

Wireless Networks of Bounded Capacity

Grace Villacrés and Tobias Koch

Abstract

We consider a noncoherent wireless network, where the transmitters and receivers are cognizant of the statistics of the fading coefficients, but are ignorant of their realizations. We demonstrate that if the nodes do not cooperate, if they transmit symbols that all follow the same distribution, and if the variances of the fading coefficients decay exponentially or more slowly, then the channel capacity is bounded in the SNR. This confirms a similar result by Lozano, Heath, and Andrews.

I. INTRODUCTION

THE information-theoretical limits of wireless networks have been studied extensively in the past and it has been demonstrated that cooperation between nodes in wireless networks can substantially increase the channel capacity. Indeed, theoretical findings suggest that throughput gains of up to two orders of magnitude are obtainable by exploiting techniques such as Coordinated Multi-Point (CoMP), network Multiple Input Multiple Output (MIMO) or Interference Alignment (IA), provided that the communicating nodes have access to the fading coefficients in the network [1]–[3]. Knowledge of these fading coefficients is usually referred to as *channel-state information (CSI)*, and the assumption that the nodes have perfect knowledge of the fading coefficients is usually referred to as *perfect CSI*. However, it is *prima facie* unclear whether perfect CSI can be obtained in practical systems. In fact, over-the-air trials have only demonstrated disappointingly low throughput improvements (not exceeding 30% in some scenarios) [4].

The purpose of this work is to investigate the channel capacity of wireless networks when the nodes neither have perfect CSI nor do they perform a channel estimation to obtain information on the fading coefficients. This is relevant, e.g., if the nodes do not have the computational resources to perform an accurate channel estimation. Moreover, in dense wireless networks, where many nodes share the same resources, an accurate channel estimation may not be feasible, even if the nodes have sufficient computing power.

Our work is along the lines of the work by Lozano, Heath, and Andrews [5]. Indeed, Lozano et al. argued that cooperation can have severe limitations in the sense that the channel capacity for wireless networks can be bounded in the signal-to-noise (SNR). In other words, cooperation cannot convert an interference-limited network into a noise-limited one. The main results in [5] are based on the analysis of a block-fading channel that models the channel within a cluster and takes out-of-cluster interference into account. Specifically, in [5] two analyses are carried out:

- The analysis of a block-fading channel in a clustered system with pilot-assisted channel estimation and out-of-cluster interference that has a Gaussian distribution and grows linearly with the SNR. Due to this dependence of the interference on the SNR, the signal-to-interference-plus-noise ratio (SINR) is bounded in the SNR, resulting in the capacity being bounded in the SNR, too.
- The analysis of a fully cooperative system, where all transmitters and all receivers cooperate, resulting effectively in MIMO transmission. It is assumed that the number of transmitters is greater than the number of time instants L over which the block-fading channel stays constant. This precludes an accurate channel estimation. For this scenario, the maximum achievable rate is studied when the time- k channel input is of the form $\sqrt{\text{SNR}}X_k$ (where the distribution of X_k does not depend on the SNR), and it is demonstrated that this rate is bounded in the SNR.

For both analyses the conclusions are the same: in the absence of perfect CSI, the capacity of wireless networks is bounded in the SNR, hence, transmission over such networks is highly power-inefficient. This accentuates the limitation of cooperation in wireless networks.

However, the analyses by Lozano et al. are based on simplifying assumptions that weaken their conclusions. Indeed, in the former analysis of a clustered system with pilot-assisted channel estimation, the authors assume that the interference has a Gaussian distribution, which seems pessimistic. Indeed, it is well-known that the capacity of the additive-noise channel is minimized, among all power-constrained noise distributions, by *identically and independently distributed (i.i.d.)* Gaussian noise [6], [7]. This result has been generalized to wireless networks [8]. Note, however, that Gaussian noise does not constitute the worst-case scenario when the input distribution is not Gaussian; see, e.g., [9]. Based on the above observations, it may seem plausible that also for the considered channel model, Gaussian interference is the worst-case interference. Since the distribution

The authors are with the Signal Theory and Communications Department, Universidad Carlos III de Madrid, 28911, Leganés, Spain and with the Gregorio Marañón Health Research Institute (e-mail: {gvillacres, koch}@tsc.uc3m.es).

This work has been partly supported by the Ministerio de Economía of Spain (projects COMONSENS, id. CSD2008-00010 and OTOSiS, id. TEC2013-41718-R), the Comunidad de Madrid (project CASI-CAM- CM, id. S2013/ICE-2845) and the European Community's Seventh Framework Programme (FP7/2007-2013) under Grant Nr. 333680.

of the interference depends on the distribution of the symbols transmitted by the interfering nodes, one may argue that the capacity may be unbounded if the interferers would use codebooks that give rise to non-Gaussian interference.

As for the latter analysis of a fully cooperative system, here the authors assume that the time- k channel inputs are of the form $\sqrt{\text{SNR}}X_k$, where X_k has a distribution that does not depend on the SNR. One may argue that the corresponding information rate is bounded in the SNR because of the suboptimal input distribution and not because of the limitations of cooperation. In fact, it has been demonstrated by Lapidot and Moser in [10, Th. 4.3] that for a memoryless channel and in the absence of perfect CSI such inputs give rise to a bounded information rate also in the point-to-point case. For noncoherent point-to-point memoryless fading channels, an input distribution that changes with the SNR is thus necessary in order to achieve an unbounded information rate.¹ Since the block-fading channel specializes to the memoryless fading channel when $L = 1$, the observation that channel inputs of the form $\sqrt{\text{SNR}}X_k$ yield a bounded information rate is perhaps not very surprising.

In this work, we explore whether the capacity of noncoherent wireless networks is still bounded in the SNR if we allow the input distribution to depend on the SNR. In contrast to the analysis by Lozano et al. [5], we assume for the sake of simplicity that the nodes do not cooperate. We do not know whether cooperation would give rise to an unbounded capacity. For simplicity, we further consider a flat-fading channel with an infinite number of interferers. The locations of these interferers enter the channel model through the variance of the fading coefficients corresponding to the paths between the interferers and the intended receiver. Without loss of generality, we order the interferers with respect to the variances of the corresponding fading coefficients: the fading coefficient of the first interferer has the largest variance, denoted by α_1 , the fading coefficient of the second interferer has the second-largest variance, denoted by α_2 , and so on. We consider a noncoherent scenario where transmitter and receiver are cognizant of the statistics of the fading coefficients, but are ignorant of their realization. We demonstrate that the observation by Lozano et al. continues to hold even if the input distribution is allowed to depend on the SNR, provided that the sequence of variances $\{\alpha_\ell\}$ decays at most exponentially and each interfering node transmits symbols that follow the same distribution as the symbols transmitted by the transmitting node. To this end, we derive an upper bound on the channel capacity that is bounded in the SNR.

The rest of this paper is organized as follows. Section II describes the channel model. Section III is devoted to channel capacity and the main result of our work. Section IV shows an upper bound on the channel capacity for the special case in which the variances of the channel coefficients decay exponentially. Section V contains the proof of our main result. Section VI concludes the paper with a summary and discussion of the results.

II. CHANNEL MODEL

A network is composed of a number of users that are communicating with each other. A possible scenario is depicted in Figure 1a, where n transmitting nodes (denoted in the figure by “A”) communicate with m receiving nodes (denoted by “B”). This network is sometimes referred to as an *interference network*. For simplicity, we assume that the set of transmitting nodes and the set of receiving nodes are disjoint and both the transmitting nodes and the receiving nodes do not cooperate.

