Bounds on the Fading Number of Multiple-Input Single-Output Fading Channels with Memory

Stefan M. Moser*
Department of Communication Engineering
National Chiao Tung University (NCTU)
Hsinchu 300, Taiwan
stefan.moser@ieee.org

ABSTRACT

We derive new upper and lower bounds on the fading number of multiple-input single-output (MISO) fading channels of general (not necessarily Gaussian) regular law with spatial and temporal memory. The fading number is the second term, after the double-logarithmic term, of the high signal-to-noise ratio (SNR) expansion of channel capacity.

In case of an isotropically distributed fading vector it is proven that the upper and lower bound coincide, *i.e.*, the general MISO fading number with memory is known precisely.

The upper and lower bounds show that a type of beam-forming is asymptotically optimal.

Categories and Subject Descriptors

H.1.1 [**Systems and Information Theory**]: Information Theory

General Terms

Theory

Keywords

Beam-forming, channel capacity, fading number, flat fading, high SNR, memory, MISO, multiple-antenna, non-coherent.

1. INTRODUCTION

It has been recently shown in [1], [2] that, whenever the matrix-valued fading process is of finite differential entropy rate (so-called *regular fading*), the non-coherent capacity of multiple-input multiple-output (MIMO) fading channels typically grows only double-logarithmically in the signal-to-noise ratio (SNR).

This is in stark contrast to both, the coherent fading channel where the receiver has *perfect* knowledge about the channel state, and to the non-coherent fading channel where the channel law is *non-regular*, *i.e.*, its differential entropy rate is not finite. In the former case the capacity grows logarithmically in the SNR with a factor in front of the logarithm equal to the minimum of the number of receiver and transmitter antennas [3].

In the latter case the asymptotic growth rate of the capacity depends highly on the specific details of the fading process. In the case of Gaussian fading, non-regularity means that the present fading realization can be predicted *precisely* from the past realizations. However, note that in any practical system the past realizations are not known *a priori*, but need to be estimated either by known past channel inputs and outputs or by means of special training signals. Depending on the spectral distribution of the fading process, the dependence of such estimations on the available power can vary largely which gives rise to a huge variety of possible high-SNR capacity behaviors: it is shown in [4], [5], and [6] that depending on the spectrum $\mathsf{F}(\cdot)$ of the non-regular Gaussian fading process, the asymptotic behavior of the channel capacity can be varied, *e.g.*, double-logarithmic, logarithmic, or a fractional power thereof.

Similarly, Liang and Veeravalli show in [7] that the capacity of a Gaussian block-fading channel depends critically on the assumptions one makes about the time-correlation of the fading process: if the correlation matrix is rank deficient, the capacity grows logarithmically in the SNR, otherwise double-logarithmically.

In this paper we will only consider regular fading processes, *i.e.*, the capacity at high SNR will be growing double-logarithmically. To quantify the rates at which this poor power efficiency begins, [1], [2] introduce the *fading number* as the second term in the high-SNR asymptotic expansion of channel capacity. Hence, the capacity can be written as

$$C(\text{SNR}) = \log(1 + \log(1 + \text{SNR})) + \chi + o(1) \tag{1}$$

where $o(1)$ tends to zero as the SNR tends to infinity.

Explicit expressions for the fading number are known for a number of fading models. For channels with memory, the fading number of single-input single-output (SISO) fading channels is derived in [1], [2]. The single-input multiple-output (SIMO) case is derived in [8], [9], [2].

*This work was supported in part by the National Science Council under NSC 94-2218-E-009-037.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
IWCMC'06, July 3–6, 2006, Vancouver, British Columbia, Canada.
Copyright 2006 ACM 1-59593-306-9/06/0007 ...$5.00.

The fading number of the multiple-input single-output (MISO) fading channel has been derived in general only for the memoryless case [1], [2]:

$$\chi(\mathbf{H}^{\mathsf{T}}) = \sup_{\|\hat{\mathbf{x}}\|=1} \left\{ \log\pi + \mathsf{E}\left[\log|\mathbf{H}^{\mathsf{T}}\hat{\mathbf{x}}|^2\right] - h(\mathbf{H}^{\mathsf{T}}\hat{\mathbf{x}}) \right\}. \quad (2)$$

This fading number is achievable by inputs that can be expressed as the product of a constant unit vector in $\mathbb{C}^{n_\mathrm{T}}$ and a circularly symmetric, scalar, complex random variable of the same law that achieves the memoryless SISO fading number [1]. Hence, the memoryless MISO fading number is achieved by beam-forming where the beam-direction is chosen not to maximize the SNR, but the fading number.

In [10] and [11] Koch & Lapidoth investigate the fading number of MISO fading channels with memory where the fading process is Gaussian. For the case of a mean-$\mathbf{d}$ Gaussian vector process with memory where $\{\mathbf{H}_k - \mathbf{d}\}$ is spatially independent and identically distributed (IID) and where each component is a zero-mean unit-variance circularly symmetric complex Gaussian process, the fading number is shown to be[1]

$$\chi_{\text{Gauss, spat. IID}}(\{\mathbf{H}_k^{\mathsf{T}}\}) = -1 + \log\|\mathbf{d}\|^2 - \mathrm{Ei}\bigl(-\|\mathbf{d}\|^2\bigr) + \log\frac{1}{\epsilon^2}, \quad (3)$$

where ϵ^2 denotes the prediction error when predicting one of the components of the fading vector based on the observation of its past.

