An Induced Additive-Noise Model for Non-Coherent Discrete-Time Memoryless Rayleigh Fading Channels

Wenyi Zhang and J. Nicholas Laneman[1]
Department of Electrical Engineering
University of Notre Dame
Notre Dame, Indiana 46556
e-mail: {wzhang1, jnl}@nd.edu

Abstract — **This paper investigates a non-coherent discrete-time memoryless Rayleigh fading channel. A logarithmic transform converts it into an induced channel with additive noise that is independent of the channel input. From this perspective, it is natural and convenient for us to revisit many known results and gain several new insights. In particular, we specify a class of simple log-scale uniform channel input distributions that performs well for moderate to high signal-to-noise ratio (SNR). Furthermore, for this class of input distributions, its limiting continuous distribution is asymptotically capacity-achieving in the sense that the gap between the resulting mutual information and the channel capacity approaches zero as SNR becomes large. The induced channel modeling approach can also be applied to analyzing other performance metrics such as the cutoff rate and the error exponent, as well as be extended to certain multiple-antenna channels and multiple access channels.**

I. Introduction

In contrast to the thorough understanding of the additive white Gaussian noise (AWGN) channel, much less is understood about the non-coherent discrete-time memoryless Rayleigh fading channel. The simplest case is that for which the channel is scalar, *i.e.*,

$$\mathbf{x} = \mathbf{s} \cdot \mathbf{h} + \mathbf{z}, \tag{1}$$

where $\mathbf{s} \in \mathcal{C}$ is the channel input, and $\mathbf{x} \in \mathcal{C}$ is the channel output. The fading coefficient $\mathbf{h} \in \mathcal{C}$ and the additive noise $\mathbf{z} \in \mathcal{C}$ are both zero-mean circular complex Gaussian random variables, and independent for different channel uses. Hence in (1) we suppress all of the time indexes. For sake of simplicity and without loss of generality, we let $\mathbf{h}$ and $\mathbf{z}$ both be of unit variance, and $\mathbf{s}$ be average-power-limited so that

$$\mathcal{E}\{|\mathbf{s}|^2\} = P. \tag{2}$$

[1]This work has been supported in part by the State of Indiana through the 21st Century Research and Technology Fund, and by the National Science Foundation through contract ECS03-29766.

Consequently, the average signal-to-noise ratio (SNR) of this channel becomes

$$\rho = \frac{\mathcal{E}\{|\mathbf{h}|^2\} \cdot \mathcal{E}\{|\mathbf{s}|^2\}}{\mathcal{E}\{|\mathbf{z}|^2\}} = \mathcal{E}\{|\mathbf{s}|^2\} = P. \tag{3}$$

Throughout the paper we consider the case that the realization of $\mathbf{h}$, the fading coefficient, is available to neither the transmitter nor the receiver.

Neither the capacity nor the capacity-achieving input distribution is fully known. The authors of [1] prove that the capacity of channel (1) is achieved by a discrete input distribution with a finite number of mass points, including a mass point at $\mathbf{s} = 0$. However, they do not provide any specific signaling design, partly because the proof technique employed is non-constructive in nature. Furthermore, the capacity-achieving input distribution changes for different SNR. The number of mass points, their locations, and their probabilities, all depend on the operating SNR, and can only be determined through numerical nonlinear optimization. In [2] the authors obtain upper and lower bounds on the channel capacity, constructing an explicit discrete input distribution to give the lower bound. They show that, at high SNR, the channel capacity grows double-logarithmically, that is, $C = \mathcal{O}(\log \log \rho)$ as $\rho \to \infty$. This double-logarithmic growth rate is refined in [3] [4] for more general models such as Ricean fading, non-memoryless fading processes, and multiple-antenna channels. It is shown that for very general channel models, the channel capacity always behaves like $C = \log \log \rho + \chi + o(1)$ as $\rho \to \infty$, where χ is called the fading number varying for different channel configurations. Another relevant work is [5], in which the authors derive necessary and sufficient conditions for continuous input distributions that lead to unbounded mutual information as SNR goes to infinity.