Since a characterization of all achievable rates in the network is unfeasible when the number of nodes is large, it is common to study the *sum-rate capacity* of the network, which is defined as the supremum of the sum of all achievable rates in the network. However, it is *prima facie* unclear whether the transmission strategy that achieves the sum-rate capacity will also be optimal in practice. Indeed, it may well be that the optimal transmission strategy consists of turning off all but one of the transmitting nodes, thereby minimizing the interference. Such a strategy prevents $n - 1$ nodes from transmitting anything and is probably not very practical. In fact, practical constraints may demand that each node is offered roughly the same transmission rate or, at least, a transmission rate that is above a given threshold. In order to enforce such a solution, one could study the sum-rate capacity of the network under the constraint that all transmitting nodes transmit at the same rate, but obtaining an expression for such a capacity seems challenging. To model the problem in a way that it can be analyzed analytically, we simplify the original problem as follows:

Firstly, we consider the case where one transmitting node communicates with one receiving node and the interfering nodes emit symbols that interfere with this communication; see Figure 1(b). To model a large network, we assume that there are an infinite number of interfering nodes. As performance measure we consider the capacity of the channel between the transmitting and receiving node. Secondly, to avoid transmission strategies for which the interfering nodes are turned off (which would, in fact, maximize the capacity), we assume that the symbols transmitted by the interfering nodes must follow the same distribution as the symbols transmitted by the transmitting node. By the channel coding theorem [11], this implies that each node (transmitting and interfering) is transmitting at the same rate, while at the same time it keeps the analysis accessible.

Note that, while the above simplifications permit a mathematical analysis of the channel capacity of the network, they are more restrictive than the analysis of the sum-rate capacity under the constraint that all nodes transmit at the same rate. Indeed, one can find transmission strategies for which each node transmits at the same rate, but the symbols produced by them do not follow the same distribution. One such strategy would be time-division multiple access (TDMA), where only one pair of nodes communicates with each other and the other nodes remain silent, and where the pairs of communicating nodes change after a given period until all nodes have transmitted their data. Clearly, in this case a node that transmits data produces symbols that

¹However, in contrast to the case of perfect CSI, in its absence the capacity only grows double-logarithmically with the SNR [10, Th. 4.2].

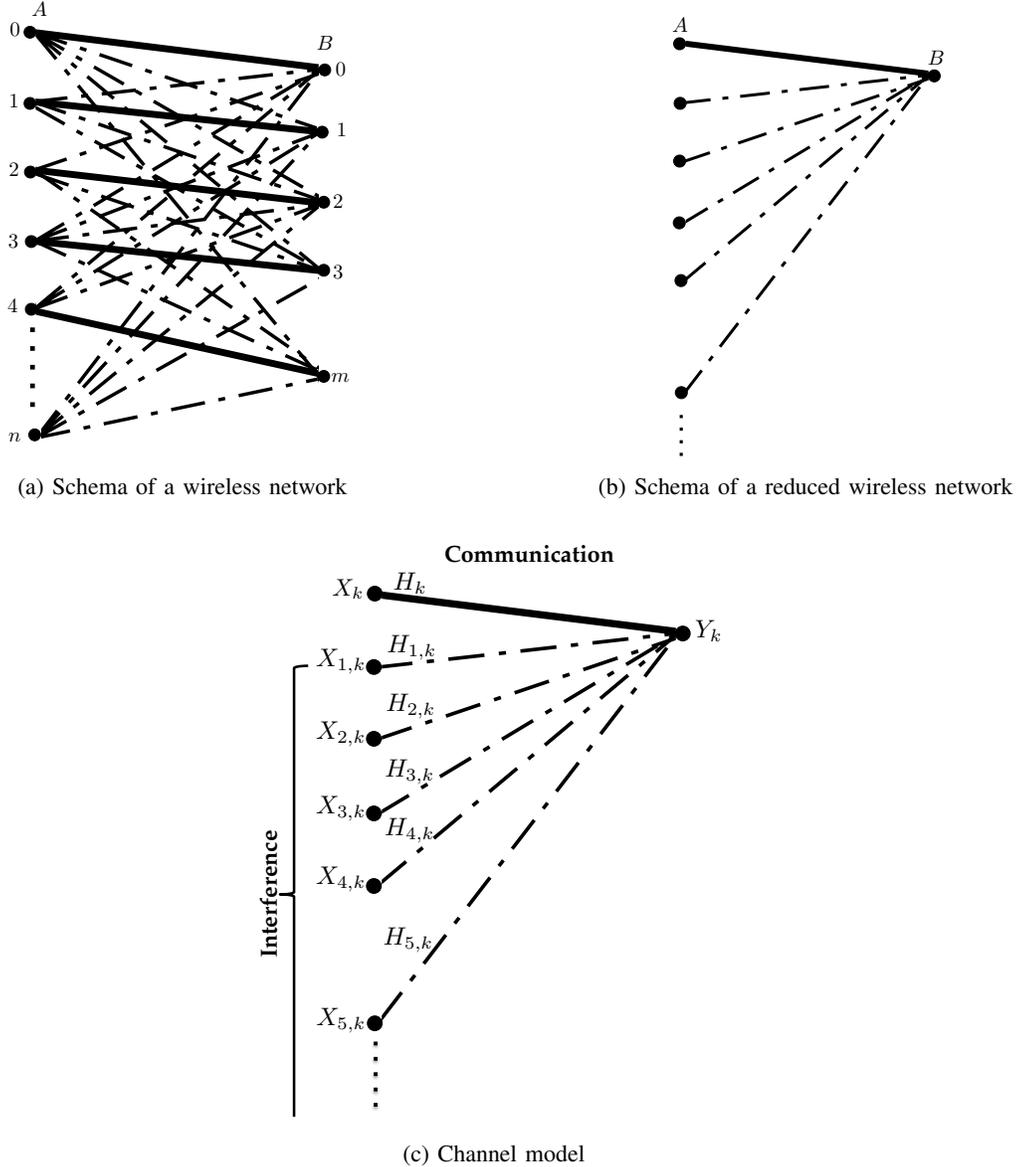


Figure 1: Wireless Networks

follow a different distribution than the symbols emitted by silent nodes (for which the symbol transmitted at time k is equal to zero). Thus, our constraint that all nodes shall produce symbols that follow the same distribution precludes, e.g., strategies such as TDMA. That said, the rates achievable with TDMA vanish as the number of users in the network tends to infinity, so in a large network TDMA seems not very attractive anyway.

We model the channel between the transmitting and receiving node by a discrete-time flat-fading channel whose complex-valued output Y_k at time $k \in \mathbb{Z}$ (where \mathbb{Z} denotes the set of integers) corresponding to the time- k channel input X_k and the time- k interfering symbols $X_{\ell,k}$, $\ell = 1, 2, \dots$ is given by

$$Y_k = H_k X_k + \sum_{\ell=1}^{\infty} H_{\ell,k} X_{\ell,k} + Z_k, \quad k \in \mathbb{Z}. \quad (1)$$

Here Z_k models the time- k additive noise; H_k denotes the time- k fading coefficient of the channel between the transmitter and receiver; and $H_{\ell,k}$, $\ell = 1, 2, \dots$ denotes the time- k fading coefficient of the link between the ℓ -th interfering node and the receiver; see Figure 1(c). We assume that the sequences $\{Z_k, k \in \mathbb{Z}\}$, $\{H_k, k \in \mathbb{Z}\}$, and $\{H_{\ell,k}, k \in \mathbb{Z}\}$, $\ell = 1, 2, \dots$ are all sequences of i.i.d. complex random variables that they are independent of each other. We further assume that $Z_k \sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2)$, $H_k \sim \mathcal{N}_{\mathbb{C}}(0, 1)$ and $H_{\ell,k} \sim \mathcal{N}_{\mathbb{C}}(0, \alpha_{\ell})$ for some $\alpha_{\ell} > 0$, where we use the notation $U \sim \mathcal{N}_{\mathbb{C}}(\mu, \sigma^2)$ to indicate that U is a circularly-symmetric, complex Gaussian random variable of mean μ and variance σ^2 (\mathbb{C} denotes the set of complex numbers).

We consider a noncoherent scenario where transmitter and receiver are cognizant of the statistics of the fading coefficients, but are ignorant of their realization.

We assume that the interfering nodes do neither cooperate with each other nor with the transmitting node. This implies that the sequences $\{X_k, k \in \mathbb{Z}\}$ and $\{X_{\ell,k}, k \in \mathbb{Z}\}$, $\ell = 1, 2, \dots$ are independent of each other. Since we also demand that each node produces symbols according to the same distribution, this implies that $\{X_k, k \in \mathbb{Z}\}$ and $\{X_{\ell,k}, k \in \mathbb{Z}\}$, $\ell = 1, 2, \dots$ are also identically distributed. Note, however, that this does not mean that the sequences of transmitted symbols are i.i.d. in time—to compute the capacity we will optimize over all distributions of $\{X_k, k \in \mathbb{Z}\}$ that satisfy the power constraint.

