# ON THE BOUNDS OF THE NON-COHERENT CAPACITY OF GAUSS-MARKOV FADING CHANNELS

*Jianqiu Zhang*

Dept. of ECE, University of New Hampshire, Durham, NH 03824
e-mail: jianqiu.zhang@unh.edu

### ABSTRACT

In this paper, we discuss the upper and lower bounds of the flat fading Gauss-Markov channel capacity. It is shown that due to phase variations on correlated channel, the amount of information carried by phase will not grow with SNR asymptotically. As a result, optimal coherent designs or designs based on quasi-static assumptions can not be optimal for correlated fading channels. We also show that channel correlation may provide significant capacity gain over the capacity of independent fading channels. It is also shown that if the input is i.i.d., the asymptotic growth rate of the capacity is the same as that of independent Rayleigh flat fading channels. This indicates that the asymptotic region is very power inefficient.

## 1. INTRODUCTION

Recently, the calculation of non-coherent fading channel capacities has received a lot of attention. In [3], an upper and lower bound on independent Rayleigh flat fading channels are derived. It is shown that the memoryless Rayleigh flat fading channel capacity will grow with SNR double logarithmically. The upper and lower bounds are further tightened in [1]. In [4], it is proven that the input distribution that achieves the capacity on a memoryless Rayleigh flat fading channel should be discrete with a finite number of mass points. While the memoryless fading model is appropriate in certain channels such as the frequency hopping channels, or when large time guards are inserted between symbols, the Gauss-Markov flat fading model is more appropriate for a larger class of channels with memory. In [5], it is shown that Gaussian input distributions will generate bounded mutual information on Gauss-Markov channels. Most recently, in [2], an upper and lower bound are derived for correlated Rayleigh fading channel in a single-input-single-output (SISO) system. However the detailed capacity characteristics of correlated fading channel is still not known.

In this paper, we focus on the capacity of Rayleigh flat fading channels with memory modeled as Gauss-Markov channels. Only the average power constraint is considered. We derived a tighter upper bound on channel capacity. Through our analysis, we have gained some insight to the structure of the capacity. We conclude that phase variations on correlated fading channels will limit the amount of capacity carried by the phase. The phase capacity will not grow with SNR asymptotically. This shows that previous designs based on the assumption of constant phase within a block of data can not be optimal for correlated fading channels.

On the other hand, the analysis also points out that the capacity due to channel correlation on correlated channels could be significant when compared to the capacity of independent fading channels

We also show that the asymptotic growth rate of the capacity must be the same as that of a memoryless Rayleigh flat fading channel if the input is i.i.d. which is very power inefficient. Therefore asymptotic regions should be avoided if possible.

This paper is structured as the following. In section 2, we describe the formulation of the problem. In section 3, we derive an upper and lower bound of the Gauss-Markov channel capacity. It is shown that the growth rate with SNR is the same as that of a memoryless Rayleigh flat fading channel.

## 2. PROBLEM FORMULATION

Consider the following simple model of a flat fading channel with memory.

$$Y_i = S_i X_i + V_i, \tag{1}$$

where $X_i$ is the input signal at symbol time $i$, $V_i$ is the additive complex circular Gaussian noise at the receive antenna with zero mean and unit variance, and $Y_i$ is the output signal at the receiver. The channel states or the fading coefficients, $S_i$s, are modeled as samples of a complex Gauss-Markov process with zero mean and a covariance matrix $\Sigma_{Sn}$. The marginal distribution of the channel states should be complex Gaussian with zero mean and unit variance. The capacity of the channel described above is given by the supreme of the averaged mutual information over $n$ channel uses as

$n \to \infty$. We can write the average mutual information as

$$I_n = \frac{1}{n} I(X^n, Y^n) = \frac{1}{n}[h(Y^n) - h(Y^n|X^n)], \quad (2)$$

where $X^n = \{X_1, \cdots, X_n\}$ represents the input sequence and $Y^n = \{Y_1, \cdots, Y_n\}$ represents the sequence of output symbols. We use the same convention to represent other sequences throughout of the paper. Note that the input sequence has an average power constraint $P$, $\frac{1}{n}\sum_{i=1}^{n} X_i^2 \leq P$. Let $n_1 = \{i_1, i_2, \cdots, i_{n_1}\}$ represent the set of indices such that if $i \in n_1$, then $X_i \neq 0$, let $n_2$ represent the set of indices such that if $i \in n_2$, then $X_i = 0$. In the following text, $|\cdot|$ represents the size of the set of indices if the operand is a set, and if the operand is a complex number, it represents the amplitude of the complex number. Using the above notations, the conditional differential entropy $h(Y^n|X^n)$ can be written as

$$\begin{aligned} h(Y^n|X^n) &= h(Y^{n_1}, Y^{n_2}|X^n) \\ &= h(Y^{n_1}|X^n) + h(Y^{n_2}|X^n, Y^{n_1}) \\ &= h(Y^{n_1}|X^n) + h(V^{n_2}). \quad (3) \end{aligned}$$