Furthermore, Koch & Lapidoth derive an upper bound to the fading number for the general Gaussian case, *i.e.*, for a mean vector $\mathbf{d}$, $\{\mathbf{H}_k - \mathbf{d}\}$ is a zero-mean circularly symmetric stationary ergodic complex Gaussian process with matrix-valued spectral distribution function $\mathsf{F}(\cdot)$ and with covariance matrix K. Assuming that the prediction error covariance matrix $\boldsymbol{\Sigma}$ is non-singular (regularity assumption) they show that

$$\chi_{\text{Gauss}}(\{\mathbf{H}_k^{\mathsf{T}}\}) \le -1 + \log d_*^2 - \mathrm{Ei}\bigl(-d_*^2\bigr) + \log\frac{\|\mathsf{K}\|}{\sigma_{\min}}, \quad (4)$$

where

$$d_* = \max_{\|\hat{\mathbf{x}}\|=1} \frac{\left|\mathbf{d}^{\mathsf{T}}\hat{\mathbf{x}}\right|}{\sqrt{\mathsf{Var}(\mathbf{H}_k^{\mathsf{T}}\hat{\mathbf{x}})}}; \quad (5)$$

$\sigma_{\min}$ denotes the smallest eigenvalue of $\boldsymbol{\Sigma}$; and where $\|\cdot\|$ denotes the Euclidean operator norm of matrices, *i.e.*, the largest singular value.

In this paper we extend these results to general (not necessarily Gaussian) fading channels.

The remaining of this paper is structured as follows: after defining the channel model in detail in the following section, we will present the main results, *i.e.*, the new upper and lower bound on the MISO fading number, in Section 3.

We then specialize these results to the case of isotropically distributed fading processes in Section 4 and to Gaussian fading in Section 5. For isotropically distributed fading we will show that the upper and lower bound coincide. In the Gaussian case we shall derive the above mentioned results of Koch & Lapidoth as special cases of our bounds.

We conclude in Section 6.

[1]Note that all results in this paper are in nats.

2. THE CHANNEL MODEL

We consider a MISO fading channel whose time-k output $Y_k \in \mathbb{C}$ is given by

$$Y_k = \mathbf{H}_k^{\mathsf{T}}\mathbf{x}_k + Z_k \quad (6)$$

where $\mathbf{x}_k \in \mathbb{C}^{n_\mathrm{T}}$ denotes the time-k channel input; the random vector $\mathbf{H}_k$ is the time-k fading vector; $\mathbf{H}_k^{\mathsf{T}}$ denotes the transpose of the vector $\mathbf{H}_k$; and where Z_k denotes additive noise. Here $\mathbb{C}$ denotes the complex field, $\mathbb{C}^{n_\mathrm{T}}$ denotes the n_T-dimensional complex Euclidean space, and n_T is the number of transmitter antennas. We assume that the additive noise is an IID zero-mean white Gaussian process of variance $\sigma^2 > 0$.

As for the multi-variate fading process $\{\mathbf{H}_k\}$, we shall only assume that it is stationary, ergodic, of finite second moment

$$\mathsf{E}\left[\|\mathbf{H}_k\|^2\right] < \infty, \quad (7)$$

and of finite differential entropy rate (regularity condition)

$$h(\{\mathbf{H}_k\}) > -\infty \quad (8)$$

Finally, we assume that the fading process $\{\mathbf{H}_k\}$ and the additive noise process $\{Z_k\}$ are independent and of a joint law that does not depend on the channel input $\{\mathbf{x}_k\}$.

As for the input, we consider two different constraints: a peak-power constraint and an average-power constraint. We use $\mathcal{E}$ to denote the maximal allowed instantaneous power in the former case, and to denote the allowed average power in the latter case. For both cases we set

$$\mathrm{SNR} \triangleq \frac{\mathcal{E}}{\sigma^2}. \quad (9)$$

The capacity $\mathsf{C}(\mathrm{SNR})$ of the channel (6) is given by

$$\mathsf{C}(\mathrm{SNR}) = \lim_{n\to\infty}\frac{1}{n}\sup I\left(\mathbf{X}_1^n; Y_1^n\right) \quad (10)$$

where we use $\mathbf{X}_j^k$ to denote $\mathbf{X}_j, \ldots, \mathbf{X}_k$ and where the supremum is over the set of all probability distributions on $\mathbf{X}_1^n$ satisfying the constraints, *i.e.*,

$$\|\mathbf{X}_k\|^2 \le \mathcal{E}, \quad \text{almost surely}, \quad k = 1, 2, \ldots, n \quad (11)$$

for a peak-power constraint, or

$$\frac{1}{n}\sum_{k=1}^{n}\mathsf{E}\left[\|\mathbf{X}_k\|^2\right] \le \mathcal{E} \quad (12)$$

for an average-power constraint.

Specializing [1, Theorem 4.2], [2, Theorem 6.10], to MISO fading, we have

$$\varlimsup_{\mathrm{SNR}\uparrow\infty}\Bigl\{\mathsf{C}(\mathrm{SNR}) - \log\log\mathrm{SNR}\Bigr\} < \infty. \quad (13)$$

The fading number χ is now defined as in [1, Definition 4.6], [2, Definition 6.13] by

$$\chi(\{\mathbf{H}_k^{\mathsf{T}}\}) = \varlimsup_{\mathrm{SNR}\uparrow\infty}\Bigl\{\mathsf{C}(\mathrm{SNR}) - \log\log\mathrm{SNR}\Bigr\}. \quad (14)$$

Prima facie the fading number depends on whether a peak-power constraint (11) or an average-power constraint (12) is imposed on the input. Since a peak-power constraint is more stringent than an average-power constraint, we will derive the upper bound using the average-power constraint and the lower bound using the peak-power constraint. In case of an isotropically distributed fading process we shall see that both constraints lead to identical fading numbers.

