International Symposium on Information Theory and its Applications, ISITA2004
Parma, Italy, October 10–13, 2004

The Fading Number of SIMO Fading Channels with Memory

Amos Lapidoth Stefan M. Moser

Signal and Information Processing Laboratory
Department of Information Technology and Electrical Engineering
Swiss Federal Institute of Technology (ETH) Zurich
CH-8092 Zurich, Switzerland
e-mail: {lapidoth,moser}@isi.ee.ethz.ch

Abstract

We derive the fading number of a general (not necessarily Gaussian) single-input multiple-output (SIMO) fading channel with memory, where the transmitter and receiver—while fully cognizant of the probability law governing the fading process—have no access to the fading realization.

It is demonstrated that the fading number is achieved by IID circularly-symmetric inputs of log squared-magnitude that is uniformly distributed over a signal-to-noise (SNR) dependent interval. The upper limit of the interval is the logarithm of the allowed transmit power, and the lower limit tends to infinity sub-logarithmically in the SNR. Among the new ingredients in the proof is a new theorem regarding input distributions that escape to infinity.

Upper and lower bounds on the fading number for SIMO Gaussian fading are also presented. Those are computed explicitly for stationary m-th order auto-regressive AR(m) Gaussian fading processes.

Keywords: Auto-regressive process, channel capacity, fading, fading number, high SNR, memory, multiple-antenna, SIMO.

1. INTRODUCTION

We consider a single-input multiple-output (SIMO) fading channel whose time-k output $\mathbf{Y}_k \in \mathbb{C}^{n_\mathrm{R}}$ is given by

$$\mathbf{Y}_k = \mathbf{H}_k x_k + \mathbf{Z}_k \tag{1}$$

where $x_k \in \mathbb{C}$ denotes the time-k channel input; the random vector $\mathbf{H}_k \in \mathbb{C}^{n_\mathrm{R}}$ denotes the time-k fading vector; and where $\mathbf{Z}_k$ denotes additive noise. Here $\mathbb{C}$ denotes the complex field, $\mathbb{C}^{n_\mathrm{R}}$ denotes the n_R-dimensional complex Euclidean space, and n_R denotes the number of receive antennas. We assume that the additive noise is a zero-mean temporally and spatially white Gaussian process of covariance matrix $\sigma^2 \mathsf{I}_{n_\mathrm{R}}$, where $\sigma^2 > 0$ and where $\mathsf{I}_{n_\mathrm{R}}$ denotes the $n_\mathrm{R} \times n_\mathrm{R}$ identity matrix. Thus, $\{\mathbf{Z}_k\}$ is a zero-mean circularly-symmetric stationary multi-variate Gaussian process such that $\mathsf{E}\left[\mathbf{Z}_k \mathbf{Z}_{k+m}^\dagger\right]$ is the zero matrix if $m \neq 0$, and is $\sigma^2 \mathsf{I}_{n_\mathrm{R}}$ for $m = 0$. Here $()^\dagger$ denotes Hermitian conjugation.

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

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

and of finite differential entropy rate

$$h(\{\mathbf{H}_k\}) > -\infty. \tag{3}$$

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

As for the input, we consider two different constraints: a peak-power constraint and an average-power constraint. We use $\mathcal{E}_\mathrm{s}$ 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

$$\mathrm{SNR} = \frac{\mathcal{E}_\mathrm{s}}{\sigma^2}.$$

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

$$\mathsf{C}(\mathrm{SNR}) = \lim_{n\to\infty} \frac{1}{n} \sup I\left(X_1^n; \mathbf{Y}_1^n\right),$$

where we use X_j^k to denote $X_j, \ldots, X_k$ and where the supremum is over the set of all probability distributions on X_1^n satisfying the constraint, *i.e.*,

$$|X_k|^2 \leq \mathcal{E}_\mathrm{s}, \quad k = 1, 2, \ldots, n \quad \text{for a peak constraint,}$$

The work of S. M. Moser was supported in part by the ETH under TH-23 02-2.

or

$$\frac{1}{n}\sum_{k=1}^{n} \mathsf{E}\left[|X_k|^2\right] \le \mathcal{E}_{\mathrm{s}} \quad \text{for an average constraint.}$$

It follows from the general results of [1] that—contrary to the case where the receiver has access to the fading realization—at high SNR the capacity of this channel grows only double-logarithmically in the SNR

$$\varlimsup_{\mathrm{SNR}\uparrow\infty} \Big\{ \mathsf{C}(\mathrm{SNR}) - \log\log \mathrm{SNR} \Big\} < \infty.$$

In an attempt to quantify the rates at which the double-logarithmic asymptotics begin, [1] introduced the fading number χ as

$$\chi(\{\mathbb{H}_k\}) = \varlimsup_{\mathrm{SNR}\uparrow\infty} \Big\{ \mathsf{C}(\mathrm{SNR}) - \log\log \mathrm{SNR} \Big\}. \tag{4}$$

