Achievable Rates and Scaling Laws of Power-constrained Wireless Sensory Relay Networks *

Bo Wang and Junshan Zhang
Dept. of E.E.
A.S.U.
Tempe, AZ 85287
{bo.wang, junshan.zhang}@asu.edu

Lizhong Zheng
Dept. of E.E.C.S.
M.I.T.
Cambridge, MA 02139
lizhong@mit.edu

Abstract

A wireless sensory relay network consists of one source node, one destination node and multiple intermediate relay nodes. In this paper, we study the achievable rates and the scaling laws of power-constrained wireless relay networks in the wideband regime, assuming that relay nodes have no *a priori* knowledge of channel state information (CSI) for both the backward channels and the forward channels. We examine the achievable rates in the joint asymptotic regime of the number of relay nodes n, the channel coherence interval L, and the bandwidth W (or the SNR per link ρ). We first study narrowband relay networks in the low SNR regime. We investigate a relaying scheme, namely amplify-and-forward (AF) with network training, in which the source node and the destination node broadcast training symbols and each relay node carries out channel estimation and then applies AF relaying to relay information. We provide an equivalent source-to-destination channel model, and characterize the corresponding achievable rate. Our findings show that when ρL, proportional to the transmission energy in each fading block, is bounded below, the achievable rate has the same scaling order as in coherent relaying, thus enabling us to characterize the scaling law of the relay networks in the low SNR regime. We then generalize the study to power-constrained wideband relay networks, where frequency-selective fading is taken into account. Again, the focus is on the achievable rates by using AF with network training for information relaying. In particular, we examine the scaling behavior of the achievable rates corresponding to two power allocation policies across the frequency subbands at relay nodes, namely, a simple equal power allocation policy and the optimal power allocation policy. We identify the conditions under which the scaling law of the wideband relay networks can be achieved by both power allocation policies. Somewhat surprising, our findings indicate that these two power allocation policies result in achievable rates of the same scaling order, and the scaling law can be characterized under the condition that L/W, proportional to the energy per fading block per subband, is bounded below, and that W is sub-linear in n.

Keywords: scaling law, wireless sensory relay network, cooperative relaying, network training.

*This research is supported in part by Office of Naval Research through the grant N00014-05-1-0636 and National Science Foundation through the CAREER award ANI-0208135.

1 Introduction

The last decade has witnessed a tremendous growth of interest in wireless sensor networks. Current and anticipated technological advances in sensor design, coupled with advances in data processing and wireless communications, are creating a new realm of possibilities for the development of sensor networks in many application domains. A sensor network is often designed to carry out a set of high-level data acquisition and information processing tasks. The raw sensor measurements are obtained and then need to be transported to a data collecting/fusion center to be analyzed, which requires reliable networking capabilities in disadvantaged wireless environments. There are many sensor network models developed for different applications: multi-hop sensor networks, many-to-one sensor networks, sensory relay networks - to name a few.

In this paper, we consider power constrained wireless sensory relay networks in the wideband regime, because of their ability of overlay with other legacy networks and the advantage that the use of a larger bandwidth can offer significant power savings. In sensor networks, most sensor devices have limited power supply (e.g., battery supply), and lower power consumption of sensor devices is of critical importance. Wideband communications has recently garnered much attention (e.g., ultra-wide-band (UWB) systems) [24]. Indeed, UWB radios are expected to be inexpensive and low-power, and are ideal for sensor network applications. It is of great interest to investigate the fundamental limits in such systems, such as the capacity and energy efficiency. Thus motivated, we focus on the sensory relay networks in which each node is power constrained and the frequency bandwidth for communication may be large. Following [21], "the wideband regime" here encompasses all scenarios where the information bits transmitted per receive dimension are small. Two key challenges in power constrained wideband relay networks are as follows: 1) coherent relaying may not be feasible due to the low SNR and the corresponding cooperative relaying strategies should be pursued; and 2) when frequency bandwidth is large, it remains open what power allocation policies across the subbands work well at relay nodes.

As depicted in Fig. 1, a wireless sensory relay network consists of one source-destination pair and n relay nodes. When the measurement data at the source node are sent to the destination node, many intermediate sensors can serve as relay nodes to cooperatively relay the information. (This model is well motivated by event-driven sensor networks, in which the data gathered by the source node can be relayed by other sensors in a cooperative manner.) The cooperation among relay nodes can yield diversity gain to enhance the data transport capacity between the source node and the destination node. As is standard (e.g., [1, 4, 7, 8, 14]), we consider a two-hop relaying strategy, in which each relay node first receives the signal from the source node in the first hop and then forwards a processed signal to the destination node in the second hop. We refer the links between the source node and the relay nodes as the backward channels, composing the first hop; and those between the relay nodes and the destination node as the forward channels, corresponding to the second hop. We assume that every link experiences Rayleigh fading. We note that multiple source-relay-destination links offer multiple spatial

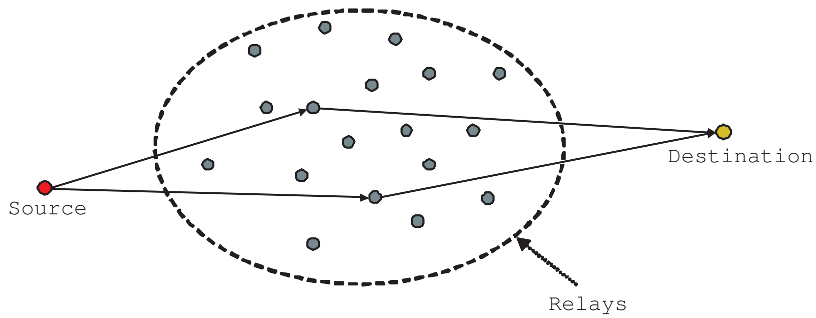

Figure 1: A wireless sensory relay network model

routing paths for information relaying, pointing to potential cooperative diversity gain.

Clearly, The key to reap the inherent cooperative diversity gain in such networks is to devise good relaying strategies. Recently, "decode-and-forward (DF)" and "amplify-and-forward (AF)" have emerged as two popular candidates for two-hop information relaying. In the decode-and-forward strategy, the relay nodes first extract the information bits from the received signals, re-encode it, and then transmit the new encoded bits in the second hop. For instance, one DF scheme can be to select the most reliable source-relay-destination path to transmit the signal. In general, the optimal DF strategy is highly non-trivial and the complexity increases significantly as the number of relay nodes grows. Alternatively, simple amplify-and-forward relaying can be used to relay information and has low complexity. Intuitively speaking, to fully reap the cooperative diversity gain, a good AF relaying scheme should reduce the phase-offset across different source-relay-destination links as much as possible, and this phase-aligning operation should enable coherent combining of transmitted signals from relay nodes at the destination node. The focus of this study is to investigate the "right" relaying strategies for achieving high throughput in wireless relay networks in the wideband regime. In particular, we attempt to address the following critical issues:

- When would "amplify-and-forward" work well in relay networks in the wideband regime? What is the "right" amplify-and-forward scheme to close the gap between coherent relaying and non-coherent relaying?
- What is the "right" power allocation strategy at the relay nodes that would maximize the scaling order of the achievable rates of relay networks in the wideband regime?
- When would it be possible to achieve the same scaling laws as in coherent relay networks?

Recall that the wideband regime encompasses all scenarios where the information bits transmitted per receive dimension are small. One key challenge in relay networks in the wideband regime is that coherent communications may not be feasible due to the low SNR. To obtain a clear understanding of the impact of low SNR on cooperative relaying, we make a more realistic assumption that at relay nodes there is no *a priori* knowledge of channel state information (CSI) for both the backward channels and the forward channels, and this is one of key features distinguishing our model from the existing ones. As a result, each relay node needs to estimate the channel conditions prior to the data transmission, in order to carry out the phase-alignment as in coherent relaying. In particular, the channel estimation of the forward links is more challenging because it is a multi-access channel. A naive approach is to carry out channel estimation of the forward links separately, which would lead to a significant reduction of the throughput and is clearly not energy-efficient. To overcome this difficulty, we study the following strategy for AF relaying. As depicted in Fig. 2, at the beginning of the first hop transmission, the source node first broadcasts common pilot symbols, followed by its data transmission. Based on the received signals corresponding to the pilot symbols, each relay nodes estimates its own backward channel condition. Then at the beginning of the second hop, the destination node (not the relay nodes) sends out pilot symbols (We assume that the forward links are reciprocal in both directions in each subband.) Then, each relay node estimates its forward channel condition. After the channel estimation is done, each relay node carries out phase alignment by using its channel estimates of both the backward channel and the forward channel, amplifies the received data signal under the given power constraints, and forwards the processed signal to the destination node. For convenience, we refer this strategy as amplify-and-forward (AF) with *network training*. We emphasize that the above channel estimation approach makes use of the broadcast nature of wireless transmissions, and is an energy-efficient candidate for network training. For instance, this network training approach can be applied by all source-destination pairs in a collocated network [6].

The main thrust of this paper is devoted to characterizing the achievable rates and scaling laws by using AF with network training in power-constrained relay networks in the wideband regime. Simply put, the scaling law here is concerned with the order of the achievable rates as the number of the relay nodes n grows, i.e., the scaling behavior of the throughput (and capacity) as $n \to \infty$. Let ρ denote the SNR per link. It is understood that $\rho \to 0$ in the wideband regime. Furthermore, it is clear that the coherence interval of the fading channel, denoted as L, plays a key role in channel estimation. To obtain a clear understanding of the continuum between coherent relaying and non-coherent relaying, we examine the achievable rates and scaling laws in the joint asymptotic regime of the number of relay nodes n, the channel coherence interval L, the SNR per link ρ (or the bandwidth W). We note that although there is no physical connection among these parameters, the achievable rates depend on how they scale together, rather than on each of them in isolation [18, 26].

Since a wideband channel can be decomposed into multiple orthogonal narrowband channels, each with flat fading, we first investigate achievable rates and scaling laws by using AF relaying with network training in narrowband relay networks in the low SNR regime [25]. Our findings show that when ρL (which is proportional to the transmission energy per fading block) is bounded below, AF with network training achieves the same scaling law as in the coherent relay networks where the CSI is known a priori. We then generalize the study to power-constrained wideband relay networks in which frequency-selective fading is inevitable. We model the wideband relay networks as a set of multiple narrowband relay networks under an overall power constraint. We examine the achievable rates and scaling laws by using AF with network training, under two power allocation policies, respectively. Specifically, we first assume

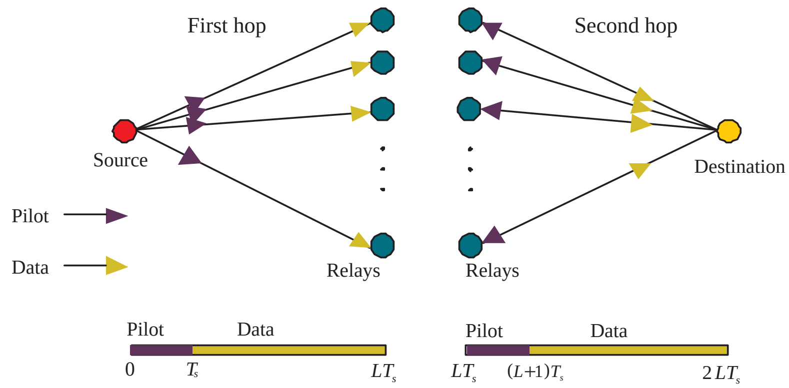

Figure 2: Amplify-and-forward with network training

that an equal power allocation policy across the frequency subbands[1] is applied at each relay node. Then we examine the optimal power allocation policy for wideband relay networks, where each relay node is allowed to allocate its power across the subchannels and the fading blocks. In general, one would expect a higher achievable rate by using bursty flashing signals (more generally, water-filling techniques). Somewhat surprising, our findings show that the achievable rates by using AF with network training have the same scaling order as those given by the equal power allocation policy, and both policies can achieve the scaling law under the condition that L/W is bounded below and that W is sub-linear in n. That is to say, the equal power allocation policy at relay nodes is order optimal in the sense of achieving the scaling law asymptotically. Roughly speaking, the condition that W is sub-linear in n is equivalent to that the degree of node multiplexing per unit bandwidth (n/W) grows as n increases.

In related work, the capacity of wireless networks and the capacity of relay channels have been studied extensively in the past few years (see, [1, 4, 5, 7–13, 16, 17, 20, 22, 23] and the references therein). In their seminal work [10], Gupta and Kumar investigate the throughput scaling for dense many-to-many multi-hop network models, and they prove that an aggregated throughput of $O(\sqrt{n})$ can be achieved if nodes can be arbitrarily located. Grossglauser and Tse study the impact of mobility on the throughput in mobile ad hoc networks [9]. They show a $O(n)$ throughput can be achieved by using two-hop relaying and allowing possibly unbounded delay. Recently, wireless sensory relay networks has received much attention. In [7,8], Gastpar and Vetterli show that the asymptotic capacity of wireless relay networks behaves like $\log(n)$ under the relay traffic pattern, assuming that power allocation can be carried out across the relay nodes. Bölcskei, Nabar, Oyman and Paulraj find in [1] that the throughput of wireless relay networks, where each node is equipped with M antennas, scales as $M\log(n)$. Recent work [4] by Dana and Hassibi reveals that the power efficiency of the AF scheme scales at least by a factor of $\sqrt{n}$. Very recent works [16, 17] by Oyman and Paulraj find that a higher power efficiency of order n can be achieved by using bursty transmission. Dousse, Franceschetti and Thiran discover in [5] that the per-node throughput remains constant as the size of the wireless

[1]Throughout, we use the notation of subchannel and subband interchangeably.

network increases, if a "small" fraction of nodes are allowed to be disconnected.

Most relevant to our work is perhaps [1,7,8], which study the scaling laws assuming coherent AF relaying. It is also pointed out in [1] that the achievable rates by using non-coherent AF relaying does not scale well at all. The study in this paper is intended to examine the continuum between non-coherent AF relaying and coherent AF relaying and to close the gap in between.

The rest of this paper is organized as follows. In Section 2, we present the models for narrowband relay networks in the low SNR regime and power-constrained wideband relay networks, respectively. We summarize in Section 3 the main results of this paper. In Section 4, we obtain an equivalent source-to-destination channel model for narrowband relay networks using AF with network training, and examine the scaling order of the achievable rates in different scenarios. We characterize the conditions for achieving the scaling laws accordingly. In Section 5, we generalize the study to power-constrained wideband relay networks. We analyze the achievable rates and the scaling order corresponding to the equal power allocation policy at the relay nodes, and compare the scaling behavior with that by using the optimal power allocation policy at the relay nodes. Section 6 contains numerical examples which are used to illustrate the achievable rates and capacity bounds. Section 7 concludes this paper.

2 System Model

In this section, we present the model for narrowband relay networks in the low SNR regime and that for power constrained wideband relay networks.

2.1 Narrowband Relay Networks in the Low SNR Regime

We impose the following assumptions for the narrowband relay network model.

- there are n relay nodes between the source node and the destination node. There is no direct link between the source node and the destination node. In the first hop, each relay node listens to the transmissions from the source node, and then forwards the processed signal to the destination node in the second hop;
- the communication bandwidth is B (Hz);
- the ambient noise at the relay nodes and the destination node has power spectral density N_0;
- all nodes have an average transmit power constraint P (watts) in one fading block;
- all channels experience i.i.d. frequency-flat block fading. Denote the backward channel coefficients as $\{\alpha_i\}$, and the forward channel coefficients as $\{\beta_i\}$. We assume that $\{\alpha_i\}$ and $\{\beta_i\}$ are complex Gaussian random variables, with zero mean and unit variance;
- the backward channels have a coherence interval of L symbol periods, so do the forward channels. Throughout, we assume that $B = 1/T_s$. Accordingly, the channel coherence time is LT_s (seconds);

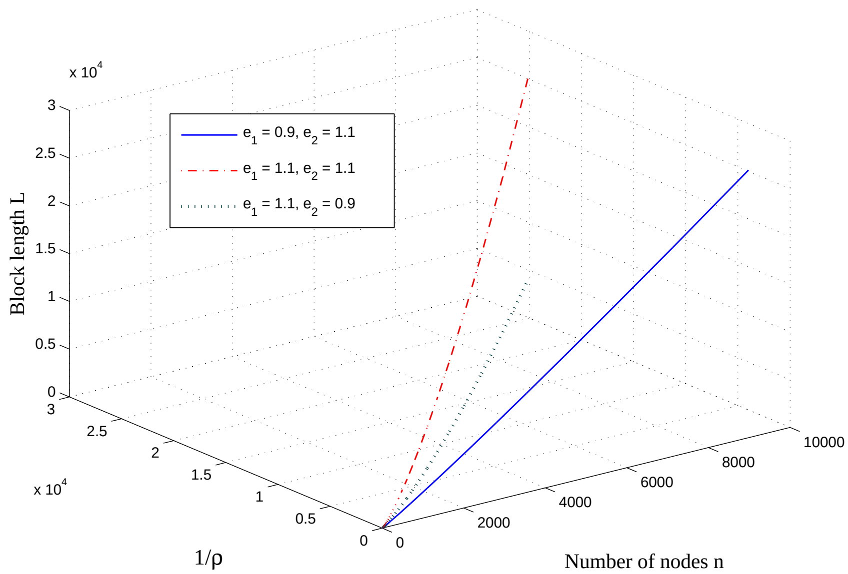

Figure 3: Asymptotic regime of $(n, 1/\rho, L)$

- within one fading block, the forward channel conditions are symmetric in both directions, that is, the channel condition from the relay node to the destination node is the same as that from the destination node to the relay node.

