The Asymptotic Capacity of Noncoherent Single-Input Multiple-Output Fading Channels with Memory and Feedback

Final Report of NSC Project

"Capacity Analysis of Multiple-Users Communication Systems using Joint Estimation and Detection"

Date:	31 Oct 2013
Project-Numbers:	NSC 99–2628–E–009–002
	NSC 100–2628–E–009–003
	NSC 101–2628–E–009–002
Project Duration:	1 August 2010 – 31 July 2013
Funded by:	National Science Council, Taiwan
Author:	Stefan M. Moser
Organization:	Information Theory Laboratory
	Department of Electrical and Computer Engineering
	National Chiao Tung University
Address:	Engineering Building IV, Office 727
	1001 Daxue Rd.
	Hsinchu 30010, Taiwan
E-mail:	stefan.moser@ieee.org

Abstract

The channel capacity of a noncoherent single-input multiple-output regular fading channel with memory and with feedback is investigated. The fading process is assumed to be a general stationary and ergodic random process of finite energy and finite differential entropy rate. The feedback is assumed to be noisefree (i.e., it is of infinite capacity), but causal. It is shown that the asymptotic capacity grows double-logarithmically in the power and that the second term in the asymptotic expansion, the *fading number*, is unchanged with respect to the same channel without feedback.

Keywords: Asymptotic analysis, channel capacity, fading, fading number, feedback, noncoherent.

Contents

1 Introduction

Noncoherent multiple-antenna fading channel models have attracted a lot of attention for quite some years because they realistically describe the omnipresent mobile wireless communication channels. Here, *noncoherent* refers to the fundamental assumption that transmitter and receiver only have knowledge about the distribution of the fading process, but have no direct access to the current realization. Hence, the communication system needs to provide some means of measuring the current channel state, thereby using part of the available bandwidth, power, and computational efforts for the channel state estimation.

This is in stark contrast to the coherent fading models where it is assumed that the receiver has *free* and *noiseless* access to the current fading realization [1]. It is particularly the latter assumption of *perfect* knowledge of the fading realization that leads to overly optimistic capacity results for coherent channel models with respect to what can be expected to be seen in practice.

The noncoherent channel models can be split into different families. For so-called *underspread fading channels*, it is assumed that the fading process is wide-sense stationary and uncorrelated in the delay, where the product of the delay and Doppler spread is small (for more details, see [2] and references therein). The *block-fading* models assume that for a certain time, the fading realization remains unchanged before a new (potentially dependent) value is taken on [3], [4], [5]. In *nonregular* fading, the fading process is assumed to be stationary with strong memory that permits a quite precise prediction of the present fading values from the past [6], [7]. It might be even the case that one can perfectly compute the current values from the infinite past with a zero prediction error. Note, however, that due to the noncoherence assumption and due to the additive noise, the receiver never has access to the exact past fading values, but only to a noisy observation of them.

In this paper we investigate the family of noncoherent *regular* fading channels. In contrast to nonregular fading, here it is assumed that the stationary fading process has a *finite* differential entropy rate. In [8] it has been shown that the capacity of multiple-antenna regular fading channels only grows *double-logarithmically* in the available power at high signal-to-noise ratios (SNR). This is much slower than the common logarithmic growth, e.g., of coherent fading channels, and it persists independently of the number of antennas used at transmitter and receiver and independently of the memory in the fading process.

For a more precise description of this phenomena, [8] defined the *fading number* χ as the second term in the high-SNR asymptotic expansion of the channel capacity:

$$\chi(\{\mathbb{H}_k\}) \triangleq \varlimsup_{\mathsf{E}_\mathrm{s}\uparrow\infty} \{\mathsf{C}(\mathsf{E}_\mathrm{s}) - \log\log \mathsf{E}_\mathrm{s}\}. \tag{1}$$

An analytic expression for its value for general multiple-input multiple-output fading channels with memory has been derived in [8], [9].

While the assumption of a noncoherent communication system is realistic, we also should take into account that many practical communication systems are bidirectional allowing to send feedback from the receiver back to the transmitter. Such a feedback link will help to simplify the necessary coding scheme and it even has the potential to increase the channel capacity. In this paper, we investigate the impact of feedback in the situation of a general single-input multiple-output (SIMO) regular fading channel with memory. We do not restrict the exact distribution of the fading process, apart from it being stationary and ergodic. Concerning the feedback, we assume the rather unrealistic situation of a feedback link that has infinite capacity.