3. MAIN RESULTS

We first derive an upper bound to the fading number of a MISO fading channel:

THEOREM 1. *Consider a MISO fading channel with memory* (6) *where the stationary and ergodic fading process* $\{\mathbf{H}_k\}$ *takes value in* $\mathbb{C}^{n_\mathrm{T}}$ *and satisfies*

$$h(\{\mathbf{H}_k\}) > -\infty \tag{15}$$

and

$$\mathsf{E}\left[\|\mathbf{H}_k\|^2\right] < \infty. \tag{16}$$

Then, irrespective of whether a peak-power constraint (11) *or an average-power constraint* (12) *is imposed on the input, the fading number* $\chi(\{\mathbf{H}_k^\mathsf{T}\})$ *is upper bounded by*

$$\chi\big(\{\mathbf{H}_k^\mathsf{T}\}\big) \leq \sup_{\hat{\mathbf{x}}} \Big\{ \log\pi + \mathsf{E}\big[\log|\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}|^2\big] - h(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}) \Big\} + \sup_{\hat{\mathbf{x}}_{-\infty}^0} \Big\{ h(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0) - h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0 \,\big|\, \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=-\infty}^{-1}\big) \Big\} \tag{17}$$

where $\hat{\mathbf{x}}_\ell \triangleq \frac{\mathbf{x}_\ell}{\|\mathbf{x}_\ell\|}$ *denotes a vector of unit length.*

PROOF. We use the chain rule to get

$$\frac{1}{n} I\big(\mathbf{X}_1^n; Y_1^n\big) = \frac{1}{n}\sum_{k=1}^n I\big(\mathbf{X}_1^n; Y_k \,\big|\, Y_1^{k-1}\big), \tag{18}$$

and upper bound each term on the right-hand side (RHS) of the above as follows:

$$\begin{aligned}
&I\big(\mathbf{X}_1^n; Y_k \,\big|\, Y_1^{k-1}\big)\\
&= I\big(\mathbf{X}_1^n, Y_1^{k-1}; Y_k\big) - I\big(Y_k; Y_1^{k-1}\big) &(19)\\
&\leq I\big(\mathbf{X}_1^n, Y_1^{k-1}; Y_k\big) &(20)\\
&= I\big(\mathbf{X}_1^{k-1}, Y_1^{k-1}, \mathbf{X}_k; Y_k\big) &(21)\\
&\leq I\big(\mathbf{X}_1^{k-1}, Y_1^{k-1}, \{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=1}^{k-1}, \mathbf{X}_k; Y_k\big) &(22)\\
&= I\big(\mathbf{X}_1^{k-1}, \{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=1}^{k-1}, \mathbf{X}_k; Y_k\big) &(23)\\
&= I(\mathbf{X}_k; Y_k) + \underbrace{I\big(\mathbf{X}_1^{k-1}; Y_k \big| \mathbf{X}_k\big)}_{=0}\\
&\quad + I\big(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=1}^{k-1}; Y_k \,\big|\, \mathbf{X}_1^k\big) &(24)\\
&= I(\mathbf{X}_k; Y_k) + I\big(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=1}^{k-1}; Y_k \,\big|\, \mathbf{X}_1^k\big) &(25)\\
&= I(\mathbf{X}_k; Y_k) + I\big(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=1}^{k-1}; Y_k, \mathbf{X}_k \,\big|\, \mathbf{X}_1^{k-1}\big)\\
&\quad - I\big(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=1}^{k-1}; \mathbf{X}_k \,\big|\, \mathbf{X}_1^{k-1}\big) &(26)\\
&= I(\mathbf{X}_k; Y_k) + I\big(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=1}^{k-1}; Y_k, \mathbf{X}_k \,\big|\, \mathbf{X}_1^{k-1}\big) &(27)\\
&\leq I(\mathbf{X}_k; Y_k) + I\big(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=1}^{k-1}; Y_k, \mathbf{X}_k, \mathbf{H}_k^\mathsf{T}\mathbf{X}_k \,\big|\, \mathbf{X}_1^{k-1}\big) &(28)\\
&= I(\mathbf{X}_k; Y_k) + I\big(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=1}^{k-1}; \mathbf{X}_k, \mathbf{H}_k^\mathsf{T}\mathbf{X}_k \,\big|\, \mathbf{X}_1^{k-1}\big) &(29)\\
&= I(\mathbf{X}_k; Y_k) + \underbrace{I\big(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=1}^{k-1}; \mathbf{X}_k \,\big|\, \mathbf{X}_1^{k-1}\big)}_{=0}\\
&\quad + I\big(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=1}^{k-1}; \mathbf{H}_k^\mathsf{T}\mathbf{X}_k \,\big|\, \mathbf{X}_1^k\big) &(30)\\
&= I(\mathbf{X}_k; Y_k) + I\big(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=1}^{k-1}; \mathbf{H}_k^\mathsf{T}\mathbf{X}_k \,\big|\, \mathbf{X}_1^k\big) &(31)\\
&= I(\mathbf{X}_k; Y_k) + I\big(\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=1}^{k-1}; \mathbf{H}_k^\mathsf{T}\mathbf{X}_k \,\big|\, \{\|\mathbf{X}_\ell\|\}_{\ell=1}^k, \hat{\mathbf{X}}_1^k\big) &(32)\\
&= I(\mathbf{X}_k; Y_k) + I\big(\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{X}}_\ell\}_{\ell=1}^{k-1}; \mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k \,\big|\, \hat{\mathbf{X}}_1^k\big). &(33)
\end{aligned}$$

Here the first equality follows from the chain rule; the subsequent inequality by the non-negativity of mutual information; the subsequent equality follows because we prohibit feedback; the subsequent inequality from the inclusion of the additional random variables $\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=1}^{k-1}$ in the mutual information term; (23) follows because, conditional on the past terms $\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=1}^{k-1}$ and on the present and past inputs $\mathbf{X}_1^k$, the past outputs Y_1^{k-1} are independent of the present output Y_k; the subsequent equality follows by the chain rule; the subsequent equality from the independence of the past inputs and the present output when conditioning on the present input; the subsequent equality again from the chain rule; in (27) we use the fact that based on the past inputs the present input and the random variables $\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=1}^{k-1}$ are independent; (28) follows from the inclusion of the random variable $\mathbf{H}_k^\mathsf{T}\mathbf{X}_k$ in the mutual information term; the subsequent equality because, conditional on $\mathbf{H}_k^\mathsf{T}\mathbf{X}_k$ and the present and past inputs, the past terms $\{\mathbf{H}_\ell^\mathsf{T}\mathbf{X}_\ell\}_{\ell=1}^{k-1}$ are independent of the present output; the subsequent equality from the chain rule; the subsequent equality by the independence of the inputs and the fading; and the final equality follows by dividing each term by the magnitude of the input vectors.

