A Riemannian Metric for Noncoherent Constellation Design and Its Application to Multiple Access Channel

Khac-Hoang Ngo*, Sheng Yang†
*Department of Electrical Engineering, Chalmers University of Technology, 41296 Gothenburg, Sweden
†Laboratory of Signals and Systems, CentraleSupélec, Paris-Saclay University, 91190 Gif-sur-Yvette, France
Emails: ngok@chalmers.se, sheng.yang@centralesupelec.fr

***Abstract*—We extend the study of the joint constellation design problem for noncoherent multiple-input multiple-output multiple-access channels in Rayleigh block fading. First, we derive the pairwise error probability (PEP) exponent of the noncoherent maximum-likelihood detector to within a multiplicative factor of two. In particular, the lower bound of this exponent is obtained from a Chernoff upper bound of the error probability. Then, we show that the PEP exponent scales linearly with the Riemannian distance between the shifted Gram matrices of the symbols. This gives a geometric interpretation for our proposed metrics: a pair of joint symbols achieves a low PEP if the corresponding shifted Gram matrices are well separated, i.e., joined by a long geodesic, in the manifold of Hermitian positive definite matrices. Finally, we run numerical results to show that our metrics are meaningful for joint constellation design and evaluation, and result in constellations that outperform the constellations optimized with existing metrics and a pilot-based scheme.**

I. Introduction

In noncoherent wireless communications, the channel state information (CSI) is assumed to be unknown *a priori* [1], [2]. This is a reasonable assumption since the wireless channel varies dynamically over time, especially in highly mobile environment. This paper considers the noncoherent multiple-input multiple-output (MIMO) multiple-access channel (MAC) with Rayleigh block fading. Within a coherence block, the matrix-valued signal transmitted by each user is taken from a finite discrete constellation. We study the design of the joint constellation so as to minimize the maximum-likelihood (ML) detection error.

In the single-user case, a noncoherent approach, so-called unitary space-time modulation (USTM), consists in transmitting isotropically distributed and truncated unitary signal matrices within a coherence block [3]. Information is carried by the position of the signal matrix subspace in the Grassmann manifold. This approach was shown to be close to capacity-achieving in the high signal-to-noise-ratio (SNR) regime, which motivates the design of the constellation as a set of points on the Grassmann manifold [4]. A common design criterion for the so-called Grassmannian constellation is to maximize the minimum pairwise chordal distance between the symbols [5], [6]. This is equivalent to maximizing the *sum* of the squared sines of the principle angles between the subspaces represented by the symbols. Another criterion consisting in maximizing the *product* of those squared sines was proposed in [7]. These criteria have been shown to be effective, and inspire various designs of Grassmannian constellations for the point-to-point channel, e.g., [8]–[10].

In the multi-user case, however, the joint constellation design problem has not been as well investigated. A baseline approach is to regard the MAC as a MIMO point-to-to-point channel and thus treat the joint constellation as a Grassmannian constellation (although the joint symbols are not guaranteed to be truncated unitary). Following this approach, one can adopt the aforementioned criteria. In [11], a new design criterion consisting in maximizing the expected pairwise log-likelihood ratio (PLLR) between the symbols was proposed, followed by a simplified version of this criterion. These criteria were shown to be more effective than the baseline and a pilot-based strategy. The design metrics therein are obtained from asymptotic approximations of the worst-case pairwise error probability (PEP) of the joint ML detector.

In this paper, we extend our study of the noncoherent joint constellation design in [11]. We derive new design criteria based on nonasymptotic bounds of the worst-case PEP exponent. Specifically, we first lower bound the PEP exponent using the Chernoff bound. We further upper bound the PEP exponent and obtain a metric which is within a multiplicative factor of 2 from the exponent. We show that this metric, and hence the PEP exponent, for two symbols $\boldsymbol{X}$ and $\boldsymbol{X}'$ scale linearly with the Riemannian distance between the respective shifted Gram matrices $\boldsymbol{I}+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}$ and $\boldsymbol{I}+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}$. This provides a geometric interpretation: a pair of joint symbols $\boldsymbol{X}$ and $\boldsymbol{X}'$ are less likely to be misdetected for each other if $\boldsymbol{I}+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}$ and $\boldsymbol{I}+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}$ are well separated in the manifold of Hermitian positive definite matrices. That is, the geodesic joining $\boldsymbol{I}+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}$ and $\boldsymbol{I}+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}$ in the manifold should have a great length. Finally, we use the proposed metric to numerically optimize the joint constellation. Numerical results show that our metric is effective and meaningful for constellation design and evaluation in both the single-user and multi-user cases.

Notation: Random quantities are denoted with non-italic letters with sans-serif fonts, e.g., a scalar x, a vector $\mathbf{v}$, and a matrix $\mathbf{M}$. Deterministic quantities are denoted with italic letters, e.g., a scalar x, a vector $\boldsymbol{v}$, and a matrix $\boldsymbol{M}$. The Euclidean norm is denoted by $\|\cdot\|$ and the Frobenius norm by $\|\cdot\|_{\mathrm{F}}$. The trace, transpose, and conjugate transpose of $\boldsymbol{M}$ are respectively $\operatorname{tr}(\boldsymbol{M})$, $\boldsymbol{M}^{\mathsf{T}}$, and $\boldsymbol{M}^{\mathsf{H}}$. We denote $[n] := \{1,2,\dots,n\}$.