Without loss of generality, we assume that the interfering nodes are ordered according to the variances of the corresponding fading coefficients, i.e., $\alpha_\ell \geq \alpha_{\ell'}$ for any $\ell < \ell'$. We further assume that there exists a $0 < \rho < 1$ such that

$$\frac{\alpha_{\ell+1}}{\alpha_\ell} \geq \rho, \quad \ell = 1, 2, \dots \quad (2)$$

The condition (2) states that the variances of the fading coefficients corresponding to the interfering nodes decay at most exponentially. If we relate these variances to the path losses between the receiving node and the interfering nodes, then (2) essentially requires that the distance between neighboring interfering nodes grows at most exponentially as we move farther away from the receiving node. We believe that this assumption is reasonably mild.

III. CHANNEL CAPACITY AND MAIN RESULT

We shall denote sequences such as $A_n, A_{n+1}, A_{n+2}, \dots, A_m$ by A_n^m . We further denote by $\overline{\lim}$ the limit superior. We define the capacity of the above channel (1) as²

$$C(\text{SNR}) \triangleq \overline{\lim}_{n \rightarrow \infty} \frac{1}{n} \sup_{Q^n} I(X_1^n; Y_1^n) \quad (3)$$

where we assume that the sequences X_1^n and $X_{\ell,1}^n$, $\ell = 1, 2, \dots$ are independent of each other and that each such sequence has distribution Q^n . This is consistent with the simplifying assumptions introduced in Section II. The supremum in (3) is over all n -dimensional, complex-valued, probability distributions Q^n satisfying the power constraint

$$\int \frac{1}{n} \sum_{k=1}^n |x_k|^2 dQ^n(x_1^n) \leq P. \quad (4)$$

The SNR is defined as

$$\text{SNR} \triangleq \frac{P}{\sigma^2}. \quad (5)$$

We do not claim that there is a coding theorem associated with (3), i.e., we do not claim that for any rate below $C(\text{SNR})$ there exists an encoding and decoding scheme for which the decoding error probability tends to zero as n tends to infinity. (See [11] for more details.) However, by Fano's inequality [11, Sec. 7.9], we do know that for any rate above $C(\text{SNR})$ there exists no such encoding and decoding scheme. Thus, by demonstrating that $C(\text{SNR})$ is bounded in the SNR, we also demonstrate that there exists no encoding and decoding scheme that has a rate that tends to infinity as $\text{SNR} \rightarrow \infty$ and for which the decoding error probability vanishes as n tends to infinity.

Our main result is the following.

Theorem 1 (Main Result). *For the channel model presented in Section II, the capacity is bounded in the SNR, i.e.,*

$$\sup_{\text{SNR} > 0} C(\text{SNR}) < \infty. \quad (6)$$

Proof: See Section V. ■

Thus, Theorem 1 demonstrates that if the nodes do not cooperate, if they emit symbols that all follow the same distribution, and if the variances of the fading coefficients satisfy (2), then the capacity is bounded in the SNR, even if the distribution of the channel inputs is allowed to depend on the SNR. This result is more general than the result by Lozano et al. [5] in the sense that we allow the distribution of the channel inputs to depend on the SNR, but it is less general than [5] in the sense that we assume that the variances of the fading coefficients, i.e., $\{\alpha_\ell\}$, decay at most exponentially and that we do not allow cooperation between the nodes.

The condition (2) is satisfied, for example, if

$$\alpha_\ell = \rho^\ell, \quad \text{for some } 0 < \rho < 1. \quad (7)$$

In general, we can say that if $\{\alpha_\ell\}$ decays *exponentially or more slowly*, then $C(\text{SNR})$ is bounded in the SNR. This result is reminiscent of a result obtained by Koch and Lapidoth [12] that states that the capacity of single-user, frequency-selective,

²The logarithms used in this paper are natural logarithms. The capacity has thus the dimension ‘‘nats per channel use’’.

noncoherent fading channels is bounded in the SNR if the variances of the path gains decay exponentially or more slowly. Note, however, that the proof given in Section V is different from the one given in [12].

In order to prove the above result, we derive an upper bound on the capacity that does not depend on the SNR. Since this bound may also be of independent interest, we summarize it in the following proposition.

Proposition 1 (Upper Bound). *Consider the channel model introduced in Section II. The capacity is upper-bounded by*

$$C(\text{SNR}) \leq \log \frac{\pi}{2} + 2 \left(\frac{\eta_{\max}^2}{\eta_{\min}^2} - \frac{1}{2} \right) + \log \eta_{\max}^2 + \frac{1}{2} \log \frac{\eta_{\max}^2}{\eta_{\min}^2}, \quad \text{SNR} > 0. \quad (8)$$

where η_{\min}^2 and η_{\max}^2 are defined as

$$\eta_{\min}^2 \triangleq \min \left(1, \frac{\alpha_0}{\alpha_1} \right) \quad (9)$$

$$\eta_{\max}^2 \triangleq \max \left(\frac{\alpha_0}{\alpha_1}, \frac{1}{\rho} \right). \quad (10)$$

Proof: Follows directly from the upper bound derived in the proof of Theorem 1; cf. (58) and (56). ■

Assuming that $\{\alpha_\ell\}$ satisfies (7), we have $\eta_{\min}^2 = 1$ and $\eta_{\max}^2 = \frac{1}{\rho}$. In this, the capacity is thus upper-bounded by

$$C(\text{SNR}) \leq \log \frac{\pi}{2} + 2 \left(\frac{1}{\rho} - \frac{1}{2} \right) + \frac{3}{2} \log \frac{1}{\rho}, \quad \text{SNR} > 0. \quad (11)$$

IV. EXPONENTIAL PATH LOSS

In the special case where the variances of the fading coefficients satisfy (7), i.e., where they decay exponentially, we obtain an upper bound that is tighter than the upper bound in (11). Indeed, it can be shown that in this case the capacity is upper-bounded by

$$C(\text{SNR}) \leq \frac{1}{\rho}, \quad \text{SNR} > 0. \quad (12)$$

As we shall see, the proof of (12) is much simpler than the proof of Theorem 1.

To obtain (12), we begin by deriving an upper bound on the mutual information $I(X_1^n; Y_1^n)$ as follows. Since conditioning reduces entropy [11, Th. 8.6.1], we have

$$\begin{aligned} I(X_1^n; Y_1^n) &= h(Y_1^n) - h(Y_1^n | X_1^n) \\ &\leq h(Y_1^n) - h(Y_1^n | H_1^n, X_1^n) \\ &= h(Y_1^n) - h(Y_1^n - H_1^n X_1^n) \end{aligned} \quad (13)$$

where $Y_1^n - H_1^n X_1^n$ refers to sequence of symbols $Y_1 - H_1 X_1, \dots, Y_n - H_n X_n$. In the last step of (13), we use that the symbols emitted by the nodes are independent, so X_1^n and $X_{\ell,1}^n$, $\ell = 1, 2, \dots$ are independent.

