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

Stefan M. Moser[1]

Department of Communication Engineering
National Chiao Tung University (NCTU)
Hsinchu, Taiwan
Email: stefan.moser@ieee.org

Abstract— We derive new upper and lower bounds on the fading number of non-coherent 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.

I. INTRODUCTION

It has been recently shown in [1], [2] that, whenever the matrix-valued fading process is of finite differential entropy rate, the capacity of non-coherent 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 with *non-regular* channel law, *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 receive and transmit 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 non-coherent channels with 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

$$\mathsf{C}(\textsc{snr}) = \log(1 + \log(1 + \textsc{snr})) + \chi + o(1) \qquad (1)$$

where $o(1)$ tends to zero as the SNR tends to infinity, and where χ is a constant, denoted *fading number*, that does not depend on the SNR.

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] and the single-input multiple-output (SIMO) case is derived in [8], [2].

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

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

Here $\mathbf{H}^\mathsf{T}$ denotes the transpose of the vector $\mathbf{H}$. This fading number is achievable by inputs that can be expressed as the product of a constant unit vector in $\mathbb{C}^{n_\mathsf{T}}$ and a circularly symmetric, scalar, complex random variable of the same law that achieves the memoryless SISO fading number [1], [2]

$$\chi_{\text{IID}}(H) = \log \pi + \mathsf{E}\left[\log |H|^2\right] - h(H), \qquad (3)$$

and the memoryless SISO fading number with partial[2] receiver side-information S

$$\chi_{\text{IID}}(H|S) = \log \pi + \mathsf{E}\left[\log |H|^2\right] - h(H|S). \qquad (4)$$

---

[2]A random variable S is said to contain only *partial* information about H if $h(H|S) > -\infty$. Since we assume that $\mathsf{E}\left[|H|^2\right] < \infty$, this is equivalent to assuming $I(S; H) < \infty$.

Hence, the asymptotic capacity of a memoryless MISO fading channel is achieved by beam-forming where the beam-direction is chosen not to maximize the SNR, but the fading number.

In [9], [10] 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 (spatially IID) with each component being a zero-mean, unit-variance, circularly symmetric, complex Gaussian process, the fading number is shown to be[3]

$$\chi_{\text{Gauss, spatially IID}}(\{\mathbf{H}_k^\mathsf{T}\})$$
$$= -1 + \log \|\mathbf{d}\|^2 - \text{Ei}(-\|\mathbf{d}\|^2) + \log \frac{1}{\epsilon^2}, \quad (5)$$

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 Σ is non-singular (regularity assumption) they show that

$$\chi_{\text{Gauss}}(\{\mathbf{H}_k^\mathsf{T}\}) \leq -1 + \log d_*^2 - \text{Ei}(-d_*^2) + \log \frac{\|\mathsf{K}\|}{\sigma_{\min}}, \quad (6)$$

where

$$d_* = \max_{\|\hat{\mathbf{x}}\|=1} \frac{|\mathbf{d}^\mathsf{T}\hat{\mathbf{x}}|}{\sqrt{\text{Var}(\mathbf{H}_k^\mathsf{T}\hat{\mathbf{x}})}}; \quad (7)$$

$\sigma_{\min}$ denotes the smallest eigenvalue of Σ; 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 (*i.e.*, not necessarily Gaussian) fading channels.

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

We then specialize these results to the case of isotropically distributed fading processes in Section IV and to Gaussian fading in Section V. 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 VI.

II. THE CHANNEL MODEL

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

$$Y_k = \mathbf{H}_k^\mathsf{T}\mathbf{x}_k + Z_k \quad (8)$$

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

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

As for the multi-variate fading process $\{\mathbf{H}_k\}$, we shall only assume that it is stationary, ergodic, of finite second moment $\mathsf{E}[\|\mathbf{H}_k\|^2] < \infty$, and of finite differential entropy rate

$$h(\{\mathbf{H}_k\}) \triangleq \lim_{n \uparrow \infty} \frac{1}{n} h(\mathbf{H}_1, \ldots, \mathbf{H}_n) > -\infty \quad (9)$$

(the *regularity* assumption). Neither transmitter nor receiver knows the realization of the fading vector; they only know its law.

