How Good is a Uniform Power Allocation on a MIMO Ricean Channel?

Daniel Hösli and Amos Lapidoth
Signal and Information Processing Laboratory
Swiss Federal Institute of Technology (ETH) Zurich
e-mail: {hoesli, lapidoth}@isi.ee.ethz.ch

Abstract

We consider a MIMO Ricean fading channel with perfect side information at the receiver. We derive an analytic upper bound on the difference between the capacity of this channel and the mutual information that is induced by an isotropic circularly-symmetric Gaussian input. This bound is based on a dual expression for mutual information. If the number of receiver antennas m is at least equal to the number of transmitter antennas n, i.e., $m \geq n$, this bound tends to zero as the signal-to-noise ratio tends to infinity. This shows that for this case a uniform power allocation is asymptotically optimal. If $m < n$ such a uniform power allocation need not be asymptotically optimal.

1 Introduction

We consider a discrete-time memoryless multiple-input multiple-output (MIMO) channel whose output $\mathbf{Y}$ takes value in the m-dimensional complex Euclidean space $\mathbb{C}^m$ and is given by

$$\mathbf{Y} = \mathbb{H}\mathbf{x} + \mathbf{Z} \tag{1}$$

where $\mathbf{x} \in \mathbb{C}^n$ is the channel input; the random vector $\mathbf{Z}$ has a $\mathcal{N}_{\mathbb{C}}\big(\mathbf{0}, \sigma^2 \mathsf{I}_m\big)$ distribution; and the random matrix $\mathbb{H} \in \mathbb{C}^{m\times n}$ can be written as

$$\mathbb{H} = \widetilde{\mathbb{H}} + \mathsf{D} \tag{2}$$

where $\mathsf{D} \in \mathbb{C}^{m\times n}$ is a deterministic $m \times n$ matrix and where the $m \cdot n$ random components of the random matrix $\widetilde{\mathbb{H}} \in \mathbb{C}^{m\times n}$ are IID $\mathcal{N}_{\mathbb{C}}(0, 1)$. It is assumed that $\widetilde{\mathbb{H}}$ and $\mathbf{Z}$ are independent, and that their joint law does not depend on the input $\mathbf{x}$. Hereafter we refer to $\widetilde{\mathbb{H}}$ and D as the channel state matrix and the line-of-sight matrix, respectively.

Here $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathsf{K})$ denotes the zero-mean circularly-symmetric multivariate Gaussian distribution of covariance matrix K, and I_m denotes the $m \times m$ identity matrix.

We shall consider the capacity of this channel when the realization of the fading matrix $\mathbb{H}$ is known to the receiver, but only its probability law is known at the transmitter. We assume that the transmitted signal is subject to an average power constraint

$$\mathsf{E}\Big[\mathbf{X}^\dagger \mathbf{X}\Big] \leq \mathcal{E}_{\mathrm{s}} \tag{3}$$

where we use $\mathsf{A}^\dagger$ to denote the Hermitian conjugate of A.

The capacity C of this channel is achieved by a multivariate circularly-symmetric Gaussian input [1]. Using the fact that the capacity-achieving input meets (3) with equality and combining the input power and the noise power to a single "signal-to-noise ratio" parameter

$$\rho = \frac{\mathcal{E}_{\mathrm{s}}}{\sigma^2} \tag{4}$$

one can show that capacity can be expressed as

$$C = \sup_{\hat{\mathsf{K}}} \mathsf{E}_{\widetilde{\mathbb{H}}}\left[\log\det\left(\mathsf{I}_m + \rho(\widetilde{\mathbb{H}}+\mathsf{D})\hat{\mathsf{K}}(\widetilde{\mathbb{H}}+\mathsf{D})^\dagger\right)\right] \tag{5}$$

where the supremum is over all diagonal positive semi-definite matrices $\hat{\mathsf{K}}$ with

$$\operatorname{tr}\left(\hat{\mathsf{K}}\right) = 1. \tag{6}$$

The supremum is achieved by a unique matrix $\hat{\mathsf{K}}^*$.

