log\rho}=\lim_{\rho\to\infty}\frac{\rho}{P_e(\mathbf{X}_i,\mathbf{X}_j)}\cdot\frac{d}{d\rho}\left(\frac{1}{\pi}\int_0^{\pi/2}\prod_{m=M-D+1}^{M}\left(\frac{1}{1+\frac{(\rho T/M)^2(1-d_m^2)}{4(1+\rho T/M)\cos^2\theta}}\right)^N d\theta\right).$$

(note that we only need to consider the product for m from $M-D+1$ to M since the inner parenthesis equals 1 if $d_m=1$). There exists some hypotheses under which the interchange of derivative and integral signs are allowed: these are sometimes called Leibniz's rule for differentiation under the integral sign, see for example [35, Th. 16.11]. In our case, since we are integrating a smooth function such that both it and its first derivative are bounded above by a constant in a compact interval, we can make this interchange getting

$$\begin{aligned}&\frac{d}{d\rho}\left(\frac{1}{\pi}\int_0^{\pi/2}\prod_{m=M-D+1}^{M}\left(\frac{1}{1+\frac{(\rho T/M)^2(1-d_m^2)}{4(1+\rho T/M)\cos^2\theta}}\right)^N d\theta\right)\\ &=-\frac{N}{\pi}\sum_{m=M-D+1}^{M}\int_0^{\pi/2}\prod_{m=M-D+1}^{M}\\ &\left(\frac{1}{1+\frac{(\rho T/M)^2(1-d_m^2)}{4(1+\rho T/M)\cos^2\theta}}\right)^N\frac{\frac{d}{d\rho}\left(1+\frac{(\rho T/M)^2(1-d_i^2)}{4(1+\rho T/M)\cos^2\theta}\right)}{1+\frac{(\rho T/M)^2(1-d_i^2)}{4(1+\rho T/M)\cos^2\theta}}\,d\theta.\end{aligned}$$

Computing the derivative and multiplying the numerator and denominator by ρ^{ND} we then get

$$\begin{aligned}&\lim_{\rho\to\infty}\frac{d\log P_e(\mathbf{X}_i,\mathbf{X}_j)}{d\log\rho}=-N\sum_{m=M-D+1}^{M}\lim_{\rho\to\infty}\\ &\frac{\int_0^{\pi/2}\prod_{m=M-D+1}^{M}\left(\frac{\rho}{1+\frac{(\rho T/M)^2(1-d_i^2)}{4(1+\rho T/M)\cos^2\theta}}\right)^N Q_i(\rho,\theta)\,d\theta}{\int_0^{\pi/2}\prod_{m=M-D+1}^{M}\left(\frac{\rho}{1+\frac{(\rho T/M)^2(1-d_m^2)}{4(1+\rho T/M)\cos^2\theta}}\right)^N d\theta},\end{aligned}$$

where

$$Q_i(\rho,\theta)=\frac{4(T/M)\cos^2\theta+2\rho^2T^2/M^2(1-d_i)^2}{4(1+\rho T/M)\cos^2\theta+(\rho T/M)^2(1-d_i^2)}-\frac{\rho T/M}{1+\rho T/M}.$$

Exactly as in the proof of Corollary 2, we can apply Lebesgue's Dominated Convergence Theorem again to interchange limit and integral both in the numerator and the denominator, since all the functions are bounded above by a constant independent of θ and ρ. Now, we have $\lim_{\rho\to\infty}Q_i(\rho,\theta)=1$ for all i, $1\le i\le D$, which implies

$$\lim_{\rho\to\infty}\frac{d\log P_e(\mathbf{X}_i,\mathbf{X}_j)}{d\log\rho}=-N\sum_{i=1}^{D}1=-ND,$$

as claimed.