Without loss of generality, we assume that the average receive SNR for each backward (forward) link is P/BN_0, i.e, $\rho = P/(BN_0)$. It is understood that $\rho \to 0$ in the wideband regime. On the other hand, the coherence interval L points to how fast the channel changes. Clearly, a larger L is "favorable" for more accurate channel estimation, and we would expect that when L exceeds a certain threshold, the achievable rates of the relay networks can scale the same as in coherent relaying case where a priori knowledge of CSI is available at relay nodes. In contrast, a low SNR makes channel estimation more challenging.

To obtain a clear understanding of the continuum between coherent relaying and non-coherent relaying, we let ρ and L scale with n. We investigate the achievable rates in the joint asymptotic regime of n, L and ρ. Accordingly, we can obtain a set of achievable rates corresponding to different values of ρ and L. Define

$$e_1 \triangleq \frac{\log(\frac{1}{\rho})}{\log n} = -\frac{\log \rho}{\log n} \quad \text{and} \quad e_2 \triangleq \frac{\log L}{\log n}. \tag{1}$$

Remarks: Because the focus of this study is on the scaling law, i.e., the scaling order of the capacity, it is natural to define e_1 and e_2 on the logarithmic scale. Although there is no physical connection among these three parameters, the achievable rates and their scaling behavior depend on the relation among these three key parameters, rather than on each of them alone. Simply put, the asymptotic regime is along a curve determined by (n, e_1, e_2) in the three-dimensional space of $(n, 1/\rho, L)$, as illustrated in Fig. 3. Indeed, we show in Section 4 that the simultaneous scaling of the three parameters enables the characterization of the achievable rates and their scaling in terms of how they are related through the exponents.

2.2 Amplify-and-Forward (AF) with Network Training

As aforementioned, the channel estimation at relay nodes is carried out in the network training phase (see Fig. 2). More specifically, the source node first broadcasts common pilot symbols to relay nodes at the beginning of the first hop transmission, followed by its data transmission. Then the destination node broadcasts common pilot symbols at the beginning of the second hop. Based on the received signals corresponding to the pilot symbols, each relay nodes estimates its backward channel condition and its forward channel condition, respectively. After the channel estimation is done, each relay node carries out phase alignment by using its channel estimates, amplifies the received data signal under the given power constraints, and forwards the processed signal to the destination node.

2.2.1 Channel Estimation via Network Training

We assume that the minimum mean-square error (MMSE) estimation method is applied to estimate $\{\alpha_i\}$ and $\{\beta_i\}$. Let E_{total} denote the total energy for one hop transmission over L symbols, and E_{tr} be the energy for the training with $E_{tr} = \eta E_{total}$. Since the estimate error of the MMSE estimator depends only on E_{tr} [19, p. 599], we further assume that the training can be done using one pilot sample (symbol). We assume that $L \gg 1$.

Consider the channel estimation for the i-th channel in the first hop:

$$Y_i^p = \sqrt{E_{tr}}\alpha_i + Z_i \ , \tag{2}$$

where Y_i^p is the received pilot symbol, and $Z_i \sim \mathcal{CN}(0, \ N_0)$ is the noise at the i-th relay node. Let $\tilde{\alpha_i}$ be the estimate error and $\hat{\alpha}_i$ be the channel estimator at the i-th relay node., i.e., $\tilde{\alpha_i} \triangleq \alpha_i - \hat{\alpha}_i$. By the orthogonality principle of the MMSE estimation, the estimate error and the channel estimator are uncorrelated. Since all channel coefficients are complex Gaussian, the estimate error is in fact *independent* from the channel estimator. Recall that $T_s = 1/B$, and thus $E_{tr} = \eta LP/B$. We conclude that the variance of $\hat{\alpha}_i$ is

$$\mathrm{var}(\tilde{\alpha_i}) = \frac{N_0}{\frac{\eta LP}{B} + N_0} = \frac{1}{\eta\rho L + 1}. \tag{3}$$

Similarly, we have that the variance of $\hat{\beta}_i$ is

$$\mathrm{var}(\tilde{\beta_i}) = \frac{1}{\eta\rho L + 1}. \tag{4}$$

2.2.2 Amplify-and-Forward at Relay Nodes

In what follows, we establish the channel and signal models corresponding to AF with network training. Recall that the network training is done via using one pilot symbol in the first hop and another one in the second hop. During the data transmission period in the first hop, the information signal received by the i-th relay node is given by

$$Y_i(t) = \alpha_i X(t) + Z_i(t), \quad \text{for} \quad T_s \leq t \leq LT_s. \tag{5}$$

where $X(t)$is the transmitted signal from the source node, and $Z_i(t)$ is complex Gaussian noise at the i-th relay node. In the second hop, the i-th relay node applies AF with network training to relay the received signal $Y_i(t)$. In particular, it carries out phase alignment via multiplying $Y_i(t)$ with $\frac{\hat{\alpha_i}^* \hat{\beta_i}^*}{|\hat{\alpha_i}||\hat{\beta_i}|}$. The corresponding transmitted signal at the i-th relay node is given by

$$S_i(t) = A_i[\alpha_i X(t - LT_s) + Z_i(t - LT_s)] \ , \quad \text{for} \quad (L+1)T_s \leq t \leq 2LT_s \tag{6}$$

where A_i is the amplification factor to meet the power constraint for each relay node, and is given by

$$A_i = \frac{\hat{\alpha_i}^* \hat{\beta_i}^*}{|\hat{\alpha_i}||\hat{\beta_i}|} \sqrt{\frac{P}{|\alpha_i|^2 \frac{P(1-\eta)L}{L-1} + BN_0}} \ , \tag{7}$$

The destination node collects signals from n relay nodes, and the received signal is given as below:

$$\begin{aligned} Y(t) &= \sum_{i=1}^{n} \beta_i S_i(t) + Z(t) \\ &= \sum_{i=1}^{n} \beta_i A_i \alpha_i X(t - LT_s) + \sum_{i=1}^{n} \beta_i A_i Z_i(t - LT_s) + Z(t), \ \text{for} \ (L+1)T_s \leq t \leq 2LT_s, \end{aligned} \tag{8}$$

where $Z(t)$ is complex Gaussian noise at the destination node. Moreover, as $\rho \to 0$, α_i, $BN_0 \gg |\alpha_i|^2 \frac{P(1-\eta)L}{L-1}$ with probability 1, and hence in the low SNR regime,

$$A_i \approx \frac{\hat{\alpha_i}^* \hat{\beta_i}^*}{|\hat{\alpha_i}||\hat{\beta_i}|} \sqrt{\rho} \ . \tag{9}$$

2.3 Power-Constrained Wideband Relay Networks

In wideband relay networks, frequency-selective fading becomes inevitable when the frequency bandwidth increases. To study the scaling laws of wideband relay networks, we investigate an equivalent model where each wideband channel is decomposed into a set of frequency-flat sub-channels. Let the bandwidth of each subchannel in wideband relay networks be B and the total number of subchannels is K. Then, the overall frequency bandwidth $W = KB$. Define

$$e_3 \triangleq \frac{\log K}{\log n}. \tag{10}$$

Since B is fixed, we have that the total frequency bandwidth $W = \Theta(K) = \Theta(n^{e_3})$. Again, we are interested in how the three parameters n, L and W jointly determine the achievable rates; and the above setting is used to obtain a clear understanding how the simultaneous scaling of the three key parameters impacts the scaling laws.

We have a few words on the exponents e_1 and e_3. Recall that

$$e_1 \triangleq \frac{\log(\frac{1}{\rho})}{\log n} = \frac{-\log \frac{P}{N_o} + \log B}{\log n}.$$

It is clear that e_1 is intimately related to e_3. Indeed, the "low SNR" in the narrowband relay network can be viewed as the result of a large bandwidth and fixed power. On the other hand, it should be cautioned that the wideband relay network model is not a simple aggregation of many narrowband relay network models, because power allocation across the subchannels at relay nodes does play an important role to determine the capacity. We elaborate further in Section 5 the power allocation across the subchannels.

Due to frequency-selective fading, the backward (forward) links may have different channel statistics in different frequency subbands. To take this into account, we assume that for the k-th subchannel,

- let $\{\alpha_{i,k}\}$ and $\{\beta_{i,k}\}$ denote the i.i.d. Rayleigh fading coefficients corresponding to the backward link and the forward link, respectively. Assume that $\alpha_{i,k}$ and $\beta_{i,k}$ have zero mean and common variance σ_k^2;
- assume that all relay nodes have an average power constraint p_k within any fading block (for the k-th subchannel);
- assume that the source node allocates its average power uniformly across the subchannels within one fading block, namely, P/K for each subchannel.
- denote the portion of energy allocated for training as η_k, which could be different across the subchannels.

We assume uniform power allocation at the source node, mainly because the optimal power allocation at the source node requires the knowledge of the channel gains of all subchannels and also depends on $\{p_k\}$'s. It is clear that the optimization of $\{p_k\}$'s is highly non-trivial and it is unlikely to be available in practical systems.

Letting X_k be the transmitted signal in the k-th subband from the source node, we have that the received signal at the destination node in the k-th subband is given by (cf. (8))

$$Y(t) = \sum_{i=1}^{n} \beta_{i,k} A_{i,k} \alpha_{i,k} X_k(t - LT_s) + \sum_{i=1}^{n} \beta_{i,k} A_{i,k} Z_{i,k}(t - LT_s) + Z_k(t) \ , \tag{11}$$

for $(L+1)T_s \le t \le 2LT_s$. The amplification factor in the k-th subband at the i-th relay node $A_{i,k}$ is given as

$$A_{i,k} \approx \frac{\hat{\alpha}_{i,k}^* \hat{\beta}_{i,k}^*}{|\hat{\alpha}_{i,k}||\hat{\beta}_{i,k}|} \sqrt{\frac{p_k}{BN_0}}. \tag{12}$$

Along the same line as in the study for the narrowband relay network model, we have that for the k-th subchannel, the variance of $\hat{\alpha}_{i,k}$ and $\hat{\beta}_{i,k}$ is

$$\mathrm{var}(\tilde{\alpha}_{i,k}) = \frac{\sigma_k^2 W N_0}{\eta_k \sigma_k^2 P L + W N_0}, \quad (13)$$

$$\mathrm{var}(\tilde{\beta}_{i,k}) = \frac{\sigma_k^2 W N_0}{\eta_k \sigma_k^2 P L + W N_0}. \quad (14)$$

3 Summary of Main Results on Scaling Laws and Achievable Rates

For convenience, we introduce the following natation for asymptotic order relationship. For two sequences $\{f(n)\}$ and $\{g(n)\}$, we write

$$f(n) \overset{\bullet}{\geq} g(n) \quad (15)$$

if

$$\liminf_{n\to\infty} \frac{f(n)}{g(n)} \geq a$$

where $a > 0$ is some positive constant, and vice verse. Intuitively speaking, $f(n) \overset{\bullet}{\geq} g(n)$ indicates that $f(n)$ grows at least at the same rate as $g(n)$.

In the following, we summarize the main results on the scaling laws in the joint asymptotic regimes.

Theorem 3.1 *[**Narrowband Relay Networks**] If there exists $\zeta \geq 0$ such that $\rho L \overset{\bullet}{\geq} n^{\zeta}$, then as $n \to \infty$, the capacity of the narrowband relay networks in the low SNR regime scales as*

$$C = \begin{cases} \Theta(\log(n\rho)) & \text{if} \quad \liminf_{n\to\infty} \log(n\rho) > 0 \\ \mathrm{O}(1) & \text{if} \quad \limsup_{n\to\infty} \log(n\rho) \leq 0 \end{cases}. \quad (16)$$

Theorem 3.2 *[**Wideband Relay Networks**] a). If there exist $\delta \geq 0$ and $0 < \epsilon < 1$ such that $\frac{L}{W} \overset{\bullet}{\geq} n^{\delta}$ and $W \overset{\bullet}{\leq} n^{1-\epsilon}$, then as $n \to \infty$, the capacity of power-constrained wideband relay networks scales as*

$$C^{WB} = \Theta\left(W \log\left(\frac{n}{W}\right)\right). \quad (17)$$

b). The equal power allocation policy at relay nodes can achieve the same scaling order as that by the optimal power allocation policy at relay nodes.

Remarks on scaling laws: Theorem 3.1 provides a sufficient condition to achieve the scaling law of the narrowband relay network model in the joint asymptotic regime of n, L and ρ. Roughly speaking, the scaling law can be characterized by (16) provided that the transmission energy per fading block (proportional to ρL) is bounded below as the number of relay nodes n grows. That is to say, "almost" coherent relaying can be made possible based on good channel estimation, provided that the amount of energy available within each fading block is at least

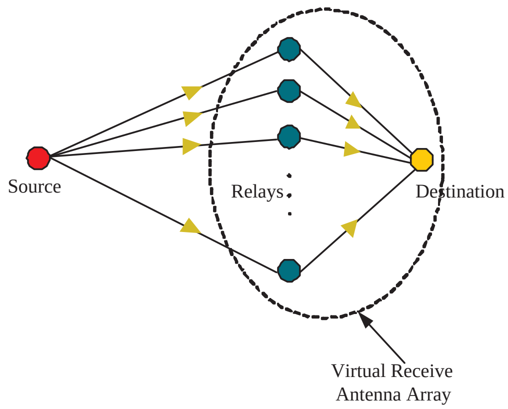

Figure 4: W is sublinear in n

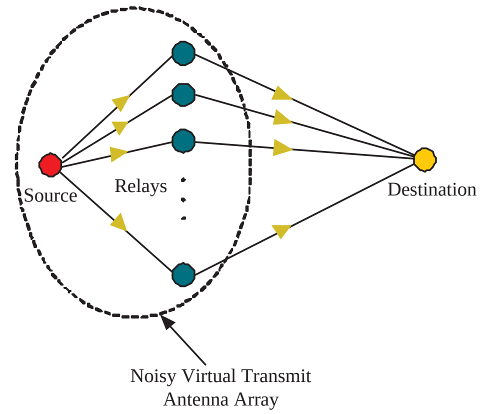

Figure 5: W is suplinear in n

constant. Observe that $\log(n\rho) = (1 - e_1)\log(n)$. It turns out that if ρ is bounded below, the capacity of narrowband relay networks is of $\Theta(\log(n))$ (cf. [7]).

Theorem 3.2 provides a sufficient condition to achieve the scaling law for power-constrained wideband relay networks in the joint asymptotic regime of n, L and W. Since $L/W = LT_sP/K$, we can view L/W as the (normalized) energy per fading block in each subband, and n/W as the degree of node diversity per unit bandwidth. Intuitively speaking, the above condition reveals that if the energy per fading block in each subband, proportional to L/W, is bounded below and W is sub-linear in n, then the scaling law can be characterized by (17). Note that under the above conditions, (17) indicates that the capacity of wideband relay networks is of $\Theta(n^{1-e_3}\log(n))$.