In addition to the high-SNR analyses mentioned above, for sufficiently low SNR it is known that the channel capacity is achieved by using on-off keying (OOK), for which the non-zero input has amplitude that becomes unbounded as SNR approaches zero [1] [6]. Numerical results [1] indicate that, for SNR less than roughly −3.5 dB, optimized OOK is capacity-achieving, and for SNR less than 10 dB, there is at most 4% gap to capacity from the mutual information achieved by optimized OOK.

The main idea of this paper stems from the following observation: As SNR approaches infinity, most of the channel randomness results from the fading coefficient $\mathbf{h}$ rather than the additive noise $\mathbf{z}$. Intuitively we may approximate the limiting channel behavior at high SNR as $\log|\mathbf{x}| = \log|\mathbf{s}| + \log|\mathbf{h}|$, *i.e.*, a channel with additive noise only. This limiting approximation is validated and employed as a proof technique in [5]. In this paper, we further observe that such an additive-noise channel interpretation holds for *all* SNR. Specifically, by taking the logarithm of the channel output's magnitude $|\mathbf{x}|$, we can decompose this transformed channel output into the sum of two terms. One term is merely a deterministic function of the channel input $\mathbf{s}$, and the other term is independent of the channel input and can therefore be viewed as an additive noise. This equivalent additive-noise channel model is conceptually more intuitive to communication engineers, and yields a natural geometric interpretation for many earlier results. The effects of both the multiplicative fading and the additive noise in the original channel (1) are now described by a single induced additive noise. The inefficiency of communication over the original fading channel (1) can be intuitively understood as the log-scale transform rescales the original SNR ρ to roughly $\log\rho$. Consequently, the double-logarithmic capacity growth behavior immediately follows.

We briefly summarize the main results of the paper as follows.

1) In Section II we derive the induced additive-noise channel model from the original fading channel model (1). The two channel models are equivalent, thus the problem is converted to communication over an additive-noise channel with certain non-Gaussian noise.

2) Based upon the induced additive-noise channel perspective, in Section III we establish new necessary and sufficient conditions for channel input distributions to achieve unbounded mutual information as SNR $\rho \to \infty$, and to asymptotically achieve the channel capacity. These conditions are qualitatively similar to those established in [5]. However, we feel that our treatment using the induced channel model involves much less mathematical machinery, and appears more straightforward.

3) In Section III we further show that a simple input distribution called the log-scale (continuous) uniform distribution is asymptotically capacity-achieving. To provide practical signaling design, we re-interpret a class of discrete input distributions, which, in the induced channel perspective, turn out to be nothing but pulse amplitude modulation (PAM) with uniformly spaced pulses. For moderate SNR we demonstrate that these discrete input constellations with appropriate size achieve mutual information higher than that achieved by the log-scale (continuous) uniform input.

Some extensions of this work are included in [7]. Specifically, we study the issue of maximum-likelihood (ML) decoding and derive commonly used lower bounds on the cutoff rate and the error exponent, again based upon the induced additive-noise channel perspective. These results may be useful in the design of practical communication systems. Furthermore, as initial attempts of generalization, we extend the induced channel modeling approach to certain multiple-antenna channels and multiple-access channels (MAC).

We use bold font to represent random variables, for which capital letters distinguish random matrices from random scalars and vectors. For a matrix A, A^{T} ($A^{\dagger}$) represents its transpose (conjugate transpose). $\mathcal{N}(m, \Sigma)$ ($\mathcal{CN}(m, \Sigma)$) represents Gaussian (circular complex Gaussian) distributions with mean vector m and covariance matrix Σ. We frequently use Euler's constant $\gamma = 0.5772\ldots$ Throughout the paper, all logarithms are taken with natural base e.