Among the fading numbers computed in [1] are the fading number for single-input single-output (SISO) fading channels with memory

$$\chi(\{H_k\}) = \log\pi + \mathsf{E}\left[\log|H_1|^2\right] - h\big(\{H_k\}\big), \tag{5}$$

where $h\big(\{H_k\}\big)$ is the differential entropy rate of $\{H_k\}$; and the fading number for memoryless SIMO fading

$$\begin{aligned}\chi_{\mathrm{IID}}(\mathbf{H}) &= I\big(\hat{U}; \mathbf{H}\hat{U}\big) + \mathsf{E}[\log\|\mathbf{H}\|] \\ &\quad - h\big(\|\mathbf{H}\| \,\big|\, \hat{\mathbf{H}}\hat{U}\big) - \log 2, \end{aligned} \tag{6}$$

where $\hat{U}$ is independent of $\mathbf{H}$ and uniformly distributed over the complex sphere $\{z \in \mathbb{C} : |z| = 1\}$, and where $\hat{\mathbf{H}} = \mathbf{H}/\|\mathbf{H}\|$. Alternatively, $\chi_{\mathrm{IID}}(\mathbf{H})$ can be expressed as

$$\chi_{\mathrm{IID}}(\mathbf{H}) = h_\lambda(\hat{\mathbf{H}}\hat{U}) - h(\mathbf{H}) + n_{\mathrm{R}}\mathsf{E}\left[\log\|\mathbf{H}\|^2\right] - \log 2, \tag{7}$$

where h_λ is the differential entropy on the sphere, so that if a random vector $\mathbf{G}$ takes value on the unit-sphere and has the density $f_{\mathbf{G}}^{\lambda}(\mathbf{g})$ with respect to the surface-area measure λ, then

$$h_\lambda(\mathbf{G}) = -\mathsf{E}\left[\log f_{\mathbf{G}}^{\lambda}(\mathbf{G})\right].$$

The above is extended [1] to the case where the receiver has access to some side-information $\mathbf{S}$ such that $(\mathbf{H}, \mathbf{S})$ are independent of $\mathbf{Z}$, the joint law of $(\mathbf{H}, \mathbf{S}, \mathbf{Z})$ does not depend on the input, and the mutual information $I(\mathbf{H}; \mathbf{S})$ is finite,

$$I(\mathbf{H}; \mathbf{S}) < \infty. \tag{8}$$

In this case

$$\begin{aligned}\chi_{\mathrm{IID}}(\mathbf{H}|\mathbf{S}) &= h_\lambda(\hat{\mathbf{H}}\hat{U}|\mathbf{S}) - h(\mathbf{H}|\mathbf{S}) \\ &\quad + n_{\mathrm{R}}\mathsf{E}\left[\log\|\mathbf{H}\|^2\right] - \log 2. \end{aligned} \tag{9}$$

It is further shown in [1] that in the case of memoryless multiple-input multiple-output (MIMO) fading where the $n_{\mathrm{R}} \times n_{\mathrm{T}}$ random fading matrix $\mathbb{H}$ is of the form

$$\mathbb{H} = \mathsf{D} + \tilde{\mathbb{H}},$$

where D is a deterministic $n_{\mathrm{R}} \times n_{\mathrm{T}}$ matrix and $\tilde{\mathbb{H}}$ is a random $n_{\mathrm{R}} \times n_{\mathrm{T}}$ matrix of IID $\mathcal{N}_{\mathbb{C}}(0,1)$ components, the fading number can be bounded as

$$\chi_{\mathrm{IID}}(\mathsf{D} + \tilde{\mathbb{H}}) \ge \log\|\mathsf{D}\|^2 - \mathrm{Ei}\big(-\|\mathsf{D}\|^2\big) - 1, \tag{10}$$

$$\begin{aligned}\chi_{\mathrm{IID}}(\mathsf{D} + \tilde{\mathbb{H}}) &\le \min\{n_{\mathrm{R}}, n_{\mathrm{T}}\}\log\left(1 + \frac{\|\mathsf{D}\|^2}{\min\{n_{\mathrm{R}}, n_{\mathrm{T}}\}}\right) \\ &\quad + n_{\mathrm{R}}\log n_{\mathrm{R}} - n_{\mathrm{R}} - \log\Gamma(n_{\mathrm{R}}), \end{aligned} \tag{11}$$

where $\|\cdot\|$ denotes the matrix operator norm. This specializes for the SIMO case to

$$\chi_{\mathrm{IID}}(\mathbf{d} + \tilde{\mathbf{H}}) \ge \log\|\mathbf{d}\|^2 - \mathrm{Ei}\big(-\|\mathbf{d}\|^2\big) - 1, \tag{12}$$

$$\begin{aligned}\chi_{\mathrm{IID}}(\mathbf{d} + \tilde{\mathbf{H}}) &\le \log(1 + \|\mathbf{d}\|^2) + n_{\mathrm{R}}\log n_{\mathrm{R}} \\ &\quad - n_{\mathrm{R}} - \log\Gamma(n_{\mathrm{R}}), \end{aligned} \tag{13}$$