II. System Model and Problem Formulation

We consider a MIMO MAC consisting of a receiver equipped with N antennas and $K \geq 1$ users, where user k has M_k antennas, $k \in [K]$. The channel is assumed to be flat and block fading with equal-length and synchronous (across the users) coherence intervals of length $T \geq 2$. That is, the channel matrix $\mathsf{H}_k \in \mathbb{C}^{N\times M_k}$ of user k remains constant within each coherence block of T channel uses and changes between blocks. Furthermore, the *distribution* of H_k is assumed to be known, but its *realizations* are unknown to both the users and the receiver. We consider independent and identically distributed (i.i.d.) Rayleigh fading, namely, the rows of $\mathsf{H} := [\mathsf{H}_1\ \mathsf{H}_2\ \ldots\ \mathsf{H}_K]$ are independent and follow $\mathcal{CN}\big(\mathbf{0}, \boldsymbol{I}_{M_{\rm tot}}\big)$ where $M_{\rm tot} := \sum_{k=1}^K M_k$. Within a coherence block, each user k sends a signal matrix symbol $\mathsf{X}_k \in \mathbb{C}^{T\times M_k}$, and the receiver observes

$$\mathsf{Y} = \sum_{k=1}^K \mathsf{X}_k \mathsf{H}_k^{\mathsf{T}} + \mathsf{Z},$$

where the additive noise $\mathsf{Z} \in \mathbb{C}^{T\times N}$ has i.i.d. $\mathcal{CN}(0,1)$ entries independent of $\{\mathsf{H}_k\}$.

We assume that X_k takes value from a *finite constellation* $\mathcal{X}_k$ of fixed size $|\mathcal{X}_k| = 2^{B_k}$ with equally likely symbols, where B_k is the number of bits per symbol for user k. Let $P_k := \frac{1}{T|\mathcal{X}_k|}\sum_{\boldsymbol{X}_k\in\mathcal{X}_k}\|\boldsymbol{X}_k\|_{\rm F}^2$ be the average normalized symbol power of user k. We consider the power constraint $\max_k\{P_k\} = P$, where P is referred to as the SNR. Let us rewrite the channel output as

$$\mathsf{Y} = [\mathsf{X}_1\ \mathsf{X}_2\ \ldots\ \mathsf{X}_K][\mathsf{H}_1\ \mathsf{H}_2\ \ldots\ \mathsf{H}_K]^{\mathsf{T}} + \mathsf{Z} = \mathsf{X}\mathsf{H}^{\mathsf{T}} + \mathsf{Z},$$

where the concatenated signal matrix $\mathsf{X} := [\mathsf{X}_1\ \ \mathsf{X}_2\ \ \ldots\ \ \mathsf{X}_K]$ takes value from the *joint constellation* $\mathcal{X} := \big\{[\boldsymbol{X}_1\ \ \boldsymbol{X}_2\ldots\boldsymbol{X}_K] : \ \boldsymbol{X}_k \in \mathcal{X}_k\big\} = \prod_{k=1}^K \mathcal{X}_k$. Our goal is to derive the desirable properties of the set tuple $(\mathcal{X}_1,\ldots,\mathcal{X}_K)$ for a given rate tuple $(R_1,\ldots,R_K)$ to achieve low *symbol detection error probability*, i.e.,

$$\mathcal{X}^* = \arg\max_{\mathcal{X}} P_e(\mathcal{X}).$$

where $P_e(\mathcal{X}) := \mathbb{P}\left(\Xi(\mathsf{Y}) \neq \mathsf{X}\right)$ with $\Xi(\mathsf{Y})$ denoting the output of the maximum-likelihood (ML) detector. Specifically, $\Xi(\mathsf{Y}) = \arg\max_{\boldsymbol{X}\in\mathcal{X}} p_{\mathsf{Y}\,|\,\mathsf{X}}(\boldsymbol{Y}\,|\,\boldsymbol{X})$ with

$$p_{\mathsf{Y}|\mathsf{X}}(\boldsymbol{Y}|\boldsymbol{X}) = \frac{\exp(-{\rm tr}(\boldsymbol{Y}^{\mathsf{H}}(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1}\boldsymbol{Y}))}{\pi^{NT}{\det}^N(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})}. \quad (1)$$

Since $p_{\mathsf{Y}|\mathsf{X}}(\boldsymbol{Y}|\boldsymbol{X})$ depends on $\boldsymbol{X}$ only through $\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}$, the following proposition is straightforward.

Proposition 1 (Identifiability condition)**.** *For the joint ML detection error probability $P_e(\mathcal{X})$ to vanish at high SNR, the joint constellation $\mathcal{X}$ must satisfy $\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}} \neq \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}$ for any pair of distinct symbols $\boldsymbol{X}$ and $\boldsymbol{X}'$ in $\mathcal{X}$.*

Remark 1. *The analysis in this paper holds for the single-user case (i.e., the MIMO point-to-point channel) by letting $K = 1$. Hereafter, for notational simplicity, we drop the user's index k whenever the single-user case is considered.*

State of the Art: In the single-user case ($K = 1$), following USTM, the constellation symbols are (scaled) truncated unitary matrix representative of points on the Grassman manifold. A common design criterion is $\mathcal{X}^* = \arg\min_{\mathcal{X}} \max_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} {\rm tr}(\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X})$, which coincides with maximizing the pairwise chordal distance [5], [6], [10]. Another criterion derived based on an asymptotic analysis of the pairwise ML error probability is $\mathcal{X}^* = \arg\min_{\mathcal{X}} \ln\sum_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} {\det}^{-N}\big(\boldsymbol{I} - \big(\frac{M}{PT}\big)^2\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}\big)$ [7]. Note that the former criterion is equivalent to maximizing the *sum* of the squared sines of the principle angles between the subspaces represented by the symbols, while the latter criterion aims to maximize the *product* of those squared sines. When $K \geq 2$, $\boldsymbol{X}$ is not guaranteed to be truncated unitary. One can choose to treat the joint constellation as a constellation of an $M_{\rm tot} \times N$ MIMO point-to-point channel and adopt the aforementioned criteria, i.e.,

$$(\text{Min-}m_i)\colon\ \mathcal{X}^* = \arg\min_{\mathcal{X}} m_i(\mathcal{X}), \quad i \in \{1,2\},$$

where $m_1(\mathcal{X}) := \max_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} {\rm tr}\left(\frac{\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}}{\|\boldsymbol{X}\|_{\rm F}^2\|\boldsymbol{X}'\|_{\rm F}^2}\right)$ and $m_2(\mathcal{X}) := \ln\sum_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} {\det}^{-N}\left(\boldsymbol{I} - M_{\rm tot}^2\frac{\boldsymbol{X}^{\mathsf{H}}\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}\boldsymbol{X}}{\|\boldsymbol{X}\|_{\rm F}^2\|\boldsymbol{X}'\|_{\rm F}^2}\right)$.