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

$$\text{SNR} \triangleq \frac{\mathcal{E}}{\sigma^2}. \quad (10)$$

The capacity $\mathsf{C}(\text{SNR})$ of the channel (8) is given by

$$\mathsf{C}(\text{SNR}) = \lim_{n \to \infty} \frac{1}{n} \sup I(\mathbf{X}_1^n; Y_1^n) \quad (11)$$

where the supremum is over the set of all probability distributions on $\mathbf{X}_1^n$ satisfying the constraints, *i.e.*,

$$\|\mathbf{X}_k\|^2 \leq \mathcal{E}, \quad \text{almost surely}, \quad k = 1, 2, \ldots, n \quad (12)$$

for a peak-power constraint, or

$$\frac{1}{n} \sum_{k=1}^n \mathsf{E}[\|\mathbf{X}_k\|^2] \leq \mathcal{E} \quad (13)$$

for an average-power constraint.

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

$$\overline{\lim_{\text{SNR} \uparrow \infty}} \left\{ \mathsf{C}(\text{SNR}) - \log \log \text{SNR} \right\} < \infty. \quad (14)$$

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

$$\chi(\{\mathbf{H}_k^\mathsf{T}\}) \triangleq \overline{\lim_{\text{SNR} \uparrow \infty}} \left\{ \mathsf{C}(\text{SNR}) - \log \log \text{SNR} \right\}. \quad (15)$$

Prima facie the fading number depends on whether a peak-power constraint (12) or an average-power constraint (13) 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.

III. MAIN RESULTS

We first state a new upper bound to the fading number of a MISO fading channel:

Theorem 1: Consider a non-coherent MISO fading channel with memory (8) where the stationary and ergodic fading process $\{\mathbf{H}_k\}$ takes value in $\mathbb{C}^{n_\mathrm{T}}$ and satisfies $h(\{\mathbf{H}_k\}) > -\infty$ and $\mathsf{E}\left[\|\mathbf{H}_k\|^2\right] < \infty$. Then, irrespective of whether a peak-power constraint (12) or an average-power constraint (13) is imposed on the input, the fading number $\chi(\{\mathbf{H}_k^\mathsf{T}\})$ is upper-bounded by

$$\chi\big(\{\mathbf{H}_k^\mathsf{T}\}\big) \leq \sup_{\hat{\mathbf{x}}_{-\infty}^0} \Big\{ \log \pi + \mathsf{E}\big[\log |\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0|^2\big]$$
$$- h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0 \,\big|\, \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=-\infty}^{-1}\big) \Big\} \quad (16)$$

where $\hat{\mathbf{x}}_\ell \triangleq \frac{\mathbf{x}_\ell}{\|\mathbf{x}_\ell\|}$ denote a vectors of unit length.

Proof: The proof is rather long and in part pretty technical. We therefore give here only an outline.

Similar to the proof of the SIMO fading number with memory [8], [2] we need a lemma that limits the possible joint input distributions on $\mathbf{X}_1, \ldots, \mathbf{X}_n$ to such under which each random vector $\mathbf{X}_k$ has the same law with an average power equal to the constraint $\mathcal{E}$. Note that we are not allowed to assume a stationary input *a priori* because—even though it feels very natural—we are not aware of any proof showing that the capacity-achieving input distribution of a stationary channel is stationary. The lemma does not prove this either, but at least it allows us to restrict the input to have the same marginals.

Unfortunately, the proof is complicated by the fact that this lemma does not guarantee equal marginals for the time epochs k on the border of a block. However, these edge effects wash out once we let the blocklength n tend to infinity.

The proof then proceeds as follows: the mutual information between joint input and joint output is split up into a term describing the memoryless case and a term that takes care of the memory:

$$\lim_{n\to\infty} \frac{1}{n} I(\mathbf{X}_1^n; Y_1^n) \leq \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^n \Big(I(\mathbf{X}_k; Y_k)$$
$$+ I\big(\mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k; \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{X}}_\ell\}_{\ell=1}^{k-1} \,\big|\, \hat{\mathbf{X}}_1^k\big)\Big). \quad (17)$$

Here $\hat{\mathbf{X}}_\ell$ denotes the unit vector $\frac{\mathbf{X}_\ell}{\|\mathbf{X}_\ell\|}$. Note that the memory in the fading process cannot be fully exploited because the receiver has only one antenna, *i.e.*, it only gains the knowledge about the fading vector when it is multiplied with a beam-direction.

