Approximate Ergodic Capacity of a Class of Fading 2-user 2-hop Networks

Sang-Woon Jeon, *Member, IEEE*, Chien-Yi Wang, *Student Member, IEEE*, and
Michael Gastpar, *Member, IEEE*

Abstract

We consider a fading AWGN 2-user 2-hop network where the channel coefficients are independent and identically distributed (i.i.d.) drawn from a continuous distribution and vary over time. For a broad class of channel distributions, we characterize the ergodic sum capacity to within a constant number of bits/sec/Hz, independent of signal-to-noise ratio. The achievability follows from the analysis of an interference neutralization scheme where the relays are partitioned into M pairs, and interference is neutralized separately by each pair of relays. When $M = 1$, the proposed ergodic interference neutralization characterizes the ergodic sum capacity to within 4 bits/sec/Hz for i.i.d. uniform phase fading and approximately 4.7 bits/sec/Hz for i.i.d. Rayleigh fading. We further show that this gap can be tightened to $4 \log \pi - 4$ bits/sec/Hz (approximately 2.6) for i.i.d. uniform phase fading and $4 - 4 \log(\frac{3\pi}{8})$ bits/sec/Hz (approximately 3.1) for i.i.d. Rayleigh fading in the limit of large M.[1]

Index Terms

Amplify-and-forward, approximate capacity, ergodic capacity, fading, interference neutralization, two unicast, two-user two-hop networks.

I. INTRODUCTION

In recent years, there has been significant progress towards understanding fundamentals of multi-source single-hop networks [1]–[3].[2] Following up on these successes for single-hop networks, more recent and emerging work has considered *multi-source multi-hop networks* [4]–[8]. For multi-source multi-hop networks, interference can be cancelled by aligning multiple paths through the network, a technique referred to as *interference neutralization*. Proper exploitation of such interference neutralization is the key for an approximate capacity [4] and the optimal degrees of freedom (DoF) characterization [5]–[8]. Recently, for 2-user 2-hop networks, interference neutralization combining with symbol extension was used to show that two relays suffice to achieve the optimal DoF [5]. In spite of recent progress in this area, the best known capacity characterization for fully connected 2-user 2-hop networks is to within $o(\log(\text{SNR}))$ bits/sec/Hz [5], which can be arbitrarily large as the signal-to-noise ratio (SNR) increases.

The aim of this paper is to *tighten the capacity gap of 2-user 2-hop networks to within a constant number of bits/sec/Hz, independent of SNR*. Our achievability is based on *ergodic interference neutralization* [6], which is similar to ergodic interference alignment [3] applied to multi-source single-hop networks. Suppose that the sources transmit their signals at time t through the first-hop channel matrix $\mathbf{H}[t]$. Then the relays amplify and forward their received signals with an appropriate delay τ through the second-hop channel matrix $\mathbf{G}[t + \tau]$ such that $\mathbf{G}[t + \tau]\mathbf{H}[t]$ becomes an approximately diagonal matrix with non-zero diagonal elements. This approach can completely neutralize interference in the finite SNR regime.

Assuming independent and identically distributed (i.i.d.) channel coefficients, the proposed ergodic interference neutralization characterizes the ergodic sum capacity to within a constant number of bits/sec/Hz for a broad class of channel distributions. For instance, when the number of relays L is equal to two, it achieves the ergodic sum

[1]Throughout the paper, $\log(\cdot)$ denotes the logarithm of base two.

[2]Unless otherwise specified, we assume Gaussian networks throughout the paper.

TABLE I
NEW APPROXIMATE CAPACITY RESULTS AND THE EXISTING DoF AND APPROXIMATE CAPACITY RESULTS.

	K-user interference channel	2-user 2-hop network with 2 relays	K-user 2-hop network with L relays	K-user K-hop network with K relays at each layer
DoF	$\frac{K}{2}$ [2]	2 [5]	Generally unknown K if $L \geq K(K-1)+1$ [9] L if $K \to \infty$ [10]	Generally unknown K for isotropic fading [6]
Ergodic capacity for uniform phase fading	Exact capacity [11]	4 bits/sec/Hz gap	Generally unknown 2.6 bits/sec/Hz gap if $K=2$ and $L \to \infty$	Unknown
Ergodic capacity for Rayleigh fading	$1.3K$ bits/sec/Hz gap	4.7 bits/sec/Hz gap	Generally unknown 3.1 bits/sec/Hz gap if $K=2$ and $L \to \infty$	Unknown

capacity to within 4 bits/sec/Hz for uniform phase fading and approximately 4.7 bits/sec/Hz for Rayleigh fading. As L increases, we narrow the corresponding gap in our analysis. Specifically, this gap is given as $4 \log \pi - 4$ bits/sec/Hz (approximately 2.6) for i.i.d. uniform phase fading and $4 - 4\log(\frac{3\pi}{8})$ bits/sec/Hz (approximately 3.1) for i.i.d. Rayleigh fading in the limit of large L. We also notice that a similar analysis is applicable for the K-user interference channel and show that ergodic interference alignment in [3] characterizes the ergodic sum capacity assuming that all sources employ uniform power allocation across time to within $(\frac{1}{2}\log 6)K$ bits/sec/Hz (approximately $1.3K$) for i.i.d. Rayleigh fading. Table I summarizes the new approximate ergodic capacity results of this paper and the existing DoF and approximate capacity results.

A. Related Work

1) Degrees of freedom: In seminal work [2], interference alignment has been proposed to achieve the optimal DoF of the K-user interference channel with time-varying channel coefficients. The concept of this signal space alignment has been successfully adapted to various network environments, e.g., see [10], [12]–[18] and the references therein. It was shown in [19], [20] that interference alignment can also be attained on fixed (not time-varying) interference channels.

In spite of recent achievements on interference channels or multi-source single-hop networks, understanding of multi-source multi-hop networks is still in progress. The work [5] has exploited interference alignment to neutralize interference at final destinations, which is referred to as aligned interference neutralization, and showed that the optimal 2 DoF is achievable for 2-user 2-hop networks with 2 relays. This result has been recently generalized to two unicast networks [7], [8]. For more than two unicast, the optimal DoF is in general unknown except for a certain class of networks. For the K-user 2-hop network with L relays, interference can be completely neutralized if $L \geq K(K-1)+1$ [9]. Similar concept of ergodic interference alignment has been proposed for interference neutralization in [6] showing that ergodic interference neutralization achieves the optimal DoF of K-user K-hop isotropic fading networks with K relays in each layer.

2) Beyond degrees of freedom: The DoF discussed previously is a fundamental metric of multi-source networks especially for high SNR, which characterizes capacity to within $o(\log \text{SNR})$ bits/sec/Hz. Depending on the operational regime, however, the gap of $o(\log \text{SNR})$ bits/sec/Hz in practice can be significant and achieving the optimal DoF may not be enough. For the 2-user interference channel, for instance, time-sharing between the two users can also achieve the optimal one DoF. On the other hand, a simple Han–Kobayashi scheme can tighten the gap to within one bit/sec/Hz [1], which provides an arbitrarily larger rate compared with the time-sharing for a certain operational regime and channel parameters. Consequently, several works have recently established tighter bounds on the gap from capacity [4], [21]–[26] to provide a universal performance guarantee, independent of SNR and channel parameters.

A similar flavor of such bounds on the gap from capacity concerns time-varying channel models. The recently proposed ergodic interference alignment in [3] makes interference aligned in the finite SNR regime and, as a result, provides significant rate improvement compared with the conventional time-sharing strategy in the finite SNR regime. Ergodic interference alignment was shown to achieve the ergodic sum capacity of the K-user interference channel for i.i.d. uniform phase fading [3]. For the K-user finite field interference channel (with time-varying channel coefficients), the idea of ergodic interference alignment was independently proposed by Nazer *et al.* [11]

TABLE II
SUMMARY OF NOTATION

$\mathbf{A}^T$ (or $\mathbf{a}^T$)	Transpose of $\mathbf{A}$(or $\mathbf{a}$)
$\mathbf{A}^\dagger$ (or $\mathbf{a}^\dagger$)	Conjugate transpose of $\mathbf{A}$(or $\mathbf{a}$)
$\det(\mathbf{A})$	Determinant of $\mathbf{A}$
$\mathbf{I}$	Identity matrix
$\jmath$	$\sqrt{-1}$
$\mathrm{re}(a)$(or $\mathrm{im}(a)$)	Real (or imaginary) part of a
$|a|$	Absolute value of a
a^*	Complex conjugate of a
$\lfloor a \rfloor$	Floor of a ($\lfloor a \rfloor = \max\{x \in \mathbb{Z} | x \le a\}$)
$\mathrm{card}(\mathcal{A})$	Cardinality of $\mathcal{A}$
$\mathcal{CN}(\mu, \sigma^2)$	Circularly symmetric complex Gaussian distribution with mean μ and variance σ^2

and Jeon and Chung [27] in two slightly different versions. In [28], ergodic channel pairing was applied to tighten the gap from the ergodic capacity for fading multihop networks showing a gap depending only on the number of nodes in a layer, instead of the total number of nodes in a network.

B. Paper Organization

The rest of the paper is organized as follows. In Section II, we introduce the fading 2-user 2-hop network model considered in this paper and formally define its ergodic sum capacity. In Section III, we first state the main results of the paper, approximate ergodic sum capacities of fading 2-user 2-hop networks. In Section IV, we explain ergodic interference neutralization and its achievable rate. In Section V, we prove the approximate ergodic sum capacity results in Section III based on the achievability in Section IV. Finally, we conclude in Section VI and refer some technical proofs to the appendices.

II. PROBLEM FORMULATION

In this section, we explain our network model and define its sum capacity. Throughout the paper, we will use $\mathbf{A}$, $\mathbf{a}$, and $\mathcal{A}$ to denote a matrix, vector, and set, respectively. The notation used in the paper is summarized in Table II.

A. Fading 2-User 2-Hop Networks

We study the 2-user 2-hop network depicted in Fig. 1 in which each source wishes to transmit an independent message to its destination with the help of L relays, where $L \ge 2$. The input–output relation of the first hop at time t is given by

$$\mathbf{y}_R[t] = \mathbf{H}[t]\mathbf{x}[t] + \mathbf{z}_R[t], \tag{1}$$

where

$$\mathbf{H}[t] = \begin{bmatrix} h_{1,1}[t] & h_{1,2}[t] \\ h_{2,1}[t] & h_{2,2}[t] \\ \vdots & \vdots \\ h_{L,1}[t] & h_{L,2}[t] \end{bmatrix} \tag{2}$$

is the $L \times 2$ dimensional complex channel matrix of the first hop at time t, $\mathbf{y}_R[t] = [y_{R,1}[t], \cdots, y_{R,L}[t]]^T$ is the $L \times 1$ dimensional received signal vector of the relays at time t, $\mathbf{x}[t] = [x_1[t], x_2[t]]^T$ is the 2×1 dimensional transmit signal vector of the sources at time t, and $\mathbf{z}_R[t] = [z_{R,1}[t], \cdots, z_{R,L}[t]]^T$ is the $L \times 1$ dimensional noise vector of the relays at time t. Similarly, the input–output relation of the second hop at time t is given by

$$\mathbf{y}[t] = \mathbf{G}[t]\mathbf{x}_R[t] + \mathbf{z}[t], \tag{3}$$

where

$$\mathbf{G}[t] = \begin{bmatrix} g_{1,1}[t] & g_{1,2}[t] & \cdots & g_{1,L}[t] \\ g_{2,1}[t] & g_{2,2}[t] & \cdots & g_{2,L}[t] \end{bmatrix} \tag{4}$$
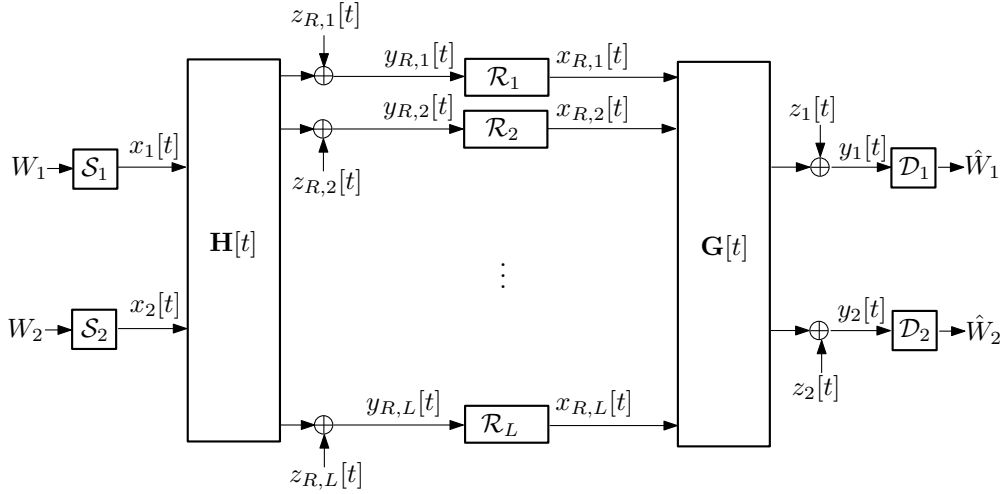
Fig. 1. Gaussian 2-user 2-hop network with L relays.

is the the $2 \times L$ dimensional complex channel matrix of the second hop at time t, $\mathbf{y}[t] = [y_1[t], y_2[t]]^T$ is the 2×1 dimensional received signal vector of the destinations at time t, $\mathbf{x}_R[t] = [x_{R,1}[t], \cdots, x_{R,L}[t]]^T$ is the $L \times 1$ dimensional transmit signal vector of the relays at time t, and $\mathbf{z}[t] = [z_1[t], z_2[t]]^T$ is the 2×1 dimensional noise vector of the destinations at time t. We assume that the elements of $\mathbf{z}_R[t]$ and $\mathbf{z}[t]$ are i.i.d. drawn from $\mathcal{CN}(0, 1)$. Each source and relay should satisfy the average power constraint P, i.e., $\mathsf{E}[|x_i[t]|^2] \leq P$ for $i \in \{1, 2\}$ and $\mathsf{E}[|x_{R,j}[t]|^2] \leq P$ for $j \in \{1, \cdots, L\}$.