In [11], a design criterion consisting in maximizing the minimal expected pairwise log-likelihood ratio (PLLR) between the joint symbols was proposed. Specifically, the criterion is

$$(\text{Max-}e_{\min})\colon\ \ \mathcal{X}^* = \arg\max_{\mathcal{X}} e_{\min}(\mathcal{X}),$$

where $e_{\min}(\mathcal{X}) := \frac{1}{N}\min_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} \mathbb{E}\left[\ln\frac{p_{\mathsf{Y}|\mathsf{X}}(\mathsf{Y}|\boldsymbol{X})}{p_{\mathsf{Y}|\mathsf{X}}(\mathsf{Y}|\boldsymbol{X}')}\right]$. A simplified and approximative version of this criteria was also provided

$$(\text{Max-}d_{\min})\colon\ \ \mathcal{X}^* = \arg\max_{\mathcal{X}} d_{\min}(\mathcal{X}),$$

where $d_{\min}(\mathcal{X}) := \min_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} {\rm tr}\big((\boldsymbol{I}_T + \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}\big)$. The criteria Max-$e_{\min}$ and Max-$d_{\min}$ were shown to be more effective than Min-m_1 and Min-m_2 when $K > 1$.

In this paper, we propose new design criteria based on new bounds of the ML error probability, and give a geometric interpretation with the Riemannian distance. We first provide some preliminaries on this distance in the next section.

III. Preliminaries on the Riemannian Distance

Let $\mathcal{P}_T$ be the set of $T \times T$ Hermitian and positive definite matrices. This set is a differential manifold. At a point $\boldsymbol{A}$ of $\mathcal{P}_T$, define the differential $\|\boldsymbol{A}^{-\frac{1}{2}}\,{\rm d}\boldsymbol{A}\boldsymbol{A}^{-\frac{1}{2}}\|_{\rm F} = \big[{\rm tr}\left(\boldsymbol{A}^{-1}\,{\rm d}\boldsymbol{A}\right)^2\big]^{\frac{1}{2}}$. This differential is used to compute the length of a piecewise differential path in $\mathcal{P}_T$. Specifically, the length of a path $\gamma\colon [a,b] \to \mathcal{P}_T$ is given by $L(\gamma) = \int_a^b \|\gamma^{-\frac{1}{2}}(t)\gamma'(t)\gamma^{-\frac{1}{2}}(t)\|_{\rm F}\,{\rm d}t$ [12, Chapter 6]. Define a distance between any two points $\boldsymbol{A}$ and $\boldsymbol{B}$ in $\mathcal{P}_T$ as

$$\delta_{\rm R}(\boldsymbol{A},\boldsymbol{B}) := \inf\{L(\gamma)\colon \gamma \text{ is a path from } \boldsymbol{A} \text{ to } \boldsymbol{B}\}. \quad (2)$$

That is, $\delta_{\rm R}(\boldsymbol{A},\boldsymbol{B})$ is the minimum length of a path between $\boldsymbol{A}$ and $\boldsymbol{B}$. According to [12, Chapter 6], the infimum in (2) is achieved by a unique path joining $\boldsymbol{A}$ and $\boldsymbol{B}$, which is called a *geodesic* from $\boldsymbol{A}$ to $\boldsymbol{B}$. This geodesic has a parameterization

$$\gamma(t) = \boldsymbol{A}^{\frac{1}{2}}\big(\boldsymbol{A}^{-\frac{1}{2}}\boldsymbol{B}\boldsymbol{A}^{-\frac{1}{2}}\big)^t\boldsymbol{A}^{\frac{1}{2}}, \quad 0 \le t \le 1.$$

Furthermore, $\delta_{\mathrm{R}}(\boldsymbol{A},\boldsymbol{B})$ is explicitly given by

$$\delta_{\mathrm{R}}(\boldsymbol{A},\boldsymbol{B}) = \|\ln(\boldsymbol{A}^{-\frac{1}{2}}\boldsymbol{B}\boldsymbol{A}^{-\frac{1}{2}})\|_{\mathrm{F}} = \left(\sum_{i=1}^{T}\ln^2\sigma_i(\boldsymbol{A}^{-1}\boldsymbol{B})\right)^{\frac{1}{2}}$$

where $\{\sigma_i(\boldsymbol{M})\}_i$ denote the eigenvalues of a matrix $\boldsymbol{M}$. The distance $\delta_{\mathrm{R}}(\boldsymbol{A},\boldsymbol{B})$ is called the *Riemannian distance* on the manifold $\mathcal{P}_T$. The readers are referred to [12, Chapter 6] for a further description of this distance and its relation to the geometry of the manifold $\mathcal{P}_T$.

IV. Constellation Design Criteria

We denote the pairwise error event as $\{\boldsymbol{X} \to \boldsymbol{X}'\} := \{p_{\mathsf{Y}|\mathsf{X}}(\mathsf{Y}|\boldsymbol{X}) \le p_{\mathsf{Y}|\mathsf{X}}(\mathsf{Y}|\boldsymbol{X}')|\mathsf{X} = \boldsymbol{X}\}$. As shown in [11, Sec. III], for a given constellation size $|\mathcal{X}|$, $P_e(\mathcal{X})$ vanishes if and only if the *worst-case PEP*, $\max_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} \mathbb{P}(\boldsymbol{X}\to\boldsymbol{X}')$, vanishes. Therefore, our goal is to minimize the worst-case PEP, or equivalently, maximize the worst-case PEP exponent:

$$\mathcal{X}^* = \arg\max_{\mathcal{X}} \min_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} \Big(-\frac{1}{N}\ln\mathbb{P}(\boldsymbol{X}\to\boldsymbol{X}')\Big).$$