The first term in the sum is then further upper bounded as follows:

$$I(\mathbf{X}_k; Y_k) = I\big(R_k; \mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k R_k + Z_k\big)$$
$$+ I\big(\hat{\mathbf{X}}_k; \mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k R_k + Z_k \,\big|\, R_k\big) \quad (18)$$
$$\leq \mathsf{C}_{\mathrm{SISO,IID},H=\mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k}(\mathcal{E})$$
$$+ I\big(\hat{\mathbf{X}}_k; \mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k R_k + Z_k \,\big|\, R_k\big), \quad (19)$$

where R_k stands for $\|\mathbf{X}_k\|$ and where $\mathsf{C}_{\mathrm{SISO,IID},H=\mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k}(\mathcal{E})$ denotes the capacity of a memoryless SISO fading channel with fading $H = \mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k$. Note that due to the stationarity of $\{\mathbf{H}_k\}$ and to the fact that we may restrict ourselves to inputs with equal marginals, this SISO capacity actually does not depend on time. The total bound on the MISO capacity, however, still does:

$$\mathsf{C}(\mathcal{E}) \leq \lim_{n\to\infty} \sup_{\substack{Q_{\mathbf{X}_1^n} \\ \mathrm{eq.\ marg.}}} \frac{1}{n} \sum_{k=1}^n \Big(\mathsf{C}_{\mathrm{SISO,IID},H=\mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k}(\mathcal{E})$$
$$+ I\big(\hat{\mathbf{X}}_k; \mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k R_k + Z_k \,\big|\, R_k\big)$$
$$+ I\big(\mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k; \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{X}}_\ell\}_{\ell=1}^{k-1} \,\big|\, \hat{\mathbf{X}}_1^k\big)\Big). \quad (20)$$

This is partially mended in a next step: it is shown that the capacity-achieving input distribution can be written as IID blocks of a given fixed length $\kappa + 1$ independent of n. This allows us to get rid of the limit of n. We then let the power tend to infinity in order to be able to invoke the fading number of a SISO fading channel. Note that for this we need to swap a supremum and a limit which needs justification. The bound now looks as follows:

$$\chi(\{\mathbf{H}_k^\mathsf{T}\})$$
$$\leq \sup_{\substack{Q_{\hat{\mathbf{X}}_0^\kappa} \\ \mathrm{eq.\ marg.}}} \frac{1}{\kappa + 1} \sum_{j=0}^\kappa \Bigg(\log \pi$$
$$+ \mathsf{E}\Big[\log \big|\mathbf{H}_0^\mathsf{T}\hat{\mathbf{X}}_{(\kappa+j) \bmod (\kappa+1)}\big|^2\Big]$$
$$- h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{X}}_{(\kappa+j) \bmod (\kappa+1)} \,\big|\, \mathbf{H}_{-1}^\mathsf{T}\hat{\mathbf{X}}_{(\kappa-1+j) \bmod (\kappa+1)},$$
$$\ldots, \mathbf{H}_{-\kappa}^\mathsf{T}\hat{\mathbf{X}}_{(0+j) \bmod (\kappa+1)}, \hat{\mathbf{X}}_0^\kappa\big)\Bigg). \quad (21)$$

From this the result follows by replacing the supremum over the distribution of $\hat{\mathbf{X}}_0^\kappa$ by the supremum over all possible values of $\hat{\mathbf{x}}_0^\kappa$ and then let κ tend to infinity. ∎

Next we state a lower bound to the fading number:

Theorem 2: Consider a non-coherent MISO fading channel with memory (8) where the stationary and ergodic fading process $\{\mathbf{H}_k\}$ takes value in $\mathbb{C}^{n_\mathrm{T}}$ and satisfies $h(\{\mathbf{H}_k\}) > -\infty$ and $\mathsf{E}\left[\|\mathbf{H}_k\|^2\right] < \infty$. Then the fading number $\chi\big(\{\mathbf{H}_k^\mathsf{T}\}\big)$ is lower-bounded by

$$\chi\big(\{\mathbf{H}_k^\mathsf{T}\}\big) \geq \sup_{\hat{\mathbf{x}}} \Big\{ \log \pi + \mathsf{E}\big[\log |\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}|^2\big]$$
$$- h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}} \,\big|\, \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=-\infty}^{-1}\big) \Big\} \quad (22)$$

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 $\hat{\mathbf{x}} \in \mathbb{C}^{n_\mathrm{T}}$ and a circularly symmetric, scalar, complex IID random process $\{X_k\}$ such that

$$\log |X_k|^2 \sim \mathcal{U}\left([\log \log \mathcal{E}, \log \mathcal{E}]\right). \quad (23)$$

Note that this input satisfies the peak-power constraint (12) (and therefore also the average-power constraint (13)).