## 3. UPPER BOUND

Lemma 1:

$$h(Y^{n_1}|X^n) \geq h((SX)^{n_1}|X^n), \quad (4)$$

where $(SX)^{n_1}$ represents the sequence of squared output amplitude before the observation noise.

Proof: We can write $h(Y^{n_1}|X^n)$ as $h((SX+V)^{n_1}|X^n)$, and the sequence $(SX+V)^{n_1}$ is complex Gaussian with zero mean and a covariance matrix

$$\Sigma_y = 1_{|n_1|} + \mathbf{X}'\Sigma_{Sn_1}\mathbf{X},$$

where $\mathbf{X} = [X_1, \cdots, X_{|n_1|}]^T = X^{n_1}$ is the vector representation of the non-zero input sequence, $\Sigma_{Sn_1}$ represent the covariance matrix of the channel states when the inputs are non-zero, $1_{|n_1|}$ stands for an identity matrix with $|n_1|$ dimensions, and $()'$ stands for the Hermitian transpose of the operand. Similarly, $(SX)^{n_1}$ is also a complex Gaussian sequence with zero mean and a covariance matrix

$$\Sigma_{sx} = \mathbf{X}'\Sigma_{Sn_1}\mathbf{X}.$$

Since $\Sigma_y - \Sigma_{sx} = 1_{|n_1|}$ is positive definite, using Lemma [10.6.2] in [6], we have $|\Sigma_y| \geq |\Sigma_{sx}|$. Consequently, we have

$$\begin{aligned} & h((SX+V)^{n_1}|X^n) = \ln(\pi)^{|n_1|}|\Sigma_y| \\ \geq \; & h((SX)^{n_1}|X^n) = \ln(\pi)^{|n_1|}|\Sigma_{sx}|. \quad (5) \end{aligned}$$

Now the mutual information in (2) can be upper bounded as

$$I_n \leq \frac{1}{n}[h(Y^n) - h((SX)^{n_1}|X^n) - h(V^{n_2})]. \quad (6)$$

Let $(|Y|^2)^n$ represent the sequence of squared output amplitude, and ${\theta_Y}^n$ represent the sequence of the output phase. It can be verified that the differential entropy of the output can be written as

$$h(Y^n) = h((|Y|^2)^n) + h(\theta_Y^n|(|Y|^2)^n) + n\ln 2,$$

where $\ln 2$ is due to the Jacobian of the transformation of $Y^n$ from complex numbers to phase and squared amplitude sequences. Similarly the conditional differential entropy as

$$\begin{aligned} h((SX)^{n_1}|X^n) \;=\; & h((|SX|^2)^{n_1}|X^n) \\ & + h(\theta_{SX}^{n_1}|(|SX|^2)^{n_1}, X^n) + |n_1|\ln 2, \end{aligned}$$

where $(|SX|^2)^{n_1}$ stands for the squared amplitude sequence of the output before adding the observation noise, and $\theta_{SX}^{n_1}$ represents the phase sequence of the output before adding the observation noise. Now (6) can be expressed as

$$\begin{aligned} nI_n \;\leq\; & h((|Y|^2)^n) - h((|SX|^2)^{n_1}|X^n) - h((|V|^2)^{n_2}) \\ & + h(\theta_Y^n|(|Y|^2)^n) - h(\theta_{SX}^{n_1}|(|SX|^2)^{n_1}, X^n) \\ & - h((\theta_V)^{n_2}|(|V|^2)^n). \quad (7) \end{aligned}$$

The upper bound of the averaged mutual information $I_n$ can be analyzed from two parts,

$$nI_{An} = h((|Y|^2)^n) - h((|SX|^2)^{n_1}|X^n) - h((|V|^2)^{n_2}) \quad (8)$$

and

$$\begin{aligned} nI_{\theta n} \;=\; & h(\theta_Y^n|(|Y|^2)^n) - h(\theta_{SX}^{n_1}|(|SX|^2)^{n_1}, X^n) \\ & - h((\theta_V)^{n_2}|(|V|^2)^n), \quad (9) \end{aligned}$$

which correspond to the mutual information between the amplitude and between the phase.