A. Pairwise Error Exponent Analysis

Following [13], the PEP can be computed in closed form (see [14, Appendix B]). However, the closed-form expression is rather involved and cannot be directly used to optimize the constellations. Therefore, one needs to resort to tractable bounds. To this end, we analyze the PEP and propose some bounds in the following. Let us first rewrite the PEP as

$$\mathbb{P}(\boldsymbol{X}\to\boldsymbol{X}') = \mathbb{P}\Big(\ln\frac{p_{\mathsf{Y}|\mathsf{X}}(\mathsf{Y}|\boldsymbol{X})}{p_{\mathsf{Y}|\mathsf{X}}(\mathsf{Y}|\boldsymbol{X}')} \le 0\Big) = \mathbb{P}\big(\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}') \le 0\big)$$

with the PLLR $\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}') := \ln\frac{p_{\mathsf{Y}|\mathsf{X}}(\mathsf{Y}|\boldsymbol{X})}{p_{\mathsf{Y}|\mathsf{X}}(\mathsf{Y}|\boldsymbol{X}')}$. Using (1), we obtain

$$\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}') = N\ln\frac{\det\big(\boldsymbol{I}_T + \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}\big)}{\det\big(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}\big)} - \operatorname{tr}\Big(\big((\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1} - (\boldsymbol{I}_T + \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}\big)\mathsf{Y}\mathsf{Y}^{\mathsf{H}}\Big). \quad (3)$$

Hereafter, we denote by $\{\lambda_i\}_{i=1}^T$ the eigenvalues of the matrix $\boldsymbol{\Gamma} := (\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})(\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}$. Note that $\lambda_i \ge 0$, $\forall i\in[T]$.

Lemma 1. *The PEP can be expressed as*

$$\mathbb{P}(\boldsymbol{X}\to\boldsymbol{X}') = \mathbb{P}\left(\sum_{i=1}^{T}(\lambda_i - 1)\mathsf{g}_i \le N\sum_{i=1}^{T}\ln\lambda_i\right),$$

where $\{\mathsf{g}_i\}_{i=1}^T$ are independent Gamma random variables with shape N and scale 1.

Proof. See Appendix A. □

A Chernoff lower bound on the PEP exponent is given in the following proposition.

Proposition 2. *The PEP exponent is lower-bounded as $-\frac{1}{N}\ln\mathbb{P}(\boldsymbol{X}\to\boldsymbol{X}') \ge J_s(\boldsymbol{X},\boldsymbol{X}')$ $\forall s\in[0,1]$, where*

$$J_s(\boldsymbol{X},\boldsymbol{X}') := \ln\det(s(\boldsymbol{I}_T + \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1} + (1-s)(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1}) - \big[s\ln\det((\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}) + (1-s)\ln\det((\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1})\big].$$

Proof. See Appendix B. □

In particular, with $s=\frac{1}{2}$, after some manipulations, we obtain

$$J_{1/2}(\boldsymbol{X},\boldsymbol{X}') = \frac{1}{2}\ln\det\big(2\boldsymbol{I}_T + (\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})^{-1}(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}) + (\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{-1}(\boldsymbol{I}_T + \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})\big) - T\ln 2. \quad (4)$$

The bounds of the PEP exponent can be tightened with an upper bound as follows.

Proposition 3. *The PEP exponent is upper and lower-bounded as*

$$\beta(\boldsymbol{X},\boldsymbol{X}') + T \ge -\frac{1}{N}\ln\mathbb{P}(\boldsymbol{X}\to\boldsymbol{X}') \ge \frac{1}{2}\beta(\boldsymbol{X},\boldsymbol{X}') - T\ln 2. \quad (5)$$

where $\beta(\boldsymbol{X},\boldsymbol{X}')$ is defined through $\{\lambda_i\}$ as $\beta(\boldsymbol{X},\boldsymbol{X}') := \sum_{i=1}^T|\ln\lambda_i|$.

Proof. See Appendix C □

Proposition 3 states that the PEP exponent scales linearly with $\beta(\boldsymbol{X},\boldsymbol{X}')$ where the multiplicative factor is tightly bounded between $\frac{1}{2}$ and 1. Note that the lower limit factor $\frac{1}{2}$ can be improved by optimizing the parameter s in Proposition 2. For the purpose of this paper, however, we neglect the multiplicative and additive factors and focus on the key part $\beta(\boldsymbol{X},\boldsymbol{X}')$ in both upper and lower bounds.

B. Proposed Criteria and Its Relation to the Riemannian Distance

Define $\beta_{\min}(\mathcal{X}) := \min_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}}\beta(\boldsymbol{X},\boldsymbol{X}')$. It follows from Proposition 3 that the worst-case PEP exponent is sandwiched between $\beta_{\min}(\mathcal{X}) + T$ and $\frac{1}{2}\beta_{\min}(\mathcal{X}) - T\ln 2$. Motivated by this, we propose the following design criterion

$$(\text{Max-}\beta_{\min})\colon \mathcal{X}^* = \arg\max_{\mathcal{X}}\beta_{\min}(\mathcal{X}).$$

Furthermore, since the metric $\beta_{\min}(\mathcal{X})$ provides tight bounds on the PEP exponent, it can also be used to evaluate the error performance of a given joint constellation. The higher the value of $\beta_{\min}(\mathcal{X})$, the lower the joint ML detection error is expected to be. Computing $\beta_{\min}(\mathcal{X})$ is more efficient than evaluating the empirical joint ML symbol error rate.

We now present a relation between our β-metric and the Riemannian distance. Since the matrices $\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}$ and $\boldsymbol{I}_T + \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}$ are Hermitian and positive definite, we can define the Riemannian distance between them as $\delta_{\mathrm{R}}(\boldsymbol{I}_T + \boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}, \boldsymbol{I}_T + \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}) = \big(\sum_{i=1}^T\ln^2\lambda_i\big)^{\frac{1}{2}}$.

