Presented at the 41st Annual Allerton Conference on Communication, Control, and Computing, Monticello IL, October 1 – 3, 2003

On the High SNR Capacity of Stationary Gaussian Fading Channels

Amos Lapidoth

Signal and Information Processing Laboratory
Swiss Federal Institute of Technology (ETH), Zurich
lapidoth@isi.ee.ethz.ch

Abstract

We present results on the high signal-to-noise ratio (SNR) asymptotic capacity of peak-power limited single-antenna discrete-time stationary complex-Gaussian fading channels with memory, where the transmitter and receiver, while fully cognizant of the fading law, have no access to its realization. Complementing recent results of Lapidoth & Moser about the case where the fading process is *regular*, we consider here the *non-regular* case, i.e., the case where the entropy-rate of the fading process is negative infinity.

It is demonstrated that while in the former case capacity grows double logarithmically in the SNR with a fading number that is determined by the prediction error in predicting the present fading from *noiseless observations* of its past, here the asymptotics require a careful analysis of the *noisy* prediction error, i.e., the asymptotic functional dependence of the prediction error in predicting the present fading from *noisy observations* of its past. This functional dependence, which can be made explicit in terms of the spectrum of the fading process, may lead to dramatically different asymptotic dependencies of capacity on SNR, e.g., double-logarithmic, logarithmic, or fractional powers of the logarithm of the SNR.

The "pre-log", i.e., the asymptotic ratio of channel capacity to the logarithm of the SNR takes on a particularly simple form: it is the Lebesgue measure of the set of harmonics where the spectral density is zero. The pre-log is unrelated to the "bandwidth" of the process or its "coherence time".

In the light of these results we re-examine some of the models in the literature on fading channels and the asymptotic behaviors associated with them. It is found that what may appear as slight changes in the channel model may lead to dramatically different high-SNR asymptotics.

1 Introduction

This paper addresses the capacity of a single-antenna fading channel whose time-k complex-valued output $Y_k \in \mathbb{C}$ is given by

$$Y_k = \big(d + H_k\big)x_k + Z_k, \tag{1}$$

where $x_k \in \mathbb{C}$ is the time-k complex-valued channel input; the complex valued "fading process" $\{H_k\}$ models multiplicative noise; the "specular component" $d \in \mathbb{C}$ is deterministic; and the complex process $\{Z_k\}$ models additive noise. The processes $\{H_k\}$ and

$\{Z_k\}$ are assumed to be independent and of a joint law that does not depend on the input sequence $\{x_k\}$.

We shall assume that the additive noise $\{Z_k\}$ is independent and identically distributed (IID) circularly-symmetric complex Gaussian random variables of zero mean and variance σ^2. Thus $Z_k \sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2)$ where we use the notation $W \sim \mathcal{N}_{\mathbb{C}}(\mu, \sigma^2)$ to indicate that $W - \mu$ has a zero-mean variance-σ^2 circularly-symmetric complex-Gaussian distribution, i.e., to indicate that the density $f_W(w)$ of W is given by

$$f_W(w) = \frac{1}{\pi\sigma^2} e^{-\frac{|w-\mu|^2}{\sigma^2}}, \quad w \in \mathbb{C}. \tag{2}$$

The fading process $\{H_k\}$ is assumed to be a unit-variance zero-mean stationary circularly-symmetric Gaussian process of general *spectral distribution function* $F(\lambda)$, $\lambda \in [-1/2, 1/2]$. Thus, $F(\cdot)$ is a monotonically non-decreasing function on the interval $[-1/2, 1/2]$ [1, Theorem 3.2, p. 474],

$$\mathsf{E}\left[H_{k+m} H_k^*\right] = \int_{-1/2}^{1/2} e^{i2\pi m\lambda} \, \mathrm{d}F(\lambda), \quad k, m \in \mathbb{Z}, \tag{3}$$

and

$$\mathsf{E}\left[|H_k|^2\right] = 1. \tag{4}$$

Being monotonic, the function $F(\cdot)$ is almost everywhere differentiable, and we denote its derivative by $F'(\cdot)$. If $F(\cdot)$ is absolutely continuous with respect to the Lebesgue measure on $[-1/2, 1/2]$ then its derivative $F'(\lambda)$ is called the *spectral density* of the fading process. In this case the process $\{H_k\}$ is also guaranteed to be ergodic.