II. Log-Scale Transform: from Fading Channel to Additive-Noise Channel

For the scalar channel (1), because the fading coefficient $\mathbf{h}$ totally distorts the phase information contained in the channel input $\mathbf{s}$, we can simply let $\mathbf{s}$ be real and nonnegative. The channel output $\mathbf{x}$ has a sufficient statistic $|\mathbf{x}|^2 = \mathbf{x}\mathbf{x}^{\dagger}$, *i.e.*, its energy, or equivalently, $|\mathbf{x}|$, *i.e.*, its magnitude. Conditioned on $\mathbf{s} = s$, we have $\mathbf{x} \sim \mathcal{CN}(0, s^2+1)$, and its energy $\mathbf{r} = |\mathbf{x}|^2$ is exponentially distributed with probability density function (PDF)

$$f_{\mathbf{r}|\mathbf{s}}(r|s) = \begin{cases} \frac{1}{s^2+1} e^{-\frac{r}{s^2+1}} & \text{if } r \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

Taking the logarithm of $\sqrt{\mathbf{r}} = |\mathbf{x}|$, and denoting the resulting random variable by $\mathbf{t} = \frac{1}{2}\log\mathbf{r}$, we have

$$\begin{aligned} f_{\mathbf{t}|\mathbf{s}}(t|s) &= \frac{2e^{2t}}{s^2+1} \exp\left\{-\frac{e^{2t}}{s^2+1}\right\} & (5) \\ &= 2 \cdot e^{2\left(t - \frac{1}{2}\log(s^2+1)\right)} \cdot \exp\{-e^{2\left(t-\frac{1}{2}\log(s^2+1)\right)}\} & (6) \end{aligned}$$

for $t \in (-\infty, \infty)$.

From (5), we observe a convenient shift-invariance property of $f_{\mathbf{t}|\mathbf{s}}(t|s)$, *i.e.*,

$$f_{\mathbf{t}|\mathbf{s}}\left(t + \frac{1}{2}\log(s_1^2+1) \mid s_1\right) = f_{\mathbf{t}|\mathbf{s}}\left(t + \frac{1}{2}\log(s_2^2+1) \mid s_2\right) \quad (7)$$

for all $t \in (-\infty, \infty)$ and any $s_1, s_2 \in [0, \infty)$.

The shift-invariance property (7) immediately suggests an additive-noise perspective for the original fading channel (1) after the log-scale transform. If we introduce a new random variable

$$\mathbf{u} \stackrel{\text{def}}{=} \frac{1}{2}\log(\mathbf{s}^2+1), \quad (8)$$

then, via simple manipulations, we establish an induced additive-noise channel equivalent to the original fading channel (1) as

$$\mathbf{t} = \mathbf{u} + \mathbf{w}. \quad (9)$$

The additive noise $\mathbf{w}$ is independent of $\mathbf{u}$, and has PDF

$$f_{\mathbf{w}}(w) = 2e^{2w} \exp\{-e^{2w}\} \tag{10}$$

for $w \in (-\infty, \infty)$. From the average power constraint (2) for $\mathbf{s}$, the equivalent channel input $\mathbf{u}$ should satisfy

$$\mathcal{E}\{e^{2\mathbf{u}}\} = \rho + 1, \tag{11}$$

and has a support set $\mathbf{u} \in [0, \infty)$.

Remark: The induced channel (9) is equivalent to the normalized channel (1), in which the fading coefficient $\mathbf{h}$ and the noise $\mathbf{z}$ are both $\mathcal{CN}(0,1)$ random variables. For general parametrization, *i.e.*, $\mathbf{h} \sim \mathcal{CN}(0, \sigma_{\mathbf{h}}^2)$ and $\mathbf{z} \sim \mathcal{CN}(0, \sigma_{\mathbf{z}}^2)$, the induced channel (9) still holds, except that the induced channel input $\mathbf{u}$ becomes

$$\mathbf{u} \stackrel{\text{def}}{=} \frac{1}{2} \log(\sigma_{\mathbf{h}}^2 \mathbf{s}^2 + \sigma_{\mathbf{z}}^2), \tag{12}$$

with constraint

$$\mathcal{E}\{e^{2\mathbf{u}}\} = \sigma_{\mathbf{h}}^2 \cdot P + \sigma_{\mathbf{z}}^2 \tag{13}$$

and support set $\mathbf{u} \in [\frac{1}{2} \log \sigma_{\mathbf{z}}^2, \infty)$. We note that the definitions of $\mathbf{t}$ and $\mathbf{w}$ do not change under different parametrization.

Figure 1 plots the PDF of the additive noise $\mathbf{w}$. We observe that, the noise density is not symmetric, and it decreases much slower on the left side than on the right side. These qualitative observations are more precisely quantified by the following lemma.