Putting (33) into (18) we get

$$\begin{aligned}
&\frac{1}{n} I\big(\mathbf{X}_1^n; Y_1^n\big)\\
&\leq \frac{1}{n}\sum_{k=1}^n \Big(I(\mathbf{X}_k; Y_k) + I\big(\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{X}}_\ell\}_{\ell=1}^{k-1}; \mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k \,\big|\, \hat{\mathbf{X}}_1^k\big)\Big) &(34)\\
&\leq \frac{1}{n}\sum_{k=1}^n \mathsf{C}_{\mathrm{IID}}(\mathsf{P}_k(\mathcal{E})) + \frac{1}{n}\sum_{k=1}^n I\big(\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{X}}_\ell\}_{\ell=1}^{k-1}; \mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k \,\big|\, \hat{\mathbf{X}}_1^k\big) &(35)\\
&\leq \mathsf{C}_{\mathrm{IID}}\left(\frac{1}{n}\sum_{k=1}^n \mathsf{P}_k(\mathcal{E})\right)\\
&\quad + \frac{1}{n}\sum_{k=1}^n I\big(\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{X}}_\ell\}_{\ell=1}^{k-1}; \mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k \,\big|\, \hat{\mathbf{X}}_1^k\big) &(36)\\
&\leq \mathsf{C}_{\mathrm{IID}}(\mathcal{E}) + \frac{1}{n}\sum_{k=1}^n I\big(\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{X}}_\ell\}_{\ell=1}^{k-1}; \mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k \,\big|\, \hat{\mathbf{X}}_1^k\big) &(37)
\end{aligned}$$

where $\mathsf{C}_{\mathrm{IID}}(\mathsf{P}_k(\mathcal{E}))$ denotes the capacity of the memoryless MISO fading channel for a given power $\mathsf{P}_k(\mathcal{E})$ which must satisfy the average-power constraint (12)

$$\frac{1}{n}\sum_{k=1}^n \mathsf{P}_k(\mathcal{E}) \leq \mathcal{E}; \tag{38}$$

where the third inequality follows by the concavity of capacity in the power-constraint and by Jensen's inequality; and where the last inequality follows from (38) and the fact that $\mathsf{C}(\mathrm{SNR})$ is monotonically non-decreasing.

We continue to bound the second term of (37):

$$\begin{aligned}
&\frac{1}{n} I\big(\mathbf{X}_1^n; Y_1^n\big)\\
&\leq \mathsf{C}_{\mathrm{IID}}(\mathcal{E})\\
&\quad + \frac{1}{n}\sum_{k=1}^n \mathsf{E}_{\hat{\mathbf{X}}_1^k}\Big[I\big(\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=1}^{k-1}; \mathbf{H}_k^\mathsf{T}\hat{\mathbf{x}}_k \,\big|\, \hat{\mathbf{X}}_1^k = \hat{\mathbf{x}}_1^k\big)\Big] &(39)\\
&\leq \mathsf{C}_{\mathrm{IID}}(\mathcal{E}) + \frac{1}{n}\sum_{k=1}^n \sup_{\hat{\mathbf{x}}_1^k} I\big(\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=1}^{k-1}; \mathbf{H}_k^\mathsf{T}\hat{\mathbf{x}}_k\big) &(40)\\
&\leq \mathsf{C}_{\mathrm{IID}}(\mathcal{E})\\
&\quad + \frac{1}{n}\sum_{k=1}^n \sup_{\hat{\mathbf{x}}_1^n} I\big(\mathbf{H}_k^\mathsf{T}\hat{\mathbf{x}}_k; \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=1}^{k-1}, \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_{\ell+n}\}_{\ell=k-n+1}^{0}\big) &(41)
\end{aligned}$$

$$= \mathsf{C}_{\text{IID}}(\mathcal{E}) + \frac{1}{n}\sum_{k=1}^{n} \sup_{\hat{\mathbf{x}}_1^n} I\big(\mathbf{H}_n^\mathsf{T}\hat{\mathbf{x}}_n; \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=1}^{n-1}\big) \quad (42)$$

$$= \mathsf{C}_{\text{IID}}(\mathcal{E}) + \sup_{\hat{\mathbf{x}}_1^n} I\big(\mathbf{H}_n^\mathsf{T}\hat{\mathbf{x}}_n; \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=1}^{n-1}\big) \quad (43)$$

$$\leq \mathsf{C}_{\text{IID}}(\mathcal{E}) + \sup_{\hat{\mathbf{x}}_{-\infty}^0} I\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0; \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=-\infty}^{-1}\big). \quad (44)$$

Here the second inequality follows from replacing the expectation by the supremum; the subsequent inequality from the inclusion of additional random variables in the mutual information term; the subsequent equality from stationarity; the subsequent equality from taking the constant terms out of the sum; and the final inequality from the inclusion of additional random variables.

The theorem now follows by letting n tend to infinity, taking the supremum over all input distribution that fulfill the average-power constraint (both steps having no impact on (44)), and using the MISO fading number without memory (2). □

Next we derive a lower bound to the fading number of a MISO fading channel:

Theorem 2. *Consider a MISO fading channel with memory* (6) *where the stationary and ergodic fading process* $\{\mathbf{H}_k\}$ *takes value in* $\mathbb{C}^{n_\text{T}}$ *and satisfies* (15) *and* (16). *Then the fading number* $\chi(\{\mathbf{H}_k^\mathsf{T}\})$ *is lower bounded by*

$$\chi\big(\{\mathbf{H}_k^\mathsf{T}\}\big) \geq \sup_{\hat{\mathbf{x}}} \Big\{ \log\pi + \mathsf{E}\big[\log|\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}|^2\big] - h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}} \,\big|\, \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=-\infty}^{-1}\big)\Big\} \quad (45)$$

where $\hat{\mathbf{x}} \triangleq \frac{\mathbf{x}}{\|\mathbf{x}\|}$ *denotes a vector of unit length.*

Moreover, this lower bound is achievable by IID inputs that can be expressed as the product of a constant unit vector $\hat{\mathbf{x}} \in \mathbb{C}^{n_\text{T}}$ *and a circularly symmetric, scalar, complex IID random process* $\{X_k\}$ *such that*

$$\log|X_k|^2 \sim \mathcal{U}\left([\log\log\mathcal{E}, \log\mathcal{E}]\right). \quad (46)$$