The conditions in Theorem 3.2 suggest that when W is sublinear in n, i.e., $W \overset{\bullet}{<} n$, the wideband relay networks using AF with network training yields achievable rates of the same scaling order as that of the Cut-Set upper bound. The underlying rationale is that the scaling law for the wideband relay networks using AF relaying is limited by the node diversity order n. In general, in a power-constrained wideband relay network using AF relaying, a larger bandwidth would lead to a smaller amplification factor for relaying, which in turn may reduce the power of the desired signal embedded in the amplified signal from each relay node, and therefore the equivalent overall SNR at the destination node may or may not increase, depending on how W and n scale together. It can be shown that the power of the aggregated noise from the relay nodes is of $\Theta(n)$ because every relay node has an average power constraint P, and the power of the ambient noise at the destination node is of $\Theta(W)$. In the case when W is sublinear in n, the aggregated "amplified" noise from all the relay nodes asymptotically dominates the ambient noise at the destination node, and all forward links are virtually "noise-free". Such "noise-free" links effectively convert all the relay nodes and the destination node into an antenna array of n receive antennas, as shown in Fig. 4. Observe that the relay nodes carry out phase alignment, and as a result, the equivalent virtual antenna array effectively conducts maximal-ratio combining (MRC). Thus, the relay network functions as if it were a single-input-multiple-output (SIMO) system with optimal decoding at the receiver. We note that the upper bound on the capacity of the relay network can be obtained by using the Cut-Set Theorem which corresponds to the capacity of a SIMO system when receive CSI is available. Then it can be expected that under the "noise-free" condition mentioned above, the scaling law can be achieved.

In contrast, when W is suplinear in n, the ambient noise at the destination node asymptotically dominates the aggregated "amplified" noise from all relay nodes. More specifically, it turns out that in this case the smaller amplification factor at the relay nodes would decrease the overall SNR at the destination node as n and W grow together, indicating that the achievable rate of AF does not scale well. We also note that in this case the source node and all relay nodes virtually function as a "noisy" antenna array of n transmit antennas, as depicted in Fig. 5; and the virtual antenna array is "noisy" in the sense that the transmitted signals from the relay nodes consist of the desired signals and the amplified noise.

In summary, the scaling law can be achieved if W is sublinear in n, i.e., $W \overset{\bullet}{<} n$. When W is suplinear in n, i.e., $W \overset{\bullet}{>} n$, there exits a gap between the scaling order of the upper bound and that of the achievable rates. For the case when W is linear in n, the achievable rate would depend on other parameters. We elaborate further on this in Sections 4.5 and 5.4.

The second part of Theorem 3.2 indicates that AF with network training, under simple equal power allocation across the frequency subbands at the relay nodes, can achieve the scaling law. Our intuition is as follows. When the AF relaying strategy is applied at a relay node, it "amplifies" both the signals and the received noise simultaneously, regardless of what power allocation scheme is applied. As a result, the optimal power allocation scheme would achieve the scaling order under the same conditions as the equal power allocation one. We elaborate further on this in Section 5.3.

Remarks on achievable rates: We have also investigated the achievable rates and their scaling behavior when the conditions for achieving scaling laws in Theorems 3.1 and 3.2 are not satisfied. In particular, for the narrowband relay network model, when $e_1 > e_2$, the corresponding achievable rates are summarized in Proposition 4.2. For the wideband relay network model, when $e_3 > e_2$, the corresponding achievable rates under equal power allocation at relay nodes can be found in Section 5.2. We show in Sections 4 and 5 that in both cases the estimate error cannot be made small and that the corresponding scaling orders lie in between the scaling order for non-coherent relaying and that for coherent relaying.

4 Narrowband Relay Networks in the Low SNR Regime

4.1 Equivalent Source-to-Destination Channel Model

In the following, we first study the equivalent source-to-destination model for the narrowband relay networks. Recall that in the first hop, the information data signal received by the i-th relay node is given by

$$Y_i(t) = \alpha_i X(t) + Z_i(t) \ , \quad \text{for} \quad T_s \le t \le LT_s. \tag{18}$$

In the second hop, the i-th relay node applies AF with network training to relay the received signal $Y_i(t)$. In particular, it carries out phase alignment, and the corresponding transmitted signal at the i-th relay node is given by

$$S_i(t) = A_i[\alpha_i X(t - LT_s) + Z_i(t - LT_s)] \ , \quad \text{for} \quad (L+1)T_s \le t \le 2LT_s, \tag{19}$$

where A_i is the amplification factor to meet the power constraint for each relay node, and is given by

$$A_i = \frac{\hat{\alpha}_i{}^* \hat{\beta}_i{}^*}{|\hat{\alpha}_i||\hat{\beta}_i|} \sqrt{\frac{P}{|\alpha_i|^2 \frac{P(1-\eta)L}{L-1} + BN_0}} \ . \tag{20}$$

The destination node collects signals from n relay nodes, and the received signal is given as:

$$\begin{aligned} Y(t) &= \sum_{i=1}^{n} \beta_i S_i(t) + Z(t) \\ &= \sum_{i=1}^{n} \beta_i A_i \alpha_i X(t - LT_s) + \sum_{i=1}^{n} \beta_i A_i Z_i(t - LT_s) + Z(t), \ \text{for} \ (L+1)T_s \le t \le 2LT_s. \end{aligned} \tag{21}$$

After some algebra, we can rewrite (21) as

$$\begin{aligned} Y(t) = \quad & \sum_i \sqrt{\rho} |\hat{\alpha}_i||\hat{\beta}_i| X(t - LT_s) \\ & + \underbrace{\sum_i \sqrt{\rho} \left(\frac{|\hat{\alpha}_i|^2 \tilde{\beta}_i \hat{\beta}_i{}^* + |\hat{\beta}_i|^2 \tilde{\alpha}_i \hat{\alpha}_i{}^* + \tilde{\alpha}_i \hat{\alpha}_i{}^* \tilde{\beta}_i \hat{\beta}_i{}^*}{|\hat{\alpha}_i||\hat{\beta}_i|} \right) X(t - LT_s)}_{\Omega_1(t)} \\ & + \underbrace{\sum_i \sqrt{\rho} \left(\frac{\hat{\alpha}_i{}^* |\hat{\beta}_i|^2 + \hat{\alpha}_i{}^* \tilde{\beta}_i \hat{\beta}_i{}^*}{|\hat{\alpha}_i||\hat{\beta}_i|} \right) Z_i(t - LT_s)}_{\Omega_2(t)} \\ & + Z(t), \quad \text{for} \quad (L+1)T_s \le t \le 2LT_s \ . \end{aligned} \tag{22}$$

In the above signal model, the received signal at the destination node consists of

- the information-bearing signal, $\sum_i \sqrt{\rho}|\hat{\alpha}_i||\hat{\beta}_i| X(t - LT_s)$,
- the signal due to the estimate error, $\Omega_1(t)$,
- the aggregated amplified noise from all relay nodes, $\Omega_2(t)$,
- and the noise at the destination node, $Z(t)$.

We observe that in the above model the information is embedded in the received signal with unknown parameters, because the destination node does not have the knowledge of the instantaneous CSI of the backward and forward channels. Define $H \triangleq \sqrt{\rho} \sum_i |\hat{\alpha}_i||\hat{\beta}_i|$. We can rewrite the equivalent source-to-destination model as

$$Y(t) = \mathbb{E}[H] X(t - LT_s) + \Omega_1(t) + \Omega_2(t) + \Omega_3(t) + Z(t), \tag{23}$$

where $\Omega_3(t) \triangleq (H - \mathbb{E}[H])X(t - LT_s)$. In this equivalent source-to-destination channel model (23), it is clear that $\mathbb{E}[H]$ can be easily made available at the destination node, because it is determined by the channel statistics only. Furthermore, both $\Omega_1(t)$ and $\Omega_3(t)$ are signal-dependent, and $\Omega_2(t)$ is non-Gaussian. This equivalent model is then a deterministic channel with non-Gaussian signal-dependent noise, provided that $\Omega_1(t)+\Omega_2(t)+\Omega_3(t)+Z(t)$ is treated as the aggregated noise. In general, it is non-trivial to characterize the exact capacity of the channel model (23).

In what follows, we turn to characterize an achievable rate for (23) instead, which serves as a lower bound. Observe that the channel side information is imperfect in the channel model in (23), and the mutual information between the input and the output depends heavily on the variance of the channel estimate error. Along the lines of [15], an achievable rate is given by (we relegate the detailed derivation to the Appendix B)

$$C \geq B\frac{L-1}{L}\log\left(1+\frac{\mathbb{E}\left[|\mathbb{E}[H]X(t-LT_s)|^2\right]}{\mathbb{E}[|\Omega_1(t)|^2]+\mathbb{E}[|\Omega_2(t)|^2]+\mathbb{E}[|\Omega_3(t)|^2]+WN_0}\right). \tag{24}$$

The RHS of (24) is an achievable rate for the equivalent model in (23). Interestingly, it can be obtained by treating $\Omega_1(t) + \Omega_2(t) + \Omega_3(t) + Z(t)$ as white Gaussian noise with the same variance. Note that $\Omega_2(t) + Z(t)$ is a dominating term in the equivalent noise due to the low SNR. As noted in Section 3, if $\Omega_2(t)$ is dominant, it can be shown that the relay networks using AF with network training act as if it were a single-input-multiple-output (SIMO) system.

For notational convenience, define

$$S \triangleq \frac{\mathbb{E}\left[|\mathbb{E}[H]X(t-LT_s)|^2\right]}{\mathbb{E}[|\Omega_1(t)|^2]+\mathbb{E}[|\Omega_2(t)|^2]+\mathbb{E}[|\Omega_3(t)|^2]+WN_0} \tag{25}$$

as the equivalent receive SNR corresponding to the equivalent model in (23). It remains to characterize the power of the desired signal term and the noise terms in (23). After some algebra (the detailed derivations have been relegated to Appendix B), we have that

$$S = \frac{n\frac{\pi^2}{16}\left(\frac{\eta\rho L}{\eta\rho L+1}\right)^2\frac{(1-\eta)L}{L-1}P}{\left(1-\frac{\pi^2}{16}\left(\frac{\eta\rho L}{\eta\rho L+1}\right)^2\right)\frac{(1-\eta)L}{L-1}P+BN_0+\frac{(BN_0)^2}{nP}}\,. \tag{26}$$

Assuming that the coherence interval $L \geq 2$, we have that

$$0 \leq \left(1-\frac{\pi^2}{16}\left(\frac{\eta\rho L}{\eta\rho L+1}\right)^2\right)\frac{(1-\eta)L}{L-1}P \leq 2P\,. \tag{27}$$

Since the receive SNR per link ρ, is small, we have that

$$\left(1-\frac{\pi^2}{16}\left(\frac{\eta\rho L}{\eta\rho L+1}\right)^2\right)\frac{(1-\eta)L}{L-1}P = \mathrm{o}\left(\mathrm{BN_0}+\frac{(\mathrm{BN_0})^2}{\mathrm{nP}}\right)\,. \tag{28}$$

Therefore, we conclude that when $n \to \infty$,

$$S = \Theta\left(\frac{n\frac{\pi^2}{16}\left(\frac{\eta\rho L}{\eta\rho L+1}\right)^2\frac{(1-\eta)L}{L-1}\rho}{1+\frac{1}{n\rho}}\right)\,. \tag{29}$$

Then, the achievable rate of narrowband relay networks using AF with network training at relay nodes is given by

$$R = B\frac{L-1}{L}\log(1+S) \tag{30}$$

$$= \Theta\left(\log\left(1+\frac{n\frac{\pi^2}{16}\left(\frac{\eta\rho L}{\eta\rho L+1}\right)^2\frac{(1-\eta)L}{L-1}\rho}{1+\frac{1}{n\rho}}\right)\right). \tag{31}$$

It is clear from (29) that S depends on n, L and ρ through $n\rho$ and ρL. Since $\rho L = LPT_s/N_0$, we note that ρL can be regarded as the normalized transmission energy per fading block, and $\eta\rho L$ corresponds to the normalized energy for training per fading block. Observe $\rho n = nPT_s/N_0$, where nPT_s can be viewed as the total energy across the relay nodes to relay one symbol.

4.2 An Upper Bound on the Capacity

We next present an upper bound on the capacity of narrowband relay networks in the low SNR regime. By the Cut-Set Theorem in [3], the capacity of the relay network is upper bounded by the information rate corresponding to the "broadcast cut", regardless of the transmission schemes. That is,

$$C < 2B\log\left(1+\frac{\sum_i|\alpha_i|^2P}{BN_0}\right). \tag{32}$$

Define $C_{upper} \triangleq 2B\log\left(1+\frac{\sum_i|\alpha_i|^2P}{BN_0}\right)$. By the Law of Large Numbers, $\sum_i|\alpha_i|^2$ grows linearly in n as $n\to\infty$, it follows that C_{upper} is $\Theta(\log(1+n\rho))$. Accordingly, we have the following lemma.

Lemma 4.1 *As $n\to\infty$, the upper bound on the capacity of narrowband relay networks, C_{upper}, scales as*

$$C_{upper} = \begin{cases}\Theta(\log(n\rho)) & \textit{if} \ \ \liminf_{n\to\infty}\log(n\rho) > 0 \\ \mathrm{O}(1) & \textit{if} \ \ \limsup_{n\to\infty}\log(n\rho) \le 0\end{cases}. \tag{33}$$

4.3 The Perfect CSI Case

Needless to say, the highest possible achievable rate of the AF strategy occurs when there is perfect CSI at the relay nodes, i.e., the i-th relay node has a priori knowledge on α_i and β_i. The corresponding achievable rate is given by [4]

$$R_{pft} = B\log\left(1+\frac{\frac{\pi^2}{16}n\rho}{1+\frac{1}{n\rho}}\right). \tag{34}$$

Note that the above result can also be obtained by letting $\mathrm{var}(\tilde{\alpha})=0$, $\mathrm{var}(\tilde{\beta})=0$ and $L\to\infty$ while we prove (31).

We have the following lemma on R_{pft}.

Lemma 4.2 *If perfect CSI on the backward channels and the forward channels is available at the corresponding relay node, then as $n \to \infty$, the achievable rate by using AF scales as*

$$R_{pft} = \begin{cases} \Theta(\log(n\rho)) & if \quad \liminf_{n\to\infty} \log(n\rho) > 0 \\ \mathrm{O}(1) & if \quad \limsup_{n\to\infty} \log(n\rho) \leq 0 \end{cases} . \tag{35}$$

Note that in the perfect CSI case, the achievable rate (a lower bound on the capacity) has the same scaling order as the upper bound, thus yielding the scaling law. In what follows, we investigate the achievable rates of narrowband relay network in the low SNR regime, where there is no a priori knowledge on the CSI at the relay nodes.

4.4 Optimal Energy Allocation for Channel Estimation

As aforementioned, $E_{tr} = \eta E_{total}$ is allocated for training. It is clear that $\eta^* < 1$, and that the larger η is, the more accurate the channel estimation would be, at the cost of the energy for data transmission. That is to say, there is a trade-off between the energy for training and that for data transmission. It is easy to see that the optimal η is the one that maximizes the equivalent receive SNR in (29). Hence, taking derivative of S with respect to η, we have that the optimal η is given by

$$\eta^* = \frac{-3 + \sqrt{8\rho L + 9}}{2\rho L} , \tag{36}$$

where $\rho L = PLT_s/N_0$, as aforementioned, can be viewed as the total energy available within each coherence time, normalized with respect to N_0. Since $1/\rho$ and L can grow unbounded as $n \to \infty$, ρL can range from 0 to ∞. We investigate the achievable rates for the following three cases:

- I) ρL grows unbounded (equivalently $e_1 < e_2$);
- II)ρL approaches zero (equivalently $e_1 > e_2$);
- III) ρL is of $\Theta(1)$ (equivalently $e_1 = e_2$).

4.5 The Achievable Rates of Narrowband Relay Networks in the Low SNR Regime

4.5.1 The Case with $e_1 < e_2$

In this case, the coherence interval L grows faster than $1/\rho$. As a result, the energy per fading block ρL increases as n grows, indicating that there exists $\epsilon > 0$, such that

$$\rho L \overset{\bullet}{\geq} n^{\epsilon}. \tag{37}$$

It can be shown that

$$\eta^* = \Theta\left(n^{\frac{e_1 - e_2}{2}}\right) \qquad \text{and} \qquad \eta^* \rho L = \Theta\left(n^{\frac{e_2 - e_1}{2}}\right). \tag{38}$$

(38) reveals that when $e_1 < e_2$, the energy for training within each fading block also grows as $n \to \infty$, even though $\eta^* \to 0$. Indeed, $\eta\rho L = \Theta(\sqrt{\rho \mathrm{L}})$, and both the total energy and the energy for training per fading block grows unbounded. Since the MMSE estimation depends on $\eta\rho L$ only, it follows that channel estimation can be accurate in this scenario, thereby enabling coherent relaying.