We next define a new channel model, for which the channel output \bar{Y}_k at time $k \in \mathbb{Z}$ corresponding to the time- k channel inputs $X_{\ell,1}, \dots, X_{\ell,k}$ is given by

$$\bar{Y}_k = \sum_{\ell=0}^{\infty} H_{\ell+1,k} X_{\ell,k} + Z_k, \quad k \in \mathbb{Z}. \quad (14)$$

Observe that $(H_{\ell+1,1}, \dots, H_{\ell+1,n})$ can be obtained from $(H_{\ell,1}, \dots, H_{\ell,n})$ by changing ℓ to $\ell + 1$. It thus follows from (7) and the independence of the fading processes that $(H_{\ell+1,1}, \dots, H_{\ell+1,n})$, $\ell = 0, 1, \dots$ has the same joint distribution as $(\sqrt{\rho}H_{\ell,1}, \dots, \sqrt{\rho}H_{\ell,n})$, $\ell = 0, 1, \dots$. Similarly, the additive Gaussian noise (Z_1, \dots, Z_n) has the same distribution as $(\sqrt{\rho}Z_1^{(1)} + \sqrt{1-\rho}Z_1^{(2)}, \dots, \sqrt{\rho}Z_n^{(1)} + \sqrt{1-\rho}Z_n^{(2)})$, where $Z_1^{(1)}, \dots, Z_n^{(1)}, Z_1^{(2)}, \dots, Z_n^{(2)}$ are i.i.d., circularly-symmetric, complex Gaussian random variable of zero-mean and variance σ^2 . It follows by comparing (14) with (1) that

$$\bar{Y}_1^n \stackrel{\mathcal{L}}{=} \sqrt{\rho}Y_1^n + \sqrt{1-\rho}Z_1^{n(2)} \quad (15)$$

where we use $\stackrel{\mathcal{L}}{=}$ to denote equivalence in the probability law. With this, we can compute the entropy of the second term in (13) as

$$\begin{aligned} h(\bar{Y}_1^n) &= h \left(\sqrt{\rho}Y_1^n + \sqrt{1-\rho}Z_1^{n(2)} \right) \\ &\geq h \left(\sqrt{\rho}Y_1^n + \sqrt{1-\rho}Z_1^{n(2)} | Z_1^{n(2)} \right) \\ &= n \log \rho + h(Y_1^n) \end{aligned} \quad (16)$$

where the second step follows from the behavior of differential entropy under translation [11, Th.8.6.3] and scaling [11, Th.8.6.4], and because conditioning reduces entropy; the last step follows because Y_1^n is independent of $Z_1^{n(2)}$.

Replacing (16) in (13), we obtain

$$\begin{aligned} \frac{1}{n}I(X_1^n; Y_1^n) &\leq \frac{1}{n}(h(Y_1^n) - n \log \rho - h(Y_1^n)) \\ &= \log \frac{1}{\rho}. \end{aligned} \quad (17)$$

This proves (12).

V. PROOF OF THEOREM 1

To prove Theorem 1, we derive an upper bound on the channel capacity that is independent of the SNR. We begin with the same steps as in Section IV to obtain (13). Thus, we derive an upper bound on the mutual information $I(X_1^n; Y_1^n)$ as follows

$$\begin{aligned} I(X_1^n; Y_1^n) &= h(Y_1^n) - h(Y_1^n | X_1^n) \\ &\leq h(Y_1^n) - h(Y_1^n | H_1^n, X_1^n) \\ &= h(Y_1^n) - h(Y_1^n - H_1^n X_1^n). \end{aligned} \quad (18)$$

For the first entropy on the right-hand side (RHS) of (18) we rewrite the channel model in (1) as

$$Y_k = \sum_{\ell=0}^{\infty} H_{\ell,k} X_{\ell,k} + Z_k, \quad k \in \mathbb{Z} \quad (19)$$

where we define $H_{0,k} \triangleq H_k$. For the second entropy in (18) we define again a new channel model, as we have done in the last section. Specifically, the time- k channel output \tilde{Y}_k of the new channel corresponding to the time- k channel inputs $X_{\ell,1}, \dots, X_{\ell,k}$ is given by

$$\tilde{Y}_k = \sum_{\ell=0}^{\infty} \tilde{H}_{\ell+1,k} X_{\ell,k} + \tilde{Z}_k, \quad k \in \mathbb{Z} \quad (20)$$

where the fading coefficients $\{\tilde{H}_{\ell,k}\}$ have the same distribution as $\{H_{\ell,k}\}$ but are independent of $\{H_{\ell,k}\}$ and, likewise, the additive-noise terms $\{\tilde{Z}_k\}$ have the same distribution as $\{Z_k\}$ but are independent of $\{Z_k\}$. It follows that the sequences $\{\tilde{Z}_k\}$ and $\{\tilde{H}_{\ell,k}\}$ are i.i.d. and independent of each other, with $\tilde{Z}_k \sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2)$ and $\tilde{H}_{\ell,k} \sim \mathcal{N}_{\mathbb{C}}(0, \alpha_{\ell})$. Since $\{X_k\}$ and $\{X_{\ell,k}\}$, $\ell = 1, 2, \dots$ have the same distribution, it follows that \tilde{Y}_1^n has the same distribution as $Y_1^n - H_1^n X_1^n$, hence $h(Y_1^n - H_1^n X_1^n) = h(\tilde{Y}_1^n)$.

To find an upper bound on (18) we use the identity $h(A) - h(B) = h(A|B) - h(B|A)$ to obtain

$$\begin{aligned} h(Y_1^n) - h(\tilde{Y}_1^n) &= h(Y_1^n | \tilde{Y}_1^n) - h(\tilde{Y}_1^n | Y_1^n) \\ &\leq \sum_{k=1}^n h(Y_k | \tilde{Y}_k) - \sum_{k=1}^n h(\tilde{Y}_k | \tilde{Y}_1^{k-1}, Y_1^n) \\ &= \sum_{k=1}^n \left(h\left(\tilde{Y}_k \frac{Y_k}{\tilde{Y}_k} \middle| \tilde{Y}_k\right) - h(\tilde{Y}_k | \tilde{Y}_1^{k-1}, Y_1^n) \right) \\ &\leq \sum_{k=1}^n \left(\mathbb{E} \left[\log |\tilde{Y}_k|^2 \right] + h\left(\frac{Y_k}{\tilde{Y}_k}\right) - h(\tilde{Y}_k | \tilde{Y}_1^{k-1}, Y_1^n) \right) \end{aligned} \quad (21)$$

where the second step follows from the chain rule for entropy and because conditioning reduces entropy; the third step follows because if we multiply and divide by \tilde{Y}_k we do not change the conditional differential entropy; the last inequality follows from the behavior of differential entropy under scaling by a complex random variable and because conditioning reduces entropy.

In order to find an upper bound on $h(Y_1^n) - h(\tilde{Y}_1^n)$, we calculate bounds for each term in (21). For the first term we have

$$\begin{aligned} \mathbb{E} \left[\log |\tilde{Y}_k|^2 \right] &= \mathbb{E} \left[\mathbb{E} \left[\log \left(|\tilde{Y}_k|^2 \right) \middle| \{X_{\ell,k}\} \right] \right] \\ &\leq \mathbb{E} \left[\log \left(\mathbb{E} \left[|\tilde{Y}_k|^2 \middle| \{X_{\ell,k}\} \right] \right) \right] \end{aligned} \quad (22)$$

where last step follows from Jensen's inequality.

For the third term on the RHS of (21), we use that conditioning reduces entropy to obtain

$$\begin{aligned} h\left(\tilde{Y}_k|\tilde{Y}_1^{k-1}, Y_1^n\right) &\geq h\left(\tilde{Y}_k|\tilde{Y}_1^{k-1}, Y_1^n, \{X_{\ell,k}\}\right) \\ &= h\left(\tilde{Y}_k|\{X_{\ell,k}\}\right) \\ &= \log(\pi e) + \mathbf{E}\left[\log\left(\mathbf{E}\left[|\tilde{Y}_k|^2|\{X_{\ell,k}\}\right]\right)\right] \end{aligned} \quad (23)$$

where the first equality follows because $\{\tilde{Z}_k\}$ and $\{\tilde{H}_{\ell,k}\}$ are i.i.d. and independent of $(\{H_{\ell,k}\}, \{Z_k\})$, so conditioned on $\{X_{\ell,k}\}$, \tilde{Y}_k is independent of $(\tilde{Y}_1^{k-1}, Y_1^n)$; the second equality follows because, conditioned on the inputs $\{X_{\ell,k}\}$, \tilde{Y}_k has a Gaussian distribution.

