

Monotonicity Results for Coherent MIMO Rician Channels

Daniel Hoesli, *Student Member, IEEE*,
Young-Han Kim, *Student Member, IEEE*, and
Amos Lapidoth, *Fellow, IEEE*

Abstract—The dependence of the Gaussian input information rate on the line-of-sight (LOS) matrix in multiple-input multiple-output (MIMO) coherent Rician fading channels is explored. It is proved that the outage probability and the mutual information induced by a multivariate circularly symmetric Gaussian input with any covariance matrix are monotonic in the LOS matrix \mathbf{D} , or more precisely, monotonic in $\mathbf{D}^\dagger \mathbf{D}$ in the sense of the Loewner partial order. Conversely, it is also demonstrated that this ordering on the LOS matrices is a necessary condition for the uniform monotonicity over all input covariance matrices. This result is subsequently applied to prove the monotonicity of the isotropic Gaussian input information rate and channel capacity in the singular values of the LOS matrix. Extensions to multiple-access channels (MAC) are also provided.

Index Terms—Capacity, line of sight, multiple-access channels (MAC), multiple-input multiple-output (MIMO), outage, partial order, preorder, Rician fading.

I. INTRODUCTION AND MAIN RESULT

It is well known that the capacity of a single-input single-output coherent Rician fading channel is monotonic in the magnitude of the line-of-sight (LOS) component. This can be easily deduced from the facts that channel capacity is achieved by a zero-mean circularly symmetric Gaussian input and that a noncentral chi-square random variable is stochastically monotonic in the noncentrality parameter [1, Lemma 6.2 (b)], [2]. This result extends easily to the single-input multiple-output and, with a little more work, to multiple-input single-output scenarios, from the similar stochastic monotonicity for the noncentral chi-square random variable of more degrees of freedom.

The extension to the multiple-input multiple-output (MIMO) case, which may look straightforward at first, requires some extra care. The first difficulty one encounters is that in order to demonstrate the monotonicity, one has to introduce an ordering on the LOS matrices and it is *a priori* unclear what the natural ordering is for the problem at hand. The second difficulty is that there is no closed-form expression for the capacity achieving input distribution. It is straightforward to demonstrate that the capacity is achieved by a circularly symmetric multivariate Gaussian input, but no closed-form expression for the eigenvalues of the optimal covariance matrix is known. Finally, as in the single-input case, under a fixed input distribution, one LOS matrix may give rise to a larger information rate than another LOS matrix for a particular fading

realization, but it may actually perform worse when averaged over all realizations.

In this correspondence, we show that the natural ordering on the LOS matrices \mathbf{D} is given by the Loewner partial order on $\mathbf{D}^\dagger \mathbf{D}$, and through this ordering we extend the monotonicity results to MIMO Rician channels. More specifically, we say that an $m \times n$ LOS matrix \mathbf{D} is “larger than or equal to” another $m \times n$ LOS matrix $\tilde{\mathbf{D}}$, if $\mathbf{D}^\dagger \mathbf{D}$ is greater than or equal to $\tilde{\mathbf{D}}^\dagger \tilde{\mathbf{D}}$ in the Loewner sense, i.e., if $\mathbf{D}^\dagger \mathbf{D} - \tilde{\mathbf{D}}^\dagger \tilde{\mathbf{D}}$ is a positive semidefinite $n \times n$ matrix.¹ (Here \mathbf{D}^\dagger is the Hermitian conjugate of \mathbf{D} .) Under this ordering on the LOS matrices, we shall show the monotonicity of channel capacity, the monotonicity of the information rate corresponding to the isotropic Gaussian input, and the monotonicity of outage probability.

We shall also extend the discussion to the multiple-access channel (MAC). The MAC poses an additional challenge in that the capacity region depends not only on the LOS matrices of the different users individually, but also on how these matrices relate to each other. This requires a joint preorder on LOS matrices, as will be made clear in Section II.

It should be emphasized that our monotonicity results are proved when the distribution of the granular component is held fixed. Consequently, as we vary the LOS matrix the output power is not held fixed. See [5]–[8] for studies where the output power is held fixed.

We state our main theorem, from which the monotonicity results will follow.

Theorem 1.1: Let \mathbf{H} be a random $m \times n$ matrix whose components are independent and identically distributed (i.i.d.) according to a zero-mean unit-variance circularly symmetric complex Gaussian distribution. If two deterministic complex $m \times n$ matrices \mathbf{D} and $\tilde{\mathbf{D}}$ are such that

$$\mathbf{D}^\dagger \mathbf{D} \succeq \tilde{\mathbf{D}}^\dagger \tilde{\mathbf{D}} \quad (1)$$

then

$$\Pr \left[\log \det \left(\mathbf{I}_m + (\mathbf{H} + \mathbf{D})\mathbf{K}(\mathbf{H} + \mathbf{D})^\dagger \right) \leq t \right] \leq \Pr \left[\log \det \left(\mathbf{I}_m + (\mathbf{H} + \tilde{\mathbf{D}})\mathbf{K}(\mathbf{H} + \tilde{\mathbf{D}})^\dagger \right) \leq t \right] \quad (2)$$

for any $t \geq 0$ and any positive semidefinite $n \times n$ matrix \mathbf{K} .

In this theorem and throughout, the notation $\mathbf{A} \succeq \mathbf{B}$ indicates that $\mathbf{A} - \mathbf{B}$ is positive semidefinite. The notation \mathbf{I}_m stands for the $m \times m$ identity matrix. We use $\mathcal{H}^+(n)$ to denote the set of all $n \times n$ positive semidefinite Hermitian matrices and use $\mathcal{U}(n)$ for the set of all unitary $n \times n$ matrices. For a complex matrix \mathbf{A} , \mathbf{A}^\top denotes its transpose while \mathbf{A}^\dagger denotes its Hermitian conjugate (i.e., elementwise complex conjugate of \mathbf{A}^\top). We extend the usual notion of diagonality to nonsquare matrices by saying that a matrix \mathbf{A} of any size is *diagonal* if $A_{ij} = 0$ for all $i \neq j$. All vectors are column vectors unless specified otherwise. All logarithms are natural, i.e., to the base e .

In Section II, we describe the single-user and the multiple-access Rician fading channels and present the main corollaries of Theorem 1.1. The proof of Theorem 1.1 is given in Section III.

¹We point out that the Loewner partial order on $\mathbf{D}^\dagger \mathbf{D}$ induces a *preorder* on the LOS matrices \mathbf{D} , for $\mathbf{D}^\dagger \mathbf{D} \succeq \tilde{\mathbf{D}}^\dagger \tilde{\mathbf{D}}$ and $\tilde{\mathbf{D}}^\dagger \tilde{\mathbf{D}} \succeq \mathbf{D}^\dagger \mathbf{D}$ implies $\mathbf{D}^\dagger \mathbf{D} = \tilde{\mathbf{D}}^\dagger \tilde{\mathbf{D}}$, but *not* $\mathbf{D} = \tilde{\mathbf{D}}$. It only implies $\tilde{\mathbf{D}} = \mathbf{U}\mathbf{D}$ for some unitary matrix \mathbf{U} (see, for instance, [3], [4]).