Further, if D is "diagonal"[1], then $\hat{\mathsf{K}}^*$ is diagonal [1, 2]. Hereafter we will consider only "diagonal" line-of-sight matrices D and therefore diagonal covariance matrices $\hat{\mathsf{K}}$.

2 Main Results

As the optimal power allocation, i.e., $\hat{\mathsf{K}}^*$ is unknown, it is natural to ask how much we lose with respect to capacity by applying a uniform power allocation, i.e., by using an isotropic circularly-symmetric Gaussian input

$$\mathbf{X}_{\mathrm{GI}} \sim \mathcal{N}_{\mathbb{C}}\left(\mathbf{0}, \frac{\mathcal{E}_{\mathrm{s}}}{n}\mathsf{I}_n\right). \tag{7}$$

This question was also addressed in [3].

Lemma 1. *Consider the MIMO Ricean channel (1) with a "diagonal" line-of-sight matrix* D. *If the input* $\mathbf{X}$ *is* $\mathcal{N}_{\mathbb{C}}\left(\mathbf{0}, \mathcal{E}_{\mathrm{s}}\hat{\mathsf{K}}\right)$ *distributed with diagonal* $\hat{\mathsf{K}}$ *satisfying* $\operatorname{tr}\left(\hat{\mathsf{K}}\right) = 1$, *then the difference between capacity and the induced mutual information is upper bounded by*

$$C - I(\mathbf{X};\mathbf{Y},\widetilde{\mathbb{H}}) \leq \max_i \mathsf{M}_{i,i} - \operatorname{tr}\left(\hat{\mathsf{K}}\mathsf{M}\right) \tag{8}$$

where the matrix M *is defined by*

$$\mathsf{M} = \mathsf{E}_{\mathbb{H}}\left[\mathbb{H}^\dagger\left(\mathbb{H}\hat{\mathsf{K}}\mathbb{H}^\dagger + \frac{1}{\rho}\mathsf{I}_m\right)^{-1}\mathbb{H}\right]. \tag{9}$$

The bound is met with equality if and only if $\hat{\mathsf{K}}$ *achieves the maximum in (5), in which case both sides of (8) are equal to zero.*

Figure 1(a) shows numerical examples of the upper bound (8) for the case $m = 1$, $n = 2$. In this figure the normalized input covariance matrix is $\hat{\mathsf{K}} = \operatorname{diag}(1/2, 1/2)$, and the bounds are plotted for the two line-of-sight matrices $\mathsf{D} = [1, 0]$ and $\mathsf{D} = [10, 0]$. Also depicted is the upper bound (13) that can be computed analytically. It is described in Theorem 1 ahead.

In general, if $n > m$ one can show that if $\hat{\mathsf{K}}$ is non-singular, then

$$I(\mathbf{X};\mathbf{Y},\widetilde{\mathbb{H}}) = m\cdot\log\rho + \mathsf{E}_{\mathbb{H}}\left[\log\det\left(\mathbb{H}\hat{\mathsf{K}}\mathbb{H}^\dagger\right)\right] + o(1) \tag{10}$$

[1] By a "diagonal" $m \times n$ matrix we refer to a matrix whose (i, j) component is zero whenever $i \neq j$, for $1 \leq i \leq m$ and $1 \leq j \leq n$