Now we bound the second term in (21). To this end, we first note that, conditioned on the inputs $\{X_{\ell,k}\}$, the ratio $\frac{Y_k}{\tilde{Y}_k}$ is a ratio of independent, circularly-symmetric complex random variables. This implies that $\frac{Y_k}{\tilde{Y}_k}$ is also circularly symmetric unconditioned on $\{X_{\ell,k}\}$. Indeed, express Y_k and \tilde{Y}_k as

$$Y_k = |Y_k|e^{j\phi_Y} \quad (24)$$

$$\tilde{Y}_k = |\tilde{Y}_k|e^{j\phi_{\tilde{Y}}} \quad (25)$$

for some $\phi_Y, \phi_{\tilde{Y}} \in [0, 2\pi)$. By circular symmetry, it follows that

$$\frac{Y_k}{\tilde{Y}_k} = \left|\frac{Y_k}{\tilde{Y}_k}\right| e^{j(\phi_Y - \phi_{\tilde{Y}})} \quad (26)$$

where, conditioned on $\{X_{\ell,k}\}$, ϕ_Y and $\phi_{\tilde{Y}}$ are uniformly distributed and independent of $\left|\frac{Y_k}{\tilde{Y}_k}\right|$. Since this is true irrespective of $\{X_{\ell,k}\}$, it follows that $\frac{Y_k}{\tilde{Y}_k}$ is circularly symmetric.

Once we have determined that the ratio $\frac{Y_k}{\tilde{Y}_k}$ is circularly symmetric, we can apply [10, Lemma 6.16] to obtain

$$\begin{aligned} h\left(\frac{Y_k}{\tilde{Y}_k}\right) &= \log 2\pi + h\left(\left|\frac{Y_k}{\tilde{Y}_k}\right|\right) + \mathbf{E}\left[\log\left(\left|\frac{Y_k}{\tilde{Y}_k}\right|\right)\right] \\ &\leq \log 2\pi + h\left(\left|\frac{Y_k}{\tilde{Y}_k}\right|\right) + \log\left(\mathbf{E}\left[\left|\frac{Y_k}{\tilde{Y}_k}\right|\right]\right) \end{aligned} \quad (27)$$

where the inequality follows from Jensen's inequality.

Let $R \triangleq \left|\frac{Y_k}{\tilde{Y}_k}\right|$. Conditioned on $\{X_{\ell,k}\}$, R is the ratio of two independent Rayleigh distributed random variables, whose probability density function is given by [13, p. 68, Eq. (7.44)]

$$f_{R|\{X_{\ell,k}\}}(r|\{x_{\ell,k}\}) = \frac{2\mathbf{E}\left[|\tilde{Y}_k|^2|\{X_{\ell,k}\} = \{x_{\ell,k}\}\right] \mathbf{E}\left[|Y_k|^2|\{X_{\ell,k}\} = \{x_{\ell,k}\}\right] r}{\left(\mathbf{E}\left[|\tilde{Y}_k|^2|\{X_{\ell,k}\} = \{x_{\ell,k}\}\right] r^2 + \mathbf{E}\left[|Y_k|^2|\{X_{\ell,k}\} = \{x_{\ell,k}\}\right]\right)^2}, \quad (28)$$

for $r > 0$.

To compute the needed differential entropy of R , we need to marginalize over $\{X_{\ell,k}\}$. To this end, we define

$$\eta^2(\{x_{\ell,k}\}) \triangleq \frac{\mathbf{E}\left[|Y_k|^2|\{X_{\ell,k}\} = \{x_{\ell,k}\}\right]}{\mathbf{E}\left[|\tilde{Y}_k|^2|\{X_{\ell,k}\} = \{x_{\ell,k}\}\right]} \quad (29)$$

which can be evaluated as

$$\eta^2(\{x_{\ell,k}\}) = \frac{\sum_{\ell=0}^{\infty} \alpha_{\ell} |x_{\ell,k}|^2 + \sigma^2}{\sum_{\ell=0}^{\infty} \alpha_{\ell+1} |x_{\ell,k}|^2 + \sigma^2}. \quad (30)$$

With this, we obtain

$$f_R(r) = 2r \mathbf{E}\left[\frac{\eta^2(\{X_{\ell,k}\})}{(r^2 + \eta^2(\{X_{\ell,k}\}))^2}\right] \quad (31)$$

where the expectation is over $\{X_{\ell,k}\}$. This yields for the differential entropy of R

$$\begin{aligned}
h\left(\left|\frac{Y_k}{\tilde{Y}_k}\right|\right) &\triangleq -\int_0^\infty f_R(r) \log f_R(r) dr \\
&= -\int_0^\infty 2r \mathbf{E} \left[\frac{\eta^2(\{X_{\ell,k}\})}{(r^2 + \eta^2(\{X_{\ell,k}\}))^2} \right] \log \left(2r \mathbf{E} \left[\frac{\eta^2(\{X_{\ell,k}\})}{(r^2 + \eta^2(\{X_{\ell,k}\}))^2} \right] \right) dr \\
&= -\int_0^\infty 2r \mathbf{E} \left[\frac{\eta^2(\{X_{\ell,k}\})}{(r^2 + \eta^2(\{X_{\ell,k}\}))^2} \right] \log(2r) dr \\
&\quad - \int_0^\infty 2r \mathbf{E} \left[\frac{\eta^2(\{X_{\ell,k}\})}{(r^2 + \eta^2(\{X_{\ell,k}\}))^2} \right] \log \left(\mathbf{E} \left[\frac{\eta^2(\{X_{\ell,k}\})}{(r^2 + \eta^2(\{X_{\ell,k}\}))^2} \right] \right) dr.
\end{aligned} \tag{32}$$

Computing the first integral yields

$$\begin{aligned}
-\int_0^\infty 2r \mathbf{E} \left[\frac{\eta^2(\{X_{\ell,k}\})}{(r^2 + \eta^2(\{X_{\ell,k}\}))^2} \right] \log(2r) dr &= -\mathbf{E} \left[\int_0^\infty 2r \frac{\eta^2(\{X_{\ell,k}\})}{(r^2 + \eta^2(\{X_{\ell,k}\}))^2} \log(2r) dr \right] \\
&= -\log(2) - \mathbf{E} \left[\int_0^\infty 2r \frac{\eta^2(\{X_{\ell,k}\})}{(r^2 + \eta^2(\{X_{\ell,k}\}))^2} \log(r) dr \right] \\
&= -\log(2) - \frac{1}{2} \mathbf{E} [\log(\eta^2(\{X_{\ell,k}\}))]
\end{aligned} \tag{33}$$

where the first step follows by Fubini's theorem [14, p. 108, Th. 2.6.6], the last step follows from computing the integral using the change of variable $t = r^2$ and [15, p. 535, Sec. 4.23, Eq. (5)].

To bound the second integral in (32), we require bounds on $\eta^2(\{x_{\ell,k}\})$, which we shall present in the following lemma.

Lemma 1. *Assume that $\{\alpha_\ell\}$ satisfies (2), namely, $\frac{\alpha_{\ell+1}}{\alpha_\ell} \geq \rho$, $\ell = 1, 2, \dots$. Then, $\eta^2(\{x_{\ell,k}\})$ is bounded by*

$$\eta_{\min}^2 \leq \eta^2(\{x_{\ell,k}\}) \leq \eta_{\max}^2 \tag{34}$$

where

$$\eta_{\min}^2 \triangleq \min \left(1, \frac{\alpha_0}{\alpha_1} \right) \tag{35}$$

$$\eta_{\max}^2 \triangleq \max \left(\frac{\alpha_0}{\alpha_1}, \frac{1}{\rho} \right). \tag{36}$$

Proof: By (30), we have

$$\eta^2(\{x_{\ell,k}\}) = \frac{\sum_{\ell=0}^\infty \alpha_\ell |x_{\ell,k}|^2 + \sigma^2}{\sum_{\ell=0}^\infty \alpha_{\ell+1} |x_{\ell,k}|^2 + \sigma^2}. \tag{37}$$