Manuscript received December 13, 2004; revised June 22, 2005. The work of Y.-H. Kim was supported in part by the National Science Foundation under Grant CCR-0311633. This collaboration started when Y.-H. Kim was visiting the Signal and Information Processing Laboratory (ISI), ETH Zurich, Switzerland, in summer 2002. The material in this correspondence was presented in part at the IEEE International Symposium on Information Theory, Yokohama, Japan, June/July 2003, at the International ITG Conference on Source and Channel Coding, Erlangen, Germany, January 2004, and at the IEEE International Symposium on Information Theory, Adelaide, Australia, September 2005.

D. Hoesli and A. Lapidoth are with the Signal and Information Processing Laboratory (ISI), Swiss Federal Institute of Technology (ETH), Zurich, Switzerland (e-mail: hoesli@isi.ee.ethz.ch; lapidoth@isi.ee.ethz.ch).

Y.-H. Kim is with the Information Systems Laboratory, Stanford University, Stanford, CA 94305 USA (e-mail: yhk@stanford.edu).

Communicated by Y. Steinberg, Associate Editor for Shannon Theory. Digital Object Identifier 10.1109/TIT.2005.858968

II. APPLICATIONS

We introduce two functions that will simplify the notation in our subsequent discussion. In the notation of Theorem 1.1, we define for any $t \geq 0$ and $\mathbf{K} \in \mathcal{H}^+(n)$

$$F(t; \mathbf{K}, \mathbf{D}) \triangleq \Pr \left[\log \det \left(\mathbf{I}_m + (\mathbf{H} + \mathbf{D})\mathbf{K}(\mathbf{H} + \mathbf{D})^\dagger \right) \leq t \right] \quad (3)$$

and

$$\mathcal{I}(\mathbf{K}, \mathbf{D}) \triangleq \mathbb{E} \left[\log \det \left(\mathbf{I}_m + (\mathbf{H} + \mathbf{D})\mathbf{K}(\mathbf{H} + \mathbf{D})^\dagger \right) \right]. \quad (4)$$

Noting that

$$\mathcal{I}(\mathbf{K}, \mathbf{D}) = \int_0^\infty (1 - F(t; \mathbf{K}, \mathbf{D})) dt \quad (5)$$

we obtain the following corollary of Theorem 1.1.

Corollary 2.1: If $\mathbf{D}^\dagger \mathbf{D} \succeq \tilde{\mathbf{D}}^\dagger \tilde{\mathbf{D}}$, then

$$\mathcal{I}(\mathbf{K}, \mathbf{D}) \geq \mathcal{I}(\mathbf{K}, \tilde{\mathbf{D}}), \quad \forall \mathbf{K} \in \mathcal{H}^+(n). \quad (6)$$

The following converse to Corollary 2.1 shows that the preorder on the LOS matrices is natural.

Proposition 2.2: If $\mathcal{I}(\mathbf{K}, \mathbf{D}) \geq \mathcal{I}(\mathbf{K}, \tilde{\mathbf{D}})$ for all $\mathbf{K} \in \mathcal{H}^+(n)$, then $\mathbf{D}^\dagger \mathbf{D} \succeq \tilde{\mathbf{D}}^\dagger \tilde{\mathbf{D}}$.

Proof: See the Appendix. \square

We further note the rotational symmetry in $F(t; \mathbf{K}, \mathbf{D})$ and $\mathcal{I}(\mathbf{K}, \mathbf{D})$. First observe that the law of \mathbf{H} is invariant under left and right rotations, i.e., for any $\mathbf{U} \in \mathcal{U}(m)$ and $\mathbf{V} \in \mathcal{U}(n)$,

$$\mathbf{U}\mathbf{H}\mathbf{V}^\dagger \stackrel{\mathcal{L}}{\cong} \mathbf{H}. \quad (7)$$

Consequently, we have for any $\mathbf{U} \in \mathcal{U}(m)$ and $\mathbf{V} \in \mathcal{U}(n)$

$$\begin{aligned} F(t; \mathbf{K}, \mathbf{U}\mathbf{D}\mathbf{V}^\dagger) &= \Pr \left[\log \det \left(\mathbf{I}_m + (\mathbf{H} + \mathbf{U}\mathbf{D}\mathbf{V}^\dagger)\mathbf{K}(\mathbf{H} + \mathbf{U}\mathbf{D}\mathbf{V}^\dagger)^\dagger \right) \leq t \right] \\ &= \Pr \left[\log \det \left(\mathbf{U} \left(\mathbf{I}_m + (\mathbf{H} + \mathbf{D})\mathbf{V}^\dagger \mathbf{K} \mathbf{V} (\mathbf{H} + \mathbf{D})^\dagger \right) \mathbf{U}^\dagger \right) \leq t \right] \end{aligned} \quad (8)$$

$$= \Pr \left[\log \det \left(\mathbf{I}_m + (\mathbf{H} + \mathbf{D})\mathbf{V}^\dagger \mathbf{K} \mathbf{V} (\mathbf{H} + \mathbf{D})^\dagger \right) \leq t \right] \quad (9)$$

$$= \Pr \left[\log \det \left(\mathbf{I}_m + (\mathbf{H} + \mathbf{D})\mathbf{V}^\dagger \mathbf{K} \mathbf{V} (\mathbf{H} + \mathbf{D})^\dagger \right) \leq t \right] \quad (10)$$

$$= F(t; \mathbf{V}^\dagger \mathbf{K} \mathbf{V}, \mathbf{D}). \quad (11)$$

From this and (5), we also have

$$\mathcal{I}(\mathbf{K}, \mathbf{U}\mathbf{D}\mathbf{V}^\dagger) = \mathcal{I}(\mathbf{V}^\dagger \mathbf{K} \mathbf{V}, \mathbf{D}). \quad (12)$$

A. The Single-User Rician Fading Channel

The output (\mathbf{H}, \mathbf{Y}) of the coherent single-user Rician (or Ricean in certain dialects) fading channel consists of a random $m \times n$ matrix \mathbf{H} whose components are i.i.d. according to the zero-mean unit-variance circularly symmetric complex Gaussian distribution $\mathcal{N}_{\mathbb{C}}(0, 1)$, and of a random m -vector $\mathbf{Y} \in \mathbb{C}^m$ given by

$$\mathbf{Y} = (\mathbf{H} + \mathbf{D})\mathbf{x} + \mathbf{Z} \quad (13)$$

where $\mathbf{x} \in \mathbb{C}^n$ is the channel input; \mathbf{D} is a deterministic $m \times n$ complex LOS matrix; and $\mathbf{Z} \in \mathbb{C}^m$ is drawn according to the zero-

mean circularly symmetric complex multivariate Gaussian distribution $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sigma^2 \mathbf{I}_m)$ for some $\sigma^2 > 0$. It is assumed that \mathbf{H} and \mathbf{Z} are independent of each other, and that their joint law does not depend on the channel input \mathbf{x} .

Since the law of \mathbf{H} does not depend on \mathbf{x} , we can express the mutual information between the channel input and output as

$$I(\mathbf{X}; \mathbf{H}, \mathbf{Y}) = I(\mathbf{X}; \mathbf{Y} | \mathbf{H}). \quad (14)$$

Among all input distributions of a given covariance matrix, the zero-mean circularly symmetric multivariate complex Gaussian maximizes the conditional mutual information $I(\mathbf{X}; \mathbf{Y} | \mathbf{H} = \mathbf{H})$, irrespective of the realization $\mathbf{H} = \mathbf{H}$. Consequently, it also maximizes the average mutual information $I(\mathbf{X}; \mathbf{Y} | \mathbf{H})$. We shall, therefore, consider zero-mean circularly symmetric Gaussian input distributions $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{K})$ only and focus on the dependence of mutual information on the LOS matrix \mathbf{D} when the input covariance matrix \mathbf{K} is held fixed. Also, since we can absorb the dependence on σ^2 into \mathbf{K} , we assume $\sigma^2 = 1$ without loss of generality.

