Bursty Relay Networks in Low-SNR Regimes

Tony Q. S. Quek, Member, IEEE, and Hyundong Shin, Member, IEEE

Abstract—In a wireless network, the use of cooperation among nodes can significantly improve capacity and robustness to fading. Node cooperation can take many forms, including relaying and coordinated beamforming. However, many cooperation techniques have been developed for operation in narrowband systems for high signal-to-noise ratio (SNR) applications. It is important to study how relay networks perform in a low-SNR regime, where the available degrees of freedom is large and the resulting SNR per degree of freedom is small. In this paper, taking into account either low-power narrowband transmissions ($P \to 0$) or wideband transmissions with fixed power ($W \to \infty$), we investigate the achievable rates and scaling laws of bursty amplify-and-forward relay networks in the low-SNR regime. Specifically, our results allow us to understand the effect of different system parameters on the achievable rates and scaling laws in the low-SNR regime, and highlight the role of bursty transmissions in this regime. These results entirely depend on the geographic locations of the nodes and are applicable for both fixed and random networks. We identify four scaling regimes that depend on the growth of the number of relay nodes and the increase of burstiness relative to the SNR. We characterize the achievable rates and the scaling laws in the joint asymptotic regime of the number of relay nodes, SNR, and duty-cycle parameter. These results can serve as design guidelines to indicate when bursty transmissions are most useful.

Index Terms—Wideband relay networks, amplify-and-forward, low-SNR regime, bursty, achievable rates, scaling law.

I. INTRODUCTION

Most previous work in relay channels has primarily focused on narrowband systems for high signal-to-noise ratio (SNR) applications [1], [2]. However, it is important to study how relay communications perform in the low-SNR regime, where the available degrees of freedom is large and the resulting SNR per degree of freedom is low [3]–[5]. For example, wireless ad-hoc networks, sensor networks, and ultrawide bandwidth (UWB) networks are possible applications. In such networks, the nodes are usually inexpensive and low-power, and are not degree-of-freedom-limited. In the low-SNR regime, coherent amplify-and-forward (AF) relaying has been shown to perform poorly even in the presence of perfect knowledge of channel state information (CSI) at the relay nodes due to excessive noise amplification [3], [4]. One possible way to mitigate this effect is to use bursty transmissions at the expense of lower spectral efficiency. This idea was first proposed and studied in [6] for Gaussian parallel AF relay networks in the low-SNR regime. Subsequently, the idea was further extended to study the outage capacity in relay fading channels [4], the energy efficiency in Gaussian relay channels [3], and multi-antenna relay networks [7]. However, all these works either assume a priori local CSI available at the relay nodes [4], [7], or absence of fading [3], [6]. In the low-SNR regime, channel estimation can be challenging and acquiring perfect CSI at the relay nodes can be extremely difficult. This motivates current work to study the effect of imperfect channel estimates on the achievable rates of bursty AF relay networks in the low-SNR regime. Most relevant to our work is [5], where the achievable rates and scaling laws are studied for wideband relay networks using coherent AF with network training. However, our work essentially differs from [5] in two ways. First, we consider that transmissions occur in a bursty fashion, rather than continuous transmissions as in [5], and this eventually affects the asymptotic rates. Second, we consider a realistic path loss model and an asymmetric network model that takes into account the network geometry rather than symmetric network as in [5].

In this work, we derive the achievable rates of bursty AF relay networks in the low-SNR regime. The motivation for imposing bursty transmissions is twofold: (i) excessive noise amplification at the relay nodes can be reduced in the low-SNR regime; and (ii) channel uncertainty during the network training process can be significantly reduced due to higher training power. We consider both narrowband and wideband relay networks in the low-SNR regime. We identify four scaling regimes — depending on the growth of the number of relay nodes and the increase of burstiness in transmissions relative to the SNR — and characterize their scaling laws in the joint asymptotic regime of the number of relay nodes, SNR, and the duty-cycle parameter.

The paper is organized as follows. In Section II, we describe the network and system models. In Section III, we introduce the system model for bursty AF narrowband relay networks. In addition, we derive the achievable rates and scaling laws of bursty AF narrowband relay networks in the joint asymptotic

---

1By low-SNR regime, we refer to the regime that either results from low-power narrowband transmissions ($P \to 0$) or wideband transmissions with fixed power ($W \to \infty$). Note that in both scenarios, the SNR per degree of freedom tends to zero.
that \( g \to \infty \) the following notation is used. We give some concluding remarks in the last section. The regime. Numerical examples are given in Section V. Finally, the achievable rates and scaling laws in the joint asymptotic regime assumption, the distance \( d \) is assumed to be much larger than \( d_{\text{min}} \). Although such a network geometry may seem to oversimplify realistic wireless networks, it allows tractable analysis and has shown to provide valuable insights and design guidelines for general networks [10]–[13].

To obtain a clear understanding of the relation among three key parameters, namely, the number of relay nodes \( K \), SNR per fading block per degree of freedom \( \text{SNR} \), and duty-cycle \( \theta \), and their effect on the achievable rates and scaling behavior, we define

\[
\theta \triangleq \text{SNR}^{-1-\delta} \quad \text{and} \quad K \triangleq \text{SNR}^{-\epsilon}
\]

(2) for notational convenience. Although there is no interpretation on the physical correspondence between these parameters, these definitions allow us to capture the joint scaling between these parameters, which affects the achievable rates and scaling behavior of the networks. In this work, we focus on the low-SNR regime (SNR \( \leq 1 \)) and let \( \delta \in [0, 1] \) and \( \epsilon \geq 0 \).

III. LOW-POWER NARROWBAND RELAY NETWORKS

A. Bursty Channel Estimation

For small-scale fading, we denote the channel gains from the source to \( k \)th relay node and from \( k \)th relay to destination node as \( h_{B,k} \) and \( h_{F,k} \), respectively. Typically, we can model each random channel gain as a circularly symmetric complex Gaussian random variable (r.v.), i.e., \( h_{B,k} \sim \mathcal{CN}(0, 1) \) and \( h_{F,k} \sim \mathcal{CN}(0, 1) \) for all \( k \), where \( \mathcal{CN}(\mu, \sigma^2) \) denotes a complex circularly symmetric Gaussian distribution with mean \( \mu \) and variance \( \sigma^2 \). Furthermore, we consider a block fading channel model parameterized by a finite and fixed coherence time with duration \( T_{\text{coh}} \) [15]. The channel is partitioned into multiple flat fading blocks, where \( h_{B,k} \) and \( h_{F,k} \) within each block of duration \( T_{\text{coh}} \) for the \( k \)th relay node is constant. Each symbol duration is \( T_s \), which is assumed to be equal to \( 1/B \). Each fading block consists of \( N \) symbols, so the coherence time is fixed as \( T_{\text{coh}} = NT_s \). At each node, we impose an average transmitted power constraint \( P \) over each fading block.

The coherence block length is divided into training, which is placed at the beginning, and data transmission. Channel estimation is carried out at the relay nodes, where each

\[
\text{PL}(d) = \begin{cases} 
1/d_{\text{coh}} & \text{if } d > d_{\text{coh}} \\
1/d_{\text{min}} & \text{if } d < d_{\text{min}}
\end{cases}
\]

(1) where \( d \in \{1, 2, \ldots, K\} \) and \( d_{\text{min}} \) ensures that we have a bounded attenuation when nodes come close together, i.e., the average received power does not exceed the average transmit power. Similarly, the above model also holds for the forward channels, where the distance between the \( k \)th relay node and the destination node is denoted by \( d_{F,k} \).