We next maximize and minimize

$$\eta^2(\{x_{\ell,k}\}) = \frac{\alpha_0 |x_{0,k}|^2 + \alpha_1 |x_{1,k}|^2 + \dots + \sigma^2}{\alpha_1 |x_{0,k}|^2 + \alpha_2 |x_{1,k}|^2 + \dots + \sigma^2} \tag{38}$$

over $x_{0,k}, x_{1,k}, \dots$. To this end, we first fix $x_{1,k}, x_{2,k}, \dots$ and optimize over $x_{0,k}$. Thus, we write $\eta^2(\{x_{\ell,k}\})$ as

$$f(x) \triangleq \frac{\alpha_0 x + \beta_1}{\alpha_1 x + \beta_2} \tag{39}$$

where x corresponds to $|x_{0,k}|^2$, and β_1 and β_2 are the remaining terms in the numerator and denominator in (38), respectively, i.e., $\beta_1 = \alpha_1 |x_{1,k}|^2 + \dots + \sigma^2$ and $\beta_2 = \alpha_2 |x_{1,k}|^2 + \dots + \sigma^2$. It is easy to show that

$$\min \left(\frac{\alpha_0}{\alpha_1}, \frac{\beta_1}{\beta_2} \right) \leq f(x) \leq \max \left(\frac{\alpha_0}{\alpha_1}, \frac{\beta_1}{\beta_2} \right). \tag{40}$$

We next express $\frac{\beta_1}{\beta_2}$ as

$$\tilde{f}(x) = \frac{\alpha_1 x + \tilde{\beta}_1}{\alpha_2 x + \tilde{\beta}_2} \tag{41}$$

where x corresponds to $|x_{1,k}|^2$, and $\tilde{\beta}_1$ and $\tilde{\beta}_2$ are the remaining terms in the numerator and denominator, respectively, i.e., $\tilde{\beta}_1 = \alpha_2 |x_{2,k}|^2 + \dots + \sigma^2$ and $\tilde{\beta}_2 = \alpha_3 |x_{2,k}|^2 + \dots + \sigma^2$. Again, $\min \left(\frac{\alpha_1}{\alpha_2}, \frac{\tilde{\beta}_1}{\tilde{\beta}_2} \right) \leq \tilde{f}(x) \leq \max \left(\frac{\alpha_1}{\alpha_2}, \frac{\tilde{\beta}_1}{\tilde{\beta}_2} \right)$, so

$$\min \left(\frac{\alpha_0}{\alpha_1}, \frac{\alpha_1}{\alpha_2}, \frac{\tilde{\beta}_1}{\tilde{\beta}_2} \right) \leq f(x) \leq \max \left(\frac{\alpha_0}{\alpha_1}, \frac{\alpha_1}{\alpha_2}, \frac{\tilde{\beta}_1}{\tilde{\beta}_2} \right). \tag{42}$$

Repeating these steps and considering that at the end we have $\frac{\beta_1}{\beta_2} = \frac{\sigma}{\sigma} = 1$, we obtain

$$\eta^2(\{x_{\ell,k}\}) \leq \max \left(\sup_{k=0,1,\dots} \left(\frac{\alpha_k}{\alpha_{k+1}} \right), 1 \right) \triangleq \eta_{\max}^2, \quad (43)$$

$$\eta^2(\{x_{\ell,k}\}) \geq \min \left(1, \inf_{k=0,1,\dots} \left(\frac{\alpha_k}{\alpha_{k+1}} \right) \right) \triangleq \eta_{\min}^2. \quad (44)$$

Using the assumption that $\alpha_1 \geq \alpha_2 \geq \dots$ and (2), we have

$$1 \leq \frac{\alpha_k}{\alpha_{k+1}} \leq \frac{1}{\rho}, \quad k = 1, 2, \dots \quad (45)$$

from which we obtain

$$\eta_{\min}^2 = \min \left(1, \frac{\alpha_0}{\alpha_1} \right) \quad (46)$$

$$\eta_{\max}^2 = \max \left(\frac{\alpha_0}{\alpha_1}, \frac{1}{\rho} \right). \quad (47)$$

■

By the continuity of $x \rightarrow \frac{x}{(r^2+x)^2}$, it follows that we can define

$$\eta_*^2 = \operatorname{argmin}_{\eta_{\min}^2 \leq \eta^2 \leq \eta_{\max}^2} \frac{\eta^2}{(r^2 + \eta^2)^2} \quad (48)$$

which satisfies $\eta_{\min}^2 \leq \eta_*^2 \leq \eta_{\max}^2$.

We next apply (48) to compute an upper bound on the second integral in (32). By definition, we have that

$$\mathbb{E} \left[\left(\frac{\eta^2(\{X_{\ell,k}\})}{(r^2 + \eta^2(\{X_{\ell,k}\}))^2} \right) \right] \geq \frac{\eta_*^2}{(r^2 + \eta_*^2)^2} \quad (49)$$

so by the monotonicity of the logarithm function

$$\begin{aligned} & - \int_0^\infty 2r \mathbb{E} \left[\frac{\eta^2(\{X_{\ell,k}\})}{(r^2 + \eta^2(\{X_{\ell,k}\}))^2} \right] \log \left(\mathbb{E} \left[\frac{\eta^2(\{X_{\ell,k}\})}{(r^2 + \eta^2(\{X_{\ell,k}\}))^2} \right] \right) dr \\ & \leq - \int_0^\infty \mathbb{E} \left[\frac{2r\eta^2(\{X_{\ell,k}\})}{(r^2 + \eta^2(\{X_{\ell,k}\}))^2} \right] \log \left(\frac{\eta_*^2}{(r^2 + \eta_*^2)^2} \right) dr \\ & = - \mathbb{E} \left[\int_0^\infty \frac{2r\eta^2(\{X_{\ell,k}\})}{(r^2 + \eta^2(\{X_{\ell,k}\}))^2} \log \left(\frac{\eta_*^2}{(r^2 + \eta_*^2)^2} \right) dr \right] \\ & = - \log(\eta_*^2) + \mathbb{E} \left[\int_0^\infty \frac{4r\eta^2(\{X_{\ell,k}\})}{(r^2 + \eta^2(\{X_{\ell,k}\}))^2} \log(r^2 + \eta_*^2) dr \right] \end{aligned} \quad (50)$$

where the second step follows by Fubini's theorem.

To derive an upper bound on the second term in (50), we integrate by parts. To this end, we have to distinguish between the cases $\eta_*^2 \geq \eta^2(\{x_{\ell,k}\})$ and $\eta_*^2 < \eta^2(\{x_{\ell,k}\})$. The details are carried in Appendix A, the results are as follows:

If $\eta_*^2 \geq \eta^2(\{x_{\ell,k}\})$, then

$$\begin{aligned} \mathbb{E} \left[\int_0^\infty \frac{4r\eta^2(\{X_{\ell,k}\}) \log(r^2 + \eta_*^2)}{(r^2 + \eta^2(\{X_{\ell,k}\}))^2} dr \mathbb{1}\{\eta_*^2 \geq \eta^2(\{X_{\ell,k}\})\} \right] & \leq \mathbb{E} \left[(2 \log(\eta_*^2) + 2) \mathbb{1}\{\eta_*^2 \geq \eta^2(\{X_{\ell,k}\})\} \right] \\ & = (2 \log(\eta_*^2) + 2) \Pr(\eta_*^2 \geq \eta^2(\{X_{\ell,k}\})) \end{aligned} \quad (51)$$

where the last step follows because η_*^2 is deterministic. Here $\mathbb{1}\{\cdot\}$ denotes the indicator function.

If $\eta_*^2 < \eta^2(\{x_{\ell,k}\})$, then

$$\begin{aligned} \mathbb{E} \left[\int_0^\infty \frac{4r\eta^2(\{X_{\ell,k}\}) \log(r^2 + \eta_*^2)}{(r^2 + \eta^2(\{X_{\ell,k}\}))^2} dr \mathbb{1}\{\eta_*^2 < \eta^2(\{X_{\ell,k}\})\} \right] & \leq \mathbb{E} \left[\left(2 \log(\eta_*^2) + 2 \frac{\eta^2(\{X_{\ell,k}\})}{\eta_*^2} \right) \mathbb{1}\{\eta_*^2 < \eta^2(\{X_{\ell,k}\})\} \right] \\ & \leq \left(2 \log(\eta_*^2) + 2 \frac{\eta_{\max}^2}{\eta_*^2} \right) \Pr(\eta_*^2 < \eta^2(\{X_{\ell,k}\})) \end{aligned} \quad (52)$$

where the last step follows because $\eta^2(\{x_{\ell,k}\}) \leq \eta_{\max}^2$ and because $\log(\eta_*^2)$ and η_{\max}^2 are deterministic.

