On Multipath Fading Channels at High SNR

Tobias Koch, *Member, IEEE*, and Amos Lapidoth, *Fellow, IEEE*

***Abstract*—A noncoherent multipath fading channel is considered, where neither the transmitter nor the receiver is cognizant of the realization of the path gains, but both are cognizant of their statistics. It is shown that if the delay spread is large in the sense that the variances of the path gains decay exponentially or slower, then capacity is bounded in the signal-to-noise ratio (SNR). For such channels, capacity does not tend to infinity as the SNR tends to infinity. In contrast, if the variances of the path gains decay faster than exponentially, then capacity is unbounded in the SNR. It is further demonstrated that if the number of paths is finite, then at high SNR capacity grows double-logarithmically with the SNR, and the capacity pre-loglog—defined as the limiting ratio of capacity to loglog(SNR) as the SNR tends to infinity—is 1 irrespective of the number of paths. The results demonstrate that at high SNR multipath fading channels with an infinite number of paths cannot be approximated by multipath fading channels with only a finite number of paths. The number of paths that are needed to approximate a multipath fading channel typically depends on the SNR and may grow to infinity as the SNR tends to infinity.**

***Index Terms*—Channel capacity, channels with memory, fading channels, frequency-selective fading, high signal-to-noise ratio, multipath, noncoherent.**

I. Introduction

WE STUDY the capacity of discrete-time *multipath fading channels*. In such channels the transmitted signal propagates along a multitude of paths, and the gains and delays of these paths vary over time. In general, the path delays differ from each other, and the receiver thus observes a weighted sum of delayed replicas of the transmitted signal, where the weights are random. We shall slightly abuse nomenclature and refer to each summand in the received signal as a path, and to the corresponding weight as its path gain, even if it is in fact composed of a multitude of paths. We consider a *noncoherent* channel model, where transmitter and receiver are cognizant of the statistics of the path gains, but are ignorant of their realization.

Multipath fading channels arise in wireless communications, where obstacles in the surroundings reflect the transmitted signal and cause it to propagate along multiple paths, and where relative movements of transmitter, receiver, and obstacles lead to time-variations of the path gains and delays. Examples of wireless communication scenarios where the receiver observes typically more than one path include *radio communications* (particularly if the transmitted signal is of large bandwidth as, for example, in *Ultra-Wideband* or in *CDMA*) and *underwater acoustic communications*.

The capacity of noncoherent multipath fading channels has been investigated extensively in the wideband regime, where the signal-to-noise ratio (SNR) is typically small. It was shown by Kennedy that, in the limit as the available bandwidth tends to infinity, the capacity of the fading channel is the same as the capacity of the additive white Gaussian noise (AWGN) channel of equal received power; see [1, Sec. 8.6] and references therein.

To the best of our knowledge, not much is known about the capacity of noncoherent multipath fading channels at high SNR. For the special case of noncoherent *frequency-flat* fading channels (where we only have *one* path), it was shown by Lapidoth and Moser [2] that if the fading process is of finite entropy rate, then at high SNR capacity grows double-logarithmically in the SNR. This is much slower than the logarithmic growth of the AWGN capacity [3].

In this work, we study the high-SNR behavior of the capacity of noncoherent *multipath* fading channels (where the number of paths is typically greater than one). We demonstrate that the capacity of such channels does not merely grow more slowly with the SNR than the capacity of the AWGN channel, but it may be even *bounded* in the SNR. In other words, for such channels the capacity does not necessarily tend to infinity as the SNR tends to infinity.

We derive a necessary and a sufficient condition for the capacity to be bounded in the SNR. We show that if the variances of the path gains decay *exponentially or slower*, then capacity is bounded in the SNR. In contrast, if the variances of the path gains decay *faster than exponentially*, then capacity is unbounded in the SNR. We further show that if the number of paths is finite, then at high SNR capacity increases double-logarithmically with the SNR, and the capacity pre-loglog—defined as the limiting ratio of the capacity to $\log\log \mathrm{SNR}$ as the SNR tends to infinity—is 1 irrespective of the number of paths.

The rest of this paper is organized as follows. Section II describes the channel model. Section III is devoted to channel capacity. Section IV summarizes our main results. Sections V and VI derive the upper bounds and the lower bounds on channel capacity that are used to prove these results. Section VII concludes the paper with a brief discussion of our results.

II. Channel Model

Let $\mathbb{C}$ and $\mathbb{N}$ denote the set of complex numbers and the set of positive integers, respectively. We consider a discrete-time multipath fading channel whose channel output $Y_k \in \mathbb{C}$ at time

Manuscript received January 05, 2009. Date of current version November 19, 2010. The material in this paper was presented in part at the IEEE Information Theory Workshop (ITW), Porto, Portugal, May 2008, at the IEEE International Symposium on Information Theory (ISIT), Toronto, ON, Canada, July 2008, and at the IEEE 25th Convention of Electrical and Electronics Engineers in Israel, Eilat, Israel, December 2008.

T. Koch was with the Department of Information Technology and Electrical Engineering, ETH Zurich, 8092 Zurich, Switzerland. He is now with the Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, U.K. (e-mail: tobi.koch@eng.cam.ac.uk).

A. Lapidoth is with the Department of Information Technology and Electrical Engineering, ETH Zurich, 8092 Zurich, Switzerland (e-mail: lapidoth@isi.ee.ethz.ch).

Communicated by L. Zheng, Associate Editor for Communications.

Digital Object Identifier 10.1109/TIT.2010.2080995

$k \in \mathbb{N}$ corresponding to the time-1 through time-k channel inputs $x_1, x_2, \ldots, x_k \in \mathbb{C}$ is given by

$$Y_k = \sum_{\ell=0}^{k-1} H_k^{(\ell)} x_{k-\ell} + Z_k, \qquad k \in \mathbb{N}. \tag{1}$$

Here $\{Z_k, k \in \mathbb{Z}\}$ (where $\mathbb{Z}$ denotes the set of integers) models additive noise, and $H_k^{(\ell)}$ denotes the time-k gain of the ℓth path. We assume that $\{Z_k, k \in \mathbb{Z}\}$ is a sequence of independent and identically distributed (i.i.d.), zero-mean, variance-σ^2, circularly-smmmetric, complex Gaussian random variables. For each path $\ell \in \mathbb{N}_0$ (where $\mathbb{N}_0$ denotes the set of nonnegative integers), we assume that $\{H_k^{(\ell)}, k \in \mathbb{Z}\}$ is a zero-mean, complex stationary process. We denote its variance and its differential entropy rate by

$$\alpha_\ell \triangleq \mathsf{E}\left[\left|H_k^{(\ell)}\right|^2\right], \qquad \ell \in \mathbb{N}_0 \tag{2}$$

$$h_\ell \triangleq \lim_{n\to\infty} \frac{1}{n} h\left(H_1^{(\ell)}, H_2^{(\ell)}, \ldots, H_n^{(\ell)}\right), \qquad \ell \in \mathbb{N}_0. \tag{3}$$

We shall say that the channel has a *finite number of paths*, if for some finite integer $\mathsf{L} \in \mathbb{N}_0$

$$H_k^{(\ell)} = 0, \qquad \ell > \mathsf{L}, \quad k \in \mathbb{N}. \tag{4}$$

We assume that $\alpha_0 > 0$. We further assume that

$$\sup_{\ell\in\mathbb{N}_0} \alpha_\ell < \infty \tag{5}$$

$$\inf_{\ell\in\mathcal{L}} h_\ell > -\infty \tag{6}$$

where the set $\mathcal{L}$ is defined as $\mathcal{L} \triangleq \{\ell \in \mathbb{N}_0 : \alpha_\ell > 0\}$. (If the path gains are Gaussian, then (6) is equivalent to the mean-square error in predicting the present path gain from its past being strictly positive, i.e., the present path gain cannot be predicted perfectly from its past.) We finally assume that the processes

$$\{H_k^{(0)}, k \in \mathbb{Z}\}, \{H_k^{(1)}, k \in \mathbb{Z}\}, \ldots$$

are independent ("uncorrelated scattering"); that they are jointly independent of $\{Z_k, k \in \mathbb{Z}\}$; and that the joint law of

$$\left(\{Z_k, k \in \mathbb{Z}\}, \{H_k^{(0)}, k \in \mathbb{Z}\}, \{H_k^{(1)}, k \in \mathbb{Z}\}, \ldots\right)$$

does not depend on the input sequence $\{x_k, k \in \mathbb{Z}\}$. We consider a noncoherent channel model where neither transmitter nor receiver is cognizant of the realization of $\{H_k^{(\ell)}, k \in \mathbb{Z}\}$, $\ell \in \mathbb{N}_0$, but both are aware of their law. We do not assume that the path gains are Gaussian.

III. Channel Capacity

Let A_m^n denote the sequence $A_m, A_{m+1}, \ldots, A_n$. We define the capacity (in nats per channel use) as

$$C(\mathrm{SNR}) \triangleq \varliminf_{n\to\infty} \frac{1}{n} \sup I\left(X_1^n; Y_1^n\right) \tag{7}$$

where the supremum is over all joint distributions on $X_1, X_2, \ldots, X_n$ satisfying the power constraint

$$\frac{1}{n}\sum_{k=1}^{n} \mathsf{E}\left[|X_k|^2\right] \le \mathsf{P} \tag{8}$$

and where SNR is defined as

$$\mathrm{SNR} \triangleq \frac{\mathsf{P}}{\sigma^2}. \tag{9}$$

Here $\varliminf$ denotes the *limit inferior* and $\varlimsup$ the *limit superior*. By Fano's inequality, no rate above $C(\mathrm{SNR})$ is achievable. (See [4] for a definition of an achievable rate.) We do not claim that there is a coding theorem associated with (7), i.e., that $C(\mathrm{SNR})$ is achievable. A coding theorem will hold, for example, if the number of paths is finite, and if the processes corresponding to these paths $\{H_k^{(0)}, k \in \mathbb{Z}\}, \{H_k^{(1)}, k \in \mathbb{Z}\}, \ldots, \{H_k^{(\mathsf{L})}, k \in \mathbb{Z}\}$ are jointly ergodic (see [5, Th. 2]).

The special case of noncoherent frequency-flat fading channels (where we have only one path) was studied by Lapidoth and Moser [2]. They showed that if the fading process $\{H_k^{(0)}, k \in \mathbb{Z}\}$ is ergodic, then the capacity satisfies [2, Th. 4.41]

$$\lim_{\mathrm{SNR}\to\infty} \left\{C(\mathrm{SNR}) - \log\log \mathrm{SNR}\right\} = \log \pi + \mathsf{E}\left[\log\left|H_1^{(0)}\right|^2\right] - h_0 \tag{10}$$

where $\log(\cdot)$ denotes the natural logarithm function. Thus, at high SNR the capacity of noncoherent frequency-flat fading channels grows double-logarithmically with the SNR. Lapidoth and Moser concluded that communicating over noncoherent frequency-flat fading channels at high SNR is extremely power-inefficient, as one should expect to square the SNR for every additional bit per channel use.[1]

In this paper, we show *inter alia* that communicating over noncoherent multipath fading channels at high SNR is not merely power-inefficient, but may be even worse: if the delay spread is large in the sense that the sequence $\{\alpha_\ell, \ell \in \mathbb{N}_0\}$ (which describes the variances of the path gains) decays exponentially or slower, then capacity is *bounded* in the SNR. For such channels, capacity does not tend to infinity as the SNR tends to infinity. The main results of this paper are presented in the following section.

IV. Main Results

Our main results are a sufficient and a necessary condition on $\{\alpha_\ell, \ell \in \mathbb{N}_0\}$ for $C(\mathrm{SNR})$ to be bounded in the SNR, as well as a characterization of the capacity pre-loglog when the number of paths is finite.