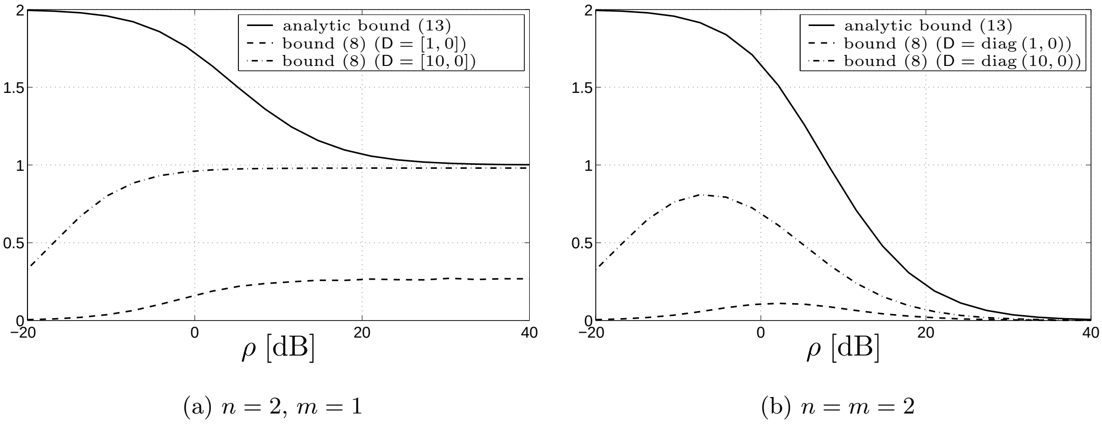

Figure 1: Examples of the upper bound (8) for $\hat{\mathsf{K}} = \operatorname{diag}(1/2, 1/2)$ and the analytic upper bound (13). (The units on the vertical axis are nats.)

where $o(1)$ tends to zero as $\rho \to \infty$. Let $\hat{\mathsf{K}}^*$ denote the matrix that achieves

$$\sup_{\hat{\mathsf{K}}} \mathsf{E}_{\mathbb{H}}\Big[\log\det\Big(\mathbb{H}\hat{\mathsf{K}}\mathbb{H}^\dagger\Big)\Big] \tag{11}$$

where the supremum is over all diagonal positive definite matrices $\hat{\mathsf{K}}$ with $\operatorname{tr}\Big(\hat{\mathsf{K}}\Big) = 1$. By numerical examples one can show that if $n > m$, then $\hat{\mathsf{K}} = \mathsf{I}_n/n$ need not fulfill the Karush-Kuhn-Tucker optimality conditions [4] for the concave optimization problem (11). This shows that if $n > m$, then with $\mathbf{X}^* \sim \mathcal{N}_{\mathbb{C}}\Big(\mathbf{0}, \mathcal{E}_{\mathrm{s}}\hat{\mathsf{K}}^*\Big)$ the difference

$$I(\mathbf{X}^*; \mathbf{Y}, \widetilde{\mathbb{H}}) - I(\mathbf{X}_{\mathrm{GI}}; \mathbf{Y}, \widetilde{\mathbb{H}}) \tag{12}$$

need not tend to zero as $\rho \to \infty$, and hence that a uniform power allocation can be asymptotically suboptimal if $n > m$.

We next turn to an upper bound that is looser than (8) but easier to compute. It does not require any Monte-Carlo simulations to compute M.

Theorem 1. *If the MIMO Ricean channel (1) has a "diagonal" line-of-sight matrix and if* $\mathbf{X}_{\mathrm{GI}} \sim \mathcal{N}_{\mathbb{C}}\big(\mathbf{0}, \frac{\mathcal{E}_{\mathrm{s}}}{n}\mathsf{I}_n\big)$ *is used as the input to the channel, then the difference between capacity and the induced mutual information is upper bounded by*

$$C - I(\mathbf{X}_{\mathrm{GI}}; \mathbf{Y}, \widetilde{\mathbb{H}}) \le n - l - \frac{nl^2}{\rho} \cdot e^{\frac{nl}{\rho}} \cdot \operatorname{Ei}\left(-\frac{nl}{\rho}\right) \tag{13}$$

where $l = \min\{m, n\}$ *and* $\operatorname{Ei}(-x) = -\int_x^\infty \frac{e^{-t}}{t}\,\mathrm{d}t,\ x > 0$.

Figure 1(b) compares the analytic upper bound (13) with the upper bound (8) for the examples $\mathsf{D} = \operatorname{diag}(1, 0)$ and $\mathsf{D} = \operatorname{diag}(10, 0)$ (i.e., $m = n = 2$). It can be seen that generally if $n \le m$, then the analytic upper bound is asymptotically tight in that both the RHS and LHS of (13) tend to zero as $\rho \to \infty$.