Combining (33) and (50)-(52) with (32), we obtain

$$\begin{aligned}
h\left(\frac{Y_k}{\tilde{Y}_k}\right) &\leq -\log(2) - \frac{1}{2}\mathbf{E}[\log(\eta^2(\{X_{\ell,k}\}))] - \log(\eta_*^2) + (2\log(\eta_*^2) + 2)\Pr(\eta_*^2 \geq \eta^2(\{X_{\ell,k}\})) \\
&\quad + \left(2\log(\eta_*^2) + 2\frac{\eta_{\max}^2}{\eta_*^2}\right)\Pr(\eta_*^2 < \eta^2(\{X_{\ell,k}\})) \\
&\leq -\log(2) - \frac{1}{2}\mathbf{E}[\log(\eta^2(\{X_{\ell,k}\}))] + \log(\eta_*^2) + 2\frac{\eta_{\max}^2}{\eta_*^2} \\
&\leq -\log(2) - \frac{1}{2}\log(\eta_{\min}^2) + \log(\eta_*^2) + 2\frac{\eta_{\max}^2}{\eta_*^2}
\end{aligned} \tag{53}$$

where the second step follows by upper-bounding $1 \leq \frac{\eta_{\max}^2}{\eta_*^2}$ and the last step follows by lower-bounding $\mathbf{E}[\log(\eta^2(\{X_{\ell,k}\}))] \geq \log(\eta_{\min}^2)$.

Together with (27), this yields

$$\begin{aligned}
h\left(\frac{Y_k}{\tilde{Y}_k}\right) &\leq \log 2\pi + h\left(\frac{Y_k}{\tilde{Y}_k}\right) + \log\left(\mathbf{E}\left[\left|\frac{Y_k}{\tilde{Y}_k}\right|\right]\right) \\
&\leq \log(\pi) - \frac{1}{2}\log(\eta_{\min}^2) + \log(\eta_*^2) + 2\frac{\eta_{\max}^2}{\eta_*^2} + \log\left(\mathbf{E}\left[\left|\frac{Y_k}{\tilde{Y}_k}\right|\right]\right) \\
&= \log(\pi) - \frac{1}{2}\log(\eta_{\min}^2) + \log(\eta_*^2) + 2\frac{\eta_{\max}^2}{\eta_*^2} + \log\left(\mathbf{E}\left[\frac{\pi}{2}\eta(\{X_{\ell,k}\})\right]\right) \\
&= 2\log(\pi) - \log(2) - \frac{1}{2}\log(\eta_{\min}^2) + \log(\eta_*^2) + 2\frac{\eta_{\max}^2}{\eta_*^2} + \frac{1}{2}\log\mathbf{E}[\eta^2(\{X_{\ell,k}\})] \\
&\leq 2\log(\pi) - \log(2) - \frac{1}{2}\log(\eta_{\min}^2) + \log(\eta_*^2) + 2\frac{\eta_{\max}^2}{\eta_*^2} + \frac{1}{2}\log(\eta_{\max}^2) \\
&= 2\log(\pi) - \log(2) + \log(\eta_*^2) + \frac{1}{2}\log\left(\frac{\eta_{\max}^2}{\eta_{\min}^2}\right) + 2\frac{\eta_{\max}^2}{\eta_*^2}
\end{aligned} \tag{54}$$

where we have evaluated $\mathbf{E}\left[\left|\frac{Y_k}{\tilde{Y}_k}\right|\right]$ using [13, p.69, Eq. (7.46)] (see Appendix B) and we have upper-bounded $\mathbf{E}[\eta^2(\{X_{\ell,k}\})] \leq \eta_{\max}^2$.

Replacing (22), (23) and (54) in (21), we finally obtain

$$\begin{aligned}
h(Y_1^n) - h(\tilde{Y}_1^n) &\leq \sum_{k=1}^n \left(\mathbf{E}\left[\log\left(\mathbf{E}\left[|\tilde{Y}_k|^2|\{X_{\ell,k}\}\right]\right)\right] - \log(\pi e) - \mathbf{E}\left[\log\left(\mathbf{E}\left[|\tilde{Y}_k|^2|\{X_{\ell,k}\}\right]\right)\right] \right) \\
&\quad + 2\log(\pi) - \log(2) + \log(\eta_*^2) + \frac{1}{2}\log\left(\frac{\eta_{\max}^2}{\eta_{\min}^2}\right) + 2\frac{\eta_{\max}^2}{\eta_*^2} \\
&= \sum_{k=1}^n \left(\log(\pi) - 1 - \log(2) + \log(\eta_*^2) + \frac{1}{2}\log\left(\frac{\eta_{\max}^2}{\eta_{\min}^2}\right) + 2\frac{\eta_{\max}^2}{\eta_*^2} \right) \\
&= n\mathsf{K}
\end{aligned} \tag{55}$$

where

$$\mathsf{K} \triangleq \log(\pi) - 1 - \log(2) + \log(\eta_*^2) + \frac{1}{2}\log\left(\frac{\eta_{\max}^2}{\eta_{\min}^2}\right) + 2\frac{\eta_{\max}^2}{\eta_*^2}. \tag{56}$$

Back to the mutual information (18), we have

$$\frac{1}{n}I(X_1^n; Y_1^n) \leq \mathsf{K} \tag{57}$$

hence the capacity is upper-bounded by

$$C(\text{SNR}) \leq \mathsf{K}. \tag{58}$$

This proves Theorem 1.

VI. CONCLUSIONS

We analyzed the channel capacity of a noncoherent wireless network. We demonstrated that the channel capacity is bounded in the SNR, provided that the nodes do not cooperate and the path gains $\{\alpha_\ell\}$ decay exponentially or more slowly. This confirms the observation made by Lozano et al. [5] that in the absence of perfect CSI the capacity of wireless networks is bounded in the SNR. Our analysis is more general than the analysis in [5] in the sense that we allow that the distribution of the channel inputs may depend on the SNR, but it is less general than the analysis in [5] in the sense that we assumed that the variances of the fading coefficients decay at most exponentially and that the nodes do not cooperate.

The behavior of the channel capacity observed in this paper is similar to the one observed for frequency-selective fading channels by Koch and Lapidath in [12]. Specifically, Koch and Lapidath showed that if the path gains $\{\alpha_\ell\}$ decay exponentially or more slowly, then the capacity is bounded in the SNR; conversely, if the path gains decay faster than exponentially, then the capacity is unbounded in the SNR. The question then arises as to whether the channel capacity of noncoherent wireless networks is also unbounded in the SNR if we assume that the variances of the fading coefficients decay faster than exponentially. This will be the subject of future studies.

APPENDIX A UPPER BOUND ON (50)

We solve the second integral in (50) divided by $2\eta^2(\{x_{\ell,k}\})$, namely,

$$\int_0^\infty \frac{2r}{(r^2 + \eta^2(\{x_{\ell,k}\}))^2} \log(r^2 + \eta_*^2) dr. \quad (59)$$

In the following, we first apply the change of variable $t = r^2$ to obtain

$$\int_0^\infty \frac{\log(t + \eta_*^2)}{(t + \eta^2(\{x_{\ell,k}\}))^2} dt. \quad (60)$$

We then integrate by parts. To this end, we distinguish between the two cases $\eta_*^2 \geq \eta^2(\{x_{\ell,k}\})$ and $\eta_*^2 < \eta^2(\{x_{\ell,k}\})$.