For a given realization $\mathbf{H} = \mathbf{H}$, we can express the conditional mutual information $I(\mathbf{X}; \mathbf{Y} | \mathbf{H} = \mathbf{H})$ for a $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{K})$ input as

$$I(\mathbf{X}; \mathbf{Y} | \mathbf{H} = \mathbf{H}) = \log \det \left(\mathbf{I}_m + (\mathbf{H} + \mathbf{D})\mathbf{K}(\mathbf{H} + \mathbf{D})^\dagger \right). \quad (15)$$

By taking the expectation with respect to \mathbf{H} , we can express the average conditional mutual information as an explicit function of \mathbf{K} and \mathbf{D} as

$$I(\mathbf{X}; \mathbf{Y} | \mathbf{H}) = \mathbb{E} \left[\log \det \left(\mathbf{I}_m + (\mathbf{H} + \mathbf{D})\mathbf{K}(\mathbf{H} + \mathbf{D})^\dagger \right) \right] \quad (16)$$

$$= \mathcal{I}(\mathbf{K}, \mathbf{D}). \quad (17)$$

Thus Corollary 2.1 can be interpreted as the monotonicity of the average conditional mutual information over the Rician fading channel (13) induced by Gaussian inputs with fixed input covariance matrix. We can also give a more direct interpretation of Theorem 1.1 through the notion of *outage probability*. Consider the probability

$$\Pr \left[\log \det \left(\mathbf{I}_m + (\mathbf{H} + \mathbf{D})\mathbf{K}(\mathbf{H} + \mathbf{D})^\dagger \right) \leq R \right] = F(R; \mathbf{K}, \mathbf{D}). \quad (18)$$

We can interpret this quantity as the probability that the realization \mathbf{H} of \mathbf{H} will be such that the information rate on the Gaussian channel $\mathbf{Y} = (\mathbf{H} + \mathbf{D})\mathbf{x} + \mathbf{Z}$ for the input distribution $\mathcal{N}_{\mathbb{C}}(0, \mathbf{K})$ does not exceed R . Under this interpretation, Theorem 1.1 can be viewed as the monotonicity of the outage probability in the channel LOS matrix.

These monotonicity results can be used to study the power- \mathcal{E} isotropic Gaussian input information rate

$$I^{\text{IG}}(\mathcal{E}, \mathbf{D}) \triangleq \mathcal{I} \left(\frac{\mathcal{E}}{n} \mathbf{I}_n, \mathbf{D} \right) \quad (19)$$

and the capacity $C(\mathcal{E}, \mathbf{D})$ of the Rician channel under the average input power constraint $\mathbb{E}[\mathbf{X}^\dagger \mathbf{X}] \leq \mathcal{E}$, i.e.,

$$C(\mathcal{E}, \mathbf{D}) \triangleq \max_{\mathbf{K}} \mathcal{I}(\mathbf{K}, \mathbf{D}) \quad (20)$$

where the maximum is taken over the set of all input covariance matrices \mathbf{K} satisfying the trace constraint

$$\text{tr}(\mathbf{K}) \leq \mathcal{E}. \quad (21)$$

It follows immediately from Corollary 2.1 that, if $\mathbf{D}^\dagger \mathbf{D} \succeq \tilde{\mathbf{D}}^\dagger \tilde{\mathbf{D}}$, then $I^{\text{IG}}(\mathcal{E}, \mathbf{D}) \geq I^{\text{IG}}(\mathcal{E}, \tilde{\mathbf{D}})$ and $C(\mathcal{E}, \mathbf{D}) \geq C(\mathcal{E}, \tilde{\mathbf{D}})$.

Theorem 1.1 can also be used to study the rate- R outage probability corresponding to the isotropic Gaussian input of power- \mathcal{E}

$$P_{\text{out}}^{\text{IG}}(R, \mathcal{E}, \mathbf{D}) \triangleq F\left(R; \frac{\mathcal{E}}{n} \mathbf{I}_n, \mathbf{D}\right) \quad (22)$$

and the optimal power- \mathcal{E} rate- R outage probability $P_{\text{out}}^*(R, \mathcal{E}, \mathbf{D})$, which is the smallest outage probability that can be achieved for the rate R and the average power \mathcal{E} , i.e.,

$$P_{\text{out}}^*(R, \mathcal{E}, \mathbf{D}) \triangleq \min_{\mathbf{K}} F(R; \mathbf{K}, \mathbf{D}) \quad (23)$$

where the minimum is over all positive semidefinite matrices \mathbf{K} satisfying (21). From Theorem 1.1 we now obtain that $\mathbf{D}^\dagger \mathbf{D} \succeq \tilde{\mathbf{D}}^\dagger \tilde{\mathbf{D}}$ implies that $P_{\text{out}}^{\text{IG}}(R, \mathcal{E}, \mathbf{D}) \leq P_{\text{out}}^{\text{IG}}(R, \mathcal{E}, \tilde{\mathbf{D}})$ and $P_{\text{out}}^*(R, \mathcal{E}, \mathbf{D}) \leq P_{\text{out}}^*(R, \mathcal{E}, \tilde{\mathbf{D}})$.²

Using the rotational invariance (12), we can strengthen these results by stating them in terms of the singular values of the LOS matrices. Indeed, for any unitary matrix \mathbf{V} , we have $\text{tr}(\mathbf{V}^\dagger \mathbf{K} \mathbf{V}) = \text{tr}(\mathbf{K})$, and hence it follows from (12) that for any $\mathbf{U} \in \mathcal{U}(m)$ and $\mathbf{V} \in \mathcal{U}(n)$

$$I^{\text{IG}}(\mathcal{E}, \mathbf{U} \mathbf{D} \mathbf{V}^\dagger) = I^{\text{IG}}(\mathcal{E}, \mathbf{D}) \quad (24)$$

and

$$C(\mathcal{E}, \mathbf{U} \mathbf{D} \mathbf{V}^\dagger) = C(\mathcal{E}, \mathbf{D}). \quad (25)$$

In other words, the isotropic Gaussian input information rate and channel capacity depend on the LOS matrix only via its singular values. By a similar argument, it can be verified from (11) that both the outage probability $P_{\text{out}}^{\text{IG}}(R, \mathcal{E}, \mathbf{D})$ corresponding to the isotropic Gaussian input and the optimal outage probability $P_{\text{out}}^*(R, \mathcal{E}, \mathbf{D})$ depend on the LOS matrix \mathbf{D} only via its singular values. Consequently, all these quantities are monotonic in the singular values of the LOS matrix.