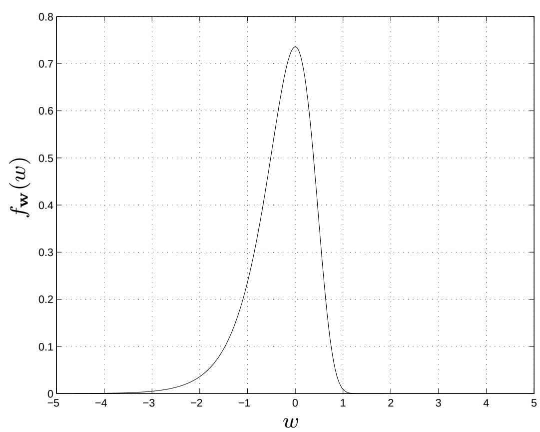

Figure 1: The probability density function (PDF) of the induced additive noise $\mathbf{w}$

Lemma II.1 *(Properties of the induced additive noise $\mathbf{w}$, the scalar channel case)*

1. *The cumulative distribution function (CDF) of $\mathbf{w}$ is*

$$F_{\mathbf{w}}(w) = 1 - \exp\{-e^{2w}\}. \tag{14}$$

2. *The mean of $\mathbf{w}$ is*

$$\mathcal{E}\{\mathbf{w}\} = -\frac{\gamma}{2}, \tag{15}$$

and the maximum of $f_{\mathbf{w}}(w)$ is attained at $w = 0$.

3. *The variance of $\mathbf{w}$ is*

$$\sigma_{\mathbf{w}}^2 = \frac{\pi^2}{24}. \tag{16}$$

4. *The differential entropy of $\mathbf{w}$ is*

$$\begin{aligned} h(\mathbf{w}) &= -\int f_{\mathbf{w}}(w) \log f_{\mathbf{w}}(w) dw \\ &= 1 - \log 2 + \gamma. \end{aligned} \tag{17}$$

III. Channel Mutual Information

The induced additive-noise channel perspective (9) is a convenient tool for studying the channel behavior. In this section we investigate the channel mutual information, focusing on the high SNR regime.

For non-coherent channels, an asymptotically tight characterization of the channel capacity is given by [3]

$$\lim_{\rho \to \infty} (C - \log \log \rho) = \chi, \tag{18}$$

where χ is called the *fading number*, and $\chi = -1 - \gamma$ for the scalar memoryless Rayleigh case (1).

Using the equivalent induced additive-noise channel (9), we can write the channel mutual information as

$$\begin{aligned} \mathcal{I}(\mathbf{s}; \mathbf{x}) &= \mathcal{I}(\mathbf{u}; \mathbf{t}) \\ &= h(\mathbf{t}) - h(\mathbf{t}|\mathbf{u}) = h(\mathbf{t}) - h(\mathbf{w}). \end{aligned} \tag{19}$$

From Lemma II.1 $h(\mathbf{w}) = 1 - \log 2 + \gamma$, hence the channel capacity corresponds to maximizing $h(\mathbf{t})$, differential entropy of the channel output $\mathbf{t}$.

A. A Capacity Upper Bound Revisited

The problem of maximizing $h(\mathbf{t})$ is non-trivial, because one of the constraints is implicitly given by $\mathbf{t} = \mathbf{u} + \mathbf{w}$, where the support set of $\mathbf{u}$ is $[0, \infty)$. If we relax the above constraint to

$$\mathcal{E}\{\mathbf{t}\} = \mathcal{E}\{\mathbf{u}\} + \mathcal{E}\{\mathbf{w}\} = \lambda \geq -\frac{\gamma}{2}, \tag{20}$$

and incorporating the input constraint (11) to have that

$$\mathcal{E}\{e^{2\mathbf{t}}\} = \mathcal{E}\{e^{2\mathbf{u}}\}\mathcal{E}\{e^{2\mathbf{w}}\} = \rho + 1, \tag{21}$$

then, following the standard maximum entropy analysis [8], the resulting entropy-maximizing distribution of $\mathbf{t}$ has pdf

$$f_{\mathbf{t}}^*(t) = 2\left(\frac{\mu}{\rho+1}\right)^{\mu} \frac{1}{\Gamma(\mu)} e^{2\mu t} \exp\left\{-\frac{\mu}{\rho+1} e^{2t}\right\}, \tag{22}$$

where μ is determined by the relation

$$\psi(\mu) - \log\mu = 2\lambda - \log(\rho+1), \tag{23}$$

and $\psi(x) \stackrel{\text{def}}{=} \frac{d}{dx}\log\Gamma(x)$ is the psi function [9]. Thus, we can obtain an upper bound on the channel capacity by letting $\lambda = -\frac{\gamma}{2}$,

$$\begin{aligned} C = \sup_{f_{\mathbf{t}}(t)} \mathcal{I}(\mathbf{u};\mathbf{t}) &< h(f_{\mathbf{t}}^*(t))|_{\lambda=-\frac{\gamma}{2}} - h(\mathbf{w}) \\ &= \log\mu - \mu\psi(\mu) + \mu - \gamma - 1, \end{aligned} \tag{24}$$

where

$$\log\mu - \psi(\mu) = \log(\rho+1). \tag{25}$$

This is exactly the same upper bound obtained in [2] using variational methods. Here we derive it from the induced additive-noise channel. We note that a tighter upper bound on the channel capacity is given in [3] using duality methods.

B. Conditions for Good Continuous Input Distributions and Log-Scale Uniform Inputs

In addition to the aforementioned capacity upper bounds, we have the following upper/lower bounds for any given input distribution:

$$\mathcal{I}(\mathbf{u};\mathbf{t}) \leq \frac{1}{2}\log\sigma_{\mathbf{u}}^2 + \log\sqrt{2\pi e(1+\frac{\sigma_{\mathbf{w}}^2}{\sigma_{\mathbf{u}}^2})} - h(\mathbf{w}) \tag{26}$$

$$\mathcal{I}(\mathbf{u};\mathbf{t}) \geq h(\mathbf{u}) + \log\sqrt{1+e^{2[h(\mathbf{w})-h(\mathbf{u})]}} - h(\mathbf{w}) \tag{27}$$

The upper bound (26) is a straightforward application of the property that Gaussian distribution maximizes the entropy over all distributions with the same variance [8, Theorem 9.6.5], and the lower bound (27) is obtained using the entropy power inequality [8, Theorem 16.7.1]. From (27) we observe that if $h(\mathbf{u}) \to \infty$ as $\rho \to \infty$ then $\mathcal{I}(\mathbf{u};\mathbf{t}) \to \infty$; on the other hand, from (26) we observe that if $\mathcal{I}(\mathbf{u};\mathbf{t}) \to \infty$ as $\rho \to \infty$ then $\sigma_{\mathbf{u}}^2 \to \infty$. We thus obtain necessary and sufficient conditions for the unboundedness of channel mutual information as SNR becomes unbounded.

Proposition III.1 *The channel input* $\mathbf{u}$ *achieves* $\mathcal{I}(\mathbf{u};\mathbf{t}) \to \infty$ *as* $\rho \to \infty$ *if*

$$h(\mathbf{u}) \to \infty \text{ as } \rho \to \infty; \tag{28}$$

on the other hand, if $\mathbf{u}$ *achieves* $\mathcal{I}(\mathbf{u};\mathbf{t}) \to \infty$ *as* $\rho \to \infty$*, then*

$$\sigma_{\mathbf{u}}^2 \to \infty \text{ as } \rho \to \infty. \tag{29}$$

Neither of the above sufficient and necessary conditions can be simultaneously necessary and sufficient. Two simple counterexamples using discrete channel inputs are given as follows. On one hand, since the channel capacity is always achieved by discrete inputs whose differential entropy is $-\infty$, $h(\mathbf{u}) \to \infty$ is not necessary. On the other hand, a symmetric OOK input leads to unbounded variance as $\rho \to \infty$, but the resulting channel mutual information is always upper bounded by $\log 2$ nats.