A. $\eta_*^2 \geq \eta^2(\{x_{\ell,k}\})$

We use another change of variable, namely, $x = t + \eta^2(\{x_{\ell,k}\})$, to obtain

$$\int_0^\infty \frac{\log(t + \eta_*^2)}{(t + \eta^2(\{x_{\ell,k}\}))^2} dt = \int_{\eta^2(\{x_{\ell,k}\})}^\infty \frac{\log(x + \eta_*^2 - \eta^2(\{x_{\ell,k}\}))}{x^2} dx. \quad (61)$$

Introducing the auxiliary variable $\gamma = \eta_*^2 - \eta^2(\{x_{\ell,k}\})$, this can be written as

$$\int_{\eta^2(\{x_{\ell,k}\})}^\infty \frac{\log(x + \eta_*^2 - \eta^2(\{x_{\ell,k}\}))}{x^2} dx = \int_{\eta^2(\{x_{\ell,k}\})}^\infty \frac{1}{x^2} \log(x + \gamma) dx \quad (62)$$

which we now integrate by parts:

$$\begin{aligned} \int_{\eta^2(\{x_{\ell,k}\})}^\infty \frac{1}{x^2} \log(x + \gamma) dx &= -\frac{1}{x} \log(x + \gamma) \Big|_{\eta^2(\{x_{\ell,k}\})}^\infty + \int_{\eta^2(\{x_{\ell,k}\})}^\infty \frac{1}{x(x + \gamma)} dx \\ &= \frac{\log(\eta^2(\{x_{\ell,k}\}) + \gamma)}{\eta^2(\{x_{\ell,k}\})} + \frac{1}{\gamma} \log\left(1 + \frac{\gamma}{\eta^2(\{x_{\ell,k}\})}\right) \\ &\leq \frac{\log(\eta^2(\{x_{\ell,k}\}) + \gamma)}{\eta^2(\{x_{\ell,k}\})} + \frac{1}{\eta^2(\{x_{\ell,k}\})} \\ &= \frac{1}{\eta^2(\{x_{\ell,k}\})} (\log(\eta_*^2) + 1) \end{aligned} \quad (63)$$

where the second step follows from [15, p. 70, Sec 2.118, Eq. (1)], the third step follows from $\log(1+b) \leq b$ being $b = \frac{\gamma}{\eta^2(\{x_{\ell,k}\})}$ and the last step follows by replacing $\eta^2(\{x_{\ell,k}\}) + \gamma = \eta_*^2$. Multiplying (63) by $2\eta^2(\{x_{\ell,k}\})$ and averaging over all values of $\{x_{\ell,k}\}$ yields (51).

B. $\eta_*^2 < \eta^2(\{x_{\ell,k}\})$

We use again a change of variable, namely, $x = t + \eta_*^2$, to express the integral as

$$\int_0^\infty \frac{\log(t + \eta_*^2)}{(t + \eta^2(\{x_{\ell,k}\}))^2} dt = \int_{\eta_*^2}^\infty \frac{\log(x)}{(x + \eta^2(\{x_{\ell,k}\}) - \eta_*^2)^2} dx. \quad (64)$$

Introducing the auxiliary variable $\gamma = \eta^2(\{x_{\ell,k}\}) - \eta_*^2$ and integrating by parts yields

$$\begin{aligned} \int_{\eta_*^2}^\infty \frac{\log(x)}{(x + \gamma)^2} dx &= -\frac{\log(x)}{(x + \gamma)} \Big|_{\eta_*^2}^\infty + \int_{\eta_*^2}^\infty \frac{1}{x(x + \gamma)} dx \\ &= \frac{\log(\eta_*^2)}{(\eta_*^2 + \gamma)} + \frac{1}{\gamma} \log\left(1 + \frac{\gamma}{\eta_*^2}\right) \\ &\leq \frac{1}{(\eta_*^2 + \gamma)} \log(\eta_*^2) + \frac{1}{\eta_*^2} \\ &= \frac{1}{\eta^2(\{x_{\ell,k}\})} \left(\log(\eta_*^2) + \frac{\eta^2(\{x_{\ell,k}\})}{\eta_*^2} \right) \end{aligned} \quad (65)$$

where the second step follows from [15, p. 70, Sec. 2.118, Eq. (1)], the third step follows from $\log(1 + b) \leq b$ being $b = \frac{\gamma}{\eta_*^2}$, and the last step follows by replacing $\eta^2(\{x_{\ell,k}\}) - \gamma = \eta_*^2$. Multiplying (65) by $2\eta^2(\{x_{\ell,k}\})$ and averaging over all values of $\{x_{\ell,k}\}$ yields (52).

APPENDIX B EXPECTED VALUE OF THE RATIO $\left| \frac{Y_k}{\tilde{Y}_k} \right|$

In this section we calculate the expected value of the ratio $R = \left| \frac{Y_k}{\tilde{Y}_k} \right|$. Using the probability density function in (31), we have

$$\begin{aligned} \mathbb{E}[R] &= \int_0^\infty 2r^2 \mathbb{E} \left[\frac{\eta^2(\{x_{\ell,k}\})}{(r^2 + \eta^2(\{x_{\ell,k}\}))^2} \right] dr \\ &= \mathbb{E} \left[\int_0^\infty \frac{2r^2 \eta^2(\{x_{\ell,k}\})}{(r^2 + \eta^2(\{x_{\ell,k}\}))^2} dr \right] \\ &= \mathbb{E} \left[2\eta^2(\{x_{\ell,k}\}) \int_0^\infty \frac{r^2}{(r^2 + \eta^2(\{x_{\ell,k}\}))^2} dr \right] \end{aligned} \quad (66)$$

where the second step follows by applying Fubini's theorem. Computing the integral yields

$$\begin{aligned} \int_0^\infty \frac{r^2}{(r^2 + \eta^2(\{x_{\ell,k}\}))^2} dr &= \frac{1}{2\eta^4(\{x_{\ell,k}\})} (\eta^3(\{x_{\ell,k}\})) \frac{\Gamma(\frac{3}{2}) \Gamma(\frac{1}{2})}{\Gamma(2)} \\ &= \frac{1}{2\eta(\{x_{\ell,k}\})} \frac{\pi}{2}, \end{aligned} \quad (67)$$

where the first step follows from [15, p. 322, Sec 3.241, Eq. (4)]. Combining (67) with (66), we obtain the expected value of $\left| \frac{Y_k}{\tilde{Y}_k} \right|$:

$$\mathbb{E} \left[\left| \frac{Y_k}{\tilde{Y}_k} \right| \right] = \mathbb{E}[R] = \frac{\pi}{2} \mathbb{E}[\eta(\{x_{\ell,k}\})]. \quad (68)$$

REFERENCES

- [1] O. Simeone, O. Somekh, H. Poor, and S. Shamai, "Local base station cooperation via finite-capacity links for the uplink of linear cellular networks," *IEEE Transactions on Information Theory*, vol. 55, no. 1, pp. 190–204, Jan. 2009.
- [2] A. Ozgur, O. Leveque, and D. Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," *IEEE Transactions on Information Theory*, vol. 53, no. 10, pp. 3549–3572, Oct. 2007.
- [3] V. Cadambe and S. Jafar, "Interference alignment and degrees of freedom of the k-user interference channel," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.
- [4] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H.-P. Mayer, L. Thiele, and V. Jungnickel, "Coordinated multipoint: Concepts, performance, and field trial results," *IEEE Communications Magazine*, vol. 49, no. 2, pp. 102–111, Feb. 2011.
- [5] A. Lozano, R. Heath, and J. Andrews, "Fundamental limits of cooperation," *IEEE Transactions on Information Theory*, vol. 59, no. 9, pp. 5213–5226, Sep. 2013.
- [6] N. M. Blachman, "Communication as a game," in *Proc. IRE WESCON Conf*, 1957, pp. 61–66.
- [7] R. Dobrushin, "Optimum information transmission through a channel with unknown parameters," *Radio Eng. Electron*, vol. 4, no. 12, pp. 1–8, 1959.

- [8] I. Shomorony and A. Avestimehr, "Worst-case additive noise in wireless networks," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3833–3847, Jun. 2013.
- [9] S. Shamai and S. Verdú, "Worst-case power-constrained noise for binary-input channels," *IEEE Transactions on Information Theory*, vol. 38, no. 5, pp. 1494–1511, 1992.
- [10] A. Lapidoth and S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2426–2467, 2003.
- [11] T. M. Cover and J. A. Thomas, *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 2006.
- [12] T. Koch and A. Lapidoth, "On multipath fading channels at high SNR," *IEEE Transactions on Information Theory*, vol. 56, no. 12, pp. 5945–5957, Dec. 2010.
- [13] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables: A Handbook for Engineers, Scientists and Mathematicians*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.
- [14] R. B. Ash and C. Doleans-Dade, *Probability and Measure Theory*. Academic Press, 2000.
- [15] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. Elsevier/Academic Press, Amsterdam, 2007.