As we increase \( \delta \) from 0 to 1, the fraction of transmission time increases from SNR to 1 so the transmission changes from bursty to continuous [14]. Throughout our analysis, we have intentionally separate topology dependent factor, \( \text{PL}(d) \), from the SNR. In the low SNR regimes, as long as \( \text{SNR} \to 0, \text{PL}(d)\text{SNR} \) will also approaches zero since \( \text{PL}(d) \) is always less than one (see eq. (1)).

Each is assumed to be independent and identically distributed (i.i.d.) across all the relay nodes. The independence assumption arises due to the presence of different propagation paths and scatterers for each relay node.

This model is different from [5] where the coherence time is assumed to scale with the SNR, i.e., the channel can become more and more underspread as the SNR decreases. However, this restriction on the channel is not physically realizable in practice.
relay node obtains minimum mean squared error (MMSE) estimates of the backward and forward channels, based on the pilot symbols transmitted during the training phase. However, different from [5], we allow the nodes to transmit bursty signals as illustrated in Fig. 2. Specifically, transmission occurs over one fading block and remains silent for the rest of the time. We refer to these time intervals as ‘bursty’. The bursty transmission in (3) takes a duration of $T_{coh}/\theta$ to transmit one fading block. We assume that the bursty channel estimation is done using a pilot symbol with $\eta$ fraction of the total transmission energy, and this training energy is denoted as

$$E_p = \frac{\eta P T_{coh}}{\theta} = \frac{\eta NP}{\theta B}$$

where $\eta \in (0, 1)$. As discussed in [16], [17], the reliability of the channel estimates depends on the effective energy per degrees of freedom, which directly depends on the coherence length $N$, the training power allocation $\eta$, and the duty-cycle parameter $\theta$ for a given power $P$ and bandwidth $B$ as shown in (3).7

During the first hop, the source node first broadcasts a bursty common pilot signal to all the relay nodes, followed by its bursty data transmission, and remains silent for the rest of the time. At the beginning of the second hop, the destination node broadcasts a common pilot signal to all the relay nodes.8 Similarly to the first hop, the transmission in the second hop over a fading block also occurs in a bursty manner with duty-cycle $\theta$. Based on this bursty network training phase, the received pilot symbols at the $k$th relay node after appropriate sampling and filtering are given by

$$Y_{P,k} = \begin{cases} \frac{1}{N_0} \text{PL}(d_{B,k}) h_{B,k} X_p + Z_{B,k}, & \text{for backward channel} \\ \frac{1}{N_0} \text{PL}(d_{F,k}) h_{F,k} X_p + Z_{F,k}, & \text{for forward channel} \end{cases}$$

where $X_p$ is the pilot symbol with energy normalized to one, $Z_{B,k} \sim \mathcal{CN}(0, 1)$ and $Z_{F,k} \sim \mathcal{CN}(0, 1)$ are independent across all $k \in \mathcal{K}$. Let $\hat{h}_{B,k}$ and $\hat{h}_{F,k}$ be the channel estimators at the $k$th relay node for the backward and forward channel coefficients, respectively. Then, we can write their estimation errors as $\tilde{h}_{B,k} = h_{B,k} - \hat{h}_{B,k}$ and $\tilde{h}_{F,k} = h_{F,k} - \hat{h}_{F,k}$, respectively. Applying the MMSE estimation, the variances of the zero-mean channel estimation error associated with $\tilde{h}_{B,k}$ and $\tilde{h}_{F,k}$ are given by

$$\sigma_{\tilde{h}_{B,k}}^2 = \frac{1}{\eta SNR} \text{PL}(d_{B,k}) + 1$$

$$\sigma_{\tilde{h}_{F,k}}^2 = \frac{1}{\eta SNR} \text{PL}(d_{F,k}) + 1$$

(5)

where $\text{SNR} = \frac{P/(N_0 B)}{\delta}$ is the SNR per fading block per degree of freedom for the narrowband relay network. Clearly, from (4), we can observe that sending pilot signals with higher $\delta$ or having a smaller path loss attenuation would result in better channel estimates due to smaller estimation error variances.

B. Bursty Coherent AF Relaying

During a data transmission period in the first hop, the received signal at the $k$th relay node is given by

$$Y_{R,k}(t) = \begin{cases} \sqrt{\frac{(1-\eta)\text{PL}(d_{B,k}) h_{B,k} X(t) + Z_{R,k}(t)}{\eta P}}, & \text{for } T_s \leq t \leq T_{coh} \\ 0, & \text{for } T_{coh} \leq t \leq T_{coh}/\theta \end{cases}$$

(6)

where $X(t)$ is the transmitted signal from the source node with normalized power equal to one and $Z_{R,k}(t)$ is complex Gaussian noise at the $k$th relay node with one-sided power spectral density (PSD) of $N_0$. In the beginning of the second hop, the relay nodes carry out the channel estimation for the forward channels and apply coherent AF relaying while satisfying the average power constraint. The phase alignment by using the coherent AF relaying is crucial for achieving the full diversity combining gain so that the signal parts of the received signal at the destination node can add up coherently. During the bursty data transmission period, the transmitted signal at the $k$th relay node is given by

$$X_{R,k}(t) = \beta_k Y_{R,k}(t - T_{coh}/\theta)$$

(7)

for $T_{coh}/\theta + T_s \leq t \leq T_{coh}(1 + 1/\theta)$ with

$$\beta_k = \frac{\hat{h}_{B,k}^*}{|\hat{h}_{B,k}|} \sqrt{\frac{\text{SNR}^\delta}{(N-1)\text{PL}(d_{B,k})}}$$

(8)

where $\beta_k$ is the amplification gain for the $k$th relay node. In the low-SNR regime, when $0 < \delta \leq 1$, we have $\frac{\text{SNR}^\delta}{\text{PL}(d_{B,k})} \ll 1$ with probability 1 in (8). Therefore, we can approximate $\beta_k$ as

$$\beta_k \approx \frac{\hat{h}_{B,k}^*}{|\hat{h}_{B,k}|} \frac{\tilde{h}_{F,k}^*}{|\tilde{h}_{F,k}|} \sqrt{\text{SNR}^\delta}$$

(9)

where $\text{SNR}^\delta$ can be interpreted as the normalized transmitted energy per fading block per degree of freedom. From (9), we observe that a smaller $\text{SNR}^\delta$ would lead to a smaller
amplification factor for relaying. However, it is not obvious whether the overall SNR at the destination node will decrease as it depends on how the number of relay nodes scale with respect to SNR

On the other hand, when \( \delta = 0 \), we have \( \text{SNR}^{\delta} \to 1 \) and \( N(1-\eta)\text{SNR}^{\delta} \approx (1-\eta)\text{SNR} \) for large \( N \) in (8). In this case, we can approximate \( \beta_k \) as

\[
\beta_k \approx \frac{\hat{h}_{B,k}}{|h_{B,k}|} \frac{1}{\sqrt{1+(1-\eta)\text{SNR}}}.
\]

Notice that if the amplification factor is independent of the SNR and depends only on the location of the relay nodes and \( \eta \). This indicates that bursty transmissions can alleviate the problem of noise amplification in AF relaying. Similar observations have been made by [3], [6], [8].