D. Proof of Lemma 3

As shown in Lemma 4, the real function UB is defined in a neighborhood of the product of Stiefel manifolds and is well defined as a mapping in the product of Grassmmanians. Then, it suffices to consider the gradient of UB defined as a mapping in $\mathbb{C}^{T\times M}\times\cdots\times\mathbb{C}^{T\times M}$, which is a vector in the Cartesian product of the ambient tangent spaces. Denote this gradient by $(\dot{\mathbf{Y}}_1,\ldots,\dot{\mathbf{Y}}_K)\in\mathbb{C}^{T\times M}\times\cdots\times\mathbb{C}^{T\times M}$, whose defining property is given by its relation with the directional derivative of φ:

$$\begin{aligned}D\varphi(\mathbf{A})(\dot{\mathbf{A}})&=\Re(\langle(\dot{\mathbf{Y}}_1,\ldots,\dot{\mathbf{Y}}_K),\dot{\mathbf{A}}\rangle_F),\\ &\forall\dot{\mathbf{A}}=(\dot{\mathbf{A}}_1,\ldots,\dot{\mathbf{A}}_K)\in\mathbb{C}^{T\times M}\times\cdots\times\mathbb{C}^{T\times M}.\end{aligned} \quad (25)$$

From Section A, in order to get the gradient of UB in $\mathbb{G}(M, \mathbb{C}^T)$ we just need to project each $\dot{\mathbf{Y}}_i$ onto the tangent space of the Grassmann manifold $T_{[\mathbf{X}_i]}\mathbb{G}(M, \mathbb{C}^T)$ by $\dot{\mathbf{X}}_i = (\mathbf{I}_T - \mathbf{X}_i\mathbf{X}_i^H)\dot{\mathbf{Y}}_i$, so that $\nabla\mathrm{UB}([\mathbf{X}_1], \ldots, [\mathbf{X}_K]) = (\dot{\mathbf{X}}_1, \ldots, \dot{\mathbf{X}}_K) \in T_{[\mathbf{X}_1]}\mathbb{G}(M, \mathbb{C}^T) \times \cdots \times T_{[\mathbf{X}_K]}\mathbb{G}(M, \mathbb{C}^T)$.

Now we obtain each $\dot{\mathbf{Y}}_i$ by vector calculus standard routine using (25): begin by computing the derivative using Jacobi's formula for the derivative of the determinant,

$$\begin{aligned} & D\mathrm{UB}(\mathbf{X}_1, \ldots, \mathbf{X}_K)(\dot{\mathbf{Z}}_1, 0, \ldots, 0) \\ & = \frac{d}{dt}\bigg|_{t=0} \mathrm{UB}(\mathbf{X}_1 + t\dot{\mathbf{Z}}_1, \mathbf{X}_2, \ldots, \mathbf{X}_K) \\ & = \sum_{j>1} N \det(\mathbf{I}_M - \mathbf{X}_j^H\mathbf{X}_1\mathbf{X}_1^H\mathbf{X}_j)^{-N} \\ & \qquad \cdot \operatorname{tr}\left((\mathbf{I}_M - \mathbf{X}_j^H\mathbf{X}_1\mathbf{X}_1^H\mathbf{X}_j)^{-1}\mathcal{L}_{1j}\right), \end{aligned}$$

where $\mathcal{L}_{1j} = \mathbf{X}_j^H\dot{\mathbf{Z}}_1\mathbf{X}_1^H\mathbf{X}_j + \mathbf{X}_j^H\mathbf{X}_1\dot{\mathbf{Z}}_1^H\mathbf{X}_j$. The inner trace has 2 summands a, b, corresponding to these two parts of $\mathcal{L}_{1j}$ (we use the circular property of the trace operator $\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA})$):

$$\begin{aligned} a &= \operatorname{tr}\left(\mathbf{X}_1^H\mathbf{X}_j(\mathbf{I}_M - \mathbf{X}_j^H\mathbf{X}_1\mathbf{X}_1^H\mathbf{X}_j)^{-1}\mathbf{X}_j^H\dot{\mathbf{Z}}_1\right), \\ b &= \operatorname{tr}\left(\dot{\mathbf{Z}}_1^H\mathbf{X}_j(\mathbf{I}_M - \mathbf{X}_j^H\mathbf{X}_1\mathbf{X}_1^H\mathbf{X}_j)^{-1}\mathbf{X}_j^H\mathbf{X}_1\right). \end{aligned}$$