Proof: We only give an outline of the proof. The details are omitted.

The lower bound is based on the assumption of a specific input distribution which is chosen to be of the form

$$\mathbf{X}_k = X_k \cdot \hat{\mathbf{x}} \tag{24}$$

where $\hat{\mathbf{x}}$ is a deterministic unit vector (the beam-direction) and where $\{X_k\}$ is IID circularly symmetric satisfying (23). This choice is motivated by our intuition which tells that for a stationary channel model a stationary input should be sufficient for achieving the capacity and by the fact that in the SISO and SIMO case such an IID input actually suffices to achieve capacity at high SNR [1], [2], [8]. Hence, this choice for $\{X_k\}$ achieves the fading number of the SISO fading channel

$$Y_k = (\mathbf{H}_k^\mathsf{T}\hat{\mathbf{x}}) \cdot X_k + Z_k \tag{25}$$

with fading process $\{H_k\} = \{\mathbf{H}_k^\mathsf{T}\hat{\mathbf{x}}\}$. The lower bound is then derived by proving

$$\frac{1}{n} I(\mathbf{X}_1^n; Y_1^n) \approx \frac{1}{n} \sum_{k=1}^{n} I\big(X_k; Y_k \,\big|\, \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=1}^{k-1}\big). \tag{26}$$

Here the approximation is due to some correction terms that will tend to zero once we let the power tend to infinity. Therefore, letting the power grow to infinity and using the fading number (4) of a memoryless SISO fading channel with side-information, we get

$$\chi\big(\{\mathbf{H}_k^\mathsf{T}\}\big) \geq \chi_{\mathrm{IID}}\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}} \,\big|\, \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=-\infty}^{-1}\big). \tag{27}$$

∎

IV. Special Case of Isotropically Distributed Fading

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

$$\mathbf{H}_k \stackrel{\mathscr{L}}{=} \mathsf{U}\mathbf{H}_k, \tag{28}$$

where we use "$\stackrel{\mathscr{L}}{=}$" to denote *equal in law*. In this case we have the following corollary:

Corollary 3: Consider a non-coherent MISO fading channel with memory (8) where the stationary and ergodic fading process $\{\mathbf{H}_k\}$ takes value in $\mathbb{C}^{n_T}$, satisfies $h(\{\mathbf{H}_k\}) > -\infty$ and $\mathsf{E}\big[\|\mathbf{H}_k\|^2\big] < \infty$, and is *isotropically distributed*. Then the upper bound (16) and the lower bound (22) coincide and the fading number $\chi_{\mathrm{iso}}\big(\{\mathbf{H}_k^\mathsf{T}\}\big)$ is given by

$$\chi_{\mathrm{iso}}\big(\{\mathbf{H}_k^\mathsf{T}\}\big) = \log \pi + \mathsf{E}\big[\log |\mathbf{H}_0^\mathsf{T}\hat{\mathbf{e}}|^2\big] \\ - h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{e}} \,\big|\, \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{e}}\}_{\ell=-\infty}^{-1}\big) \tag{29}$$

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

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

$$\mathbf{H}_k^\mathsf{T}\hat{\mathbf{e}} \stackrel{\mathscr{L}}{=} \mathbf{H}_k^\mathsf{T}\mathsf{U}^\mathsf{T}\hat{\mathbf{e}} = \mathbf{H}_k^\mathsf{T}\hat{\mathbf{e}}' \tag{30}$$

where the first equality in law follows from (28) 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] or [2, Definition 6.37]. ∎

V. 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 $\{\mathbf{H}_k - \mathbf{d}\}$ 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 Σ is non-singular (regularity assumption).