Theorem 1: $nI_{An}$ can be upper bounded as

$$nI_{An} \leq C_{iid} + \ln \sum_{i:X_i \neq 0} \frac{1+|X_i|^2}{|X_i|^2} + n - h((|S|^2)^{n_1}), \quad (10)$$

Proof: From Theorem [9.6.3] in [6], $h((|SX|^2)^{n_1}|X^n)$ can be written as

$$\begin{aligned} & h((|SX|^2)^{n_1}|X^n) \\ = \; & h((|S|^2)^{n_1}) + \sum_{i:i\in n_1} \ln|X_i|^2 \\ = \; & h((|S|^2)^{n_1}) + \sum_{i:i\in n_1} \ln(1+|X_i|^2) - \\ & \sum_{i:i\in n_1} \ln\frac{1+|X_i|^2}{|X_i|^2}. \quad (11) \end{aligned}$$

Now $nI_{An}$ in (8) can be expressed as

$$\begin{aligned}
nI_{An} &= h((|Y|^2)^n) - h((|V|^2)^{n_2}) \\
&\quad - \sum_{i:i\in n_1} \ln(1+|X_i|^2) - |n_1| + |n_1| \\
&\quad -h((|S|^2)^{n_1}) + \sum_{i:i\in n_1} \ln\frac{1+|X_i|^2}{|X_i|^2} \\
&\overset{(a)}{=} h((|Y|^2)^n) - \sum_{i:i\in n} \ln(1+|X_i|^2) - n + |n_1| \\
&\quad -h((|S|^2)^{n_1}) + \sum_{i:i\in n_1} \ln\frac{1+|X_i|^2}{|X_i|^2} \\
&\overset{(b)}{\leq} h_{iid}((|Y|^2)^n) - \sum_{i:i\in n} \ln(1+|X_i|^2) \\
&\quad -n + |n_1| - h((|S|^2)^{n_1}) \\
&\quad + \sum_{i:i\in n_1} \ln\frac{1+|X_i|^2}{|X_i|^2} \\
&\overset{(c)}{\leq} nC_{iid} + |n_1| - h((|S|^2)^{n_1}) \\
&\quad + \sum_{i:i\in n_1} \ln\frac{1+|X_i|^2}{|X_i|^2} \\
&\leq nC_{iid} + n - h((|S|^2)^{n_1}) \\
&\quad + \sum_{i:i\in n_1} \ln\frac{1+|X_i|^2}{|X_i|^2} \qquad (12)
\end{aligned}$$

where (a) follows because

$$h((|V|^2)^{n_2}) = \sum_{i:i\in n_2} \ln(1+|X_i|^2) + |n_2|,$$

merging this term with

$$\sum_{i:i\in n_1} \ln(1+|X_i|^2) + |n_1|,$$

we have the expression in (a). In Step (b), $h_{iid}((|Y|^2)^n)$ stands for the differential entropy of the squared output amplitude given the same i.i.d. input and if the channel is i.i.d. distributed according to the marginal distribution of the Gauss-Markov channel. The proof for the inequality

$$h((|Y|^2)^n) \leq h_{iid}((|Y|^2)^n) \qquad (13)$$

is provided in Appendix A. Consequently maximizing

$$h_{iid}((|Y|^2)^n) - \sum_{i:i\in n} \ln(1+|X_i|^2) - n$$

is equivalent to maximizing the mutual information of an independent Rayleigh flat fading channel with the conditional output differential entropy

$$h((|Y|^2)^n|X^n) = \sum_{i:i\in n} \ln(1+|X_i|^2) + n.$$

The result is the Rayleigh fading channel capacity $nC_{iid}$. In fact, as $n \to \infty$, it can be proved (the proof is trivial and is not shown here.) that $\frac{1}{n}\sum_{i:i\in n_1} \ln\frac{1+|X_i|^2}{|X_i|^2} \to 0$ and the bound on $I_{An}$ can be written as $I_{An} \leq C_{iid} + 1 - h((|S|^2))$. Note that these results are proven for the case when the channel input is i.i.d., which is a reasonable assumption. The inequality in (13) should also hold in correlated input cases, and our result could be generalized. This is an area that we will look into in the future.