In light of the results in (35) for the perfect CSI case, we expect that the achievable rates are different for different values of e_1. Indeed, if $e_1 < 1$, it can be shown that

$$\begin{aligned} \lim_{n\to\infty} S &= \lim_{n\to\infty} n\frac{\pi^2}{16}\left(\frac{\eta\rho L}{\eta\rho L+1}\right)^2 \frac{(1-\eta)L}{L-1}\rho \\ &= \Theta\left(n^{1-e_1}\right) \ , \end{aligned} \tag{39}$$

indicating that the achievable rate of the relay network is

$$R = \Theta\left(\log(n\rho)\right) \ . \tag{40}$$

Interestingly, the achievable rate in the case where $e_1 < 1$ can scale the same as in the perfect CSI case. Note that if ρ is bounded below, the achievable rate is of $\Theta\left(\log(n)\right)$ (cf. [7]).

In contrast, if $e_1 \geq 1$, we have that $\rho \leq \frac{1}{n}$ as $n \to \infty$, i.e., the SNR goes to zero faster than $\frac{1}{n}$. Accordingly,

$$\begin{aligned} \lim_{n\to\infty} S &= \lim_{n\to\infty} \frac{n\frac{\pi^2}{16}\left(\frac{\eta\rho L}{\eta\rho L+1}\right)^2 \frac{(1-\eta)L}{L-1}\rho}{\frac{1}{n\rho}} \\ &= \Theta\left(n^{2-2e_1}\right) \ . \end{aligned} \tag{41}$$

It follows that the achievable rate for this case is

$$R = \mathrm{O}(1) \ . \tag{42}$$

Our intuition for this case is as follows: even if coherent relaying is possible in this case, the achievable rate would grow slower than $\log n$ because total power across the relay nodes $n\rho \,\dot{\leq}\, 1$.

The following proposition summarizes the results on the achievable rate by using AF with network training for the case $e_1 < e_2$.

Proposition 4.1 *If there exists $\epsilon > 0$, such that $\rho L \,\dot{\geq}\, n^{\epsilon}$, then as $n \to \infty$, the achievable rate by using AF with network training scales as*

$$R = \begin{cases} \Theta(\log(n\rho)) & \text{if} \quad e_1 < 1 \\ \mathrm{O}(1) & \text{if} \quad e_1 \geq 1 \end{cases} \ . \tag{43}$$

4.5.2 The Case with $e_1 > e_2$

In this scenario, $1/\rho$ grows faster than L. Accordingly, ρL approaches 0 as n goes to infinity. Along the same lines as above, we study the optimal energy allocation and the achievable rates for different values of e_1. It can be shown that the optimal energy allocation for using AF with network training is given by

$$\begin{aligned}\lim_{n\to\infty}\eta^* &= \lim_{n\to\infty}\left(\frac{-3+3\sqrt{1+\frac{8\rho L}{9}}}{2\rho L}\right)\\ &= \frac{2}{3}\ ,\end{aligned} \tag{44}$$

and

$$\eta^*\rho L = \Theta\left(n^{e_2-e_1}\right). \tag{45}$$

We observe that in this case $\eta\rho L = \Theta(\rho\mathrm{L})$, whereas $\eta\rho L = \Theta(\sqrt{\rho\mathrm{L}})$ for the case with $e_1 < e_2$. We note that in this case the total normalized energy per fading block goes to zero, so as the training energy.

When $e_1 < 1$, we have that

$$\begin{aligned}\lim_{n\to\infty} S &= \lim_{n\to\infty} n\frac{\pi^2}{16}\left(\frac{2}{3}n^{e_2-e_1}\right)^2\frac{(1-\frac{2}{3})L}{L-1}\rho\\ &= \Theta\left(n^{1+2e_2-3e_1}\right)\ .\end{aligned} \tag{46}$$

In contrast, if $e_1 \geq 1$, then

$$\begin{aligned}\lim_{n\to\infty} S &= \lim_{n\to\infty}\frac{n\frac{\pi^2}{16}\left(\frac{2}{3}n^{e_2-e_1}\right)^2\frac{(1-\frac{2}{3})L}{L-1}\rho}{\frac{1}{n\rho}}\\ &= \Theta\left(n^{2+2e_2-4e_1}\right)\ .\end{aligned} \tag{47}$$

In summary, we have the following proposition.

Proposition 4.2 *If $\rho L \to 0$ as $n \to \infty$, then the achievable rate by using AF with network training scales as*

- *If $e_1 < 1$, then*

$$R = \begin{cases}\Theta\left(\log(n^{1+2e_2-3e_1})\right) & \textit{for} \quad e_1 < \frac{1+2e_2}{3}\\ \mathrm{O}(1) & \textit{for} \quad e_1 \geq \frac{1+2e_2}{3}\end{cases}\ ; \tag{48}$$

- *if $e_1 \geq 1$, then*

$$R = \begin{cases}\Theta\left(\log(n^{2+2e_2-4e_1})\right) & \textit{for} \quad e_1 < \frac{1+e_2}{2}\\ \mathrm{O}(1) & \textit{for} \quad e_1 \geq \frac{1+e_2}{2}\end{cases}\ . \tag{49}$$

4.5.3 The Case with $e_1 = e_2$

This case is equivalent to $\rho L = \Theta(1)$. It can be shown that $\eta^*\rho L = \Theta(1)$, i.e.,, ρL and $\eta^*\rho L$ are bounded. Along the same line as above, we have that when $e_1 < 1$,

$$\begin{aligned}\lim_{n\to\infty} S &= \lim_{n\to\infty} n\frac{\pi^2}{16}\left(\frac{\eta^*\rho L}{\eta^*\rho L+1}\right)^2\frac{(1-\eta^*)L}{L-1}\rho\\ &= \Theta\left(n^{1-e_1}\right)\ .\end{aligned} \tag{50}$$

For the case $e_1 \geq 1$, it can be shown that

$$\lim_{n\to\infty} S = \lim_{n\to\infty} \frac{n\frac{\pi^2}{16}\left(\frac{\eta^*\rho L}{\eta^*\rho L+1}\right)^2 \frac{(1-\eta^*)L}{L-1}\rho}{\frac{1}{n\rho}} = \Theta\left(n^{2-2e_1}\right) . \quad (51)$$

Correspondingly, we have the following proposition.

Proposition 4.3 *If the normalized energy per fading block ρL is bounded, then as $n \to \infty$, the achievable rate by using AF with network training scales as*

$$R = \begin{cases} \Theta(\log(n\rho)) & \textit{if} \quad e_1 < 1 \\ \mathrm{O}(1) & \textit{if} \quad e_1 \geq 1 \end{cases} . \quad (52)$$

Remarks: Summarizing the above three cases in terms of e_1 and e_2, we conclude that if $e_1 \leq e_2$, the achievable rates have the same scaling order as the upper bound in (4.1), thereby yielding the scaling law in Theorem 3.1. More specifically, as $n \to \infty$, if there exist $\zeta \geq 0$ and $a > 0$, such that $\rho L \stackrel{\bullet}{\geq} an^\zeta$, then as $n \to \infty$, the capacity of the relay networks scales as

$$C = \begin{cases} \Theta(\log(n\rho)) & \text{if} \quad \lim_{n\to\infty} \inf \log(n\rho) > 0 \\ \mathrm{O}(1) & \text{if} \quad \lim_{n\to\infty} \sup \log(n\rho) \leq 0 \end{cases} . \quad (53)$$

5 Power-Constrained Wideband Relay Networks

Since a frequency-selective fading channel can be decomposed as multiple parallel frequency-flat fading subchannels [19], we first examine the equivalent source-to-destination channel model corresponding to the k-th subband in wideband relay networks. Along the same line of Section 4.1, the received signal in equivalent model for the k-th subchannel is given by

$$Y_k(t) = \mathbb{E}[H_k]X_k(t - LT_s) + \Omega_{1k}(t) + \Omega_{2k}(t) + \Omega_{3k}(t) + Z_k(t), \quad (54)$$

where

- $\mathbb{E}[H_k]X_k(t-LT_s)$ is the received signal used for information extracting, and $\mathbb{E}[H_k]$ denotes the equivalent channel gain;
- $\Omega_{1k}(t)$ is due to the estimate error;
- $\Omega_{2k}(t)$ corresponds to the aggregated amplified noise at the relay nodes;
- $\Omega_{3k}(t)$ stands for the received signal weighted by unknown channel strength.

In what follows, we characterize the achievable rate of each equivalent subchannel in (54), along the lines for narrowband relay network models. As aforementioned, the achievable rate can be obtained by treating $\Omega_{1k}(t)+\Omega_{2k}(t)+\Omega_{3k}(t)+Z_k(t)$ as equivalent Gaussian noise of the same power. Since the SNR per link in each subband is small, $\Omega_{2k}(t) + Z_k(t)$ is a dominating term in the equivalent noise. In particular, when $\Omega_{2k}(t)$ dominates, the relay network using AF with network training acts as if it were a single-input-multiple-output (SIMO) system.

For convenience, define

$$S_k \triangleq \frac{\mathbb{E}\left[|\mathbb{E}[H_k]X_k(t-LT_s)|^2\right]}{\mathbb{E}[|\Omega_{1k}(t)|^2]+\mathbb{E}[|\Omega_{2k}(t)|^2]+\mathbb{E}[|\Omega_{3k}(t)|^2]+WN_0} \tag{55}$$

as the equivalent overall SNR corresponding to the equivalent model in (54) Next, we examine the power of the terms in (54) (the derivation is relegated to Appendix B). After some algebra, the equivalent SNR in the equivalent network model for the k-th subband can be shown to be

$$S_k = \frac{n\sigma_k^4 \frac{\pi^2}{16}\left(\frac{\eta_k\sigma_k^2 PL}{\eta_k\sigma_k^2 PL+WN_0}\right)^2 \frac{(1-\eta_k)L}{L-1}\frac{P}{K}}{\sigma_k^4\left(1-\frac{\pi^2}{16}\left(\frac{\eta_k\sigma_k^2 PL}{\eta_k\sigma_k^2 PL+WN_0}\right)^2\right)\frac{(1-\eta_k)L}{L-1}\frac{P}{K}+\sigma_k^2 BN_0+\frac{(BN_0)^2}{np_k}} \,. \tag{56}$$

Note that as $K\rightarrow\infty$, $P/K \ll BN_0$. It follows that in the joint asymptotic regime of n, L and W,

$$S_k = \Theta\left(\frac{n\sigma_k^4 \frac{\pi^2}{16}\left(\frac{\eta_k\sigma_k^2 PL}{\eta_k\sigma_k^2 PL+WN_0}\right)^2 \frac{(1-\eta_k)L}{L-1}\frac{P}{K}}{\sigma_k^2 BN_0+\frac{(BN_0)^2}{np_k}}\right) \,. \tag{57}$$

The optimal η_k maximizing S_k is given by

$$\eta_k^* = \frac{-3+\sqrt{\frac{8\sigma_k^2 PL}{WN_0}+9}}{\frac{2\sigma_k^2 PL}{WN_0}} \,. \tag{58}$$

We emphasize that η_k^* can be different across different subbands, simply because their channel statistics are different.

In a nutshell, the achievable rate by using AF with network training in the power-constrained wideband relay network can be given as

$$R^{WB} = \sum_{k=1}^{K}\frac{B(L-1)}{L}\log\left(1+S_k\right) \,. \tag{59}$$

5.1 An Upper Bound on the Capacity

Similar to the narrowband case, by the Cut-Set Theorem, one upper bound on the capacity of the wideband relay networks is given by the information rate through the “broadcast-cut”, i.e.,

$$C^{WB} < \sum_{k=1}^{K}\left[2B\log\left(1+\frac{\sum_i|\alpha_{i,k}|^2\frac{P}{K}}{BN_0}\right)\right] \tag{60}$$

$$\overset{(a)}{\leq} 2W\log\left(1+\frac{1}{K}\frac{\sum_k\sum_i|\alpha_{i,k}|^2 P}{WN_0}\right) \,, \tag{61}$$

where (a) follows Jensen's Inequality. For convenience, define

$$C_{upper}^{WB} \triangleq 2W\log\left(1+\frac{1}{K}\frac{\sum_k\sum_i|\alpha_{i,k}|^2 P}{WN_0}\right) \,. \tag{62}$$

By the Laws of Large Numbers, we have that $\sum_k \sum_i |\alpha_{i,k}|^2$ increases linearly with respect to n and K in the asymptotic regime. Then it follows that

$$C_{upper}^{WB} = \Theta\left(W \log\left(1 + \frac{nP}{WN_0}\right)\right) \ . \tag{63}$$

Accordingly, we have the following lemma.

Lemma 5.1 *As $n \to \infty$, the scaling order of C_{upper}^{WB} is given by*

$$C_{upper}^{WB} = \begin{cases} \Theta(W \log(\frac{n}{W})) & \text{if} \ \ \liminf_{n\to\infty} \log(n/W) \geq 0 \\ \Theta(n) & \text{if} \ \ \limsup_{n\to\infty} \log(n/W) < 0 \end{cases} \ . \tag{64}$$

In the next section, we first study the network model where a simple power allocation policy, namely the equal power allocation policy, is applied at relay nodes.

5.2 Achievable Rates with the Equal Power Allocation Policy at Relay Nodes

When the equal power allocation policy is applied at relay nodes, the relaying power for the k-th subband p_k is simply P/K. From (54), the SNR corresponding to the k-th equivalent subchannel is

$$\begin{aligned} S_{Eq,k} \ &\triangleq \ \frac{n\sigma_k^4 \frac{\pi^2}{16}\left(\frac{\eta_k \sigma_k^2 PL}{\eta_k \sigma_k^2 PL + WN_0}\right)^2 \frac{(1-\eta_k)L}{L-1}\frac{P}{K}}{\sigma_k^2 BN_0 + \frac{(BN_0)^2}{n\frac{P}{K}}} \\ &= \ \frac{n\sigma_k^4 \frac{\pi^2}{16}\left(\frac{\eta_k \sigma_k^2 PL}{\eta_k \sigma_k^2 PL + WN_0}\right)^2 \frac{(1-\eta_k)L}{L-1} P}{\sigma_k^2 WN_0 + \frac{(WN_0)^2}{nP}} \ . \end{aligned} \tag{65}$$

Following the same line as in the study on the narrowband relay network model, we next investigate the scaling behavior of $S_{Eq,k}$ in three scenarios in terms of e_3 and e_2.