At the destination node, the received signal can be written as

\[
Y_D(t) = \sum_{k \in \mathcal{K}} \beta_k \sqrt{\text{PL}(d_{f,k})h_{F,k}Y_{R,k}(t-T_{coh}/\theta)} + Z(t)
\]

for \( T_{coh}/\theta + T_c \leq t \leq T_{coh}(1+1/\theta) \), where \( Z(t) \) is complex Gaussian noise at the destination node with one-sided PSD of \( N_0 \). By sampling (11) at a rate \( 1/B \), the discrete-time equivalent of (11) for \( i = 0, \ldots, N-1 \) is given by (12), shown at the top of this page with

\[
G_k(\delta) = \begin{cases} 
\left( \frac{N(1-\eta)}{(N-1)} \right) \text{SNR}^{\delta} \text{PL}(d_{f,k}) \text{PL}(d_{p,k}) & 0 < \delta \leq 1 \\
\left( \frac{N(1-\eta)}{(N-1)} \right) \frac{\text{PL}(d_{f,k})\text{PL}(d_{p,k})}{1+(1-\eta)\text{SNR}} & \delta = 0
\end{cases}
\]

\[
G'_k(\delta) = \begin{cases} 
\frac{\text{PL}(d_{f,k})\text{SNR}^{\delta}}{1+(1-\eta)\text{SNR}} & 0 < \delta \leq 1 \\
\frac{\text{PL}(d_{f,k})\text{SNR}^{\delta}}{1+(1-\eta)\text{SNR}} & \delta = 0
\end{cases}
\]

where \( Z_{R,k}[i] \) and \( Z[i] \) are normalized such that \( Z_{R,k}[i] \sim \mathcal{CN}(0,1) \) and \( Z[i] \sim \mathcal{CN}(0,1) \).

C. Achievable Rates of Bursty AF Narrowband Relay Networks

Before we present the achievable rates, we first introduce an useful result that characterizes the upper bound on the capacity of narrowband relay networks.

**Lemma 1** ([5, Lemma 4.11]): The capacity of narrowband relay networks is upper bounded by

\[
C_{\text{upper}} \triangleq 2B \log \left( 1 + \text{SNR} \sum_{k \in \mathcal{K}} |h_{B,k}|^2 \right)
\]

and \( C_{\text{upper}} \) scales as

\[
C_{\text{upper}} = \begin{cases} 
\Theta(\log(K\text{SNR})), & \epsilon < 1 \\
O(1), & \epsilon \geq 1.
\end{cases}
\]

Next, we derive two achievable rates of the bursty AF narrowband relay networks. The first achievable rate corresponds to the case when the relay nodes have perfect bidirectionally local CSI, namely, \( h_{B,k} \) and \( h_{F,k} \) at each \( k \)th relay node. Assuming that the destination node has knowledge of the channel statistics of \( \{h_{B,k}\}_{k \in \mathcal{K}} \) and \( \{h_{F,k}\}_{k \in \mathcal{K}} \), the maximum achievable rate of bursty AF narrowband relay networks is given in the following proposition.

**Proposition 1**: For a regular linear network, the maximum achievable rate of bursty AF narrowband relay networks with duty-cycle \( \theta = \text{SNR}^{\delta} \) satisfies (17) when \( 0 < \delta \leq 1 \) and (18) when \( \delta = 0 \).

**Proof**: See Appendix A.

**Remark 1**: Although the results above are derived for regular linear networks, we can easily obtain bounds for random networks. For example, consider a dense random network where the node locations are chosen to be i.i.d. from a disk of unit area. We can obtain similar bounds for this random network by replacing \( \text{PL}(Kd) \) and \( \text{PL}(1-d) \) with \( \text{PL}(2/\sqrt{\pi}) \), where \( 2/\sqrt{\pi} \) is the largest distance in the network. Such an approach has also been used in [2]. Therefore, even for a simple network topology such as regular linear networks, we can capture similar interaction between \( K \), \( \theta \), and SNR as in a random network. In the following, we give further insights on the scaling behavior of these rates in different regimes according to the joint scaling of relevant parameters.

**Definition 1**: In bursty AF relay networks, we can characterize the asymptotic achievable rates into four regimes:

- **Regime 1**: \( 0 < \delta \leq 1 \) and \( \epsilon < \delta \).
- **Regime 2**: \( 0 < \delta \leq 1 \) and \( \epsilon = \delta \).
- **Regime 3**: \( 0 < \delta \leq 1 \) and \( \epsilon > \delta \).
- **Regime 4**: \( \delta = 0 \) and \( \epsilon \geq 0 \).
Remark 2: Clearly, we can see that all the above regimes depend on how SNR$^3$ and $K$SNR$^3$ scale with respect to SNR as SNR → 0. In Regime 1, the normalized transmitted energy per fading block per degree of freedom SNR$^3$ as well as the aggregate normalized relay energy per fading block per degree of freedom $K$SNR$^3$ go to zero as SNR → 0. In Regime 2, we have SNR$^3$ → 0 as well as $K$SNR$^3$ approaches a non-zero constant. In Regime 3, we have SNR$^3$ → 0 as well as $K$SNR$^3$ grows unbounded. Lastly, in Regime 4, SNR$^3$ does not diminish to zero and $K$SNR$^3$ remains a non-zero constant.

Theorem 1: In the low-SNR regime and perfect local CSI at the relay nodes, the scaling behavior of the achievable rate of the bursty AF narrowband relay networks with duty-cycle $\theta = \text{SNR}^{1-\delta}$ is given by

$$ R = \begin{cases} 
\Theta(\text{SNR}^{1+\delta-2\epsilon}), & \text{Regime 1} \\
\Theta(\text{SNR}^{-\delta}), & \text{Regime 2} \\
\Theta(\text{SNR}^{1-\delta} \log \text{SNR}^{-\epsilon}), & \text{Regime 3} \\
\Theta(\text{SNR} \log K), & \text{Regime 4} 
\end{cases} \quad (19) $$

Proof: First, we consider the case when $0 < \delta \leq 1$. In this case, we have SNR$^3$ approaches zero as SNR → 0. It is clear that the bounds in (17) depends on the joint scaling of $K$SNR$^3$ with respect to SNR and $K$. In the low-SNR regime, depending on the rate at which SNR and $K$ scales, the aggregate normalized relay energy per fading block per degree of freedom can be asymptotically zero, bounded below or unbounded according to the different regimes shown as follows:

$$ \lim_{\text{SNR} \to 0} K\text{SNR}^3 = \lim_{\text{SNR} \to 0} \text{SNR}^{3-\epsilon} = \begin{cases} 
0, & 0 < \epsilon < \delta \leq 1 \\
1, & 0 < \epsilon = \delta \leq 1 \\
\infty, & 0 < \delta < \epsilon.
\end{cases} $$

Using (20), we can characterize the scaling behavior of the bounds in (17) as follows:

1) When $K$SNR$^3$ is asymptotically zero, we have $K$SNR$^3$PL$(1-d) \ll 1$ and $K$SNR$^3$PL$(d_{\min}) \ll 1$ as $K \to \infty$ and SNR → 0. In this case, the achievable rate scales like

$$ R = \Theta(\text{SNR}^{1-\delta} K^2 \text{SNR}^{2\delta}) = \Theta(\text{SNR}^{1+\delta-2\epsilon}). \quad (21) $$

2) When $K$SNR$^3$ is bounded below, the achievable rate scales like $R = \Theta(\text{SNR}^{-\delta})$.

3) When $K$SNR$^3$ grows unbounded, we have $K$SNR$^3$PL$(1-d) \gg 1$ and $K$SNR$^3$PL$(d_{\min}) \gg 1$ as $K \to \infty$ and SNR → 0. Thus, the achievable rate for this case scales like $R = \Theta(\text{SNR}^{1-\delta} \log(K\text{SNR}^3))$.