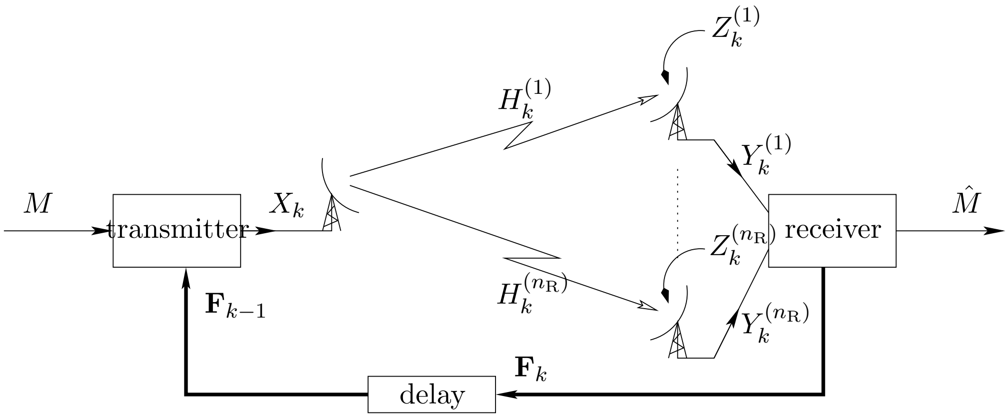

Figure 1: SIMO regular fading channel with n_R receive antennas and with noiseless causal feedback.

This will lead to an upper bound on the capacity in the presence of any practical type of feedback. The only constraint we make is *causality*, i.e., the feedback will arrive at the transmitter delayed by one time-step.

In [10], [11], [12] it has been shown that the potential improvement in capacity offered by feedback is only very limited even in the situation of multiple-input multiple-output (MIMO) regular fading channels. There it was also proven that for the case of single-input single-output (SISO) regular fading channels, the asymptotic high-SNR capacity remains unchanged by any type of causal feedback. This work now presents a generalization of these results.

The remainder of this paper is structured as follows: in Section 2 we will specify the channel model in detail and give some comments about the used notation. Section 3 summarizes the results for the channel model without feedback including some required definitions and some explanations about the meaning of the fading number. The main result, i.e., the exact asymptotic capacity of SIMO fading channels with noiseless feedback, are then presented in Section 4. In Section 5 we give an overview over the derivations, and Section 6 contains some concluding remarks.

2 Channel Model

We consider a communication system as shown in Figure 1. A message M is transmitted over a SIMO fading channel with memory where the transmitter has one antenna and the receiver has n_R antennas. The channel output vector $\mathbf{Y}_k \in \mathbb{C}^{n_R}$ at time k is given by

$$\mathbf{Y}_k = \mathbf{H}_k x_k + \mathbf{Z}_k, \tag{2}$$

where $x_k \in \mathbb{C}$ denotes the time-k channel input; the random vector $\mathbf{H}_k \in \mathbb{C}^{n_R}$ denotes the time-k fading vector with n_R components corresponding to the n_R antennas at the receiver; and where the random vector $\mathbf{Z}_k \in \mathbb{C}^{n_R}$ models additive noise.

We assume that the additive noise process $\{\mathbf{Z}_k\}$ is spatially and temporally independent and identically distributed (IID), circularly-symmetric, and complex Gaussian with zero mean and with variance $\sigma^2 > 0$:

$$\{\mathbf{Z}_k\} \text{ IID} \sim \mathcal{N}_{\mathbb{C}}\big(\mathbf{0}, \sigma^2 \mathsf{I}_{n_R}\big). \tag{3}$$

Here, I_{n_R} denotes the $n_R \times n_R$ identity matrix.

The fading process $\{\mathbf{H}_k\}$ is statistically independent of $\{\mathbf{Z}_k\}$ and is assumed to be stationary, ergodic, of finite energy $\mathsf{E}\left[\|\mathbf{H}_k\|^2\right] < \infty$, and of finite differential entropy rate

$$h(\{\mathbf{H}_k\}) > -\infty. \tag{4}$$

A random process satisfying this later condition (4) is usually called *regular*. Note that we do not make any further assumptions about $\{\mathbf{H}_k\}$, i.e., we do not assume a particular law (like, e.g., a Gaussian distribution). In particular we do allow for arbitrary dependences between the different components $\{H_k^{(j)}\}$ corresponding to the different antennas (spatial memory) and over time (temporal memory).