We assume that *channel coefficients are i.i.d. drawn from a continuous distribution $f(x)$, $x \in \mathbb{C}$, and vary independently over time*. Without loss of generality, we assume that $\mathsf{E}[|h_{i,j}[t]|^2] = 1$ and $\mathsf{E}[|g_{j,i}[t]|^2] = 1$ for all $i \in \{1, \cdots, L\}$ and $j \in \{1, 2\}$. We further assume that the sources do not know any channel state information (CSI) and the relays and the destinations know global CSI. That is, at time t, each relay and destination knows $\mathbf{H}[t]$ and $\mathbf{G}[t]$.

B. Ergodic Sum Capacity

Based on the network model, we consider a set of length-n block codes. Let W_i be the message of source i uniformly distributed over $\{1, \cdots, 2^{nR_i}\}$, where R_i is the rate of source i. A rate pair (R_1, R_2) is said to be *achievable* if there exists a sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes such that the probabilities of error for W_1 and W_2 converge to zero as n increases. Notice that since channel coefficients are i.i.d. varying over time, an achievable rate pair (R_1, R_2) is given as in the ergodic sense, i.e., the expectation over random channel coefficients. The ergodic sum capacity C_{sum} is defined as the maximum achievable ergodic sum rate. Unless otherwise specified, an achievable sum rate or the sum capacity in this paper mean an achievable ergodic sum rate or the ergodic sum capacity, respectively.

III. MAIN RESULTS

In this section, we first introduce our main results. Let $M := \lfloor \frac{L}{2} \rfloor$. As will be explained in Section IV, we only use $2M$ relays among the total number L of relays for the achievability. That is, the achievability is based on an even number of relays. Without loss of generality, we assume that relay 1 to relay $2M$ are used for relaying. The achievability follows from ergodic interference neutralization based on amplify-and-forward relaying in which $2M$ relays are partitioned into M pairs and interference is neutralized separately by each pair of relays. In order to describe the proposed ergodic interference neutralization and its achievable sum rate, for $m \in \{1, \cdots, M\}$, we denote

$$\mathbf{H}_m[t] := \begin{bmatrix} h_{2m-1,1}[t] & h_{2m-1,2}[t] \\ h_{2m,1}[t] & h_{2m,2}[t] \end{bmatrix} \tag{5}$$

and

$$\mathbf{G}_m[t] := \begin{bmatrix} g_{1,2m-1}[t] & g_{2,2m-1}[t] \\ g_{1,2m}[t] & g_{2,2m}[t] \end{bmatrix}, \tag{6}$$

which are the 2×2 dimensional channel matrices at time t from the sources to relays $2m-1$ and $2m$ and from relays $2m-1$ and $2m$ to the destinations, respectively.

A. Achievable Sum Rate

The following theorem states an achievable symmetric rate of the fading 2-user 2-hop network.

Theorem 1: For the fading 2-user 2-hop network with L relays,

$$R_i = \mathsf{E}\left[\log\left(1 + \frac{P\gamma^2\left(\sum_{m=1}^{M}|\det(\mathbf{H}_m)|\right)^2}{1+\sigma_{\mathrm{AF},i}^2}\right)\right] \tag{7}$$

is achievable for $i \in \{1,2\}$, where $M = \lfloor\frac{L}{2}\rfloor, \gamma = \sqrt{\frac{P}{1+2P}}, \sigma_{\mathrm{AF},i}^2 = \gamma^2\sum_{m=1}^{M}(|h_{2(m-1)+3-i,3-i}|^2+|h_{2(m-1)+i,3-i}|^2)$, and the expectation is over the channel coefficients.

Proof: The proof is in Section IV. ∎

The most important aspect is that there is no residual interference after ergodic interference neutralization, meaning that interference can completely be neutralized at finite SNR. Moreover, from the block-wise coherent combining gain shown as $(\sum_{m=1}^{M}|\det(\mathbf{H}_m)|)^2$ in (7), the received signal power increases as the number of pairs M increases. Although there is noise amplification due to amplify-and-forward relaying given as $\sigma_{\mathrm{AF},i}^2$ in (7), this additional noise results in a constant number of bits/sec/Hz loss for a broad class of channel distributions, which will be proved in Section V.

For notational convenience, let

$$R_{\mathrm{in}} := \sum_{i=1}^{2}\mathsf{E}\left[\log\left(1 + \frac{P\gamma^2\left(\sum_{m=1}^{M}|\det(\mathbf{H}_m)|\right)^2}{1+\sigma_{\mathrm{AF},i}^2}\right)\right], \tag{8}$$

which is the achievable sum rate from Theorem 1. For comparison, we consider the ergodic capacity of the multiple-input multiple-output (MIMO) channel from the sources to the relays, that is

$$R_{\mathrm{mimo}} := \mathsf{E}\left[\log\det(\mathbf{I} + P\mathbf{H}\mathbf{H}^\dagger)\right]. \tag{9}$$

Since the channel coefficients are i.i.d. and the sources do not know CSI, C_{sum} is upper bounded by R_{mimo} [29]. The following example illustrates R_{in} and R_{mimo} for i.i.d. Rayleigh fading, i.e., $f(x)$ follows $\mathcal{CN}(0,1)$.

Example 1 (Sum rate: Rayleigh fading): Figure 2 plots R_{in} and R_{mimo} for i.i.d. Rayleigh fading. Two important aspects can be observed in the figure. First, for a fixed number of relays L, the sum rate gap $R_{\mathrm{mimo}}-R_{\mathrm{in}}$ appears to be upper bounded by some constant independent of power P, which suggests that the proposed ergodic interference neutralization can achieve the ergodic sum capacity to within a constant number of bits/sec/Hz independent of P. Second, for a fixed P, the sum rate gap $R_{\mathrm{mimo}}-R_{\mathrm{in}}$ appears to decrease with increasing L, which suggests that this approximate capacity characterization can be tightened as the number of relays L increases.

Both observations in Example 1 are established in this paper and shown to hold beyond the case of Rayleigh fading for any fading model for which $f(x)$ is only a function of $|x|$. The following two subsections describe our approximate capacity results characterizing the ergodic sum capacity to within a constant number of bits/sec/Hz, independent of P.

B. Approximate Ergodic Sum Capacity for $L = 2$

In this subsection, we assume $L = 2$. We first consider i.i.d. uniform phase fading in which $h_{i,j}[t] = \exp(\jmath\theta_{i,j}[t])$ and $g_{j,i}[t] = \exp(\jmath\varphi_{j,i}[t])$, where $\theta_{i,j}[t]$ and $\varphi_{j,i}[t]$ are uniformly distributed over $[0,2\pi)$ for all $i,j \in \{1,2\}$. Although uniform phase fading violates the channel assumption in Section II-A, i.e., $f(x)$ is continuous over $x \in \mathbb{C}$, we can slightly modify the proposed ergodic interference neutralization and show that Theorem 1 still
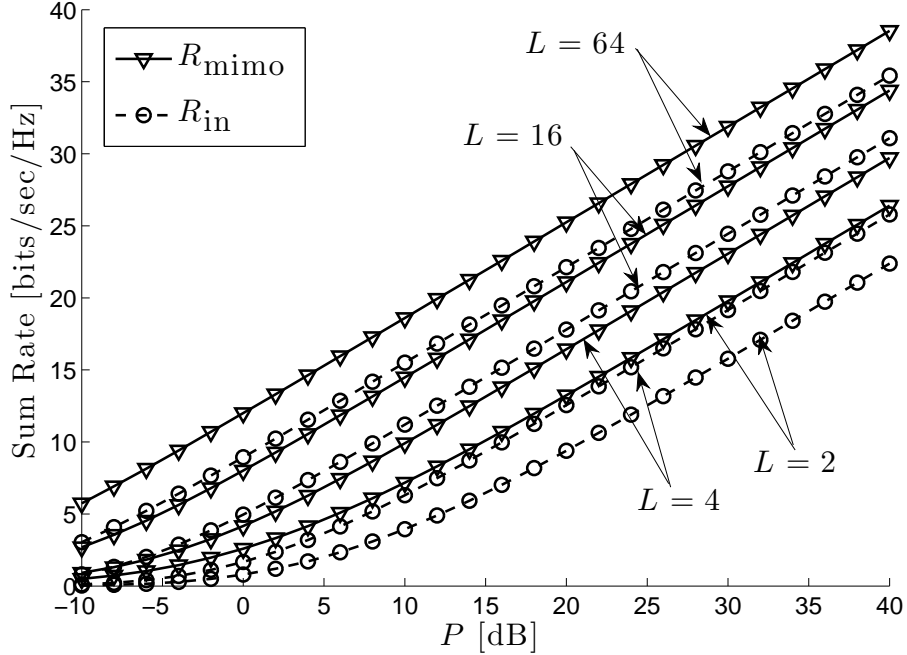
Fig. 2.   The achievable sum rate R_{in} and its upper bound R_{mimo} for i.i.d. Rayleigh fading when $L = 2, 4, 16, 64$.

holds. The detailed modification is given in Appendix I. The following theorem characterizes an approximate ergodic sum capacity for i.i.d. uniform phase fading.

Theorem 2: Consider the fading 2-user 2-hop network with $L = 2$ relays. If $h_{i,j}[t] = \exp(\jmath\theta_{i,j}[t])$ and $g_{j,i}[t] = \exp(\jmath\varphi_{j,i}[t])$, where $\theta_{i,j}[t]$ and $\varphi_{j,i}[t]$ are uniformly distributed over $[0, 2\pi)$ for all $i, j \in \{1, 2\}$, then

$$C_{\mathrm{sum}} - R_{\mathrm{in}} \leq 4 \tag{10}$$

for any $P > 0$.

Proof: The proof is in Section V-A. ■

Example 2 (Gap for $L = 2$: Uniform phase fading): Figure 3 plots $R_{\mathrm{mimo}} - R_{\mathrm{in}}$ with respect to P for i.i.d. uniform phase fading (the closed forms of R_{mimo} and R_{in} are given by (37) and (38), respectively). As proved by Theorem 2, the proposed ergodic interference neutralization achieves C_{sum} to within 4 bits/sec/Hz for i.i.d. uniform phase fading. This theoretical gap coincides with the actual gap $R_{\mathrm{mimo}} - R_{\mathrm{in}}$ at high SNR, i.e., $\lim_{P \to \infty} \{R_{\mathrm{mimo}} - R_{\mathrm{in}}\} = 4$.

Based on the bounding techniques used in proving Theorem 2, we characterize an approximate ergodic sum capacity for a class of channel distributions satisfying that $f(x)$ is only a function of $|x|$. Specifically, for a given set of channel amplitudes, we first upper bound the gap $R_{\mathrm{mimo}} - R_{\mathrm{in}}$ by averaging out the effect of phase fading. Then we further apply additional bounding techniques to obtain an upper bound, independent of power P.

Theorem 3: Consider the fading 2-user 2-hop network with $L = 2$ relays. If $f(x)$ is only a function of $|x|$, then

$$C_{\mathrm{sum}} - R_{\mathrm{in}} \leq 2\,\mathsf{E}\left[\log\left(\frac{\sqrt{A}(A + B^2)}{B(A + \sqrt{A^2 - G^2})}\right)\right] + 2 \tag{11}$$

for any $P > 0$, where

$$
\begin{aligned}
A &= |h_{1,1}|^2|h_{2,2}|^2 + |h_{1,2}|^2|h_{2,1}|^2, \\
B &= |h_{1,1}|^2 + |h_{2,1}|^2 + 2, \\
G &= 2|h_{1,1}||h_{1,2}||h_{2,1}||h_{2,2}|,
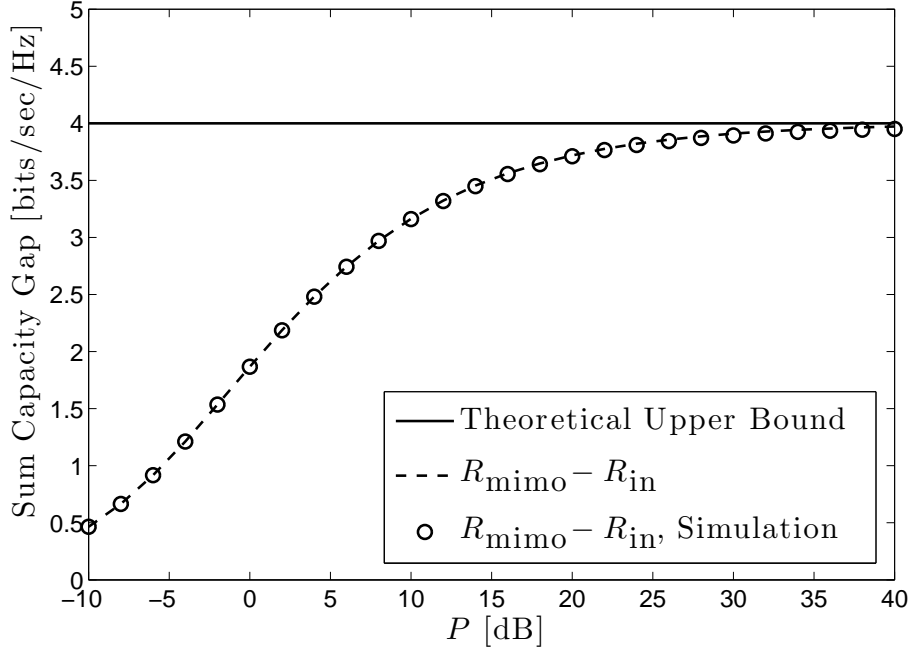\end{aligned}
\tag{12}
$$