- If $e_3 < e_2$, then the optimal ratio of the energy allocation for training is given by

$$\lim_{n\to\infty} \eta_k^* = 0 \ . \tag{66}$$

  Observing that

$$\frac{\eta_k^* PL}{WN_0} = \Theta\left(n^{\frac{e_2 - e_3}{2}}\right) \ , \tag{67}$$

  we conclude that the equivalent SNR scales as

$$S_{Eq,k} = \begin{cases} \Theta\left(n^{1-e_3}\right) & \text{if} \quad e_3 < 1 \\ \Theta\left(n^{2-2e_3}\right) & \text{if} \quad e_3 \geq 1 \end{cases} \ . \tag{68}$$

- If $e_3 > e_2$, then the optimal energy allocation for training is given by

$$\lim_{n\to\infty} \eta_k^* = \frac{2}{3}. \tag{69}$$

Since

$$\frac{\eta_k^* PL}{WN_0} = \Theta\left(n^{e_2-e_3}\right) \ , \tag{70}$$

we conclude that the equivalent SNR scales as

$$S_{Eq,k} = \begin{cases} \Theta\left(n^{1+2e_2-3e_3}\right) & \text{if} \quad e_3 < 1 \\ \Theta\left(n^{2+2e_2-4e_3}\right) & \text{if} \quad e_3 \geq 1 \end{cases} . \tag{71}$$

- If $e_3 = e_2$, then L/W is bounded, i.e., $L/W = \Theta(1)$. It follows that

$$\frac{\eta_k^* PL}{WN_0} = \Theta(1) \ . \tag{72}$$

Then, the equivalent SNR scales as

$$S_{Eq,k} = \begin{cases} \Theta\left(n^{1-e_3}\right) & \text{if} \quad e_3 < 1 \\ \Theta\left(n^{2-2e_3}\right) & \text{if} \quad e_3 \geq 1 \end{cases} . \tag{73}$$

We observe that the scaling behavior of $S_{Eq,k}$ in (68), (71) and (73) satisfies that

$$S_{Eq,k} = \Theta(S_{Eq,l}), \quad k \neq l. \tag{74}$$

That is to say, in the asymptotic regime, when the equal power allocation policy is applied at relay nodes, the equivalent receive SNRs in different subchannels have the same scaling behavior. Combining (74) with (59), the achievable rate of wideband relay networks using the equal power allocation policy at relay nodes is given by

$$R_{Eq}^{WB} = \sum_{k=1}^{K} \frac{B(L-1)}{L} \log\left(1 + S_{Eq,k}\right) \ . \tag{75}$$

$$= \Theta\left(W \log\left(1 + S_{Eq,k}\right)\right) \ . \tag{76}$$

Summarizing the above, we conclude that

- if $e_3 \leq e_2$, then

$$R_{Eq}^{WB} = \begin{cases} \Theta\left(W \log(\frac{n}{W})\right) & \text{for} \quad e_3 \leq 1 \\ \Theta\left(\frac{n^2}{W}\right) & \text{for} \quad e_3 > 1 \end{cases} ; \tag{77}$$

- if $e_3 > e_2$ and $e_3 < 1$, then

$$R_{Eq}^{WB} = \begin{cases} \Theta\left(W \log\left(\frac{nL^2}{W^3}\right)\right) & \text{for} \quad e_3 < \frac{1+2e_2}{3} \\ \Theta\left(\frac{nL^2}{W^2}\right) & \text{for} \quad e_3 \geq \frac{1+2e_2}{3} \end{cases} ; \tag{78}$$

- if $e_3 > e_2$ and $e_3 \geq 1$, then

$$R_{Eq}^{WB} = \begin{cases} \Theta\left(W \log\left(\frac{n^2L^2}{W^4}\right)\right) & \text{for} \quad e_3 < \frac{1+e_2}{2} \\ \Theta\left(\frac{n^2L^2}{W^3}\right) & \text{for} \quad e_3 \geq \frac{1+e_2}{2} \end{cases} . \tag{79}$$

5.3 Achievable Rates with the Optimal Power Allocation Policy at Relay Nodes

Under equal power allocation, the relay nodes allocate power uniformly across the subbands within one fading block and there is no power allocation across fading blocks. A natural question to ask is that "If more flexibility is available, i.e., if each relay node is able to allocate its power across different subbands (in the frequency-domain) based on the channel statistics and across the fading blocks based on the instantaneous channel conditions (in the time-domain), would this outperform the equal power allocation policy?" In general, one would expect a higher achievable rate, and this is often the case by using bursty flashing signals (more generally, water-filling techniques). Thus motivated, we next examine the optimal power allocation policies, where each relay node allocates its power across the fading blocks and across the subbands, while keeping an overall long-term power constraint across the subbands.

Let $p_{i,k,m}$ denote the relaying power of the i-th relay node in the k-th subband and in the m-th fading block[2]. Assume that an average power constraint for the i-th relay node is

$$\frac{1}{M}\sum_{m=1}^{M}\sum_{k=1}^{K} p_{i,k,m} \leq P \ , \tag{80}$$

where M is sufficiently large. Worth noting is that there is no power allocation across the relay nodes is allowed (cf. [7]).

Next, we characterize the equivalent source-to-destination model, assuming that relay nodes are allowed to carry out power allocation under the constraint in (80). More specifically, in the k-th subband and the m-th fading block, the equivalent model is given by

$$Y_{k,m}(t) = \mathbb{E}[H_{k,m}]X_k(t-LT_s) + \Omega_{1k,m}(t) + \Omega_{2k,m}(t) + \Omega_{3k,m}(t) + Z_k(t) \ , \tag{81}$$

The power of the desired signal is given by

$$\begin{aligned}\mathbb{E}\left[|\mathbb{E}[H_{k,m}]X_k(t-LT_s)|^2\right] &= \frac{\pi^2}{16}\frac{\sigma_k^4}{BN_0}\left(\sum_{i=1}^{n}\sqrt{p_{i,k,m}}\right)^2\left(\frac{\eta_k\sigma_k^2PL}{\eta_k\sigma_k^2PL+WN_0}\right)^2 \\ &\quad\cdot\frac{(1-\eta_k)L}{L-1}\frac{P}{K} \ .\end{aligned} \tag{82}$$

The power of the noise terms in the k-th subband is given by

$$\mathbb{E}[|\Omega_{1k,m}(t)|^2] = \frac{\sigma_k^4}{BN_0}\left(\sum_{i=1}^{n}p_{i,k,m}\right)\left(1-\left(\frac{\eta_k\sigma_k^2PL}{\eta_k\sigma_k^2PL+WN_0}\right)^2\right)\frac{(1-\eta_k)L}{L-1}\frac{P}{K} \ , \tag{83}$$

$$\mathbb{E}[|\Omega_{2k,m}(t)|^2] = \sigma_k^2\left(\sum_{i=1}^{n}p_{i,k,m}\right) \ , \tag{84}$$

$$\begin{aligned}\mathbb{E}[|\Omega_{3k,m}(t)|^2] &= \frac{\sigma_k^4}{BN_0}\left(\sum_{i=1}^{n}p_{i,k,m}\right)\left(\left(\frac{\eta_k\sigma_k^2PL}{\eta_k\sigma_k^2PL+WN_0}\right)^2-\frac{\pi^2}{16}\left(\frac{\eta_k\sigma_k^2PL}{\eta_k\sigma_k^2PL+WN_0}\right)^2\right) \\ &\quad\cdot\frac{(1-\eta_k)L}{L-1}\frac{P}{K} \ .\end{aligned} \tag{85}$$

[2]Note that each relay node transmits only during the second hop. Without loss of generality, we assume that the power constraint is imposed for the transmission duration for each relay node.

For convenience, define the corresponding SNR at the k-th subband and the m-th fading block as

$$S_{k,m} \triangleq \frac{\mathbb{E}\left[\left|\mathbb{E}[H_{k,m}]X_{k,m}(t-LT_s)\right|^2\right]}{\mathbb{E}[|\Omega_{1k,m}(t)|^2]+\mathbb{E}[|\Omega_{2k,m}(t)|^2]+\mathbb{E}[|\Omega_{3k,m}(t)|^2]+WN_0}$$
$$= \frac{\sigma_k^4\frac{\pi^2}{16}n\left(\frac{1}{n}\sum_{i=1}^n\sqrt{p_{i,k,m}}\right)^2\left(\frac{\eta_k\sigma_k^2PL}{\eta_k\sigma_k^2PL+WN_0}\right)^2\frac{(1-\eta_k)L}{L-1}\frac{P}{K}}{\sigma_k^2\left(\frac{1}{n}\sum_{i=1}^n p_{i,k,m}\right)BN_0+\frac{(BN_0)^2}{n}}. \tag{86}$$

Next we study the optimal power allocation across the fading blocks and across the subbands for relay nodes. We note that the optimal power allocation should exploit the degrees of freedom in both the time domain and the frequency domain, and is the solution to the following optimization problem:

$$\max_{\{p_{i,k,m}\}} \quad \frac{1}{M}\sum_{m=1}^{M}\sum_{k=1}^{K}\frac{B(L-1)}{L}\log\left(1+S_{k,m}\right)$$
$$\text{subject to} \quad \frac{1}{M}\sum_{m=1}^{M}\sum_{k=1}^{K}p_{i,k,m}\le P, \quad \text{and} \quad p_{i,k,m}\ge 0. \tag{87}$$

In general, the optimal power allocation involves water-filling non-linear functions such as $x(\cdot)^+$. Needless to say, it is highly non-trivial to characterize the exact solution.

Let R^{WB^*} be the achievable rate corresponding to the optimal power allocation policy. In what follows, we use a "sandwich" argument instead to characterize the scaling order of R^{WB^*}; that is, we first find an upper bound and a lower bound on R^{WB^*}, and then show that these two bounds have the same scaling order under the same conditions.

[An Upper Bound on R^{WB^*}]:

First, by Cauchy-Schwarz's Inequality, we have that

$$\left(\frac{1}{n}\sum_{i=1}^{n}\sqrt{p_{i,k,m}}\right)^2 \le \frac{1}{n}\sum_{i=1}^{n}p_{i,k,m}, \tag{88}$$

where the equality can be achieved if and only if $p_{i,k,m}=p_{j,k,m}$, for $i\neq j; i,j=1,2,...,n$.

For convenience, define σ_{max}^2 and σ_{min}^2 as the maximum and minimum of $\{\sigma_1^2,\sigma_2^2,...,\sigma_K^2\}$, respectively; and define η_{max} and η_{min} as the maximum and minimum of $\{\eta_1,\eta_2,...,\eta_K\}$, respectively. Using (88), we have that

$$S_{k,m}\le S'_{k,m}\ , \tag{89}$$

where

$$S'_{k,m}\triangleq\frac{\sigma_{max}^4\frac{\pi^2}{16}n\left(\frac{1}{n}\sum_{i=1}^n p_{i,k,m}\right)\left(\frac{\eta_{max}\sigma_{max}^2PL}{\eta_{min}\sigma_{min}^2PL+WN_0}\right)^2\frac{(1-\eta_{min})L}{L-1}\frac{P}{K}}{\sigma_{min}^2\left(\frac{1}{n}\sum_{i=1}^n p_{i,k,m}\right)BN_0+\frac{(BN_0)^2}{n}}. \tag{90}$$

We need the following lemma.

Lemma 5.2 *The function $f(x_1,\ x_2, ..., x_n)$ is concave in $[x_1,\ x_2, ..., x_n]$, where*

$$f(x_1,\ x_2, ..., x_n) = \log\left(1 + \frac{C\sum_{i=1}^n x_i}{D_1 \sum_{i=1}^n x_i + D_2}\right), \tag{91}$$

with $x_i \geq 0$, $i = 1, 2, ..., n$; C, D_1 and D_2 being positive constants.

Proof: The proof simply follows the facts that

- $log\left(1 + \frac{C}{D_1}x\right)$ is concave and increasing in x with $x \geq 0$;
- $\frac{x}{x+D_2/D_1}$ is concave and increasing in x with $x \geq 0$;
- $\sum_i x_i$ is concave in vector $[x_1,\ x_2, ..., x_n]$ with $x_i \geq 0$.

By the properties of the function composition that preserves the concavity [2, p.84], we conclude the proof of Lemma 5.2. ∎

Applying Lemma 5.2, we have that $S'_{k,m}$ is concave in vector $[p_{1,k,m},\ p_{2,k,m}, ..., p_{n,k,m}]$. Then, by Jensen's Inequality, it follows that

$$\begin{aligned}
&\frac{1}{MK}\sum_{m=1}^{M}\sum_{k=1}^{K}\log\left(1 + S'_{k,m}\right) \\
&\leq \log\left(1 + \frac{\sigma_{max}^4 \frac{\pi^2}{16} n \left(\frac{1}{n}\sum_{i=1}^n \frac{1}{MK}\sum_{k=1}^K\sum_{m=1}^M p_{i,k,m}\right)\left(\frac{\eta_{max}\sigma_{max}^2 PL}{\eta_{min}\sigma_{min}^2 PL + WN_0}\right)^2 \frac{(1-\eta_{min})L}{L-1}\frac{P}{K}}{\sigma_{min}^2\left(\frac{1}{n}\sum_{i=1}^n \frac{1}{MK}\sum_{k=1}^K\sum_{m=1}^M p_{i,k,m}\right) BN_0 + \frac{(BN_0)^2}{n}}\right) \quad (92) \\
&= \log\left(1 + \frac{\sigma_{max}^4 \frac{\pi^2}{16} n \left(\frac{\eta_{max}\sigma_{max}^2 PL}{\eta_{min}\sigma_{min}^2 PL + WN_0}\right)^2 \frac{(1-\eta_{min})L}{L-1} P}{\sigma_{min}^2 WN_0 + \frac{(WN_0)^2}{nP}}\right). \quad (93)
\end{aligned}$$

We note that the RHS of (93) does not depend on $p_{i,k,m}$, which indicates that the LHS of (92) can be maximized by setting $p_{i,k,m} = P/K$, $1 \leq i \leq n$, $1 \leq k \leq K$, $1 \leq m \leq M$. Then it follows that the achievable rate, R^{WB^*}, is upper bounded by

$$\begin{aligned}
R^{WB^*} &\leq \max_{\{p_{i,k,m}\}} \frac{KB(L-1)}{MKL}\sum_{m=1}^{M}\sum_{k=1}^{K}\log\left(1 + S'_{k,m}\right) \\
&= \frac{W(L-1)}{L}\log\left(1 + \frac{\sigma_{max}^4 \frac{\pi^2}{16} n \left(\frac{\eta_{max}\sigma_{max}^2 PL}{\eta_{min}\sigma_{min}^2 PL + WN_0}\right)^2 \frac{(1-\eta_{min})L}{L-1} P}{\sigma_{min}^2 WN_0 + \frac{(WN_0)^2}{nP}}\right). \quad (94)
\end{aligned}$$

For convenience, define

$$R_{upper}^{WB\ *} \triangleq W\log\left(1 + \frac{\sigma_{max}^4 \frac{\pi^2}{16} n \left(\frac{\eta_{max}\sigma_{max}^2 PL}{\eta_{min}\sigma_{min}^2 PL + WN_0}\right)^2 \frac{(1-\eta_{min})L}{L-1} P}{\sigma_{min}^2 WN_0 + \frac{(WN_0)^2}{nP}}\right). \tag{95}$$

[A Lower Bound on R^{WB*}]:

Since R^{WB*} is the maximum achievable rate when the relay node can allocation its power across the fading blocks and across the subchannels, any other achievable rate serves as a lower bound on R^{WB*}. Then it follows that

$$R^{WB*} \geq R_{Eq}^{WB} \ . \tag{96}$$

Observe that the scaling order of R_{Eq}^{WB} in (76) is in fact the same as that of ${R_{upper}^{WB}}^*$ in (95) (the proof is relegated to Appendix C), we conclude that in the asymptotic regime of n, L and W,

$$R_{Eq}^{WB} = \Theta({R_{upper}^{WB}}^*) \ . \tag{97}$$

Somewhat surprisingly, as shown in (97), the scaling order of the "optimal" achievable rate R^{WB*} is exactly the same as that of R_{Eq}^{WB}. That is to say, even if the relay nodes can optimally allocate the power across the fading blocks and among the subbands, the scaling order of the corresponding achievable rate is the same as that by applying the equal power allocation policy at the relay nodes.

5.4 Scaling Laws and Discussions

We have a few important observations are in order. First, the findings in Section 5.2 reveal that if $e_2 \geq e_3$ and $e_3 < 1$, then the achievable rate by using AF with network training, under the equal power allocation policy at relay nodes, has the same scaling order as the upper bound given by Lemma 5.1. That is to say, the scaling laws can be characterized accordingly. More specifically, as $n \to \infty$, if there exist $\delta \geq 0$, and $0 < \epsilon < 1$, such that $\frac{L}{W} \overset{\bullet}{\geq} n^{\delta}$ and $\frac{n}{W} \overset{\bullet}{\geq} n^{\epsilon}$, then as $n \to \infty$, the capacity of wideband relay networks scales as

$$C^{WB} = \Theta\left(W \log\left(\frac{n}{W}\right)\right) . \tag{98}$$

Under this sufficient condition, assuming equal power allocation across the subbands, AF with network training can yield an achievable rate of the same scaling order as that of the upper bound, thus enabling us to characterize the scaling law. Furthermore, combing the results in Sections 5.2 and 5.3, we conclude that the conditions for achieving the scaling law under optimal power allocation are the same as that under equal power allocation. In a nutshell, in power-constrained wideband relay networks using AF with network training, the equal power allocation policy at relay nodes can result in the same scaling order as the optimal power allocation policy at relay nodes. However, it should be cautioned that from the point of view of exact achievable rates (not only the scaling orders), the wideband relay networks are not a simple collection of narrowband relay networks with equal power allocation.