We denote the mean squared-error in predicting H_0 from the infinite past $\{H_k\}_{k=-\infty}^{-1}$ by ϵ^2_{pred}. It is given by

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

We say that the process is *regular* if $\epsilon^2_{\text{pred}} > 0$. Otherwise we say that the process is *non-regular*.

In analyzing the capacity of this channel we shall assume that while the statistics of the channel — namely d, $F(\cdot)$, and σ^2 — are fully known at the transmitter and receiver, neither transmitter nor receiver has knowledge of the realization of the fading process.

Without any constraints on the input the capacity of this channel is infinite. We shall thus be interested in two different input constraints: a peak power constraint and an average power constraint. In the former we require that, at no time k, will the magnitude of the input exceed the peak power A:

$$|x_k| \leq A, \quad k \in \mathbb{Z}. \tag{5}$$

In this case the capacity $C_{\text{PP}}(A)$ is given by

$$C_{\text{PP}}(A) = \lim_{n\to\infty} \frac{1}{n} \sup I(X_1, \ldots, X_n; Y_1, \ldots Y_n) \tag{6}$$

where the supremum is over all joint distributions on $X_1, \ldots, X_n$ satisfying the peak constraint (5), and where the limit exists because $\{H_k\}$ was assumed stationary. If instead of a peak-power constraint we impose an average power constraint, the capacity C_{Avg} is given by

$$C_{\text{Avg}}(\mathcal{E}_{\text{s}}) = \lim_{n\to\infty} \frac{1}{n} \sup I(X_1,\ldots,X_n;Y_1,\ldots Y_n) \tag{7}$$

where the supremum is over all joint distributions on $X_1,\ldots,X_n$ satisfying the average power constraint

$$\sum_{k=1}^{n} \mathsf{E}\left[|X_k|^2\right] \leq n \cdot \mathcal{E}_{\text{s}}.$$

In the case of a peak power constraint we define the signal-to-noise ratio SNR as A^2/σ^2 whereas in the average power constraint we define the SNR as $\mathcal{E}_{\text{s}}/\sigma^2$. Thus,

$$\text{SNR} = \begin{cases} \frac{A^2}{\sigma^2} & \text{peak power} \\ \frac{\mathcal{E}_{\text{s}}}{\sigma^2} & \text{average power} \end{cases}. \tag{8}$$

2 Regular Fading

The high SNR channel asymptotics for regular fading were studied in [2]. There it was shown that for *general* finite entropy rate stationary fading processes capacity is given by

$$C = \log\log \text{SNR} + \log \pi + \mathsf{E}\left[\log |H_k|^2\right] - h\left(\{H_k\}\right) + o(1) \tag{9}$$

where the $o(1)$ term tends to zero as the SNR tends to infinity, and where — with the SNR being defined as in (8) — the asymptotics hold irrespective of whether a peak power or an average power input constraint is imposed. Specializing this result to the case of Gaussian fading [2] one obtains the expansion

$$C = \log\log \text{SNR} + \log |d|^2 - \text{Ei}(-|d|^2) - 1 + \log \frac{1}{\epsilon^2_{\text{pred}}} + o(1) \tag{10}$$

where $\text{Ei}(\cdot)$ denotes the exponential integral function

$$\text{Ei}(-x) = -\int_x^{\infty} \frac{e^{-t}}{t}\, \mathrm{d}t, \qquad x > 0. \tag{11}$$

It is based on (10) that Lapidoth & Moser proposed that the key to the high SNR asymptotics of regular Gaussian fading is the prediction error ϵ^2_{pred}.