Fig. 3.   Gap from the sum capacity for i.i.d. uniform phase fading when $L = 2$.
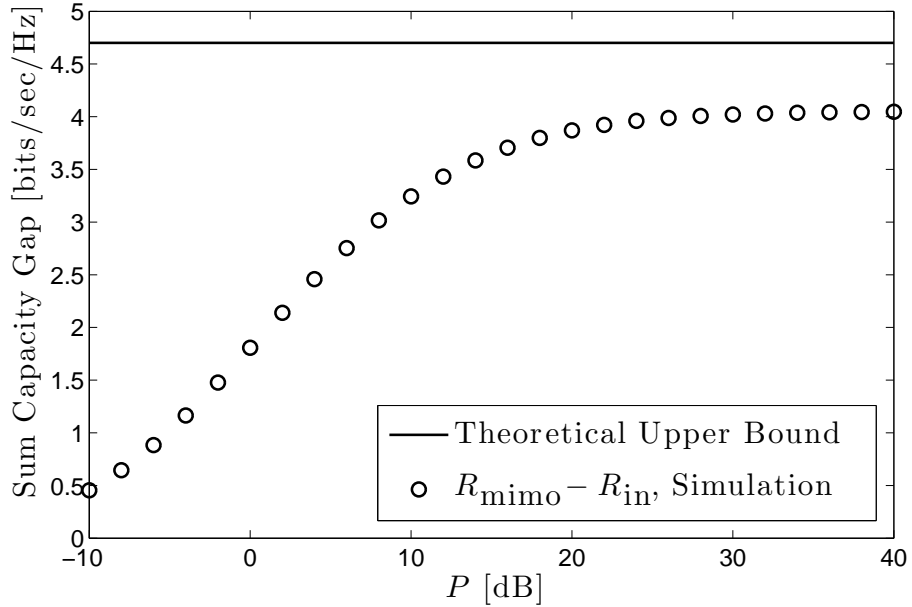

Fig. 4.   Gap from the sum capacity for i.i.d. Rayleigh fading when $L = 2$.

and the expectation is over the channel coefficients.

   Proof: The proof is in Section V-A.   ∎

   The presented gap in Theorem 3 only depends on the amplitude distribution of channel coefficients, which provides universal performance guarantee regardless of power P. The following example evaluates the presented gap for i.i.d. Rayleigh fading.

   Example 3 (Gap for $L = 2$: Rayleigh fading): Figure 4 plots $R_{\mathrm{mimo}} - R_{\mathrm{in}}$ with respect to P and also plots its upper bound in Theorem 3 for i.i.d. Rayleigh fading. Since there is no closed form, we evaluate the bound in Theorem 3 by simulation, which approximately provides 4.7 bits/sec/Hz gap. Simulation result shows that the
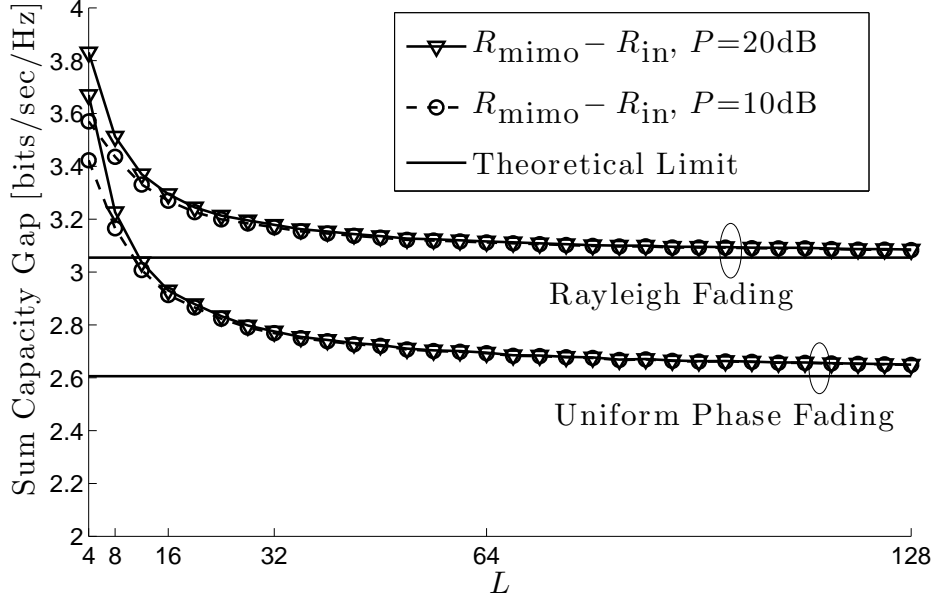
Fig. 5. Gap from the sum capacity with respect to the number of relays.

proposed scheme achieves at least 71%, 79%, 84%, 87%, and 89% percent of the ergodic sum capacity at SNR 20, 30, 40, 50, and 60 dB, respectively.

C. Approximate Ergodic Sum Capacity as $L \to \infty$

In this subsection, we focus on an approximate ergodic sum capacity as the number L of relays increases. Again, we first consider i.i.d. uniform phase fading and then consider a class of channel distributions satisfying that $f(x)$ is only a function of $|x|$.

Theorem 4: Consider the fading 2-user 2-hop network with L relays. If $h_{i,j}[t] = \exp(\jmath\theta_{i,j}[t])$, $g_{j,i}[t] = \exp(\jmath\varphi_{j,i}[t])$, and $\theta_{i,j}[t]$ and $\varphi_{j,i}[t]$ are uniformly distributed over $[0, 2\pi)$ for all $i \in \{1, \cdots, L\}$ and $j \in \{1, 2\}$, then

$$\lim_{L \to \infty} \{C_{\mathrm{sum}} - R_{\mathrm{in}}\} \le 4 \log \pi - 4 \tag{13}$$

for any $P > 0$.

Proof: The proof is in Section V-B. ∎

Example 4 (Gap as $L \to \infty$: Uniform phase fading): Figure 5 plots the gap $R_{\mathrm{mimo}} - R_{\mathrm{in}}$ for i.i.d. uniform phase fading with respect to L. As shown in the figure, this gap decreases as L increases and eventually converges to $4 \log \pi - 4$ (approximately 2.6) regardless of P, which was proved in Theorem 4. Therefore the proposed ergodic interference neutralization characterizes C_{sum} to within $4 \log \pi - 4$ bits/sec/Hz in the limit of large L. Compared to 4 bits/sec/Hz, the sum capacity gap for $L = 2$ in Theorem 2, the result shows that the sum capacity gap can be tightened as L increases.

Theorem 5: Consider the fading 2-user 2-hop network with L relays. If $f(x)$ is only a function of $|x|$, then

$$\lim_{L \to \infty} \{C_{\mathrm{sum}} - R_{\mathrm{in}}\} \le 4 - 4 \log \left(\mathsf{E}[|\det(\mathbf{H}_1)|] \right) \tag{14}$$

for any $P > 0$.

Proof: The proof is in Section V-B. ∎

Example 5 (Gap as $L \to \infty$: Rayleigh fading): Figure 5 plots $R_{\mathrm{mimo}} - R_{\mathrm{in}}$ for i.i.d. Rayleigh fading with respect to L. That is, $f(x)$ follows $\mathcal{CN}(0, 1)$. For this case, it can be shown that $\mathsf{E}[|\det(\mathbf{H}_1)|] = \frac{3\pi}{8}$ and, thus, the theoretical limit in Theorem 5 leads $4 - 4 \log(\frac{3\pi}{8})$ (approximately 3.1). The detailed proof of $\mathsf{E}[|\det(\mathbf{H}_1)|] = \frac{3\pi}{8}$ is in Appendix
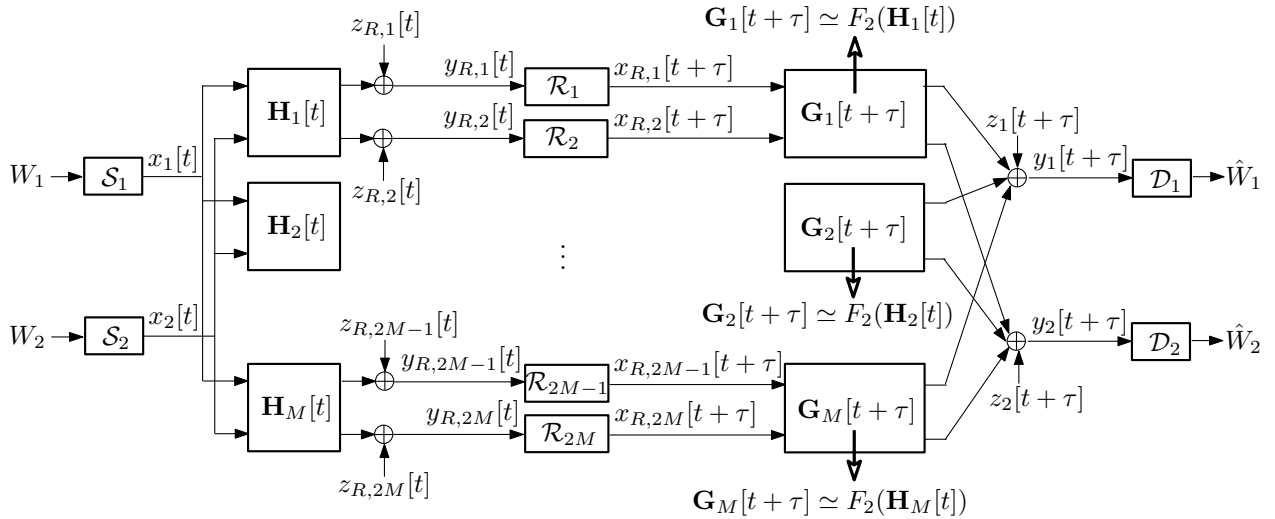
Fig. 6. Block-wise ergodic interference neutralization based on amplify-and-forward relaying.

II. As shown in the figure, $R_{\mathrm{mimo}} - R_{\mathrm{in}}$ quickly converges to the theoretical limit as L increases. Considering that the sum capacity gap is approximately given by 4.7 bits/sec/Hz when $L = 2$ (Theorem 3 and Example 3), the sum capacity gap can be tightened as L increases.

D. Approximate Ergodic Capacity for Fading Interference Channel

We notice that a similar analysis used in Theorems 2 and 3 is applicable to show an approximate ergodic capacity for fading K-user interference channel. The achievability follows from ergodic interference alignment in [3]. Assuming that all sources employ uniform power allocation across time, we show that ergodic interference alignment characterizes an approximate ergodic per-user capacity, i.e., ergodic sum capacity divided by K, for a broad class of channel distributions. The detailed statement is given in Theorem 6 in Appendix III. For i.i.d. Rayleigh fading, for instance, our analysis characterizes the ergodic per-user capacity to within $\frac{1}{2} \log 6$ bits/sec/Hz (approximately 1.3 bits/sec/Hz).

IV. ERGODIC INTERFERENCE NEUTRALIZATION

For the achievability, we propose ergodic interference neutralization using an even number of relays. Let $M := \lfloor \frac{L}{2} \rfloor$. Then we can choose $2M$ relays among the total number L of relays and apply the proposed ergodic interference neutralization by using these $2M$ relays. For simplicity, we assume L is even in the rest of this section. That is, $L = 2M$.

A. High-Level View

Before the detailed description and analysis, we begin by providing a high-level view of the proposed ergodic interference neutralization. Consider length-n sequences of matrices $\{\mathbf{H}[t]\}_{t=1}^{n}$ and $\{\mathbf{G}[t]\}_{t=1}^{n}$, drawn i.i.d. according to a certain probability density function. We partition these sequences judiciously into pairs of matrices $(\mathbf{H}[t_1], \mathbf{G}[t_2])$ such that $\mathbf{G}[t_2]$ and $F(\mathbf{H}[t_1])$ are *almost equal*, where $F(\cdot)$ is a cleverly chosen mapping to be discussed below. The main argument is that by considering a longer and longer sequence of matrices, we can make these two matrices arbitrarily close. The formal and technical details of this argument can be found in Sections IV-B and IV-C. For notational convenience, we introduce the notation $\mathbf{G}[t_2] \simeq F(\mathbf{H}[t_1])$ for the two matrices that are almost equal.