We observe that in general, in a power-constrained wideband relay network where AF is employed, a larger bandwidth corresponds to a smaller amplification factor for relaying, which in turn may reduce the power of the desired signal embedded in the amplified signal from each relay node, and therefore the equivalent overall SNR at the destination node may or may not

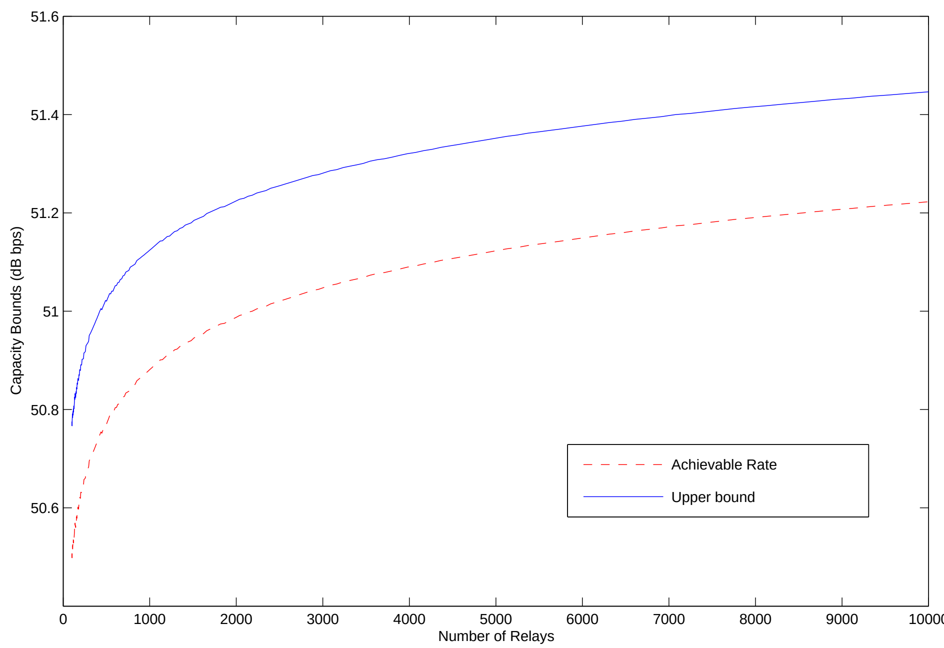

Figure 6: Narrowband relay networks: achievable rates and upper bounds

increase, depending on how W and n scale together. As shown in the above sections, the power of the aggregated "amplified" noise from the relay nodes is of $\Theta(n)$, and the power of the ambient noise at the destination node is of $\Theta(W)$. When W is sublinear in n, the "amplified" noise from all the relay nodes dominates the ambient noise at the destination node. As a result, all forward links are virtually "noise-free". Such "noise-free" links effectively convert all the relay nodes and the destination node into an antenna array of n receive antennas. Furthermore, observe that the relay nodes carry out phase alignment, and the equivalent virtual antenna array effectively conducts maximal-ratio combining (MRC). Thus, the relay network functions as if it were a single-input-multiple-output system (SIMO) with optimal decoding at the receiver. Recall that the upper bound in Lemma 5.1, given by the Cut-Set Theorem, assumes that an optimal joint decoding is conducted by all receivers, i.e., the upper bound corresponds to the capacity of a SIMO system. It is now not surprising that under the "noise-free" conditions mentioned above, the lower bound can meet the upper bound.

6 Numerical Examples

We now illustrate via numerical examples the achievable rates and upper bounds in Theorem 3.1, Theorem 3.2, Lemma 4.1 and Lemma 5.1. The key parameters used in the examples are listed as below:

- transmission power of the source node and the relay nodes: $P = 1$ dBm;

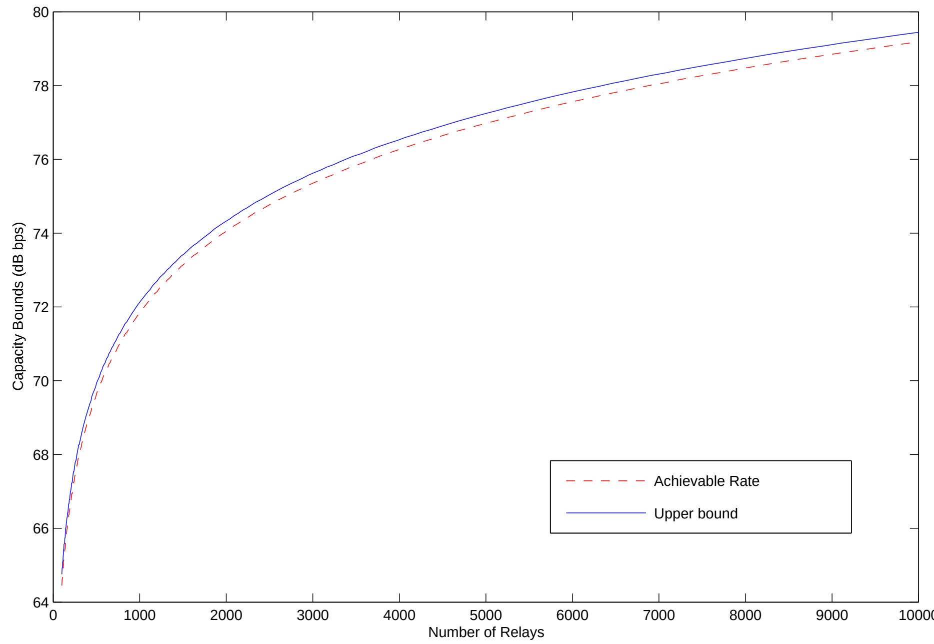

Figure 7: Wideband relay networks: achievable rates and upper bounds

- bandwidth of each subband: $B = 10$ kHz;
- noise power spectral density: $N_0 = 10^{-10}$ watts/Hz;
- the number of relay nodes n ranges from 10^2 to 10^4;
- the SNR per link in narrowband relay networks: $\rho = n^{-0.7}$; the number of subbands in wideband relay networks: $K = n^{0.7}$; and coherence interval: $L = n^{1.5}$. That is, $e_1 = e_3 = 0.7$ and $e_2 = 1.5$, which meet the scaling-law-achieving conditions given in Theorem 3.1 and Theorem 3.2;
- the unit of all capacity bounds is dB bits/second.

Fig. 6 depicts the lower bounds and the upper bounds on the capacity for a narrowband relay network model. Worth noting is that "dB bits/second" is used as the unit, because the main goal is to study the scaling order. It can be seen that when the number of relay nodes exceeds some threshold (2500 in this example), the two curves have the same slope, i.e., the lower bound and the upper bound have the same scaling order, corroborating the results in Theorem 3.1.

Fig. 7 illustrates the lower bound and the upper bound on the capacity for a power-constrained wideband relay network. It can be seen that the upper bounds and the lower bounds become "parallel" when n exceeds approximately 3000, indicating that the lower bounds and the upper bounds have the same scaling order asymptotically.

7 Conclusion

In this paper, we study the achievable rates and scaling laws of large-scale wireless relay networks in the wideband regime. Assuming that relay nodes have no *a priori* knowledge of channel state information (CSI) for both the backward channels and the forward channels, we examine the achievable rates in the joint asymptotic regime of the number of relay nodes n, the channel coherence interval L, and the SNR per link ρ, or the system bandwidth W. We first study narrowband relay networks in the low SNR regime. We study a relaying scheme, namely amplify-and-forward (AF) with network training, in which the source node and the destination node broadcast training symbols and each relay node carries out channel estimation and then uses AF relaying to relay information. We provide an equivalent source-to-destination channel model, and characterize the corresponding achievable rate. Our findings show that when ρL, proportional to the transmission energy in each fading block, is bounded below, the achievable rate has the same scaling order as in coherent relaying, thus enabling us to characterize the scaling law of the relay networks in the low SNR regime.

We then generalize the study to power-constrained wideband relay networks, where frequency-selective fading is taken into account. In particular, we examine the scaling behavior of the achievable rates corresponding to two power allocation policies applied at relay nodes, namely, a simple equal power allocation policy and the optimal power allocation policy. We identify the conditions under which the scaling law of the wideband relay networks can be achieved by both power allocation policies. Somewhat surprising, our findings indicate that these two power allocation policies result in achievable rates of the same scaling behavior, and the scaling law can be characterized under the condition that L/W, proportional to the energy per fading block per subband, is bounded below and that W is sub-linear in n.

Throughout this paper, we have focused on the achievable rates using amplify-and-forward (AF) with network training. Another popular candidate for information relaying is decode-and-forward (DF). In the DF strategy, the relay nodes first extract the information bits from the received signals, re-encode it, and then transmit the new encoded bits in the second hop. In general, the optimal DF strategy is highly non-trivial and corresponding complexity can increase significantly as the number of relay nodes grows. We are currently pursuing the scaling behavior by using DF relaying.

Acknowledgements

The authors would like to thank Professor P.R. Kumar at UIUC for helpful comments, and Amir F. Dana at Caltech for useful discussion on Appendix A. They would also like to thank the anonymous reviewers for their helpful comments that greatly improved the presentation of the paper.

A Proof of (24)

The derivation is a modification of [15, (46)], where the author assumed the noise is Gaussian. In what follows, we show that letting the noise be non-Gaussian and signal-dependent can yield the same result.

For convenience, define

$$F \triangleq H + \sum_i \sqrt{\rho} \left(\frac{|\hat{\alpha}_i|^2 \tilde{\beta}_i \hat{\beta}_i^* + |\hat{\beta}_i|^2 \tilde{\alpha}_i \hat{\alpha}_i^* + \tilde{\alpha}_i \hat{\alpha}_i^* \tilde{\beta}_i \hat{\beta}_i^*}{|\hat{\alpha}_i||\hat{\beta}_i|} \right). \quad (99)$$

Then, we have the equivalent source-to-destination signal model as

$$Y(t) = FX(t - LT_s) + \Omega_2(t) + Z(t). \quad (100)$$

Observe that in the above equivalent signal model ,

- the channel gain F and the transmitted signal $X(t-LT_s)$ are independent, and $X(t-LT_s)$ and the noise term $\Omega_2(t) + Z(t)$ are uncorrelated;
- $\mathbb{E}[F] = \mathbb{E}[H]$;
- the noise term $\Omega_2(t) + Z(t)$ is non-Gaussian and signal-dependent, and is zero-mean and of bounded power.

Using standard vector representation for band-limited signals [3], the continuous-time channel model in (100) can be mapped onto a set of discrete-time subchannels. That is, the signal model in the j-th dimension is given by

$$Y^{(j)} = FX^{(j)} + \Omega_2^{(j)} + Z^{(j)} \quad (101)$$

Next, we examine the achievable rate corresponding to the channel model in (101). Note that the mutual information between $X^{(j)}$ and $Y^{(j)}$ is

$$I\left(X^{(j)}; Y^{(j)}\right) = h\left(X^{(j)}\right) - h\left(X^{(j)}|Y^{(j)}\right). \quad (102)$$

Along the lines in [15], we fix the distribution of $X^{(j)}$ as complex Gaussian with zero mean, even though this choice may not be optimal for maximizing the mutual information. Next, we find an upper bound on $h\left(X^{(j)}|Y^{(j)}\right)$, which holds for all possible distributions for $X^{(j)}$. We have that

$$\begin{aligned} h\left(X^{(j)}|Y^{(j)}\right) &= h\left(X^{(j)} - bY^{(j)}|Y^{(j)}\right) \\ &\overset{\text{(i)}}{\leq} h\left(X^{(j)} - bY^{(j)}\right) \\ &\overset{\text{(ii)}}{\leq} \log\left(2\pi e \cdot \text{var}\left(X^{(j)} - bY^{(j)}\right)\right), \end{aligned} \quad (103)$$

where (i) follows the fact that conditioning always decreases entropy; (ii) follows that the entropy of a random variable of bounded variance is upper-bounded by the entropy of a Gaussian random variable with the same variance; and the above bound holds for any b. Let b be the coefficient, for which $bY^{(j)}$ is the linear minimum mean-square error(LMMSE) estimate of $X^{(j)}$ in terms of $Y^{(j)}$. That is,

$$b = \frac{\mathbb{E}\left[X^{(j)}Y^{(j)}\right]}{\mathrm{var}\left(Y^{(j)}\right)}. \tag{104}$$

We note that F and $X(t-LT_s)$ are independent, and $X(t-LT_s)$ and $\Omega_2(t)+Z(t)$ are uncorrelated and zero-mean. After some algebra, we have that

$$\begin{aligned}\max_{p(X^{(j)})} I\left(X^{(j)};Y^{(j)}\right) &\geq \log\left(2\pi e\cdot\mathrm{var}\left(X^{(j)}\right)\right)-\log\left(2\pi e\cdot\mathrm{var}\left(X^{(j)}-bY^{(j)}\right)\right)\\ &= \log\left(1+\frac{\mathbb{E}\left[\left|\mathbb{E}[F]X^{(j)}\right|^2\right]}{\mathbb{E}[|\Omega_1^{(j)}|^2]+\mathbb{E}[|\Omega_2^{(j)}|^2]+\mathbb{E}[|\Omega_3^{(j)}|^2]+\mathbb{E}[|Z^{(j)}|^2]}\right)\end{aligned} \tag{105}$$

Combing all the dimensions and taking into account the fact that the relay nodes work in half-duplex mode, we have that the lower bound on the capacity of (23) is given by

$$C \geq B\frac{L-1}{L}\log\left(1+\frac{\mathbb{E}\left[\left|\mathbb{E}[H]X(t-LT_s)\right|^2\right]}{\mathbb{E}[|\Omega_1(t)|^2]+\mathbb{E}[|\Omega_2(t)|^2]+\mathbb{E}[|\Omega_3(t)|^2]+WN_0}\right), \tag{106}$$

where $\frac{L-1}{L}$ denotes the portion of time for information transmission.

B Proof of (82), (83), (84) and (85)

Recall that in (81),

$$H_{k,m} \triangleq \sum_i \sqrt{\frac{p_{i,k,m}}{BN_0}}\left|\hat{\alpha}_{i,k}\hat{\beta}_{i,k}\right| , \tag{107}$$

$$\Omega_{1k,m}(t) \triangleq \sum_i \sqrt{\frac{p_{i,k,m}}{BN_0}}\left(\frac{|\hat{\alpha}_{i,k}|^2\tilde{\beta}_{i,k}\hat{\beta}^*_{i,k}+|\hat{\beta}_{i,k}|^2\tilde{\alpha}_{i,k}\hat{\alpha}^*_{i,k}+\tilde{\alpha}_{i,k}\hat{\alpha}^*_{i,k}\tilde{\beta}_{i,k}\hat{\beta}^*_{i,k}}{|\hat{\alpha}_{i,k}||\hat{\beta}_{i,k}|}\right)X_k(t-LT_s) , \tag{108}$$

$$\Omega_{2k,m}(t) \triangleq \sum_i \sqrt{\frac{p_{i,k,m}}{BN_0}}\left(\frac{\hat{\alpha}^*_{i,k}|\hat{\beta}_{i,k}|^2+\hat{\alpha}^*_{i,k}\tilde{\beta}_{i,k}\hat{\beta}^*_{i,k}}{|\hat{\alpha}_{i,k}||\hat{\beta}_{i,k}|}\right)Z_{i,k}(t-LT_s) , \tag{109}$$

$$\Omega_{3k,m}(t) \triangleq (H_{k,m}-\mathbb{E}[H_{k,m}])X_k(t-LT_s) , \tag{110}$$

where $Z_{i,k}(t-LT_s)$ is the noise at the i-th relay node within the k-th subband.

first, note that the power of the signal used for decoding is given by

$$\mathbb{E}\left[\left|\mathbb{E}[H_{k,m}]X_k(t-LT_s)\right|^2\right] = \left|\mathbb{E}[H_{k,m}]\right|^2\mathbb{E}[|X_k(t-LT_s)|^2] . \tag{111}$$