where the n_{R} components of $\tilde{\mathbf{H}}$ are IID $\mathcal{N}_{\mathbb{C}}(0,1)$. More generally, if

$$\mathbf{H} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{d}, \mathsf{\Sigma}),$$

where $\mathsf{\Sigma}$ has eigenvalues $\lambda^{(1)}, \ldots, \lambda^{(n_{\mathrm{R}})}$, then

$$\begin{aligned}\chi_{\mathrm{IID}}(\mathbf{H}) &\ge \log\sum_{r=1}^{n_{\mathrm{R}}}\left|\frac{(\mathsf{V}^{\mathsf{T}}\mathbf{d})^{(r)}}{\sqrt{\lambda^{(r)}}}\right|^2 \\ &\quad - \mathrm{Ei}\left(-\sum_{r=1}^{n_{\mathrm{R}}}\left|\frac{(\mathsf{V}^{\mathsf{T}}\mathbf{d})^{(r)}}{\sqrt{\lambda^{(r)}}}\right|^2\right) - 1, \end{aligned} \tag{14}$$

$$\begin{aligned}\chi_{\mathrm{IID}}(\mathbf{H}) &\le \log\left(1 + \sum_{r=1}^{n_{\mathrm{R}}}\left|\frac{(\mathsf{V}^{\mathsf{T}}\mathbf{d})^{(r)}}{\sqrt{\lambda^{(r)}}}\right|^2\right) \\ &\quad + n_{\mathrm{R}}\log n_{\mathrm{R}} - n_{\mathrm{R}} - \log\Gamma(n_{\mathrm{R}}), \end{aligned} \tag{15}$$

where V is unitary and diagonalizes $\mathsf{\Sigma}$:

$$\mathsf{\Sigma}\mathsf{V} = \mathsf{V}\,\mathrm{diag}\left(\lambda^{(1)}, \ldots, \lambda^{(n_{\mathrm{R}})}\right).$$

This follows because for any non-singular deterministic $n_{\mathrm{R}} \times n_{\mathrm{R}}$ matrix G,

$$\chi(\mathsf{G}\mathbb{H}) = \chi(\mathbb{H}),$$

see [1, Lemma 4.7]. The choice

$$\mathsf{G} = \mathrm{diag}\left(\frac{1}{\sqrt{\lambda^{(1)}}}, \ldots, \frac{1}{\sqrt{\lambda^{(n_{\mathrm{R}})}}}\right) \cdot \mathsf{V}^{\mathsf{T}}$$

leads to a fading vector $\mathsf{G}\mathbf{H}$ with components that are IID $\mathcal{N}_{\mathbb{C}}(0,1)$.

In particular, if $\mathbf{\Sigma}$ is diagonal,

$$\mathbf{\Sigma} = \text{diag}\left(\lambda^{(1)}, \dots, \lambda^{(n_\text{R})}\right),$$

then

$$\chi_\text{IID}(\mathbf{H}) \geq \log \sum_{r=1}^{n_\text{R}} \frac{\left|d^{(r)}\right|^2}{\lambda^{(r)}} - \text{Ei}\left(-\sum_{r=1}^{n_\text{R}} \frac{\left|d^{(r)}\right|^2}{\lambda^{(r)}}\right) - 1, \tag{16}$$

$$\chi_\text{IID}(\mathbf{H}) \leq \log\left(1 + \sum_{r=1}^{n_\text{R}} \frac{\left|d^{(r)}\right|^2}{\lambda^{(r)}}\right) + n_\text{R} \log n_\text{R} - n_\text{R} - \log \Gamma(n_\text{R}). \tag{17}$$

In this contribution we shall extend these results and compute the fading number for SIMO fading with memory.

2. THE MAIN RESULT

Our main result is the fading number of general (not necessarily Gaussian) SIMO channels with memory.

Theorem 1. *Consider the SIMO fading channel with memory* (1)*, and assume that the stationary and ergodic multi-variate fading process* $\{\mathbf{H}_k\}$ *satisfies the finite energy and finite differential entropy rate conditions* (2) *and* (3)*. Then irrespective of whether a peak- or an average-power constraint is imposed, the capacity of the channel is given at high SNR by*

$$\mathsf{C}(\text{SNR}) = \log\log \text{SNR} + \chi\big(\{\mathbf{H}_k\}\big) + o(1), \tag{18}$$

where the $o(1)$ *term tends to zero as the SNR tends to infinity, and where the fading number* $\chi\big(\{\mathbf{H}_k\}\big)$ *is given by*

$$\boxed{\chi\big(\{\mathbf{H}_k\}\big) = \chi_\text{IID}\left(\mathbf{H}_0 \,\middle|\, \mathbf{H}_{-\infty}^{-1}, \{\hat{\mathbf{H}}_\ell \hat{U}_\ell\}_{\ell=1}^{\infty}\right).} \tag{19}$$