A. Upper Bound for Gaussian Fading

We start with a new derivation of the upper bound (6) based on Theorem 1. We will see that (6) is in general less tight than (16).

We start by loosening the upper bound (16) as follows:

$$\chi\big(\{\mathbf{H}_k^\mathsf{T}\}\big)$$
$$\leq \sup_{\hat{\mathbf{x}}_{-\infty}^0} \Big\{ \log \pi + \mathsf{E}\big[\log |\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0|^2\big] - h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0\big)$$
$$+ h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0\big) - h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0 \,\big|\, \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=-\infty}^{-1}\big) \Big\} \tag{31}$$
$$\leq \sup_{\hat{\mathbf{x}}_0} \Big\{ \log \pi + \mathsf{E}\big[\log |\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0|^2\big] - h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0\big) \Big\}$$
$$+ \sup_{\hat{\mathbf{x}}_{-\infty}^0} I\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0; \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=-\infty}^{-1}\big). \tag{32}$$

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

$$\chi(\mathbf{H}^\mathsf{T}) = \sup_{\hat{\mathbf{x}}} \Big\{ \log \pi + \mathsf{E}\big[\log |\mathbf{H}^\mathsf{T}\hat{\mathbf{x}}|^2\big] - h\big(\mathbf{H}^\mathsf{T}\hat{\mathbf{x}}\big) \Big\} \tag{33}$$
$$= -1 + \log d_*^2 - \mathrm{Ei}\big(-d_*^2\big) \tag{34}$$

where d_* is given in (7). This proves the equivalence of the first supremum in (32) with the first three terms of (6). It therefore only remains to prove that

$$\sup_{\hat{\mathbf{x}}} I\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}; \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=-\infty}^{-1}\big) \leq \log \frac{\|\mathsf{K}\|}{\sigma_{\min}}, \tag{35}$$

where $\sigma_{\min}$ is the smallest eigenvalue of the prediction error covariance matrix Σ. To this goal note that

$$\sup_{\hat{\mathbf{x}}} I\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}; \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=-\infty}^{-1}\big)$$
$$\leq \sup_{\hat{\mathbf{x}}} I\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}; \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}\}_{\ell=-\infty}^{-1}, \mathbf{H}_{-\infty}^{-1}\big) \tag{36}$$
$$= \sup_{\hat{\mathbf{x}}} I\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}; \mathbf{H}_{-\infty}^{-1}\big) \tag{37}$$
$$= \sup_{\hat{\mathbf{x}}} \Big\{ h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}\big) - h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}} \,\big|\, \mathbf{H}_{-\infty}^{-1}\big) \Big\} \tag{38}$$
$$= \sup_{\hat{\mathbf{x}}} \Big\{ \log \big(\pi e \hat{\mathbf{x}}^\dagger \mathsf{K}\hat{\mathbf{x}}\big) - h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}} \,\big|\, \mathbf{H}_{-\infty}^{-1}\big) \Big\} \tag{39}$$
$$\leq \sup_{\hat{\mathbf{x}}} \log \big(\pi e \hat{\mathbf{x}}^\dagger \mathsf{K}\hat{\mathbf{x}}\big) - \inf_{\hat{\mathbf{x}}} h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}} \,\big|\, \mathbf{H}_{-\infty}^{-1}\big). \tag{40}$$

To compute the second term in (40), we express the fading $\mathbf{H}_0$ as $\mathbf{H}_0 = \bar{\mathbf{H}}_0 + \check{\mathbf{H}}_0$ with $\bar{\mathbf{H}}_0$ being the best estimate of $\mathbf{H}_0$

based on the past realizations. Note that the remaining error $\check{\mathbf{H}}_0 \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \Sigma)$ is independent of $\bar{\mathbf{H}}_0$ and $\mathbf{H}_{-\infty}^{-1}$. Hence

$$h\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}} \mid \mathbf{H}_{-\infty}^{-1}\big) = h\big((\bar{\mathbf{H}}_0^\mathsf{T} + \check{\mathbf{H}}_0^\mathsf{T})\hat{\mathbf{x}} \mid \mathbf{H}_{-\infty}^{-1}\big) \qquad (41)$$

$$= h\big(\check{\mathbf{H}}_0^\mathsf{T}\hat{\mathbf{x}} \mid \mathbf{H}_{-\infty}^{-1}\big) \qquad (42)$$

$$= h\big(\check{\mathbf{H}}_0^\mathsf{T}\hat{\mathbf{x}}\big) \qquad (43)$$

$$= \log\big(\pi e \hat{\mathbf{x}}^\dagger \Sigma \hat{\mathbf{x}}\big). \qquad (44)$$