We assume noncoherent communication, i.e., neither transmitter nor receiver know the realization of $\{\mathbf{H}_k\}$, they only know its law.

From the receiver to the transmitter we have a noiseless feedback link (i.e., the link has infinite capacity and allows the receiver to send everything it knows back to the transmitter). However, to preserve causality of the system, we require the feedback to be delayed by one discrete time-step. So the feedback vector $\mathbf{F}_k$ that is available at the transmitter at time k consists of all past channel output vectors:

$$\mathbf{F}_k = \mathbf{Y}_1^{k-1}. \tag{5}$$

The channel input X_k at time k therefore is a deterministic function of the message M and the feedback $\mathbf{Y}_1^{k-1}$. Note that we assume M to be uniformly distributed.

We consider two types of power constraints: an average-power constraint and a peak-power constraint. Under the former we require that for every message m

$$\frac{1}{n}\sum_{k=1}^{n} \mathsf{E}\left[\left|X_k\left(m, \mathbf{Y}_1^{k-1}\right)\right|^2\right] \le \mathsf{E}_{\mathrm{s}}, \tag{6}$$

where n denotes the blocklength. Under the peak-power constraint we replace (6) with the almost-sure constraint that for every message m

$$\left|X_k\left(m, \mathbf{Y}_1^{k-1}\right)\right|^2 \le \mathsf{E}_{\mathrm{s}}, \quad \text{a.s.,} \;\; k = 1, \ldots, n. \tag{7}$$

To clarify notation we will use a subscript "FB" whenever feedback is available, while the subscript "IID" refers to a situation without memory or feedback. RHS stands for 'right-hand side'. Note that log refers to the natural logarithm and all rates are specified in nats.

3 Capacity and Fading Number without Feedback

It has been shown in [8] that the capacity of general SIMO regular fading channels under either an average-power constraint or a peak-power constraint is

$$\mathsf{C}(\mathsf{E}_{\mathrm{s}}) = \log(1 + \log(1 + \mathsf{E}_{\mathrm{s}})) + \chi(\{\mathbf{H}_k\}) + o(1), \tag{8}$$

where $o(1)$ denotes terms that tend to zero as E_{s} tends to infinity, and where the fading number χ is defined in (1) and is given by [8]

$$\chi(\{\mathbf{H}_k\}) = h_\lambda\left(\hat{\mathbf{H}}_0 e^{\mathrm{i}\Theta_0} \,\middle|\, \left\{\hat{\mathbf{H}}_\ell e^{\mathrm{i}\Theta_\ell}\right\}_{\ell=-\infty}^{-1}\right) - \log 2 \\ + n_{\mathrm{R}}\,\mathsf{E}\left[\log \|\mathbf{H}_0\|^2\right] - h\left(\mathbf{H}_0 \,\middle|\, \mathbf{H}_{-\infty}^{-1}\right). \tag{9}$$

Here, $\{\Theta_k\}$ is IID $\sim \mathcal{U}((-\pi, \pi])$ and independent of $\{\mathbf{H}_k\}$, and $\hat{\mathbf{H}}_k$ denotes the unit vector

$$\hat{\mathbf{H}}_k \triangleq \frac{\mathbf{H}_k}{\|\mathbf{H}_k\|}. \tag{10}$$

Note that since the unit vectors $\hat{\mathbf{H}}_k$ only take value on the unit sphere in $\mathbb{C}^{n_\mathrm{R}}$ and since the surface of this unit sphere has zero measure over $\mathbb{C}^{n_\mathrm{R}}$, we define a differential entropy-like quantity $h_\lambda(\cdot)$ that only lives on the surface of the unit sphere in $\mathbb{C}^{n_\mathrm{R}}$:

$$h_\lambda(\hat{\mathbf{V}}) \triangleq \mathsf{E}\left[-\log p^\lambda_{\hat{\mathbf{V}}}(\hat{\mathbf{V}})\right], \tag{11}$$

if the expectation exists. Here $p^\lambda_{\hat{\mathbf{V}}}(\hat{\mathbf{v}})$ denotes the PDF of the random unit-vector $\hat{\mathbf{V}}$ with respect to the $\mathbb{C}^{n_\mathrm{R}}$-surface measure λ. For more details we refer to [9, Sec. II].