Proposition 4. *The metric $\beta(\boldsymbol{X},\boldsymbol{X}')$ is bounded in terms of the Riemannian distance $\delta_{\mathrm{R}}(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}},\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})$ as*

$$\sqrt{T}\delta_{\mathrm{R}}(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}},\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}) \geq \beta(\boldsymbol{X},\boldsymbol{X}') \geq \delta_{\mathrm{R}}(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}},\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}).$$

Proof. The lower bound follows from $\beta(\boldsymbol{X},\boldsymbol{X}') = \sum_{i=1}^T |\ln\lambda_i| = \sqrt{\left(\sum_{i=1}^T |\ln\lambda_i|\right)^2} \geq \sqrt{\sum_{i=1}^T \ln^2\lambda_i} = \delta_{\mathrm{R}}(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}},\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})$, where the inequality holds because the terms $|\ln\lambda_i|$ are nonnegative. The upper bound follows directly from the Cauchy-Schwarz inequality. □

This proposition says that the metric $\beta(\boldsymbol{X},\boldsymbol{X}')$ is within a multiplicative factor from the Riemannian distance $\delta_{\mathrm{R}}(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}},\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})$, where the factor is between 1 and $\sqrt{T}$. Therefore, $\beta(\boldsymbol{X},\boldsymbol{X}')$ is large if and only if $\delta_{\mathrm{R}}(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}},\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}})$ is large. It follows that *a pair of joint symbols $\boldsymbol{X}$ for $\boldsymbol{X}'$ are less likely to be misdetected for each other if the geodesic joining $\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}}$ and $\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}$ in $\mathcal{P}_T$ is longer.* If $\boldsymbol{X}\boldsymbol{X} = \boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}$, this geodesic has length zero, thus $\beta(\boldsymbol{X},\boldsymbol{X}')=0$ and the PEP exponent is upper bounded by a constant. Note that this agrees with the identifiability condition in Proposition 1.

In numerical optimization using the proposed metric, one has to compute the gradient of $\beta(\boldsymbol{X},\boldsymbol{X}')$ with respect to the symbols. This can be challenging since $\beta(\boldsymbol{X},\boldsymbol{X}')$ involves the eigenvalues of $\boldsymbol{\Gamma}$. In this regard, it can be more convenient to maximize the bound $J_s(\boldsymbol{X},\boldsymbol{X}')$ given in Proposition 2:

$$(\text{Max-}J_{s,\min}):\ \ \mathcal{X}^* = \arg\max_{\mathcal{X}} J_{s,\min}(\mathcal{X})$$

for some $s\in[0,1]$, where $J_{s,\min}(\mathcal{X}) := \min_{\boldsymbol{X}\neq\boldsymbol{X}'\in\mathcal{X}} J_s(\boldsymbol{X},\boldsymbol{X}')$. In the single-user SIMO case, let $s=\frac{1}{2}$ and consider Grassmannian signaling $\|\boldsymbol{x}\| = PT, \forall \boldsymbol{x}\in\mathcal{X}$, then Max-$J_{s,\min}$ is equivalent to the max-min chordal distance criterion

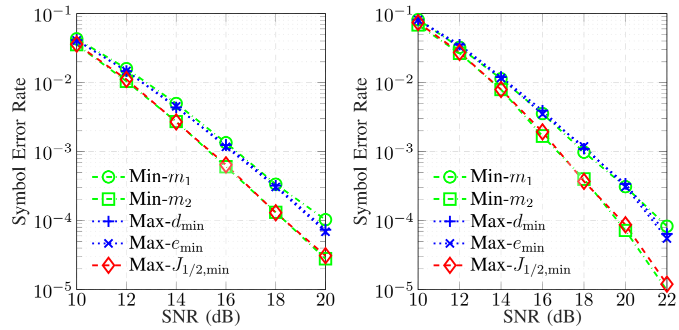
$$\mathcal{X}^* = \arg\max_{\mathcal{X}} \min_{\boldsymbol{x}\neq\boldsymbol{x}'\in\mathcal{X}} \sqrt{1-\frac{1}{P^2T^2}|\boldsymbol{x}^{\mathsf{H}}\boldsymbol{x}'|^2}.$$

V. Numerical Results

We generate the constellation according to the proposed criterion and compare it with the constellations optimized with existing criteria. We assume that $M_k=M$, $B_k=B$, $k\in[K]$ and consider USTM, i.e., $\boldsymbol{X}_k^{\mathsf{H}}\boldsymbol{X}_k = \frac{PT}{M}\boldsymbol{I}_M$, $\forall\boldsymbol{X}_k\in\mathcal{X}_k$, $k\in[K]$. We solve the constellation optimization at $P=30$ dB although the performance of the resulting constellations is benchmarked at other SNR values. We employ the Manopt toolbox [15], and the optimization technique is similar to that described in [14, Sec. VI-A]. We compare the constellation optimized with Max-$J_{1/2,\min}$ and that with Min-m_1, Min-m_2, Max-$e_{\min}$, and Max-$d_{\min}$ in terms of joint-ML symbol error rate (SER) and the metric $\beta_{\min}(\mathcal{X})$.

A. The Single-User Case

We first consider the single-user case, i.e., $K=1$, with coherent interval $T=4$, $B\in\{5,6\}$ bits/symbol, $M=2$ transmit antennas, and $N=2$ receive antennas. In Fig. 1, we show the SER as a function of the SNR P for the constellations obtained by optimizing different metrics. We see that the constellations optimized with the proposed criteria Max-$J_{1/2,\min}$ is on par with that with Min-m_2, and outperforms the constellations optimized with the other metrics. Note that although the metric $m_2(\mathcal{X})$ results from a bound on the PEP as for our metric, it relies on the truncated unitary structure of the symbols. The constellation obtained with Max-$e_{\min}$ and Max-$d_{\min}$ are slightly better than that with Min-m_1.



(a) $B=5$ bits/symbol (b) $B=6$ bits/symbol

Fig. 1. The SER of the constellations optimized with different criteria for $K=1$ user, coherent interval $T=4$, $B\in\{5,6\}$ bits/symbol, $M=2$ transmit antennas, and $N=2$ receive antennas.