Next, when $\delta = 0$, we have SNR$^3$ approaches one as SNR → 0. It is clear that the bounds in (18) only depends on scaling behavior of $K$ and SNR. For large $K$, we have $K$SNR$^3$PL$(1-2d_{\min}) \gg 1$ and $K$SNR$^3$PL$(d_{\min}) \gg 1$, and the achievable rate in (18) scales like $R = \Theta(\text{SNR} \log K)$. □

Remark 3:

- When the normalized transmitted energy per fading block per degree of freedom approaches zero in the low-SNR regime, cooperation is key to maintain a favorable scaling behavior of the achievable rate. Specifically, from Theorem 1, we see that the growth rate of the number of relay nodes is important since it directly affects the asymptotic behavior of the normalized aggregate relay energy per fading block per degree of freedom.
- When $\epsilon < \delta$ as in Regime 1, the normalized aggregate relay energy per fading block per degree of freedom still approaches zero in the limit of large $K$ and low SNR. As a result, the achievable rate scales poorly as $\Theta(\text{SNR}^{1+\delta-2\epsilon})$. In this regime, the burstiness only helps to provide a better vanishing rate.
- When $\epsilon = \delta$ as in Regime 2, the normalized aggregate relay energy per fading block per degree of freedom remains bounded in the limit of large $K$ and low SNR. Therefore, the achievable rate scales better than Regime 1 and we obtain $\Theta(\text{SNR}^{1-\delta})$. With cooperation, we achieve a higher achievable rate compared to point-to-point low-SNR capacity of $O(\text{SNR})$.
- When $\epsilon > \delta$ as in Regime 3, the normalized aggregate relay energy per fading block per degree of freedom grows with the number of relay nodes. In this regime, we observe that cooperation can offset the effect of vanishing SNR in the low-SNR regime. Note that when $\delta = 1$, we obtain the same scaling law of $\Theta(\log K)$. In [5], it has the same scaling order as $C_{\text{upper}}$ in Lemma 1. Unlike in the previous two regimes, bursty transmissions are not preferable since it hurts us in terms of spectral efficiency.

- When $\delta = 0$ as in Regime 4, the achievable rate scales as $\Theta(\text{SNR} \log K)$ in the limit of large $K$ and low SNR. However, we do not need to constrain the way we scale $K$ compared to Regimes 1–3. Any increasing rate of $K$ will contribute to a non-zero multiplicative term in the low-SNR capacity. As a result, this shows another example how we can achieve a higher rate compared to point-to-point low-SNR capacity through the effective use of cooperation.

Note that the low-SNR scaling laws for the achievable rate of bursty AF narrowband relay networks in Theorem 1 assume perfect local CSI at each relay node. Assuming that the destination node only has partial knowledge of the global CSI in terms of the channel statistics of $\{h_{B,k}\}$ and $\{h_{F,k}\}$, the achievable rate of bursty AF narrowband relay networks using network channel estimation is given in the following proposition.

Proposition 2: For a fixed deployment of the nodes, the achievable rate of bursty AF narrowband relay networks with duty-cycle $\theta = \text{SNR}^{1-\delta}$ and network training is given by

$$ R = \frac{\theta(N-1)}{N} B \log (1 + \text{SNR}_{\text{eff}}) $$

where $\text{SNR}_{\text{eff}} \triangleq \sigma_X^2 / (\sigma_{V_1}^2 + \sigma_{V_2}^2 + \sigma_{V_3}^2 + 1)$ such that $\sigma_X^2$, $\sigma_{V_1}^2$, $\sigma_{V_2}^2$, and $\sigma_{V_3}^2$ are given by (23), (24), (25), and (26), respectively, as shown at the top of next page.

Proof: See Appendix B. □

Proposition 3: For a regular linear network, the effective SNR at the destination node in Proposition 2 satisfies (27) and (28) when $0 < \delta \leq 1$, and (29) and (30) when $\delta = 0$.

Proof: See Appendix C. □

Theorem 2: In the low-SNR regime, the scaling behavior of the achievable rate of bursty AF narrowband relay networks with duty-cycle $\theta = \text{SNR}^{1-\delta}$ and network training is given
\[
\sigma_X^2 = \frac{\pi^2}{16} (\eta N \text{SNR}^3)^2 \sum_{k \in \mathcal{K}} \frac{G_k(\delta)PL(d_{b,k})PL(d_{f,k})}{(\eta N \text{PL}(d_{b,k})\text{SNR}^3 + 1)(\eta N \text{PL}(d_{f,k})\text{SNR}^3 + 1)}
\]

(23)

\[
\sigma_{V_1}^2 = \sum_{k \in \mathcal{K}} G_k(\delta) \frac{\eta N \text{PL}(d_{b,k})\text{SNR}^3 + 1}{(\eta N \text{PL}(d_{b,k})\text{SNR}^3 + 1)(\eta N \text{PL}(d_{f,k})\text{SNR}^3 + 1)}
\]

(24)

\[
\sigma_{V_2}^2 = \sum_{k \in \mathcal{K}} G_k^2(\delta)
\]

(25)

\[
\sigma_{V_3}^2 = \left(1 - \frac{\pi^2}{16}\right) (\eta N \text{SNR}^3)^2 \sum_{k \in \mathcal{K}} \frac{G_k(\delta)PL(d_{b,k})PL(d_{f,k})}{(\eta N \text{PL}(d_{b,k})\text{SNR}^3 + 1)(\eta N \text{PL}(d_{f,k})\text{SNR}^3 + 1)}
\]

(26)

\[
\text{SNR}_{\text{eff}} \leq \frac{\pi^2 K^2 \text{SNR}^3 N(1-\eta)}{16} \left(\frac{\eta N \text{PL}(d_{\min})\text{SNR}^3}{(\eta N \text{PL}(d_{\min})\text{SNR}^3 + 1)}\right)^2
\]

(27)

\[
\text{SNR}_{\text{eff}} \geq \frac{K^2 \text{SNR}^3 \text{PL}(Kd)\text{PL}(1-d)}{16} \left(\frac{\eta N \text{PL}(d_{\min})\text{SNR}^3}{(\eta N \text{PL}(d_{\min})\text{SNR}^3 + 1)}\right)^2
\]

(28)

\[
\text{SNR}_{\text{eff}} \leq \frac{K^2 \text{SNR}^3 \text{PL}(1-d)}{16} \left(\frac{\eta N \text{PL}(d_{\min})\text{SNR}^3}{(\eta N \text{PL}(d_{\min})\text{SNR}^3 + 1)}\right)^2
\]

(29)

\[
\text{SNR}_{\text{eff}} \geq \frac{K^2 \text{SNR}^3 \text{PL}(1-d)}{16} \left(\frac{\eta N \text{PL}(d_{\min})\text{SNR}^3}{(\eta N \text{PL}(d_{\min})\text{SNR}^3 + 1)}\right)^2
\]

(30)

by

\[
R = \begin{cases} 
\Theta(\text{SNR}^{1+3\delta-2\epsilon}), & \text{Regime 1} \\
\Theta(\text{SNR}^{1-\epsilon}), & \text{Regime 2} \\
\Theta(\text{SNR}^{1-\delta} \log K \text{SNR}^{3\delta}), & \text{Regime 3} \\
\Theta(\text{SNR} \log K), & \text{Regime 4}
\end{cases}
\]

(31)

Proof: Using the asymptotic behavior of \(K\text{SNR}^{3\delta}\) characterized in (20), the scaling behavior of the effective SNR is given as follows:

1) When \(K\text{SNR}^{3\delta}\) is asymptotically zero, the denominators in (27) and (28) tend to 1 as \(K \to \infty\) and SNR \(\to 0\) since the first two terms in each denominator approach zero asymptotically. In this case, the effective SNR behaves like \(\text{SNR}_{\text{eff}} = \Theta(K^2\text{SNR}^{3\delta})\). Using (22), the achievable rate scales like \(R = \Theta(\text{SNR}^{1+3\delta-2\epsilon})\).