Corollary 2.3: Let $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min\{m,n\}}$ and $\tilde{\sigma}_1 \geq \tilde{\sigma}_2 \geq \dots \geq \tilde{\sigma}_{\min\{m,n\}}$ be the singular values of the LOS matrices \mathbf{D} and $\tilde{\mathbf{D}}$, respectively. Suppose that $\sigma_i \geq \tilde{\sigma}_i$ for all i . Then

$$I^{\text{IG}}(\mathcal{E}, \mathbf{D}) \geq I^{\text{IG}}(\mathcal{E}, \tilde{\mathbf{D}}) \quad (26)$$

$$C(\mathcal{E}, \mathbf{D}) \geq C(\mathcal{E}, \tilde{\mathbf{D}}) \quad (27)$$

$$P_{\text{out}}^{\text{IG}}(R, \mathcal{E}, \mathbf{D}) \leq P_{\text{out}}^{\text{IG}}(R, \mathcal{E}, \tilde{\mathbf{D}}) \quad (28)$$

and

$$P_{\text{out}}^*(R, \mathcal{E}, \mathbf{D}) \leq P_{\text{out}}^*(R, \mathcal{E}, \tilde{\mathbf{D}}). \quad (29)$$

We can obtain an alternative proof (cf. [9]) of this corollary based on the observation that, if the LOS matrix \mathbf{D} is diagonal, the capacity achieving covariance matrix \mathbf{K} is also diagonal. (See also [10].) Since this structural theorem on the capacity achieving input distribution is of independent interest, we restate it here.

²Note that from the definition of power- \mathcal{E} ϵ -outage capacity

$$C_{\text{out}}^*(\epsilon, \mathcal{E}, \mathbf{D}) \triangleq \sup \{R : P_{\text{out}}^*(R, \mathcal{E}, \mathbf{D}) < \epsilon\}$$

we immediately get the monotonicity

$$C_{\text{out}}^*(\epsilon, \mathcal{E}, \mathbf{D}) \geq C_{\text{out}}^*(\epsilon, \mathcal{E}, \tilde{\mathbf{D}}), \quad \text{if } \mathbf{D}^\dagger \mathbf{D} \succeq \tilde{\mathbf{D}}^\dagger \tilde{\mathbf{D}}.$$

A similar monotonicity holds for

$$C_{\text{out}}^{\text{IG}}(\epsilon, \mathcal{E}, \mathbf{D}) \triangleq \sup \left\{R : P_{\text{out}}^{\text{IG}}(R, \mathcal{E}, \mathbf{D}) < \epsilon\right\}.$$

Theorem 2.4: Let $\mathbf{D}^\dagger \mathbf{D}$ have the eigenvalue decomposition $\mathbf{D}^\dagger \mathbf{D} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^\dagger$ for some unitary matrix \mathbf{V} and diagonal matrix $\mathbf{\Lambda}$. Then the capacity achieving covariance matrix \mathbf{K}_* is given by

$$\mathbf{K}_* = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^\dagger \quad (30)$$

for some diagonal matrix $\mathbf{\Lambda}$.

Proof: We show that if \mathbf{D} is diagonal, the capacity achieving input covariance matrix \mathbf{K}_* is diagonal. The general case follows from (12) and (20).

Fix some $1 \leq j \leq n$. Let $\mathbf{V} \in \mathcal{U}(n)$ be a diagonal matrix with all diagonal entries equal to 1 except the j th entry, which is -1 . Similarly, let $\mathbf{U} \in \mathcal{U}(m)$ be diagonal with all diagonal entries equal to 1 except for the j th entry being -1 . (In case $j > m$, $\mathbf{U} = \mathbf{I}_m$.) Since \mathbf{D} is diagonal, we have

$$\mathbf{U} \mathbf{D} \mathbf{V}^\dagger = \mathbf{D}. \quad (31)$$

Let $\tilde{\mathbf{K}} = \mathbf{V}^\dagger \mathbf{K} \mathbf{V}$. From (31) and the rotational invariance (12), we have

$$\mathcal{I}(\tilde{\mathbf{K}}, \mathbf{D}) = \mathcal{I}(\mathbf{V}^\dagger \mathbf{K} \mathbf{V}, \mathbf{D}) \quad (32)$$

$$= \mathcal{I}(\mathbf{K}, \mathbf{U} \mathbf{D} \mathbf{V}^\dagger) \quad (33)$$

$$= \mathcal{I}(\mathbf{K}, \mathbf{D}). \quad (34)$$

Now consider the matrix $\hat{\mathbf{K}} = \frac{1}{2}(\mathbf{K} + \tilde{\mathbf{K}})$. We note that the entries of $\hat{\mathbf{K}}$ are identical to those of \mathbf{K} except that its off-diagonal elements in the j th row and in the j th column are zero. In particular, $\text{tr}(\hat{\mathbf{K}}) = \text{tr}(\mathbf{K})$. On the other hand, it follows from (34) and the strict concavity of $\mathcal{I}(\mathbf{K}, \mathbf{D})$ in \mathbf{K} that

$$\mathcal{I}(\hat{\mathbf{K}}, \mathbf{D}) \geq \frac{1}{2}(\mathcal{I}(\mathbf{K}, \mathbf{D}) + \mathcal{I}(\tilde{\mathbf{K}}, \mathbf{D})) \quad (35)$$

$$= \mathcal{I}(\mathbf{K}, \mathbf{D}) \quad (36)$$

with equality if, and only if, $\mathbf{K} = \hat{\mathbf{K}}$. Repeating this procedure for each $j = 1, \dots, n-1$ shows that an optimal covariance matrix must be diagonal. \square

B. The Rician Multiple-Access Fading Channel

The coherent MIMO Rician MAC with k senders is modeled as follows. The channel output consists of k independent random matrices $\mathbf{H}_1, \dots, \mathbf{H}_k$, where \mathbf{H}_i is a random $m \times n_i$ matrix whose components are i.i.d. $\mathcal{N}_{\mathbb{C}}(0, 1)$, and of a random vector $\mathbf{Y} \in \mathbb{C}^m$ of the form

$$\mathbf{Y} = \sum_{i=1}^k (\mathbf{H}_i + \mathbf{D}_i) \mathbf{x}_i + \mathbf{Z} \quad (37)$$

where $\mathbf{x}_i \in \mathbb{C}^{n_i}$ is the i th transmitter's input vector, \mathbf{D}_i is a deterministic $m \times n_i$ complex matrix corresponding to the LOS matrix of the i th transmitter, and $\mathbf{Z} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sigma^2 \mathbf{I}_m)$ corresponds to the additive noise vector. We assume that all fading matrices $\{\mathbf{H}_i\}_{i=1}^k$ are independent of \mathbf{Z} and that the joint distribution of $(\mathbf{H}_1, \dots, \mathbf{H}_k, \mathbf{Z})$ does not depend on the inputs $\{\mathbf{x}_i\}_{i=1}^k$. We will also assume $\sigma^2 = 1$ without loss of generality.