Theorem 1: Consider the above channel model. Then

(i) $$\left(\varliminf_{\ell\to\infty} \frac{\alpha_{\ell+1}}{\alpha_\ell} > 0\right) \Longrightarrow \left(\sup_{\mathrm{SNR}>0} C(\mathrm{SNR}) < \infty\right)$$

[1]Note that the capacity of coherent fading channels (where the fading realization is known to the receiver) grows logarithmically with the SNR [6]. Thus in the coherent case it suffices to double the SNR for every additional bit per channel use.

$$\text{(ii)}\quad \left(\lim_{\ell\to\infty}\frac{1}{\ell}\log\frac{1}{\alpha_\ell}=\infty\right)\Longrightarrow\left(\sup_{\mathrm{SNR}>0}C(\mathrm{SNR})=\infty\right)$$

where we define $a/0 \triangleq \infty$ for every $a>0$ and $0/0 \triangleq 0$.

Proof: Part (i) is proven in Section V-A, and Part (ii) is proven in Sections VI-A and B. ■

By noting that

$$\left(\lim_{\ell\to\infty}\frac{\alpha_{\ell+1}}{\alpha_\ell}=0\right)\Longrightarrow\left(\lim_{\ell\to\infty}\frac{1}{\ell}\log\frac{1}{\alpha_\ell}=\infty\right)$$

we obtain the following.

Corollary 2: Consider the above channel model. Then

$$\text{(i)}\quad \left(\varliminf_{\ell\to\infty}\frac{\alpha_{\ell+1}}{\alpha_\ell}>0\right)\Longrightarrow\left(\sup_{\mathrm{SNR}>0}C(\mathrm{SNR})<\infty\right)$$
$$\text{(ii)}\quad \left(\lim_{\ell\to\infty}\frac{\alpha_{\ell+1}}{\alpha_\ell}=0\right)\Longrightarrow\left(\sup_{\mathrm{SNR}>0}C(\mathrm{SNR})=\infty\right)$$

where we define $a/0 \triangleq \infty$ for every $a>0$ and $0/0 \triangleq 0$.

For example, if

$$\alpha_\ell = e^{-\ell},\qquad \ell\in\mathbb{N}_0$$

then

$$\lim_{\ell\to\infty}\frac{\alpha_{\ell+1}}{\alpha_\ell}=\frac{1}{e}>0$$

and it follows from Part (i) of Corollary 2 that the capacity is bounded in the SNR. However, if

$$\alpha_\ell=\exp\bigl(-\ell^\kappa\bigr),\qquad \ell\in\mathbb{N}_0$$

for some $\kappa>1$, then

$$\lim_{\ell\to\infty}\frac{\alpha_{\ell+1}}{\alpha_\ell}=\lim_{\ell\to\infty}\exp\bigl(\ell^\kappa-(\ell+1)^\kappa\bigr)=0$$

and it follows from Part (ii) of Corollary 2 that the capacity is unbounded in the SNR. Roughly speaking, we can say that if $\{\alpha_\ell,\ell\in\mathbb{N}_0\}$ decays *exponentially or slower*, then $C(\mathrm{SNR})$ is bounded in the SNR, and if $\{\alpha_\ell,\ell\in\mathbb{N}_0\}$ decays *faster than exponentially*, then $C(\mathrm{SNR})$ is unbounded in the SNR. The condition for unbounded capacity (Part (ii) of Corollary 2) is clearly satisfied if the channel has a finite number of paths, since in this case

$$H_k^{(\ell)}=0,\qquad \ell>\mathsf{L},\quad k\in\mathbb{N}$$

which implies

$$\alpha_\ell=0,\qquad \ell>\mathsf{L}\quad\text{and}\quad\frac{\alpha_{\ell+1}}{\alpha_\ell}=\frac{0}{0}\triangleq 0,\qquad \ell>\mathsf{L}.$$

Consequently, it follows from Corollary 2 that if the number of paths is finite, then $C(\mathrm{SNR})$ is unbounded in the SNR. This is not surprising, because if there are only $(\mathsf{L}+1)$ paths, then transmitting only at times that are integer multiples of $(\mathsf{L}+1)$, and measuring the channel outputs only at these times, reduces the channel to a frequency-flat fading channel and demonstrates [using (10)] the achievability of

$$\frac{1}{\mathsf{L}+1}\left(\log\log\mathrm{SNR}+\log\pi+\mathsf{E}\left[\log\left|H_1^{(0)}\right|^2\right]-h_0\right)+o(1)$$

(where $o(1)$ tends to zero as $\mathrm{SNR}\to\infty$), which is unbounded in the SNR.

The above achievable rate suggests that at high SNR the capacity is decreasing in the number of paths. However, Theorem 3 ahead shows that if the number of paths is finite, then the capacity pre-loglog, defined as

$$\Lambda\triangleq\varlimsup_{\mathrm{SNR}\to\infty}\frac{C(\mathrm{SNR})}{\log\log\mathrm{SNR}}\tag{11}$$

is 1 irrespective of the number of paths. The pre-loglog in this case is thus the same as for frequency-flat fading.

Theorem 3: Consider the above channel model. Further assume that the number of paths is finite. Then the limit on the right-hand side (RHS) of (11) exists, and the capacity pre-loglog is given by

$$\Lambda=\lim_{\mathrm{SNR}\to\infty}\frac{C(\mathrm{SNR})}{\log\log\mathrm{SNR}}=1.\tag{12}$$

Proof: See Section V-B for the converse and Sections VI-A and C for the direct part. ■

When studying multipath fading channels at low or at moderate SNR, it is often assumed that the channel has a finite number of paths, even if the number of paths is in reality infinite. This assumption is commonly justified by saying that only the first $(\mathsf{L}+1)$ paths are relevant, since the variances of the remaining paths are typically small and hence the influence of these paths on the capacity is marginal. As we see from Theorems 1 and 3, this argument is not valid at high SNR. In fact, if for example the sequence of variances $\{\alpha_\ell,\ell\in\mathbb{N}_0\}$ decays exponentially, then according to Part (i) of Theorem 1 the capacity is bounded in the SNR. However, if we consider only the first $(\mathsf{L}+1)$ paths and set the other paths to zero, then it follows from Theorem 3 that the capacity is unbounded: irrespective of L it increases double-logarithmically with the SNR. Thus, even though the variances of the remaining paths α_ℓ, $\ell>\mathsf{L}$ can be made arbitrarily small by choosing L sufficiently large, these paths have a significant influence on the capacity behavior at high SNR.

The reason why paths with a small variance can affect the capacity behavior is that the capacity depends on the variance of the product of the path gains and the transmitted signal and not on the variance of the path gains only. Since at high SNR the variance of $\sum_{\ell=\mathsf{L}+1}^\infty H_k^{(\ell)}X_{k-\ell}$ might be huge even if the variance of $\sum_{\ell=\mathsf{L}+1}^\infty H_k^{(\ell)}$ is small, the relevance of a path is determined not only by its own variance but also by the power available at the transmitter. The number of paths that are needed to approximate a multipath fading channel typically depends on the SNR and may grow to infinity as the SNR tends to infinity.

In order to prove the above results, we derive upper and lower bounds on the capacity. Since these bounds may also be of independent interest, we summarize them in the following propositions.

Proposition 4 (Upper Bounds):

i) Consider the above channel model. Further assume that for some $0 < \rho < 1$ and some $\ell_0 \in \mathbb{N}$

$$\alpha_{\ell_0} > 0 \quad \text{and} \quad \frac{\alpha_{\ell+1}}{\alpha_\ell} \geq \rho, \qquad \ell \geq \ell_0.$$

Then the capacity is upper bounded by

$$C(\mathrm{SNR}) \leq \log \frac{2\pi^2}{\sqrt{\tilde{\rho}}} - \inf_{\ell \in \mathcal{L}} (h_\ell - \log \alpha_\ell) \tag{13}$$

for all SNR ≥ 0, where

$$\tilde{\rho} = \min \left\{ \rho^{\ell_0 - 1} \frac{\alpha_{\ell_0}}{\max_{0 \leq \ell' < \ell_0} \alpha_{\ell'}}, \rho^{\ell_0} \right\}. \tag{14}$$

ii) Consider the above channel model. Further assume that

$$\sum_{\ell=0}^{\infty} \alpha_\ell < \infty.$$

Then

$$\varlimsup_{\mathrm{SNR} \to \infty} \{ C(\mathrm{SNR}) - \log\log \mathrm{SNR} \} \leq 1 + \log \pi - \inf_{\ell \in \mathcal{L}} (h_\ell - \log \alpha_\ell). \tag{15}$$

Proof: Part (i) is proven in Section V-A, and Part (ii) is proven in Section V-B. ■

For example, if $\{\alpha_\ell, \ell \in \mathbb{N}_0\}$ is a geometric sequence, i.e.,

$$\alpha_\ell = \rho^\ell, \qquad \ell \in \mathbb{N}_0$$

for some $0 < \rho < 1$, and if the path gains are Gaussian and memoryless so

$$h_\ell = \log(\pi e \alpha_\ell), \qquad \ell \in \mathbb{N}_0$$

then Part (i) of Proposition 4 yields

$$C(\mathrm{SNR}) \leq \log \frac{2\pi}{\sqrt{\rho}} - 1, \qquad \mathrm{SNR} \geq 0. \tag{16}$$

Part (ii) of Proposition 4 combines with (10) to show that the pre-loglog of a multipath fading channel can never be larger than the pre-loglog of a frequency-flat fading channel. This result is consistent with the intuition that at high SNR the multipath behavior is detrimental.

Our last result is a lower bound on the capacity. This bound is the basis for the proof of Part (ii) of Theorem 1 and for the direct part of Theorem 3.

Proposition 5 (Lower Bound): Consider the above channel model. Further assume that

$$\alpha \triangleq \sum_{\ell=0}^{\infty} \alpha_\ell < \infty. \tag{17}$$

Let $\mathsf{L}(\mathsf{P}) \in \mathbb{N}$ be some positive integer that satisfies

$$\sum_{\ell=\mathsf{L}(\mathsf{P})+1}^{\infty} \alpha_\ell \, \mathsf{P} \leq \sigma^2 \tag{18}$$

(if the number of paths $(\mathsf{L}+1)$ is finite, then $\mathsf{L}(\mathsf{P}) = \mathsf{L}$ satisfies (18); otherwise $\mathsf{L}(\mathsf{P})$ grows with P) and let $\tau \in \mathbb{N}$ be some arbitrary positive integer that is allowed to depend on $\mathsf{L}(\mathsf{P})$. Then the capacity $C(\mathrm{SNR})$ is lower bounded by

$$C(\mathrm{SNR}) \geq \frac{\tau}{\mathsf{L}(\mathsf{P}) + \tau} \log\log \mathsf{P}^{1/\tau} + \frac{\tau}{\mathsf{L}(\mathsf{P}) + \tau} \Upsilon, \qquad \mathsf{P} > 1 \tag{19}$$

where

$$\Upsilon = \mathsf{E}\left[\log \left| H_1^{(0)} \right|^2 \right] - 1 - 2\log\left(\sqrt{\alpha_0} + \sqrt{\alpha + 2\sigma^2}\right). \tag{20}$$

Proof: See Section VI-A. ■

V. Proofs of the Upper Bounds

In this section, we prove Proposition 4, which in turn is used to prove Part (i) of Theorem 1 and the converse to Theorem 3.

Part (i) of Proposition 4 is proven in Section V-A, where it is also demonstrated that it implies Part (i) of Theorem 1. Section V-B proves Part (ii) of Proposition 4. This part provides an upper bound on the capacity pre-loglog and will be used later, together with a capacity lower bound that is derived in Section VI, to establish Theorem 3.