2) When \(K\text{SNR}^{3\delta}\) approaches one for large \(K\) and SNR \(\to 0\), the denominators in (27) and (28) approach PL\((d_{\min}) + 1\) as \(K \to \infty\) and SNR \(\to 0\) since the first term in each denominator approaches zero asymptotically. In this case, the effective SNR behaves like \(\text{SNR}_{\text{eff}} = \Theta(\text{SNR}^{3\delta})\). Thus, the achievable rate scales like \(R = \Theta(\text{SNR}^{1-\delta} \log K \text{SNR}^{3\delta})\).

3) When \(K\text{SNR}^{3\delta} \to \infty\), the denominators in (27) and (28) approach \(K\text{SNR}^{3\delta}\text{PL}(1-d)\) and \(K\text{SNR}^{3\delta}\text{PL}(d_{\min})\), respectively, as \(K \to \infty\) and SNR \(\to 0\) since the first term in each denominator approaches zero asymptotically. In this case, the effective SNR behaves like \(\text{SNR}_{\text{eff}} = \Theta(\text{SNR}^{3\delta})\). Thus, the achievable rate scales like \(R = \Theta(\text{SNR}^{1+3\delta-2\epsilon})\).

When \(\delta = 0\), we have \(\text{SNR}^{3\delta} \to 1\) as SNR \(\to 0\). From (29) and (30), it is clear that the effective SNR depends only on the scaling of \(K\). As a result, the effective SNR behaves like \(\text{SNR}_{\text{eff}} = \Theta(K)\). Thus, the achievable rate scales like \(R = \Theta(\text{SNR} \log K)\).

Remark 4:

- Similar to the results in Theorem 1, the rate at which the number of relay nodes grows is important in determining the asymptotic scaling behavior of the achievable rates as stated in Theorem 2.

- When \(\epsilon < \delta\) as in Regime 1, we obtain a poorer scaling of \(\Theta(\text{SNR}^{1+3\delta-2\epsilon})\) due to the effect of channel estimation errors. Similar to the result in Theorem 1, the burstiness is also useful in providing a better order.

- When \(\epsilon = \delta\) as in Regime 2, the achievable rate scales as \(\Theta(\text{SNR}^{1+\delta})\), which is worse than Theorem 1 due to the effect of imperfect channel estimation. In this regime, the burstiness can also provide a better order. However, we cannot exceed the point-to-point low-SNR capacity of \(O(\text{SNR})\).
SNR$^{1-\delta}W \log \left(1 + \frac{\frac{\pi^2 K^2 \text{SNR}^3}{16} \text{PL}(Kd) \text{PL}(d_{\text{min}})}{K \text{SNR}^3 \text{PL}(d_{\text{min}}) + 1}\right) \leq R \leq \text{SNR}^{1-\delta}W \log \left(1 + \frac{\frac{\pi^2 K^2 \text{SNR}^3}{16} \text{PL}(d_{\text{min}})^2}{K \text{SNR}^3 \text{PL}(1-d) + 1}\right) (33)

\frac{P}{N_0} \log \left(1 + \frac{\frac{\pi^2 K^2 \text{PL}(Kd) \text{PL}(d_{\text{min}})}{1 + \text{PL}(d_{\text{min}})}K \text{PL}(d_{\text{min}}) + 1}\right) \leq R \leq \frac{P}{N_0} \log \left(1 + \frac{\frac{\pi^2 K^2 \text{PL}(d_{\text{min}})^2}{1 + \text{PL}(Kd)}K \text{PL}(1-d) + 1}\right) (34)

- In Regime 3, we see that $K$ is required to scale as $\text{SNR}^{-\epsilon}$ in a way such that $\epsilon \in (3\delta, +\infty)$ to achieve the rate of $\Theta(\text{SNR}^{1-\delta} \log K \text{SNR}^3)$. When $\delta = 1$, the achievable rate becomes $\Theta(\log K \text{SNR}^3)$. Similar to Theorem 1, bursty transmissions are not preferable in this regime due to lower spectral efficiency. However, if $\epsilon \in (3\delta, 3\delta^2]$, then the burstiness helps in improving the achievable rate, but the achievable rate no longer scales as $\Theta(\log K \text{SNR})$.
- When $\delta = 0$ as in Regime 4, we obtain the same result of $\Theta(\text{SNR} \log K)$ as in Theorem 1. Thus, even in the presence of imperfect channel estimation, we can achieve a higher rate compared to point-to-point low-SNR capacity through cooperation.

**Corollary 1:** In the low-SNR regime and perfect local CSI at the relay nodes, the capacity of bursty AF narrowband relay networks scales like $\Theta(\log K \text{SNR})$ when duty-cycle $\theta$ is set at one, provided that the normalized aggregate relay energy per fading block per degree of freedom grows unbounded.

**Proof:** From Lemma 1 and Theorem 1 (Regime 3), the achievable rate has the same scaling order as the upper bound on the capacity, therefore yielding the desired scaling law. □

**Corollary 2:** In the low-SNR regime, the achievable rate of bursty AF narrowband relay networks scales like $\Theta(\text{SNR} \log K)$ when the normalized transmitted energy per fading block per degree of freedom remains constant.

**Proof:** It follows straightforwardly from Theorems 1 and 2. □

**Remark 5:** Corollary 1 agrees with the result in [5]. Moreover, we show that even with burstiness we cannot further improve the capacity scaling law regardless of the quality of the CSI at the relay nodes.

### IV. WIDEBAND RELAY NETWORKS

For small-scale fading, it is important to take into account the presence of frequency selective fading due to increasing bandwidth $W$ in wideband relay networks. For simplicity, we adopt a well-known doubly block fading model that decomposes the time-frequency space into blocks of finite duration $T_{\text{coh}}$ and finite bandwidth $B_{\text{coh}}$, where $T_{\text{coh}}$ and $B_{\text{coh}}$ are the coherence time and coherence bandwidth, respectively [15]. Following this model, we can model a wideband channel as a set of $L$ parallel subchannels with bandwidth $B_{\text{coh}}$, and $L$ determines the number of resolvable multipaths, i.e., $W = LB_{\text{coh}}$. We further assume that the Doppler spread is negligible, leading to i.i.d. subchannels. In what follows, we let $B = B_{\text{coh}}$. With block fading in the frequency domain, we consider that the source transmits over $L$ subchannels with bandwidth $B$ (equivalent to the coherence bandwidth) to the relay nodes in the first hop. In addition, we assume that the source distributes the power uniformly over the $L$ subchannels while satisfying the average transmit power constraint $P$ over each block. As a result, the SNR per fading block per degree of freedom for the wideband relay network becomes $\text{SNR} = P/(LN_0B)$. In the low-SNR regime of wideband networks, we have $\text{SNR} \rightarrow 0$ when $W \rightarrow \infty$ for fixed $P$. Similar to Section III-A, we consider bursty channel estimation with duty-cycle parameter $\theta$, except that this process is performed over $L$ orthogonal subchannels and the pilot symbol energy for each subchannel is equal to $\eta NPT/(L\theta)$.

#### A. Achievable Rates of Bursty AF Wideband Relay Networks

We now establish the scaling laws of the achievable rate of bursty AF wideband relay networks. For brevity, we omit the proofs since they are straightforward from the counterpart theorem for the narrowband case and the same argument as in the proof of Proposition 4.

**Proposition 4:** For a regular linear network, the maximum achievable rate of bursty AF wideband relay networks with duty-cycle $\theta = \text{SNR}^{1-\delta}$ satisfies (33) when $0 < \delta \leq 1$, and (34) when $\delta = 0$.