3 Outline of Proofs

3.1 Proof of Lemma 1

The upper bound (8) is based on a dual expression for mutual information, which was applied in [5] to MIMO fading channels. Here we apply it to the MIMO Ricean channel (1) with input $\mathbf{x} \in \mathbb{C}^n$, output $\mathbf{Y} \in \mathbb{C}^m$, state $\widetilde{\mathbb{H}} \in \mathbb{C}^{m\times n}$, "diagonal" line-of-sight matrix D, and with perfect side information at the receiver. We use this channel with the random input $\mathbf{X} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathsf{K})$ of diagonal covariance matrix K with $\operatorname{tr}(\mathsf{K}) = \mathcal{E}_{\mathrm{s}}$ and consider the pair $(\mathbf{Y}, \widetilde{\mathbb{H}})$ as a compound output of the channel. Then, using the dual expression, we have the following upper bound on the mutual information:

$$I(\mathbf{X}; \mathbf{Y}, \widetilde{\mathbb{H}}) \leq \mathsf{E}_{\mathbf{X}}[D\left(W(\cdot, \cdot|\mathbf{X}) \| R(\cdot, \cdot)\right)] \tag{14}$$

where $D(\cdot\|\cdot)$ denotes relative entropy; $W(\cdot, \cdot|\cdot)$ is the channel law; and $R(\cdot, \cdot)$ is any joint probability distribution on $(\mathbf{Y}, \widetilde{\mathbb{H}})$. Equality in (14) holds if and only if $R(\cdot, \cdot)$ is the joint distribution on the pair $(\mathbf{Y}, \widetilde{\mathbb{H}})$ induced by $\mathbf{X}$. Since $\mathbf{X}$ and $\widetilde{\mathbb{H}}$ are independent, the channel law can be written as

$$W(\mathbf{y}, \widetilde{\mathsf{H}}|\mathbf{x}) = p_{\mathbf{Y}|\widetilde{\mathbb{H}},\mathbf{X}}(\mathbf{y}|\widetilde{\mathsf{H}}, \mathbf{x}) \cdot p_{\widetilde{\mathbb{H}}}(\widetilde{\mathsf{H}}), \quad \mathbf{y} \in \mathbb{C}^m,\ \mathbf{x} \in \mathbb{C}^n,\ \widetilde{\mathsf{H}} \in \mathbb{C}^{m\times n}. \tag{15}$$

Here $p_{\mathbf{Y}|\widetilde{\mathbb{H}},\mathbf{X}}(\cdot|\widetilde{\mathsf{H}}, \mathbf{x})$ denotes the probability density of $\mathbf{Y}$ given that $\widetilde{\mathbb{H}} = \widetilde{\mathsf{H}}$ and $\mathbf{X} = \mathbf{x}$, and $p_{\widetilde{\mathbb{H}}}(\cdot)$ denotes the probability density of $\widetilde{\mathbb{H}}$.

To arrive at the upper bound (8) we choose $R(\cdot, \cdot)$ to be of the form

$$R(\tilde{\mathbf{y}}, \widetilde{\mathsf{H}}) = p_{\tilde{\mathbf{Y}}|\widetilde{\mathbb{H}}}(\tilde{\mathbf{y}}|\widetilde{\mathsf{H}}) \cdot p_{\widetilde{\mathbb{H}}}(\widetilde{\mathsf{H}}), \quad \tilde{\mathbf{y}} \in \mathbb{C}^m,\ \widetilde{\mathsf{H}} \in \mathbb{C}^{m\times n} \tag{16}$$

where $\tilde{\mathbf{Y}}$ denotes the output of the channel that is induced by a multivariate circularly-symmetric Gaussian input $\tilde{\mathbf{X}}$ of diagonal covariance matrix $\tilde{\mathsf{K}}$ with $\operatorname{tr}\left(\tilde{\mathsf{K}}\right) = \mathcal{E}_{\mathrm{s}}$, i.e.,

$$\tilde{\mathbf{X}} \sim \mathcal{N}_{\mathbb{C}}\left(\mathbf{0}, \tilde{\mathsf{K}}\right) \tag{17}$$

and $p_{\tilde{\mathbf{Y}}|\widetilde{\mathbb{H}}}(\cdot|\widetilde{\mathsf{H}})$ is the probability density of $\tilde{\mathbf{Y}}$ given that the channel state is $\widetilde{\mathbb{H}} = \widetilde{\mathsf{H}}$. Note however that any other joint distribution $R(\cdot, \cdot)$ would also lead to an upper bound (14).