From (8) it is obvious that the capacity of the fading channel (2) grows extremely slowly at large power. Indeed, $\log(1 + \log(1 + \mathsf{E}_\mathrm{s}))$ grows so slowly that, for the smallest values of E_s for which $o(1) \approx 0$, the (constant!) fading number χ usually is much larger than $\log(1 + \log(1 + \mathsf{E}_\mathrm{s}))$. Hence, the threshold between the low-power regime and the capacity-inefficient high-power regime is strongly related to the fading number: the larger the fading number is, the higher the rate can be chosen without operating the system in the inefficient double-logarithmic regime.

Also note that even though the double-logarithmic term on the RHS of (8) does not depend on $\{\mathbf{H}_k\}$ or, particularly, on the number of antennas n_R, it is still beneficial to have multiple antennas because the fading number χ does depend strongly on the fading process and the number of antennas.

From (9) one also sees that in the case of a memoryless SIMO fading channel, the fading number is given by

$$\chi_{\mathrm{IID}}(\mathbf{H}) = h_\lambda\left(\hat{\mathbf{H}}e^{\mathrm{i}\Theta}\right) - \log 2 + n_\mathrm{R}\,\mathsf{E}\left[\log \|\mathbf{H}\|^2\right] - h(\mathbf{H}), \tag{12}$$

and that therefore the fading number in (9) also can be written as

$$\begin{aligned}\chi(\{\mathbf{H}_k\}) = \chi_{\mathrm{IID}}(\mathbf{H}_0) &+ I\left(\mathbf{H}_0; \mathbf{H}_{-\infty}^{-1}\right) \\ &- I\left(\hat{\mathbf{H}}_0 e^{\mathrm{i}\Theta_0}; \left\{\hat{\mathbf{H}}_\ell e^{\mathrm{i}\Theta_\ell}\right\}_{\ell=-\infty}^{-1}\right).\end{aligned} \tag{13}$$

In [8], it has also been shown that for an arbitrary value of the power E_s, the channel capacity can be bounded as follows:

$$\mathsf{C}(\mathsf{E}_\mathrm{s}) \leq \mathsf{C}_{\mathrm{IID}}(\mathsf{E}_\mathrm{s}) + I\left(\mathbf{H}_0; \mathbf{H}_{-\infty}^{-1}\right), \quad \mathsf{E}_\mathrm{s} \geq 0. \tag{14}$$

From (13) we see that this upper bound may not be tight. In particular, asymptotically for $\mathsf{E}_\mathrm{s} \to \infty$ it is strictly loose.

4 Capacity and Fading Number with Feedback

While it is well-known that feedback has no effect on the capacity of a memoryless channel, in general feedback does increase capacity for channels with memory. The reason for this is that the combination of feedback and memory allows the transmitter to predict the current channel state and thereby adapt to it. Unfortunately, for regular fading channels this increase in capacity due to the feedback turns out to be very limited.

Theorem 1 (Capacity Increase by Feedback is Bounded by a Constant). *Let a general SIMO fading channel with memory be defined as in (2) and consider a noiseless causal feedback link as described in (5) (see Figure 1). Then the channel capacity under either an average-power constraint (6) or a peak-power constraint (7) is upper-bounded as follows:*

$$\mathsf{C}_{\mathrm{FB}}(\mathsf{E}_\mathrm{s}) \leq \mathsf{C}_{\mathrm{IID}}(\mathsf{E}_\mathrm{s}) + I\big(\mathbf{H}_0; \mathbf{H}_{-\infty}^{-1}\big), \quad \mathsf{E}_\mathrm{s} \geq 0. \tag{15}$$

Proof: This theorem has been proven in [10] for the general case of a MIMO regular fading channel with memory and noiseless feedback. We omit the details. □

We note that the RHS of (15) is identical to the RHS of (14). Hence the same (alas potentially loose) bound holds both for the channel capacity with and without feedback. Moreover, also note that $\mathsf{C}(\mathsf{E}_\mathrm{s})$ trivially is a lower bound to $\mathsf{C}_{\mathrm{FB}}(\mathsf{E}_\mathrm{s})$ since the transmitter can simply ignore the feedback and achieve the same results as without feedback.