**Proof:** It follows straightforwardly from the proof for Proposition 2, except that the achievable rate of the wideband relay networks is given by the sum rate of $L$ parallel subchannels. Moreover, since the power allocation over the subchannels is uniform, the achievable rate of bursty AF wideband relay networks simply becomes $L$ times the achievable rate of bursty AF narrowband relay networks in the low-SNR regime. □

**Theorem 3:** In the low-SNR regime, the scaling behavior of the achievable rate of bursty AF wideband relay networks with duty-cycle $\theta = \text{SNR}^{1-\delta}$ and perfect local CSI at the relay nodes is given by

$$R = \begin{cases} \Theta(W^{2\epsilon - \delta}), & \text{Regime 1} \\ \Theta(W^\delta), & \text{Regime 2} \\ \Theta(W^{\delta} \log K \text{SNR}^3), & \text{Regime 3} \\ \Theta(\log K), & \text{Regime 4} \end{cases} (35)$$

where $\text{SNR} = P/(N_0W)$ and $K \approx \text{SNR}^{-\epsilon}$ such that $\epsilon \geq 0$.

**Theorem 4:** In the low-SNR regime, the scaling behavior of the achievable rate of bursty AF wideband relay networks with duty-cycle $\theta = \text{SNR}^{1-\delta}$ and network training is given by

$$R = \begin{cases} \Theta(W^{2\epsilon - 3\delta}), & \text{Regime 1} \\ \Theta(W^\delta), & \text{Regime 2} \\ \Theta(W^{\delta} \log K \text{SNR}^3), & \text{Regime 3} \\ \Theta(\log K), & \text{Regime 4} \end{cases} (36)$$

**Remark 6:**

- When the normalized transmitted energy per fading block per degree of freedom approaches zero, the growth rate of the number of relay nodes determines the scaling
behavior of the effective SNR at the destination node and eventually affects the scaling order of the achievable rate. Moreover, this rate approaches zero when $\epsilon < \delta/2$ and grows unbounded when $\epsilon > \delta/2$.

- In Regime 3, cooperation can offset the vanishing SNR in the wideband regime. When $\delta = 1$, the achievable rate scales as $\Theta(W \log K \text{SNR})$ consistent with [5]. Since we are in the logarithm region, spectral efficiency is more important so bursty transmissions are not preferable.

- Interestingly, when $\delta = 0$ as in Regime 4, we obtain similar scaling law as in [2] even in the wideband regime. Intuitively, this is parallel to the point-to-point capacity result in fading channels, which shows that the capacity of AWGN can be achieved through peaky signaling scheme [18]. However, we show that equivalent results also hold for wideband relay networks.

- Even in the presence of imperfect channel estimation, the scaling order $\Theta(\log K)$ can be achieved in the wideband regime through bursty transmissions with $\delta = 0$.

V. NUMERICAL EXAMPLES

In the section, we give numerical examples of the achievable rates in the low-SNR regime. Figs. 3 and 4 show the asymptotic achievable rates as a function of SNR for bursty AF narrowband relay networks in Regimes 1 and 2, respectively. In these figures, we have set $\delta$ to be 0.01, 0.5 and 0.9 for comparison. It can be seen that imperfect channel estimation with network training leads to lower rates in both figures. However, this reduction becomes smaller as $\delta$ becomes smaller, showing that the burstiness improves the performance of network channel estimation. As a result, the achievable rates with network training improve with the burstiness for both cases.

In Fig. 5, we plot the asymptotic achievable rates as a function of SNR for bursty AF narrowband relay networks in Regime 3 with $\epsilon = 4$. In addition, we plot the cut-set bound $C_{\text{upper}}$ given in Lemma 1 for comparison. Similar to previous observations in Figs. 3 and 4, the gap between the rates with and without perfect local CSI reduces as $\delta$ becomes smaller. Although the burstiness improves channel estimation performance, it leads to a loss in spectral efficiency as seen from the figure. As a result, bursty transmissions are not preferable in this regime. Next, we show in Fig. 6 the asymptotic achievable rates as a function of SNR for bursty AF narrowband relay networks in Regime 4 with $\epsilon = 2, 4$ and 6. For comparison, we also plot the cut-set bound $C_{\text{upper}}$ in Fig. 6. In this regime, the effect of channel estimation is not an issue. However, the rate greatly suffers from the loss in spectral efficiency due to the bursty transmissions. Thus, to recover this spectral efficiency loss, we need to scale the number of relay nodes significantly.

Lastly, in Fig. 7, we compare the asymptotic achievable rates for bursty AF narrowband and wideband relay networks in Regime 4. Clearly, we see that higher rates can be achieved in the wideband relay networks for all values of $\epsilon$, compared to the narrowband case. This is due to the absence of SNR
term in front of the log function for the wideband case. In addition, a larger $\epsilon$ leads to a larger rate due to the effective cooperation. Thus, we see that the bursty transmission can be very efficient in the low-SNR regime, provided the peak power constraint is not very severe.

VI. CONCLUSION

In this paper, we studied the achievable rates and scaling laws of bursty AF relay networks in the low-SNR regime. We identified four scaling regimes that depend on the growth of the number of relay nodes and the increase of burstiness relative to the SNR. We characterized the achievable rates and the scaling laws in the joint asymptotic regime of the number of relay nodes, SNR, and duty-cycle parameter. These results can serve as design guidelines to indicate when bursty transmissions are most useful. Specifically, we showed that:

- When the normalized transmitted energy per fading block per degree of freedom approaches zero in the low-SNR regime, cooperation is key to maintain a favorable scaling behavior of the achievable rate. In addition, the burstiness can help to improve the vanishing rate behavior. However, there is a tradeoff since bursty transmissions hurt in terms of spectral efficiency. This tradeoff depends on how $K$ and $\theta$ scale with SNR.

- When the normalized transmitted energy per fading block per degree of freedom remains constant due to the burstiness, we can achieve a higher rate compared to point-to-point low-SNR capacity through the cooperation.

- The effect of imperfect channel estimation is only detrimental in the scenario when the normalized transmitted energy per fading block per degree of freedom approaches zero in the low-SNR regime.

- When the normalized transmitted energy per fading block per degree of freedom remains constant due to the burstiness, the capacity of bursty wideband AF relay networks scales like $\Theta(\log K)$. This result holds even with imperfect channel estimation.

APPENDIX A
PROOF OF PROPOSITION 1

With a priori knowledge of $h_{B,k}$ and $h_{F,k}$ at the $k$th relay node, the maximum achievable rate is given by $R = \theta B \log \left(1 + \sigma_X^2/\epsilon \sigma_V^2 \right)$, where $\sigma_X^2$ is the received signal energy at the destination node given the knowledge of the backward and forward channel statistics, and $\sigma_V^2$ is the variance of the amplified noise from all the relay nodes [5]. We first derive $\sigma_X^2$ as follows:

$$
\sigma_X^2 = \mathbb{E} \left\{ \left[ \mathbb{E} \left\{ \sum_{k \in K} \sqrt{G_k(\delta)} |h_{B,k}| |h_{F,k}| \right\} X[i] \right]^2 \right\} 
$$

$$
= \frac{\pi^2}{16} \left\{ \sum_{k \in K} \sqrt{G_k(\delta)} \right\}^2,
$$

$$
= \frac{\pi^2}{16} \sum_{k \in K} \sqrt{G_k(\delta)} \left( \sum_{\delta} \frac{\text{SNR}^2}{\sum_{\delta} \sqrt{PL(d_{b,k})PL(d_{f,k})}} \right)^2, 0 < \delta \leq 1
$$

$$
= \frac{\pi^2}{16} \sum_{k \in K} \sqrt{G_k(\delta)} \left( \sum_{\delta} \frac{PL(d_{b,k})PL(d_{f,k})}{1 + PL(d_{b,k})} \right)^2, \delta = 0
$$

where (a) follows from the fact that $|h_{B,k}|$ and $|h_{F,k}|$ are Rayleigh distributed with mean $\sqrt{\pi}/4$; (b) follows by substituting (13) without network channel estimation. For regular linear networks, $\text{PL}(d_{B,k})$ and $\text{PL}(d_{F,k})$ are bounded by

$$
\text{PL}(Kd) \leq \text{PL}(d_{b,k}) \leq \text{PL}(d_{\text{min}})
$$

$$
\text{PL}(1 - d) \leq \text{PL}(d_{f,k}) \leq \text{PL}(d_{\text{min}}).
$$

(37)