As in the single-user scenario, it can be shown [11], [12] that Gaussian inputs achieve the capacity region of the MIMO Rician

MAC. The rate region $\mathcal{R}(\mathbf{K}_1, \dots, \mathbf{K}_k; \mathbf{D}_1, \dots, \mathbf{D}_k)$ achieved by independent Gaussian inputs $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{K}_i)$, $i = 1, \dots, k$, over the MIMO Rician MAC with LOS matrices $\{\mathbf{D}_i\}_{i=1}^k$ is the set of all rate vectors (R_1, \dots, R_k) satisfying

$$\sum_{i \in \mathcal{S}} R_i \leq \mathbb{E} \left[\log \det \left(\mathbf{I}_m + \sum_{i \in \mathcal{S}} (\mathbf{H}_i + \mathbf{D}_i) \mathbf{K}_i (\mathbf{H}_i + \mathbf{D}_i)^\dagger \right) \right] \quad (38)$$

for all $\mathcal{S} \subseteq \{1, \dots, k\}$. The capacity region of the MIMO Rician MAC, denoted as an explicit function of the input power constraints on the different users and of their corresponding LOS matrices, can be written as

$$\mathcal{C}(\mathcal{E}_1, \dots, \mathcal{E}_k; \mathbf{D}_1, \dots, \mathbf{D}_k) = \bigcup_{\{\mathbf{K}_i\}_{i=1}^k} \mathcal{R}(\mathbf{K}_1, \dots, \mathbf{K}_k; \mathbf{D}_1, \dots, \mathbf{D}_k) \quad (39)$$

where the union is over all input covariance matrices $\{\mathbf{K}_i\}_{i=1}^k$ that satisfy the trace constraints $\text{tr}(\mathbf{K}_i) \leq \mathcal{E}_i$, $i = 1, \dots, k$.

For each set $\mathcal{S} \subseteq \{1, \dots, k\}$ of elements $1 \leq i_1 < i_2 < \dots < i_s \leq k$, define the block matrices

$$\mathbf{D}_{\mathcal{S}} \triangleq [\mathbf{D}_{i_1}, \dots, \mathbf{D}_{i_s}] \quad (40)$$

$$\mathbf{H}_{\mathcal{S}} \triangleq [\mathbf{H}_{i_1}, \dots, \mathbf{H}_{i_s}] \quad (41)$$

and

$$\mathbf{K}_{\mathcal{S}} \triangleq \text{diag}(\mathbf{K}_{i_1}, \dots, \mathbf{K}_{i_s}). \quad (42)$$

Further define $\mathbf{D} = [\mathbf{D}_1, \dots, \mathbf{D}_k]$. Under this simplified notation, the rate region (38) can be expressed as

$$\sum_{i \in \mathcal{S}} R_i \leq \mathbb{E} \left[\log \det \left(\mathbf{I}_m + (\mathbf{H}_{\mathcal{S}} + \mathbf{D}_{\mathcal{S}}) \mathbf{K}_{\mathcal{S}} (\mathbf{H}_{\mathcal{S}} + \mathbf{D}_{\mathcal{S}})^\dagger \right) \right] \quad (43)$$

$$= \mathcal{I}(\mathbf{K}_{\mathcal{S}}, \mathbf{D}_{\mathcal{S}}). \quad (44)$$

Since the condition $\mathbf{D}^\dagger \mathbf{D} \succeq \tilde{\mathbf{D}}^\dagger \tilde{\mathbf{D}}$ implies that $\mathbf{D}_{\mathcal{S}}^\dagger \mathbf{D}_{\mathcal{S}} \succeq \tilde{\mathbf{D}}_{\mathcal{S}}^\dagger \tilde{\mathbf{D}}_{\mathcal{S}}$ for all $\mathcal{S} \subseteq \{1, \dots, k\}$, it follows from Corollary 2.1 that it also implies

$$\mathcal{R}(\mathbf{K}_1, \dots, \mathbf{K}_k; \mathbf{D}_1, \dots, \mathbf{D}_k) \supseteq \mathcal{R}(\mathbf{K}_1, \dots, \mathbf{K}_k; \tilde{\mathbf{D}}_1, \dots, \tilde{\mathbf{D}}_k) \quad (45)$$

and consequently, from (39),

$$\mathcal{C}(\mathcal{E}_1, \dots, \mathcal{E}_k; \mathbf{D}_1, \dots, \mathbf{D}_k) \supseteq \mathcal{C}(\mathcal{E}_1, \dots, \mathcal{E}_k; \tilde{\mathbf{D}}_1, \dots, \tilde{\mathbf{D}}_k). \quad (46)$$

We can strengthen this result using the symmetry of the problem as in the single-user case. The utility of the rotational invariance (12) is, however, rather limited since the LOS matrices cannot be assumed to be jointly diagonalizable. Thus, the monotonicity cannot be simply stated in terms of the singular values of LOS matrices. Instead, we have the following.

Corollary 2.5: Let $\mathbf{D} = [\mathbf{D}_1, \dots, \mathbf{D}_k]$ and $\tilde{\mathbf{D}} = [\tilde{\mathbf{D}}_1, \dots, \tilde{\mathbf{D}}_k]$ be LOS matrices such that

$$[\mathbf{D}_1 \mathbf{U}_1, \dots, \mathbf{D}_k \mathbf{U}_k]^\dagger [\mathbf{D}_1 \mathbf{U}_1, \dots, \mathbf{D}_k \mathbf{U}_k] \succeq \tilde{\mathbf{D}}^\dagger \tilde{\mathbf{D}} \quad (47)$$

for some $\mathbf{U}_i \in \mathcal{U}(n_i)$, $i = 1, \dots, k$. Then

$$\mathcal{C}(\mathcal{E}_1, \dots, \mathcal{E}_k; \mathbf{D}_1, \dots, \mathbf{D}_k) \supseteq \mathcal{C}(\mathcal{E}_1, \dots, \mathcal{E}_k; \tilde{\mathbf{D}}_1, \dots, \tilde{\mathbf{D}}_k). \quad (48)$$

III. PROOF OF THEOREM 1.1

Given any $\mathbf{K} \in \mathcal{H}^+(n)$ and $\mathbf{D}, \tilde{\mathbf{D}} \in \mathbb{C}^{m \times n}$ satisfying

$$\mathbf{D}^\dagger \mathbf{D} \succeq \tilde{\mathbf{D}}^\dagger \tilde{\mathbf{D}} \quad (49)$$

we wish to show that for all $t \geq 0$,

$$F(t; \mathbf{K}, \mathbf{D}) \leq F(t; \mathbf{K}, \tilde{\mathbf{D}}) \quad (50)$$

where

$$F(t; \mathbf{K}, \mathbf{D}) = \Pr \left[\log \det \left(\mathbf{I}_m + (\mathbf{H} + \mathbf{D}) \mathbf{K} (\mathbf{H} + \mathbf{D})^\dagger \right) \leq t \right]. \quad (51)$$

Without loss of generality, we can assume that the matrices \mathbf{D} and $\tilde{\mathbf{D}}$ satisfy

$$\tilde{\mathbf{D}} = \Phi \mathbf{D}, \quad \Phi = \text{diag}(\alpha, 1, \dots, 1) \quad (52)$$

for some $0 \leq \alpha \leq 1$. We justify this reduction as follows. Suppose that the desired inequality (50) holds under the condition (52). Then for any $m \times m$ permutation matrix \mathbf{P} and any Φ as in (52) we have

$$F(t; \mathbf{K}, \mathbf{P} \Phi \mathbf{P}^\dagger \mathbf{D}) = F(t; \mathbf{K}, \Phi \mathbf{P}^\dagger \mathbf{D}) \quad (53)$$