A. Bounded Capacity

We provide a proof of Part (i) of Proposition 4 by deriving an upper bound on channel capacity that holds under the assumption that for some $0 < \rho < 1$ and some $\ell_0 \in \mathbb{N}$

$$\alpha_{\ell_0} > 0 \quad \text{and} \quad \frac{\alpha_{\ell+1}}{\alpha_\ell} \geq \rho, \qquad \ell \geq \ell_0. \tag{21}$$

As this bound does not depend on the SNR, Part (i) of Theorem 1 follows immediately from Part (i) of Proposition 4 by noting that if

$$\varliminf_{\ell \to \infty} \frac{\alpha_{\ell+1}}{\alpha_\ell} > 0$$

then we can find a $0 < \rho < 1$ and an $\ell_0 \in \mathbb{N}$ satisfying (21).

The proof of the desired upper bound is akin to the proof of an upper bound that was derived in [7, Sec. VI-A]. (However, [7] studies a channel whose inputs and outputs take value in the set of real numbers rather than in $\mathbb{C}$.) It is based on (7) and on an upper bound on $\frac{1}{n} I(X_1^n; Y_1^n)$.

We begin with the chain rule for mutual information [4, Th. 2.5.2]

$$\frac{1}{n} I(X_1^n; Y_1^n) = \frac{1}{n} \sum_{k=1}^{\ell_0} I(X_1^n; Y_k \mid Y_1^{k-1}) + \frac{1}{n} \sum_{k=\ell_0+1}^{n} I(X_1^n; Y_k \mid Y_1^{k-1}). \tag{22}$$

The first sum on the RHS of (22) consists of ℓ_0 summands, which increase all at most logarithmically with n. Consequently, also the sum increases at most logarithmically with n, and its ratio to n vanishes as n tends to infinity. Indeed, we have for $k = 1, 2, \ldots, \ell_0$

$$\begin{aligned} I\big(X_1^n; Y_k \bigm| Y_1^{k-1}\big) &\le h(Y_k) - h\Big(Y_k \Bigm| Y_1^{k-1}, X_1^n, H_k^{(0)}, H_k^{(1)}, \ldots, H_k^{(k-1)}\Big) \\ &\le \log\left(\pi e\left(\sigma^2 + \sum_{\ell=0}^{k-1} \alpha_\ell \mathsf{E}\left[|X_{k-\ell}|^2\right]\right)\right) - \log\big(\pi e \sigma^2\big) \\ &\le \log\left(1 + \sup_{\ell\in\mathbb{N}_0} \alpha_\ell n \mathrm{SNR}\right) \end{aligned} \tag{23}$$

where the first step follows because conditioning cannot increase differential entropy [4, Th. 9.6.1]; the second step follows from the entropy maximizing property of Gaussian random variables [4, Th. 9.6.5]; and the last step follows by upper bounding

$$\alpha_\ell \le \sup_{\ell'\in\mathbb{N}_0} \alpha_{\ell'}, \qquad \ell = 0, 1, \ldots, k-1$$

and from the power constraint (8).

Before we continue by upper bounding the second sum on the RHS of (22), we would like to pause for some intuition. Firstly, for every $k = \ell_0+1, \ell_0+2, \ldots, n$ there exists a distribution on $X_1, X_2, \ldots, X_n$ for which $I\big(X_1^n; Y_k \bigm| Y_1^{k-1}\big)$ grows double-logarithmically with the SNR. (This is, for example, the case if $X_1 = X_2 = \cdots = X_{k-1} = 0$ and if X_k is such that it achieves the high-SNR asymptotic capacity of noncoherent flat-fading channels [2].) In order to obtain a tight upper bound, we therefore cannot upper bound each summand individually but have to upper bound the sum as a whole. Secondly, the mutual information $I\big(X_1^n; Y_k \bigm| Y_1^{k-1}\big)$ is difficult to evaluate. For example, we could express $I\big(X_1^n; Y_k \bigm| Y_1^{k-1}\big)$ as

$$I\big(X_1^n; Y_k \bigm| Y_1^{k-1}\big) = h\big(Y_k \bigm| Y_1^{k-1}\big) - h\big(Y_k \bigm| X_1^n, Y_1^{k-1}\big)$$

and evaluate

$$h\big(Y_k \bigm| X_1^n, Y_1^{k-1}\big) \approx \mathsf{E}\left[\log\left(\sigma^2 + \sum_{\ell=0}^{k-1} \alpha_\ell |X_{k-\ell}|^2\right)\right].$$

However, the standard approach to upper bound $h\big(Y_k \bigm| Y_1^{k-1}\big)$ by using the entropy maximizing property of Gaussian random variables and that conditioning cannot increase entropy, i.e.,

$$h\big(Y_k \bigm| Y_1^{k-1}\big) \le \log\left(\pi e\left(\sigma^2 + \sum_{\ell=0}^{k-1} \alpha_\ell \mathsf{E}\left[|X_{k-\ell}|^2\right]\right)\right) \tag{24}$$

is not very promising, since it results in an upper bound on $I\big(X_1^n; Y_k \bigm| Y_1^{k-1}\big)$ that is given by the logarithm of a sum of expectations minus the expectation of the logarithm of a sum, which is still difficult to evaluate. Moreover, it is not clear whether (24) results in a tight upper bound on $I\big(X_1^n; Y_k \bigm| Y_1^{k-1}\big)$.

We shall circumvent the latter problem by using the general upper bound on mutual information that was proposed in [2, Th. 5.1], namely

$$I(X;Y) \le \int D\big(W(\cdot|x) \,\|\, R(\cdot)\big) \mathrm{d}Q(x) \tag{25}$$

where $D(\cdot\|\cdot)$ denotes relative entropy, i.e.,

$$D(P_1\|P_0) = \begin{cases} \int \log \frac{\mathrm{d}P_1}{\mathrm{d}P_0} \mathrm{d}P_1 & \text{if } P_1 \ll P_0 \\ +\infty & \text{otherwise} \end{cases}$$

(where $P_1 \ll P_0$ indicates that P_1 is absolutely continuous with respect to P_0), $W(\cdot|\cdot)$ is the channel law, $Q(\cdot)$ denotes the distribution on the channel input X, and $R(\cdot)$ is any distribution on the output alphabet.[2] Thus, any choice of output distribution $R(\cdot)$ yields an upper bound on the mutual information. Note that (25) is tight if $R(\cdot)$ is equal to the capacity-achieving output distribution.

For any given $Y_1^{k-1} = y_1^{k-1}$, we choose the output distribution $R(\cdot)$ to be of density

$$\frac{\sqrt{\beta}}{\pi^2 |y_k|} \frac{1}{1+\beta|y_k|^2}, \qquad y_k \in \mathbb{C} \tag{26}$$

with $\beta = 1/(\tilde{\rho}|y_{k-\ell_0}|^2)$ and[3]

$$\tilde{\rho} = \min\left\{\rho^{\ell_0-1} \frac{\alpha_{\ell_0}}{\max_{0\le\ell'<\ell_0} \alpha_{\ell'}}, \rho^{\ell_0}\right\}. \tag{27}$$

The density (26) corresponds to the density of a circularly-symmetric complex random variable whose magnitude is Cauchy distributed. With this choice

$$0 < \tilde{\rho} < 1 \quad \text{and} \quad \tilde{\rho}\, \alpha_\ell \le \alpha_{\ell+\ell_0}, \qquad \ell \in \mathbb{N}_0. \tag{28}$$

For example, if $\{\alpha_\ell, \ell \in \mathbb{N}_0\}$ is a geometric sequence, then $\ell_0 = 1$ and (27) yields $\tilde{\rho} = \rho$, which clearly satisfies (28). Note that, by choosing β to be dependent on $y_{k-\ell_0}$, the output distribution $R(\cdot)$ has memory in the sense that the present output symbol Y_k depends on the previous symbols Y_1^{k-1}.

Using (26) in (25), and averaging over Y_1^{k-1}, we obtain for $k = \ell_0+1, \ell_0+2, \ldots, n$

$$\begin{aligned} I\big(X_1^n; Y_k \bigm| Y_1^{k-1}\big) \le& \frac{1}{2}\mathsf{E}\left[\log|Y_k|^2\right] + \frac{1}{2}\mathsf{E}\left[\log\left(\tilde{\rho}|Y_{k-\ell_0}|^2\right)\right] \\ &+ \mathsf{E}\left[\log\left(1 + \frac{|Y_k|^2}{\tilde{\rho}|Y_{k-\ell_0}|^2}\right)\right] \\ &- h\big(Y_k \big| X_1^n, Y_1^{k-1}\big) + \log \pi^2 \\ =& \frac{1}{2}\mathsf{E}\left[\log|Y_k|^2\right] - \frac{1}{2}\mathsf{E}\left[\log|Y_{k-\ell_0}|^2\right] \\ &+ \mathsf{E}\left[\log\big(\tilde{\rho}|Y_{k-\ell_0}|^2 + |Y_k|^2\big)\right] \\ &- h\big(Y_k \big| X_1^n, Y_1^{k-1}\big) + \log\frac{\pi^2}{\sqrt{\tilde{\rho}}}. \end{aligned} \tag{29}$$

[2]For channels with finite input and output alphabets this inequality follows by Topsøe's identity [8]; see also [9, Th. 3.4].

[3]If $y_{k-\ell_0} = 0$, then the density (26) is undefined. However, this event is of zero probability and has therefore no impact on the mutual information $I\Big(X_1^n; Y_k \Big| Y_1^{k-1}\Big)$.

By choosing $R(\cdot)$ to be of density (26), we thus obtain an upper bound on $I\left(X_1^n; Y_k \mid Y_1^{k-1}\right)$ that contains only expectations of logarithms and not a mixture of expectations of logarithms and logarithms of expectations. This will facilitate the analysis.

We bound the third and the fourth term in (29) individually. We begin by upper bounding

$$\begin{aligned}\mathsf{E}\left[\log\left(\tilde{\rho}|Y_{k-\ell_0}|^2+|Y_k|^2\right)\right] &= \mathsf{E}\left[\mathsf{E}\left[\log\left(\tilde{\rho}|Y_{k-\ell_0}|^2+|Y_k|^2\right)\middle| X_1^k\right]\right]\\ &\leq \mathsf{E}\left[\log\left(\tilde{\rho}\mathsf{E}\left[|Y_{k-\ell_0}|^2\middle| X_1^{k-\ell_0}\right]+\mathsf{E}\left[|Y_k|^2\middle| X_1^k\right]\right)\right] \quad (30)\end{aligned}$$

using Jensen's inequality. We further note that

$$\mathsf{E}\left[|Y_k|^2\middle| X_1^k\right] = \sigma^2 + \sum_{\ell=0}^{k-1}\alpha_\ell|X_{k-\ell}|^2, \qquad k\in\mathbb{N}$$

which yields (31), shown at the bottom of the page, where the second step follows from (28); the third step follows by substituting $\ell' = \ell + \ell_0$; and the last step follows because

$$\sum_{\ell=\ell_0}^{k-1}\alpha_\ell|X_{k-\ell}|^2 \leq \sum_{\ell=0}^{k-1}\alpha_\ell|X_{k-\ell}|^2.$$

We thus obtain from (30) and (31)

$$\begin{aligned}\mathsf{E}\left[\log\left(\tilde{\rho}|Y_{k-\ell_0}|^2+|Y_k|^2\right)\right] &\leq \log 2 + \mathsf{E}\left[\log\left(\sigma^2+\sum_{\ell=0}^{k-1}\alpha_\ell|X_{k-\ell}|^2\right)\right]. \quad (32)\end{aligned}$$