Using the above setting we obtain the following upper bound on the mutual information:

$$\begin{aligned} I(\mathbf{X}; \mathbf{Y}, \widetilde{\mathbb{H}}) \leq {} & -m + \mathsf{E}_{\widetilde{\mathbb{H}}}\left[\log\det\left(\frac{1}{\sigma^2}\tilde{\mathsf{L}}\right)\right] + \operatorname{tr}\left(\mathsf{E}_{\widetilde{\mathbb{H}}}\left[\sigma^2\tilde{\mathsf{L}}^{-1}\right]\right) \\ & + \operatorname{tr}\left(\mathsf{K}\mathsf{E}_{\widetilde{\mathbb{H}}}\left[(\widetilde{\mathbb{H}} + \mathsf{D})^\dagger\tilde{\mathsf{L}}^{-1}(\widetilde{\mathbb{H}} + \mathsf{D})\right]\right) \end{aligned} \tag{18}$$

where

$$\tilde{\mathsf{L}} = (\widetilde{\mathsf{H}} + \mathsf{D})\tilde{\mathsf{K}}(\widetilde{\mathsf{H}} + \mathsf{D})^\dagger + \sigma^2\mathsf{I}_m \tag{19}$$

and where equality holds if and only if $\mathbf{X} = \tilde{\mathbf{X}}$, i.e., if and only if $\mathbf{X}$ is circularly-symmetric Gaussian with covariance matrix $\mathsf{K} = \tilde{\mathsf{K}}$.

We now define the normalized covariance matrices $\hat{\mathsf{K}} = \mathsf{K}/\mathcal{E}_{\mathrm{s}}$ and $\hat{\mathsf{K}}^* = \mathsf{K}^*/\mathcal{E}_{\mathrm{s}}$, where K^* denotes the capacity-achieving input covariance matrix. Let

$$\mathsf{M} = \mathsf{E}_{\mathbb{H}}\left[\mathbb{H}^\dagger\left(\mathbb{H}\hat{\mathsf{K}}\mathbb{H}^\dagger + \frac{1}{\rho}\mathsf{I}_m\right)^{-1}\mathbb{H}\right]. \tag{20}$$

By the Karush-Kuhn-Tucker optimality conditions of (5) one can show that $\operatorname{tr}\left(\hat{\mathsf{K}}^*\mathsf{M}\right) = \max_i \mathsf{M}_{i,i}$. Using the bound (18) once for the capacity-achieving input $\mathbf{X} = \mathbf{X}^*$, i.e., for $\mathsf{K} = \mathsf{K}^*$, and another time for $\mathbf{X} = \tilde{\mathbf{X}}$, i.e., for $\mathsf{K} = \tilde{\mathsf{K}}$, we get

$$C - I(\mathbf{X}; \mathbf{Y}, \widetilde{\mathbb{H}}) \leq \max_i \mathsf{M}_{i,i} - \operatorname{tr}\left(\hat{\mathsf{K}}\mathsf{M}\right) \tag{21}$$

where equality holds if and only if $\mathbf{X} = \mathbf{X}^*$, in which case both sides are equal to zero.

3.2 Proof of Theorem 1

Applying Lemma 1 with $\hat{\mathsf{K}} = \mathsf{I}_n/n$, i.e., for $\mathbf{X} = \mathbf{X}_{\mathrm{GI}}$ (7) we obtain

$$C - I(\mathbf{X}_{\mathrm{GI}}; \mathbf{Y}, \widetilde{\mathbb{H}}) \leq \max_i \mathsf{M}_{i,i} - \frac{1}{n}\mathrm{tr}\,(\mathsf{M}) \tag{22}$$

where

$$\mathsf{M} = n \cdot \mathsf{E}_{\mathbb{H}}\left[\mathbb{H}^\dagger \left(\mathbb{H}\mathbb{H}^\dagger + \frac{n}{\rho}\mathsf{I}_m\right)^{-1} \mathbb{H}\right]. \tag{23}$$