$$\geq F(t; \mathbf{K}, \mathbf{P}^\dagger \mathbf{D}) \quad (54)$$

$$= F(t; \mathbf{K}, \mathbf{D}) \quad (55)$$

where the first equality follows from (11) (with the substitution $\mathbf{U} = \mathbf{P}$ and $\mathbf{V} = \mathbf{I}_n$) because any permutation matrix is unitary; the subsequent inequality from our assumption that we have already proved the theorem for Φ as in (52); and the last equality again from (11) (with the substitution $\mathbf{U} = \mathbf{P}^\dagger$ and $\mathbf{V} = \mathbf{I}_n$). By choosing \mathbf{P} so that $\mathbf{P} \Phi \mathbf{P}^\dagger = \text{diag}(1, \dots, 1, \alpha, 1, \dots, 1)$ it follows from (55) that if the desired inequality holds for all Φ as in (52) then it must also hold for all Φ of the form $\text{diag}(1, \dots, 1, \alpha, 1, \dots, 1)$. Writing $\text{diag}(\alpha_1, \alpha_2, \dots, \alpha_m) = \Phi_1 \cdot \Phi_2 \cdots \Phi_m$ where $\Phi_i = \text{diag}(1, \dots, 1, \alpha_i, 1, \dots, 1)$ we have by the associativity of matrix multiplication and the theorem for the case where Φ is of the form $\text{diag}(1, \dots, 1, \alpha, 1, \dots, 1)$

$$F\left(t; \mathbf{K}, \prod_{i=1}^m \Phi_i \mathbf{D}\right) = F\left(t; \mathbf{K}, \Phi_1 \prod_{i=2}^m \Phi_i \mathbf{D}\right) \quad (56)$$

$$\geq F\left(t; \mathbf{K}, \prod_{i=2}^m \Phi_i \mathbf{D}\right) \quad (57)$$

\vdots

$$\geq F(t; \mathbf{K}, \mathbf{D}) \quad (58)$$

thus establishing the result for Φ any arbitrary diagonal contraction matrix with diagonal elements $0 \leq \alpha_i \leq 1$, $i = 1, \dots, m$. Now applying the rotational invariance (11) once again to arbitrary unitary matrices $\mathbf{U} \in \mathcal{U}(m)$, $\mathbf{V} \in \mathcal{U}(m)$ and nonnegative diagonal contraction matrix Φ , we obtain

$$F(t; \mathbf{K}, \mathbf{U} \Phi \mathbf{V}^\dagger \mathbf{D}) = F(t; \mathbf{K}, \Phi \mathbf{V}^\dagger \mathbf{D}) \quad (59)$$

$$\geq F(t; \mathbf{K}, \mathbf{V}^\dagger \mathbf{D}) \quad (60)$$

$$= F(t; \mathbf{K}, \mathbf{D}). \quad (61)$$

Thus, if the desired inequality (50) holds for all $\mathbf{D}, \tilde{\mathbf{D}}$, and Φ satisfying (52) then it must also hold whenever

$$\tilde{\mathbf{D}} = \Phi \mathbf{D}, \quad \Phi^\dagger \Phi \preceq \mathbf{I}_m. \quad (62)$$

But (62) is equivalent to the original condition (49) (see, for example, [13]). Therefore, in order to prove the theorem, it suffices to establish the inequality (50) under the simplified condition (52).

For the rest of our discussion, we need the following result by T. W. Anderson [14], [15, Theorem 8.10.5].

Lemma 3.1: (Anderson's Theorem): Let \mathcal{H} be a convex set in \mathbb{C}^n , symmetric about the origin (i.e., $\boldsymbol{\xi} \in \mathcal{H}$ implies $-\boldsymbol{\xi} \in \mathcal{H}$). Let $f(\boldsymbol{\xi}) \geq 0$ be a function on \mathbb{C}^n such that (i) $f(-\boldsymbol{\xi}) = f(\boldsymbol{\xi})$ for all $\boldsymbol{\xi}$, (ii) the set $\{\boldsymbol{\xi} \in \mathbb{C}^n : f(\boldsymbol{\xi}) \geq u\}$ is convex for every $u > 0$; and (iii) $\int_{\mathcal{H}} f(\boldsymbol{\xi}) d\boldsymbol{\xi} < \infty$. Then

$$\int_{\mathcal{H}} f(\boldsymbol{\xi} + \alpha \boldsymbol{\eta}) d\boldsymbol{\xi} \geq \int_{\mathcal{H}} f(\boldsymbol{\xi} + \boldsymbol{\eta}) d\boldsymbol{\xi} \quad (63)$$

for every vector $\boldsymbol{\eta} \in \mathbb{C}^n$ and $0 \leq \alpha \leq 1$.

The proof of this celebrated result is based on the Brunn–Minkowski inequality [16]. An interested reader is referred to a nice review by Perlman [17] for further generalizations and applications in multivariate statistics.

Returning to our problem, for any $t \geq 0$, we define a set of matrices

$$\mathcal{G}_t = \left\{ \mathbf{G} \in \mathbb{C}^{m \times n} : \log \det \left(\mathbf{I}_m + \mathbf{G} \mathbf{G}^\dagger \right) \leq t \right\}. \quad (64)$$

For any fixed vectors $\mathbf{g}_2, \dots, \mathbf{g}_m \in \mathbb{C}^n$, let

$$\mathcal{H}_t(\mathbf{g}_2, \dots, \mathbf{g}_m) = \left\{ \boldsymbol{\xi} \in \mathbb{C}^n : [\boldsymbol{\xi}, \mathbf{g}_2, \dots, \mathbf{g}_m]^\top \in \mathcal{G}_t \right\}. \quad (65)$$

In other words, $\mathcal{H}_t(\mathbf{g}_2, \dots, \mathbf{g}_m)$ is the set of the first rows $\boldsymbol{\xi}^\top$ that belong to \mathcal{G}_t with given values of other rows $\mathbf{g}_2^\top, \dots, \mathbf{g}_m^\top$. As will be checked later at the end of this section, for any $\mathbf{g}_2, \dots, \mathbf{g}_m$, the set $\mathcal{H}_t(\mathbf{g}_2, \dots, \mathbf{g}_m)$ is convex and symmetric about the origin.

The rest of the proof proceeds along lines similar to those of Das Gupta, Anderson, and Mudholkar [18]. We represent \mathbb{H} as $[\mathbf{H}_1, \dots, \mathbf{H}_m]^\top$, where \mathbf{H}_j^\top is the j th row of \mathbb{H} . Similarly, let \mathbf{d}_j^\top denote the j th row of \mathbf{D} . Let $f(\boldsymbol{\xi} | \mathbf{h}_2, \dots, \mathbf{h}_m)$ be the conditional density of \mathbf{H}_1 conditioned on $\mathbf{H}_j = \mathbf{h}_j, j = 2, \dots, m$. Since the rows of \mathbb{H} are mutually independent, $f(\boldsymbol{\xi} | \mathbf{h}_2, \dots, \mathbf{h}_m) = f(\boldsymbol{\xi})$ is the density of the multivariate Gaussian $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{I}_n)$, which satisfies the conditions (i) to (iii) of Anderson's Theorem. Combining the conditions on f and \mathcal{H}_t with the standing assumption (52), we can invoke Anderson's Theorem for the first row of \mathbb{H} after conditioning on the other rows $\mathbf{H}_2^\top, \dots, \mathbf{H}_m^\top$ as shown in (66)–(68) at the bottom of the page. By taking the expectation on both sides of (68) with respect to the joint density of $\mathbf{H}_2, \dots, \mathbf{H}_m$, we establish the desired inequality (50).