Note that this input satisfies the peak-power constraint (11) *(and therefore also the average-power constraint* (12)*).*

Proof. To derive a lower bound we choose a specific input distribution which naturally yields a lower bound to channel capacity. Let $\{\mathbf{X}_k\}$ be of the form

$$\mathbf{X}_k = X_k \cdot \hat{\mathbf{x}} \quad (47)$$

where $\hat{\mathbf{x}}$ is a deterministic unit vector (which is therefore known to both the receiver and transmitter) and where $\{X_k\}$ is an IID circularly symmetric random process with

$$\log|X_k|^2 \sim \mathcal{U}\left([\log x_{\min}^2, \log\mathcal{E}]\right), \quad (48)$$

where we choose $x_{\min}^2$ as

$$x_{\min}^2 = \log\mathcal{E}. \quad (49)$$

Fix some (large) positive integer κ and use the chain rule and the non-negativity of mutual information to obtain:

$$\frac{1}{n} I\big(\mathbf{X}_1^n; Y_1^n\big) = \frac{1}{n}\sum_{k=1}^{n} I\big(\mathbf{X}_k; Y_1^n \,\big|\, \mathbf{X}_1^{k-1}\big) \quad (50)$$

$$\geq \frac{1}{n}\sum_{k=\kappa+1}^{n-\kappa} I\big(\mathbf{X}_k; Y_1^n \,\big|\, \mathbf{X}_1^{k-1}\big). \quad (51)$$

Then for any $\kappa+1 \leq k \leq n-\kappa$, we can use the fact that $\{X_k\}$ is IID and circularly symmetric to lower bound $I\big(\mathbf{X}_k; Y_1^n \,\big|\, \mathbf{X}_1^{k-1}\big)$ as follows:

$$I\big(\mathbf{X}_k; Y_1^n \,\big|\, \mathbf{X}_1^{k-1}\big)$$
$$= I\big(X_k\hat{\mathbf{x}}; Y_1^n \,\big|\, \{X_\ell\hat{\mathbf{x}}\}_{\ell=1}^{k-1}\big) \quad (52)$$
$$= I\big(X_k\hat{\mathbf{x}}; \{X_\ell\hat{\mathbf{x}}\}_{\ell=1}^{k-1}, Y_1^n\big) \quad (53)$$
$$= I\big(X_k; X_1^{k-1}, Y_1^n\big) \quad (54)$$
$$\geq I\big(X_k; X_{k-\kappa}^{k-1}, Y_{k-\kappa}^{k-1}, Y_k\big) \quad (55)$$
$$= I\big(X_k; X_{k-\kappa}^{k-1}, Y_{k-\kappa}^{k-1}, Z_{k-\kappa}^{k-1}, Y_k\big) - \underbrace{I\big(X_k; Z_{k-\kappa}^{k-1} \,\big|\, X_{k-\kappa}^{k-1}, Y_{k-\kappa}^{k-1}, Y_k\big)}_{\leq \epsilon(x_{\min},\kappa)} \quad (56)$$
$$\geq I\big(X_k; X_{k-\kappa}^{k-1}, Y_{k-\kappa}^{k-1}, Z_{k-\kappa}^{k-1}, Y_k\big) - \epsilon(x_{\min},\kappa) \quad (57)$$
$$= I\big(X_k; \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=k-\kappa}^{k-1}, Y_k\big) - \epsilon(x_{\min},\kappa) \quad (58)$$
$$= I\big(X_{\kappa+1}; \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=1}^{\kappa}, Y_{\kappa+1}\big) - \epsilon(x_{\min},\kappa) \quad (59)$$
$$= I\big(X_{\kappa+1}; Y_{\kappa+1} \,\big|\, \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=1}^{\kappa}\big) - \epsilon(x_{\min},\kappa) \qquad \kappa+1 \leq k \leq n-\kappa. \quad (60)$$

Here the second equality follows because $\{X_k\}$ is chosen to be IID; in (55) we drop some arguments which reduces the mutual information; next we use the chain rule; in the subsequent inequality we lower bound the second term by $-\epsilon(x_{\min},\kappa)$ which can be shown to depend on $x_{\min}$ and κ only and to tend to zero as $x_{\min} \uparrow \infty$ (details omitted); in the subsequent equality we use $X_{k-\kappa}^{k-1}$ and $Z_{k-\kappa}^{k-1}$ in order to extract $\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=k-\kappa}^{k-1}$ from $Y_{k-\kappa}^{k-1}$ and then drop $(\{X_\ell, Y_\ell, Z_\ell\}_{\ell=k-\kappa}^{k-1})$ since given $\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=k-\kappa}^{k-1}$ it is independent of the other random variables; and the equality before last follows from stationarity.

Plugging (60) into (51) we get

$$\frac{1}{n} I\big(X_1^n; Y_1^n\big)$$
$$\geq \frac{1}{n}\sum_{k=\kappa+1}^{n-\kappa} \Big(I\big(X_{\kappa+1}; Y_{\kappa+1} \,\big|\, \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=1}^{\kappa}\big) - \epsilon(x_{\min},\kappa)\Big) \quad (61)$$
$$= \left(1 - \frac{2\kappa}{n}\right)\Big(I\big(X_{\kappa+1}; Y_{\kappa+1} \,\big|\, \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=1}^{\kappa}\big) - \epsilon(x_{\min},\kappa)\Big). \quad (62)$$

Letting n tend to infinity we obtain

$$\mathsf{C}(\text{SNR}) \geq I\big(X_{\kappa+1}; Y_{\kappa+1} \,\big|\, \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=1}^{\kappa}\big) - \epsilon(x_{\min},\kappa) \quad (63)$$

where the first term on the RHS can be viewed as mutual information across a memoryless SISO fading channel with fading $\mathbf{H}_{\kappa+1}^\mathsf{T}\hat{\mathbf{x}}$ in the presence of the side-information $\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=1}^{\kappa}$.