Next we derive a lower bound on $h\left(Y_k\middle|X_1^n, Y_1^{k-1}\right)$. Let

$$\mathbf{H}_{\ell,1}^{k-1} \triangleq \left(H_1^{(\ell)}, H_2^{(\ell)}, \ldots, H_{k-1}^{(\ell)}\right), \qquad \ell\in\mathbb{N}_0$$

and let

$$\mathbf{H}_1^{k-1} \triangleq \left(\mathbf{H}_{0,1}^{k-1}, \mathbf{H}_{1,1}^{k-1}, \ldots, \mathbf{H}_{k-1,1}^{k-1}\right).$$

We have

$$\begin{aligned}h\left(Y_k \mid X_1^n, Y_1^{k-1}\right) &\geq h\left(Y_k \mid X_1^n, Y_1^{k-1}, \mathbf{H}_1^{k-1}\right)\\ &= h\left(Y_k \mid X_1^n, \mathbf{H}_1^{k-1}\right) \quad (33)\end{aligned}$$

where the first step follows because conditioning cannot increase differential entropy; and where the second step follows because, conditional on $\left(X_1^n, \mathbf{H}_1^{k-1}\right)$, Y_k is independent of Y_1^{k-1}. Let $\mathcal{S}_k$ be defined as

$$\mathcal{S}_k \triangleq \{\ell = 0, 1, \ldots, k-1 : (|x_{k-\ell}| > 0) \cap (\alpha_\ell > 0)\}. \quad (34)$$

Using the entropy power inequality [4, Th. 16.6.3], and using that the processes

$$\left\{H_k^{(0)}, k\in\mathbb{Z}\right\}, \left\{H_k^{(1)}, k\in\mathbb{Z}\right\}, \ldots$$

are independent and jointly independent of X_1^n, it is shown in Appendix A that for any given $X_1^n = x_1^n$

$$\begin{aligned}h\left(Y_k \mid X_1^n = x_1^n, \mathbf{H}_1^{k-1}\right) &= h\left(\sum_{\ell=0}^{k-1} H_k^{(\ell)}X_{k-\ell} + Z_k \,\middle|\, X_1^n = x_1^n, \mathbf{H}_1^{k-1}\right)\\ &\geq \log\left(\sum_{\ell\in\mathcal{S}_k} e^{h\left(H_k^{(\ell)}X_{k-\ell}\middle|X_{k-\ell}=x_{k-\ell},\mathbf{H}_{\ell,1}^{k-1}\right)} + e^{h(Z_k)}\right). \quad (35)\end{aligned}$$

We lower bound the differential entropies on the RHS of (35) as follows. The differential entropies in the sum are lower bounded by

$$\begin{aligned}h\left(H_k^{(\ell)}X_{k-\ell}\,\middle|\, X_{k-\ell} = x_{k-\ell}, \mathbf{H}_{\ell,1}^{k-1}\right) &= \log\left(\alpha_\ell|x_{k-\ell}|^2\right) + h\left(H_k^{(\ell)}\,\middle|\,\mathbf{H}_{\ell,1}^{k-1}\right) - \log\alpha_\ell\\ &\geq \log\left(\alpha_\ell|x_{k-\ell}|^2\right) + \inf_{\ell\in\mathcal{L}}\left(h_\ell - \log\alpha_\ell\right), \qquad \ell\in\mathcal{S}_k \quad (36)\end{aligned}$$

where the first step follows from the behavior of differential entropy under scaling [4, Th. 9.6.4]; and where the second step follows by the stationarity of the process $\left\{H_k^{(\ell)}, k\in\mathbb{Z}\right\}$, which implies that the differential entropy

$$h\left(H_k^{(\ell)}\,\middle|\,\mathbf{H}_{\ell,1}^{k-1}\right), \qquad \ell\in\mathcal{S}_k$$

cannot be smaller than the differential entropy rate h_ℓ [4, Th. 4.2.1 and Th. 4.2.2], and by lower bounding $(h_\ell - \log\alpha_\ell)$ by $\inf_{\ell\in\mathcal{L}}(h_\ell - \log\alpha_\ell)$ (which holds for each $\ell\in\mathcal{S}_k$ because

$$\begin{aligned}\tilde{\rho}\mathsf{E}\left[|Y_{k-\ell_0}|^2\middle| X_1^{k-\ell_0}\right] + \mathsf{E}\left[|Y_k|^2\middle| X_1^k\right] &= \tilde{\rho}\sigma^2 + \sum_{\ell=0}^{k-\ell_0-1}\tilde{\rho}\,\alpha_\ell|X_{k-\ell_0-\ell}|^2 + \sigma^2 + \sum_{\ell=0}^{k-1}\alpha_\ell|X_{k-\ell}|^2\\ &\leq 2\sigma^2 + \sum_{\ell=0}^{k-\ell_0-1}\alpha_{\ell+\ell_0}|X_{k-\ell_0-\ell}|^2 + \sum_{\ell=0}^{k-1}\alpha_\ell|X_{k-\ell}|^2\\ &= 2\sigma^2 + \sum_{\ell'=\ell_0}^{k-1}\alpha_{\ell'}|X_{k-\ell'}|^2 + \sum_{\ell=0}^{k-1}\alpha_\ell|X_{k-\ell}|^2\\ &\leq 2\left(\sigma^2 + \sum_{\ell=0}^{k-1}\alpha_\ell|X_{k-\ell}|^2\right) \quad (31)\end{aligned}$$

$\mathcal{S}_k \subseteq \mathcal{L}$). The last differential entropy on the RHS of (35) is lower bounded by

$$h(Z_k) = \log(\pi e \sigma^2) \geq \inf_{\ell \in \mathcal{L}} (h_\ell - \log \alpha_\ell) + \log \sigma^2 \qquad (37)$$

which follows because conditioning cannot increase differential entropy, and because Gaussian random variables maximize differential entropy, so

$$\begin{aligned} \inf_{\ell \in \mathcal{L}} (h_\ell - \log \alpha_\ell) &\leq \inf_{\ell \in \mathcal{L}} \left(h\left(H_k^{(\ell)}\right) - \log \alpha_\ell \right) \\ &\leq \inf_{\ell \in \mathcal{L}} \left(\log(\pi e \alpha_\ell) - \log \alpha_\ell \right) \\ &= \log(\pi e). \end{aligned} \qquad (38)$$

Applying (36) and (37) to (35), and averaging over X_1^n, yields

$$h\left(Y_k \mid X_1^n, Y_1^{k-1}\right) \geq \mathsf{E}\left[\log\left(\sigma^2 + \sum_{\ell=0}^{k-1} \alpha_\ell |X_{k-\ell}|^2\right)\right] + \inf_{\ell \in \mathcal{L}} (h_\ell - \log \alpha_\ell). \qquad (39)$$

Returning to the analysis of (29), we obtain from (32) and (39)

$$\begin{aligned} &I\left(X_1^n; Y_k \mid Y_1^{k-1}\right) \\ &\leq \frac{1}{2}\mathsf{E}\left[\log|Y_k|^2\right] - \frac{1}{2}\mathsf{E}\left[\log|Y_{k-\ell_0}|^2\right] \\ &\quad + \log 2 + \mathsf{E}\left[\log\left(\sigma^2 + \sum_{\ell=0}^{k-1} \alpha_\ell |X_{k-\ell}|^2\right)\right] \\ &\quad - \mathsf{E}\left[\log\left(\sigma^2 + \sum_{\ell=0}^{k-1} \alpha_\ell |X_{k-\ell}|^2\right)\right] \\ &\quad - \inf_{\ell \in \mathcal{L}} (h_\ell - \log \alpha_\ell) + \log \frac{\pi^2}{\sqrt{\tilde{\rho}}} \\ &= \frac{1}{2}\mathsf{E}\left[\log|Y_k|^2\right] - \frac{1}{2}\mathsf{E}\left[\log|Y_{k-\ell_0}|^2\right] + \mathsf{K} \end{aligned} \qquad (40)$$

where we define

$$\mathsf{K} \triangleq \log \frac{2\pi^2}{\sqrt{\tilde{\rho}}} - \inf_{\ell \in \mathcal{L}} (h_\ell - \log \alpha_\ell). \qquad (41)$$

Applying (40) and (23) to (22), we have

$$\begin{aligned} &\frac{1}{n} I(X_1^n; Y_1^n) \\ &\leq \frac{1}{n} \sum_{k=1}^{\ell_0} \log\left(1 + \sup_{\ell \in \mathbb{N}_0} \alpha_\ell n \mathrm{SNR}\right) \\ &\quad + \frac{1}{n} \sum_{k=\ell_0+1}^{n} \left(\frac{1}{2}\mathsf{E}\left[\log|Y_k|^2\right] - \frac{1}{2}\mathsf{E}\left[\log|Y_{k-\ell_0}|^2\right] + \mathsf{K} \right) \\ &= \frac{\ell_0}{n} \log\left(1 + \sup_{\ell \in \mathbb{N}_0} \alpha_\ell n \mathrm{SNR}\right) + \frac{n-\ell_0}{n}\mathsf{K} \\ &\quad + \frac{1}{2n} \sum_{k=\ell_0+1}^{n} \left(\mathsf{E}\left[\log|Y_k|^2\right] - \mathsf{E}\left[\log|Y_{k-\ell_0}|^2\right] \right). \end{aligned} \qquad (42)$$

Note that for every $k = \ell_0 + 1, \ell_0 + 2, \ldots, n$ there exists a distribution on $X_1, X_2, \ldots, X_n$ for which the difference $\mathsf{E}\left[\log|Y_k|^2\right] - \mathsf{E}\left[\log|Y_{k-\ell_0}|^2\right]$ is unbounded in the SNR. However, the average

$$\frac{1}{2n} \sum_{k=\ell_0+1}^{n} \left(\mathsf{E}\left[\log|Y_k|^2\right] - \mathsf{E}\left[\log|Y_{k-\ell_0}|^2\right] \right)$$

vanishes as n tends to infinity, since the term $\mathsf{E}\left[\log|Y_{k-\ell_0}|^2\right]$ cancels $\mathsf{E}\left[\log|Y_{k'}|^2\right]$ at $k' = k - \ell_0$ for every $k = 2\ell_0 + 1, 2\ell_0 + 2, \ldots, n$. To show this mathematically, we note that for any sequences $\{a_k, k \in \mathbb{N}\}$ and $\{b_k, k \in \mathbb{N}\}$

$$\sum_{k=\ell_0+1}^{n} (a_k - b_k) = \sum_{k=n-\ell_0+1}^{n} (a_k - b_{k-n+2\ell_0}) + \sum_{k=\ell_0+1}^{n-\ell_0} (a_k - b_{k+\ell_0}). \qquad (43)$$