Using now $\operatorname{tr}(\mathbf{B}^H\mathbf{A}^H) = \operatorname{tr}((\mathbf{AB})^H) = (\operatorname{tr}(\mathbf{AB}))^*$, where $()^*$ denotes complex conjugate,

$$b^* = \operatorname{tr}\left(\dot{\mathbf{Z}}_1^H\mathbf{X}_j(\mathbf{I}_M - \mathbf{X}_j^H\mathbf{X}_1\mathbf{X}_1^H\mathbf{X}_j)^{-1}\mathbf{X}_j^H\mathbf{X}_1\right)^* = a,$$

we thus have

$$\begin{aligned} a + b &= 2\Re\left(\operatorname{tr}\left(\dot{\mathbf{Z}}_1^H\mathbf{X}_j(\mathbf{I}_M - \mathbf{X}_j^H\mathbf{X}_1\mathbf{X}_1^H\mathbf{X}_j)^{-1}\mathbf{X}_j^H\mathbf{X}_1\right)\right) \\ &= \Re\left\langle 2\mathbf{X}_j(\mathbf{I}_M - \mathbf{X}_j^H\mathbf{X}_1\mathbf{X}_1^H\mathbf{X}_j)^{-1}\mathbf{X}_j^H\mathbf{X}_1, \dot{\mathbf{Z}}_1\right\rangle_F. \end{aligned}$$

All together, we have proved that

$$\begin{aligned} & D\mathrm{UB}(\mathbf{X}_1, \ldots, \mathbf{X}_K)(\dot{\mathbf{Z}}_1, 0, \ldots, 0) \\ & \quad = \Re\left\langle \sum_{j>1} 2N \det(\mathbf{I}_M - \mathbf{X}_j^H\mathbf{X}_1\mathbf{X}_1^H\mathbf{X}_j)^{-N} \right. \\ & \qquad \left. \cdot \mathbf{X}_j(\mathbf{I}_M - \mathbf{X}_j^H\mathbf{X}_1\mathbf{X}_1^H\mathbf{X}_j)^{-1}\mathbf{X}_j^H\mathbf{X}_1, \dot{\mathbf{Z}}_1 \right\rangle_F, \end{aligned}$$

which by Eq. (25) yields

$$\begin{aligned} \dot{\mathbf{Y}}_1 = \sum_{j>1} & 2N \det(\mathbf{I}_M - \mathbf{X}_j^H\mathbf{X}_1\mathbf{X}_1^H\mathbf{X}_j)^{-N} \\ & \cdot \mathbf{X}_j(\mathbf{I}_M - \mathbf{X}_j^H\mathbf{X}_1\mathbf{X}_1^H\mathbf{X}_j)^{-1}\mathbf{X}_j^H\mathbf{X}_1. \end{aligned}$$

There is nothing special in having computed $\dot{\mathbf{Y}}_1$ or $\dot{\mathbf{Y}}_i$. We just change 1 to i and $j > 1$ to $j \neq i$ to get the general expression for $\dot{\mathbf{Y}}_i$, and then compute $\dot{\mathbf{X}}_i = (\mathbf{I}_T - \mathbf{X}_i\mathbf{X}_i^H)\dot{\mathbf{Y}}_i$. This is what we claim in the lemma.

References

[1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585–595, Dec. 1999.

[2] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, Mar. 1998.

[3] L. Jing, E. Carvalho, P. Popovski, and A. Oliveras Martinez, "Design and performance analysis of noncoherent detection systems with massive receiver arrays," *IEEE Trans. Signal Processing*, vol. 64, pp. 5000–5010, Oct. 2016.

[4] K. T. Truong and R. W. Heath, "Effects of channel aging in massive MIMO systems," *Journal of Communications and Networks*, vol. 15, pp. 338–351, Aug. 2013.

[5] T. Marzetta and B. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. on Inf. Theory*, vol. 45, pp. 139–157, Jan. 1999.

[6] B. Hochwald and T. Marzetta, "Unitary space-time modulation for multiple-antenna communication in Rayleigh flat-fading," *IEEE Trans. on Inf. Theory*, vol. 46, pp. 1962–1973, Mar. 2000.

[7] M. Brehler and M. Varanasi, "Asymptotic error probability analysis of quadratic receivers in Rayleigh fading channels with applications to a unified analysis of coherent and noncoherent space-time receivers," *IEEE Trans. Inf. Theory*, vol. 47, pp. 2383–2399, Sept. 2001.

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

[9] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian spaces," *Experimental Mathematics*, vol. 5, pp. 139–159, Apr. 1996.

[10] R. H. Gohary and T. N. Davidson, "Noncoherent MIMO communication: Grassmannian constellations and efficient detection," *IEEE Trans. on Inf. Theory*, vol. 55, pp. 1176–1205, Mar. 2009.

[11] M. Beko, J. Xavier, and V. A. N. Barroso, "Noncoherent communications in multiple-antenna systems: receiver design and codebook construction," *IEEE Trans. Signal Processing*, vol. 55, pp. 5703–5715, Dec. 2007.

[12] I. Kammoun, A. M. Cipriano, and J. Belfiore, "Non-coherent codes over the Grassmannian," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 3657–3667, Oct. 2007.

[13] K. Ngo, A. Decurninge, M. Guillaud, and S. Yang, "Cube-split: A structured Grassmannian constellation for non-coherent SIMO communications," *IEEE Trans. on Wireless Commun.*, vol. 19, pp. 1948–1964, Mar. 2020.

[14] I. Dhillon, R. Heath, T. Strohmer, and J. Tropp, "Constructing packings in Grassmannian manifolds via alternating projection," *Experimental Mathematics*, vol. 17, Oct. 2007.

[15] W. Zhao, G. Leus, and G. B. Giannakis, "Orthogonal design of unitary constellations for uncoded and trellis-coded noncoherent space-time systems," *IEEE Trans. on Inf. Theory*, vol. 50, pp. 1319–1327, June 2004.

[16] B. Hochwald, T. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design of unitary space-time constellations," *IEEE Trans. on Inf. Theory*, vol. 48, pp. 1962–1973, Sept. 2000.

[17] B. Hughes, "Differential space-time modulation," *IEEE Trans. on Inf. Theory*, vol. 46, pp. 2567–2578, Nov. 2000.

[18] R. Pitaval and O. Tirkkonen, "Grassmannian packings from orbits of projective group representations," in *46th Asilomar Conference on Signals, Systems and Computers (Asilomar 2012)*, (Pacific Grove, CA, USA), pp. 478–482, Nov. 2012.

[19] G. Han and J. Rosenthal, "Geometrical and numerical design of structured unitary space-time constellations," *IEEE Trans. Inf. Theory*, vol. 52, pp. 3722–3735, Aug. 2006.

[20] 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*, vol. 48, pp. 1186–1194, May 2002.

[21] J. Alvarez-Vizoso, D. Cuevas, C. Beltrán, I. Santamaria, V. Tucek, and G. Peters, "Coherence-based subspace packings for MIMO noncoherent communications," in *30th European Signal Processing Conference (EUSIPCO 2022)*, (Belgrade, Serbia), Aug. 2022.

[22] A. Alvarado, F. Brännström, and T. Koch, "High-SNR asymptotics of mutual information for discrete constellations with applications to BICM," *IEEE Trans. Inf. Theory*, vol. 60, pp. 1061–1076, Feb. 2014.

[23] N. H. Tran, H. H. Nguyen, and T. Le-Ngoc, "Coded unitary space-time modulation with iterative decoding: Error performance and mapping design," *IEEE Trans. Commun.*, vol. 55, pp. 956–965, Apr. 2007.

[24] A. Ashikhmin and A. R. Calderbank, “Grassmannian packings from operator Reed-Muller codes,” *IEEE Trans. Inf. Theory*, vol. 56, pp. 5689–5714, Nov. 2010.

[25] Y. Qin and R. Pitaval, “Structured quasi-Gray labeling for Reed-Muller Grassmannian constellations,” in *IEEE Int. Symp. on Information Theory (ISIT)*, (Los Angeles, CA, USA), June 2020.

[26] Y. Li and X. G. Xia, “Constellation mapping for space-time matrix modulation with iterative demodulation/decoding,” *IEEE Trans. Commun.*, vol. 53, pp. 764–768, May 2005.

[27] S. Baluja and M. Covel, “Neighborhood preserving codes for assigning point labels: Application to stochastic search,” in *Proc. Int. Conf. Comput. Sci.,*, pp. 956–965, June 2013.

[28] G. W. K. Colman, R. H. Gohary, T. J. El-Azizy, T. J. Willink, and T. N. Davidson, “Quasi-gray labeling for Grassmannian constellations,” *IEEE Trans. Wireless Commun.*, vol. 10, pp. 626–636, Feb. 2011.

[29] D. Cuevas, C. Beltrán, I. Santamaria, V. Tucek, and G. Peters, “A fast algorithm for designing Grassmannian constellations,” in *25th International ITG Workshop on Smart Antennas (WSA 2021)*, (EURECOM, France), Nov. 2021.

[30] B. Hassibi and B. M. Hochwald, “How much training is needed in multiple-antenna wireless links?,” *IEEE Transactions on Information Theory*, vol. 49, pp. 951–963, Apr. 2003.

[31] K. Chen-Hu, G. C. Alexanfropoulos, and A. G. Armada, “Non-coherent MIMO-OFDM uplink empowered by the spatial diversity in reflecting surfaces,” Feb. 2022.

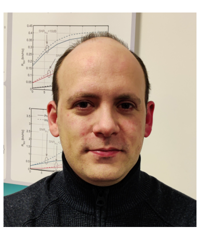
[32] K. Chen-Hu, G. C. Alexandropoulos, and A. G. Armada, “Differential data-aided beam training for RIS-empowered multi-antenna communications,” *IEEE Access*, vol. 10, pp. 113200–113213, Oct. 2022.

[33] A. Edelman, T. A. Arias, and S. T. Smith, “The geometry of algorithms with orthogonality constraints,” *SIAM J. Matrix Anal. Appl.*, vol. 20, pp. 303–353, Oct. 1998.

[34] P. A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Dec. 2007.

[35] J. Jost, *Postmodern analysis*. Universitext, Springer-Verlag, Berlin, Third ed., Aug. 2005.

Diego Cuevas received the B.E. degree in telecommunications engineering from the Universidad de Cantabria (UC), Spain, in 2018, and the M.Sc. degree in telecommunications engineering from the Universidad de Cantabria, Spain, in 2020. In October 2020, he joined the Department of Communications Engineering, Universidad de Cantabria, where he is currently working toward the Ph.D. degree in electrical engineering. His current research interests include noncoherent wireless communications, signal processing algorithms for multiple-antenna wireless communications and machine learning.



Javier Álvarez-Vizoso received his Licenciatura degree in Physics at Universidad de Oviedo, Spain, in 2013, and graduated with a Ph.D. in mathematics at Colorado State University, USA, in 2018. He has held a postdoctoral position in applied mathematics at the Max Planck Institute of Solar System Research, Germany. In October 2021 he joined the Department of Communications Engineering, Universidad de Cantabria, as postdoctoral fellow, where he is currently working for the Advanced Signal Processing Group. His research interests range from Riemannian optimization to machine learning and their applications to communications.

Carlos Beltrán received the Ph.D. degree in mathematics from the Universidad de Cantabria (UC), Spain, in 2006. He held a Postdoctoral Fellowship at the University of Toronto during 2007 and 2008, and is currently Full Professor on Mathematical Analysis at the Universidad de Cantabria. He has published over 40 research papers including some which describe a probabilistic solution to Smale’s 17th problem and other well-known open problems in the frontiers of Mathematics and Computation. His research interests include numerical analysis, complexity and approximation theory, as well as applied problems. He received the Jose Luis Rubio de Francia 2010 prize (Real Sociedad Matemática Española) and the Stephen Smale 2014 prize (Society for the Foundations of Computational Mathematics) for his contributions. He has delivered more than 50 talks including several plenaries at different international conferences.

Ignacio Santamaria (M’96–SM’05) received the Telecommunication Engineer degree and the Ph.D. degree in electrical engineering from the Universidad Politécnica de Madrid (UPM), Spain, in 1991 and 1995, respectively. Since 2007 he is Full Professor at the Universidad de Cantabria (UC). He has coauthored the book “Coherence in Signal Processing and Machine Learning”, Springer, 2022, and more than 200 publications in refereed journals and international conference papers. His research interests include signal processing algorithms and information-theoretic aspects of multiuser multiantenna wireless communication systems, multivariate statistical techniques and machine learning theories. He has been involved in numerous national and international research projects on these topics. He has been a visiting researcher at the University of Florida (in 2000 and 2004), at the University of Texas at Austin (in 2009), and at the Colorado State University (in 2015). Prof. Santamaria was General Co-Chair of the 2012 IEEE Workshop on Machine Learning for Signal Processing (MLSP 2012). He has been a member of the IEEE Machine Learning for Signal Processing Technical Committee (2009-2014), member of the IEEE Signal Processing Theory and Methods (SPTM) Technical Committee (2020-2022), and member of the IEEE Data Science Initiative (DSI) steering committee (2020-2022). He has been Associate Editor of the IEEE Transactions on Signal Processing (2011-2015), and Senior Area Editor of the IEEE Transactions on Signal Processing (2013-2015). Prof. Santamaria was a co-recipient of the 2008 EEEfCOM Innovation Award for the development of a new RF-MIMO transceiver, coauthor of a paper that received the 2012 IEEE Signal Processing Society Young Author Best Paper Award, and he has recently received the 2022 IHP International Wolfgang Mehr Fellowship Award.

Vít Tuček received his PhD degree in differential geometry and global analysis at the Charles University in Prague, Czech Republic in 2017. In 2018 he joined the Scientific Center of Excellence for Quantum and Complex Systems, and Representations of Lie Algebras (QuantiXLie) at the University of Zagreb, Croatia. Currently he is affiliated with Huawei Sweden R&D where he works on noncoherent communication schemes. His main focus are homogeneous spaces and related representation theory.

Gunnar Peters received the Ph.D. degree in mathematics and the Master of Engineering degree in applied physics from the Royal Institute of Technology. He is a Senior Expert of Network Algorithms with Huawei, Sweden. He has worked in telecom for 19 years and started with Ericsson in 2000, where he was an Expert in Radio Performance. In 2010, he joined the Wireless Organization, Huawei, where he is the Director of Wireless Network Algorithm Lab. He has also worked with the University of South Carolina, Uppsala University, and the Royal Institute of Technology. His research covers stochastic state mechanics, signal processing, and radio resource management. He has received the Individual Gold Medal Award and the Team Gold Medal Award for his work in massive MIMO, advanced receivers, and the application of machine learning to RAN.