An immediate consequence of Theorem 1 is that $\mathsf{C}_{\mathrm{FB}}(\mathsf{E}_\mathrm{s})$ only grows double-logarithmically in the power at high power and that therefore there exists a fading number $\chi_{\mathrm{FB}}(\{\mathbf{H}_k\})$ with a definition corresponding to (1). Theorem 1 can then be applied to $\chi_{\mathrm{FB}}(\{\mathbf{H}_k\})$. However, we will not explicitly spell out this bound on the fading number, but directly state a stronger statement.

Theorem 2 (SIMO Fading Number with Memory and Feedback). *Let a general SIMO fading channel with memory be defined as in (2) and consider a noiseless causal feedback link as described in (5) (see Figure 1). Then the asymptotic channel capacity under either an average-power constraint (6) or a peak-power constraint (7) is identical to the asymptotic channel capacity for the channel without feedback:*

$$\mathsf{C}_{\mathrm{FB}}(\mathsf{E}_\mathrm{s}) = \log(1 + \log(1 + \mathsf{E}_\mathrm{s})) + \chi_{\mathrm{FB}}(\{\mathbf{H}_k\}) + o(1) \tag{16}$$

where the fading number is

$$\begin{aligned} \chi_{\mathrm{FB}}(\{\mathbf{H}_k\}) &= \chi(\{\mathbf{H}_k\}) \\ &= h_\lambda\Big(\hat{\mathbf{H}}_0 e^{\mathrm{i}\Theta_0} \Big| \big\{\hat{\mathbf{H}}_\ell e^{\mathrm{i}\Theta_\ell}\big\}_{\ell=-\infty}^{-1}\Big) - \log 2 \\ &\quad + n_\mathrm{R}\, \mathsf{E}\big[\log \|\mathbf{H}_0\|^2\big] - h\big(\mathbf{H}_0 \big| \mathbf{H}_{-\infty}^{-1}\big). \end{aligned} \tag{17}$$

We would like to point out that this result even holds in the (hypothetical) case when the feedback is improved in the sense that in addition to the past channel outputs the transmitter also is informed about the past fading realizations $\mathbf{H}_1^{k-1}$. Note further that since we have assumed the most optimistic form of causal feedback, any type of realistic feedback will yield the same result.

Next, we try to give a hand-waving explanation of this behavior. Since the fading process is assumed to be regular with a finite differential entropy rate, it is not possible to perfectly predict the future realizations of the process even if one is presented with the exact realizations of the infinite past. Nevertheless, the feedback allows the transmitter to make an estimate of future realizations. Based on these estimates, the transmitter can then perform elaborate schemes of optimal power allocation over time: if the channel state is likely to be poor, it saves power and uses it once the channel state is likely to be good again. Unfortunately, due to the double-logarithmic behavior of capacity, such power allocation has no effect at all: for any constant $\beta > 0$ (β can be chosen arbitrarily large!),

$$\begin{aligned}
&\varlimsup_{\mathsf{E_s}\uparrow\infty} \{\log\log \beta\mathsf{E_s} - \log\log \mathsf{E_s}\} \\
&\quad = \varlimsup_{\mathsf{E_s}\uparrow\infty} \{\log(\log\beta + \log\mathsf{E_s}) - \log\log\mathsf{E_s}\} && (18)\\
&\quad = \varlimsup_{\mathsf{E_s}\uparrow\infty} \{\log(\log\mathsf{E_s}) - \log\log\mathsf{E_s}\} && (19)\\
&\quad = 0. && (20)
\end{aligned}$$

So not only the double-logarithmic growth is left untouched, but also the second term, i.e., the fading number, remains unchanged.

5 Outline of a Proof of Theorem 2

Since the channel capacity of the system without feedback trivially is a lower bound on the channel capacity with feedback, and since the capacity under a peak-power constraint is a lower bound on the capacity with an average-power constraint, it is sufficient to derive an upper bound on $\chi_{\text{FB}}(\{\mathbf{H}_k\})$ under the assumption of the average-power constraint (6) and to show that it is identical to the fading number without feedback and under the assumption of a peak-power constraint.