B. The Two-User Case

Next, we consider the two-user case, i.e., $K=2$, with coherent interval $T=5$, $B=4$ bits/symbol, $M=2$ transmit antennas, and $N=4$ receive antennas. In Fig. 2, we plot the joint SER of the joint constellation optimized with different criteria. We also consider a pilot-based scheme corresponding to the scenario where K users transmit orthogonal pilots, followed by spatially multiplexed QAM symbols. Note that this pilot-based scheme does not follow USTM. We observe again that optimizing the joint constellation according to Min-$J_{1/2,\min}$ results in the best performance. Our previous criteria Max-$d_{\min}$ and Max-$e_{\min}$ are also effective and results in joint SER that are only slightly higher than that of Min-$J_{1/2,\min}$. On the other hand, the joint SER of the joint constellations optimized with the Min-m_1 and Min-m_2 criteria are significantly higher and go down more slowly with the SNR. This confirms that these baseline criteria become ineffective for the multi-user case, where the truncated unitary structure of the joint symbols is not guaranteed. The pilot-based scheme outperforms these baselines, but entails about 2 dB power loss to achieve the same joint SER with respect to our Min-$J_{1/2,\min}$ criteria.

Finally, considering the same setting, we depict the values of the metrics $J_{1/2,\min}(\mathcal{X})$ and $\beta_{\min}(\mathcal{X})$ evaluated for the considered constellations in Fig. 3. As can be seen, the constellations with lower joint SER have a larger value of these metrics. In other words, the order of the metric values reveals the order of the SER performance. This confirms that our proposed metrics are meaningful for constellation design and evaluation.

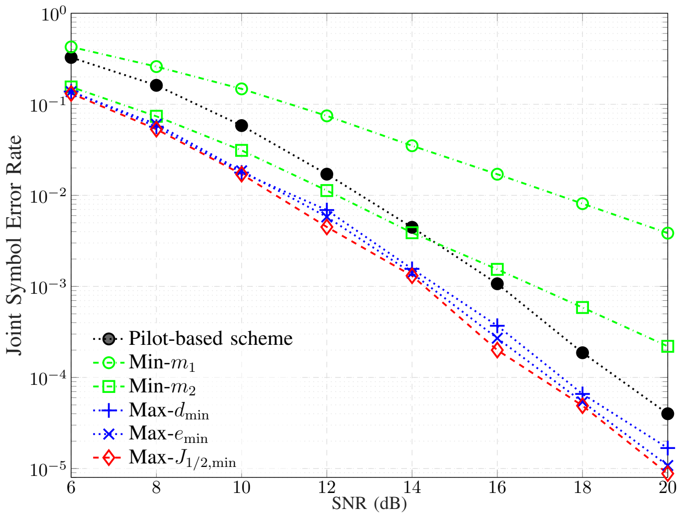
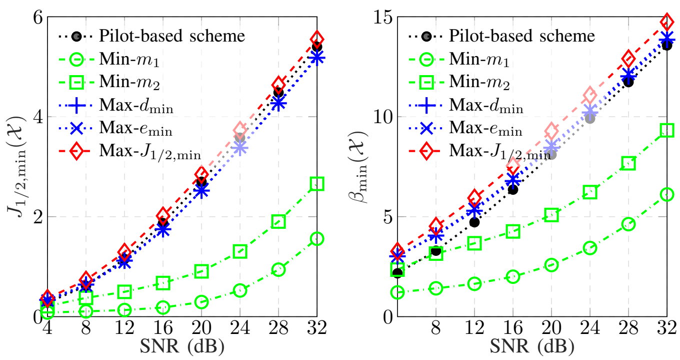


Fig. 2. The SER of the constellations optimized with different criteria for $K = 2$ users, coherent interval $T = 5$, $B = 4$ bits/symbol, $M = 2$ transmit antennas, and $N = 4$ receive antennas.



(a) $J_{1/2,\min}(\mathcal{X})$ (b) $\beta_{\min}(\mathcal{X})$

Fig. 3. The value of the design metrics $J_{1/2,\min}(\mathcal{X})$ and $\beta_{\min}(\mathcal{X})$ for the considered constellations for the same setting as in Fig. 2.

VI. Conclusion

We investigate the joint constellation design for noncoherent MIMO MAC in Rayleigh block fading. By analyzing the exponent of the worst-case pairwise joint-ML detection error, we have derived closed-form metrics which are effective for designing joint constellations that achieve a low error rate. We provided a geometric interpretation of our metric using the Riemannian distance.

Appendix A
Proof of Lemma 1

It suffices to show that the PLLR can be written as $-N\sum_{i=1}^{T}\ln\lambda_i+\sum_{i=1}^{T}(\lambda_i-1)\mathsf{g}_i$. Let $\mathbf{Y}_0 := (\boldsymbol{I}_T+\mathbf{X}\mathbf{X}^{\mathsf{H}})^{-1/2}\mathbf{Y}$ be a "whitened" version of $\mathbf{Y}$, then $\mathbf{Y}_0$ is a Gaussian matrix with T independent rows following $\mathcal{CN}(\mathbf{0},\boldsymbol{I}_N)$. From (3), the PLLR $\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')$ can be expressed as

$$
\begin{aligned}
&\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}') = -N\ln\det(\boldsymbol{\Gamma})\\
&+\operatorname{tr}\Big(\big((\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{\frac12}\big(\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}\big)^{-1}(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{\frac12}-\boldsymbol{I}_T\big)\\
&\qquad\times\mathbf{Y}_0\mathbf{Y}_0^{\mathsf{H}}\Big).
\end{aligned}
$$

Since $\boldsymbol{\Gamma}$ and $(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{\frac12}\big(\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}\big)^{-1}(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{\frac12}$ share the same eigenvalues $\{\lambda_i\}_{i=1}^T$, we can decompose $(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{\frac12}\big(\boldsymbol{I}_T+\boldsymbol{X}'\boldsymbol{X}'^{\mathsf{H}}\big)^{-1}(\boldsymbol{I}_T+\boldsymbol{X}\boldsymbol{X}^{\mathsf{H}})^{\frac12}-\boldsymbol{I}_T = \bar{\boldsymbol{U}}\operatorname{diag}(\lambda_1-1,\lambda_2-1,\dots,\lambda_T-1)\bar{\boldsymbol{U}}^{\mathsf{H}}$ where $\bar{\boldsymbol{U}}$ is a $T\times T$ unitary matrix. We further expand the PLLR as