Here $\chi_\text{IID}\big(\mathbf{H}_0 \big| \mathbf{S}\big)$ *is given in* (9)*, and the random process* $\{\hat{U}_k\}$ *is independent of* $\{\mathbf{H}_k\}$ *and constitutes of IID random variables that are uniformly distributed over the complex sphere,* i.e.*,*

$$\hat{U}_k \sim \text{Uniform on } \{z \in \mathbb{C} : |z| = 1\}.$$

Moreover, this asymptotic behavior is achievable at high SNR by IID circularly-symmetric inputs $\{X_k\}$ *such that*

$$\log|X_k|^2 \sim \text{Uniform on } [\log\log \mathcal{E}_\text{s}, \log \mathcal{E}_\text{s}].$$

Corollary 2. *From Theorem 1 it follows that*

$$\chi_\text{IID}\big(\mathbf{H}_0 \big| \mathbf{H}_{-\infty}^{-1}\big) \leq \chi(\{\mathbf{H}_k\}) \leq \chi_\text{IID}\big(\mathbf{H}_0 \big| \mathbf{H}_{-\infty}^{-1}, \mathbf{H}_1^{\infty}\big). \tag{20}$$

Remark 3. *Note that we can always find a lower bound to a SIMO fading system with memory (even with correlation between the antennas) by linear combining the outputs of the* n_R *receive antennas:*

$$\chi(\{\mathbf{H}_k\}) \geq \sup\left\{\log \pi + \mathsf{E}\left[\log|\tilde{H}_1|^2\right] - h\big(\{\tilde{H}_k\}\big)\right\}, \tag{21}$$

where $\tilde{H}_k = \sum_{r=1}^{n_\text{R}} \alpha^{(r)} H_k^{(r)}$ *and where the supremum is over all* $\alpha^{(1)}, \dots, \alpha^{(n_\text{R})}$ *that fulfill* $\sum_{r=1}^{n_\text{R}} |\alpha^{(r)}|^2 = 1$*. This bound, however, is generally not tight.*

3. INPUT DISTRIBUTIONS THAT ESCAPE TO INFINITY

An important ingredient in the proof of Theorem 1 is a new theorem regarding "input distributions that escape to infinity". Under slightly more restrictive conditions on the asymptotic behavior of channel capacity, we strengthen the results of [1] in the following sense. When specialized to the problem at hand, Theorem 4.13 of [1] demonstrates that the fading number *can* be achieved by input distributions that escape to infinity. That is, *there exist* input distributions satisfying the cost constraint and escaping to infinity that induce mutual informations whose *difference* from capacity tends to zero. Our present result, when specialized to the present setting, strengthens [1, Theorem 4.13] by showing that a sequence of inputs distributions satisfying the cost constraint *must* escape to infinity if it is to induce mutual informations whose *ratio* to $\log\log \text{SNR}$ is to approach one. The new result implies the old one, because if a sequence of input distributions induces mutual informations whose *difference* to capacity tends to zero, then the *ratios* to $\log\log\text{SNR}$ must approach one, and hence, by the new result, *must* escape to infinity.

It should, however, be noted that while the new result—like [1, Theorem 4.13]— extends to general cost constrained channels, the required assumptions on the functional form of the capacity-cost function are somewhat more stringent.

To state the new result, assume that the input and output alphabets $\mathcal{X}$ and $\mathcal{Y}$ of a memoryless channel $W : \mathcal{X} \mapsto \mathcal{Y}$ are separable metric spaces, and that for any set $\mathcal{B} \subset \mathcal{Y}$ the mapping $x \mapsto W(\mathcal{B}|x)$ from $\mathcal{X}$ to $[0, 1]$ is Borel measurable. Further assume that the cost function $g : \mathcal{X} \to [0, \infty)$ is measurable. Recall the following standard definition of the capacity-cost function:

Definition 4. *Given a channel* $W : \mathcal{X} \mapsto \mathcal{Y}$ *and given some non-negative cost function* $g : \mathcal{X} \to \mathbb{R}^+$*, we define the capacity-cost function* $\mathsf{C} : \mathbb{R}^+ \to \mathbb{R}^+$ *by*

$$\mathsf{C}(\Upsilon) = \sup_{Q : \mathsf{E}_Q[g(X)] \leq \Upsilon} I(Q, W), \qquad \Upsilon \geq 0. \tag{22}$$

Definition 5. *Let* $\{Q_\Upsilon, \Upsilon \geq 0\}$ *be a family of input distributions on* $\mathcal{X}$ *parameterized by the cost* Υ *such that*

$$\mathsf{E}_{Q_\Upsilon}[g(X)] \leq \Upsilon, \qquad \Upsilon \geq 0. \tag{23}$$

We say that the input distributions $\{Q_\Upsilon, \Upsilon \geq 0\}$ *escape to infinity if for any* $\Upsilon_0 > 0$

$$\lim_{\Upsilon\uparrow\infty} Q_\Upsilon\big(\{x \in \mathcal{X} : g(x) < \Upsilon_0\}\big) = 0. \tag{24}$$

We now present the theorem that demonstrates that if the ratio of mutual information to channel capacity is to approach one, then the input distributions must escape to infinity.