The proof is very lengthy and we therefore will only outline the main ideas. The basic structure follows the proof of the general fading number of MIMO fading channels with memory given in [9]. However, there are many details that need to be adapted and taken care of. Particularly, we have to consider the following challenges:

- Due to the feedback, the channel input, the fading, and the additive noise become dependent.
- We cannot rely on the important auxiliary result given in [9, Th. 3] that shows that the optimal input is stationary.
- We cannot rely on the important auxiliary result given in [13, Th. 8] that shows that the capacity-achieving input distribution *escapes to infinity*.

To handle the first challenge, we often rely on the concept of *causal interpretations* [14], [15]. This is a tool that allows to graphically proof the independence of random variables when conditioned on certain other random variables.

The missing auxiliary result concerning the capacity-achieving input distribution escaping to infinity can be proven indirectly inside of the derivation.

The biggest difficulty is caused by the nonstationarity of the channel input that is inherent to the given context because the transmitter continuously learns more about the fading process through the feedback and thereby changes the optimal distribution of the input.

The proof starts with Fano's inequality, which states that any given sequence of communication systems with rate R_{FB} and power $\mathsf{E_s}$ must satisfy

$$\mathsf{R}_{\text{FB}}(\mathsf{E_s}) \le \frac{1}{n} I\big(M; \mathbf{Y}_1^n\big) + \frac{\log 2}{n} + P_{\text{e}}^{(n)} \mathsf{R}_{\text{FB}}(\mathsf{E_s}) + \frac{\epsilon_n}{n}, \tag{21}$$

where $P_{\text{e}}^{(n)}$ denotes the error probability, which (for a reliable system) must tend to zero as n tends to infinity.

The mutual information term is then bounded as follows:

$$\frac{1}{n}I\big(M;\mathbf{Y}_1^n\big)$$
$$= \frac{1}{n}\sum_{k=1}^{n} I\big(M;\mathbf{Y}_k\big|\mathbf{Y}_1^{k-1}\big) \tag{22}$$
$$= \frac{1}{n}\sum_{k=1}^{n}\Big(I\big(M,\mathbf{Y}_1^{k-1};\mathbf{Y}_k\big) - I\big(\mathbf{Y}_1^{k-1};\mathbf{Y}_k\big)\Big) \tag{23}$$
$$\leq \frac{1}{n}\sum_{k=1}^{n}\Big(I\big(M,\mathbf{Y}_1^{k-1},X_k,\mathbf{H}_1^{k-1};\mathbf{Y}_k\big) - I\big(\mathbf{Y}_1^{k-1};\mathbf{Y}_k\big)\Big) \tag{24}$$
$$= \frac{1}{n}\sum_{k=1}^{n}\Big(I\big(X_k,\mathbf{H}_1^{k-1};\mathbf{Y}_k\big) + \underbrace{I\big(M,\mathbf{Y}_1^{k-1};\mathbf{Y}_k\big|X_k,\mathbf{H}_1^{k-1}\big)}_{=0 \text{ by [14], [15]}} - I\big(\mathbf{Y}_1^{k-1};\mathbf{Y}_k\big)\Big) \tag{25}$$
$$\leq \frac{1}{n}\sum_{k=1}^{n}\Big(I\big(E_k,X_k,\mathbf{H}_1^{k-1};\mathbf{Y}_k\big) - I\big(\mathbf{Y}_1^{k-1};\mathbf{Y}_k\big)\Big) \tag{26}$$
$$= \frac{1}{n}\sum_{k=1}^{n}\Big(\underbrace{I(E_k;\mathbf{Y}_k)}_{\leq H_\mathrm{b}(\beta_k)} + I(X_k;\mathbf{Y}_k|E_k) + I\big(\mathbf{H}_1^{k-1};\mathbf{Y}_k\big|X_k,E_k\big) - I\big(\mathbf{Y}_1^{k-1};\mathbf{Y}_k\big)\Big) \tag{27}$$
$$\leq \frac{1}{n}\sum_{k=1}^{n}\Big(H_\mathrm{b}(\beta_k) + \beta_k I(X_k;\mathbf{Y}_k|E_k=1) + \beta_k I\big(\mathbf{H}_1^{k-1};\mathbf{Y}_k\big|X_k,E_k=1\big) - I\big(\mathbf{Y}_1^{k-1};\mathbf{Y}_k\big) + (1-\beta_k) I\big(X_k,\mathbf{H}_1^{k-1};\mathbf{Y}_k\big|E_k=0\big)\Big). \tag{28}$$