Next, we focus on the analysis of $I_{\theta n}$ in (9),

$$\begin{aligned}
nI_{\theta n} &= h(\theta_Y^n|(|Y|^2)^n) - h(\theta_{SX}^{n_1}|(|SX|^2)^{n_1}, X^n) \\
&\quad -h((\theta_V)^{n_2}|(|V|^2)^n) \\
&\overset{(a)}{\leq} h(\theta_Y^n) - h(\theta_{SX}^{n_1}|(|SX|^2)^{n_1}, X^n) \\
&\quad -h((\theta_V)^{n_2}), \qquad (14)
\end{aligned}$$

where (a) follows because condition reduces entropy, and the for the i.i.d. complex Gaussian noise, the phase is independent of the amplitude. The term $h(\theta_{SX}^{n_1}|(|SX|^2)^{n_1}, X^n)$ can be further bounded,

$$\begin{aligned}
& h(\theta_{SX}^{n_1}|(|SX|^2)^{n_1}, X^n) \\
&= h((\theta_X+\theta_S)^{n_1}|(|SX|^2)^{n_1}, X^n) \\
&\overset{(a)}{=} h(\theta_S^{n_1}|(|SX|^2)^{n_1}, X^n) \\
&\overset{(b)}{\geq} h(\theta_S^{n_1}||S|^{n_1}, (|SX|^2)^{n_1}, X^n) \\
&\overset{(c)}{=} h(\theta_S^{n_1}||S|^{n_1}), \qquad (15)
\end{aligned}$$

where (a) follows because translation does not change entropy, (b) follows because condition reduces entropy, and (c) follows because given $|S|^{n_1}$, $\theta_S^{n_1}$ becomes independent from $((|SX|^2)^{n_1}, X^n)$.

Using (15), we can further upper bound $nI_{\theta n}$ in (14) as

$$\begin{aligned}
nI_{\theta n} &\leq h(\theta_Y^n) - h(\theta_S^{n_1}|(|S|)^{n_1}) - h((\theta_V)^{n_2}) \\
&\overset{(a)}{\leq} |n_1|\ln(2\pi) - h(\theta_S^{n_1}|(|S|)^{n_1}) \\
&\leq n\ln(2\pi) - h(\theta_S^{n_1}|(|S|)^{n_1}), \qquad (16)
\end{aligned}$$

where (a) follows because $h(\theta_Y^n) \leq n\ln(2\pi)$, and

$$h((\theta_V)^{n_2}) = n_2\ln(2\pi).$$

This shows that the amount of information that can be carried by phase is upper bounded by a constant that does not change with SNR. As long as there is uncertainty about the fading channel, $h(\theta_S^{n_1}|(|S|)^{n_1})$ will be limited, and $I_{\theta n}$ will be limited. If the channel model is coherent or quasi-static, then $h(\theta_S^{n_1}|(|S|)^{n_1})$ would approaches to $\infty$. Modulation schemes designed for these channels have to explore the unlimited capacity of the phase in order to be considered optimal. However, since in a correlated fading channel, the phase only provides limited capacity, it is safe to speculate that optimal designs based on the coherent or quasi-static fading assumptions can not be optimal.

A combined upper bound becomes

$$\begin{aligned} I_n &\leq C_{iid} + \frac{|n_1|}{n}\ln 2\pi + 1 - \frac{1}{n}h(S^{n_1}) - \frac{|n_1|}{n}\ln 2 \\ &= C_{iid} + \frac{|n_1|}{n}\ln \pi + 1 - \frac{1}{n}h(S^{n_1}). \end{aligned} \tag{17}$$

Note that this result is similar to the result in [2] where the upper bound is given as $I_n \leq C_{iid} + \ln\lambda_{\min}$, where $\lambda_{\min}$ stands for the minimum eigenvalue of the channel correlation matrix $\Sigma_{Sn}$. We can prove that our bound is tighter when the channel is Gauss-Markov. Assume that the determinate of the channel covariance matrix is $K = |\Sigma_{Sn}|$, then our bound can be reduced to the form $I_n \leq C_{iid} - \frac{1}{n}\ln K$. Since $\lambda_{\min} \leq K^{\frac{1}{n}}$, this proves that our bound is tighter.