Theorem 6. *Let the cost function* $g(\cdot)$ *and the channel* $W : \mathcal{X} \mapsto \mathcal{Y}$ *be as above. Let the capacity-cost function* $\mathsf{C}(\cdot)$ *be finite but unbounded. Let* $\mathsf{C}_{\text{asy}}(\cdot)$ *be a function of the cost that captures the asymptotic behavior of the capacity-cost function* $\mathsf{C}(\cdot)$ *in the sense that*

$$\lim_{\Upsilon\uparrow\infty} \frac{\mathsf{C}(\Upsilon)}{\mathsf{C}_{\text{asy}}(\Upsilon)} = 1. \tag{25}$$

Assume that $\mathsf{C}_{\text{asy}}(\cdot)$ *satisfies the growth condition*

$$\varlimsup_{\Upsilon\uparrow\infty} \left\{ \sup_{\alpha\in(0,\alpha_0]} \frac{\alpha \mathsf{C}_{\text{asy}}\left(\frac{\Upsilon}{\alpha}\right)}{\mathsf{C}_{\text{asy}}(\Upsilon)} \right\} < 1, \qquad \forall\, 0 < \alpha_0 < 1. \tag{26}$$

Let $\{Q_\Upsilon, \Upsilon \geq 0\}$ *be a family of input distributions satisfying* (23) *and*

$$\lim_{\Upsilon\uparrow\infty} \frac{I(Q_\Upsilon, W)}{\mathsf{C}_{\text{asy}}(\Upsilon)} = 1. \tag{27}$$

Then $\{Q_\Upsilon, \Upsilon \geq 0\}$ *escapes to infinity.*

In the proof of Theorem 1 we use Theorem 6 with $\mathsf{C}_{\text{asy}}(\mathcal{E}_{\text{s}}) = \log\log \mathcal{E}_{\text{s}}$.

4. GAUSSIAN FADING WITH MEMORY

Since it is difficult to analytically evaluate the fading number given in (19) even for Gaussian fading, we will next use the bounds (20) to approximate it.

Let then $\boldsymbol{\mu}$ denote the mean-vector of the stationary vector-valued fading process $\{\mathbf{H}_k\}$, and assume that $\{\mathbf{H}_k - \boldsymbol{\mu}\}$ is a stationary circularly symmetric vector-valued Gaussian process with a diagonal spectral distribution matrix

$$\mathsf{F} = \operatorname{diag}\left(\mathsf{F}^{(1)}, \ldots, \mathsf{F}^{(n_{\text{R}})}\right).$$

Thus, the n_{R} components of the vector-valued process $\{\mathbf{H}_k\}$ are independent (corresponding to the case where the fading $\{H_k^{(r)}\}_{k=-\infty}^{\infty}$ experienced by the link between the transmitter antenna and the r-th receiver antenna is statistically independent of the fading experienced by the links with the other receiver antennas), and for each $1 \leq r \leq n_{\text{R}}$ the process $\{H_k^{(r)} - \mu^{(r)}\}_{k=-\infty}^{\infty}$ is a stationary circularly symmetric scalar Gaussian process of spectral distribution $\mathsf{F}^{(r)}$ so that

$$\mathsf{E}\left[\left(H_k^{(r)} - \mu^{(r)}\right)\left(H_{k+m}^{(r)} - \mu^{(r)}\right)^*\right] = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{i2\pi m\lambda}\, \mathrm{d}\mathsf{F}^{(r)}(\lambda).$$

Denote the derivative of $\mathsf{F}^{(r)}(\cdot)$ by $\mathsf{F}'^{(r)}(\cdot)$.

The optimum *prediction error* in estimating $H_0^{(r)}$ from its infinite past $\{H_\ell^{(r)}\}_{\ell=-\infty}^{-1}$ is the optimum *linear prediction error* which is given by (see, *e.g.*, [2], [3])

$$\epsilon^2_{\text{pred},r} = \exp\left(\int_{-\frac{1}{2}}^{\frac{1}{2}} \log \mathsf{F}'^{(r)}(\lambda)\, \mathrm{d}\lambda\right). \tag{28}$$

The optimum *interpolation error* in estimating $H_0^{(r)}$ from its infinite past and future

$$(\{H_\ell^{(r)}\}_{\ell=-\infty}^{-1}, \{H_\ell^{(r)}\}_{\ell=1}^{\infty})$$

is the optimum *linear interpolation error* given by (see [3, Sec. 37.2], [4], [5])

$$\epsilon^2_{\text{int},r} = \frac{4\pi^2}{\int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{\mathsf{F}'^{(r)}(\lambda)}\, \mathrm{d}\lambda}. \tag{29}$$