We next let the power grow to infinity $\mathcal{E} \uparrow \infty$. Since the circularly symmetric law (48) achieves the fading number of IID SISO fading with side-information [1, Proposition 4.23], [2, Proposition 6.23] and since our choice (49) guarantees that $\epsilon(x_{\min},\kappa)$ tend to zero as $\mathcal{E} \uparrow \infty$ we obtain the bound

$$\chi(\{\mathbf{H}_k^\mathsf{T}\}) \geq \chi_{\text{IID}}\big(\mathbf{H}_{\kappa+1}^\mathsf{T}\hat{\mathbf{x}} \,\big|\, \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=1}^{\kappa}\big) \quad (64)$$
$$= \chi_{\text{IID}}\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}} \,\big|\, \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=-\kappa}^{-1}\big) \quad (65)$$
$$= \log\pi + \mathsf{E}\big[\log|\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}|^2\big] - h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}} \big| \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=-\kappa}^{-1}\big). \quad (66)$$

Finally, we let κ go to infinity. The theorem now follows by choosing $\hat{\mathbf{x}}$ such as to maximize the fading number. □

4. SPECIAL CASE OF ISOTROPICALLY DISTRIBUTED FADING

We next consider the special case of isotropically distributed fading processes, *i.e.*, for every deterministic unitary $n_\text{T} \times n_\text{T}$ matrix U

$$\mathbf{H}_k \overset{\mathscr{L}}{=} \mathsf{U}\mathbf{H}_k, \tag{67}$$

where we use "$\overset{\mathscr{L}}{=}$" to denote *equal in law.*

In this case we have the following corollary:

Corollary 3. *Consider a MISO fading channel with memory* (6) *where the stationary and ergodic fading process* $\{\mathbf{H}_k\}$ *takes value in* $\mathbb{C}^{n_\text{T}}$*, satisfies* $h(\{\mathbf{H}_k\}) > -\infty$ *and* $\mathsf{E}\big[\|\mathbf{H}_k\|^2\big] < \infty$*, and is* isotropically distributed. *Then the upper bound* (17) *and the lower bound* (45) *coincide and the fading number* $\chi_\text{iso}(\{\mathbf{H}_k^\mathsf{T}\})$ *is given by*

$$\chi_\text{iso}\big(\{\mathbf{H}_k^\mathsf{T}\}\big) = \log\pi + \mathsf{E}\big[\log|\mathbf{H}_0^\mathsf{T}\hat{\mathbf{e}}|^2\big] - h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{e}}\,\big|\,\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{e}}\}_{\ell=-\infty}^{-1}\big) \tag{68}$$

where $\hat{\mathbf{e}}$ *is any deterministic unit vector.*

Proof. This follows immediately from Theorem 1 and 2 by noting that for any $\hat{\mathbf{e}}$

$$\mathbf{H}_k^\mathsf{T}\hat{\mathbf{e}} \overset{\mathscr{L}}{=} \mathbf{H}_k^\mathsf{T}\mathsf{U}^\mathsf{T}\hat{\mathbf{e}} = \mathbf{H}_k^\mathsf{T}\hat{\mathbf{e}}', \tag{69}$$

where the first equality in law follows from (67) and the second equality by defining a new unit vector $\hat{\mathbf{e}}' \triangleq \mathsf{U}^\mathsf{T}\hat{\mathbf{e}}$. Note that for the MISO case *isotropically distributed* is equivalent to *rotation commutative in the generalized sense* as defined in [1, Definition 4.37], [2, Definition 6.37]. □

5. SPECIAL CASE OF GAUSSIAN FADING

In this section we assume that the fading process $\{\mathbf{H}_k\}$ is a mean-$\mathbf{d}$ Gaussian process such that $\{\tilde{\mathbf{H}}_k\} = \{\mathbf{H}_k - \mathbf{d}\}$ is a zero-mean, circularly symmetric, stationary, ergodic, complex Gaussian process with matrix-valued spectral distribution function $\mathsf{F}(\cdot)$, and with covariance matrix K. Furthermore, we assume that the prediction error covariance matrix $\boldsymbol{\Sigma}$ is non-singular (regularity assumption).

5.1 Upper Bound for Gaussian Fading

We start with a new derivation of the upper bound (4) based on Theorem 1. We will see that (4) is in general less tight than (17).

In [1, Corollary 4.28], [2, Corollary 6.28] it has been shown that the IID MISO fading number (2) for Gaussian fading is given by

$$\chi(\mathbf{H}^\mathsf{T}) = \sup_{\|\hat{\mathbf{x}}\|=1}\big\{\log\pi + \mathsf{E}\big[\log|\mathbf{H}^\mathsf{T}\hat{\mathbf{x}}|^2\big] - h(\mathbf{H}^\mathsf{T}\hat{\mathbf{x}})\big\} \tag{70}$$
$$= -1 + \log d_*^2 - \text{Ei}\big(-d_*^2\big) \tag{71}$$

where d_* is given in (5). This proves the equivalence of the first supremum in (17) with the first three terms of (4). It therefore only remains to prove that

$$\sup_{\hat{\mathbf{x}}_{-\infty}^0} I\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0;\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=-\infty}^{-1}\big) \le \log\frac{\|\mathsf{K}\|}{\sigma_\text{min}}, \tag{72}$$

where σ_min is the smallest eigenvalue of the prediction error covariance matrix $\boldsymbol{\Sigma}$. To this goal note that

$$\begin{aligned}
&\sup_{\hat{\mathbf{x}}_{-\infty}^0} I\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0;\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=-\infty}^{-1}\big)\\
&\quad\le \sup_{\hat{\mathbf{x}}_{-\infty}^0} I\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0;\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=-\infty}^{-1},\mathbf{H}_{-\infty}^{-1}\big) &(73)\\
&\quad= \sup_{\hat{\mathbf{x}}_0} I\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0;\mathbf{H}_{-\infty}^{-1}\big) &(74)\\
&\quad= \sup_{\hat{\mathbf{x}}_0}\Big\{h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0\big) - h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0\,\big|\,\mathbf{H}_{-\infty}^{-1}\big)\Big\} &(75)\\
&\quad= \sup_{\hat{\mathbf{x}}_0}\Big\{\log\Big(\pi e\hat{\mathbf{x}}_0^\dagger\mathsf{K}\hat{\mathbf{x}}\Big) - h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0\,\big|\,\mathbf{H}_{-\infty}^{-1}\big)\Big\} &(76)\\
&\quad\le \sup_{\hat{\mathbf{x}}_0}\log\Big(\pi e\hat{\mathbf{x}}_0^\dagger\mathsf{K}\hat{\mathbf{x}}_0\Big) - \inf_{\hat{\mathbf{x}}_0} h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0\,\big|\,\mathbf{H}_{-\infty}^{-1}\big). &(77)
\end{aligned}$$