From the upper bound, we can also infer that the correlation of the channel can provide significant capacity gain over $C_{iid}$. For example, at $50dB$, we have $C_{iid}(SNR) \leq 1.5$ bits according to the optimized upper bound on $C_{iid}$ in [1]. However, at $50dB$, where our bound in (17) is expected to be quite tight, the channel correlation can provide upto $2.3279$ additional bits in capacity when assuming the first order Gauss-Markov model with a correlation coefficient of $a = 0.95$. When $a = 0.95$, the Gauss-Markov model approximates a fast fading channel with a normalized Doppler frequency shift of approximately $f_dT_s = 0.05$. Even with fading this fast, it is still possible to double the transmission rate if we utilizes the channel correlation.

It is easy to show that the asymptotic growth rate of the upper bound in (17) of the Gauss-Markov channel capacity is the growth rate of $C_{iid}$. A natural lower bound on the Gauss-Markov channel capacity is $C_{iid}$, because by using interleaving techniques, a correlated Gauss-Markov channel can be turned into an approximate i.i.d. Rayleigh fading channel. Therefore, the growth rate of the lower bound is $\frac{\partial C_{iid}}{\partial P}$. Combining the analysis on the growth rate of the upper and lower bound, we conclude that the asymptotic growth rate of the channel capacity on a Gauss-Markov flat fading channel must be $\frac{\partial C_{iid}}{\partial P}$ which is proportional to $\log\log(SNR)$. This indicates that the asymptotic region is very power inefficient and the region should be avoided is possible.

## 4. APPENDIX A

Proof of (13), $h((|Y|^2)^n) \leq h_{iid}((|Y|^2)^n)$.
Let $p(|Y|^2)$ represent the marginal distribution of the squared amplitude of the output of a Gauss-Markov channel. Let $p(S)$ represent the marginal distribution of a Gauss-Markov channel, it is a common knowledge that $p(S)$ is complex Gaussian with zero mean and unit variance. Hence, $p(S) = p_{iid}(S)$, where $p_{iid}$ represent the channel state distribution of an independent Rayleigh fading channel with zero mean and unit variance. As a reference, please see the discussion in section [13.3] in [7]. Suppose the input is i.i.d, we have

$$\begin{aligned} p(|Y|^2) &= \int p(|Y|^2|S,X)p(S)p(X)dsdx \\ &= \int p_{iid}(|Y|^2|S,X)p_{iid}(S)p(X)dsdx \\ &= p_{iid}(|Y|^2), \end{aligned} \tag{18}$$

where $p(|Y|^2|S,X)$ and $p_{iid}(|Y|^2|S,X)$ represent the output conditional distributions given the channel state and the input on the Gauss-Markov and the independent Rayleigh channels. Since these conditional distribution is determined by the nature of the observation noise only if given the same channel state and input, we must have $p_{iid}(|Y|^2|S,X) = p(|Y|^2|S,X)$. Hence, we have

$$h((|Y|^2)^n) \overset{(a)}{\leq} nh(|Y|^2) = h_{iid}((|Y|^2)^n), \tag{19}$$

where (a) follows due to the independence bound on entropy.

## 5. REFERENCES


[1] A. Lapidoth and S. Moser, "Capacity bounds via duality with applications to multiple-antenna system on flat-fading channels," *IEEE Trans. Inform. Theory,* vol. 49, pp. 2426,2467, Oct. 2003.

[2] Y. Liang, and V. V. Veeravalli, "Capacity of Noncoherent Time-Selective Rayleigh-Fading Channels," *IEEE Trans. Inform. Theory*, vol. 50, pp. 3095-3110, Dec., 2004.

[3] G. Taricco and M. Elia, "Capacity of fading channels with no side information," *Electron. Lett.,*vol. 33, pp. 1368-1370, July 31, 1997.

[4] I. Abou-Faycal, M. Trott, and S. Shamai (Shitz), "The capacity of discrete time Rayleigh fading channels," *IEEE Trans. Inform. Theory,* vol. 47, pp. 1290-1301, May 2001.

[5] R. Chen, B. Hajek, R. Koetter and U. Madhow, "On fixed input distributions for noncoherent communication over high SNR Rayleigh fading channels" , to appear *IEEE Transactions on Information Theory*, Dec. 2004.

[6] T. M. Cover, and J. A. Thomas, *Elements of Information Theory.* Wiley, New York, 1991.

[7] Steven M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Prentice Hall PTR, New Jersey.