Conditional on $\{H_\ell^{(r)} = h_\ell^{(r)}\}_{\ell=-\infty}^{-1}$ the distribution of $H_0^{(r)}$ is Gaussian of mean

$$h_{\text{pred}}^{(r)}\big(\{h_\ell^{(r)}\}_{\ell=-\infty}^{-1}\big) = \mathsf{E}\left[H_0^{(r)} \,\middle|\, \{H_\ell^{(r)} = h_\ell^{(r)}\}_{\ell=-\infty}^{-1}\right]$$

and of variance $\epsilon^2_{\text{pred},r}$. Unconditionally, $H_{\text{pred}}^{(r)}$ is Gaussian of mean $\mu^{(r)}$ and of variance

$$\mathsf{Var}\Big(H_{\text{pred}}^{(r)}\Big) = \mathsf{Var}\Big(H_0^{(r)}\Big) - \epsilon^2_{\text{pred},r} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathsf{F}'^{(r)}(\lambda)\, \mathrm{d}\lambda - \epsilon^2_{\text{pred},r}.$$

Similarly, conditional on

$$\big(\{H_\ell^{(r)} = h_\ell^{(r)}\}_{\ell=-\infty}^{-1}, \{H_\ell^{(r)} = h_\ell^{(r)}\}_{\ell=1}^{\infty}\big)$$

the distribution of $H_0^{(r)}$ is Gaussian of mean

$$h_{\text{int}}^{(r)}\big(\{h_\ell^{(r)}\}_{\ell=-\infty}^{-1}, \{h_\ell^{(r)}\}_{\ell=1}^{\infty}\big)$$

$$= \mathsf{E}\Big[H_0^{(r)} \,\Big|\, \{H_\ell^{(r)} = h_\ell^{(r)}\}_{\ell=-\infty}^{-1}, \{H_\ell^{(r)} = h_\ell^{(r)}\}_{\ell=1}^{\infty}\Big]$$

and of variance $\epsilon^2_{\text{int},r}$. Unconditionally, $H_{\text{int}}^{(r)}$ is Gaussian of mean $\mu^{(r)}$ and of variance

$$\begin{aligned}\mathsf{Var}\Big(H_{\text{int}}^{(r)}\Big) &= \mathsf{Var}\Big(H_0^{(r)}\Big) - \epsilon^2_{\text{int},r} \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathsf{F}'^{(r)}(\lambda)\,\mathrm{d}\lambda - \epsilon^2_{\text{int},r}.\end{aligned}$$

Since we have assumed that the components of $\mathbf{H}_k$ are independent, we can use (16) and (17) to further bound the expressions in (20). We start with the upper bound:

$$\begin{aligned}&\chi(\{\mathbf{H}_k\}) \\ &\quad\le \chi_{\text{IID}}\big(\mathbf{H}_0\big|\mathbf{H}_{-\infty}^{-1}, \mathbf{H}_1^{\infty}\big) \\ &\quad\le \mathsf{E}\left[\log\left(1 + \sum_{r=1}^{n_{\text{R}}} \frac{\left|H_{\text{int}}^{(r)}\right|^2}{\epsilon^2_{\text{int},r}}\right)\right] \\ &\qquad + n_{\text{R}}\log n_{\text{R}} - n_{\text{R}} - \log\Gamma(n_{\text{R}}) \\ &\quad\le \log\left(1 + \sum_{r=1}^{n_{\text{R}}} \frac{\mathsf{E}\left[\left|H_{\text{int}}^{(r)}\right|^2\right]}{\epsilon^2_{\text{int},r}}\right) \\ &\qquad + n_{\text{R}}\log n_{\text{R}} - n_{\text{R}} - \log\Gamma(n_{\text{R}}) \\ &\quad= \log\left(1 + \sum_{r=1}^{n_{\text{R}}} \frac{\mathsf{Var}\Big(H_0^{(r)}\Big) - \epsilon^2_{\text{int},(r)} + |\mu^{(r)}|^2}{\epsilon^2_{\text{int},r}}\right) \\ &\qquad + n_{\text{R}}\log n_{\text{R}} - n_{\text{R}} - \log\Gamma(n_{\text{R}}). \qquad (30)\end{aligned}$$

Here the first inequality is due to (20); the second inequality follows from (17); and the third inequality follows from Jensen's inequality.