As a result, we can bound $\sigma_X^2$ as given in (38) and (39), shown on the top of next page. Similarly, we can derive $\sigma_V^2$ as follows:

$$
\sigma_V^2 = \mathbb{E} \left\{ \left[ \sum_{k \in K} \sqrt{G_k(\delta)} \left( \frac{h_{B,k}^* |h_{F,k}|^2}{|h_{B,k}||h_{F,k}|} \right) Z_{B,k}[i] \right]^2 \right\}
$$

$$
= \sum_{k \in K} G_k(\delta) \mathbb{E} \left\{ |h_{F,k}|^2 \right\}
$$

Fig. 6. Bursty AF narrowband relay networks: Achievable rates in Regime 4. The dashed and solid lines denote the achievable rate and cut-set bound, respectively.

Fig. 7. Achievable rates for bursty AF narrowband and wideband relay networks.
\[
\frac{\pi^2}{16} K^2 \text{SNR}^2 \text{PL}(Kd)\text{PL}(d_{\text{min}}) \leq \sigma_X^2 \leq \frac{\pi^2}{16} K^2 \text{SNR}^2 \text{PL}(d_{\text{min}})^2, \quad \text{when } 0 < \delta \leq 1 \\
\frac{\pi^2}{16} K^2 \text{PL}(Kd)\text{PL}(d_{\text{min}}) \leq \sigma_X^2 \leq \frac{\pi^2}{16} K^2 \left(\text{PL}(d_{\text{min}})^2\right), \quad \text{when } \delta = 0.
\]

\[
\sigma_X^2 = \mathbb{E} \left\{ \mathbb{E} \left\{ \sum_{k \in \mathcal{K}} \sqrt{G_k(\delta)} \hat{h}_{B,k} \hat{h}_{F,k} \right\} X[i]^2 \right\} \\
= \frac{\pi^2}{16} \eta \text{NSNR}^2 \left( \sum_{k \in \mathcal{K}} \sqrt{G_k(\delta)\text{PL}(d_{B,k})\text{PL}(d_{F,k})} \right)^2 
\]

\[
\sigma_Y^2 = \mathbb{E} \left\{ \sum_{k \in \mathcal{K}} \sqrt{G_k(\delta)} \left( \frac{\hat{h}_{B,k} \hat{h}_{F,k}}{|\hat{h}_{B,k}|^2|\hat{h}_{F,k}|} + \frac{\hat{h}_{B,k} \hat{h}_{F,k}}{|\hat{h}_{B,k}|^2|\hat{h}_{F,k}|} + \frac{\hat{h}_{B,k} \hat{h}_{F,k}}{|\hat{h}_{B,k}|^2|\hat{h}_{F,k}|} \right) \right\}^2 \\
\stackrel{(a)}{=} \sum_{k \in \mathcal{K}} G_k(\delta) \left( \mathbb{E} \{ |\Phi_{1,k}|^2 \} + \mathbb{E} \{ |\Phi_{2,k}|^2 \} + \mathbb{E} \{ |\Phi_{3,k}|^2 \} \right) \\
\stackrel{(b)}{=} \sum_{k \in \mathcal{K}} G_k(\delta) \left( \mathbb{E} \{ |\hat{h}_{B,k}|^2 \} \mathbb{E} \{ |\hat{h}_{F,k}|^2 \} + \mathbb{E} \{ |\hat{h}_{F,k}|^2 \} \mathbb{E} \{ |\hat{h}_{B,k}|^2 \} + \mathbb{E} \{ |\hat{h}_{B,k}|^2 \} \mathbb{E} \{ |\hat{h}_{F,k}|^2 \} \right) \\
\stackrel{(c)}{=} \sum_{k \in \mathcal{K}} G_k(\delta) \left( \frac{\eta \text{NPL}(d_{B,k}) \text{SNR}^2}{\eta \text{NPL}(d_{B,k}) \text{SNR}^2 + 1} \cdot \frac{1}{\eta \text{NPL}(d_{F,k}) \text{SNR}^2 + 1} + \frac{1}{\eta \text{NPL}(d_{B,k}) \text{SNR}^2 + 1} \cdot \frac{1}{\eta \text{NPL}(d_{F,k}) \text{SNR}^2 + 1} \right) \\
\stackrel{(d)}{=} \left\{ \begin{array}{ll} \text{SNR}^2 \sum_{k \in \mathcal{K}} \text{PL}(d_{F,k}), & 0 < \delta \leq 1 \\
\sum_{k \in \mathcal{K}} \frac{\text{PL}(d_{F,k})}{1 + \text{PL}(d_{B,k})}, & \delta = 0 \end{array} \right.
\]

where we can further bound \( \sigma_Y^2 \) using (37) as in the case for \( \sigma_X^2 \).

\section*{Appendix B
Proof of Proposition 2
}

It is well known that for a given variance, the worst-case noise among all additive noises uncorrelated with the signal is the zero-mean independent Gaussian noise from a capacity perspective (see, e.g., [19]). As such, we obtain the achievable rate of the bursty AF narrowband relay networks with network training by the capacity of an AWGN channel with the same variance as given in (22), where \( \text{SNR}_{\text{eff}} \) is the effective SNR of the source-to-destination link, \( \theta \) accounts for the duty-cycle transmission, \( (N - 1)/N \) accounts for the training within each transmission block, and 1/2 accounts for the half-duplex mode of the relay nodes.

The received signal energy at the destination node given the knowledge of the backward and forward channel statistics is denoted by \( \sigma_X^2 \), and it is given by (40), shown at the top of this page, where (a) follows from the fact that \( |\hat{h}_{B,k}| \) and \( |\hat{h}_{F,k}| \) are Rayleigh distributed with means

\[
\mathbb{E} \{ |\hat{h}_{B,k}| \} = \sqrt{\frac{\pi}{4} \eta \text{NPL}(d_{B,k}) \text{SNR}^2 + 1} \quad \mathbb{E} \{ |\hat{h}_{F,k}| \} = \sqrt{\frac{\pi}{4} \eta \text{NPL}(d_{F,k}) \text{SNR}^2 + 1}.
\]