Here, the first inequality follows from the inclusion of additional random variables in the mutual information; the subsequent equality from the fact that given the past realization of the fading, $\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0$ is independent of $\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=-\infty}^{-1}$; and in the second last step we have used the expression for the differential entropy of a Gaussian random variable with K denoting the covariance matrix of $\{\mathbf{H}_k\}$.

Note that the first inequality in general is not tight, *i.e.*, (17) is in general tighter than (4).

To compute the second term in (77), we express the fading $\mathbf{H}_0$ as

$$\mathbf{H}_0 = \bar{\mathbf{H}}_0 + \check{\mathbf{H}}_0 \tag{78}$$

with $\bar{\mathbf{H}}_0$ being the best estimate of $\mathbf{H}_0$ based on the past realizations. Note that the remaining error $\check{\mathbf{H}}_0$ is independent of $\bar{\mathbf{H}}_0$ and $\mathbf{H}_{-\infty}^{-1}$ and that $\check{\mathbf{H}}_0 \sim \mathcal{N}_\mathbb{C}(\mathbf{0},\boldsymbol{\Sigma})$ where $\boldsymbol{\Sigma}$ denotes the prediction error covariance matrix. Hence

$$\begin{aligned}
h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0\,\big|\,\mathbf{H}_{-\infty}^{-1}\big) &= h\big((\bar{\mathbf{H}}_0^\mathsf{T}+\check{\mathbf{H}}_0^\mathsf{T})\hat{\mathbf{x}}_0\,\big|\,\mathbf{H}_{-\infty}^{-1}\big) &(79)\\
&= h\big(\check{\mathbf{H}}_0^\mathsf{T}\hat{\mathbf{x}}_0\,\big|\,\mathbf{H}_{-\infty}^{-1}\big) &(80)\\
&= h\big(\check{\mathbf{H}}_0^\mathsf{T}\hat{\mathbf{x}}_0\big) &(81)\\
&= \log\Big(\pi e\hat{\mathbf{x}}_0^\dagger\boldsymbol{\Sigma}\hat{\mathbf{x}}_0\Big). &(82)
\end{aligned}$$

The bound (72) now follows by the Rayleigh-Ritz Theorem [12, Theorem 4.2.2], [2, Theorem A.9]

$$\sigma_\text{min} = \min_{\hat{\mathbf{x}}_0}\hat{\mathbf{x}}_0^\dagger\boldsymbol{\Sigma}\hat{\mathbf{x}}_0, \tag{83}$$

and the definition of the Euclidean norm of matrices combined with Rayleigh-Ritz:

$$\max_{\hat{\mathbf{x}}_0}\hat{\mathbf{x}}_0^\dagger\mathsf{K}\hat{\mathbf{x}}_0 = \|\mathsf{K}\|. \tag{84}$$

5.2 Spatially IID Gaussian Fading

We next specialize the assumptions to the case where $\{\tilde{\mathbf{H}}_k\} = \{\mathbf{H}_k - \mathbf{d}\}$ is a spatially IID process where each component is a zero-mean unit-variance circularly symmetric complex Gaussian process of spectral distribution function $\mathsf{F}(\cdot)$. For this case we will now present a new derivation of the result (3) based on our new bounds.

Note that we cannot apply Corollary 3 here: even though $\{\tilde{\mathbf{H}}_k\}$ is isotropically distributed, $\{\mathbf{H}_k\}$ is not due to its mean vector $\mathbf{d}$.

However, the term $I\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0;\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=-\infty}^{-1}\big)$ does not depend on the particular choice of $\hat{\mathbf{x}}_\ell$:

$$\begin{aligned}
&I\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0;\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=-\infty}^{-1}\big)\\
&\quad= I\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0 - \mathbf{d}^\mathsf{T}\hat{\mathbf{x}}_0;\{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell - \mathbf{d}^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=-\infty}^{-1}\big) &(85)
\end{aligned}$$

$$= I\big(\tilde{\mathbf{H}}_0^{\mathsf{T}}\hat{\mathbf{x}}_0; \{\tilde{\mathbf{H}}_\ell^{\mathsf{T}}\hat{\mathbf{x}}_\ell\}_{\ell=-\infty}^{-1}\big) \tag{86}$$

$$= I\big(\tilde{\mathbf{H}}_0^{\mathsf{T}}\hat{\mathbf{e}}; \{\tilde{\mathbf{H}}_\ell^{\mathsf{T}}\hat{\mathbf{e}}\}_{\ell=-\infty}^{-1}\big) \tag{87}$$

$$= I\big(H_0^{(1)}; \{H_\ell^{(1)}\}_{\ell=-\infty}^{-1}\big) \tag{88}$$

$$= \log \frac{1}{\epsilon^2}. \tag{89}$$

Equation (3) now follows from (71), Theorem 1, and Theorem 2 by noting that

$$\max_{\|\hat{\mathbf{x}}\|=1} \frac{\left|\mathsf{E}\left[\mathbf{H}_k^{\mathsf{T}}\right]\hat{\mathbf{x}}\right|}{\sqrt{\mathsf{Var}(\mathbf{H}_k^{\mathsf{T}}\hat{\mathbf{x}})}} = \max_{\|\hat{\mathbf{x}}\|=1} \left|\mathbf{d}^{\mathsf{T}}\hat{\mathbf{x}}\right| = \|\mathbf{d}\|, \tag{90}$$

where the maximum is achieved for $\hat{\mathbf{x}} = \mathbf{d}/\|\mathbf{d}\|$.