For the lower bound we get

$$\begin{aligned}\chi(\{\mathbf{H}_k\}) &\ge \chi_{\text{IID}}\big(\mathbf{H}_0\big|\mathbf{H}_{-\infty}^{-1}\big) \\ &\ge \mathsf{E}\left[\log\sum_{r=1}^{n_{\text{R}}} \frac{\left|H_{\text{pred}}^{(r)}\right|^2}{\epsilon^2_{\text{pred},r}}\right] \\ &\quad - \mathsf{E}\left[\mathrm{Ei}\left(-\sum_{r=1}^{n_{\text{R}}} \frac{\left|H_{\text{pred}}^{(r)}\right|^2}{\epsilon^2_{\text{pred},r}}\right)\right] - 1 \\ &\ge \mathsf{E}\left[\log\sum_{r=1}^{n_{\text{R}}} \frac{\left|H_{\text{pred}}^{(r)}\right|^2}{\epsilon^2_{\text{pred},r}}\right] \\ &\quad - \mathrm{Ei}\left(-\sum_{r=1}^{n_{\text{R}}} \frac{\mathsf{E}\left[\left|H_{\text{pred}}^{(r)}\right|^2\right]}{\epsilon^2_{\text{pred},r}}\right) - 1, \quad (31)\end{aligned}$$

where the first inequality is due to (20); the second inequality follows from (16); and where the last inequality follows from Jensen's inequality. The analytic computation of the RHS of (31) is greatly simplified if each component process $\{H_k^{(r)}\}$ of the vector-valued fading process $\{\mathbf{H}_k\}$ is of an identical law, which in our case means that

$$\mathsf{F} = \mathrm{diag}\Big(\tilde{\mathsf{F}}, \ldots, \tilde{\mathsf{F}}\Big) \qquad (32)$$

for a scalar spectral distribution function $\tilde{\mathsf{F}}$, and

$$\boldsymbol{\mu} = \Big(\mu^{(1)}, \ldots, \mu^{(n_{\text{R}})}\Big)^{\mathsf{T}} = (\tilde{\mu}, \ldots, \tilde{\mu})^{\mathsf{T}} \qquad (33)$$

for a mean $\tilde{\mu}$. In that case (using the expectation of the logarithm of a non-central χ^2-distribution [1, Appendix X]) we obtain

$$\begin{aligned}\chi(\{\mathbf{H}_k\}) \ge \log\frac{\mathsf{Var}\Big(H_0^{(1)}\Big) - \epsilon^2_{\text{pred}}}{\epsilon^2_{\text{pred}}} - 1 \\ + g_{n_{\text{R}}}\left(\frac{n_{\text{R}}|\tilde{\mu}|^2}{\Big(\mathsf{Var}\Big(H_0^{(1)}\Big) - \epsilon^2_{\text{pred}}\Big)^2}\right) \\ - \mathrm{Ei}\left(-n_{\text{R}}\frac{\mathsf{Var}\Big(H_0^{(1)}\Big) - \epsilon^2_{\text{pred}} + |\tilde{\mu}|^2}{\epsilon^2_{\text{pred}}}\right), \quad (34)\end{aligned}$$

where $g_m(\cdot)$ is defined as [1]

$$\begin{aligned}&g_m(z) = \log z - \mathrm{Ei}(-z) \\ &\quad + \sum_{j=1}^{m-1} (-1)^j \left(e^{-z}(j-1)! - \frac{(m-1)!}{j(m-1-j)!}\right) z^{-j}, \\ &\qquad\qquad\qquad\qquad z \ge 0. \quad (35)\end{aligned}$$

Note that the simplifying assumptions (32) and (33) are not necessary if one resorts to the weaker lower bound described in Remark 3.

Example 7. *Suppose the fading process $\{\mathbf{H}_k\}$ is spatially IID so that the processes*

$$\{H_k^{(1)}\}_{k=-\infty}^{\infty}, \{H_k^{(2)}\}_{k=-\infty}^{\infty}, \ldots, \{H_k^{(n_{\text{R}})}\}_{k=-\infty}^{\infty}$$

are independent of each other and of identical (not necessarily temporally IID) law. Suppose that under this law $\{H_k^{(r)}\}$ is a stationary, unit-variance, zero-mean, circularly symmetric, m-th order auto-regressive AR(m) Gaussian process. We thus assume that for all $1 \le r \le n_{\text{R}}$,

$$H_k^{(r)} = W_k^{(r)} - a_1 H_{k-1}^{(r)} - a_2 H_{k-2}^{(r)} - \cdots - a_m H_{k-m}^{(r)}. \quad (36)$$

Here $\{W_k^{(r)}\}$ is temporally IID $\mathcal{N}_{\mathbb{C}}(0, \varepsilon^2)$, where ε^2 denotes the innovation variance; the coefficients $a_1, \ldots, a_m$ satisfy the stability condition [6]

$$\sum_{j=1}^{m} a_j z^j \neq -1 \qquad \forall |z| \leq 1; \tag{37}$$

and ε^2 and $a_1, \ldots, a_m$ are such that

$$\mathsf{Var}\left(H_k^{(r)}\right) = 1. \tag{38}$$

Then [6]

$$\epsilon^2_{\text{pred},r} = \varepsilon^2, \tag{39}$$

$$H^{(r)}_{\text{pred}} \sim \mathcal{N}_{\mathbb{C}}\left(0, 1 - \epsilon^2_{\text{pred}}\right), \tag{40}$$