It remains to check the convexity and symmetry of the set $\mathcal{H}_t = \mathcal{H}_t(\mathbf{g}_2, \dots, \mathbf{g}_m)$. Let $\mathbf{G} = [\boldsymbol{\xi}, \mathbf{g}_2, \dots, \mathbf{g}_m]^\top$. We show that $\det(\mathbf{I}_m + \mathbf{G} \mathbf{G}^\dagger)$ is convex and symmetric in $\boldsymbol{\xi}$, which clearly implies the convexity and symmetry of \mathcal{H}_t . For the symmetry, observe that

$$\det(\mathbf{I}_m + \mathbf{G} \mathbf{G}^\dagger) = \det(\mathbf{I}_m + \mathbf{U} \mathbf{G} \mathbf{G}^\dagger \mathbf{U}^\dagger) \quad (69)$$

for any unitary matrix \mathbf{U} , and in particular for $\mathbf{U} = \text{diag}(-1, 1, \dots, 1)$.

For the convexity, let $\mathbf{F} = \mathbf{G} \mathbf{K}^{\frac{1}{2}}$ where $\mathbf{K}^{\frac{1}{2}}$ is any matrix satisfying $\mathbf{K}^{\frac{1}{2}} (\mathbf{K}^{\frac{1}{2}})^\dagger = \mathbf{K}$. Recall the identity

$$\det(\mathbf{I}_k + \mathbf{A} \mathbf{B}) = \det(\mathbf{I}_j + \mathbf{B} \mathbf{A}) \quad (70)$$

for any $\mathbf{A} \in \mathbb{C}^{k \times j}, \mathbf{B} \in \mathbb{C}^{j \times k}$. Then we have

$$\det(\mathbf{I}_m + \mathbf{G} \mathbf{K} \mathbf{G}^\dagger) = \det(\mathbf{I}_m + \mathbf{F} \mathbf{F}^\dagger) \quad (71)$$

$$= \det(\mathbf{I}_n + \mathbf{F}^\dagger \mathbf{F}) \quad (72)$$

$$= \det\left(\mathbf{I}_n + \sum_{j=2}^m (\mathbf{f}_j^\top)^\dagger \mathbf{f}_j^\top + (\mathbf{f}_1^\top)^\dagger \mathbf{f}_1^\top\right) \quad (73)$$

$$= \det\left(\mathbf{M} + (\mathbf{f}_1^\top)^\dagger \mathbf{f}_1^\top\right) \quad (74)$$

$$= \det(\mathbf{M}) \det\left(\mathbf{I}_n + \mathbf{M}^{-1} (\mathbf{f}_1^\top)^\dagger \mathbf{f}_1^\top\right) \quad (75)$$

$$= \det(\mathbf{M}) \left(1 + \mathbf{f}_1^\top \mathbf{M}^{-1} (\mathbf{f}_1^\top)^\dagger\right) \quad (76)$$

$$= \det(\mathbf{M}) \left(1 + \boldsymbol{\xi}^\top \mathbf{K}^{\frac{1}{2}} \mathbf{M}^{-1} \mathbf{K}^{\frac{1}{2}} (\boldsymbol{\xi}^\top)^\dagger\right) \quad (77)$$

where \mathbf{f}_j^\top denotes the j th row of \mathbf{F} and the positive definite matrix \mathbf{M} is defined as

$$\mathbf{M} = \mathbf{I}_n + \sum_{j=2}^m (\mathbf{f}_j^\top)^\dagger \mathbf{f}_j^\top.$$

The last line of (77) is a positive semidefinite quadratic form in $\boldsymbol{\xi}$, and hence, it is convex.

IV. CONCLUDING REMARKS

In this correspondence we have found a natural ordering of MIMO Rician channels via their LOS matrices. We have shown that for two LOS matrices $\mathbf{D}, \tilde{\mathbf{D}} \in \mathbb{C}^{m \times n}$

$$\mathbf{D}^\dagger \mathbf{D} \succeq \tilde{\mathbf{D}}^\dagger \tilde{\mathbf{D}} \iff \left(\mathcal{I}(\mathbf{K}, \mathbf{D}) \geq \mathcal{I}(\mathbf{K}, \tilde{\mathbf{D}}) \quad \forall \mathbf{K} \in \mathcal{H}^+(n) \right) \quad (78)$$

where $\mathcal{I}(\mathbf{K}, \mathbf{D}) = I(\mathbf{X}; \mathbf{Y} | \mathbb{H})$ is the mutual information induced by a $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{K})$ input over a coherent MIMO Rician channel with LOS matrix \mathbf{D} . From this result, we obtained monotonicity results for the isotropic Gaussian input information rate and for channel capacity, not only for the single-user channel but also for the multiple-access channel (MAC).

In some sense the results of this correspondence may not be surprising because the relation $\mathbf{D}^\dagger \mathbf{D} \succeq \tilde{\mathbf{D}}^\dagger \tilde{\mathbf{D}}$ implies $\text{tr}(\mathbf{D}^\dagger \mathbf{D}) \geq \text{tr}(\tilde{\mathbf{D}}^\dagger \tilde{\mathbf{D}})$ and hence a “larger” LOS matrix results in a larger output power. Note, however, that some care must be exercised because in MIMO communications a larger output power need not imply a larger capacity. For instance, if

$$\mathbf{D}_1 = \begin{pmatrix} 10 & 10 \\ 10 & 0 \end{pmatrix}, \quad \mathbf{D}_2 = \begin{pmatrix} 10 & 10 \\ 10 & 10 \end{pmatrix} \quad (79)$$

$$\begin{aligned} & \Pr \left[\log \det \left(\mathbf{I}_m + (\mathbf{H} + \mathbf{D}) \mathbf{K} (\mathbf{H} + \mathbf{D})^\dagger \right) \leq t \mid \mathbf{H}_i = \mathbf{h}_i, i = 2, \dots, m \right] \\ &= \int_{\mathcal{H}_t(\mathbf{h}_2 + \mathbf{d}_2, \dots, \mathbf{h}_m + \mathbf{d}_m)} f(\boldsymbol{\xi} - \mathbf{d}_1) d\boldsymbol{\xi} \quad (66) \\ &\leq \int_{\mathcal{H}_t(\mathbf{h}_2 + \mathbf{d}_2, \dots, \mathbf{h}_m + \mathbf{d}_m)} f(\boldsymbol{\xi} - \alpha \mathbf{d}_1) d\boldsymbol{\xi} \quad (67) \\ &= \Pr \left[\log \det \left(\mathbf{I}_m + (\mathbf{H} + \tilde{\mathbf{D}}) \mathbf{K} (\mathbf{H} + \tilde{\mathbf{D}})^\dagger \right) \leq t \mid \mathbf{H}_i = \mathbf{h}_i, i = 2, \dots, m \right]. \quad (68) \end{aligned}$$

then although the output power increases while changing from D_1 to D_2 , one can numerically show that the isotropic Gaussian input information rate and channel capacity are larger for the channel with LOS matrix D_1 . The intuition is that D_1 has full rank with singular values 16.18 and 6.18, whereas D_2 is rank-deficient with singular values 20 and 0, thus providing only one LOS eigenmode.