Here in (24) the current input X_k and the past fading values[1] $\mathbf{H}_1^{k-1}$ are added to the mutual information. In (26) we add the indicator random variable E_k that is defined as

$$E_k \triangleq \begin{cases} 1 & \text{if } |X_\ell| \geq \xi_{\min},\ \forall\, \ell = 1,\ldots,k, \\ 0 & \text{otherwise,} \end{cases} \tag{29}$$

for some given $\xi_{\min} > 0$. Moreover, $\beta_k \triangleq \Pr[E_k = 1]$. Finally, in (28) we bound

$$I(E_k;\mathbf{Y}_k) = H(E_k) - \underbrace{H(E_k|\mathbf{Y}_k)}_{\geq 0} \leq H(E_k) = H_\mathrm{b}(\beta_k) \tag{30}$$

with $H_\mathrm{b}(\cdot)$ denoting the binary entropy function.

One important part of the continued derivation will be to show that this upper bound strictly grows more slowly than double-logarithmically in E_s, unless the input is such that $\beta_k = 1$. Hence, in (28) eventually only three terms will contribute, and both the first and last term will disappear. These three terms ultimately will basically correspond to the three terms on the RHS of (13). However, note the difficulty of the order of limits here: before one can let E_s grow to infinity, one must

[1] This step shows that the proof still holds even if the feedback is improved in the sense that also $\mathbf{H}_1^{k-1}$ is presented to the transmitter.

handle the limit of $n \to \infty$. Therefore, one needs quite a bit of tricky bounding and the definition $\beta \triangleq \frac{1}{n}\sum_{k=1}^{n}\beta_k$ to reach a bound that is independent of k.

Using a similar approach as in [9], we bound the second term on the RHS of (28) as

$$\begin{aligned}
&\beta_k I(X_k;\mathbf{Y}_k|E_k=1)\\
&\quad\le \beta_k I\big(X_k;\|\mathbf{H}_k\||X_k|e^{\mathrm{i}\Theta_k}\big|E_k=1\big)\\
&\qquad+\beta_k I\big(X_k;\hat{\mathbf{H}}_k e^{\mathrm{i}(\Phi_k+\Theta_k)}\big|\|\mathbf{H}_k\||X_k|,\Theta_k,E_k=1\big)
\end{aligned}\tag{31}$$

with Φ_k denoting the phase of X_k. Here the first term will then be bounded with a duality-based upper bound [9].

The third term on the RHS of (28) is bounded as follows:

$$\begin{aligned}
&I\big(\mathbf{H}_1^{k-1};\mathbf{Y}_k\big|X_k,E_k=1\big)\\
&\quad\le I\big(\mathbf{H}_1^{k-1};\mathbf{Y}_k,\mathbf{H}_k\big|X_k,E_k=1\big) &(32)\\
&\quad= I\big(\mathbf{H}_1^{k-1};\mathbf{H}_k\big|X_k,E_k=1\big)\\
&\qquad+\underbrace{I\big(\mathbf{H}_1^{k-1};\mathbf{Y}_k\big|\mathbf{H}_k,X_k,E_k=1\big)}_{=0\text{ by [14], [15]}} &(33)\\
&\quad= h(\mathbf{H}_k|X_k,E_k=1)-h\big(\mathbf{H}_k\big|\mathbf{H}_1^{k-1},X_k,E_k=1\big) &(34)\\
&\quad= h(\mathbf{H}_k|X_k,E_k=1)-h\big(\mathbf{H}_k\big|\mathbf{H}_1^{k-1}\big), &(35)
\end{aligned}$$

where the last step follows because conditional on $\mathbf{H}_1^{k-1}$, $\mathbf{H}_k$ is independent of (X_k,E_k).

Finally, the fourth term on the RHS of (28) will be bounded as follows:

$$I\big(\mathbf{Y}_1^{k-1};\mathbf{Y}_k\big)\ge\beta_k I\big(\big\{\hat{\mathbf{H}}_\ell e^{\mathrm{i}\Theta_\ell}\big\}_{\ell=1}^{k-1};\hat{\mathbf{H}}_k e^{\Theta_k}\big|E_k=1\big)-\delta, \tag{36}$$

where it must be shown that δ is arbitrary small.