3 Non-Regular Fading

Recall that the fading is non-regular if $\epsilon^2_{\text{pred}} = 0$. This is, for example, the case whenever F' is zero over a finite interval, e.g., when the process is bandlimited. The key to the asymptotics of non-regular fading channels is the noisy prediction error $\epsilon^2_{\text{pred}}(\delta^2)$. It is the mean squared-error in predicting H_0 from *noisy* observations of the past

$$H_{-1} + W_{-1}, H_{-2} + W_{-2}, H_{-3} + W_{-3}, \cdots$$

where $\{W_k\}$ is independent of $\{H_k\}$ and IID $\mathcal{N}_{\mathbb{C}}\left(0,\delta^2\right)$ distributed. Explicitly

$$\epsilon^2_{\text{pred}}(\delta^2) = \exp\left\{\int_{-1/2}^{1/2} \log\bigl(F'(\lambda) + \delta^2\bigr)\, \mathrm{d}\lambda\right\} - \delta^2. \tag{12}$$

The analysis of the asymptotics is based on the following two bounds [3]. The first is the upper bound

$$C_{\mathrm{PP}}(\mathrm{SNR}) \leq C_{\mathrm{PP}}^{\mathrm{IID}}(\mathrm{SNR}) + \log \frac{1}{\epsilon^2_{\mathrm{pred}}(1/\mathrm{SNR})} \tag{13}$$

$$= \log\log \mathrm{SNR} + \log |d|^2 - \mathrm{Ei}(-|d|^2) - 1 + \log \frac{1}{\epsilon^2_{\mathrm{pred}}(1/\mathrm{SNR})} + o(1) \tag{14}$$

where $C_{\mathrm{PP}}^{\mathrm{IID}}(\mathrm{SNR})$ is the peak-limited capacity of the corresponding memoryless Ricean channel, and where the equality follows from the analysis of this term at high SNR; see [2]. The lower bound is somewhat loose, but suffices for our asymptotic analysis

$$C_{\mathrm{PP}}(\mathrm{SNR}) \geq \log \frac{1}{\epsilon^2_{\mathrm{pred}}(4/\mathrm{SNR}) + 8/(5\mathrm{SNR})} + \log |d|^2 - \mathrm{Ei}\left(-\frac{|d|^2}{1-\epsilon^2_{\mathrm{pred}}(4/\mathrm{SNR})}\right) - \log\frac{5e}{6}. \tag{15}$$

These bounds and (12) have some consequences that we next describe.

Log-Log Growth: As mentioned earlier, for regular processes capacity grows double logarithmically in the SNR. It turns out that this may be the case even for some non-regular fading processes. In fact, with the aid of the above bounds we can fully characterize the spectra that lead to a double logarithmic growth of capacity.

$$\varlimsup_{\mathrm{SNR}\to\infty} \frac{C(\mathrm{SNR})}{\log\log\mathrm{SNR}} < \infty \iff \varlimsup_{\delta^2 \downarrow 0} \frac{-\int_{-1/2}^{1/2} \log\big(F'(\lambda)+\delta^2\big)\,\mathrm{d}\lambda}{\log\log\frac{1}{\delta^2}} < \infty. \tag{16}$$

Pre-Log: The limiting value of the ratio of channel capacity to $\log \mathrm{SNR}$ has lately received much attention under the heading of "multiplexing gain". We prefer the term "pre-log" because it seems more self-explanatory and because for SISO channels it cannot exceed the value of 1, which corresponds to non-fading channels. With the aid of the new bounds we can give a simple characterization of the pre-log: it is the Lebesgue measure of the set of zeros of the derivative of the spectral distribution function:

$$\boxed{\lim_{\mathrm{SNR}\to\infty} \frac{C_{\mathrm{PP}}(\mathrm{SNR})}{\log\mathrm{SNR}} = \mu\Big(\Big\{\lambda : F'(\lambda) = 0\Big\}\Big),} \tag{17}$$

where $\mu(\cdot)$ denotes the Lebesgue measure on the interval $[-1/2, 1/2]$.