$$\epsilon^2_{\text{int},r} = \frac{\varepsilon^2}{1 + \sum_{j=1}^{m} |a_j|^2}, \tag{41}$$

$$H^{(r)}_{\text{int}} \sim \mathcal{N}_{\mathbb{C}}\left(0, 1 - \epsilon^2_{\text{int}}\right), \tag{42}$$

which yields

$$\chi(\{\mathbf{H}_k\}) \geq \log \frac{1-\varepsilon^2}{\varepsilon^2} + \psi(n_{\text{R}}) - \text{Ei}\left(-n_{\text{R}} \frac{1-\varepsilon^2}{\varepsilon^2}\right) - 1, \tag{43}$$

$$\chi(\{\mathbf{H}_k\}) \leq \log \left(1 + n_{\text{R}} \frac{1 + \sum_{j=1}^{m} |a_j|^2 - \varepsilon^2}{\varepsilon^2}\right) + n_{\text{R}} \log n_{\text{R}} - n_{\text{R}} - \log \Gamma(n_{\text{R}}), \tag{44}$$

where $\psi(\cdot)$ denotes Euler's psi function

$$\psi(n_{\text{R}}) = -\gamma + \sum_{j=1}^{n_{\text{R}}-1} \frac{1}{j}$$

and γ denotes Euler's constant.

For the case of Gaussian-Markov fading ($m = 1$, $a_1 = -\sqrt{1 - \varepsilon^2}$) the lower bound (43) is unchanged and the upper bound becomes

$$\chi(\{\mathbf{H}_k\}) \leq \log \left(1 + 2n_{\text{R}} \frac{1-\varepsilon^2}{\varepsilon^2}\right) + n_{\text{R}} \log n_{\text{R}} - n_{\text{R}} - \log \Gamma(n_{\text{R}}). \tag{45}$$

For $\varepsilon^2 \ll 1$ one obtains the asymptotic bounds

$$\chi(\{\mathbf{H}_k\}) \geq \log \frac{1}{\varepsilon^2} + \psi(n_{\text{R}}) - 1 + o(\varepsilon^2), \tag{46}$$

$$\chi(\{\mathbf{H}_k\}) \leq \log \frac{1}{\varepsilon^2} + \log 2 + (n_{\text{R}} + 1) \log n_{\text{R}} - n_{\text{R}} - \log \Gamma(n_{\text{R}}) + o(\varepsilon^2), \tag{47}$$

where $o(\varepsilon^2)$ tends to zero as ε^2 tends to zero.

For the case of two receive antennae $n_{\text{R}} = 2$ these bounds are depicted in Figure 1.

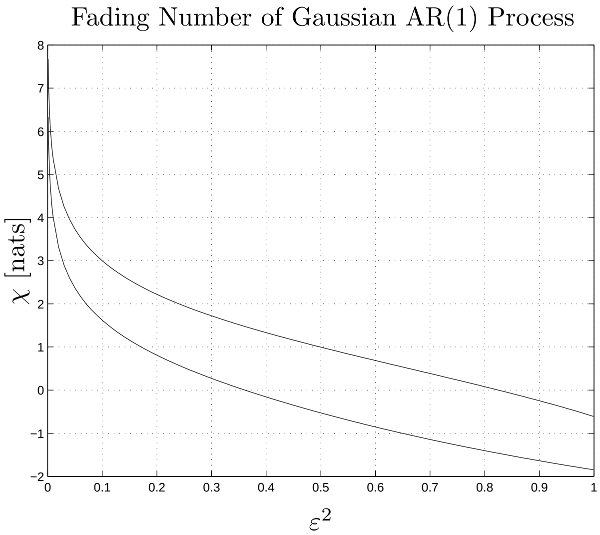

Figure 1: Upper and lower bound of a zero-mean SIMO Gaussian-Markov fading channel with memory one (AR(1)) and two receiver antennas plotted in function of the prediction error ε^2. Both components of the fading vector $\mathbf{H}_k$ are assumed to be independent and identically distributed with variance 1.

Acknowledgments

Fruitful discussions with İ. Emre Telatar are gratefully acknowledged.

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*, 49(10):2426–2467, Oct. 2003.

[2] M. Pourahmadi. *Foundations of Time Series Analysis and Prediction Theory.* J. Wiley & Sons, 2001.

[3] A. M. Yaglom. *An Introduction to the Theory of Stationary Random Functions.* Prentice Hall, 1962.

[4] A. N. Kolmogorov. Stationary sequences in Hilbert space. *Bull. Moscow State Univ.*, 2(6), 1941.

[5] A. N. Kolmogorov. Interpolation and extrapolation of stationary random sequences. *Izvestija Akademii Nauk SSSR*, 5 (Ser. Mat.):3, 1941.

[6] J. Anděl. A comparison of methods for interpolation in time series. In *Math. Statistics and Appl.*, vol. B (Bad Tatzmannsdorf), pp. 53–67, 1983.