$$
\begin{aligned}
&\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\\
&= -N\ln\det(\boldsymbol{\Gamma})+\operatorname{tr}\Big(\operatorname{diag}(\lambda_1-1,\dots,\lambda_T-1)\,\bar{\boldsymbol{U}}^{\mathsf{H}}\mathbf{Y}_0\mathbf{Y}_0^{\mathsf{H}}\bar{\boldsymbol{U}}\Big)\\
&= -N\sum_{i=1}^{T}\ln\lambda_i+\sum_{i=1}^{T}(\lambda_i-1)\mathsf{g}_i
\end{aligned}
$$

where $\mathsf{g}_i := \|\bar{\mathbf{y}}_{0,i}\|^2$ with $\bar{\mathbf{y}}_{0,i}$ being the i-th row of $\bar{\boldsymbol{U}}^{\mathsf{H}}\mathbf{Y}_0$. Since $\bar{\boldsymbol{U}}$ is unitary and deterministic, $\bar{\boldsymbol{U}}^{\mathsf{H}}\mathbf{Y}_0$ has the same distribution as $\mathbf{Y}_0$, i.e., $\bar{\mathbf{y}}_{0,i}$ are independent and follow $\mathcal{CN}(\mathbf{0},\boldsymbol{I}_N)$. Therefore, $\{\mathsf{g}_i\}_{i=1}^T$ are independent Gamma random variables with shape N and scale 1. This completes the proof.

Appendix B
Proof of Proposition 2

In this proof, for convenience, we denote $\boldsymbol{G}_{\boldsymbol{A}} := (\boldsymbol{I}+\boldsymbol{A}\boldsymbol{A}^{\mathsf{H}})^{-1}$ for a matrix $\boldsymbol{A}$. We need to show that $\mathbb{P}(\boldsymbol{X}\to\boldsymbol{X}') = \mathbb{P}\big(\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\le 0\big)\le\exp(-NJ_s(\boldsymbol{X},\boldsymbol{X}'))$. By applying the Chernoff bound [16, Th. 6.2.7], we obtain for every $s>0$ that

$$
\begin{aligned}
&\mathbb{P}(\boldsymbol{X}\to\boldsymbol{X}')\\
&\le \mathbb{E}_{\mathbf{Y}\,|\,\boldsymbol{X}}\left[\exp\left(-s\mathsf{L}(\boldsymbol{X}\to\boldsymbol{X}')\right)\right]\\
&= \mathbb{E}_{\mathbf{Y}\,|\,\boldsymbol{X}}\left[\left(\frac{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\boldsymbol{X}')}{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{Y}|\boldsymbol{X})}\right)^{s}\right]\\
&= \int_{\mathbb{C}^{T\times N}}[p_{\mathbf{Y}|\mathbf{X}}(\boldsymbol{Y}|\boldsymbol{X}')]^{s}[p_{\mathbf{Y}|\mathbf{X}}(\boldsymbol{Y}|\boldsymbol{X})]^{1-s}\,\mathrm{d}\boldsymbol{Y}\\
&= \int_{\mathbb{C}^{T\times N}}\left[\frac{\exp(-\operatorname{tr}(\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{G}_{\boldsymbol{X}'}\boldsymbol{Y}))}{\pi^{NT}\det^{-N}\boldsymbol{G}_{\boldsymbol{X}'}}\right]^{s}\left[\frac{\exp(-\operatorname{tr}(\boldsymbol{Y}^{\mathsf{H}}\boldsymbol{G}_{\boldsymbol{X}}\boldsymbol{Y}))}{\pi^{NT}\det^{-N}\boldsymbol{G}_{\boldsymbol{X}}}\right]^{1-s}\mathrm{d}\boldsymbol{Y} \qquad (6)\\
&= \left[\frac{\det^{s}(\boldsymbol{G}_{\boldsymbol{X}'})\det^{1-s}(\boldsymbol{G}_{\boldsymbol{X}})}{\det(s\boldsymbol{G}_{\boldsymbol{X}'}+(1-s)\boldsymbol{G}_{\boldsymbol{X}})}\right]^{N}\\
&\quad\times\int_{\mathbb{C}^{T\times N}}\frac{\exp(-\operatorname{tr}(\boldsymbol{Y}^{\mathsf{H}}(s\boldsymbol{G}_{\boldsymbol{X}'}+(1-s)\boldsymbol{G}_{\boldsymbol{X}})\boldsymbol{Y}))}{\pi^{NT}\det^{-N}(s\boldsymbol{G}_{\boldsymbol{X}'}+(1-s)\boldsymbol{G}_{\boldsymbol{X}})}\,\mathrm{d}\boldsymbol{Y} \qquad (7)
\end{aligned}
$$

where (6) follows from (1), and (7) follows after some simple manipulations. With $s\in[0,1]$, $(s\boldsymbol{G}_{\boldsymbol{X}'}+(1-s)\boldsymbol{G}_{\boldsymbol{X}})^{-1}$ is a covariance matrix. Therefore, the integral in (7) is an integral of a Gaussian density over the whole support, and thus equals 1. As a consequence, $\mathbb{P}(\boldsymbol{X}\to\boldsymbol{X}')$ is upper-bounded by the first term in (7), which equals $\exp(-NJ_s(\boldsymbol{X},\boldsymbol{X}'))$.

Appendix C
Proof of Proposition 3

The lower bound in (5) follows by taking $s = \frac{1}{2}$ in Proposition 2 and by bounding $J_{1/2}(\boldsymbol{X}, \boldsymbol{X}')$ in (4) as

$$\begin{aligned} J_{1/2}(\boldsymbol{X}, \boldsymbol{X}') &= \frac{1}{2}\sum_{i=1}^{T} \ln\left(2 + \lambda_i + \frac{1}{\lambda_i}\right) - T\ln 2 \\ &\ge \frac{1}{2}\sum_{i=1}^{T} \ln\max\left\{\lambda_i, \frac{1}{\lambda_i}\right\} - T\ln 2 \\ &= \frac{1}{2}\sum_{i=1}^{T} |\ln\lambda_i| - T\ln 2. \end{aligned}$$