The bound (35) now follows from the Rayleigh-Ritz Theorem [11, Theorem 4.2.2], [2, Theorem A.9]

$$\min_{\hat{\mathbf{x}}} \hat{\mathbf{x}}^\dagger \Sigma \hat{\mathbf{x}} = \sigma_{\min}, \qquad (45)$$

and the definition of the Euclidean norm of matrices

$$\max_{\hat{\mathbf{x}}} \hat{\mathbf{x}}^\dagger \mathsf{K}\hat{\mathbf{x}} = \max_{\hat{\mathbf{x}}} \hat{\mathbf{x}}^\dagger \mathsf{S}^\dagger \mathsf{S}\hat{\mathbf{x}} = \max_{\hat{\mathbf{x}}} \|\mathsf{S}\hat{\mathbf{x}}\|^2 = \|\mathsf{S}\|^2 = \|\mathsf{K}\|. \qquad (46)$$

Note that this bound is overly optimistic regarding the memory: the bound is based on the idea that the receiver can get some estimate on the channel state vector based on the knowledge of all past state vectors. It ignores the fact that we have a MISO channel where the receiver only has a single antenna and therefore never receives the channel state vector $\mathbf{H}_k$ as a whole, but only as a summed up version $\mathbf{H}_k^\mathsf{T}\hat{\mathbf{X}}_k$.

B. 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 (5) 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 $\hat{\mathbf{x}}_\ell$:

$$I\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0; \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=-\infty}^{-1}\big)$$
$$= I\big(\mathbf{H}_0^\mathsf{T}\hat{\mathbf{x}}_0 - \mathbf{d}^\mathsf{T}\hat{\mathbf{x}}_0; \{\mathbf{H}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell - \mathbf{d}^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=-\infty}^{-1}\big) \qquad (47)$$

$$= I\big(\tilde{\mathbf{H}}_0^\mathsf{T}\hat{\mathbf{x}}_0; \{\tilde{\mathbf{H}}_\ell^\mathsf{T}\hat{\mathbf{x}}_\ell\}_{\ell=-\infty}^{-1}\big) \qquad (48)$$

$$= I\big(\tilde{\mathbf{H}}_0^\mathsf{T}\hat{\mathbf{e}}; \{\tilde{\mathbf{H}}_\ell^\mathsf{T}\hat{\mathbf{e}}\}_{\ell=-\infty}^{-1}\big) \qquad (49)$$

$$= I\big(H_0^{(1)}; \{H_\ell^{(1)}\}_{\ell=-\infty}^{-1}\big) \qquad (50)$$

$$= \log\frac{1}{\epsilon^2}. \qquad (51)$$

Eq. (5) now follows from (34), Theorems 1 and 2 by noting that

$$\max_{\|\hat{\mathbf{x}}\|=1} \frac{|\mathbf{d}^\mathsf{T}\hat{\mathbf{x}}|}{\sqrt{\mathsf{Var}(\mathbf{H}_k^\mathsf{T}\hat{\mathbf{x}})}} = \max_{\|\hat{\mathbf{x}}\|=1} |\mathbf{d}^\mathsf{T}\hat{\mathbf{x}}| = \|\mathbf{d}\|. \qquad (52)$$

VI. DISCUSSION & CONCLUSION

We have derived two 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.

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

We are convinced that the lower bound is actually tight: 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 suffies to achieve capacity at high SNR [1], [2], [8]. Furthermore, considering the symmetry in the upper bound (21), the last step in the proof of the upper bound seems not tight. Unfortunately, we haven't been able yet to show that the supremum in (21) is achieved by an IID input.

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 [9] and [10] are special cases of the new bounds presented here, where the new upper bound (16) is in general tighter than (6).

The success of further attempts on deriving the MISO fading number precisely will be crucial to the investigation of the fading number of general MIMO fading channels.

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 single-input multiple-output fading channels with memory," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 437–453, Feb. 2006.

[9] 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.

[10] ——, "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.

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