Consider a singular-value decomposition (SVD) of $\mathbb{H}$, i.e.,

$$\mathbb{H} = \mathbb{U}\hat{\mathbb{H}}\mathbb{V}^\dagger \tag{24}$$

where $\mathbb{U} \in \mathbb{C}^{m\times m}$ and $\mathbb{V} \in \mathbb{C}^{n\times n}$ are unitary random matrices, and where $\hat{\mathbb{H}} \in \mathbb{C}^{m\times n}$ is a random "diagonal" matrix where the diagonal elements are the decreasingly ordered singular values of $\mathbb{H}$, i.e.,

$$\sqrt{\Lambda_1} = \hat{\mathbb{H}}_{1,1} \geq \sqrt{\Lambda_2} = \hat{\mathbb{H}}_{2,2} \geq \ldots \geq \sqrt{\Lambda_l} = \hat{\mathbb{H}}_{l,l} \geq 0 \tag{25}$$

where $l = \min\{m, n\}$. Note that for a particular realization of $\mathbb{H}$ the matrix $\hat{\mathbb{H}}$ is uniquely determined, whereas the matrices $\mathbb{U}$ and $\mathbb{V}$ are not.

We notice that the random matrix

$$\mathbb{W} = \mathbb{H}^\dagger\mathbb{H} \in \mathbb{C}^{n\times n} \tag{26}$$

has a non-central Wishart distribution and that it can be diagonalized as

$$\mathbb{W} = \mathbb{V}\mathbb{L}\mathbb{V}^\dagger \tag{27}$$

where $\mathbb{L} = \hat{\mathbb{H}}^\dagger\hat{\mathbb{H}} = \mathrm{diag}\,(\Lambda_1, \ldots, \Lambda_l, 0, \ldots, 0) \in \mathbb{C}^{n\times n}$ so that $\{\Lambda_i\}$ and $\{\mathbf{V}_i\}$ are the decreasingly ordered eigenvalues and corresponding eigenvectors of $\mathbb{W}$, respectively, where $\mathbf{V}_i$ denotes the i-th column of $\mathbb{V}$. (Note that if $n > m$, then $\mathbb{W}$ has $n - m$ eigenvalues which are deterministically zero.)

Using the SVD of $\mathbb{H}$ we can write the matrix M as

$$\begin{aligned}
\mathsf{M} &= n \cdot \mathsf{E}_{\mathbb{W}}\left[\mathbb{V}\hat{\mathbb{H}}^\dagger \left(\hat{\mathbb{H}}\hat{\mathbb{H}}^\dagger + \frac{n}{\rho}\mathsf{I}_m\right)^{-1} \hat{\mathbb{H}}\mathbb{V}^\dagger\right] && (28)\\
&= n \cdot \mathsf{E}_{\mathbb{W}}\left[\mathbb{V}\,\mathrm{diag}\left(\frac{\Lambda_1}{\Lambda_1 + \frac{n}{\rho}}, \ldots, \frac{\Lambda_l}{\Lambda_l + \frac{n}{\rho}}, 0, \ldots, 0\right)\mathbb{V}^\dagger\right] && (29)
\end{aligned}$$

so that $\mathrm{tr}\,(\mathsf{M})$ depends only on $\mathbb{L}$, i.e., only on the eigenvalues of $\mathbb{W}$:

$$\begin{aligned}
\mathrm{tr}\,(\mathsf{M}) &= n \cdot \mathrm{tr}\left(\mathsf{E}_{\mathbb{W}}\left[\mathbb{V}\,\mathrm{diag}\left(\frac{\Lambda_1}{\Lambda_1 + \frac{n}{\rho}}, \ldots, \frac{\Lambda_l}{\Lambda_l + \frac{n}{\rho}}, 0, \ldots, 0\right)\mathbb{V}^\dagger\right]\right) && (30)\\
&= n \cdot \sum_{i=1}^{l} \mathsf{E}_{\mathbb{W}}\left[\frac{\Lambda_i}{\Lambda_i + \frac{n}{\rho}}\right] && (31)
\end{aligned}$$

and the diagonal elements of M are

$$\mathsf{M}_{i,i} = n \cdot \mathsf{E}_{\mathbb{W}}\left[\sum_{j=1}^{l} |\mathbb{V}_{i,j}|^2 \frac{\Lambda_j}{\Lambda_j + \frac{n}{\rho}}\right], \quad i \in \{1, \ldots, n\}. \tag{32}$$