Defining $a_k \triangleq \mathsf{E}\left[\log|Y_k|^2\right]$ and $b_k \triangleq \mathsf{E}\left[\log|Y_{k-\ell_0}|^2\right]$ we have for the first sum on the RHS of (43)

$$\begin{aligned} &\sum_{k=n-\ell_0+1}^{n} (a_k - b_{k-n+2\ell_0}) \\ &= \sum_{k=n-\ell_0+1}^{n} \left(\mathsf{E}\left[\log|Y_k|^2\right] - \mathsf{E}\left[\log|Y_{k-n+\ell_0}|^2\right] \right) \\ &\leq \sum_{k=n-\ell_0+1}^{n} \left(\log \mathsf{E}\left[|Y_k|^2\right] - \mathsf{E}\left[\log|Y_{k-n+\ell_0}|^2\right] \right) \\ &\leq \sum_{k=n-\ell_0+1}^{n} \left(\log\left(\sigma^2 + \sup_{\ell \in \mathbb{N}_0} \alpha_\ell \, n \, \mathsf{P}\right) - \mathsf{E}\left[\log|Y_{k-n+\ell_0}|^2\right] \right) \\ &\leq \sum_{k=n-\ell_0+1}^{n} \left(\log\left(\sigma^2 + \sup_{\ell \in \mathbb{N}_0} \alpha_\ell \, n \, \mathsf{P}\right) - \mathsf{E}\left[\log|Z_{k-n+\ell_0}|^2\right] \right) \\ &= \ell_0 \log\left(1 + \sup_{\ell \in \mathbb{N}_0} \alpha_\ell n \mathrm{SNR}\right) + \ell_0 \gamma \end{aligned} \qquad (44)$$

where $\gamma \approx 0.577$ denotes Euler's constant. Here the second step follows by Jensen's inequality; the third step follows by upper bounding

$$\mathsf{E}\left[|Y_k|^2\right] = \sigma^2 + \sum_{\ell=0}^{k-1} \alpha_\ell \mathsf{E}\left[|X_{k-\ell}|^2\right] \leq \sigma^2 + \sup_{\ell \in \mathbb{N}_0} \alpha_\ell n \mathsf{P}$$

and the fourth step follows by noting that, conditional on $\sum_{\ell=0}^{k-n+\ell_0-1} H_{k-n+\ell_0}^{(\ell)} X_{k-n+\ell_0-\ell}$, we have that $|Y_{k-n+\ell_0}|^2$ is stochastically larger than $|Z_{k-n+\ell_0}|^2$, so

$$\begin{aligned} &\mathsf{E}\left[\log|Y_{k-n+\ell_0}|^2 \,\middle|\, \sum_{\ell=0}^{k-n+\ell_0-1} H_{k-n+\ell_0}^{(\ell)} X_{k-n+\ell_0-\ell}\right] \\ &\geq \mathsf{E}\left[\log|Z_{k-n+\ell_0}|^2 \,\middle|\, \sum_{\ell=0}^{k-n+\ell_0-1} H_{k-n+\ell_0}^{(\ell)} X_{k-n+\ell_0-\ell}\right] \end{aligned}$$

from which we obtain the lower bound

$$\mathsf{E}\left[\log|Y_{k-n+\ell_0}|^2\right] \geq \mathsf{E}\left[\log|Z_{k-n+\ell_0}|^2\right]$$

upon averaging over $\sum_{\ell=0}^{k-n+\ell_0-1} H_{k-n+\ell_0}^{(\ell)} X_{k-n+\ell_0-\ell}$ (see [2, Sec. VI-B] and in particular [2, Lemma 6.2b]); and the last step follows by noting that the expected logarithm of an exponentially distributed random variable of mean σ^2 is given by $\mathsf{E}\left[\log|Z_{k-n+\ell_0}|^2\right] = \log\sigma^2 - \gamma$.

For the second sum on the RHS of (43) we have

$$\sum_{k=\ell_0+1}^{n-\ell_0}(a_k - b_{k+\ell_0}) = \sum_{k=\ell_0+1}^{n-\ell_0}\left(\mathsf{E}\left[\log|Y_k|^2\right] - \mathsf{E}\left[\log|Y_k|^2\right]\right) = 0. \quad (45)$$

Thus applying (43)–(45) to (42) yields

$$\frac{1}{n}I(X_1^n;Y_1^n) \le \frac{3\ell_0}{2n}\log\left(1+\sup_{\ell\in\mathbb{N}_0}\alpha_\ell n\mathrm{SNR}\right) + \frac{n-\ell_0}{n}\mathsf{K} + \frac{\ell_0}{2n}\gamma \quad (46)$$

which tends to

$$\mathsf{K} = \log\frac{2\pi^2}{\sqrt{\tilde{\rho}}} - \inf_{\ell\in\mathcal{L}}(h_\ell - \log\alpha_\ell) < \infty$$

as n tends to infinity. This proves Part (i) of Proposition 4.

B. Unbounded Capacity

We prove Part (ii) of Proposition 4 by deriving an upper bound on capacity that holds under the assumption that

$$\sum_{\ell=0}^{\infty}\alpha_\ell < \infty.$$

From this upper bound follows that

$$\varlimsup_{\mathrm{SNR}\to\infty}\left\{C(\mathrm{SNR}) - \log\log\mathrm{SNR}\right\} < \infty \quad (47)$$

which in turn shows that the capacity pre-loglog is upper bounded by

$$\Lambda \triangleq \varlimsup_{\mathrm{SNR}\to\infty}\frac{C(\mathrm{SNR})}{\log\log\mathrm{SNR}} \le 1. \quad (48)$$

This yields the converse to Theorem 3.

As in Section V-A, the desired upper bound follows by (7) and by deriving an upper bound on $\frac{1}{n}I(X_1^n;Y_1^n)$. To this end, we begin with the chain rule for mutual information

$$I(X_1^n;Y_1^n) = \sum_{k=1}^{n} I\left(X_1^n;Y_k\middle|Y_1^{k-1}\right) \quad (49)$$

and upper bound each summand on the RHS of (49) by [2, (27)]

$$\begin{aligned} I\left(X_1^n;Y_k \mid Y_1^{k-1}\right) &\le \mathsf{E}\left[\log|Y_k|^2\right] - h\left(Y_k \mid X_1^n, Y_1^{k-1}\right) \\ &\quad + \xi\left(1+\log\mathsf{E}\left[|Y_k|^2\right] - \mathsf{E}\left[\log|Y_k|^2\right]\right) \\ &\quad + \log\Gamma(\xi) - \xi\log\xi + \log\pi \\ &= (1-\xi)\mathsf{E}\left[\log|Y_k|^2\right] - h\left(Y_k \mid X_1^n, Y_1^{k-1}\right) \\ &\quad + \xi\left(1+\log\mathsf{E}\left[|Y_k|^2\right]\right) + \log\Gamma(\xi) - \xi\log\xi + \log\pi \end{aligned} \quad (50)$$

which holds for any $\xi > 0$. Here $\Gamma(\cdot)$ denotes the Gamma function. The upper bound (50) follows from (25) by choosing $R(\cdot)$ to be the distribution of a circularly-smmmetric complex random variable whose squared magnitude is Gamma distributed. It was noted in [2, p. 2429] that this upper bound is tight enough to show that the capacity of noncoherent frequency-flat fading channels does not grow faster than double-logarithmically with the SNR.

We evaluate the terms on the RHS of (50) individually. We upper bound the first term using Jensen's inequality

$$\begin{aligned} \mathsf{E}\left[\log|Y_k|^2\right] &= \mathsf{E}\left[\mathsf{E}\left[\log|Y_k|^2 \middle| X_1^k\right]\right] \\ &\le \mathsf{E}\left[\log\mathsf{E}\left[|Y_k|^2 \middle| X_1^k\right]\right] \\ &= \mathsf{E}\left[\log\left(\sigma^2 + \sum_{\ell=0}^{k-1}\alpha_\ell|X_{k-\ell}|^2\right)\right]. \end{aligned} \quad (51)$$

The second term was already evaluated in (39)

$$h\left(Y_k\middle|X_1^n, Y_1^{k-1}\right) \ge \mathsf{E}\left[\log\left(\sigma^2 + \sum_{\ell=0}^{k-1}\alpha_\ell|X_{k-\ell}|^2\right)\right] + \inf_{\ell\in\mathcal{L}}(h_\ell - \alpha_\ell) \quad (52)$$

and the next term is readily evaluated as

$$\log\mathsf{E}\left[|Y_k|^2\right] = \log\left(\sigma^2 + \sum_{\ell=0}^{k-1}\alpha_\ell\mathsf{E}\left[|X_{k-\ell}|^2\right]\right). \quad (53)$$

Our choice of ξ will satisfy $\xi < 1$ [see (55)]. We therefore obtain, upon substituting (51)–(53) in (50)

$$\begin{aligned} I\left(X_1^n;Y_k \mid Y_1^{k-1}\right) &\le (1-\xi)\mathsf{E}\left[\log\left(\sigma^2 + \sum_{\ell=0}^{k-1}\alpha_\ell|X_{k-\ell}|^2\right)\right] \\ &\quad - \mathsf{E}\left[\log\left(\sigma^2 + \sum_{\ell=0}^{k-1}\alpha_\ell|X_{k-\ell}|^2\right)\right] \\ &\quad - \inf_{\ell\in\mathcal{L}}(h_\ell - \alpha_\ell) \\ &\quad + \xi\left\{1 + \log\left(\sigma^2 + \sum_{\ell=0}^{k-1}\alpha_\ell\mathsf{E}\left[|X_{k-\ell}|^2\right]\right)\right\} \\ &\quad + \log\Gamma(\xi) - \xi\log\xi + \log\pi \\ &= -\inf_{\ell\in\mathcal{L}}(h_\ell - \alpha_\ell) + \log\Gamma(\xi) - \xi\log\xi + \log\pi \\ &\quad + \xi\left\{1 + \log\left(\sigma^2 + \sum_{\ell=0}^{k-1}\alpha_\ell\mathsf{E}\left[|X_{k-\ell}|^2\right]\right)\right. \\ &\qquad \left. - \mathsf{E}\left[\log\left(\sigma^2 + \sum_{\ell=0}^{k-1}\alpha_\ell|X_{k-\ell}|^2\right)\right]\right\} \\ &\le -\inf_{\ell\in\mathcal{L}}(h_\ell - \alpha_\ell) + \log\Gamma(\xi) - \xi\log\xi + \log\pi \\ &\quad + \xi\left\{1 + \log\left(1 + \sum\nolimits_{\ell=0}^{k-1}\alpha_\ell\frac{\mathsf{E}\left[|X_{k-\ell}|^2\right]}{\sigma^2}\right)\right\} \end{aligned} \quad (54)$$

where the last inequality follows by lower bounding

$$\mathsf{E}\left[\log\left(\sigma^2+\sum_{\ell=0}^{k-1}\alpha_\ell|X_{k-\ell}|^2\right)\right]\geq\log\sigma^2.$$

We choose

$$\xi=\frac{1}{1+\log\bigl(1+\alpha\,\mathrm{SNR}\bigr)} \tag{55}$$

where we define

$$\alpha\triangleq\sum_{\ell=0}^{\infty}\alpha_\ell.$$

Defining

$$\Psi(\mathrm{SNR})\triangleq\left[\log\Gamma(\xi)-\log\frac{1}{\xi}-\xi\log\xi\right|_{\xi=\left(1+\log(1+\alpha\mathrm{SNR})\right)^{-1}}$$

we obtain

$$\begin{aligned}I\bigl(X_1^n;Y_k\bigm|Y_1^{k-1}\bigr)&\\ \leq-\inf_{\ell\in\mathcal{L}}&(h_\ell-\alpha_\ell)+\log\pi\\ &+\log\bigl(1+\log(1+\alpha\,\mathrm{SNR})\bigr)+\Psi(\mathrm{SNR})\\ &+\frac{1+\log\left(1+\sum_{\ell=0}^{k-1}\alpha_\ell\frac{\mathsf{E}[|X_{k-\ell}|^2]}{\sigma^2}\right)}{1+\log\left(1+\alpha\,\mathrm{SNR}\right)}.\end{aligned} \tag{56}$$

Using (56) in (49) yields then

$$\begin{aligned}\frac{1}{n}I\bigl(X_1^n;Y_1^n\bigr)&\\ \leq-\inf_{\ell\in\mathcal{L}}&(h_\ell-\alpha_\ell)+\log\pi\\ &+\log\bigl(1+\log(1+\alpha\,\mathrm{SNR})\bigr)+\Psi(\mathrm{SNR})\\ &+\frac{1+\frac{1}{n}\sum_{k=1}^{n}\log\left(1+\sum_{\ell=0}^{k-1}\alpha_\ell\frac{\mathsf{E}[|X_{k-\ell}|^2]}{\sigma^2}\right)}{1+\log\left(1+\alpha\,\mathrm{SNR}\right)}.\end{aligned} \tag{57}$$

By Jensen's inequality we obtain

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n}\log&\left(1+\sum_{\ell=0}^{k-1}\alpha_\ell\frac{\mathsf{E}\left[|X_{k-\ell}|^2\right]}{\sigma^2}\right)\\ &\leq\log\left(1+\frac{1}{n}\sum_{k=1}^{n}\sum_{\ell=0}^{k-1}\alpha_\ell\frac{\mathsf{E}\left[|X_{k-\ell}|^2\right]}{\sigma^2}\right)\\ &\leq\log\left(1+\alpha\,\mathrm{SNR}\right)\end{aligned} \tag{58}$$

where the last step follows by rearranging the double sum as

$$\frac{1}{n}\sum_{k=1}^{n}\frac{\mathsf{E}\left[|X_k|^2\right]}{\sigma^2}\sum_{\ell=0}^{n-k}\alpha_\ell$$

and by upper bounding then $\sum_{\ell=0}^{k-n}\alpha_\ell\leq\alpha$ and using the power constraint (8).