As pointed out in [6], a simple amplify-and-forward scheme with an appropriate delay $\tau \in \mathbb{Z}_{+}$ can neutralize interference by letting $\mathbf{G}[t + \tau]\mathbf{H}[t]$ approximately a diagonal matrix with non-zero diagonal elements. To satisfy this condition, we first partition L relays into $M = \frac{L}{2}$ pairs and neutralize interference separately by each pair of relays. Figure 6 illustrates the main idea of the proposed scheme. For $\mathbf{A} = \{a_{i,j}\} \in \mathbb{C}^{2 \times 2}$, define

$$F_2(\mathbf{A}) := \begin{bmatrix} a_{2,2} & a_{1,2} \\ a_{2,1} & a_{1,1} \end{bmatrix}. \tag{15}$$

The relays then amplify and forward with delay τ such that $\mathbf{G}_m[t+\tau] \simeq F_2(\mathbf{H}_m[t])$ for all $m \in \{1, \cdots, M\}$. For relaying, relays $2m-1$ and $2m$ amplify and forward with the amplification factors $\gamma\frac{\det(\mathbf{H}_m[t])^*}{|\det(\mathbf{H}_m[t])|}$ and $-\gamma\frac{\det(\mathbf{H}_m[t])^*}{|\det(\mathbf{H}_m[t])|}$, respectively. Here $\gamma = \sqrt{\frac{P}{1+2P}}$ is needed to satisfy the average power constraint P. Then the effective channel matrix of the mth pair is given by

$$\gamma\frac{\det(\mathbf{H}_m[t])^*}{|\det(\mathbf{H}_m[t])|}\mathbf{G}_m[t+\tau]\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\mathbf{H}_m[t] \simeq \gamma\frac{\det(\mathbf{H}_m[t])^*}{|\det(\mathbf{H}_m[t])|}F_2(\mathbf{H}_m[t])\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\mathbf{H}_m[t]$$

$$= \gamma|\det(\mathbf{H}_m[t])|\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{16}$$

As a consequence, the effective channel gain from each source to its destination is approximately given by $\gamma^2(\sum_{m=1}^M |\det(\mathbf{H}_m[t])|)^2$, as can be seen in the rate expression in Theorem 1. One can easily show that the additional noise power at destination i due to this amplify-and-forward relaying is given as $\sigma_{\mathrm{AF},i}^2$, as shown in the rate expression in Theorem 1. Lastly, since the probability density functions of the paired channel states are the same, i.e.,

$$f_{\mathbf{H}[t]}([\mathbf{H}_1^T, \cdots, \mathbf{H}_M^T]^T) = f_{\mathbf{G}[t]}([F_2(\mathbf{H}_1), \cdots, F_2(\mathbf{H}_M)]), \tag{17}$$

almost all channel instances can be utilized for this ergodic pairing as the block length n increases. Hence, the ergodic rate in Theorem 1 is achievable in the limit of large n.

There are two crucial facts to be observed: 1) the intended signal power received at each destination is non-zero while the interference power decreases arbitrarily close to zero at any finite power P; 2) the intended signal power received at each destination increases quadratically with increasing L. These facts make approximate capacity characterization possible for a broad class of channel distributions.

Although finding a pair of channel instances having exact prescribed values is impossible, such a pairing can be done approximately by partitioning the channel space of each hop and then pairing the partitioned channel spaces between the first and second hops. In the following subsection, we first explain channel space partition and pairing and then explain the detailed scheme.

B. Block-Wise Ergodic Interference Neutralization

1) Partitioning and pairing of channel space: We partition the channel space of each hop, i.e, $\mathbb{C}^{2M\times 2}$ space for the first hop and $\mathbb{C}^{2\times 2M}$ space for the second hop. First, consider the channel space of the first hop $\mathbb{C}^{2M\times 2}$. For $N \in \mathbb{Z}_+$ and $\Delta > 0$, define

$$\mathcal{Q}_1 := \Big\{ \mathbf{A} \in \Delta(\mathbb{Z}^{2M\times 2} + \jmath\mathbb{Z}^{2M\times 2}) \big| |\mathrm{re}(a_{i,j})| \le \Delta N, |\mathrm{im}(a_{i,j})| \le \Delta N$$
$$\text{for all } i \in \{1, \cdots, 2M\} \text{ and } j \in \{1, 2\}\Big\}, \tag{18}$$

where $\mathbf{A} = \{a_{i,j}\}$. Here, N and Δ are related to the number of quantization points and the quantization interval. For a quantized channel matrix $\mathbf{Q} \in \mathcal{Q}_1$, define

$$\mathcal{A}_1(\mathbf{Q}) := \Big\{ \mathbf{A} \in \mathbb{C}^{2M\times 2} \big| -\frac{\Delta}{2} \le \mathrm{re}(a_{i,j}) - \mathrm{re}(q_{i,j}) < \frac{\Delta}{2} \text{ and } -\frac{\Delta}{2} \le \mathrm{im}(a_{i,j}) - \mathrm{im}(q_{i,j}) < \frac{\Delta}{2}$$
$$\text{for all } i \in \{1, \cdots, 2M\} \text{ and } j \in \{1, 2\}\Big\}, \tag{19}$$

where $\mathbf{A} = \{a_{i,j}\}$ and $\mathbf{Q} = \{q_{i,j}\}$. Figure 7 illustrates the channel space partitioning with respect to $h_{i,j} \in \mathbb{C}$. We can define $\mathcal{Q}_2$ and $\mathcal{A}_2(\mathbf{Q})$ for the second hop as the same manner in (18) and (19) by substituting $\mathbf{A} \in \Delta(\mathbb{Z}^{2\times 2M} + \jmath\mathbb{Z}^{2\times 2M})$ and $\mathbf{A} \in \mathbb{C}^{2\times 2M}$, respectively. We will only use the first-hop channel instances in $\cup_{\mathbf{Q}\in\mathcal{Q}_1}\mathcal{A}_1(\mathbf{Q})$ and the second-hop channel instances in $\cup_{\mathbf{Q}\in\mathcal{Q}_2}\mathcal{A}_2(\mathbf{Q})$ for transmission.

Now consider the channel space pairing between $\mathcal{A}_1(\mathbf{Q})$ and $\mathcal{A}_2(\mathbf{Q})$. For $\mathbf{A} \in \mathbb{C}^{2M\times 2}$, define

$$F(\mathbf{A}) := [F_2(\mathbf{A}_1), F_2(\mathbf{A}_2), \cdots, F_2(\mathbf{A}_M)], \tag{20}$$

where $\mathbf{A} = [\mathbf{A}_1^T, \mathbf{A}_2^T, \cdots, \mathbf{A}_M^T]^T$ and the definition of $F_2(\cdot)$ is given by (15). For $\mathbf{H}[t] \in \mathcal{A}_1(\mathbf{Q})$, the relays will amplify and forward with delay $\tau \in \mathbb{Z}_+$ satisfying $\mathbf{G}[t+\tau] \in \mathcal{A}_2(F(\mathbf{Q}))$. Hence the channel subspace $\mathcal{A}_1(\mathbf{Q})$ of the first hop is paired with the channel subspace $\mathcal{A}_2(F(\mathbf{Q}))$ of the second hop. The detailed transmission scheme is given in the following subsection.
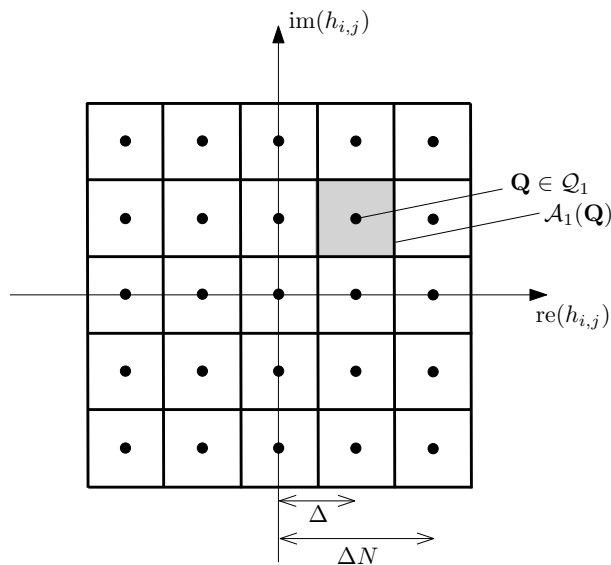
Fig. 7. Channel space partitioning with respect to the channel coefficient $h_{i,j} \in \mathbb{C}$.

2) Transmission scheme: We first divide a length-n block into B sub-blocks having length $n_B = \frac{n}{B}$ each. At the first sub-block, the sources transmit their first messages to the relays (the relays do not transmit). At the bth sub-block, $b \in \{2, \cdots, B-1\}$, the sources transmit their bth messages to the relays and the relays amplify and forward the received signals of the $(b-1)$th sub-block to the destinations. At the last sub-block, the relays amplify and forward the received signals of the $(B-1)$th sub-block to the destinations (the sources do not transmit). Hence, the number of effective sub-blocks is equal to $B-1$. Since we can set both n_B and B as large as desired as n increases, the fractional rate loss $\frac{1}{B}$ becomes negligible as n increases. For simplicity, we describe the proposed scheme based on the first message transmission and omit the sub-block index.

For $\mathbf{Q} \in \mathcal{Q}_1$, define $\mathcal{T}_1(\mathbf{Q}) := \left\{ t \in \{1, \cdots, n_B\} \middle| \mathbf{H}[t] \in \mathcal{A}_1(\mathbf{Q}) \right\}$, which is the set of time indices of the first hop whose channel instances belong to $\mathcal{A}_1(\mathbf{Q})$. Similarly, for $\mathbf{Q} \in \mathcal{Q}_2$, $\mathcal{T}_2(\mathbf{Q}) := \left\{ t \in \{n_B + 1, \cdots, 2n_B\} \middle| \mathbf{G}[t] \in \mathcal{A}_2(\mathbf{Q}) \right\}$, which is the set of time indices of the second hop whose channel instances belong to $\mathcal{A}_2(\mathbf{Q})$. The encoding, relaying, and decoding are as follows.

- (Encoding) The sources transmit their messages using Gaussian codebook with length n_B and average power P.
- (Relaying) For all $\mathbf{Q} \in \mathcal{Q}_1$, the relays amplify and forward their received signals that were received during $\mathcal{T}_1(\mathbf{Q})$ using the time indices in $\mathcal{T}_2(F(\mathbf{Q}))$. Specifically, for $t_1 \in \mathcal{T}_1(\mathbf{Q})$, the transmit signal vector of the relays is given by $\mathbf{x}_R[t_2] = \mathbf{\Gamma} \mathbf{y}_R[t_1]$, where $t_2 \in \mathcal{T}_2(F(\mathbf{Q}))$. Here

$$
\mathbf{\Gamma} = \begin{bmatrix} \gamma \frac{\det(\mathbf{Q}_1)^*}{|\det(\mathbf{Q}_1)|} \mathbf{\Lambda} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \gamma \frac{\det(\mathbf{Q}_2)^*}{|\det(\mathbf{Q}_2)|} \mathbf{\Lambda} & & \vdots \\ \vdots & & \ddots & \\ \mathbf{0} & \cdots & & \gamma \frac{\det(\mathbf{Q}_M)^*}{|\det(\mathbf{Q}_M)|} \mathbf{\Lambda} \end{bmatrix}, \tag{21}
$$

$\gamma = \sqrt{\frac{P}{1+2P}}$, and $\mathbf{\Lambda} = [[1,0]^T [0,-1]^T]^T$, where $\mathbf{Q} = [\mathbf{Q}_1^T, \cdots, \mathbf{Q}_M^T]^T$ and $\mathbf{0}$ denotes the 2×2 dimensional all-zero matrix.
- (Decoding) The destinations decode their messages based on their received signals during $t \in \{n_B + 1, \cdots, 2n_B\}$.

C. Achievable Rate Region

In this subsection, we prove Theorem 1. We first introduce the following two lemmas.

Lemma 1: For any $\mathbf{Q} \in \mathcal{Q}_1$,

$$
\mathsf{P}\left[\mathbf{H}[t] \in \mathcal{A}_1(\mathbf{Q})\right] = \mathsf{P}\left[\mathbf{G}[t] \in \mathcal{A}_2(F(\mathbf{Q}))\right]. \tag{22}
$$

Proof: Let $f_{\mathbf{H}[t]}(\cdot)$ and $f_{\mathbf{G}[t]}(\cdot)$ denote the probability density functions of $\mathbf{H}[t]$ and $\mathbf{G}[t]$, respectively. Then

$$
\begin{aligned}
\mathsf{P}\left[\mathbf{H}[t] \in \mathcal{A}_1(\mathbf{Q})\right] &= \int_{\mathbf{A} \in \mathcal{A}_1(\mathbf{Q})} f_{\mathbf{H}[t]}(\mathbf{A}) d\mathbf{A} \\
&= \int_{\mathbf{A} \in \mathcal{A}_1(\mathbf{Q})} \prod_{i \in \{1, \cdots 2M\}, j \in \{1,2\}} f(a_{i,j}) d\mathbf{A} \\
&\overset{(a)}{=} \int_{\mathbf{A} \in \mathcal{A}_1(\mathbf{Q})} f_{\mathbf{G}[t]}(F(\mathbf{A})) d\mathbf{A} \\
&\overset{(b)}{=} \int_{\mathbf{A}' \in \mathcal{A}_2(F(\mathbf{Q}))} f_{\mathbf{G}[t_2]}(\mathbf{A}') d\mathbf{A}' \\
&= \mathsf{P}\left[\mathbf{G}[t] \in \mathcal{A}(F(\mathbf{Q}))\right],
\end{aligned}
\tag{23}
$$

where $\mathbf{A} = \{a_{i,j}\}$. Here (a) follows from the definition of $F(\mathbf{A})$ and (b) follows by a change of variable $\mathbf{A}' = F(\mathbf{A})$ whose Jacobian is one and $\mathcal{A}_2(F(\mathbf{Q})) = \{F(\mathbf{A}) | \mathbf{A} \in \mathcal{A}_1(\mathbf{Q})\}$. Therefore Lemma 1 holds. ∎

Lemma 2: The probability that

$$
\left| \frac{\mathrm{card}(\mathcal{T}_1(\mathbf{Q}_1))}{n_B} - \mathsf{P}[\mathbf{H}[t] \in \mathcal{A}_1(\mathbf{Q}_1)] \right| \leq \delta
\tag{24}
$$

and

$$
\left| \frac{\mathrm{card}(\mathcal{T}_2(\mathbf{Q}_2))}{n_B} - \mathsf{P}[\mathbf{G}[t] \in \mathcal{A}_2(\mathbf{Q}_2)] \right| \leq \delta
\tag{25}
$$

for all $\mathbf{Q}_1 \in \mathcal{Q}_1$ and $\mathbf{Q}_2 \in \mathcal{Q}_2$ is greater than $1 - (\mathrm{card}(\mathcal{Q}_1) + \mathrm{card}(\mathcal{Q}_2))/(2n_B\delta^2)$.

Proof: We refer to Lemma 2.12 in [30] for the proof. ∎

Suppose that the sources transmit at time $t_1 \in \mathcal{T}_1(\mathbf{Q})$ and the relays amplify and forward their received signals at time $t_2 \in \mathcal{T}_2(F(\mathbf{Q}))$, where $\mathbf{Q} \in \mathcal{Q}_1$. For this case, from (1) and (3), the received signal vector of the destinations is given by

$$
\mathbf{y}[t_2] = \mathbf{G}[t_2]\boldsymbol{\Gamma}\mathbf{H}[t_1]\mathbf{x}[t_1] + \mathbf{G}[t_2]\boldsymbol{\Gamma}\mathbf{z}_R[t_1] + \mathbf{z}[t_2],
\tag{26}
$$

where we use $\mathbf{x}_R[t_2] = \boldsymbol{\Gamma}\mathbf{y}_R[t_1]$. Denote $\mathbf{H}[t_1] = \mathbf{H} = [\mathbf{H}_1^T, \cdots \mathbf{H}_M^T]^T$ and $\mathbf{G}[t_2] = F(\mathbf{H}) + \boldsymbol{\Delta}$, where $\boldsymbol{\Delta} = [\boldsymbol{\Delta}_1, \cdots, \boldsymbol{\Delta}_M]$ is the quantization error matrix with respect to $F(\mathbf{H})$. From (26),

$$
\mathbf{y}[t_2] = \left(\left(\gamma \sum_{m=1}^M |\det(\mathbf{H}_m)| \right) \boldsymbol{\Lambda} + \boldsymbol{\Delta}\boldsymbol{\Gamma}\mathbf{H} \right) \mathbf{x}[t_1] + (F(\mathbf{H}) + \boldsymbol{\Delta})\boldsymbol{\Gamma}\mathbf{z}_R[t_1] + \mathbf{z}[t_2],
\tag{27}
$$

where we use $F(\mathbf{H})\boldsymbol{\Gamma}\mathbf{H} = (\gamma \sum_{m=1}^M |\det(\mathbf{H}_m)|)\boldsymbol{\Lambda}$. Thus, the received signal-to-interference-and-noise ratio (SINR) of destination i is given by

$$
\mathrm{SINR}_i = \frac{P \left| (-1)^{i-1}(\gamma \sum_{m=1}^M |\det(\mathbf{H}_m)|) + [\boldsymbol{\Delta}\boldsymbol{\Gamma}\mathbf{H}]_{i,i} \right|^2}{1 + \gamma \sum_{m=1}^M \left(|[\mathbf{H}_m]_{3-i,3-i} + [\boldsymbol{\Delta}_m]_{i,i}|^2 + |[\mathbf{H}_m]_{i,3-i} + [\boldsymbol{\Delta}_m]_{i,3-i}|^2 \right) + P|[\boldsymbol{\Delta}\boldsymbol{\Gamma}\mathbf{H}]_{i,3-i}|^2}.
\tag{28}
$$

Define $R_i(\mathbf{Q}) = \min_{\mathbf{A} \in \mathcal{A}_1(\mathbf{Q})} \log(1 + \mathrm{SINR}_i)$. Then an achievable rate of destination i is lower bounded by

$$
R_i \geq \frac{1}{n_B} \sum_{\mathbf{Q} \in \mathcal{Q}_1} R_i(\mathbf{Q}) \min\{\mathrm{card}(\mathcal{T}_1(\mathbf{Q})), \mathrm{card}(\mathcal{T}_2(F(\mathbf{Q})))\}.
\tag{29}
$$

From Lemmas 1 and 2,

$$
\mathrm{card}(\mathcal{T}_1(\mathbf{Q})) \geq n_B(\mathsf{P}[\mathbf{H}[t] \in \mathcal{A}_1(\mathbf{Q})] - \delta)
\tag{30}
$$

and

$$
\begin{aligned}
\mathrm{card}(\mathcal{T}_2(F(\mathbf{Q}))) &\geq n_B(\mathsf{P}[\mathbf{G}[t] \in \mathcal{A}_2(F(\mathbf{Q}))] - \delta) \\
&= n_B(\mathsf{P}[\mathbf{H}[t] \in \mathcal{A}_1(\mathbf{Q})] - \delta)
\end{aligned}
\tag{31}
$$

for all $\mathbf{Q} \in \mathcal{Q}_1$ with probability greater than $1 - \frac{(2N+1)^{8M}}{n_B \delta^2}$, where we use $\text{card}(\mathcal{Q}_1) = \text{card}(\mathcal{Q}_2) = (2N+1)^{8M}$. Then

$$
\begin{aligned}
R_i &\geq \sum_{\mathbf{Q} \in \mathcal{Q}_1} R_i(\mathbf{Q})(\mathsf{P}[\mathbf{H}[t] \in \mathcal{A}_1(\mathbf{Q})] - \delta) \\
&\geq \sum_{\mathbf{Q} \in \mathcal{Q}_1} R_i(\mathbf{Q}) \, \mathsf{P}[\mathbf{H}[t] \in \mathcal{A}(\mathbf{Q})] - \delta 2(2N+1)^{8M} \max_{\mathbf{Q} \in \mathcal{Q}_1} R_i(\mathbf{Q})
\end{aligned}
\tag{32}
$$

is achievable with probability greater than $1 - \frac{(2N+1)^{8M}}{n_B \delta^2}$. By setting $\Delta = n_B^{-1/(3 \cdot 2^5 M)}$, $N = n_B^{1/(3 \cdot 2^4 M)}$, and $\delta = n_B^{-1/3}$, the following condition can be satisfied:

$$
\begin{aligned}
\Delta &= n_B^{-1/(3 \cdot 2^5 M)} \to 0, \\
\Delta N &= n_b^{1/(3 \cdot 2^5 M)} \to \infty, \\
\delta 2(2N+1)^{8M} \max_{\mathbf{Q} \in \mathcal{Q}_1} R_i(\mathbf{Q}) &\leq 2 \cdot 3^{8M} N^{8M} \delta \max_{\mathbf{Q} \in \mathcal{Q}_1} R_i(\mathbf{Q}) \\
&\overset{(a)}{\leq} 2 \cdot 3^{8M} N^{8M} \delta \log(1 + 2^4 M \Delta^2 N^2 P) \\
&= 2 \cdot 3^{8M} n_B^{-1/6} \log(1 + 2^4 M n_B^{1/(3 \cdot 2^4 M)} P) \to 0, \\
\frac{(2N+1)^{8M}}{n_B \delta^2} &\leq \frac{3^{8M} N^{8M}}{n_B \delta^2} = 3^{8M} n_B^{-1/6} \to 0
\end{aligned}
\tag{33}
$$

as n_B increases, where (a) follows since $|h_{ij}|^2 \leq 2\Delta^2(N + \frac{1}{2})^2 \leq 2^3 \Delta^2 N^2$ for the channel instances using the transmission (see Fig. 7).

Hence,

$$
R_i = \mathsf{E}\left[\log\left(1 + \frac{P\gamma^2 \left(\sum_{m=1}^M |\det(\mathbf{H}_m)| \right)^2}{1 + \gamma^2 \sum_{m=1}^M (|h_{2(m-1)+3-i,3-i}|^2 + |h_{2(m-1)+i,3-i}|^2)} \right) \right]
\tag{34}
$$

is achievable with probability approaching one for $i \in \{1, 2\}$, where we use the fact that

$$
\lim_{\Delta \to 0} \text{SINR}_i = \frac{P\gamma^2 \left(\sum_{m=1}^M |\det(\mathbf{H}_m)| \right)^2}{1 + \gamma^2 \sum_{m=1}^M (|h_{2(m-1)+3-i,3-i}|^2 + |h_{2(m-1)+i,3-i}|^2)}.
\tag{35}
$$

In conclusion, Theorem 1 holds.

V. APPROXIMATE CAPACITY CHARACTERIZATION

In this section, we prove Theorems 2 to 5, the approximate ergodic sum capacity characterization results. We will deal with the difference between R_{mimo} and R_{in}, which are given by (8) and (9) respectively. Throughout this section, we assume a class of channel distributions such that $f(x)$ is only a function of $|x|$. That is, for given amplitudes of the channel coefficients, their phases are i.i.d. uniformly distributed over $[0, 2\pi)$. For instance, this class of channel distributions includes i.i.d. uniform phase fading and i.i.d. Rayleigh fading as special cases. We omit the time index t in this section for notational convenience.

A. Approximate Capacity for $L = 2$

We first consider the case where $L = 2$. In order to deal with i.i.d. random phase in the rate expression in Theorem 1, we introduce the following lemma showing the exact solution of $\mathsf{E}_\phi \left[\log(1 - x \cos\phi) \right]$ for $|x| \leq 1$ when ϕ is uniformly distributed over $[0, 2\pi)$.

Lemma 3: Let ϕ be a random variable uniformly distributed over $[0, 2\pi)$. For $|x| \leq 1$,

$$
\mathsf{E}_\phi \left[\log(1 - x \cos\phi) \right] = \log\left(1 + \sqrt{1 - x^2} \right) - 1.
\tag{36}
$$

Proof: We refer to the equation (4.224 12) in [31]. ∎

1) Proof of Theorem 2: From (8),

$$R_{\text{in}} \overset{(a)}{=} 2\,\mathsf{E}_\theta\left[\log\left(1 + \frac{2P^2(1-\cos\theta)}{1+4P}\right)\right]$$

$$= 2\log\left(1 + \frac{2P^2}{1+4P}\right) + 2\,\mathsf{E}_\theta\left[\log\left(1 - \frac{2P^2}{1+4P+2P^2}\cos\theta\right)\right]$$

$$\overset{(b)}{=} 2\log\left(1 + \frac{2P^2}{1+4P}\right) + 2\log\left(1 + \sqrt{1 - \left(\frac{2P^2}{1+4P+2P^2}\right)^2}\right) - 2, \tag{37}$$

where $\theta = \theta_{1,1} + \theta_{2,2} - \theta_{1,2} - \theta_{2,1}$. Here, (a) follows since $|\det(\mathbf{H})|^2 = 2(1-\cos\theta)$ and $\sigma^2_{\text{AF},i} = \frac{2P}{1+2P}$, (b) follows since $\theta \bmod [2\pi]$ is uniformly distributed over $[0, 2\pi)$ and from Lemma 3 with $|\frac{2P^2}{1+4P+2P^2}| \le 1$. Similarly, from (9),

$$R_{\text{mimo}} = \mathsf{E}_\theta\left[\log((1+2P)^2 - 2P^2(1+\cos\theta))\right]$$

$$= \log(1 + 4P + 2P^2) + \mathsf{E}_\theta\left[\log\left(1 - \frac{2P^2}{1+4P+2P^2}\cos\theta\right)\right]$$

$$= \log(1 + 4P + 2P^2) + \log\left(1 + \sqrt{1 - \left(\frac{2P^2}{1+4P+2P^2}\right)^2}\right) - 1. \tag{38}$$

Then, from (37) and (38),

$$R_{\text{mimo}} - R_{\text{in}} = \log\left(\frac{(1+4P)^2}{1+4P+2P^2}\right) - \log\left(1 + \sqrt{1 - \left(\frac{2P^2}{1+4P+2P^2}\right)^2}\right) + 1$$

$$\overset{(a)}{\le} \log\left(\frac{(1+4P)^2}{1+4P+2P^2}\right) + 1$$

$$\overset{(b)}{\le} 4, \tag{39}$$

where (a) follows since $|\frac{2P^2}{1+4P+2P^2}| \le 1$ for any $P > 0$ and (b) follows since

$$\log\left(\frac{(1+4P)^2}{1+4P+2P^2}\right) \le \log\left(\frac{(1+4P)^2}{1+2\sqrt{2}P+2P^2}\right)$$

$$= 2\log\left(\frac{1+4P}{1+\sqrt{2}P}\right)$$

$$\le 3, \tag{40}$$

where we use the fact that $\log\left(\frac{1+4P}{1+\sqrt{2}P}\right)$ is an increasing function of $P > 0$ and $\lim_{P\to\infty}\log\left(\frac{1+4P}{1+\sqrt{2}P}\right) = \frac{3}{2}$. In conclusion, Theorem 2 holds.

2) Proof of Theorem 3: Since $f(x)$ is only a function of $|x|$, $h_{i,j}$ can be represented as $a_{i,j}\exp(\jmath\theta_{i,j})$, where $a_{i,j} \ge 0$ and $\theta_{i,j} \in [0, 2\pi)$ are independent of each other. Moreover $\theta_{i,j}$ is uniformly distributed over $[0, 2\pi)$. To simplify the notation, we denote $\mathbf{a} = \{a_{1,1}, a_{1,2}, a_{2,1}, a_{2,2}\}$, $A = a^2_{1,1}a^2_{2,2} + a^2_{1,2}a^2_{2,1}$, $B_1 = a^2_{1,1} + a^2_{2,1} + 2$, $B_2 = a^2_{1,2} + a^2_{2,2} + 2$, $G = 2a_{1,1}a_{1,2}a_{2,1}a_{2,2}$, and $S = a^2_{1,1} + a^2_{1,2} + a^2_{2,1} + a^2_{2,2}$.

From (8),

$$R_{\text{in}} \overset{(a)}{=} \sum_{i \in \{1,2\}} \mathsf{E}_{\mathbf{a}} \left[\mathsf{E}_{\theta} \left[\log \left(1 + \frac{P^2(A - G\cos\theta)}{1 + PB_i} \right) \right] \right],$$

$$= \sum_{i \in \{1,2\}} \mathsf{E}_{\mathbf{a}} \left[\log \left(1 + \frac{P^2 A}{1 + PB_i} \right) \right] + \sum_{i \in \{1,2\}} \mathsf{E}_{\mathbf{a}} \left[\mathsf{E}_{\theta} \left[\log \left(1 - \frac{P^2 G \cos\theta}{1 + PB_i + P^2 A} \right) \right] \right]$$

$$\overset{(b)}{=} \sum_{i \in \{1,2\}} \mathsf{E}_{\mathbf{a}} \left[\log \left(1 + \frac{P^2 A}{1 + PB_i} \right) \right] + \sum_{i \in \{1,2\}} \mathsf{E}_{\mathbf{a}} \left[\log \left(1 + \sqrt{1 - \left(\frac{P^2 G}{1 + PB_i + P^2 A} \right)^2} \right) \right] - 2$$

$$\overset{(c)}{\geq} \sum_{i \in \{1,2\}} \mathsf{E}_{\mathbf{a}} \left[\log \left(1 + \frac{P^2 A}{1 + PB_i} \right) \right] + 2\,\mathsf{E}_{\mathbf{a}} \left[\log \left(1 + \frac{\sqrt{A^2 - G^2}}{A} \right) \right] - 2$$

$$= \sum_{i \in \{1,2\}} \mathsf{E}_{\mathbf{a}} \left[\log \left(1 + \frac{PA}{B_i} \right) \right] + \sum_{i \in \{1,2\}} \mathsf{E}_{\mathbf{a}} \left[\log \left(\frac{B_i + PB_i^2 + P^2 AB_i}{B_i + P(A + B_i^2) + P^2 AB_i} \right) \right]$$

$$+ 2\,\mathsf{E}_{\mathbf{a}} \left[\log \left(1 + \frac{\sqrt{A^2 - G^2}}{A} \right) \right] - 2$$

$$\overset{(d)}{\geq} \sum_{i \in \{1,2\}} \mathsf{E}_{\mathbf{a}} \left[\log \left(1 + \frac{PA}{B_i} \right) \right] + \sum_{i \in \{1,2\}} \mathsf{E}_{\mathbf{a}} \left[\log \left(\frac{B_i^2}{A + B_i^2} \right) \right]$$

$$+ 2\,\mathsf{E}_{\mathbf{a}} \left[\log \left(1 + \frac{\sqrt{A^2 - G^2}}{A} \right) \right] - 2, \tag{41}$$

where $\theta = \theta_{1,1} + \theta_{2,2} - \theta_{1,2} - \theta_{2,1}$. Here (a) follows from the facts that $\mathbf{a}$ and $\{\theta_{1,1}, \theta_{1,2}, \theta_{2,1}, \theta_{2,2}\}$ are independent of each other and $|\det(\mathbf{H})|^2 = A - G\cos\theta$, (b) follows since $\theta \bmod [2\pi]$ is uniformly distributed over $[0, 2\pi)$ and from Lemma 3 with $\left| \frac{P^2 G}{1 + PB_i + P^2 A} \right| \leq 1$, (c) follows since $\frac{P^2 G}{1 + PB_i + P^2 A} \leq \frac{G}{A}$ for any $P \geq 0$, and (d) follows since $\log \left(\frac{c_1 + c_2}{c_1 + c_3} \right) \geq \log \left(\frac{c_2}{c_3} \right)$ for $c_1, c_2, c_3 > 0$ and $c_2 \leq c_3$.

From (9),

$$R_{\text{mimo}} \overset{(a)}{=} \mathsf{E}_{\mathbf{a}} \left[\mathsf{E}_{\theta} \left[\log \det(\mathbf{I} + P\mathbf{H}\mathbf{H}^{\dagger}) \right] \right]$$

$$= \mathsf{E}_{\mathbf{a}} \left[\mathsf{E}_{\theta} \left[\log \left(1 + PS + P^2(A - G\cos\theta) \right) \right] \right]$$

$$= \mathsf{E}_{\mathbf{a}} \left[\log(1 + PS + P^2 A) \right] + \mathsf{E}_{\mathbf{a}} \left[\mathsf{E}_{\theta} \left[\log \left(1 - \frac{P^2 G}{1 + PS + P^2 A} \cos\theta) \right) \right] \right]$$

$$\overset{(b)}{=} \mathsf{E}_{\mathbf{a}} \left[\log(1 + PS + P^2 A) \right] + \mathsf{E}_{\mathbf{a}} \left[\log \left(1 + \sqrt{1 - \left(\frac{P^2 G}{1 + PS + P^2 A} \right)^2} \right) \right] - 1$$

$$\overset{(c)}{\leq} \mathsf{E}_{\mathbf{a}} \left[\log(1 + PS + P^2 A) \right], \tag{42}$$

where (a) follows from the fact that $\mathbf{a}$ and $\{\theta_{1,1}, \theta_{1,2}, \theta_{2,1}, \theta_{2,2}\}$ are independent of each other, (b) follows since $\theta \bmod [2\pi]$ is uniformly distributed over $[0, 2\pi)$ and from Lemma 3 and $\left| \frac{P^2 G}{1 + PS + P^2 A} \right| \leq 1$, and (c) follows again since $\left| \frac{P^2 G}{1 + PS + P^2 A} \right| \leq 1$.

Let

$$\Delta = \log(1 + PS + P^2 A) - \sum_{i \in \{1,2\}} \log \left(1 + \frac{PA}{B_i} \right). \tag{43}$$

Then

$$\Delta = \log\left(\frac{B_1 B_2}{A}\right) + \log\left(\frac{1 + PS + P^2 A}{\frac{B_1 B_2}{A} + P(B_1 + B_2) + P^2 A}\right)$$
$$\leq \log\left(\frac{B_1 B_2}{A}\right), \tag{44}$$

where the inequality follows since $B_1 B_2 \geq A$ and $B_1 + B_2 \geq S$. Therefore, from (41) to (44),

$$R_{\mathrm{mimo}} - R_{\mathrm{in}} \leq \mathsf{E_a}[\Delta] - \sum_{i \in \{1,2\}} \mathsf{E_a}\left[\log\left(\frac{B_i^2}{A + B_i^2}\right)\right] - 2\,\mathsf{E_a}\left[\log\left(1 + \frac{\sqrt{A^2 - G^2}}{A}\right)\right] + 2$$
$$\leq \mathsf{E_a}\left[\log\left(\frac{A(A + B_1^2)(A + B_2^2)}{B_1 B_2(A + \sqrt{A^2 - G^2})^2}\right)\right] + 2$$
$$= 2\,\mathsf{E_a}\left[\log\left(\frac{\sqrt{A}(A + B_1^2)}{B_1(A + \sqrt{A^2 - G^2})}\right)\right] + 2. \tag{45}$$

In conclusion, Theorem 3 holds.

B. Approximate Capacity as $L \to \infty$

In this subsection, we characterize an approximate ergodic sum capacity in the limit of large number of relays by deriving $\lim_{L \to \infty}\{R_{\mathrm{mimo}} - R_{\mathrm{in}}\}$. For K-user 2-hop networks with L relays, it was shown in [9] that interference can be completely neutralized if $K \geq N(N-1) + 1$, which indicates that for $2 \times L \times 2$ networks interference neutralization can be achieved without channel pairing if $L \geq 3$. However, maximizing the achievable sum rate exploiting interference neutralization without channel pairing presented in [32] is non-convex and, as a result, it is unclear how to determine the sum rate gap from the cut-set upper bound. By contrast, we now show that our achievable rate expression from Theorem 1 permits to derive a finite-gap result. The rate expression R_{in} in (8) contains the sum of i.i.d. random variables, i.e., $\sum_{m=1}^{M} |\det(\mathbf{H}_m)|$, which approaches a deterministic value $M\,\mathsf{E}[|\det(\mathbf{H}_1)|]$ almost surely as $M \to \infty$ by the law of large numbers. The following lemma provides a rigorous lower bound in order to deal with R_{in} that holds for any M.

Lemma 4: Consider a sequence of i.i.d. nonnegative random variables $\{X_i, i \in \mathbb{Z}_+\}$. Let $S_m = \sum_{i=1}^{m} X_i$. If $\mathsf{E}[X_1{}^2] < \infty$, then for any $\epsilon \in (0, \mathsf{E}[X_1])$ and any $c > 0$,

$$\mathsf{E}\left[\log(1 + cS_m{}^2)\right] \geq \log\left(1 + cm^2(\mathsf{E}[X_1])^2\right) - \delta_m(c, \mathsf{E}[X_1], \mathsf{E}[X_1{}^2]), \tag{46}$$

where

$$\delta_m(c, \mathsf{E}[X_1], \mathsf{E}[X_1{}^2]) = \frac{\mathsf{E}[X_1{}^2]}{m\epsilon^2} \log\left(1 + cm^2(\mathsf{E}[X_1] - \epsilon)^2\right)$$
$$- \log\left(1 - \frac{cm^2\epsilon(2\,\mathsf{E}[X_1] - \epsilon)}{1 + cm^2(\mathsf{E}[X_1])^2}\right) \tag{47}$$

is a positive sequence of m, which converges to zero as $\epsilon \to 0$.

Proof: We have

$$\mathsf{E}\left[\log(1 + cS_m{}^2)\right]$$

$$= \mathsf{E}\left[\log\left(1 + cS_m{}^2\right)\left(1_{\{|S_m/m - \mathsf{E}[X_1]| < \epsilon\}} + 1_{\{|S_m/m - \mathsf{E}[X_1]| \geq \epsilon\}}\right)\right]$$

$$\overset{(a)}{\geq} \log\left(1 + cm^2(\mathsf{E}[X_1])^2 - 2cm^2\epsilon\,\mathsf{E}[X_1] + cm^2\epsilon^2)\right)\mathsf{E}[1_{\{|S_m/m - \mathsf{E}[X_1]| < \epsilon\}}]$$

$$\overset{(b)}{\geq} \log\left(1 + cm^2(\mathsf{E}[X_1])^2 - 2cm^2\epsilon\,\mathsf{E}[X_1] + cm^2\epsilon^2)\right)\left(1 - \frac{\mathrm{Var}(X_1)}{m\epsilon^2}\right)$$

$$\overset{(c)}{\geq} \log\left(1 + cm^2(\mathsf{E}[X_1])^2 - 2cm^2\epsilon\,\mathsf{E}[X_1] + cm^2\epsilon^2)\right)$$

$$\quad - \frac{\mathsf{E}[X_1{}^2]}{m\epsilon^2}\log\left(1 + cm^2(\mathsf{E}[X_1])^2 + cm^2\epsilon^2)\right)$$

$$= \log\left(1 + cm^2(\mathsf{E}[X_1])^2\right) - \delta_m(c, \mathsf{E}[X_1], \mathsf{E}[X_1{}^2]), \tag{48}$$

where (a) follows since $S_m > m\,\mathsf{E}[X_1] - m\epsilon$ under the condition $|S_m/m - \mathsf{E}[X_1]| < \epsilon$, (b) follows from Chebyshev's inequality, and (c) follows since $\mathrm{Var}(X_1) \leq \mathsf{E}[X_1{}^2]$. In conclusion, Lemma 4 holds. ∎

By setting δ_m arbitrarily small as m increases, Lemma 4 provides that

$$\mathsf{E}\left[\log(1 + cS_m{}^2)\right] \geq \log\left(1 + cm^2(\mathsf{E}[X_1])^2\right) \tag{49}$$

in the limit of large m. Note that this bound is asymptotically tight since $\mathsf{E}\left[\log(1 + cS_m{}^2)\right] \leq \log(1 + c\,\mathsf{E}[S_m{}^2])$ from Jensen's inequality and $\log(1 + c\,\mathsf{E}[S_m{}^2])$ is approximately given as $\log(1 + cm^2\,\mathsf{E}[X_1]^2)$ as m increases.

1) Proof of Theorem 4: Recall $M = \lfloor\frac{L}{2}\rfloor$. That is, $L \leq 2M + 1$. From (9),

$$R_{\mathrm{mimo}} \leq 2\log(1 + P(2M + 1)), \tag{50}$$

where the inequality follows from Jensen's inequality and the fact that $\log\det(\cdot)$ is a concave function [33]. Here we assume $L = 2M + 1$ to obtain an upper bound.

From (8),

$$R_{\mathrm{in}} \overset{(a)}{=} 2\,\mathsf{E}_{\{\theta_1,\cdots,\theta_M\}}\left[\log\left(1 + \frac{P^2\left(\sum_{m=1}^{M}\sqrt{2 - 2\cos\theta_m}\right)^2}{1 + P(2M + 2)}\right)\right]$$

$$\overset{(b)}{\geq} 2\log\left(1 + \frac{\frac{16}{\pi^2}P^2M^2}{1 + P(2M + 2)}\right) - 2\delta_M\left(\frac{P^2}{1 + P(2M + 2)}, \frac{4}{\pi}, 2\right), \tag{51}$$

where $\theta_m = \theta_{2m-1,1} + \theta_{2m,2} - \theta_{2m-1,2} - \theta_{2m,1}$. Here, (a) follows since $|\det(\mathbf{H}_m)| = \sqrt{2 - 2\cos\theta_m}$ and (b) follows since $\theta_m \bmod [2\pi]$ is uniformly distributed over $[0, 2\pi)$ and from Lemma 4 with the facts that $\mathsf{E}\left[\sqrt{2 - 2\cos\theta_1}\right] = \frac{4}{\pi}$, and $\mathsf{E}[2 - 2\cos\theta_1] = 2$. Then, from (50) and (51),

$$R_{\mathrm{mimo}} - R_{\mathrm{in}} \leq 2\log\left(\frac{1 + P(4M + 3) + P^2(2M + 1)(2M + 2)}{1 + P(2M + 2) + \frac{16}{\pi^2}P^2M^2}\right)$$

$$\quad + 2\delta_M\left(\frac{P^2}{1 + P(2M + 2)}, \frac{4}{\pi}, 2\right). \tag{52}$$

Hence, $\lim_{M\to\infty}\{R_{\mathrm{mimo}} - R_{\mathrm{in}}\} \leq 4\log\pi - 4 + \epsilon_1$, where

$$\epsilon_1 = \lim_{M\to\infty} 2\delta_M\left(\frac{P^2}{1 + P(2M + 2)}, \frac{4}{\pi}, 2\right)$$

$$= -2\log\left(1 - \epsilon\left(\frac{\pi}{2} - \frac{\pi^2}{16}\epsilon\right)\right) > 0, \tag{53}$$

which can be arbitrarily small as ϵ decreases. In conclusion, Theorem 4 holds.

2) Proof of Theorem 5: From (8),

$$
\begin{aligned}
R_{\text{in}} \geq\ & 2\,\mathsf{E}\left[\log\left(1 + P^2\left(\sum_{m=1}^{M}|\det(\mathbf{H}_m)|\right)^2\right)\right] \\
& - 2\,\mathsf{E}\left[\log\left(1 + P\left(\sum_{m=1}^{2M}|h_{m,1}|^2 + 2\right)\right)\right] \\
\overset{(a)}{\geq}\ & 2\log\left(\frac{1 + P^2 M^2\left(\mathsf{E}[|\det(\mathbf{H}_1)|]\right)^2}{1 + P(2M+2)}\right) \\
& - 2\delta_M\left(P^2, \mathsf{E}[|\det(\mathbf{H}_1)|], \mathsf{E}\left[|\det(\mathbf{H}_1)|^2\right]\right),
\end{aligned}
\tag{54}
$$

where (a) follows from Lemma 4 and Jensen's inequality. Hence, from (50) and (54),

$$
\begin{aligned}
R_{\text{mimo}} - R_{\text{in}} \leq\ & 2\log\left(\frac{1 + P(4M+3) + P^2(2M+1)(2M+2)}{1 + P^2 M^2\left(\mathsf{E}[|\det(\mathbf{H}_1)|]\right)^2}\right) \\
& + 2\delta_M\left(P^2, \mathsf{E}[|\det(\mathbf{H}_1)|], \mathsf{E}\left[|\det(\mathbf{H}_1)|^2\right]\right)
\end{aligned}
\tag{55}
$$

and

$$
\lim_{M\to\infty}\left\{R_{\text{mimo}} - R_{\text{in}}\right\} \leq 4 - 4\log\left(\mathsf{E}[|\det(\mathbf{H}_1)|]\right) + \epsilon_2,
\tag{56}
$$

where

$$
\begin{aligned}
\epsilon_2 &= \lim_{M\to\infty} 2\delta_M\left(P^2, \mathsf{E}[|\det(\mathbf{H}_1)|], \mathsf{E}\left[|\det(\mathbf{H}_1)|^2\right]\right) \\
&= -2\log\left(1 - \frac{\epsilon(2\,\mathsf{E}[|\det(\mathbf{H}_1)|] - \epsilon)}{(\mathsf{E}[|\det(\mathbf{H}_1)|])^2}\right) > 0,
\end{aligned}
\tag{57}
$$

which can be arbitrarily small as ϵ decreases. In conclusion, Theorem 5 holds.

VI. Conclusion

In this paper, we studied a fading 2-user 2-hop network with L relays where channel coefficients vary over time. In spite of recent achievements in this area, the best known capacity characterization is to within $o(\log \text{SNR})$ bits/sec/Hz from the ergodic sum capacity, which can be arbitrarily large as SNR increases. For a broad class of channel distributions, we tightened this gap to within a constant number of bits/sec/Hz, independent of SNR. The achievability follows from ergodic interference neutralization in which the relays are partitioned into several pairs and interference is neutralized separately by each pair of relays. The proposed scheme makes interference neutralized in the finite SNR regime and, at the same time, the intended signal power increased quadratically with L, leading that the optimal $2\log(L\text{SNR})$ rate scaling is achievable, which cannot be captured by the previous DoF work.

Appendix I
Quantization for i.i.d. Uniform Phase Fading

For i.i.d. uniform phase fading, $h_{i,j}[t]$ and $g_{j,i}[t]$ respectively are represented as $\exp(\jmath\theta_{i,j}[t])$ and $\exp(\jmath\varphi_{j,i}[t])$ for $i \in \{1, \cdots, L\}$ and $j \in \{1, 2\}$. Hence we can quantize the channel space of each hop based on angles. Specifically, the channel space of the first hop can be partitioned as follows. For $N \in \mathbb{Z}_+$, first define $\mathcal{Q}_1 := \left\{\exp(\jmath 0), \exp(\jmath\frac{2\pi}{N}), \exp(\jmath\frac{4\pi}{N}), \cdots, \exp(\jmath\frac{(N-1)2\pi}{N})\right\}^{2M\times 2}$. Let $\mathcal{U}$ denote the set of all $x \in \mathbb{C}$ satisfying $|x| = 1$. For a quantized channel matrix $\mathbf{Q} \in \mathcal{Q}_1$, define $\mathcal{A}_1(\mathbf{Q}) := \left\{\mathbf{A} \in \mathcal{U}^{2M\times 2} \mid -\frac{\pi}{N} \leq \angle a_{i,j} - \angle q_{i,j} < \frac{\pi}{N}\right.$ for all $i \in \{1, \cdots, 2M\}$ and $j \in \{1, 2\}\}$, where $\mathbf{A} = \{a_{i,j}\}$, $\mathbf{Q} = \{q_{i,j}\}$, and $\angle x$ denotes the angle of $x \in \mathcal{U}$, i.e., $x = \exp(\jmath\angle x)$. Figure 8 illustrates the channel space partitioning with respect to $h_{i,j} \in \mathcal{U}$. In a similar manner, we can define $\mathcal{Q}_2$ and $\mathcal{A}_2(\mathbf{Q})$ for the second hop. Then we can show that there exists an increasing sequence of N, which is a function of n_B, such that (7) is achievable as n_B increases using similar steps in the proof of Theorem 1.
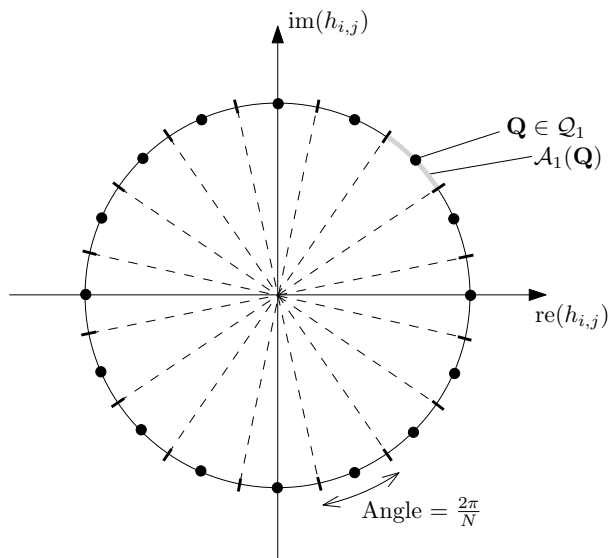
Fig. 8. Channel space partitioning with respect to the channel coefficient $h_{i,j} \in \mathcal{U}$ for i.i.d. uniform phase fading.

APPENDIX II
CLOSED FORM OF $\mathsf{E}[|\det(\mathbf{H}_1)|]$ FOR I.I.D. RAYLEIGH FADING

Let $\mathbf{A}$ be a 2×2 matrix whose entries are i.i.d. circularly symmetric complex Gaussian random variables with mean zero and unit variance and $\mathbf{W} := 2\mathbf{A}\mathbf{A}^\dagger$. Let λ_1 and λ_2, $\lambda_1 \geq \lambda_2$, be the eigenvalues of $\mathbf{W}$. Then the joint probability density function of λ_1 and λ_2 is given by [34, Equation (3.11)]

$$f(\lambda_1, \lambda_2) = \frac{1}{16} \exp\left(-\frac{1}{2}(\lambda_1 + \lambda_2)\right)(\lambda_1 - \lambda_2)^2 \mathbf{1}_{\{\lambda_1 \geq \lambda_2 \geq 0\}}(\lambda_1, \lambda_2). \tag{58}$$

Thus,

$$
\begin{aligned}
\mathsf{E}[|\det(\mathbf{H}_1)|] &= \frac{1}{2}\mathsf{E}\left[\sqrt{\det(\mathbf{W})}\right] \\
&= \frac{1}{2}\mathsf{E}\left[\sqrt{\lambda_1 \lambda_2}\right] \\
&= \frac{1}{32}\int_0^\infty \int_0^{\lambda_1} \sqrt{\lambda_1 \lambda_2} \exp\left(-\frac{1}{2}(\lambda_1 + \lambda_2)\right)(\lambda_1 - \lambda_2)^2 \, d\lambda_1 d\lambda_2 \\
&\stackrel{(a)}{=} \frac{1}{32}\int_0^\infty \int_0^u \sqrt{(u+v)(u-v)} \exp\left(-u\right)(2v)^2 \, 2dv du \\
&= \frac{1}{4}\int_0^\infty \left(\int_0^u v^2\sqrt{u^2 - v^2} \, dv\right)\exp\left(-u\right) \, du \\
&= \frac{1}{4}\int_0^\infty \frac{\pi}{16}u^4 \exp(-u)du \\
&= \frac{3\pi}{8}, \tag{59}
\end{aligned}
$$

where (a) follows by a change of variable $u = (\lambda_1 + \lambda_2)/2$ and $v = (\lambda_1 - \lambda_2)/2$. In conclusion, $\mathsf{E}[|\det(\mathbf{H})|] = \frac{3\pi}{8}$ for i.i.d. Rayleigh fading.

APPENDIX III
APPROXIMATE ERGODIC CAPACITY FOR FADING INTERFERENCE CHANNEL

A similar analysis used in Theorems 2 and 3 is applicable for fading interference channel. Specifically, consider the K-user interference channel in which the input–output relation is given by

$$\mathbf{y}[t] = \mathbf{H}[t]\mathbf{x}[t] + \mathbf{z}[t] \tag{60}$$

and the elements of $\mathbf{H}[t] = \{h_{i,j}[t]\}$ are i.i.d. drawn from a continuous distribution $f(x)$, $x \in \mathbb{C}$, and vary independently over time. The achievability follows from ergodic interference alignment in [3] showing that

$$R_i = \frac{1}{2}\mathsf{E}[\log(1 + 2|h_{i,i}|^2 P)] \tag{61}$$

is achievable for all $i \in \{1, \cdots, K\}$ [3, Theorem 2]. Theorem 6 characterizes an approximate ergodic per-user capacity, i.e., ergodic sum capacity divided by K, assuming that all sources employ uniform power allocation across time. For this case, the sum of any pair of achievable rates is upper bounded by

$$R_i + R_j \le \mathsf{E}\left[\log\left(1 + \frac{(|h_{i,j}|^2 + |h_{i,i}|^2)P}{\min\left\{1, \frac{|h_{i,j}|^2}{|h_{j,j}|^2}\right\}}\right)\right] \tag{62}$$

for all $i,j \in \{1, \cdots, K\}$, $i \ne j$ [3, Equation (99)]. From the lower bound (61) and the upper bound (62), we characterize an approximate ergodic per-user capacity in the following theorem.

Theorem 6: Consider the fading K-user interference channel. Let $R_{\mathrm{ia}} := \sum_{i=1}^{K} \frac{1}{2}\mathsf{E}[\log(1 + 2|h_{i,i}|^2 P)]$ and C_{sum} denote the sum capacity assuming that all sources employ uniform power allocation across time. Then

$$\frac{C_{\mathrm{sum}} - R_{\mathrm{ia}}}{K} \le \frac{1}{2}\log\left(\frac{3}{2}\right) + \frac{1}{2}\mathsf{E}\left[\left|\log\left(\frac{|h_{1,1}|^2}{|h_{1,2}|^2}\right)\right|\right] \tag{63}$$

for any $P > 0$.

Proof: Define

$$\Delta(|h_{i,i}|^2, |h_{j,j}|^2, |h_{i,j}|^2)$$
$$:= \log\left(1 + \frac{(|h_{i,j}|^2 + |h_{i,i}|^2)P}{\min\left\{1, \frac{|h_{i,j}|^2}{|h_{j,j}|^2}\right\}}\right) - \frac{1}{2}\log(1 + 2|h_{i,i}|^2 P) - \frac{1}{2}\log(1 + 2|h_{j,j}|^2 P). \tag{64}$$

Then, from (61) and (62) and the fact that channel coefficients are i.i.d.,

$$\frac{C_{\mathrm{sum}} - R_{\mathrm{ia}}}{K} \le \frac{1}{2}\mathsf{E}[\Delta(|h_{1,1}|^2, |h_{2,2}|^2, |h_{1,2}|^2)]. \tag{65}$$

The term $\Delta(|h_{1,1}|^2, |h_{2,2}|^2, |h_{1,2}|^2)$ can be expressed as

$$\Delta(|h_{1,1}|^2, |h_{2,2}|^2, |h_{1,2}|^2)$$
$$= \log\left(1 + \max\left\{1, \frac{|h_{2,2}|^2}{|h_{1,2}|^2}\right\}(|h_{1,2}|^2 + |h_{1,1}|^2)P\right) - \frac{1}{2}\log(1 + 2|h_{1,1}|^2 P) - \frac{1}{2}\log(1 + 2|h_{2,2}|^2 P)$$
$$= \frac{1}{2}\log\left(\frac{1 + \max\left\{1, \frac{|h_{2,2}|^2}{|h_{1,2}|^2}\right\}(|h_{1,2}|^2 + |h_{1,1}|^2)P}{1 + 2|h_{1,1}|^2 P}\right) + \frac{1}{2}\log\left(\frac{1 + \max\left\{1, \frac{|h_{2,2}|^2}{|h_{1,2}|^2}\right\}(|h_{1,2}|^2 + |h_{1,1}|^2)P}{1 + 2|h_{2,2}|^2 P}\right). \tag{66}$$

The first term of (66) is upper bounded as

$$\frac{1}{2}\log\left(\frac{1 + \max\left\{1, \frac{|h_{2,2}|^2}{|h_{1,2}|^2}\right\}(|h_{1,2}|^2 + |h_{1,1}|^2)P}{1 + 2|h_{1,1}|^2 P}\right)$$
$$\le \frac{1}{2}\log\left(\frac{1 + \max\left\{1, \frac{|h_{2,2}|^2}{|h_{1,2}|^2}\right\}|h_{1,2}|^2 P + \max\left\{1, \frac{|h_{2,2}|^2}{|h_{1,2}|^2}\right\}2|h_{1,1}|^2 P}{1 + 2|h_{1,1}|^2 P}\right)$$
$$\le \frac{1}{2}\log\left(\max\left\{1, \frac{|h_{2,2}|^2}{|h_{1,2}|^2}\right\}\frac{|h_{1,2}|^2}{2|h_{1,1}|^2} + \max\left\{1, \frac{|h_{2,2}|^2}{|h_{1,2}|^2}\right\}\right)$$
$$= \frac{1}{2}\log\left(\max\left\{1, \frac{|h_{2,2}|^2}{|h_{1,2}|^2}\right\}\right) + \frac{1}{2}\log\left(1 + \frac{|h_{1,2}|^2}{2|h_{1,1}|^2}\right). \tag{67}$$

Similarly, the second term of (66) is upper bounded as

$$\frac{1}{2}\log\left(\frac{1+\max\left\{1,\frac{|h_{2,2}|^2}{|h_{1,2}|^2}\right\}(|h_{1,2}|^2+|h_{1,1}|^2)P}{1+2|h_{2,2}|^2 P}\right)$$

$$\leq \frac{1}{2}\log\left(\frac{1+\max\left\{1,\frac{|h_{1,2}|^2}{|h_{2,2}|^2}\right\}2|h_{2,2}|^2 P+\max\left\{1,\frac{|h_{2,2}|^2}{|h_{1,2}|^2}\right\}|h_{1,1}|^2 P}{1+2|h_{2,2}|^2 P}\right)$$

$$\leq \frac{1}{2}\log\left(\max\left\{1,\frac{|h_{1,2}|^2}{|h_{2,2}|^2}\right\}+\max\left\{1,\frac{|h_{2,2}|^2}{|h_{1,2}|^2}\right\}\frac{|h_{1,1}|^2}{2|h_{2,2}|^2}\right)$$

$$\stackrel{(a)}{=}\frac{1}{2}\log\left(\max\left\{1,\frac{|h_{1,2}|^2}{|h_{2,2}|^2}\right\}+\max\left\{1,\frac{|h_{1,2}|^2}{|h_{2,2}|^2}\right\}\frac{|h_{1,1}|^2}{2|h_{1,2}|^2}\right)$$

$$=\frac{1}{2}\log\left(\max\left\{1,\frac{|h_{1,2}|^2}{|h_{2,2}|^2}\right\}\right)+\frac{1}{2}\log\left(1+\frac{|h_{1,1}|^2}{2|h_{1,2}|^2}\right), \tag{68}$$

where (a) follows since $a\max\{1,b/a\}=b\max\{1,a/b\}$ for all $a,b>0$. Therefore,

$$\Delta(|h_{1,1}|^2,|h_{2,2}|^2,|h_{1,2}|^2)$$

$$\leq \frac{1}{2}\log\left(\max\left\{1,\frac{|h_{2,2}|^2}{|h_{1,2}|^2}\right\}\right)+\frac{1}{2}\log\left(1+\frac{|h_{1,2}|^2}{2|h_{1,1}|^2}\right)$$

$$+\frac{1}{2}\log\left(\max\left\{1,\frac{|h_{1,2}|^2}{|h_{2,2}|^2}\right\}\right)+\frac{1}{2}\log\left(1+\frac{|h_{1,1}|^2}{2|h_{1,2}|^2}\right)$$

$$=\frac{1}{2}\left|\log\left(\frac{|h_{2,2}|^2}{|h_{1,2}|^2}\right)\right|+\frac{1}{2}\log\left(1+\frac{|h_{1,2}|^2}{2|h_{1,1}|^2}\right)+\frac{1}{2}\log\left(1+\frac{|h_{1,1}|^2}{2|h_{1,2}|^2}\right)$$

$$\leq \frac{1}{2}\left|\log\left(\frac{|h_{2,2}|^2}{|h_{1,2}|^2}\right)\right|+\frac{1}{2}\log\left(1+\frac{1}{2}\right)+\frac{1}{2}\log\left(1+\frac{1}{2}\max\left\{\frac{|h_{1,2}|^2}{|h_{1,1}|^2},\frac{|h_{1,1}|^2}{|h_{1,2}|^2}\right\}\right)$$

$$\leq \frac{1}{2}\log\left(\frac{3}{2}\right)+\frac{1}{2}\left|\log\left(\frac{|h_{2,2}|^2}{|h_{1,2}|^2}\right)\right|+\frac{1}{2}\log\left(\frac{3}{2}\max\left\{\frac{|h_{1,2}|^2}{|h_{1,1}|^2},\frac{|h_{1,1}|^2}{|h_{1,2}|^2}\right\}\right)$$

$$=\log\left(\frac{3}{2}\right)+\frac{1}{2}\left|\log\left(\frac{|h_{2,2}|^2}{|h_{1,2}|^2}\right)\right|+\frac{1}{2}\left|\log\left(\frac{|h_{1,1}|^2}{|h_{1,2}|^2}\right)\right|. \tag{69}$$

Finally, combining (65) and (69) shows the gap in (63), which completes the proof. ∎

For i.i.d. Rayleigh fading channel, $|h_{i,j}|^2$ has the exponential distribution and

$$f_{|h_{1,1}|^2/|h_{1,2}|^2}(x)=\frac{1}{(x+1)^2} \tag{70}$$

for $x\geq 0$. Therefore,

$$\mathsf{E}\left[\left|\log\left(\frac{|h_{1,1}|^2}{|h_{1,2}|^2}\right)\right|\right]=\int_0^\infty \frac{|\log x|}{(x+1)^2}\,dx=2$$

and the gap in Theorem 6 is given as $\frac{1}{2}\log 6$ bits/sec/Hz (approximately 1.3 bits/sec/Hz).

REFERENCES

[1] R. H. Etkin, D. N. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, pp. 5534–5562, Dec. 2008.
[2] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, pp. 3425–3441, Aug. 2008.
[3] B. Nazer, M. Gastpar, S. A. Jafar, and S. Vishwanath, "Ergodic interference alignment," *IEEE Trans. Inf. Theory*, vol. 58, pp. 6355–6371, Oct. 2012.
[4] S. Mohajer, S. Diggavi, C. Fragouli, and D. N. C. Tse, "Approximate capacity of a class of Gaussian interference-relay networks," *IEEE Trans. Inf. Theory*, vol. 57, pp. 2837–2864, May 2011.
[5] T. Gou, S. A. Jafar, C. Wang, S.-W. Jeon, and S.-Y. Chung, "Aligned interference neutralization and the degrees of freedom of the $2\times 2\times 2$ interference channel," *IEEE Trans. Inf. Theory*, vol. 58, pp. 4381–4395, Jul. 2012.

[6] S.-W. Jeon, S.-Y. Chung, and S. A. Jafar, "Degrees of freedom region of a class of multisource Gaussian relay networks," *IEEE Trans. Inf. Theory*, vol. 57, pp. 3032–3044, May 2011.

[7] C. Wang, T. Gou, and S. A. Jafar, "Multiple unicast capacity of 2-source 2-sink networks," in *arXiv:cs.IT/1104.0954*, Apr. 2011.

[8] I. Shomorony and A. S. Avestimehr, "Two-unicast wireless networks: Characterizing the degrees-of-freedom," in *arXiv:cs.IT/1102.2498*, Feb. 2011.

[9] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 379–389, Feb. 2007.

[10] V. R. Cadambe and S. A. Jafar, "Interference alignment and the degrees of freedom of wireless X networks," *IEEE Trans. Inf. Theory*, vol. 55, pp. 3893–3908, Sep. 2009.

[11] B. Nazer, M. Gastpar, S. A. Jafar, and S. Vishwanath, "Ergodic interference alignment," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Seoul, South Korea, Jun./Jul. 2009.

[12] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. Inf. Theory*, vol. 54, pp. 3457–3470, Aug. 2008.

[13] T. Gou and S. A. Jafar, "Degrees of freedom of the K user $M \times N$ MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 56, pp. 6040–6057, Dec. 2010.

[14] C. Suh, M. Ho, and D. N. C. Tse, "Downlink interference alignment," *IEEE Trans. Commun.*, vol. 59, pp. 2616–2626, Sep. 2011.

[15] C. Suh and D. N. C. Tse, "Interference alignment for cellular networks," in *Proc. 46th Annu. Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Sep. 2008.

[16] V. R. Cadambe and S. A. Jafar, "Degrees of freedom of wireless networks with relays, feedback, cooperation, and full duplex operation," *IEEE Trans. Inf. Theory*, vol. 55, pp. 2334–2344, May 2009.

[17] V. S. Annapureddy, A. El Gamal, and V. V. Veeravalli, "Degrees of freedom of interference channels with CoMP transmission and reception," *IEEE Trans. Inf. Theory*, vol. 58, pp. 5740–5760, Sep. 2012.

[18] L. Ke, A. Ramamoorthy, Z. Wang, and H. Yin, "Degrees of freedom region for an interference network with general message demands," *IEEE Trans. Inf. Theory*, vol. 58, pp. 3787–3797, Jun. 2012.

[19] A. S. Motahari, S. O. Gharan, and A. K. Khandani, "Real interference alignment with real numbers," in *arXiv:cs.IT/0908.1208*, 2009.

[20] A. S. Motahari, S. O. Gharan, M. A. Maddah-Ali, and A. K. Khandani, "Real interference alignment: Exploiting the potential of single antenna systems," in *arXiv:cs.IT/0908.2282*, 2009.

[21] G. Bresler, A. Parekh, and D. N. C. Tse, "The approximate capacity of the many-to-one and one-to-many Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 56, pp. 4566–4592, Sep. 2010.

[22] W. Nam, S.-Y. Chung, and Y. H. Lee, "Capacity of the Gaussian two-way relay channel to within $\frac{1}{2}$ bit," *IEEE Trans. Inf. Theory*, vol. 56, pp. 5488–5494, Nov. 2010.

[23] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory*, vol. 57, pp. 1872–1905, Apr. 2011.

[24] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," *IEEE Trans. Inf. Theory*, vol. 57, pp. 3132 – 3152, May 2011.

[25] U. Niesen and M. A. Maddah-Ali, "Interference alignment: From degrees-of-freedom to constant-gap capacity approximations," in *arXiv:cs.IT/1112.4879*, 2011.

[26] O. Ordentlich, U. Erez, and B. Nazer, "The approximate sum capacity of the symmetric Gaussian k-user interference channel," in *arXiv:cs.IT/1206.0197*, 2012.

[27] S.-W. Jeon and S.-Y. Chung, "Capacity of a class of multi-source relay networks," in *Information Theory and Applications Workshop*, University of California San Diego, La Jolla , CA, Feb. 2009.

[28] U. Niesen, B. Nazer, and P. Whiting, "Computation alignment: Capacity approximation without noise accumulation," in *arXiv:cs.IT/1108.6312*, 2011.

[29] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.

[30] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic Press, 1981.

[31] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*. Academic Press, 2007.

[32] C. Esli, S. Berger, and A. Wittneben, "Optimizing zero-forcing based gain allocation for wireless multiuser networks," in *Proc. IEEE International Conference on Communications (ICC)*, Beijing, China, Jun. 2007.

[33] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge Univ. Press, 2004.

[34] A. Edelman, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, Massachusetts Institute of Technology (MIT), 1989.