From the above we now derive the upper bound (13). Since $\mathbb{V}$ is unitary, so that any of its rows has Euclidean norm one, we get from (32) the upper bound

$$\max_i \mathsf{M}_{i,i} \leq n. \tag{33}$$

We next find a lower bound for the sum (31). To do so we use the fact that the decreasingly ordered eigenvalues of a non-central Wishart matrix are stochastically larger [6] than the corresponding eigenvalues of a central Wishart matrix [7], i.e., that

$$\Lambda_i \geq^{\text{st}} \tilde{\Lambda}_i, \quad \forall i \in \{1, \ldots, l\} \tag{34}$$

where $\{\Lambda_i\}$ and $\{\tilde{\Lambda}_i\}$ are the decreasingly ordered eigenvalues of $\mathbb{W} = (\widetilde{\mathbb{H}} + \mathsf{D})^\dagger(\widetilde{\mathbb{H}} + \mathsf{D})$ and $\widetilde{\mathbb{W}} = \widetilde{\mathbb{H}}^\dagger\widetilde{\mathbb{H}}$, respectively. Since the function $f(\lambda) = \frac{\lambda}{\lambda + n/\rho}$, $\lambda \geq 0$, is strictly increasing we get the bounds

$$\mathsf{E}_{\mathbb{W}}\left[\frac{\Lambda_i}{\Lambda_i + \frac{n}{\rho}}\right] \geq \mathsf{E}_{\widetilde{\mathbb{W}}}\left[\frac{\tilde{\Lambda}_i}{\tilde{\Lambda}_i + \frac{n}{\rho}}\right], \quad \forall i \in \{1, \ldots, l\}. \tag{35}$$

And because the eigenvalues are decreasingly ordered we further have

$$\mathsf{E}_{\widetilde{\mathbb{W}}}\left[\frac{\tilde{\Lambda}_i}{\tilde{\Lambda}_i + \frac{n}{\rho}}\right] \geq \mathsf{E}_{\widetilde{\mathbb{W}}}\left[\frac{\tilde{\Lambda}_l}{\tilde{\Lambda}_l + \frac{n}{\rho}}\right], \quad \forall i \in \{1, \ldots, l\} \tag{36}$$

so that we are left with finding a lower bound on $\mathsf{E}_{\widetilde{\mathbb{W}}}\left[\frac{\tilde{\Lambda}_l}{\tilde{\Lambda}_l + \frac{n}{\rho}}\right]$.

The l-th eigenvalue of $\widetilde{\mathbb{H}}^\dagger\widetilde{\mathbb{H}}$ equals the l-th eigenvalue of $\widetilde{\mathbb{H}}\widetilde{\mathbb{H}}^\dagger$. Thus, without loss of generality we only consider the case $n \geq m$. Writing $\widetilde{\mathbb{H}} = [\widetilde{\mathbb{H}}_{\text{s}}, \widetilde{\mathbb{H}}_{\text{r}}]$ where $\widetilde{\mathbb{H}}_{\text{s}} = [\tilde{\mathbf{H}}_1, \ldots, \tilde{\mathbf{H}}_m] \in \mathbb{C}^{m \times m}$ is square and $\widetilde{\mathbb{H}}_{\text{r}} = [\tilde{\mathbf{H}}_{m+1}, \ldots, \tilde{\mathbf{H}}_n] \in \mathbb{C}^{m \times (n-m)}$ (and where we write $\widetilde{\mathbb{H}} = \widetilde{\mathbb{H}}_{\text{s}}$ and define $\widetilde{\mathbb{H}}_{\text{r}} = \mathsf{0}$ if $n = m$) we have

$$\widetilde{\mathbb{H}}\widetilde{\mathbb{H}}^\dagger = \widetilde{\mathbb{H}}_{\text{s}}\widetilde{\mathbb{H}}_{\text{s}}^\dagger + \widetilde{\mathbb{H}}_{\text{r}}\widetilde{\mathbb{H}}_{\text{r}}^\dagger. \tag{37}$$