Combining (58) and (57) with (7), we obtain the upper bound

$$\begin{aligned}C(\mathrm{SNR})\leq-\inf_{\ell\in\mathcal{L}}(h_\ell-\alpha_\ell)+\log\pi+1&\\ +\log\bigl(1+\log(1+\alpha\,\mathrm{SNR})\bigr)+\Psi(\mathrm{SNR}).\end{aligned} \tag{59}$$

It follows from [2, (337)] that

$$\lim_{\mathrm{SNR}\to\infty}\Psi(\mathrm{SNR})=\lim_{\xi\downarrow0}\left\{\log\Gamma(\xi)-\log\frac{1}{\xi}-\xi\log\xi\right\}=0. \tag{60}$$

Noting that

$$\lim_{\mathrm{SNR}\to\infty}\Bigl\{\log\bigl(1+\log(1+\alpha\,\mathrm{SNR})\bigr)-\log\log\mathrm{SNR}\Bigr\}=0$$

we obtain from (59) and (60) the desired result

$$\begin{aligned}\varlimsup_{\mathrm{SNR}\to\infty}&\bigl\{C(\mathrm{SNR})-\log\log\mathrm{SNR}\bigr\}\\ &\leq1+\log\pi-\inf_{\ell\in\mathcal{L}}(h_\ell-\alpha_\ell).\end{aligned} \tag{61}$$

VI. Proofs of the Achievability Results

In Section VI-A, we derive the lower bound on channel capacity that is presented in Proposition 5. This lower bound will be used in Sections VI-B and VI-C to prove Part (ii) of Theorem 1 and to prove the direct part of Theorem 3, respectively.

A. Lower Bound

To derive the desired lower bound on capacity, we evaluate $\frac{1}{n}I(X_1^n;Y_1^n)$ for the following distribution on the inputs $\{X_k,k\in\mathbb{Z}\}$.

Let $\mathsf{L}(\mathsf{P})$ be such that

$$\sum_{\ell=\mathsf{L}(\mathsf{P})+1}^{\infty}\alpha_\ell\,\mathsf{P}\leq\sigma^2. \tag{62}$$

To shorten notation, we shall write in the following L instead of $\mathsf{L}(\mathsf{P})$. Let $\tau\in\mathbb{N}$ be some positive integer that possibly depends on L, and let $\mathbf{X}_b=(X_{b(\mathsf{L}+\tau)+1},\ldots,X_{(b+1)(\mathsf{L}+\tau)})$. We choose $\{\mathbf{X}_b,b\in\mathbb{Z}\}$ to be i.i.d. with

$$\mathbf{X}_b=\bigl(\underbrace{0,\ldots,0}_{\mathsf{L}},\tilde{X}_{b\tau+1},\ldots,\tilde{X}_{(b+1)\tau}\bigr)$$

where $\tilde{X}_{b\tau+1},\tilde{X}_{b\tau+2},\ldots,\tilde{X}_{(b+1)\tau}$ is a sequence of independent, zero-mean, circularly-symmetric, complex random variables with $\log|\tilde{X}_{b\tau+\nu}|^2$ being uniformly distributed over the interval $\bigl[\log\mathsf{P}^{(\nu-1)/\tau},\log\mathsf{P}^{\nu/\tau}\bigr]$, i.e., for each $\nu=1,2,\ldots,\tau$

$$\log|\tilde{X}_{b\tau+\nu}|^2\sim\mathcal{U}\left(\left[\log\mathsf{P}^{(\nu-1)/\tau},\log\mathsf{P}^{\nu/\tau}\right]\right).$$

(Here and throughout this proof we assume that $\mathsf{P}>1$.)

Here the first L zero-symbols in $\mathbf{X}_b$ are guard symbols and ensure that transmission in block b is not affected by transmissions that took place in previous blocks. The subsequent symbols are chosen such that for every $\nu=1,2,\ldots,\tau$ the power of $\tilde{X}_{b\tau+\nu}$ is much larger than the power of the previously (in the same block) transmitted symbols, guaranteeing that the interference from these symbols is insignificant.

In the following we analyze the information rates that are achievable with this coding scheme. Define $\kappa \triangleq \lfloor \frac{n}{\mathsf{L}+\tau} \rfloor$ (where $\lfloor a \rfloor$ denotes the largest integer that is less than or equal to a), and let $\mathbf{Y}_b$ denote the vector $(Y_{b(L+\tau)+1}, \ldots, Y_{(b+1)(\mathsf{L}+\tau)})$. By the chain rule for mutual information we have

$$\begin{aligned} I(X_1^n; Y_1^n) &\geq I(\mathbf{X}_0^{\kappa-1}; \mathbf{Y}_0^{\kappa-1}) \\ &= \sum_{b=0}^{\kappa-1} I(\mathbf{X}_b; \mathbf{Y}_0^{\kappa-1} \mid \mathbf{X}_0^{b-1}) \\ &\geq \sum_{b=0}^{\kappa-1} I(\mathbf{X}_b; \mathbf{Y}_b) \end{aligned} \tag{63}$$

where the first step follows by restricting the number of observables; and where the last step follows by restricting the number of observables and by noting that $\{\mathbf{X}_b, b \in \mathbb{Z}\}$ is i.i.d. We continue by lower bounding each summand on the RHS of (63). We use again the chain rule and that reducing observations cannot increase mutual information to obtain

$$\begin{aligned} I(\mathbf{X}_b; \mathbf{Y}_b) &= \sum_{\nu=1}^{\tau} I(\tilde{X}_{b\tau+\nu}; \mathbf{Y}_b \mid \tilde{X}_{b\tau+1}^{b\tau+\nu-1}) \\ &\geq \sum_{\nu=1}^{\tau} I(\tilde{X}_{b\tau+\nu}; Y_{b(\mathsf{L}+\tau)+\mathsf{L}+\nu} \mid \tilde{X}_{b\tau+1}^{b\tau+\nu-1}) \\ &\geq \sum_{\nu=1}^{\tau} I(\tilde{X}_{b\tau+\nu}; Y_{b(\mathsf{L}+\tau)+\mathsf{L}+\nu}) \end{aligned} \tag{64}$$

where we have additionally used in the last inequality that $\tilde{X}_{b\tau+1}, \tilde{X}_{b\tau+2}, \ldots, \tilde{X}_{(b+1)\tau}$ are independent.

Defining

$$W_{b\tau+\nu} \triangleq \sum_{\ell=1}^{b(\mathsf{L}+\tau)+\mathsf{L}+\nu-1} H_{b(\mathsf{L}+\tau)+\mathsf{L}+\nu}^{(\ell)} X_{b(\mathsf{L}+\tau)+\mathsf{L}+\nu-\ell} + Z_{b(\mathsf{L}+\tau)+\mathsf{L}+\nu} \tag{65}$$

each summand on the RHS of (64) can be written as

$$\begin{aligned} &I(\tilde{X}_{b\tau+\nu}; Y_{b(\mathsf{L}+\tau)+\mathsf{L}+\nu}) \\ &\quad = I(\tilde{X}_{b\tau+\nu}; H_{b(\mathsf{L}+\tau)+\mathsf{L}+\nu}^{(0)} \tilde{X}_{b\tau+\nu} + W_{b\tau+\nu}). \end{aligned} \tag{66}$$

A lower bound on (66) follows from the following lemma.

Lemma 6: Let the random variables X, H, and W have finite second moments. Assume that both X and H are of finite differential entropy. Finally, assume that X is independent of H; that X is independent of W; and that X —○— H —○— W forms a Markov chain. Then

$$\begin{aligned} I(X; HX+W) &\geq h(X) - \mathsf{E}\left[\log|X|^2\right] \\ &\quad + \mathsf{E}\left[\log|H|^2\right] - \mathsf{E}\left[\log\left(\pi e \left(\sigma_H + \frac{\sigma_W}{|X|}\right)^2\right)\right] \end{aligned} \tag{67}$$

where $\sigma_W^2 \geq 0$ and $\sigma_H^2 > 0$ denote the variances of W and H. (Note that the assumptions that X and H have finite second moments and are of finite differential entropy guarantee that $\mathsf{E}\left[\log|X|^2\right]$ and $\mathsf{E}\left[\log|H|^2\right]$ are finite, see [2, Lemma 6.7e].)

Proof: See [10, Lemma 4]. ■

It can be easily verified that for the channel model given in Section II and for the above coding scheme the lemma's conditions are satisfied. We therefore obtain from Lemma 6

$$\begin{aligned} &I(\tilde{X}_{b\tau+\nu}; H_{b(\mathsf{L}+\tau)+\mathsf{L}+\nu}^{(0)} \tilde{X}_{b\tau+\nu} + W_{b\tau+\nu}) \\ &\quad \geq h(\tilde{X}_{b\tau+\nu}) - \mathsf{E}\left[\log|\tilde{X}_{b\tau+\nu}|^2\right] \\ &\qquad + \mathsf{E}\left[\log\left|H_{b(\mathsf{L}+\tau)+\mathsf{L}+\nu}^{(0)}\right|^2\right] \\ &\qquad - \mathsf{E}\left[\log\left(\pi e \left(\sqrt{\alpha_0} + \frac{\sqrt{\mathsf{E}\left[|W_{b\tau+\nu}|^2\right]}}{|\tilde{X}_{b\tau+\nu}|}\right)^2\right)\right]. \end{aligned} \tag{68}$$

Using that the differential entropy of a circularly-smmmetric random variable is given by [2, Eqs. (320) and (316)]

$$h(\tilde{X}_{b\tau+\nu}) = \mathsf{E}\left[\log|\tilde{X}_{b\tau+\nu}|^2\right] + h(\log|\tilde{X}_{b\tau+\nu}|^2) + \log\pi \tag{69}$$

and evaluating $h(\log|\tilde{X}_{b\tau+\nu}|^2)$ for our choice of $\tilde{X}_{b\tau+\nu}$, yields for the first two terms on the RHS of (68)

$$h(\tilde{X}_{b\tau+\nu}) - \mathsf{E}\left[\log|\tilde{X}_{b\tau+\nu}|^2\right] = \log\log \mathsf{P}^{1/\tau} + \log\pi. \tag{70}$$