6. DISCUSSION & CONCLUSION

We have derived two new bounds for a MISO fading channel of general law including memory. Both bounds show the same structure involving the maximization of a deterministic beam-direction $\hat{\mathbf{x}}$, which suggests that beam-forming is optimal at high SNR. However, one has to be aware that the beam-direction is not chosen to maximize the SNR, but to maximize the fading number.

We would like to emphasize that even though this is an asymptotic result for the theoretical situation of infinite power, it still is of relevance for finite SNR values: it has been shown that the approximation

$$\mathsf{C}(\text{SNR}) \approx \log(1 + \log(1 + \text{SNR})) + \chi \tag{91}$$

holds already for moderate values of the SNR. Hence, the fading number is an indicator of the maximal rate at which power efficient communication is achievable on the channel. For a further discussion about the practical relevance of the fading number we also refer to [11] and [10].

The differences between the upper and lower bound lies in the details of the maximization: while in the lower bound one single direction unit vector $\hat{\mathbf{x}}$ is chosen for all time

$$\chi\big(\{\mathbf{H}_k^{\mathsf{T}}\}\big) \ge \sup_{\hat{\mathbf{x}}} \Big\{ \log\pi + \mathsf{E}\big[\log|\mathbf{H}_0^{\mathsf{T}}\hat{\mathbf{x}}|^2\big] - h(\mathbf{H}_0^{\mathsf{T}}\hat{\mathbf{x}}) + I\big(\mathbf{H}_0^{\mathsf{T}}\hat{\mathbf{x}}; \{\mathbf{H}_\ell^{\mathsf{T}}\hat{\mathbf{x}}\}_{\ell=-\infty}^{-1}\big)\Big\}, \tag{92}$$

the upper bound allows for different realizations of $\hat{\mathbf{x}}_k$ for different times k

$$\chi\big(\{\mathbf{H}_k^{\mathsf{T}}\}\big) \le \sup_{\hat{\mathbf{x}}} \Big\{ \log\pi + \mathsf{E}\big[\log|\mathbf{H}_0^{\mathsf{T}}\hat{\mathbf{x}}|^2\big] - h(\mathbf{H}_0^{\mathsf{T}}\hat{\mathbf{x}})\Big\} + \sup_{\hat{\mathbf{x}}_{-\infty}^{0}} I\big(\mathbf{H}_0^{\mathsf{T}}\hat{\mathbf{x}}_0; \{\mathbf{H}_\ell^{\mathsf{T}}\hat{\mathbf{x}}_\ell\}_{\ell=-\infty}^{-1}\big). \tag{93}$$

Moreover, in the upper bound the maximization is split up into two parts, one part maximizing the memoryless IID MISO fading number, and another part maximizing a term that contains the contribution of the memory.

We are convinced that the lower bound is actually tight for several reasons: intuition tells that for our stationary channel model a stationary input should be sufficient for achieving the capacity. As a matter of fact in the SISO and SIMO case it has been shown that actually an IID input suffices to achieve capacity at high SNR [1], [2], [9].

The splitting of the maximization into two separate maximizations seems to be an artifact of our derivation of the upper bound. Note that for any functions $f(\cdot)$ and $g(\cdot)$

$$\sup_x \big\{f(x) + g(x)\big\} \le \sup_x f(x) + \sup_x g(x). \tag{94}$$

Currently, attempts to get rid of this artifact are under study.

In the case of isotropically distributed fading the particular choice of direction has no influence on the fading process and therefore the upper and lower bounds coincide.

In the case of Gaussian fading we could show that the bounds presented in [10] and [11] are special cases of the new bounds presented here, where the new upper bound (17) is in general tighter than (4).

7. ACKNOWLEDGMENTS

Helpful comments from Amos Lapidoth and Tobias Koch are gratefully acknowledged.

8. REFERENCES

[1] A. Lapidoth and S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat fading channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2426–2467, Oct. 2003.

[2] S. M. Moser, "Duality-based bounds on channel capacity," Ph.D. dissertation, Swiss Federal Institute of Technology, Zurich, Oct. 2004, Diss. ETH No. 15769. [Online]. Available: http://moser.cm.nctu.edu.tw/

[3] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov.–Dec. 1999.

[4] A. Lapidoth, "On the high SNR capacity of stationary Gaussian fading channels," in *Proc. 41st Allerton Conf. Comm., Contr. and Comp.*, Allerton H., Monticello, Il, Oct. 1–3, 2003, pp. 410–419.

[5] T. Koch, "On the asymptotic capacity of multiple-input single-output fading channels with memory," Master's thesis, Signal and Inform. Proc. Lab., ETH Zurich, Switzerland, Apr. 2004, supervised by Prof. Dr. Amos Lapidoth.

[6] A. Lapidoth, "On the asymptotic capacity of stationary Gaussian fading channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 437–446, Feb. 2005.

[7] Y. Liang and V. V. Veeravalli, "Capacity of noncoherent time-selective Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3095–3110, Dec. 2004.

[8] A. Lapidoth and S. M. Moser, "The fading number of SIMO fading channels with memory," in *Proc. IEEE Int. Symp. Inf. Theory and its Appl.*, Parma, Italy, Oct. 10–13, 2004, pp. 287–292.

[9] ——, "The fading number of single-input multiple-output fading channels with memory," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 437–453, Feb. 2006.

[10] T. Koch and A. Lapidoth, "Degrees of freedom in non-coherent stationary MIMO fading channels," in *Proc. Winter School Cod. and Inform. Theory*, Bratislava, Slovakia, Feb. 20–25, 2005, pp. 91–97.

[11] ——, "The fading number and degrees of freedom in non-coherent MIMO fading channels: a peace pipe," in *Proc. IEEE Int. Symp. Inf. Theory*, Adelaide, Australia, Sept. 4–9, 2005, pp. 661–665.

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