How Good is an Isotropic Gaussian Input on a MIMO Ricean Channel?

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

Abstract — **For a MIMO Ricean fading channel with perfect side information at the receiver we derive an analytic upper bound on the difference between capacity and the mutual information that is induced by an isotropic Gaussian input. We show that if the number of receiver antennas is at least equal to the number of transmitter antennas, then, as the signal-to-noise ratio tends to infinity, such an input is asymptotically optimal. But otherwise such an isotropic input might be suboptimal. We also propose an iterative algorithm to calculate the optimal power allocation.**

I. Introduction

We consider a discrete-time memoryless MIMO channel whose output $\mathbf{Y} \in \mathbb{C}^m$ is given by

$$\mathbf{Y} = (\widetilde{\mathbb{H}} + \mathsf{D})\mathbf{x} + \mathbf{Z} \tag{1}$$

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[1]; the fading matrix $\mathbb{H} = \widetilde{\mathbb{H}} + \mathsf{D} \in \mathbb{C}^{m\times n}$ consists of a deterministic mean matrix D and a random matrix $\widetilde{\mathbb{H}}$, whose $m \cdot n$ random components 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 $\mathbf{x}$.

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 $\mathsf{E}\big[\mathbf{X}^\dagger\mathbf{X}\big] \le \mathcal{E}_{\mathrm{s}}$, 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. Combining the input power and the noise power to a single "signal-to-noise ratio" parameter $\rho = \frac{\mathcal{E}_{\mathrm{s}}}{\sigma^2}$, capacity can be expressed as

$$C(\rho) = \sup_{\hat{\mathsf{K}} \in \mathcal{K}} \mathsf{E}_{\widetilde{\mathbb{H}}}\Big[\log\det\Big(\mathsf{I}_m + \rho(\widetilde{\mathbb{H}} + \mathsf{D})\hat{\mathsf{K}}(\widetilde{\mathbb{H}} + \mathsf{D})^\dagger\Big)\Big] \tag{2}$$

where $\mathcal{K}$ is the set of positive semi-definite matrices $\hat{\mathsf{K}}$ with trace $\mathrm{tr}\Big(\hat{\mathsf{K}}\Big) \le 1$.

As shown in [1, 2], if $\mathsf{D} = \mathsf{U}\Sigma_{\mathsf{D}}\mathsf{V}^\dagger$ is a singular value decomposition, where $(\Sigma_{\mathsf{D}})_{i,i} = \sigma_i$ are the decreasingly ordered singular values of D, then the optimal normalized input covariance matrix is $\hat{\mathsf{K}}^* = \mathsf{V}\Lambda_{\hat{\mathsf{K}}^*}\mathsf{V}^\dagger$, where $\Lambda_{\hat{\mathsf{K}}^*} = \mathrm{diag}\,(\lambda_1^*, \ldots, \lambda_n^*)$. Since no explicit expression for the decreasingly ordered eigenvalues (power allocation) $\{\lambda_i^*\}$ is known, and since in some systems it may be advisable to choose codes that do not depend on D, it is natural to ask how much is lost w.r.t. capacity if a uniform power allocation $\lambda_i = 1/n$, $1 \le i \le n$, is used.

[1] 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.

II. Main Results

The following result is stated without proof.

Theorem 1. *The difference between capacity (2) and the mutual information between the input and output of the channel (1) induced by* $\mathbf{X}_{\mathrm{GI}} \sim \mathcal{N}_{\mathbb{C}}\big(0, \frac{\mathcal{E}_{\mathrm{s}}}{n}\mathsf{I}_n\big)$ *is upper bounded as*

$$C(\rho) - I(\mathbf{X}_{\mathrm{GI}}; \mathbf{Y}, \widetilde{\mathbb{H}}) \le n - l - \frac{nl^2}{\rho} \cdot e^{\frac{nl}{\rho}} \cdot \mathrm{Ei}\left(-\frac{nl}{\rho}\right) \tag{3}$$

where $l = \min\{m, n\}$ *and* $\mathrm{Ei}\,(-\xi) = -\int_\xi^\infty \frac{e^{-t}}{t}\,\mathrm{d}t$, $\xi > 0$, *is the exponential integral function.*

If $n \le m$, then the bound tends to zero as $\rho \to \infty$, which shows that $\mathbf{X}_{\mathrm{GI}}$ is asymptotically optimal. In fact, one can show that in this case capacity (2) can be expressed as

$$C(\rho) = n\log\rho + \mathsf{E}\Big[\log\det(\mathbb{H}^\dagger\mathbb{H})\Big] - n\log n + o(1) \tag{4}$$

and is achieved asymptotically by $\mathcal{N}_{\mathbb{C}}\Big(\mathbf{0}, \mathcal{E}_{\mathrm{s}}\hat{\mathsf{K}}(\rho)\Big)$ inputs ($\hat{\mathsf{K}}(\rho) \in \mathcal{K}$) if, and only if, $\lim_{\rho\to\infty}\hat{\mathsf{K}}(\rho) = \mathsf{I}_n/n$ componentwise.

The situation is different if $m < n$. In this case

$$C(\rho) = m\log\rho + \sup_{\hat{\mathsf{K}} \in \mathcal{K}} \mathsf{E}\Big[\log\det(\mathbb{H}\hat{\mathsf{K}}\mathbb{H}^\dagger)\Big] + o(1). \tag{5}$$

Since the cost function is strictly concave in $\hat{\mathsf{K}}$, the asymptotically optimal $\hat{\mathsf{K}}^*$ is unique. Considering the decreasingly ordered singular values $\{\sigma_i\}$ of D, one can show that if $\sigma_i = \sigma_j$ for $1 \le i, j \le n$ (where we define $\sigma_k = 0$ if $k > m$), then $\lambda_i^* = \lambda_j^*$. Also, a numerical evaluation shows that unless $\mathsf{D} = \mathsf{0}$ (i.e., Rayleigh fading) a uniform power allocation $\hat{\mathsf{K}} = \mathsf{I}_n/n$ need not be asymptotically optimal.

Such numerical results can be obtained by use of the iterative algorithm below, which is based on an application of the Blahut-Arimoto algorithm to the channel (1) and the fact that if one initializes the algorithm with a Gaussian distribution one obtains a Gaussian distribution in the next step:

- Choose $s > 0$ and $\sigma^2 > 0$, and set $\mathsf{K}^{(0)} = \mathsf{I}_n$.
- For $i \ge 0$ iterate $\mathsf{K}^{(i+1)} = \Big((\mathsf{K}^{(i)})^{-1} + s\mathsf{I}_n - \mathsf{M}^{(i)}\Big)^{-1}$ with $\mathsf{M}^{(i)} = \mathsf{E}\Big[\mathbb{H}^\dagger\Big(\mathbb{H}\mathsf{K}^{(i)}\mathbb{H}^\dagger + \sigma^2\mathsf{I}_m\Big)^{-1}\mathbb{H}\Big]$.

The algorithm converges towards the optimal input covariance matrix K^* with $\mathrm{tr}\,(\mathsf{K}^*) = \mathcal{E}_{\mathrm{s}}$ for the channel (1), where $\mathcal{E}_{\mathrm{s}}$ depends on the choice of s.

References

[1] D. Hösli, A. Lapidoth, "The Capacity of a MIMO Ricean Channel is Monotonic in the Singular Values of the Mean," *Proc. 5th Int. ITG Conf. SCC*, Erlangen, Germany, Jan. 14–16, 2004.

[2] S. Venkatesan, S. H. Simon, and R. A. Valenzuela, "Capacity of a Gaussian MIMO Channel with Nonzero Mean," *Proc. IEEE Semiann. Vehicular Technol. Conf.*, Orlando, FL, Oct. 6–9, 2003.