The matrix $\widetilde{\mathbb{H}}\widetilde{\mathbb{H}}^\dagger - \widetilde{\mathbb{H}}_{\text{s}}\widetilde{\mathbb{H}}_{\text{s}}^\dagger = \widetilde{\mathbb{H}}_{\text{r}}\widetilde{\mathbb{H}}_{\text{r}}^\dagger$ is positive semi-definite, so that by definition $\widetilde{\mathbb{H}}\widetilde{\mathbb{H}}^\dagger \succcurlyeq \widetilde{\mathbb{H}}_{\text{s}}\widetilde{\mathbb{H}}_{\text{s}}^\dagger$ [8, Definition 7.7.1]. But, since $\widetilde{\mathbb{H}}\widetilde{\mathbb{H}}^\dagger$ and $\widetilde{\mathbb{H}}_{\text{s}}\widetilde{\mathbb{H}}_{\text{s}}^\dagger$ are positive definite almost surely, this means that

$$\tilde{\Lambda}_l \geq \Sigma_l \quad \text{almost surely} \tag{38}$$

where Σ_l is the l-th eigenvalue of $\widetilde{\mathbb{H}}_{\text{s}}\widetilde{\mathbb{H}}_{\text{s}}^\dagger$, i.e., the smallest of the l decreasingly ordered eigenvalues of the square central Wishart matrix $\widetilde{\mathbb{H}}_{\text{s}}^\dagger\widetilde{\mathbb{H}}_{\text{s}}$ [8, Corollary 7.7.4].

The joint probability density of the decreasingly ordered eigenvalues $\{\Sigma_i\}$ of $\widetilde{\mathbb{H}}_{\text{s}}^\dagger\widetilde{\mathbb{H}}_{\text{s}}$ and in particular the marginal density of Σ_l are known [9]. The latter Σ_l is exponentially distributed with mean $1/l$, so that

$$\mathsf{E}_{\Sigma_l}\left[\frac{\Sigma_l}{\Sigma_l + \frac{n}{\rho}}\right] = 1 + \frac{nl}{\rho} \cdot e^{\frac{nl}{\rho}} \cdot \text{Ei}\left(-\frac{nl}{\rho}\right). \tag{39}$$

Combining (31), (35), (36), (38), and (39) we obtain

$$\frac{1}{n}\text{tr}\,(\mathsf{M}) \geq l + \frac{nl^2}{\rho} \cdot e^{\frac{nl}{\rho}}\, \text{Ei}\left(-\frac{nl}{\rho}\right). \tag{40}$$

The result now follows from (33) and (40).

References

[1] Daniel Hösli and Amos Lapidoth. The capacity of a MIMO Ricean channel is monotonic in the singular values of the mean. To appear in *Proceedings of the 5th International ITG Conference on Source and Channel Coding (SCC)*, 2004.

[2] Sivarama Venkatesan, Steven H. Simon, and Reinaldo A. Valenzuela. Capacity of a Gaussian MIMO channel with nonzero mean. In *Proceedings of the IEEE Semiannual Vehicular Technology Conference*, Orlando, FL, October 6–9 2003.

[3] Mahesh Godavarti and Alfred O. Hero-III. Multiple-antenna capacity in a deterministic Ricean fading channel. Submitted to *IEEE Transactions on Information Theory*, September 2001.

[4] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press. To appear in February 2004.

[5] Amos Lapidoth and Stefan M. Moser. Capacity bounds via duality with applications to multiple-antenna systems on flat fading channels. *IEEE Transactions on Information Theory*, 49(10):2426–2467, October 2003.

[6] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications*. Academic Press, 1979.

[7] Michael D. Perlman and Ingram Olkin. Unbiasedness of invariant test for MANOVA and other multivariate problems. *The Annals of Statistics*, 8(6):1326–1341, November 1980.

[8] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[9] Alan Edelman. Eigenvalues and condition numbers of random matrices. *SIAM Journal on Matrix Analysis and Applications*, 9(4):543–560, October 1988.