To show the upper bound, we first write the Gamma random variables g_i as $\mathsf{g}_i = \sum_{j=1}^{N} \mathsf{e}_{i,j}$, $i \in [T]$, where $\{\mathsf{e}_{i,j}\}_{i\in[T], j\in[N]}$ are independent exponential random variables with parameter 1. From this and Lemma 1, we can bound the PEP as

$$\begin{aligned} \mathbb{P}(\boldsymbol{X} \to \boldsymbol{X}') &= \mathbb{P}\bigg(\sum_{i=1}^{T}\sum_{j=1}^{N}(\lambda_i - 1)\mathsf{e}_{i,j} \le N\sum_{i=1}^{T}\ln\lambda_i\bigg) \\ &\ge \mathbb{P}\left((\lambda_i - 1)\mathsf{e}_{i,j} \le \ln\lambda_i, \forall i \in [T], j \in [N]\right) \\ &= \prod_{i=1}^{T}\prod_{j=1}^{N}\mathbb{P}\left((\lambda_i - 1)\mathsf{e}_{i,j} \le \ln\lambda_i\right) \\ &= \exp\bigg(-N\sum_{i=1}^{T} f(\lambda_i)\bigg) \qquad (8) \end{aligned}$$

where $f(\lambda) := -\ln\mathbb{P}\left((\lambda-1)\mathsf{e} \le \ln\lambda\right)$ with e being an exponential random variable with parameter 1. We shall show that

$$f(\lambda) \le |\ln\lambda| + 1, \quad \forall\lambda \ge 0. \qquad (9)$$

- If $\lambda = 1$, (9) obviously holds with equality.
- If $\lambda < 1$, we have that $\mathbb{P}\left((\lambda-1)\mathsf{e} \le \ln\lambda\right) = \mathbb{P}\big(\mathsf{e} \ge \frac{-\ln\lambda}{1-\lambda}\big) = \exp\big(\frac{\ln\lambda}{1-\lambda}\big)$. Thus $f(\lambda) = \frac{-\ln\lambda}{1-\lambda} = -\ln\lambda + \frac{\ln\lambda^{-1}}{\lambda^{-1}-1} < |\ln\lambda| + 1$ since $\ln\lambda^{-1} < \lambda^{-1} - 1$ for all $\lambda^{-1} > 1$.
- If $\lambda > 1$, we have that $\mathbb{P}\left((\lambda-1)\mathsf{e} \le \ln\lambda\right) = \mathbb{P}\left(\mathsf{e} \le \frac{\ln\lambda}{\lambda-1}\right) = 1 - \exp\left(-\frac{\ln\lambda}{\lambda-1}\right) \ge \frac{1-e^{-1}}{\lambda}$. To verify the inequality, notice that the function $\lambda\left(1 - \exp\left(-\frac{\ln\lambda}{\lambda-1}\right)\right) = \lambda - \lambda^{-\frac{1}{\lambda-1}+1}$ is increasing for $\lambda > 1$, and converges from above to $1 - e^{-1}$ as λ approaches 1 from above. We deduce that $f(\lambda) \le \ln(\frac{\lambda}{1-e^{-1}}) = \ln\lambda - \ln(1 - e^{-1}) < |\ln\lambda| + 1$.

Introducing (9) into (8), we upper bound the PEP exponent as

$$-\frac{1}{N}\ln\mathbb{P}\left(\boldsymbol{X} \to \boldsymbol{X}'\right) \le \sum_{i=1}^{T} f(\lambda_i) \le \sum_{i=1}^{T}|\ln\lambda_i| + T.$$

This completes the proof.

References

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

[2] W. Yang, G. Durisi, and E. Riegler, "On the capacity of large-MIMO block-fading channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 117–132, Feb. 2013.

[3] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 543–564, Mar. 2000.

[4] R. H. Gohary and H. Yanikomeroglu, "Noncoherent MIMO signaling for block-fading channels: Approaches and challenges," *IEEE Veh. Technol. Mag.*, vol. 14, no. 1, pp. 80–88, Mar. 2019.

[5] R. H. Gohary and T. N. Davidson, "Noncoherent MIMO communication: Grassmannian constellations and efficient detection," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1176–1205, Mar. 2009.

[6] D. Agrawal, T. J. Richardson, and R. L. Urbanke, "Multiple-antenna signal constellations for fading channels," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2618–2626, Sep. 2001.

[7] M. L. McCloud, M. Brehler, and M. K. Varanasi, "Signal design and convolutional coding for noncoherent space-time communication on the block-Rayleigh-fading channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 5, pp. 1186–1194, May 2002.

[8] W. Zhao, G. Leus, and G. B. Giannakis, "Orthogonal design of unitary constellations for uncoded and trellis-coded noncoherent space-time systems," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1319–1327, June 2004.

[9] I. Kammoun, A. M. Cipriano, and J. C. Belfiore, "Non-coherent codes over the Grassmannian," *IEEE Trans. Wireless Commun.*, vol. 6, no. 10, pp. 3657–3667, Oct. 2007.

[10] K.-H. Ngo, A. Decurninge, M. Guillaud, and S. Yang, "Cube-split: A structured Grassmannian constellation for non-coherent SIMO communications," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 1948–1964, Mar. 2020.

[11] K.-H. Ngo, S. Yang, M. Guillaud, and A. Decurninge, "Noncoherent MIMO multiple-access channels: A joint constellation design," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Riva del Garda, Italy, Apr. 2021.

[12] R. Bhatia, *Positive definite matrices*. Princeton, New Jersey: Princeton University Press, 2009.

[13] M. Brehler and M. K. Varanasi, "Noncoherent multiuser space-time communications: Optimum receivers and signal design," in *35th Annual Conference on Information Sciences and Systems (CISS)*. The Johns Hopkins University, Mar. 2001, pp. 379–383.

[14] K.-H. Ngo, S. Yang, M. Guillaud, and A. Decurninge, "Joint constellation design for non-coherent MIMO multiple-access channels," *arXiv preprint arXiv:2009.11548v1*, 2020.

[15] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre, "Manopt, a Matlab toolbox for optimization on manifolds," *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1455–1459, Jan. 2014. [Online]. Available: http://www.manopt.org

[16] M. H. DeGroot and M. J. Schervish, *Probability and Statistics*, 4th ed. Pearson, 2012.