APPENDIX

Instead of proving Proposition 2.2 directly, we will prove the equivalent statement

$$D^\dagger D \not\leq \tilde{D}^\dagger \tilde{D} \Rightarrow \mathcal{I}(K, D) < \mathcal{I}(K, \tilde{D}), \quad \text{for some } K \in \mathcal{H}^+(n). \quad (80)$$

We first note that $D^\dagger D \not\leq \tilde{D}^\dagger \tilde{D}$ means that there exists a vector $\mathbf{a} \in \mathbb{C}^n$ such that

$$\mathbf{a}^\dagger D^\dagger D \mathbf{a} < \mathbf{a}^\dagger \tilde{D}^\dagger \tilde{D} \mathbf{a}. \quad (81)$$

For such a vector \mathbf{a} , let $K_0 = \mathbf{a}\mathbf{a}^\dagger \in \mathcal{H}^+(n)$. We will show that the strict inequality $\mathcal{I}(K_0, D) < \mathcal{I}(K_0, \tilde{D})$ holds.

By (5) it suffices to show that $F(t; K_0, D) > F(t; K_0, \tilde{D})$ for all $t > 0$. Define $\mathbf{G} = \mathbf{H}\mathbf{a}$, $\mathbf{b} = D\mathbf{a}$, and $\tilde{\mathbf{b}} = \tilde{D}\mathbf{a}$. Then we have for any $t > 0$

$$F(t; K_0, D) = \Pr \left[\log \det \left(\mathbf{I}_m + (\mathbf{H} + D)\mathbf{a}\mathbf{a}^\dagger(\mathbf{H} + D)^\dagger \right) \leq t \right] \quad (82)$$

$$= \Pr \left[\log \left(1 + \mathbf{a}^\dagger (\mathbf{H} + D)^\dagger (\mathbf{H} + D) \mathbf{a} \right) \leq t \right] \quad (83)$$

$$= \Pr \left[\log \left(1 + (\mathbf{G} + \mathbf{b})^\dagger (\mathbf{G} + \mathbf{b}) \right) \leq t \right] \quad (84)$$

$$> \Pr \left[\log \left(1 + (\mathbf{G} + \tilde{\mathbf{b}})^\dagger (\mathbf{G} + \tilde{\mathbf{b}}) \right) \leq t \right] \quad (85)$$

$$= \Pr \left[\log \left(1 + \mathbf{a}^\dagger (\mathbf{H} + \tilde{D})^\dagger (\mathbf{H} + \tilde{D}) \mathbf{a} \right) \leq t \right] \quad (86)$$

$$= \Pr \left[\log \det \left(\mathbf{I}_m + (\mathbf{H} + \tilde{D})\mathbf{a}\mathbf{a}^\dagger(\mathbf{H} + \tilde{D})^\dagger \right) \leq t \right] \quad (87)$$

$$= F(t; K_0, \tilde{D}) \quad (88)$$

where (83) follows from (70), and (85) follows from the strict monotonicity result for the single-antenna case [1, Lemma 6.2 (b)]. Indeed, \mathbf{G} is $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{a}^\dagger \mathbf{a} \mathbf{I}_m)$ distributed and $(\mathbf{G} + \mathbf{b})^\dagger (\mathbf{G} + \mathbf{b})$ thus has a scaled noncentral chi-square distribution with (scaled) noncentrality parameter $\mathbf{b}^\dagger \mathbf{b}$. Now $(\mathbf{G} + \tilde{\mathbf{b}})^\dagger (\mathbf{G} + \tilde{\mathbf{b}})$ in (85) is also a scaled noncentral chi-square random variable, which, from (81), has a strictly larger noncentrality parameter $\tilde{\mathbf{b}}^\dagger \tilde{\mathbf{b}} > \mathbf{b}^\dagger \mathbf{b}$. Hence, $(\mathbf{G} + \tilde{\mathbf{b}})^\dagger (\mathbf{G} + \tilde{\mathbf{b}})$ is stochastically strictly larger than $(\mathbf{G} + \mathbf{b})^\dagger (\mathbf{G} + \mathbf{b})$, so that the strict inequality in (85) is justified for any $t > 0$.

REFERENCES

[1] A. Lapidoth and S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2426–2467, Oct. 2003.
 [2] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, 2nd ed. New York: Wiley, 1995, vol. 2.
 [3] W. Parker, "The characteristic roots of matrices," *Duke Math. J.*, vol. 12, no. 3, pp. 519–526, 1945.
 [4] B. Vinograd, "Canonical positive definite matrices under internal linear transformations," *Proc. Amer. Math. Soc.*, vol. 1, no. 2, pp. 159–161, Apr. 1950.
 [5] L. Cottatellucci and M. Debbah, "The effect of line of sight on the asymptotic capacity of MIMO systems," in *Proc. IEEE Int. Symp. Information Theory*, Chicago, IL, Jun./Jul. 2004, p. 241.

[6] S. K. Jayaweera and H. V. Poor, "On the capacity of multiple-antenna systems in Rician fading," *IEEE Trans. Wireless Commun.*, vol. 4, no. 3, pp. 1102–1111, May 2005.
 [7] G. Lebrun, M. Faulkner, M. Shafi, and P. J. Smith, "MIMO Ricean channel capacity," in *Proc. Int. Conf. Communications*, vol. 5, 2004, pp. 2939–2943.
 [8] P. Driessen and G. Foschini, "On the capacity formula for multiple input-multiple output wireless channels: A geometric interpretation," *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 173–176, Feb. 1999.
 [9] D. Hoesli and A. Lapidoth, "The capacity of a MIMO Ricean channel is monotonic in the singular values of the mean," in *Proc. 5th Int. ITG Conf. SCC*, Erlangen, Germany, Jan. 2004, pp. 381–385.
 [10] S. Venkatesan, S. H. Simon, and R. A. Valenzuela, "Capacity of a Gaussian MIMO channel with nonzero mean," in *Proc. IEEE Semiann. Vehicular Technology Conf.*, Orlando, FL, Oct. 2003, pp. 1767–1771.
 [11] A. D. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1713–1727, Nov. 1994.
 [12] S. Shamai (Shitz) and A. D. Wyner, "Information-theoretic considerations for symmetric, cellular, multiple-access fading channels—Part I," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1877–1894, Nov. 1997.
 [13] M. L. Eaton, "On group induced orderings, monotone functions, and convolution theorems," in *Inequalities in Statistics and Probability*. ser. Lecture Notes—Monograph Series, Y. L. Tong, Ed. Hayward, CA: Inst. Math. Statistics, 1984, vol. 5, pp. 13–25.
 [14] T. W. Anderson, "The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities," *Proc. Amer. Math. Soc.*, vol. 6, no. 2, pp. 170–176, 1955.
 [15] ———, *An Introduction to Multivariate Statistical Analysis*, 3rd ed. New York: Wiley, 2003.
 [16] R. Gardner, "The Brunn-Minkowski inequality," *Bull. Amer. Math. Soc.*, vol. 39, no. 3, pp. 355–405, 2002.
 [17] M. D. Perlman, "T. W. Anderson's theorem on the integral of a symmetric unimodal function over a symmetric convex set and its applications in probability and statistics," in *The Collected Papers of T. W. Anderson*, G. P. H. Styan, Ed. New York: Wiley, 1990, vol. 2.
 [18] S. Das Gupta, T. W. Anderson, and G. S. Mudholkar, "Monotonicity of the power functions of some tests of the multivariate linear hypothesis," *Ann. Math. Statist.*, vol. 35, no. 1, pp. 200–205, Mar. 1964.