Other Asymptotics: It turns out that the capacity asymptotics are not limited to logarithmic or double-logarithmic growths in the SNR. For example, using these bounds it is shown in [3] that for any $0 < \beta < 1$ and any $0 < \omega < 1/2$ there exist spectral distribution functions such that

$$\lim_{\mathrm{SNR}\to\infty} \frac{C(\mathrm{SNR})}{\big(\log\mathrm{SNR}\big)^{\beta}} = \frac{2\omega}{\beta}, \quad 0 < \beta < 1, \quad 0 < \omega < 1/2. \tag{18}$$

Where does the Log-Log Region Start?: The new bounds are not only useful in the study of non-regular fading. They can also be used for a finer analysis of the regular fading asymptotics. We mention here one such possible use. In [2] it was suggested to use (10) as an indication of the rates above which capacity grows only double-logarithmically in the SNR. Thus, it was suggested that a system operating at rates that are significantly higher than the fading number

$$\chi = \log |d|^2 - \mathrm{Ei}(-|d|^2) - 1 + \log \frac{1}{\epsilon^2_{\text{pred}}} \tag{19}$$

is operating in the log-log SNR regime. This characterization of the double-logarithmic regime in terms of *rates* gives no indication of the corresponding SNRs. With the aid of the new bounds one can give an alternative indication of the log-log regime. Namely, an SNR is in the log-log region if the prediction error $\epsilon^2_{\text{pred}}(1/\text{SNR})$ is "close" to the noiseless prediction error ϵ^2_{pred}, e.g., if

$$\epsilon^2_{\text{pred}}(1/\text{SNR}) < 2\epsilon^2_{\text{pred}}(0). \tag{20}$$

To this one should add, of course, that the SNR must be large enough so that the first term on the RHS of (13) be in the double-log regime. For example [2], [4],

$$\log\left(1 + \frac{\text{SNR} \cdot |d|^2}{1 + \text{SNR}}\right) \approx \log\left(1 + |d|^2\right).$$

4 Comments, Questions, and Other Models

In this section we shall comment on some of the above results, raise some questions, and address other models that relate to this paper.

1. *Asymptotics and Bandwidth:* The noisy prediction error $\epsilon^2_{\text{pred}}(\delta^2)$, the noiseless prediction error ϵ^2_{pred}, and the Lebesgue measure of the set $\{\lambda : F'(\lambda) = 0\}$ are all invariant under a reordering of the set of harmonics $[-1/2, 1/2]$. Consequently, there can be no direct relationship between the "bandwidth" of $\{H_k\}$ and the channel capacity asymptotics. If the process is bandlimited, capacity will grow logarithmically in the SNR, but the pre-log will depend not only on the bandwidth but also on the Lebesgue measure of the set of in-band harmonics where the spectral density is zero.

2. *Asymptotics and Snippets of the Autocorrelation:* No matter how large p is, the first $p + 1$ values of the autocorrelation

$$\{\mathsf{E}\left[H_m H_0^*\right]\}_{m=0}^{p}$$

typically say nothing about the channel asymptotics. This can be attributed to the fact that, unless these values are not of full rank, the sequence can be extended to a regular autocorrelation (e.g., to that of the max-entropy process) thus yielding a $\log\log \text{SNR}$ growth or to an autocorrelation whose spectral measure consists of a finite number of atoms and which thus leads to a $\log \text{SNR}$ growth. More precisely:

Fact 1. *The maximum entropy rate stochastic process $\{X_k\}$ satisfying the constraint*

$$\mathsf{E}\left[X_k X^*_{k+m}\right] = r_m, \qquad m = 0, \ldots, p, \qquad \textit{for all integer } k \tag{21}$$

is the p-th order Gauss-Markov process. The differential entropy rate of this maximum entropy process is finite if, and only if, the covariance matrix of the vector $X_0, \ldots, X_p$ is non-singular.

A minimum entropy rate stochastic process $\{X_k\}$ is a Gaussian process of the form

$$X_k = \sum_{\ell=1}^{p+1} \alpha_\ell X_{k-\ell}$$

for some constants $\alpha_1, \ldots, \alpha_{p+1}$. This process is typically non-ergodic and is of a spectral distribution function $F(\cdot)$ which is a step function with a finite number of steps. It's derivative $F'(\lambda)$ is thus almost everywhere zero. The entropy rate of this minimum entropy process is $-\infty$.