Observe that

$$
\begin{aligned}
|\mathbb{E}[H_{k,m}]|^2 &= \left|\mathbb{E}\left[\sum_i \sqrt{\frac{p_{i,k,m}}{BN_0}}\left|\hat{\alpha}_{i,k}\hat{\beta}_{i,k}\right|\right]\right|^2 \\
&= \left|\sum_i \sqrt{\frac{p_{i,k,m}}{BN_0}}\mathbb{E}\left[\left|\hat{\alpha}_{i,k}\hat{\beta}_{i,k}\right|\right]\right|^2 \\
&\overset{(a)}{=} \left|\sum_i \sqrt{\frac{p_{i,k,m}}{BN_0}}\mathbb{E}\left[|\hat{\alpha}_{i,k}|\right]\mathbb{E}\left[\left|\hat{\beta}_{i,k}\right|\right]\right|^2 \\
&\overset{(b)}{=} \frac{1}{BN_0}\left(\sum_i \sqrt{p_{i,k,m}}\right)^2 \left(\mathbb{E}\left[|\hat{\alpha}_{i,k}|\right]\right)^2 \left(\mathbb{E}\left[\left|\hat{\beta}_{i,k}\right|\right]\right)^2 \\
&\overset{(c)}{=} \frac{1}{BN_0}\left(\sum_i \sqrt{p_{i,k,m}}\right)^2 \left(\frac{\pi\sigma_k^2}{4}\frac{\eta_k\sigma_k^2 PL}{\eta_k\sigma_k^2 PL + WN_0}\right)^2 \qquad (112)
\end{aligned}
$$

where

(a) follows the fact that $\hat{\alpha}_{i,k}$ and $\hat{\beta}_{i,k}$ are independent;

(b) follows that within a given subband, channel coefficients are i.i.d random variables;

(c) comes from the fact that $\mathbb{E}\left[|\hat{\alpha}_{i,k}|\right] = \mathbb{E}\left[\left|\hat{\beta}_{i,k}\right|\right]$, and that

$$
\begin{aligned}
\mathbb{E}\left[|\hat{\alpha}_{i,k}|\right] &= \sqrt{\frac{\pi}{4}\mathbb{E}[|\hat{\alpha}_{i,k}|^2]} \\
&= \sqrt{\frac{\pi\sigma_k^2}{4}\frac{\eta_k\sigma_k^2 PL}{\eta_k\sigma_k^2 PL + WN_0}}\ . \qquad (113)
\end{aligned}
$$

Hence, we have that

$$
\mathbb{E}\left[\left|\mathbb{E}[H_{k,m}]X_k(t-LT_s)\right|^2\right] = \frac{\pi^2}{16}\frac{\sigma_k^4}{BN_0}\left(\sum_{i=1}^{n}\sqrt{p_{i,k,m}}\right)^2\left(\frac{\eta_k\sigma_k^2 PL}{\eta_k\sigma_k^2 PL + WN_0}\right)^2\frac{(1-\eta_k)L}{L-1}\frac{P}{K}\ . \qquad (114)
$$

Next, we characterize the power of $\Omega_{1k,m}(t)$.

$$
\begin{aligned}
\mathbb{E}\left[\left|\Omega_{1k,m}(t)\right|^2\right] &= \mathbb{E}\left[\left|\sum_i \sqrt{\frac{p_{i,k,m}}{BN_0}}\cdot\frac{|\hat{\alpha}_{i,k}|^2\tilde{\beta}_{i,k}\hat{\beta}^*_{i,k} + |\hat{\beta}_{i,k}|^2\tilde{\alpha}_{i,k}\hat{\alpha}^*_{i,k} + \tilde{\alpha}_{i,k}\hat{\alpha}^*_{i,k}\tilde{\beta}_{i,k}\hat{\beta}^*_{i,k}}{|\hat{\alpha}_{i,k}||\hat{\beta}_{i,k}|}\right|^2\right] \\
&\quad \cdot\mathbb{E}[|X_k(t-LT_s)|^2]\ . \qquad (115)
\end{aligned}
$$

We have that

$$
\begin{aligned}
& \mathbb{E}\left[\left|\sum_i \sqrt{\frac{p_{i,k,m}}{BN_0}} \cdot \frac{|\hat{\alpha}_{i,k}|^2 \tilde{\beta}_{i,k}\hat{\beta}^*_{i,k} + |\hat{\beta}_{i,k}|^2 \tilde{\alpha}_{i,k}\hat{\alpha}^*_{i,k} + \tilde{\alpha}_{i,k}\hat{\alpha}^*_{i,k}\tilde{\beta}_{i,k}\hat{\beta}^*_{i,k}}{|\hat{\alpha}_{i,k}||\hat{\beta}_{i,k}|}\right|^2\right] \\
= & \mathbb{E}\left[\left|\sum_i \sqrt{\frac{p_{i,k,m}}{BN_0}} \left(\underbrace{\frac{|\hat{\alpha}_{i,k}|^2 \tilde{\beta}_{i,k}\hat{\beta}^*_{i,k}}{|\hat{\alpha}_{i,k}||\hat{\beta}_{i,k}|}}_{U_{1i,k}} + \underbrace{\frac{|\hat{\beta}_{i,k}|^2 \tilde{\alpha}_{i,k}\hat{\alpha}^*_{i,k}}{|\hat{\alpha}_{i,k}||\hat{\beta}_{i,k}|}}_{U_{2i,k}} + \underbrace{\frac{\tilde{\alpha}_{i,k}\hat{\alpha}^*_{i,k}\tilde{\beta}_{i,k}\hat{\beta}^*_{i,k}}{|\hat{\alpha}_{i,k}||\hat{\beta}_{i,k}|}}_{U_{3i,k}}\right)\right|^2\right] \\
= & \frac{1}{BN_0}\sum_i p_{i,k,m}\ \mathbb{E}\left[\left|U_{1i,k} + U_{2i,k} + U_{3i,k}\right|^2\right] \\
& + \frac{1}{BN_0}\sum_i \sum_{j,j\neq i} \sqrt{p_{i,k,m}p_{j,k,m}}\ \mathbb{E}\left[\left(U_{1i,k} + U_{2i,k} + U_{3i,k}\right)\left(U_{1j,k}^* + U_{2j,k}^* + U_{3j,k}^*\right)\right] \\
\overset{(a)}{=} & \frac{1}{BN_0}\left(\sum_i p_{i,k,m}\left(\mathbb{E}[|U_{1i,k}|^2] + \mathbb{E}[|U_{2i,k}|^2] + \mathbb{E}[|U_{3i,k}|^2]\right)\right) + 0 \\
\overset{(b)}{=} & \frac{1}{BN_0}\left(\sum_i p_{i,k,m}\right)\left(\frac{\sigma_k^2 W W_0}{\eta_k\sigma_k^2 PL + WN_0} \cdot \frac{\eta_k \sigma_k^4 PL}{\eta_k\sigma_k^2 PL + WN_0}\right. \\
& \left. + \frac{\sigma_k^2 W N_0}{\eta_k\sigma_k^2 PL + WN_0} \cdot \frac{\eta_k \sigma_k^4 PL}{\eta_k\sigma_k^2 PL + WN_0} + \left(\frac{\sigma_k^2 W N_0}{\eta_k\sigma_k^2 PL + WN_0}\right)^2\right) \\
= & \frac{\sigma_k^4}{BN_0}\left(\sum_i p_{i,k,m}\right)\left(1 - \left(\frac{\eta_k\sigma_k^2 PL}{\eta_k\sigma_k^2 PL + WN_0}\right)^2\right), \qquad (116)
\end{aligned}
$$

where

(a) follows the facts that

1. $\tilde{\alpha}_{i,k}$ and $\tilde{\beta}_{i,k}$ are zero-mean;
2. $U_{li,k}$ and $U_{l'i,k}$ ($l \neq l'$; $l, l' = 1, 2, 3$) are orthogonal;
3. $U_{li,k}$ and $U_{l'j,k}$ ($l, l' = 1, 2, 3$ and $i \neq j$) are orthogonal;

(b) follows that

1. $\mathbb{E}[|U_{1i,k}|^2]$, $\mathbb{E}[|U_{1i,k}|^2]$ and $\mathbb{E}[|U_{1i,k}|^2]$ are independent of i, and thus the summation is only taken with respect to $p_{i,k,m}$;
2.

$$
\begin{aligned}
\mathbb{E}[|U_{1i,k}|^2] &= \mathbb{E}[|U_{2i,k}|^2] \\
&= \frac{\sigma_k^2 W W_0}{\eta_k\sigma_k^2 PL + WN_0} \cdot \frac{\eta_k \sigma_k^4 PL}{\eta_k\sigma_k^2 PL + WN_0} \qquad (117) \\
\mathbb{E}[|U_{3i,k}|^2] &= \left(\frac{\sigma_k^2 W W_0}{\eta_k\sigma_k^2 PL + WN_0}\right)^2. \qquad (118)
\end{aligned}
$$

Therefore, the power of $\Omega_{1k,m}(t)$ is given by

$$\mathbb{E}[|\Omega_{1k,m}(t)|^2] = \frac{\sigma_k^4}{BN_0}\left(\sum_{i=1}^{n} p_{i,k,m}\right)\left(1-\left(\frac{\eta_k\sigma_k^2 PL}{\eta_k\sigma_k^2 PL + WN_0}\right)^2\right)\frac{(1-\eta_k)L}{L-1}\frac{P}{K}\ . \tag{119}$$

The power of the aggregated noise $\Omega_{2k,m}(t)$ is given by

$$\begin{aligned}\mathbb{E}[|\Omega_{2k,m}(t)|^2] &= \mathbb{E}\left[\left|\sum_i \sqrt{\frac{p_{i,k,m}}{BN_0}}\left(\frac{\hat{\alpha}_{i,k}^*|\hat{\beta}_{i,k}|^2 + \hat{\alpha}_{i,k}^*\tilde{\beta}_{i,k}\hat{\beta}_{i,k}^*}{|\hat{\alpha}_{i,k}||\hat{\beta}_{i,k}|}\right) Z_{i,k}(t-LT_s)\right|^2\right] \\ &= \mathbb{E}\left[\left|\sum_i \sqrt{\frac{p_{i,k,m}}{BN_0}}\left(\frac{\hat{\alpha}_{i,k}^*|\hat{\beta}_{i,k}|^2 + \hat{\alpha}_{i,k}^*\tilde{\beta}_{i,k}\hat{\beta}_{i,k}^*}{|\hat{\alpha}_{i,k}||\hat{\beta}_{i,k}|}\right)\right|^2\right] BN_0\ . \end{aligned}\tag{120}$$

Along the same line in the proof for (116), we have that

$$\begin{aligned}&\mathbb{E}\left[\left|\sum_i \sqrt{\frac{p_{i,k,m}}{BN_0}}\left(\frac{\hat{\alpha}_{i,k}^*|\hat{\beta}_{i,k}|^2 + \hat{\alpha}_{i,k}^*\tilde{\beta}_{i,k}\hat{\beta}_{i,k}^*}{|\hat{\alpha}_{i,k}||\hat{\beta}_{i,k}|}\right)\right|^2\right] \\ =\ & \mathbb{E}\left[\left|\sum_i \sqrt{\frac{p_{i,k,m}}{BN_0}}\left(\underbrace{\frac{\hat{\alpha}_{i,k}^*|\hat{\beta}_{i,k}|^2}{|\hat{\alpha}_{i,k}||\hat{\beta}_{i,k}|}}_{V_{1i,k}} + \underbrace{\frac{\hat{\alpha}_{i,k}^*\tilde{\beta}_{i,k}\hat{\beta}_{i,k}^*}{|\hat{\alpha}_{i,k}||\hat{\beta}_{i,k}|}}_{V_{2i,k}}\right)\right|^2\right] \\ =\ & \frac{1}{BN_0}\sum_i p_{i,k,m}\ \mathbb{E}\left[\left|V_{1i,k}+V_{2i,k}\right|^2\right] \\ & + \frac{1}{BN_0}\sum_i\sum_{j,j\neq i}\sqrt{p_{i,k,m}p_{j,k,m}}\ \mathbb{E}\left[\left(V_{1i,k}+V_{2i,k}\right)\left(V_{1j,k}^*+V_{2j,k}^*\right)\right] \\ \overset{(a)}{=}\ & \frac{1}{BN_0}\sum_i p_{i,k,m}\ \left(\mathbb{E}\left[\left|V_{1i,k}\right|^2\right]+\mathbb{E}\left[\left|V_{2i,k}\right|^2\right]\right)+0 \\ =\ & \frac{1}{BN_0}\sum_i p_{i,k,m}\ \left(\mathbb{E}\left[\left|\hat{\beta}_{i,k}\right|^2\right]+\mathbb{E}\left[\left|\tilde{\beta}_{i,k}\right|^2\right]\right) \\ =\ & \frac{\sigma_k^2}{BN_0}\left(\sum_i p_{i,k,m}\right)\ , \end{aligned}\tag{121}$$

where

(a) follows the facts that

1. $V_{li,k}$ and $V_{l'j,k}$ ($i\neq j$, $l,l'=1,2$) are independent;
2. $\mathbb{E}[V_{1i,k}]=0$ and $\mathbb{E}[V_{2i,k}]=0$.

As a result, we have that

$$\mathbb{E}[|\Omega_{2k,m}(t)|^2] = \sigma_k^2\left(\sum_{i=1}^{n} p_{i,k,m}\right)\ . \tag{122}$$

$$\begin{aligned}\mathbb{E}\left[|\Omega_{3k,m}(t)|^2\right] &= \mathbb{E}\left[\left(\sum_i \sqrt{\frac{p_{i,k,m}}{BN_0}}\left|\hat{\alpha}_{i,k}\hat{\beta}_{i,k}\right| - \mathbb{E}\left[\sum_i \sqrt{\frac{p_{i,k,m}}{BN_0}}\left|\hat{\alpha}_{i,k}\hat{\beta}_{i,k}\right|\right]\right)^2\right] \\ &\cdot \mathbb{E}\left[\left|X_k(t-LT_s)\right|^2\right] \ . \qquad (123)\end{aligned}$$

Note that

$$\begin{aligned}&\mathbb{E}\left[\left(\sum_i \sqrt{\frac{p_{i,k,m}}{BN_0}}\left(\left|\hat{\alpha}_{i,k}\hat{\beta}_{i,k}\right| - \mathbb{E}\left[\left|\hat{\alpha}_{i,k}\hat{\beta}_{i,k}\right|\right]\right)\right)^2\right] \\ = \ & \frac{1}{BN_0}\sum_i p_{i,k,m}\mathbb{E}\left[\left(\left|\hat{\alpha}_{i,k}\hat{\beta}_{i,k}\right| - \mathbb{E}\left[\left|\hat{\alpha}_{i,k}\hat{\beta}_{i,k}\right|\right]\right)^2\right] \\ & + \frac{1}{BN_0}\sum_i \sum_{j,j\neq i}\sqrt{p_{i,k,m}p_{j,k,m}}\ \mathbb{E}\left[\left(\left|\hat{\alpha}_{i,k}\hat{\beta}_{i,k}\right| - \mathbb{E}\left[\left|\hat{\alpha}_{i,k}\hat{\beta}_{i,k}\right|\right]\right)\left(\left|\hat{\alpha}_{j,k}\hat{\beta}_{j,k}\right| - \mathbb{E}\left[\left|\hat{\alpha}_{j,k}\hat{\beta}_{j,k}\right|\right]\right)\right] \\ \overset{(a)}{=} \ & \frac{1}{BN_0}\sum_i p_{i,k,m}\left(\mathbb{E}\left[\left|\hat{\alpha}_{i,k}\hat{\beta}_{i,k}\right|^2\right] - \left(\mathbb{E}\left[\left|\hat{\alpha}_{i,k}\hat{\beta}_{i,k}\right|\right]\right)^2\right) + 0 \\ \overset{(b)}{=} \ & \frac{1}{BN_0}\sum_i p_{i,k,m}\left(\left(\mathbb{E}\left[|\hat{\alpha}_{i,k}|^2\right]\right)^2 - \left(\mathbb{E}\left[|\hat{\alpha}_{i,k}|\right]\right)^4\right) \\ = \ & \frac{\sigma_k^4}{BN_0}\left(\sum_i p_{i,k,m}\right)\left(\left(\frac{\eta_k\sigma_k^2 PL}{\eta_k\sigma_k^2 PL + WN_0}\right)^2 - \frac{\pi^2}{16}\left(\frac{\eta_k\sigma_k^2 PL}{\eta_k\sigma_k^2 PL + WN_0}\right)^2\right) \ , \qquad (124)\end{aligned}$$

where

(a) follows the fact that

$$\begin{aligned}&\mathbb{E}\left[\left(\left|\hat{\alpha}_{i,k}\hat{\beta}_{i,k}\right| - \mathbb{E}\left[\left|\hat{\alpha}_{i,k}\hat{\beta}_{i,k}\right|\right]\right)\left(\left|\hat{\alpha}_{j,k}\hat{\beta}_{j,k}\right| - \mathbb{E}\left[\left|\hat{\alpha}_{j,k}\hat{\beta}_{j,k}\right|\right]\right)\right] \\ = \ & \mathbb{E}\left[\left(\left|\hat{\alpha}_{i,k}\hat{\beta}_{i,k}\right| - \mathbb{E}\left[\left|\hat{\alpha}_{i,k}\hat{\beta}_{i,k}\right|\right]\right)\right]\cdot\mathbb{E}\left[\left(\left|\hat{\alpha}_{j,k}\hat{\beta}_{j,k}\right| - \mathbb{E}\left[\left|\hat{\alpha}_{j,k}\hat{\beta}_{j,k}\right|\right]\right)\right] \\ = \ & 0; \qquad (125)\end{aligned}$$

(b) follows that $\hat{\alpha}_{i,k}$ and $\hat{\beta}_{i,k}$ are i.i.d. random variables.