The equivalent AWGN has total variance \( \sigma_Y^2 + \sigma_{Y^2} + \sigma_{Y^4} + 1 \). The variance of the propagated channel estimation errors at all the relay nodes is denoted by \( \sigma_{Y^2} \) and is given by (42), shown at the top of this page, where (a) follows from the fact that estimate errors are uncorrelated with the channel estimates by the orthogonality principle of MMSE estimation. Since the channel coefficients are complex Gaussian, the estimate errors are independent from the channel estimates. Moreover, the channel estimate errors are zero-mean; (b) is obtained by simply expanding each term within the square and using the fact that estimate errors and estimates are independent; and...
\[ \sigma^2_{V_2} = \mathbb{E} \left\{ \sum_{k \in \mathcal{K}} \sqrt{G_k(\delta)} \left( \frac{\hat{h}_{B,k} \hat{h}_{F,k}}{|\hat{h}_{B,k}|^2} + \frac{\hat{h}_{B,k} \hat{h}_{F,k}}{|\hat{h}_{F,k}|} \right)^2 Z_{R,k} [i] \right\} \]

\[ = \sum_{k \in \mathcal{K}} G_k(\delta) \left( \mathbb{E} \left\{ |\hat{h}_{F,k}|^2 \right\} + \mathbb{E} \left\{ |\hat{h}_{B,k}|^2 \right\} \right) \]

\[ \stackrel{(a)}{=} \sum_{k \in \mathcal{K}} G_k(\delta) \left( \frac{\eta_{NPL}(d_{F,k})SNR^3}{\eta_{NPL}(d_{F,k})SNR^3 + 1} + \frac{1}{\eta_{NPL}(d_{F,k})SNR^3 + 1} \right) = \sum_{k \in \mathcal{K}} G_k(\delta) \]  \hspace{1cm} (44)

\[ \sigma^2_{V_3} = \mathbb{E} \left\{ \sum_{k \in \mathcal{K}} \sqrt{G_k(\delta)} \left( |\hat{h}_{B,k}|^2 - \mathbb{E} \left\{ |\hat{h}_{B,k}|^2 \right\} \right) X[i]^2 \right\} \]

\[ \stackrel{(a)}{=} \sum_{k \in \mathcal{K}} G_k(\delta) \left[ \mathbb{E} \left\{ |\hat{h}_{B,k}|^2 \right\} \mathbb{E} \left\{ |\hat{h}_{F,k}|^2 \right\} - \left( \mathbb{E} \left\{ |\hat{h}_{B,k}|^2 \right\} \mathbb{E} \left\{ |\hat{h}_{F,k}|^2 \right\} \right)^2 \right] \]

\[ \stackrel{(b)}{=} \left( 1 - \frac{\pi^2}{16} \right) \eta_{NSNR} \sum_{k \in \mathcal{K}} \frac{G_k(\delta)PL(d_{B,k})PL(d_{F,k})}{(\eta_{NPL}(d_{B,k})SNR^3 + 1)(\eta_{NPL}(d_{F,k})SNR^3 + 1)} \]  \hspace{1cm} (45)

\[ \sigma^2_X \leq \frac{\pi^2}{16} \frac{N(1-\eta)}{(N-1)} \frac{\left( \eta_{NPL}(d_{min})^2SNR^3 \right)^2}{(\eta_{NPL}(Kd)SNR^3 + 1)(\eta_{NPL}(1-d)SNR^3 + 1)} K^2SNR^{26} \]  \hspace{1cm} (46)

\[ \sigma^2_X \geq \frac{\pi^2}{16} \frac{N(1-\eta)}{(N-1)} \frac{\left( \eta_{NPL}(Kd)PL(d_{min})^2SNR^3 \right)^2}{(\eta_{NPL}(d_{min})SNR^3 + 1)} K^2SNR^{26} \]  \hspace{1cm} (47)

\[ \sigma^2_{V_1} + \sigma^2_{V_3} \leq \frac{N(1-\eta)}{(N-1)} KSNR^{26}(PL(d_{min}))^2 \left[ 1 - \frac{\pi^2}{16} \frac{PL(Kd)PL(d_{min}) - 1}{(\eta_{NPL}(Kd)SNR^3 + 1)(\eta_{NPL}(1-d)SNR^3 + 1)} \right] \]  \hspace{1cm} (48)

\[ \sigma^2_{V_1} + \sigma^2_{V_3} \geq \frac{N(1-\eta)}{(N-1)} KSNR^{26}PL(Kd)PL(1-d) \left[ 1 - \frac{\pi^2}{16} \frac{(\eta_{NPL}(d_{min})^2SNR^3)^2}{(\eta_{NPL}(Kd)SNR^3 + 1)(\eta_{NPL}(1-d)SNR^3 + 1)} \right] \]  \hspace{1cm} (49)

\[ \sigma^2_{V_2} \leq KSNR^8PL(d_{min}) \]  \hspace{1cm} (50)

\[ \sigma^2_{V_2} \geq KSNR^8PL(1-d) \]  \hspace{1cm} (51)

\[ \sigma^2_X \leq \frac{\pi^2}{16} \frac{N(1-\eta)}{(N-1)} \frac{\left( \eta_{NPL}(d_{min})^2SNR^3 \right)^2}{(\eta_{NPL}(Kd) + 1)(\eta_{NPL}(1-d) + 1) + 1} K^2 \]  \hspace{1cm} (52)

\[ \sigma^2_X \geq \frac{\pi^2}{16} \frac{N(1-\eta)}{(N-1)} \frac{\left( \eta_{NPL}(Kd)PL(d_{min})^2SNR^3 \right)^2}{(\eta_{NPL}(d_{min}) + 1)} K^2 \]  \hspace{1cm} (53)

\[ \sigma^2_{V_1} + \sigma^2_{V_3} \leq \frac{N(1-\eta)}{(N-1)} K(PL(d_{min}))^2 \left[ 1 - \frac{\pi^2}{16} \frac{PL(Kd)PL(d_{min}) - 1}{(\eta_{NPL}(Kd)SNR^3 + 1)^2} \right] \]  \hspace{1cm} (54)

\[ \sigma^2_{V_1} + \sigma^2_{V_3} \geq \frac{N(1-\eta)}{(N-1)} KPL(Kd)PL(1-d) \left[ 1 - \frac{\pi^2}{16} \frac{(\eta_{NPL}(d_{min})^2SNR^3)^2}{(\eta_{NPL}(Kd) + 1)(\eta_{NPL}(1-d) + 1)} \right] \]  \hspace{1cm} (55)

\[ \sigma^2_{V_2} \leq \frac{K}{1-\eta} \frac{PL(d_{min})}{PL(Kd)} \]  \hspace{1cm} (56)

\[ \sigma^2_{V_2} \geq \frac{K}{1-\eta} \frac{PL(1-d)}{PL(d_{min})} \]  \hspace{1cm} (57)

(c) follows from (4), and the fact that

\[ \mathbb{E} \left\{ |\hat{h}_{B,k}|^2 \right\} = \frac{\eta_{NPL}(d_{B,k})SNR^3}{\eta_{NPL}(d_{B,k})SNR^3 + 1} \]  \hspace{1cm} (43)
The variance of the amplified noise from all the relay nodes is given by (44), shown at the top of the previous page, where (a) can be obtained by substituting using (4) and (43). Lastly, the variance of the estimation errors at the destination node is given by (45), shown at the top of the previous page, where (a) follows from the fact that \( \langle [h_{B,i}]_{i} [h_{F,j}] - \mathbb{E} \{ [h_{B,i}]_{i} [h_{F,j}] \} \rangle \) and \( \langle [h_{B,j}]_{j} [h_{F,j}] - \mathbb{E} \{ [h_{B,j}]_{j} [h_{F,j}] \} \rangle \) are independent, and \( h_{B,i} \) and \( h_{F,i} \) are independent; and (b) is obtained by substituting (43).

**Appendix C**

**Proof of Proposition 3**

Using (37), we first derive the bounds on the effective SNR in (27) and (28) when \( 0 < \delta \leq 1 \). It follows from (13), (14) and Proposition 2 that we can bound \( \sigma_{V}^{2} \), \( \sigma_{V_{1}}^{2} \), \( \sigma_{V_{2}}^{2} \), and \( \sigma_{V_{3}}^{2} \) as given by (46)-(51), shown at the top of the previous page. Similarly, for \( \delta = 0 \), we have (52)-(57), shown at the top of the previous page.

**References**