Moreover, if the vector $(X_0, \ldots, X_p)^{\mathsf{T}}$ has a covariance matrix (induced by the constraint (21)) that is non-singular, then for any $M > 0$ there exists an ergodic $(p+1)$-th order Gauss-Markov process that satisfies (21) and which has a differential entropy rate smaller than $-M$.

3. *Testing for Regularity:* The above fact also demonstrates the difficulty in designing a statistical test for regularity. Indeed, there is no non-trivial unbiased hypothesis-testing procedure for determining whether a finite number of samples come from a regular Gaussian process or from a non-regular Gaussian process. In other words, in any statistical procedure that does not ignore the data, there will always be a non-regular process that is more likely to be classified as "regular" than some regular process.

This can be formalized as follows: Assume that we have no measurement noise, and that we can obtain clean samples from the realization of a fading channel. Let $p+1$ denote the number of samples $H_0, \ldots, H_p$ obtained from the fading process. Moreover assume that the fading is known to be a stationary Gaussian process. Based on these measurements, we would like to determine whether the fading process is regular or not. We shall show that this hypothesis testing problem does not admit a non-trivial unbiased test. To be more formal, denote by Ω the set of all stationary Gaussian processes, and by $\Omega_{\mathcal{R}}$ its subset of regular processes. Denote by $\Omega_{\bar{\mathcal{R}}} = \Omega \setminus \Omega_{\mathcal{R}}$ the set of non-regular processes. Recall that a hypothesis test is simply a set $\mathcal{S}_0 \subset \mathbb{C}^{p+1}$ where we declare "$\mathcal{R}$" if $(H_0, \ldots, H_p) \in \mathcal{S}_0$, and declare "$\bar{\mathcal{R}}$" otherwise. The "size of the test" α is, by definition,

$$\alpha = \sup_{\theta \in \Omega_{\mathcal{R}}} \Pr(\text{“}\bar{\mathcal{R}}\text{”} | \theta).$$

The test is *unbiased* if

$$\Pr(\text{“}\bar{\mathcal{R}}\text{”} | \theta) \geq \alpha, \qquad \text{for all } \theta \in \Omega_{\bar{\mathcal{R}}}.$$

To quote [5]: "unless this condition is satisfied, there will exist alternatives under which the acceptance of the hypothesis is more likely than in some cases in which the hypothesis is true". We shall show that the only unbiased tests for this problem declare "$\mathcal{R}$" with a probability that does not depend on $\theta \in \Omega$. These unbiased tests are, in other words, (possibly randomized) tests that ignore the data.

To prove this denote for any process $\theta \in \Omega_{\mathcal{R}}$ by $\theta_{\min} \in \Omega_{\bar{\mathcal{R}}}$ the Gaussian min-entropy process whose autocorrelation agrees with that of θ in the first $p+1$ values. Thus, under $\theta \in \Omega_{\mathcal{R}}$, and $\theta_{\min} \in \Omega_{\bar{\mathcal{R}}}$ the joint distributions of $(H_0, \ldots, H_p)$ are identical. In particular

$$\Pr(\text{“}\bar{\mathcal{R}}\text{”} | \theta) = \Pr(\text{“}\bar{\mathcal{R}}\text{”} | \theta_{\min}), \qquad \theta \in \Omega_{\mathcal{R}}. \tag{22}$$

We shall show that if a test is unbiased then

$$\Pr(\text{“}\bar{\mathcal{R}}\text{”} | \theta) = \Pr(\text{“}\bar{\mathcal{R}}\text{”} | \theta'), \qquad \theta, \theta' \in \Omega_{\mathcal{R}}. \tag{23}$$

This will imply that the probability of declaring "$\bar{\mathcal{R}}$" does not depend on the covariance matrix of $H_0, \ldots, H_p$ so that the test is trivial.

To prove that any unbiased must satisfy (23) assume by contradiction that for some $\theta, \theta' \in \Omega_{\mathcal{R}}$ we have

$$\Pr(\text{“}\bar{\mathcal{R}}\text{”}|\theta) < \Pr(\text{“}\bar{\mathcal{R}}\text{”}|\theta').$$

Then by (22) we have:

$$\begin{aligned}\Pr(\text{“}\bar{\mathcal{R}}\text{”}|\theta_{\min}) &= \Pr(\text{“}\bar{\mathcal{R}}\text{”}|\theta)\\ &< \Pr(\text{“}\bar{\mathcal{R}}\text{”}|\theta')\end{aligned}$$

but $\theta_{\min} \in \Omega_{\bar{\mathcal{R}}}$ and $\theta' \in \Omega_{\mathcal{R}}$, so that the test is biased.