Thus, the power of $\Omega_{3k,m}(t)$ is given by

$$\begin{aligned}\mathbb{E}[|\Omega_{3k,m}(t)|^2] &= \frac{\sigma_k^4}{BN_0}\left(\sum_{i=1}^n p_{i,k,m}\right)\left(\left(\frac{\eta_k\sigma_k^2 PL}{\eta_k\sigma_k^2 PL + WN_0}\right)^2 - \frac{\pi^2}{16}\left(\frac{\eta_k\sigma_k^2 PL}{\eta_k\sigma_k^2 PL + WN_0}\right)^2\right) \\ &\cdot\frac{(1-\eta_k)L}{L-1}\frac{P}{K} \ . \qquad (126)\end{aligned}$$

When $p_{i,k,m} \equiv p_k$, namely, a common power constraint in the k-th subband across fading blocks is applied at each relay node, then it follows that

$$\sum_{i=1}^n p_{i,k,m} = np_k \qquad (127)$$

and

$$\left(\sum_{i=1}^n \sqrt{p_{i,k,m}}\right)^2 = n^2 p_k, \qquad (128)$$

Therefore we have that the power of the desired signal in the k-th subband is given by

$$\mathbb{E}\left[|\mathbb{E}[H_k]X_k(t-LT_s)|^2\right] = n^2\sigma_k^4\frac{\pi^2}{16}\frac{p_k}{BN_0}\left(\frac{\eta_k\sigma_k^2PL}{\eta_k\sigma_k^2PL+WN_0}\right)^2\frac{(1-\eta_k)L}{L-1}\frac{P}{K}\ . \tag{129}$$

The noise power in the k-th subband is given by

$$\mathbb{E}[|\Omega_{1k}(t)|^2] = n\sigma_k^4\frac{p_k}{BN_0}\left(1-\left(\frac{\eta_k\sigma_k^2PL}{\eta_k\sigma_k^2PL+WN_0}\right)^2\right)\frac{(1-\eta_k)L}{L-1}\frac{P}{K}\ , \tag{130}$$

$$\mathbb{E}[|\Omega_{2k}(t)|^2] = n\sigma_k^2p_k\ , \tag{131}$$

$$\begin{aligned}\mathbb{E}[|\Omega_{3k}(t)|^2] = n\sigma_k^4\frac{p_k}{BN_0}&\left(\left(\frac{\eta_k\sigma_k^2PL}{\eta_k\sigma_k^2PL+WN_0}\right)^2-\frac{\pi^2}{16}\left(\frac{\eta_k\sigma_k^2PL}{\eta_k\sigma_k^2PL+WN_0}\right)^2\right)\\ &\cdot\frac{(1-\eta_k)L}{L-1}\frac{P}{K}\ .\end{aligned} \tag{132}$$

For the narrowband relay network ($K=1$), we have that

$$\sigma_k^2 = 1\ , \qquad \eta_k = 1 \tag{133}$$

and

$$p_k = P. \tag{134}$$

Then it follows that the power of the signal term is given by

$$\begin{aligned}\mathbb{E}\left[|\mathbb{E}[H]X(t-LT_s)|^2\right] &= \rho\left(\mathbb{E}\left[\sum_i|\alpha_i||\beta_i|\right]\right)^2\frac{(1-\eta)L}{L-1}P\\ &= n^2\frac{\pi^2}{16}\rho\left(\frac{\eta\rho L}{\eta\rho L+1}\right)^2\frac{(1-\eta)L}{L-1}P\ ,\end{aligned} \tag{135}$$

and the power of the noise terms is given by

$$\mathbb{E}[|\Omega_1(t)|^2] = n\rho\left(1-\left(\frac{\eta\rho L}{\eta\rho L+1}\right)^2\right)\frac{(1-\eta)L}{L-1}P\ , \tag{136}$$

$$\mathbb{E}[|\Omega_2(t)|^2] = \rho\mathbb{E}\left[\left|\sum_i\frac{\hat{\alpha_i}^*}{|\hat{\alpha_i}|}\left(\hat{\beta_i}+\frac{\tilde{\beta_i}\hat{\beta_i}^*}{|\hat{\beta_i}|}\right)\right|^2\right]BN_0 = nP\ ,$$

$$\mathbb{E}[|\Omega_3(t)|^2] = n\rho\left(\left(\frac{\eta\rho L}{\eta\rho L+1}\right)^2-\frac{\pi^2}{16}\left(\frac{\eta\rho L}{\eta\rho L+1}\right)^2\right)\frac{(1-\eta)L}{L-1}P. \tag{137}$$

C Proof of (97)

Recall that

$$R_{upper}^{WB\ *} = W\log\left(1+\frac{\sigma_{max}^4\frac{\pi^2}{16}n\left(\frac{\eta_{max}\sigma_{max}^2PL}{\eta_{min}\sigma_{min}^2PL+WN_0}\right)^2\frac{(1-\eta_{min})L}{L-1}P}{\sigma_{min}^2WN_0+\frac{(WN_0)^2}{nP}}\right)\ , \tag{138}$$

$$R_{Eq}^{WB} = \Theta\left(W\log\left(1+S_{Eq,k}\right)\right)\ . \tag{139}$$

In order to prove (97), it suffices to show that

$$S_{Eq,k} = \Theta(S_{k,upper}) \ , \tag{140}$$

where

$$S_{k,upper} \triangleq \frac{\sigma_{max}^4 \frac{\pi^2}{16} n \left(\frac{\eta_{max}\sigma_{max}^2 PL}{\eta_{min}\sigma_{min}^2 PL + WN_0} \right)^2 \frac{(1-\eta_{min})L}{L-1} P}{\sigma_{min}^2 WN_0 + \frac{(WN_0)^2}{nP}} \ . \tag{141}$$

Note that $\eta_k^* = \left(-3 + \sqrt{8\sigma_k^2 PL/WN_0 + 9}\right)/(2\sigma_k^2 PL/WN_0)$ is monotonically decreasing in σ_k^2. That is to say, η_{min} is corresponding to σ_{max}^2 and η_{max} is with respect to σ_{min}^2. Following the same line as in the equal power allocation policy, we have that

- If $e_3 < e_2$, then the optimal energy allocation for training is given by

$$\lim_{n\to\infty} \eta_{min} = 0 \ . \tag{142}$$

  Observing that

$$\frac{\eta_{min} PL}{WN_0} = \Theta\left(n^{\frac{e_2-e_3}{2}}\right) \quad \text{and} \quad \frac{\eta_{max} PL}{WN_0} = \Theta\left(n^{\frac{e_2-e_3}{2}}\right) \ , \tag{143}$$

  we conclude that

$$S_{k,upper} = \begin{cases} \Theta\left(n^{1-e_3}\right) & \text{if} \quad e_3 < 1 \\ \Theta\left(n^{2-2e_3}\right) & \text{if} \quad e_3 \geq 1 \end{cases} \ . \tag{144}$$

- If $e_3 > e_2$, then the optimal energy allocation for training is given by

$$\lim_{n\to\infty} \eta_{min} = \frac{2}{3}. \tag{145}$$

  Observing that

$$\frac{\eta_{min} PL}{WN_0} = \Theta\left(n^{e_2-e_3}\right) \quad \text{and} \quad \frac{\eta_{max} PL}{WN_0} = \Theta\left(n^{e_2-e_3}\right) \ , \tag{146}$$

  we conclude that

$$S_{k,upper} = \begin{cases} \Theta\left(n^{1+2e_2-3e_3}\right) & \text{if} \quad e_3 < 1 \\ \Theta\left(n^{2+2e_2-4e_3}\right) & \text{if} \quad e_3 \geq 1 \end{cases} \ . \tag{147}$$

- If $e_3 = e_2$, namely, $L/W = \Theta(1)$, then it follows that

$$\frac{\eta_{min} PL}{WN_0} = \Theta(1) \quad \text{and} \quad \frac{\eta_{max} PL}{WN_0} = \Theta(1) \ . \tag{148}$$

  Then it follows that

$$S_{k,upper} = \begin{cases} \Theta\left(n^{1-e_3}\right) & \text{if} \quad e_3 < 1 \\ \Theta\left(n^{2-2e_3}\right) & \text{if} \quad e_3 \geq 1 \end{cases} \ . \tag{149}$$

Clearly, the scaling behavior of $S_{k,upper}$ is the same as that of $S_{Eq,k}$. Indeed, $S_{k,upper}$ takes the same form as $S_{Eq,k}$ except the difference between the coefficients (η_k, η_{min}, η_{max}, σ_k^2, σ_{min}^2 and σ_{max}^2). As a result, the scaling order of $S_{Eq,k}$ and $S_{k,upper}$ depends only on the form and the order of L, W and n, thereby concluding the proof of (97).

References

[1] Helmut Bölcskei, Rohit U. Nabar, Özgü r Oyman, and Arogyaswami J. Paulraj. Capacity scaling laws in MIMO relay networks. *IEEE Transactions on Wireless Communications*, June 2006. to appear.

[2] Stephen Boyd and Lieven Vandenberghe. *Convex Opitmization.* (see http://www.stanford.edu/ boyd/cvxbook.html), 2004.

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

[4] Amir F. Dana and Babak Hassibi. On the power-efficiency of sensory and ad-hoc wireless networks. *IEEE Transactions on Information Theory*, July 2006. to appear.

[5] Olivier Dousse, Massimo Franceschetti, and Patrick Thiran. Information theoretic bounds on the throughput scaling of wireless relay networks. In *Proceedings of Infocom*, Miami, March 2005.

[6] Robert G. Gallager. Finding parity in a simple broadcast network. *IEEE Transactions on Information Theory*, 34(2):176–180, March 1988.

[7] Michael Gastpar and Martin Vetterli. On the capacity of wireless networks: The relay case. In *Proceedings of Infocom*, pages 1577–1586, June 2002.

[8] Michael Gastpar and Martin Vetterli. On the capacity of large gaussian relay networks. *IEEE Transactions on Information Theory*, 51(3):765–779, March 2005.

[9] M. Grossglauser and David N. Tse. Mobility increases the capacity of adhoc wireless networks. *IEEE/ACM Transactions on Networking*, 10(4), Aug. 2002.

[10] P. Gupta and P.R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):288–404, March 2000.

[11] Anders Høst-Madsen and Junshan Zhang. Capacity bounds and power allocation for wireless relay channel. *IEEE Transactions on Information Theory*, 51(6):2020–2040, June 2005.

[12] Gerhard. Kramer, Michael Gastpar, and Piyush Gupta. Cooperative strategies and capacity theorems for relay networks. *IEEE Transactions on Information Theory*, 51(9):3037–3063, Sept. 2005.

[13] J.Nicholas Laneman, David N. C. Tse, and Gregory W. Wornell. Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks. *IEEE Transactions on Information Theory*, 49:2415–2425, Oct. 2003.

[14] Ivana Maric and Roy D. Yates. Power and bandwidth allocation for cooperative strategies in gaussian relay networks. In *Proceedings of Asilomar Conference on Signals, Systems and Computers*, Nov. 2004. (invited).

[15] Muriel Médard. The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel. *IEEE Transactions on Information Theory*, 46(3):933–946, May 2000.

[16] Özgü r Oyman and Arogyaswami J. Paulraj. Power-bandwidth tradeoff in linear mimo interference relay network. *IEEE Transactions on Wireless Communications*, Oct. 2005. submitted.

[17] Özgü r Oyman and Arogyaswami J. Paulraj. Power-bandwidth tradeoff in linear multi-antenna interference relay networks. In *Proceedings of Allerton Conference on Communication, Control, and Computeing*, Monticello, IL, Sept. 2005.

[18] I. Emre Telatar and David N.C. Tse. Capacity and mutual information of wideband multipath fading channels. *IEEE Transactions on Information Theory*, 46(4):1384–1400, July 2000.

[19] David Tse and Pramod Viswanath. *Fundamentals of Wireless Communications*. World Wide Web, http://www.eecs.berkeley.edu/ dtse/main.pdf, 2004.

[20] Venugopal V. Veeravalli and Yingbin Liang. Gaussian orthogonal relay channels: Optimal resource allocation and capacity. *IEEE Transactions on Information Theory*, 51(6):3284 – 3289, Sept. 2005.

[21] Sergio Verdú. Spectral efficiency in the wideband regime. *IEEE Transactions on Information Theory*, 48(6):1319–1343, June 2002.

[22] Bo Wang and Junshan Zhang. Throughput scaling of wideband sensory relay networks: Cooperative relaying, power allocation and achievable rates. In *Proceedings of Infocom 06*, Barcelona, Spain, Apr. 2006. to appear.

[23] Bo Wang, Junshan Zhang, and Anders Høst-Madsen. On the capacity of MIMO relay channels. *IEEE Transactions on Information Theory*, 51(1):29–43, Jan. 2005.

[24] Liuqing Yang and G.B.Giannakis. Ultra-wideband communications: an idea whose time has come. *IEEE Signal Processing Magazine*, 21, Nov. 2004.

[25] Lizhong Zheng, David N.C. Tse, and Muried Médard. Channel coherence in the low snr regime. In *Proceedings of CISS*, pages 520–525, Princeton, March 2004.

[26] Lizhong Zheng, David N.C. Tse, and Muriel Médard. Channel coherence in the low snr regime. *IEEE Transactions on Information Theory*, Aug 2005. submitted.

Biographies

Bo Wang (S'03) received his M.S. degree from the Department of Electronic Engineering at Tsinghua University, Beijing, China, in 2002. He has been working towards his Ph.D. degree in the Department of Electrical Engineering at Arizona State University since August 2002.

His research interests are in the area of wireless relay networks and MIMO systems. His ongoing research includes information-theoretic aspects of wireless sensor networks.

Junshan Zhang was born in September 1972. He received his Ph.D. degree from the School of Electrical and Computer Engineering at Purdue University in 2000. He joined the Department of Electrical Engineering at Arizona State University in August 2000, where he is currently an Associate Professor. His research interests fall in the general area of wireless networks, spanning from the networking layer to the physical layer. His current research focuses on fundamental problems in wireless ad-hoc networks and sensor networks, including cross-layer optimization and design, network management, network information theory, stochastic analysis.

He ia a recipient of the ONR Young Investigator Award in 2005 and the NSF CAREER award in 2003. He has also received the Outstanding Research Award from the IEEE Phoenix Section in 2003. He was chair of the IEEE Communications and Signal Processing Phoenix Chapter from Jan. 2001 to Dec. 2003. he has served as a TPC co-chair for IPCCC 2006 and TPC vice chair for ICCCN 2006, and will be the general chair for IEEE Communication Theory Workshop 2007. He has been on the technical program committees for many conferences, including INFOCOM, SECON, GLOBECOM, ICC, MOBIHOC, BTOADNETS, and SPIE ITCOM. He has been an Associate Editor for IEEE Transactions on Wireless Communications since 2004.

Lizhong Zheng received the B.S and M.S. degrees, in 1994 and 1997 respectively, from the Department of Electronic Engineering, Tsinghua University, China, , and the Ph.D. degree, in 2002, from the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley. Since 2002, he has been working as an assistant professor in the Department of Electrical Engineering and Computer Sciences and the Laboratory of Information and Decision Systems at Massachusetts Institute of Technology. His research interests includes information theory, wireless communications and wireless networks. He received Eli Jury award from UC Berkeley in 2002, and IEEE Information Theory Society Paper Award in 2003.