We compare the sufficient condition in Proposition III.1 and those obtained in [5, Theorem 4.3]. In [5] it is established that, if $h(\log|\mathbf{s}|) \to \infty$ as $\rho \to \infty$, then $\mathcal{I}(\mathbf{s};\mathbf{x}) = \mathcal{I}(\mathbf{u};\mathbf{t}) \to \infty$. The two sufficient conditions are obviously similar, except for different transforms of the channel input.

Proposition III.1 immediately explains why circular complex Gaussian input $\mathbf{s}$ performs poorly for channel (1). In fact we have

Corollary III.2 *If* $\mathbf{s} \sim \mathcal{CN}(0,\rho)$*, then its resulting equivalent input* $\mathbf{u}$ *has*

$$h(\mathbf{u}) \to h(\mathbf{w}) = 1 - \log 2 + \gamma \tag{30}$$

$$\sigma_{\mathbf{u}}^2 \to \sigma_{\mathbf{w}}^2 = \frac{\pi^2}{24}, \tag{31}$$

as $\rho \to \infty$.

Comparing the mutual information lower bound (27) and the capacity asymptote (18), we further obtain a sufficient condition for input distributions that asymptotically achieve capacity.

Proposition III.3 *A sufficient condition for*

$$\lim_{\rho\to\infty}(C - \mathcal{I}(\mathbf{u};\mathbf{t})) = 0 \tag{32}$$

is

$$\lim_{\rho\to\infty}(h(\mathbf{u}) - \log\log\rho) \geq -\log 2. \tag{33}$$

Now we introduce the *log-scale (continuous) uniform* input distribution as follows.

Definition III.4 *A log-scale (continuous) uniform input distribution for the induced additive-noise channel (9) has PDF*

$$f_{\mathbf{u}}(u) = \begin{cases} \frac{1}{A} & \text{if } 0 \leq u \leq A \\ 0 & \text{otherwise} \end{cases} \tag{34}$$

where $A > 0$ *is determined by*

$$\frac{e^{2A}-1}{2A} - 1 = \rho. \tag{35}$$

For the log-scale uniform input distribution, the channel output $\mathbf{t}$ has PDF

$$f_{\mathbf{t}}(t) = \int f_{\mathbf{u}}(u) f_{\mathbf{w}}(t-u)du = \frac{(F_{\mathbf{w}}(t) - F_{\mathbf{w}}(t-A))}{A}. \tag{36}$$

We can then numerically evaluate the corresponding $\mathcal{I}(\mathbf{u};\mathbf{t})$. Furthermore, the log-scale (continuous) uniform input distribution is asymptotically capacity-achieving.

Corollary III.5 *The log-scale (continuous) uniform input achieves*

$$\lim_{\rho\to\infty} \left(C - \mathcal{I}(\mathbf{u};\mathbf{t})\right) = 0. \tag{37}$$

We compare the log-scale (continuous) uniform distribution with that used in [3], which lets $\log \mathbf{s}$ take a uniform distribution over a certain interval with width roughly $\log\rho$. Both of them are asymptotically capacity-achieving. This is intuitively reasonable since, for large $\mathbf{s}$, $\mathbf{u} = \log\sqrt{\mathbf{s}^2+1}$ and $\log \mathbf{s}$ are essentially identical.

For the log-scale (continuous) uniform input in the high SNR limit, the lower bound (27) attains the channel capacity, and the upper bound (26) cannot be met, with a gap of $\log\sqrt{\pi e/6} \approx 0.1765$ nats. For finite SNR, both these bounds may be useful in characterizing the achievable rate. In Figure 2, we plot the lower/upper bounds and actual mutual information for log-scale (continuous) uniform input. For comparison, we also plot the numerically computed channel capacity at low SNR [1], the capacity upper bound given by (24) [2], the capacity upper bound given in [3] using duality methods, and the asymptotic fading number limit (18) [3].

We make several observations based on Figure 2. First, the actual mutual information is close to its lower bound for low and high SNR, while its upper bound is a tight estimate for SNR between 10 and 30 dB. Second, at high SNR the gap between the asymptotic fading number limit (18) and the mutual information lower bound (27) is approximately 0.2 nats, even at $\rho \approx 80$ dB. This indicates that the asymptotic fading number limit may under-estimate the actual channel capacity for a rather wide range of SNR. Indeed these two curves remains separated by about 0.05 nats at $\rho \approx 400$ dB. Finally, at low SNR, there exists a significant rate loss by using the log-scale uniform input rather than optimized OOK, which is capacity-achieving in that regime. In fact, in the low SNR limit, the log-scale uniform distribution yields $A = \mathcal{O}(\rho)$, so the fourthegy [10] of the input is vanishing with SNR. As a consequence, the channel mutual information is quadratically vanishing with SNR.

C. Log-Scale Discrete Uniform Inputs

For practical purposes we are interested in discrete input distributions. Motivated by the log-scale (continuous) uniform distribution, and the idea of maximizing minimum distance in the induced additive-noise channel, we introduce the following *$\mathcal{M}$-ary log-scale discrete uniform* input distribution.

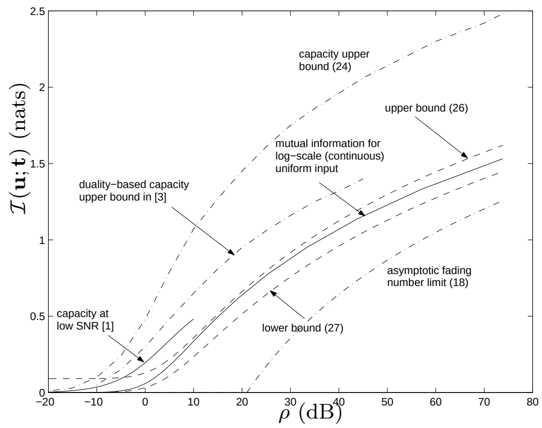

Figure 2: Comparison of different channel mutual information bounds vs. channel SNR

Definition III.6 *A $\mathcal{M}$-ary log-scale discrete uniform distribution for the induced additive-noise channel (9) has probability mass function (PMF)*

$$p_m \stackrel{\text{def}}{=} \text{Prob}\left(\mathbf{u} = m\Delta\right) = \frac{1}{\mathcal{M}} \tag{38}$$

for $m = 0, \ldots, \mathcal{M}-1$. The spacing Δ is determined by SNR through the relation

$$\frac{e^{2\mathcal{M}\Delta}-1}{e^{2\Delta}-1} = \mathcal{M}(\rho+1). \tag{39}$$

We note that this type of input is precisely that proposed in [11, Theorem 3]. In [11] it is shown that this type of input maximizes the Kullback-Leibler (KL) distance between the output PDFs conditioned on different inputs. From the induced additive-noise channel perspective, this type of input distributions is nothing but log-scale PAM with uniformly spaced pulses. For $\mathcal{M} = 2$ the log-scale discrete uniform input reduces to symmetric OOK, and for $\mathcal{M}\to\infty$ it approximates the continuous log-scale uniform input.

We numerically compute $\mathcal{I}(\mathbf{u};\mathbf{t})$ for log-scale discrete uniform inputs with different $\mathcal{M}$ and plot the results in Figure 3. We observe that, at low SNR, binary ($\mathcal{M} = 2$) inputs outperform higher-dimensional inputs. As SNR increases beyond roughly 12.5 dB, higher-dimensional inputs with $\mathcal{M} = 3, 4, 5, \ldots$ successively dominate. Although the log-scale (continuous) uniform input (*i.e.*, $\mathcal{M}\to\infty$ limit) is asymptotically capacity-achieving, for moderate SNR it is outperformed by log-scale discrete uniform inputs. At low SNR, we also plot the channel capacity, achieved by optimized OOK. We observe that the rate gain by using optimized OOK instead of symmetric OOK can be significant. Finally we note that, the capacity lower bound obtained in [2] approximately corresponds to the envelope of all the curves in Figure 3.

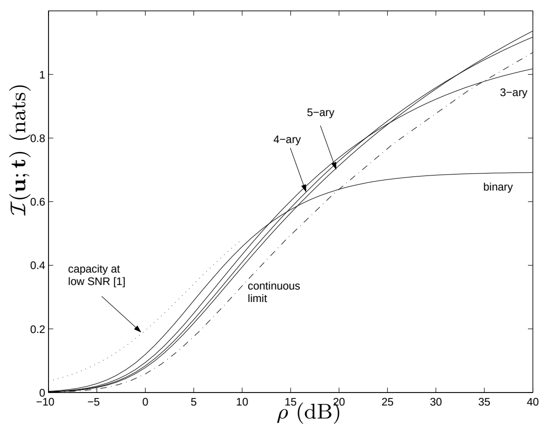

Figure 3: Mutual information vs. SNR for $\mathcal{M}$-ary log-scale discrete uniform input ($\mathcal{M} = 2, 3, 4, 5, \infty$) and optimized OOK

IV. Conclusion

In this paper we address non-coherent discrete-time memoryless Rayleigh fading channels. By taking the logarithm of the channel output's magnitude, we transform the fading channel into an equivalent induced channel with additive noise that is independent of the channel input. This additive-noise channel model holds for all SNR. Using this novel perspective, we revisit and establish a series of known and new results. We in particular examine the high SNR channel behavior. We obtain simple conditions for channel input distributions to achieve unbounded mutual information as SNR becomes unbounded, and to asymptotically achieve the channel capacity in the high SNR limit. We show that a simple input distribution called the log-scale (continuous) uniform distribution is asymptotically capacity-achieving. We further re-interpret a class of log-scale discrete uniform input distributions, and demonstrate that for moderate SNR they perform better than the log-scale (continuous) uniform input.

The log-scale transform approach in this paper naturally prompts a further question: "Is it possible to generalize this approach to other channels, say, general non-coherent fading channels?" Such a generalization, if it exists, doubtlessly would enhance our understanding of channel behavior and communication system design. Some extensions to restricted classes of multiple-antenna channels and multiple-access channels are developed in [7]. However, we note that currently this approach seems to be applicable only to memoryless Rayleigh fading channels; extensions to more general scenarios are not immediate. There are primarily two obstacles. First, for other fading distributions, the "non-Rayleighness" precludes the convenient shift-invariance property of the log-scale transform. For example, in Ricean channels or phase non-coherent Gaussian channels, the "additive noise" after the log-scale transform becomes input-dependent. More fundamentally, for channels in which the output sufficient statistic is multivariate rather than scalar, we have not yet found an immediate analog of the log-scale transform.

Acknowledgments

The authors are grateful to Prof. Amos Lapidoth for his thoughtful comments on this work.

References

[1] Ibrahim C. Abou-Faycal, Mitchell D. Trott, and Shlomo Shamai (Shitz), "The Capacity of Discrete-Time Memoryless Rayleigh-Fading Channels," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1290–1301, May 2001.

[2] Giorgio Taricco and M. Elia, "Capacity of Fading Channels with No Side Information," *Electron. Lett.*, vol. 33, no. 16, pp. 1368–1370, July 1997.

[3] Amos Lapidoth and Stefan M. Moser, "Capacity Bounds via Duality with Applications to Multiple-Antenna Systems on Flat Fading Channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2426–2567, Oct. 2003.

[4] Amos Lapidoth, "On the Asymptotic Capacity of Stationary Gaussian Fading Channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 437–446, Feb. 2005.

[5] R.-R. Chen, Bruce Hajek, Ralf Koetter, and U. Madhow, "On Fixed Input Distributions for Noncoherent Communication over High SNR Rayleigh Fading Channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3390–3396, Dec. 2004.

[6] Sergio Verdú, "Spectral Efficiency in the Wideband Regime," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1319–1343, June 2002.

[7] Wenyi Zhang and J. Nicholas Laneman, "An Induced Additive-Noise Model for Non-Coherent Discrete-Time Memoryless Rayleigh Fading Channels," *IEEE Trans. Inform. Theory*, Feb. 2005, Submitted for publication.

[8] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Inc., New York, 1991.

[9] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, Academic Press, fifth edition, 1994.

[10] Vijay Subramanian and Bruce Hajek, "Broad-Band Fading Channels: Signal Burstiness and Capacity," *IEEE Trans. Inform. Theory*, vol. 48, no. 4, pp. 809–827, Apr. 2002.

[11] M. J. Borran, A. Sabharwal, and B. Aazhang, "On Design Criteria and Construction of Noncoherent Space-Time Constellations," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2332–2351, Oct. 2003.