It should be noted that the absence of non-trivial unbiased test does not, in itself, prove that the testing is impossible. It does demonstrates some of the difficulties involved.

4. *Asymptotics via Physics:* The difficulty of using statistical procedures to test for regularity and for the corresponding asymptotics motivate considering the physics involved. It is interesting that Jakes's model for continuous-time fading [6] models the fading as a bandlimited process, i.e., a non-regular process. There is a need to study the discrete-time analog of this model and examine whether it also leads to a non-regular fading process. A step in this direction was recently taken in [7].

5. *"Real World Numbers":* Before we can answer such basic questions as whether the typically encountered fading is regular, it will be extremely difficult to come up with approximate numbers that quantify the SNRs at which the system's capacity exhibits various behaviors. An attempt at numerically quantifying the region in which systems exhibit a $\log\log \text{SNR}$ capacity dependence was recently made in [8]. The authors used such measures as "coherence bandwidth", "RMS delay spread", "coherence time", and "maximum Doppler shift" to estimate the number of samples L required before the autocorrelation of the fading process decreases below half the variance. The authors then fit a first-order Gauss-Markov process accordingly. That is, they found the innovation variance that would result in a first-order Gauss-Markov process having an autocorrelation function that decays to half the variance in the above specified number of samples L. The authors then studied the asymptotic of this Gauss-Markov model to determine the SNR at which the $\log\log \text{SNR}$ behavior of channel capacity begins.

This approach requires a leap of faith. As mentioned above, no snippet of the autocorrelation can determine the channel asymptotics. Thus, there are numerous different processes whose autocorrelation decays to half the variance in L symbols. These have arbitrary noiseless prediction errors. Some lead to logarithmic capacity growths, some to double-logarithmic, etc. In fact, as we shall see, even if we knew that the process is regular, and even if we knew its noiseless prediction error (which we cannot!), there would still be no way to determine based on this data alone the SNR at which the double-logarithmic region begins.

6. *The Marzetta-Hochwald Model:* The fading model proposed by Marzetta & Hochwald [9] is parameterized by one parameter T, which is a positive integer. The model, also known as the "block-constant" model, assumes that once every T symbols the fading level is picked from a zero-mean unit-variance circularly symmetric Gaussian distribution and remains constant at this level for the next T fading symbols. After T symbols the fading level is drawn again from the above distribution (independently of previous fading values) and the resulting value is the fading level for the next T symbols, whence a new level is drawn, etc. While [9] focused on capacity at moderate SNR, Zheng and Tse [10] derived the high SNR asymptotics. They showed that for $T > 1$ capacity grows logarithmically in the SNR, with the pre-log being $(T-1)/T$. It is thus seen that this model behaves very differently from the stationary Gaussian fading model with regular fading. Indeed, in the latter model capacity always grows double logarithmically in the SNR; in the former only when the fading is memoryless ($T = 1$).

Moreover, this model also behaves very differently from the stationary Gaussian fading model with non-regular fading. The pre-log in the Marzetta-Hochwald model is always in the set $\{0, 1/2, 2/3, 3/4, \cdots\}$, whereas $\mu(\{\lambda : F'(\lambda) = 0\})$ can be any number between zero and one.

We also note here that the Marzetta-Hochwald model is a non-stationary model. This is sometimes justified by assuming a frequency hopping system, i.e., that the system hops to a new frequency every T symbols. This justification is addressed next

7. *Frequency Hopping:* The Marzetta-Hochwald model is often justified by assuming a frequency hopping system. We note here that in addition to limiting the scope of the results, some difficulties remain. First note that in comparison to the regular stationary fading model, this model is overly optimistic (in predicting a logarithmic rather than double-logarithmic growth). Since frequency hopping destroys memory, it reduces capacity, so that in the absence of frequency hopping the discrepancy with the regular fading model would only be greater.

Secondly, if we are to seriously consider a frequency hopping system, we must conclude that the Marzetta-Hochwald model is appropriate only if we are willing to model the fading in each frequency as constant over time. This corresponds to a constant autocorrelation for which capacity grows with a unit pre-log[1]. We would thus have to conclude that if the system were not hopping we could have achieved, at each of the frequencies, a unit pre-log. Under such conditions one would have great difficulty justifying the use of frequency hopping.

In short, the Marzetta-Hochwald model makes two idealizations: that the fading is independent from block to block and that it is constant within a block. It is the latter, which is not justifyable by frequency hopping, that leads to the overly optimistic asymptotics.

8. *The Liang-Veeravalli Model:* The Liang-Veeravalli model extends the Marzetta-Hochwald model to allow for the T fading values in each duration-T block to be jointly Gaussian with a covariance matrix of rank Q. They show that for $Q = T$ one obtains a double-logarithmic capacity growth, whereas for $Q < T$ the growth is logarithmic with a $(T-Q)/T$ pre-log. This certainly allows for richer pre-logs, though never non-rational. In addition, the model does not allow for growths of the form $(\log \mathrm{SNR})^{\beta}$, for $0 < \beta < 1$.

The richness of the model makes it somewhat difficult to pick its parameters. As we have discussed, this is also a problem that haunts the stationary Gaussian fading model.

9. *Stationarity:* In both the Marzetta-Hochwald and the Liang-Veeravalli the fading is modeled as a non-stationary process. Is this physically motivated or is it done for mathematical convenience alone? We suspect the latter. Stationary models with memory are far less prevalent. See [2], [3] and the first-order Gauss-Markov model [11], [8].

10. *The First Order Gauss-Markov Model:* The capacity of a fading channel where the

[1]This corresponds to a non-ergodic channel but one for which one can achieve a unit pre-log with arbitrarily high probability.

fading process is a first order Gauss-Markov model was studied in [11] and [8]. In the former the focus was on Gaussian inputs. In [8] it was demonstrated that the capacity-SNR function of a Gauss-Markov fading model can be roughly divided into three regions:

$$C \approx \begin{cases} \log \mathrm{SNR} & \mathrm{SNR} < 1/\epsilon^2_{\mathrm{pred}} \\ \log 1/\epsilon^2_{\mathrm{pred}} & 1/\epsilon^2_{\mathrm{pred}} \leq \mathrm{SNR} < e^{1/\epsilon^2_{\mathrm{pred}}} \\ \log\log \mathrm{SNR} & \mathrm{SNR} \geq e^{1/\epsilon^2_{\mathrm{pred}}} \end{cases}. \tag{24}$$

We would like here to comment on the dangers in extrapolating the classification and boundaries of (24) to general stationary Gaussian fading models.

We begin by noting that the Gauss-Markov model falls under the category of regular stationary Gaussian fading[2]. The asymptotics of channel capacity are thus always double-logarithmic. The model will thus never predict a logarithmic growth of channel capacity, or any fractional power thereof. Indeed, for non-regular processes the asymptotics need not be double-logarithmic, so that it is questionable whether one would get such three regions.

The reason that caution must be exercised in extrapolating from this model, is that the model is not very rich. It is based on only one parameter — the variance of the innovation. Thus, quantities that for general Gaussian processes are unrelated become rigidly tied in the Gauss-Markov model. Indeed, if one is restricted to a one-parameter family, almost any non-trivial quantity determines all others. To demonstrate the need for caution consider the problem of determining the region in which capacity grows double-logarithmically in the SNR. Specializing the result of [2] to the Gauss-Markov case, one obtains that Lapidoth & Moser would characterize the region as the one where the rate is significantly higher than the RHS of (19), which for $d = 0$ equates to $-1 - \gamma - \log \epsilon^2_{\mathrm{pred}}$. (For Gauss-Markov processes the noiseless prediction error is simply the variance of the innovation process.)