We next upper bound

$$\begin{aligned} \frac{\mathsf{E}\left[|W_{b\tau+\nu}|^2\right]}{|\tilde{X}_{b\tau+\nu}|^2} &= \sum_{\ell=1}^{\mathsf{L}} \alpha_\ell \frac{\mathsf{E}\left[|X_{b(\mathsf{L}+\tau)+\mathsf{L}+\nu-\ell}|^2\right]}{|\tilde{X}_{b\tau+\nu}|^2} \\ &\quad + \sum_{\ell=\mathsf{L}+1}^{b(\mathsf{L}+\tau)+\mathsf{L}+\nu-1} \alpha_\ell \frac{\mathsf{E}\left[|X_{b(\mathsf{L}+\tau)+\mathsf{L}+\nu-\ell}|^2\right]}{|\tilde{X}_{b\tau+\nu}|^2} \\ &\quad + \frac{\sigma^2}{|\tilde{X}_{b\tau+\nu}|^2}. \end{aligned} \tag{71}$$

To this end, we note that for our choice of $\{X_k, k \in \mathbb{Z}\}$ and by the assumption that $\mathsf{P} > 1$, we have

$$\mathsf{E}\left[|X_\ell|^2\right] \leq \mathsf{P}, \quad \ell \in \mathbb{N} \tag{72}$$

$$\mathsf{E}\left[|X_{b(\mathsf{L}+\tau)+\mathsf{L}+\nu-\ell}|^2\right] \leq \mathsf{P}^{(\nu-\ell)/\tau}, \quad \ell = 1, 2, \ldots, \mathsf{L} \tag{73}$$

$$|\tilde{X}_{b\tau+\nu}|^2 \geq \mathsf{P}^{(\nu-1)/\tau} \geq 1 \tag{74}$$

from which we obtain

$$\begin{aligned} \frac{\mathsf{E}\left[|X_{b(\mathsf{L}+\tau)+\mathsf{L}+\nu-\ell}|^2\right]}{|\tilde{X}_{b\tau+\nu}|^2} &\leq \frac{\mathsf{P}^{(\nu-\ell)/\tau}}{\mathsf{P}^{(\nu-1)/\tau}} \\ &\leq 1, \quad \ell = 1, 2, \ldots, \mathsf{L} \end{aligned} \tag{75}$$

and

$$\frac{\mathsf{E}\left[|X_{b(\mathsf{L}+\tau)+\mathsf{L}+\nu-\ell}|^2\right]}{|\tilde{X}_{b\tau+\nu}|^2} \leq \mathsf{P}, \quad \mathsf{L}+1 \leq \ell < b(\mathsf{L}+\tau)+\mathsf{L}+\nu. \tag{76}$$

Applying (74)–(76) to (71) yields

$$\begin{aligned} \frac{\mathsf{E}\left[|W_{b\tau+\nu}|^2\right]}{|\tilde{X}_{b\tau+\nu}|^2} &\leq \sum_{\ell=1}^{\mathsf{L}} \alpha_\ell + \sum_{\ell=\mathsf{L}+1}^{b(\mathsf{L}+\tau)+\mathsf{L}+\nu-1} \alpha_\ell\, \mathsf{P} + \sigma^2 \\ &\leq \alpha + \sum_{\ell=\mathsf{L}+1}^{\infty} \alpha_\ell\, \mathsf{P} + \sigma^2 \\ &\leq \alpha + 2\sigma^2 \end{aligned} \tag{77}$$

where

$$\alpha \triangleq \sum_{\ell=0}^{\infty} \alpha_\ell.$$

Here the second step follows because α_ℓ, $\ell \in \mathbb{N}_0$ and P are nonnegative, and the last step follows from (62).

By combining (68) with (70) and (77), and by noting that by the stationarity of $\{H_k^{(0)}, k \in \mathbb{Z}\}$

$$\mathsf{E}\left[\log\left|H^{(0)}_{b(\mathsf{L}+\tau)+\mathsf{L}+\nu}\right|^2\right] = \mathsf{E}\left[\log\left|H_1^{(0)}\right|^2\right]$$

we obtain the lower bound

$$\begin{aligned} I\big(\tilde{X}_{b\tau+\nu}; H^{(0)}_{b(\mathsf{L}+\tau)+\mathsf{L}+\nu}\tilde{X}_{b\tau+\nu} + W_{b\tau+\nu}\big) &\\ \geq \log\log \mathsf{P}^{1/\tau} + \mathsf{E}\left[\log\left|H_1^{(0)}\right|^2\right] - 1 &\\ - 2\log\left(\sqrt{\alpha_0} + \sqrt{\alpha + 2\sigma^2}\right). & \quad (78) \end{aligned}$$

Note that the RHS of (78) neither depends on ν nor on b. We therefore obtain from (78), (64), and (63)

$$I\big(X_1^n; Y_1^n\big) \geq \kappa\tau \log\log \mathsf{P}^{1/\tau} + \kappa\tau\Upsilon \quad (79)$$

where Υ is defined as

$$\Upsilon \triangleq \mathsf{E}\left[\log\left|H_1^{(0)}\right|^2\right] - 1 - 2\log\left(\sqrt{\alpha_0} + \sqrt{\alpha + 2\sigma^2}\right).$$

Dividing the RHS of (79) by n, and computing the limit as n tends to infinity, yields the lower bound

$$C(\text{SNR}) \geq \frac{\tau}{\mathsf{L}+\tau} \log\log \mathsf{P}^{1/\tau} + \frac{\tau}{\mathsf{L}+\tau}\Upsilon, \quad \mathsf{P} > 1 \quad (80)$$

where we have used that $\lim_{n\to\infty} \kappa/n = 1/(\mathsf{L}+\tau)$. This proves Proposition 5.

B. Condition for Unbounded Capacity

We use Proposition 5 to prove Part (ii) of Theorem 1. In particular, we show that if

$$\lim_{\ell\to\infty} \frac{1}{\ell} \log \frac{1}{\alpha_\ell} = \infty \quad (81)$$

then, by cleverly choosing $\mathsf{L}(\mathsf{P})$ and τ, the lower bound (19), namely

$$C(\text{SNR}) \geq \frac{\tau}{\mathsf{L}(\mathsf{P})+\tau} \log\log \mathsf{P}^{1/\tau} + \frac{\tau}{\mathsf{L}(\mathsf{P})+\tau}\Upsilon, \quad \mathsf{P} > 1$$

can be made arbitrarily large as the SNR tends to infinity. To this end, we first note that by (81) there exists for every $0 < \varrho < 1$ an $\ell_0 \in \mathbb{N}$ such that

$$\alpha_\ell < \varrho^\ell, \quad \ell \geq \ell_0. \quad (82)$$

By choosing

$$\mathsf{L}(\mathsf{P}) = \left\lceil \frac{\log\left(\frac{\mathsf{P}}{\sigma^2}\,\frac{\varrho}{(1-\varrho)}\right)}{\log\left(\frac{1}{\varrho}\right)} \right\rceil \quad (83)$$

(where $\lceil a \rceil$ denotes the smallest integer that is greater than or equal to a) and $\tau = \mathsf{L}(\mathsf{P})$, we obtain from (19) the lower bound

$$C(\text{SNR}) \geq \frac{1}{2} \log \frac{\log \mathsf{P}}{\left\lceil \frac{\log\left(\frac{\mathsf{P}}{\sigma^2}\,\frac{\varrho}{(1-\varrho)}\right)}{\log\left(\frac{1}{\varrho}\right)} \right\rceil} + \frac{1}{2}\Upsilon, \quad \mathsf{P} > 1. \quad (84)$$

Taking the limit as the SNR (and, hence, also $\mathsf{P} = \sigma^2\text{SNR}$) tends to infinity yields

$$\lim_{\text{SNR}\to\infty} C(\text{SNR}) \geq \frac{1}{2} \log\log \frac{1}{\varrho} + \frac{1}{2}\Upsilon. \quad (85)$$

Since this holds for every $0 < \varrho < 1$, we have

$$\sup_{\text{SNR}>0} C(\text{SNR}) = \infty. \quad (86)$$

It remains to show that $\{\alpha_\ell, \ell \in \mathbb{N}_0\}$ and our choice of $\mathsf{L}(\mathsf{P})$ (83) satisfy conditions (17) and (18) of Proposition 5, namely

$$\sum_{\ell=0}^{\infty} \alpha_\ell < \infty \quad \text{and} \quad \sum_{\ell=\mathsf{L}(\mathsf{P})+1}^{\infty} \alpha_\ell\, \mathsf{P} \leq \sigma^2.$$

It follows immediately from (5) and (82) that $\{\alpha_\ell, \ell \in \mathbb{N}_0\}$ satisfies (17)

$$\begin{aligned} \sum_{\ell=0}^{\infty} \alpha_\ell &= \sum_{\ell=0}^{\ell_0-1} \alpha_\ell + \sum_{\ell=\ell_0}^{\infty} \alpha_\ell \\ &< \ell_0 \sup_{\ell\in\mathbb{N}_0} \alpha_\ell + \sum_{\ell=\ell_0}^{\infty} \varrho^\ell \\ &= \ell_0 \sup_{\ell\in\mathbb{N}_0} \alpha_\ell + \frac{\varrho^{\ell_0}}{1-\varrho} < \infty. \end{aligned} \quad (87)$$

In order to show that $\mathsf{L}(\mathsf{P})$ satisfies (18), we first note that by (82)

$$\sum_{\ell=\ell'+1}^{\infty} \alpha_\ell < \sum_{\ell=\ell'+1}^{\infty} \varrho^\ell = \varrho^{\ell'} \frac{\varrho}{1-\varrho}, \quad \ell' \geq \ell_0 - 1. \quad (88)$$

Since $\mathsf{L}(\mathsf{P})$ tends to infinity as $\mathsf{P} \to \infty$, it follows that $\mathsf{L}(\mathsf{P})$ is greater than $(\ell_0 - 1)$ for sufficiently large P. Furthermore, (83) implies

$$\varrho^{\mathsf{L}(\mathsf{P})} \frac{\varrho}{1-\varrho}\, \mathsf{P} \leq \sigma^2. \quad (89)$$

We therefore obtain from (88) and (89)

$$\sum_{\ell=\mathsf{L}(\mathsf{P})+1}^{\infty} \alpha_\ell\, \mathsf{P} < \varrho^{\mathsf{L}(\mathsf{P})} \frac{\varrho}{1-\varrho}\, \mathsf{P} \leq \sigma^2 \quad (90)$$

thus demonstrating that $\mathsf{L}(\mathsf{P})$ satisfies (18).

C. Pre-LogLog

We use Proposition 5 to prove Theorem 3. To this end, we first note that because the number of paths is finite, we have for some $\mathsf{L} \in \mathbb{N}_0$

$$\alpha_\ell = 0, \quad \ell > \mathsf{L} \tag{91}$$

which implies that

$$\sum_{\ell=0}^{\infty} \alpha_\ell = \sum_{\ell=0}^{\mathsf{L}} \alpha_\ell \leq (\mathsf{L}+1) \sup_{\ell \in \mathbb{N}_0} \alpha_\ell < \infty \tag{92}$$

and

$$\sum_{\ell=\mathsf{L}+1}^{\infty} \alpha_\ell \mathsf{P} = 0 \leq \sigma^2. \tag{93}$$

Consequently, (17) and (18) of Proposition 5 are satisfied, and it follows from (19) that the capacity is lower bounded by

$$C(\mathrm{SNR}) \geq \frac{\tau}{\mathsf{L}+\tau} \log\log \mathsf{P}^{1/\tau} + \frac{\tau}{\mathsf{L}+\tau}\Upsilon, \quad \mathsf{P} > 1. \tag{94}$$

Dividing by $\log\log \mathrm{SNR}$ and computing the limit as SNR tends to infinity (while holding L and τ fixed) yields

$$\varliminf_{\mathrm{SNR}\to\infty} \frac{C(\mathrm{SNR})}{\log\log \mathrm{SNR}} \geq \frac{\tau}{\mathsf{L}+\tau} \tag{95}$$

where we have used that for any fixed τ

$$\lim_{\mathrm{SNR}\to\infty} \frac{\log\log \mathsf{P}^{1/\tau}}{\log\log \mathrm{SNR}} = 1.$$

The lower bound on the capacity pre-loglog

$$\Lambda \triangleq \varlimsup_{\mathrm{SNR}\to\infty} \frac{C(\mathrm{SNR})}{\log\log \mathrm{SNR}} \geq \varliminf_{\mathrm{SNR}\to\infty} \frac{C(\mathrm{SNR})}{\log\log \mathrm{SNR}} \geq 1 \tag{96}$$

follows then by letting τ tend to infinity. Together with the upper bound $\Lambda \leq 1$, which was derived in Section V-B, this proves Theorem 3.