Note that we have missed out quite a few more problems. For example, one needs to condition all terms on $\|\mathbf{H}\|\le t$ for some t, in order to be able to show that $\mathsf{E}\big[\|\mathbf{H}_k\|^2|X_k|^2\big]$ is finite (and handle the case when $\|\mathbf{H}\|>t$ differently). Also, in the end, we need to be able to get rid of the conditional event $E_k=1$. For space reasons, we omit all these details and refer to a journal version of this paper that is under preparation.

6 Conclusion

We have shown that the asymptotic capacity of general SIMO regular fading channels with memory remains unchanged even if one allows causal noiseless feedback. This once again shows the extremely unattractive behavior of regular fading channels at high SNR: besides the double-logarithmic growth [8] and the very poor performance in a multiple-user setup (where the maximum sum-rate only can be achieved if all users apart from one *always* remain switched off [16]), we now see that any type of feedback does not increase capacity in spite of memory in the channel.

We would like to point out that the results presented here can be extended to the situation where both transmitter and receiver have access to causal partial side-information $\mathbf{S}_k$ about the fading, where by *partial* we mean that

$$\lim_{n\to\infty}\frac{1}{n}I\big(\mathbf{S}_1^n;\mathbf{H}_1^n\big)<\infty. \tag{37}$$

References

[1] İ. Emre Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, November–December 1999.

[2] Giuseppe Durisi, Ulrich G. Schuster, Helmut Bölcskei, and Shlomo Shamai (Shitz), "Noncoherent capacity of underspread fading channels," *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 367–395, January 2010.

[3] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Transactions on Information Theory*, vol. 45, no. 1, pp. 139–157, January 1999.

[4] Lizhong Zheng and David N. C. Tse, "Communicating on the Grassmann manifold: a geometric approach to the non coherent multiple-antenna channel," *IEEE Transactions on Information Theory*, vol. 48, no. 2, pp. 359–383, February 2002.

[5] Yingbin Liang and Venugopal V. Veeravalli, "Capacity of noncoherent time-selective Rayleigh-fading channels," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3095–3110, December 2004.

[6] Tobias Koch and Amos Lapidoth, "Degrees of freedom in non-coherent stationary MIMO fading channels," in *Proceedings Winter School on Coding and Information Theory*, Bratislava, Slovakia, February 20–25, 2005, pp. 91–97.

[7] Tobias Koch and Amos Lapidoth, "The fading number and degrees of freedom in non-coherent MIMO fading channels: a peace pipe," in *Proceedings IEEE International Symposium on Information Theory (ISIT)*, Adelaide, Australia, September 4–9, 2005, pp. 661–665.

[8] Amos Lapidoth and Stefan M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat fading channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2426–2467, October 2003.

[9] Stefan M. Moser, "The fading number of multiple-input multiple-output fading channels with memory," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2716–2755, June 2009.

[10] Stefan M. Moser, "Duality-based bounds on channel capacity," Ph.D. dissertation, ETH Zurich, October 2004, Diss. ETH No. 15769. [Online]. Available: http://moser.cm.nctu.edu.tw/publications.html

[11] Amos Lapidoth and Stefan M. Moser, "Feedback increases neither the fading number nor the pre-log," in *Proceedings Twenty-Third IEEE Convention of Electrical & Electronics Engineers in Israel (IEEEI)*, Herzlia, Israel, September 6–7, 2004, pp. 213–215.

[12] Amos Lapidoth and Stefan M. Moser, "On non-coherent fading channels with feedback," in *Proceedings Winter School on Coding and Information Theory*, Bratislava, Slovakia, February 20–25, 2005, pp. 113–118.

[13] Amos Lapidoth and Stefan M. Moser, "The fading number of single-input multiple-output fading channels with memory," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 437–453, February 2006.

[14] James L. Massey, "Causal interpretations of random variables," *Problemy Peredachi Informatsii (Problems of Information Transmission)*, vol. 32, no. 1, pp. 131–136, January–March 1996.

[15] James L. Massey, "Correction to 'Causal interpretations of random variables'," *Problemy Peredachi Informatsii (Problems of Information Transmission)*, May 16, 1997.

[16] Gu-Rong Lin and Stefan M. Moser, "The fading number of a multiple-access Rician fading channel," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 4983–4991, August 2011.