Etkin and Tse specify the region in terms of SNR. Namely, the double-logarithmic region corresponds to an SNR greater than $e^{1/\epsilon^2_{\mathrm{pred}}}$. Notice that for Gauss-Markov processes the two characterizations are by (24) essentially the same. The difficulty lies in extrapolating to general fading. The characterization in terms of SNR is problematic. The reason is that as one can see from the new bounds, the SNR at which the double-logarithmic behavior begins is related to the noisy prediction error $\epsilon^2_{\mathrm{pred}}(\delta^2)$ and not directly to its limiting value $\epsilon^2_{\mathrm{pred}} = \epsilon^2_{\mathrm{pred}}(0)$; hence the motivation for (20).

Note, however, that while for the Gauss-Markov process the value of the noiseless prediction error $\epsilon^2_{\mathrm{pred}}$ determines the noisy one, this is generally not the case. The SNR at which (20) holds with equality can be arbitrarily higher than $e^{1/\epsilon^2_{\mathrm{pred}}}$. Thus, while for Gauss-Markov fading the characterization of the double-logarithmic region in terms of rates (19) or in terms of SNR (20) both agree with the characterization as $\mathrm{SNR} > e^{1/\epsilon^2_{\mathrm{pred}}}$ this is in general not the case.

Great caution must be exercised extrapolating the $\mathrm{SNR} > e^{1/\epsilon^2_{\mathrm{pred}}}$ characterization to non Gauss-Markov fading. Ditto for the other two regions.

11. *The Gaussian Assumption:* Can the fading be modeled as Gaussian at high SNR? While (9) holds in general, our study of the non regular fading relies heavily on this assumption. This assumption is also made in all the models we discussed above.

---

[2]In this discussion we assume that the Gauss-Markov is not degenerate, i.e., that the innovation is of positive variance so that the process is not constant over time.

12. *Peak vs. Average Power:* Our study was limited to a peak power constraint. Do the asymptotics change for average power constraints? They do not for regular fading or for non-fading channels. Even if they do, one could argue that when they differ, the peak constraint is more important because in such cases the asymptotics of the average power constraint can only be achieved by input distributions with a peak-to-average ratio tending to infinity.

References

[1] J. Doob, *Stochastic Processes.* John Wiley & Sons, 1990.

[2] A. Lapidoth and S.M. Moser, "Capacity bounds via duality with applications to multi-antenna systems on flat fading channels," *IEEE Trans. on Inform. Theory*, Vol. 49, No. 10, October 2003, in press.

[3] A. Lapidoth, "On the asymptotic capacity of fading channels," `submitted 2003.` Available at:
http://www.isi.ee.ethz.ch/archive/publications/publ_sig.en.php?isipap/alap

[4] A. Lapidoth and S. Shamai(Shitz), "Fading channels: how perfect need "perfect side information" be?," *IEEE Trans. on Inform. Theory*, vol. 48, no. 5, pp. 1118–1134, May 2002.

[5] E. L. Lehmann, *Testing Statistical Hypotheses.* Pacific Grove, CA: Wadsworth & Brooks, second ed., 1991.

[6] W.C. Jakes Jr., *Microwave mobile communications.* New York: Wiley, 1974.

[7] Y. Liang and V. V. Veeravalli, "Capacity of noncoherent time-selective Rayleigh fading channels," `submitted to` *IEEE Trans. on Inform. Theory*, 2003.

[8] R. Etkin and D. Tse, "Degrees of freedom in underspread MIMO fading channels," `preprint`, 2003.

[9] T. Marzetta and B. Hochwald, "Capacity of a multiple-antenna communication link in Rayleigh flat fading environment," *IEEE Trans. on Inform. Theory*, vol. 45, pp. 139–157, Jan. 1999.

[10] L. Zheng and D. Tse, "Communicating on the Grassmann manifold: A geometric approach to the non-coherent multiple antenna channel," *IEEE Trans. on Inform. Theory*, vol. 48, pp. 359–383, Feb. 2002.

[11] R.-R. Chen, B. Hajek, R. Koetter, and U. Madhow, "On fixed input distributions for noncoherent communication over high SNR Rayleigh fading channels," `submitted to` *IEEE Trans. on Inform. Theory*, 2002.