VII. Conclusion

We studied the high-SNR behavior of the capacity of noncoherent multipath fading channels. We demonstrated that, depending on the decay rate of the sequence $\{\alpha_\ell, \ell \in \mathbb{N}_0\}$, capacity may be bounded or unbounded in the SNR. We further showed that if the number of paths is finite, then at high SNR capacity grows double-logarithmically with the SNR, and the capacity pre-loglog is 1 irrespective of the number of paths. The picture that emerges is as follows:

- If $\{\alpha_\ell, \ell \in \mathbb{N}_0\}$ decays exponentially or slower, then capacity is bounded in the SNR.
- If $\{\alpha_\ell, \ell \in \mathbb{N}_0\}$ decays faster than exponentially, then capacity is unbounded in the SNR.
- If the number of paths is finite, then the capacity pre-loglog is equal to 1, irrespective of the number of paths.

The conclusions that can be drawn from these results are twofold. Firstly, multipath fading channels with an infinite number of paths and multipath fading channels with a finite number of paths have in general completely different capacity behaviors at high SNR. Indeed, at high SNR, if the number of paths is finite, then capacity grows double-logarithmically with the SNR, whereas if the number of paths is infinite, then capacity may even be bounded in the SNR. Thus, while for low or for moderate SNR it might be reasonable to approximate a multipath fading channel with infinitely many paths by a multipath fading channel with only a finite number paths, this is not reasonable when the SNR tends to infinity. The number of paths that are needed to approximate a multipath fading channel typically depends on the SNR and may grow to infinity as the SNR tends to infinity.

Secondly, the above results indicate that the high-SNR behavior of the capacity of multipath fading channels depends critically on the assumed channel model. Thus, when studying such channels at high SNR, the channel modeling is crucial, as slight changes in the channel model might lead to completely different capacity results.

Appendix

To prove (35) we lower bound

$$h\left(\sum_{\ell=0}^{k-1} H_k^{(\ell)} X_{k-\ell} + Z_k \,\middle|\, X_1^n = x_1^n, \mathbf{H}_1^{k-1} = \mathbf{h}_1^{k-1}\right)$$

for a given $\mathbf{h}_1^{k-1}$ and average the result over $\mathbf{H}_1^{k-1}$. Let

$$\mathcal{H}_k \triangleq \left\{H_k^{(\ell)}, \ell = 0, 1, \ldots, k-1 : \alpha_\ell = 0\right\}.$$

We have

$$\begin{aligned}
&h\left(\sum_{\ell=0}^{k-1} H_k^{(\ell)} X_{k-\ell} + Z_k \,\middle|\, X_1^n = x_1^n, \mathbf{H}_1^{k-1} = \mathbf{h}_1^{k-1}\right)\\
&\geq h\left(\sum_{\ell=0}^{k-1} H_k^{(\ell)} X_{k-\ell} + Z_k \,\middle|\, X_1^n = x_1^n, \mathbf{H}_1^{k-1} = \mathbf{h}_1^{k-1}, \mathcal{H}_k\right)\\
&= h\left(\sum_{\ell\in\mathcal{S}_k} H_k^{(\ell)} X_{k-\ell} + Z_k \,\middle|\, X_1^n = x_1^n, \mathbf{H}_1^{k-1} = \mathbf{h}_1^{k-1}, \mathcal{H}_k\right)\\
&= h\left(\sum_{\ell\in\mathcal{S}_k} H_k^{(\ell)} X_{k-\ell} + Z_k \,\middle|\, X_1^n = x_1^n, \mathbf{H}_1^{k-1} = \mathbf{h}_1^{k-1}\right)\\
&\geq \log\left(\sum_{\ell\in\mathcal{S}_k} e^{h\left(H_k^{(\ell)} X_{k-\ell} \middle| X_{k-\ell} = x_{k-\ell}, \mathbf{H}_{\ell,1}^{k-1} = \mathbf{h}_{\ell,1}^{k-1}\right)} + e^{h(Z_k)}\right)
\end{aligned} \tag{97}$$

where $\mathcal{S}_k$ is defined in (34). Here the first step follows because conditioning cannot increase differential entropy; the second step follows because differential entropy is invariant under deterministic translation [4, Th. 9.6.3] and because terms for which we have $x_{k-\ell} = 0$ do not contribute to the sum; the third step follows from the "uncorrelated scattering" assumption, i.e., from the assumption that the processes

$$\left\{H_k^{(0)}, k \in \mathbb{Z}\right\}, \left\{H_k^{(1)}, k \in \mathbb{Z}\right\}, \ldots$$

are independent, which implies that, conditioned on $(X_1^n, \mathbf{H}_1^{k-1}) = (x_1^n, \mathbf{h}_1^{k-1})$, the path gains $\left\{H_k^{(\ell)}, \ell \in \mathcal{S}_k\right\}$ do not depend on $\mathcal{H}_k$; and the last step follows from the entropy power inequality [4, Th. 16.6.3] and from the "uncorrelated scattering" assumption which implies that $H_k^{(\ell)}$ depends only on $\mathbf{H}_{\ell,1}^{k-1}$ and not on $\mathbf{H}_1^{k-1}$. (Note that, for a given $\mathbf{H}_1^{k-1} = \mathbf{h}_1^{k-1}$, the conditional entropies on the RHS of (97) are possibly $-\infty$. However, by (6) this event is of zero probability and has therefore no impact on (97) when averaged over $\mathbf{H}_1^{k-1}$.)

Since the processes

$$\{H_k^{(0)}, k \in \mathbb{Z}\}, \{H_k^{(1)}, k \in \mathbb{Z}\}, \dots$$

are independent and jointly independent of X_1^n, we can compute the expectation of (97) over $\mathbf{H}_1^{k-1}$ by averaging (97) first over $\mathbf{H}_{0,1}^{k-1}$, then averaging the result over $\mathbf{H}_{1,1}^{k-1}$, and so on. To lower bound the individual expectations, we note that the function

$$x \mapsto \log\left(e^x + \zeta\right)$$

is convex for every $\zeta > 0$. Let $\zeta_{\ell'}, \ell' = 0, 1, \dots, k-1$ be defined as

$$\zeta_{\ell'} \triangleq \sum_{\substack{\ell \in \mathcal{S}_k:\\ \ell < \ell'}} e^{h\left(H_k^{(\ell)} X_{k-\ell} \middle| X_{k-\ell}=x_{k-\ell}, \mathbf{H}_{\ell,1}^{k-1}\right)} + \sum_{\substack{\ell \in \mathcal{S}_k:\\ \ell > \ell'}} e^{h\left(H_k^{(\ell)} X_{k-\ell} \middle| X_{k-\ell}=x_{k-\ell}, \mathbf{H}_{\ell,1}^{k-1}=\mathbf{h}_{\ell,1}^{k-1}\right)} + e^{h(Z_k)}. \quad (98)$$

Then it follows from Jensen's inequality that for every $\ell' \in \mathcal{S}_k$ the conditional expectation, conditioned on $\mathbf{H}_{\ell,1}^{k-1} = \mathbf{h}_{\ell,1}^{k-1}, \ell > \ell'$, is lower bounded by

$$\mathsf{E}\left[\log\left(e^{h\left(H_k^{(\ell')} X_{k-\ell'} \middle| X_{k-\ell'}=x_{k-\ell'}, \mathbf{H}_{\ell',1}^{k-1}=\mathbf{h}_{\ell',1}^{k-1}\right)} + \zeta_{\ell'}\right)\right] \geq \log\left(e^{h\left(H_k^{(\ell')} X_{k-\ell'} \middle| X_{k-\ell'}=x_{k-\ell'}, \mathbf{H}_{\ell',1}^{k-1}\right)} + \zeta_{\ell'}\right). \quad (99)$$

Averaging (97) over $\mathbf{H}_1^{k-1}$, and employing (99) to compute this average, yields the desired lower bound (35)

$$h\left(\sum_{\ell=0}^{k-1} H_k^{(\ell)} X_{k-\ell} + Z_k \middle| X_1^n = x_1^n, \mathbf{H}_1^{k-1}\right) \geq \log\left(\sum_{\ell \in \mathcal{S}_k} e^{h\left(H_k^{(\ell)} X_{k-\ell} \middle| X_{k-\ell}=x_{k-\ell}, \mathbf{H}_{\ell,1}^{k-1}\right)} + e^{h(Z_k)}\right). \quad (100)$$

ACKNOWLEDGMENT

Fruitful discussions with H. Bölcskei and G. Durisi are gratefully acknowledged. The authors also wish to thank O. Leveque and N. Jindal for their comments, which motivated the proof of Proposition 5.

REFERENCES

[1] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.
[2] A. Lapidoth and S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2426–2467, Oct. 2003.
[3] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 623–656, Oct. 1948.
[4] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, First ed. New York: Wiley, 1991.
[5] Y.-H. Kim, "A coding theorem for a class of stationary channels with feedback," *IEEE Trans. Inf. Theory*, vol. 54, no. 4, pp. 1488–1499, Apr. 2008.
[6] T. H. E. Ericson, "A Gaussian channel with slow fading," *IEEE Trans. Inf. Theory*, vol. 16, no. 5, pp. 353–355, May 1970.
[7] T. Koch, A. Lapidoth, and P. P. Sotiriadis, "Channels that heat up," *IEEE Trans. Inf. Theory*, vol. 55, pp. 3594–3612, Aug. 2009.
[8] F. Topsøe, "An information theoretical identity and a problem involving capacity," *Studia Scientarum Math. Hungarica*, vol. 2, pp. 291–292, 1967.
[9] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.
[10] A. Lapidoth, "On the high-SNR capacity of noncoherent networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3025–3036, Sep. 2005.

Tobias Koch (S'02– M'09) received the M.Sc. degree in electrical engineering (with distinction) and the Ph.D. degree in electrical engineering from ETH Zurich, Switzerland, in 2004 and 2009, respectively.

He is currently a Marie Curie Intra-European Research Fellow with the University of Cambridge, Cambridge, U.K. His research interests include digital communication theory and information theory with particular focus on wireless communication and communication in electronic circuits.

Amos Lapidoth (S'89–M'95–SM'00–F'04) received the B.A. degree in mathematics (*summa cum laude*) in 1986, the B.Sc. degree in electrical engineering (*summa cum laude*) in 1986, and the M.Sc. degree in electrical engineering in 1990, all from the Technion—Israel Institute of Technology, Haifa, Israel, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1995.

During 1995–1999, he was an Assistant and Associate Professor with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, and was the KDD Career Development Associate Professor in Communications and Technology. He is now a Professor of information theory with ETH Zurich, Switzerland. He is the author of the book *A Foundation in Digital Communication* (Cambridge Univ. Press, 2009). His research interests are in digital communications and information theory.

Dr. Lapidoth served during 2003–2004 and 2009 as an Associate Editor for Shannon Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY.