On the DMT Optimality of Time-Varying Distributed Rotation Over Slow Fading Relay Channels

Ramtin Pedarsani, Olivier Lévêque, Sheng Yang

On the DMT Optimality of Time-Varying Distributed Rotation over Slow Fading Relay Channels

Ramtin Pedarsani, Olivier Lévêque, *Member, IEEE*, Sheng Yang, *Member, IEEE*

***Abstract*—We consider a slow fading two-hop relay channel where a source terminal communicates with a destination through a layer of relays without a direct link. First, we introduce the notion of time-varying distributed rotation and propose a linear relaying scheme called rotate-and-forward (RF). The main idea is to create a time-varying channel and to convert the spatial diversity to time diversity. It is shown that this scheme achieves the optimal diversity-multiplexing tradeoff (DMT) of the channel with full-duplex relays. While more involved non-linear relaying schemes previously proposed in the literature are optimal in the same setting, we show here that simple linear relaying can also be DMT optimal. Then, we extend the RF scheme to the relay channel with multiple hops where the DMT optimality of the two-antenna case is shown. Finally, we apply the idea of distributed rotation to the decode-and-forward relays. Same diversity order as previous schemes can be achieved with low signaling complexity.**

I. Introduction

The gain of using multiple antennas for setting up communication over a wireless medium has been widely acknowledged in the literature, starting with the seminal works [1], [2]. For point-to-point channels, the performance of multiple antenna systems is quite well understood by now. In particular, the optimal tradeoff between reliability and rate (also known as diversity-multiplexing tradeoff or DMT) of such systems at high signal-to-noise ratio (SNR) was analyzed in detail in [3].

In recent years, there has been a surge of interest in cooperative diversity techniques, where spatial diversity is exploited with distributed relay antennas. Many different schemes have been proposed to improve the diversity of the channel (see, e.g., [4]–[15] and references therein). Essentially, these schemes can be divided into two categories, namely, linear and nonlinear relaying schemes. Based on message decoding (e.g., decode-and-forward) or signal compression (e.g., compress-and-forward) or a mixture of both at the relays, nonlinear relaying schemes are "intelligent" and can usually outperform the "dumb" linear relaying schemes (e.g., amplify-and-forward) where relays only forward linear combinations of individual observations. In a multi-hop MIMO relay network, the role of the relays is two-fold: to provide diversity gain with independent paths and multiplexing gains with "antenna pooling". While the traditional decode-and-forward scheme performs well in networks with single-antenna nodes ([16], [17]), it can suffer from significant loss of multiplexing gain: requiring each relay node to decode the entire source message jeopardizes the degrees of freedom of the network. This is due to the impossibility of splitting the source message into separated parts for different relays without channel state information (CSI) at the source terminal. On the other hand, with amplify-and-forward, the relays send the analog observation without decoding nor encoding. It has been shown that even this naive amplify-and-forward scheme is optimal in terms of multiplexing gain in the high SNR regime [18]. However, as pointed out in [14], the amplify-and-forward operation correlates the source-destination paths and penalizes the diversity gain. Compress-and-forward is a solution in between the above two. In this case, each relay encodes the digitized (quantized) observation and send it to the destination. The latter tries to recover the source message from the received signal. It has been shown in [19], [20] that relaying schemes of this kind (e.g., quantize-map-and-forward in [19] and noisy network coding in [20]), achieves any rate within a constant number of bits to the capacity of multi-hop relay networks. As a result, these schemes attain the optimal DMT of such networks [21]. While it appears that the multihop relaying problem has been solved with nonlinear relaying (in the high SNR regime) at this point, it is still unknown whether the same can be done with linear relaying.

Unlike nonlinear relaying, linear schemes are appealing for their low signaling and computational complexity as well as their scalability in terms of practical implementation. More importantly, it has been shown that they can also be DMT optimal in some non-trivial settings [11]. It is worth mentioning that a linear scheme, called the *flip-and-forward scheme*, was proposed in [14] and shown to achieve the maximum diversity as well as the maximum multiplexing gain for any number of antennas and hops.

As space-time codes exploit the spatial diversity in a multiple-antenna (MIMO) system, cooperative diversity can be achieved with distributed space-time code/processing [6], [10], [13]. With linear relaying, the relays perform a linear processing on the received signal in a coordinated manner and forward it. In a nutshell, while the signal is naturally mixed in space before arriving at the relays, it is then artificially transformed in the time domain by the relays in such a way to mimic the space-time codes. In most cases, the relaying creates

R. Pedarsani is with the department of EECS, UC Berkeley, USA. O. Lévêque is with the Swiss Federal Institute of Technology - Lausanne, S. Yang is with SUPELEC, France (e-mails: ramtin@eecs.berkeley.edu, olivier.leveque@epfl.ch, sheng.yang@supelec.fr). The material in this paper was presented in part at the International Symposium on Information Theory, Austin, TX, July 2010, and the 48th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, September 2010.

a "good" equivalent channel with high diversity [10], [14]. Although it is rather straightforward to conceive a scheme with maximum diversity and/or maximum multiplexing gain, designing an optimal scheme in terms of the fundamental tradeoff between these two remains challenging. As a matter of fact, the equivalent channel with distributed space-time processing can be quite involved, if not intractable.

In this work, we propose a conceptually different framework to exploit cooperative diversity. In this framework, the relays do not perform any temporal transformation that is the key ingredient for traditional cooperative diversity schemes. Instead, with the *time-varying* scalar rotation based on the so-called *distributed rotation sequences*, an artificial time-varying channel is created to recover the spatial diversity. It turns out that the proposed framework is both tractable from the theoretical point of view and simple from the practical point of view. As a main result of this paper, a linear relaying scheme called *rotate-and-forward* is proposed and shown to be DMT optimal for two-hop relay channels with arbitrary number of source/relays/destination antennas. Moreover, we show that this scheme is also DMT optimal in some cases in the multi-hop setting. Finally, we apply the idea of distributed rotation to decode-and-forward relays and show that equivalent performance as the existing schemes can be achieved with much lower relaying complexity.

The rest of the paper is organized as follows. The system models and some basic assumptions are presented in Section II. The main results on the rotate-and-forward scheme are presented in Section III. The two main theorems of the paper are proved separately in Sections IV and V. Then, the RF scheme is extended to the multi-hop case and the optimality is shown for certain settings. As another application of the distributed rotation, a variant of the decode-and-forward scheme is introduced and analyzed in Section VII. Finally, the paper is concluded in Section VIII. Some of the proofs are deferred to the appendix to make the reading fluid.

II. System Model and Assumptions

Throughout the paper, we will use the following notations. Boldface lower-case letters $\boldsymbol{v}$ and upper-case letters $\boldsymbol{M}$ are used to denote vectors and matrices, respectively. Unless otherwise is specified, vectors are column vectors. Matrix transpose, Hermitian transpose, inverse, trace, and determinant are denoted by $\boldsymbol{A}^T$, $\boldsymbol{A}^*$, $\boldsymbol{A}^{-1}$, $\operatorname{Tr}(\boldsymbol{A})$, and $\det(\boldsymbol{A})$, respectively. We let $\boldsymbol{A}_{\mathcal{I},\mathcal{J}}$ denote the submatrix of $\boldsymbol{A}$ as $[A_{ij}]_{i\in\mathcal{I},j\in\mathcal{J}}$ and $\boldsymbol{A}_{\mathcal{I}}$ denote the submatrix of $\boldsymbol{A}$ as $[A_{ij}]_{i,j\in\mathcal{I}}$. We also define $\det(\boldsymbol{A})_{\mathcal{I}} = \det(\boldsymbol{A}_{\mathcal{I}})$ for any non-empty set $\mathcal{I}$. When $\mathcal{I}$ is empty, we define $\det(\boldsymbol{A})_{\emptyset} = \det(\boldsymbol{A}_{\emptyset}) = 1$ for notational convenience. Finally, the dot-equality $\doteq$ means the equality of the SNR exponent at high SNR, i.e., $f \doteq g$ means $\lim_{\mathsf{snr}\to\infty} \frac{\log(f)}{\log \mathsf{snr}} = \lim_{\mathsf{snr}\to\infty} \frac{\log(g)}{\log \mathsf{snr}}$.

A. Signal Model

We consider a slow fading wireless channel with one source, one destination, and n relays. It is assumed that the source and destination are equipped with m and p antennas, respectively, while each of the relays has only one antenna, as shown in Fig. 1. For simplicity, we call it a (m, n, p) relay channel. In this work, we focus on distributed relaying schemes. By distributed relaying, we mean that no information on the message or channel state information (CSI) is exchanged between the relays. All terminals in the channel have perfect receiver CSI and no transmitter CSI at all. Furthermore, we assume that the relays only receive signal from the source and the destination only receives signal from the relays, i.e., there is no direct link from the source to the destination. Unless it is otherwise specified, we also assume that the n relays work in full-duplex mode, i.e., they transmit and receive simultaneously at any instant t. Finally, all terminals work with perfect synchronization.

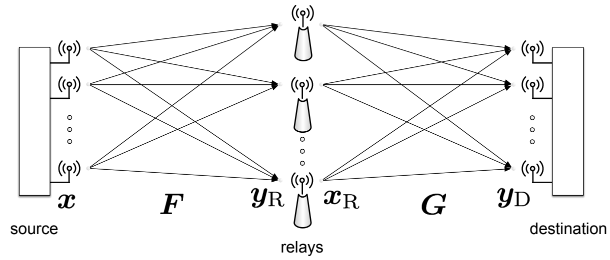

Fig. 1. The two-hop layered relay channel model.

The signal model can be described as follows

$$\begin{aligned} \boldsymbol{y}_{\mathrm{R}}[t] &= \boldsymbol{F}\,\boldsymbol{x}[t] + \boldsymbol{z}_{\mathrm{R}}[t], \\ \boldsymbol{y}_{\mathrm{D}}[t] &= \boldsymbol{G}\,\boldsymbol{x}_{\mathrm{R}}[t] + \boldsymbol{z}_{\mathrm{D}}[t], \quad t = 1, 2, \ldots \end{aligned}$$

where $\boldsymbol{x} \in \mathbb{C}^{m\times 1}$, $\boldsymbol{y}_{\mathrm{R}} \in \mathbb{C}^{n\times 1}$ and $\boldsymbol{y}_{\mathrm{D}} \in \mathbb{C}^{p\times 1}$ are the transmitted signal from the source, received signal at the relays, and received signal at the destination, respectively; $\boldsymbol{z}_{\mathrm{R}} \in \mathbb{C}^{n\times 1}$ and $\boldsymbol{z}_{\mathrm{D}} \in \mathbb{C}^{p\times 1}$ are the additive white Gaussian noise with independent and identically distributed (i.i.d.) $\mathcal{N}_{\mathbb{C}}(0, \sigma^2)$ entries at the relay and the destination, respectively; $\boldsymbol{F} \in \mathbb{C}^{n\times m}$ and $\boldsymbol{G} \in \mathbb{C}^{p\times n}$ are channel matrices defining the source-relays and relays-destination channels, respectively; for simplicity, the transmit power at the source and the relay is subject to the following average per-antenna constraints P, i.e.,

$$\begin{aligned} \mathbb{E}\left(\|\boldsymbol{x}[t]\|^2\right) &\le mP, \\ \mathbb{E}\left(\|\boldsymbol{x}_{\mathrm{R}}[t]\|^2\right) &\le nP, \end{aligned}$$

at any time instant t. Note that the above expectations are taken over all random factors including the message, channel coefficients, and noise. From now on, we define the signal-to-noise ratio as $\mathsf{snr} \triangleq P/\sigma^2$.

B. Diversity-Multiplexing Tradeoff

In this work, we are interested in the high SNR performance of this system. In order to evaluate the performance of the relaying schemes, we use the diversity-multiplexing tradeoff (DMT) [3] to characterize the fundamental interplay between reliability and throughput in the high signal-to-noise ratio (SNR) regime. A relaying scheme is said to achieve *multiplexing gain* r and *diversity gain* d if

$$d(r) = \lim_{\mathsf{snr}\to\infty} -\frac{\log(\mathbb{P}_{\mathrm{out}}(r \log \mathsf{snr}))}{\log \mathsf{snr}}$$

where $\mathbb{P}_{\text{out}}(r \log \mathsf{snr})$ denotes the outage probability[1], that is, the probability that the mutual information between the source and the destination is lower than the target rate $R = r \log \mathsf{snr}$. Alternatively, we can use the dot-equality expression

$$\mathbb{P}_{\text{out}}(r \log \mathsf{snr}) \doteq \mathsf{snr}^{-d(r)}. \tag{1}$$

Note that in order to remove the dependency on any particular coding scheme, instead of using the error probability, we use directly the outage probability to characterize the achievable DMT. This choice is justified given that we can always use an outage-optimal or DMT-achieving code for a particular relaying scheme (e.g., random [3] or universal coding [22]).

C. Main Ingredient: Distributed Rotation Sequence

Let us first define

$$\mathcal{U} \triangleq [0, 1).$$

Then, we define a set of K equally spaced values in $\mathcal{U}$ and the corresponding set of complex rotations

$$\mathcal{A}_K \triangleq \left\{0, \frac{1}{K}, \ldots, \frac{K-1}{K}\right\},$$
$$\mathcal{R}_K \triangleq \left\{e^{j2\pi\varphi} : \ \varphi \in \mathcal{A}_K\right\}.$$

A distributed rotation sequence is defined as follows.

Definition 1 (Distributed Rotation Sequence): A sequence of diagonal matrices $\{\boldsymbol{\Delta}_t\}$, $t = 1, \ldots, K^n$, is said to be a *distributed rotation sequence (DRS)* if

1) $\boldsymbol{\Delta}_t = \operatorname{diag}\left(e^{j2\pi\varphi_{1,t}}, \ldots, e^{j2\pi\varphi_{n,t}}\right)$ with $\varphi_{i,t} \in \mathcal{A}_K$ for $i = 1, \ldots n$,
2) $\boldsymbol{\Delta}_t \neq \boldsymbol{\Delta}_{t'}$, $\forall\, t \neq t'$.

In other words, any sequence of length K^n that runs through all the possibilities defined by $\mathcal{R}_K^{n\times 1}$ is a DRS. As we will show in the next sections, fixed DRS (known at all terminals) is used by the relays to create time-varying channels.

III. Distributed Linear Relaying: Rotate-and-Forward and DMT Optimality

A. Protocol description

In this section, we consider *distributed* linear relaying schemes, i.e.,

$$\boldsymbol{x}_{\text{R}}[t] = \boldsymbol{D}[t]\, \boldsymbol{y}_{\text{R}}[t-1]$$

where $\boldsymbol{x}_{\text{R}} \in \mathbb{C}^{n\times 1}$ is transmitted signal from the relays; $\boldsymbol{D}$ is a *diagonal* matrix. Note that if the relaying matrix $\boldsymbol{D}[t]$ does not vary with t, the above signal model is reduced to a naive antenna-wise amplify-and-forward scheme. For simplicity, we assume that $\boldsymbol{F}$ satisfies $\mathbb{E}\left\{|f_{ij}|^2\right\} = 1$, $\forall\, i, j$. Furthermore, we impose $|D_{ii}|^2 = \dfrac{\mathsf{snr}}{n\,\mathsf{snr}+1}$, i.e., constant relaying gain. The proposed *rotate-and-forward* (RF) scheme is based on a fixed DRS $\{\boldsymbol{\Delta}_t\}$ and works as follows. A codeword $\boldsymbol{X} \triangleq \begin{bmatrix} \boldsymbol{x}[1] & \cdots & \boldsymbol{x}[T] \end{bmatrix} \in \mathbb{C}^{m\times T}$ spanning over T symbols time is transmitted by the source with $T = K^n$. At instant t, each relay transmits a rotated version of what it received at instant $t-1$. The rotation used by the relay i is $e^{j2\pi\varphi_{i,t}}$ as in Definition 1. Note that this is equivalent to defining $\boldsymbol{D}[t] = \sqrt{\dfrac{\mathsf{snr}}{n\,\mathsf{snr}+1}}\, \boldsymbol{\Delta}_t$. Thus, we have, by defining $c \triangleq \sqrt{\dfrac{\mathsf{snr}}{n\,\mathsf{snr}+1}}$,

$$\boldsymbol{y}_{\text{D}}[t+1] = c\,\boldsymbol{G}\boldsymbol{\Delta}_t\boldsymbol{F}\boldsymbol{x}[t] + c\,\boldsymbol{G}\boldsymbol{\Delta}_t\boldsymbol{z}_{\text{R}}[t] + \boldsymbol{z}_{\text{D}}[t+1]$$

for $t = 1, \ldots, T$. Hence, the transmitted codeword $\boldsymbol{X}$ goes through an equivalent time-varying fading channel with channel matrix $c\,\boldsymbol{G}\boldsymbol{\Delta}_t\boldsymbol{F}$ and equivalent noise covariance $\sigma^2(\mathbf{I} + c^2\boldsymbol{G}\boldsymbol{\Delta}_t\boldsymbol{\Delta}_t^*\boldsymbol{G}^*) = \sigma^2(\mathbf{I} + c^2\boldsymbol{G}\boldsymbol{G}^*)$. It is important to note that, since $\boldsymbol{z}_{\text{R}}[t]$ is i.i.d. over time and circularly symmetric, $\boldsymbol{\Delta}_t\boldsymbol{z}_{\text{R}}[t]$ and therefore the equivalent noise $c\,\boldsymbol{G}\boldsymbol{\Delta}_t\boldsymbol{z}_{\text{R}}[t] + \boldsymbol{z}_{\text{D}}[t+1]$ are independent of $\boldsymbol{\Delta}_t$ and i.i.d. over time. It is thus without loss of generality to rewrite the signal model as

$$\boldsymbol{y}_{\text{D}}[t+1] = c\,\boldsymbol{G}\boldsymbol{\Delta}_t\boldsymbol{F}\boldsymbol{x}[t] + \boldsymbol{z}[t]$$

where $\boldsymbol{z}[t] \sim \mathcal{N}_{\mathbb{C}}(0, \boldsymbol{\Sigma}_{\boldsymbol{z}})$ with $\boldsymbol{\Sigma}_{\boldsymbol{z}} \triangleq \sigma^2(\mathbf{I} + c^2\boldsymbol{G}\boldsymbol{G}^*)$.

B. Outage Analysis

Since $c^2 \leq 1$ and $\boldsymbol{G}\boldsymbol{G}^* \preceq \|\boldsymbol{G}\|_{\text{F}}^2\,\mathbf{I}$, we have $\sigma^2(1 + \|\boldsymbol{G}\|_{\text{F}}^2)\,\mathbf{I} \succeq \boldsymbol{\Sigma}_{\boldsymbol{z}} \succeq \sigma^2\mathbf{I}$. As in most works on high SNR analysis in the literature, we only consider channel distributions such that the density function has exponentially decaying tail, which implies that there exists some $\mu > 0$ such that, for any $\epsilon > 0$, $\mathbb{P}\left(\|\boldsymbol{G}\|_{\text{F}}^2 > \mathsf{snr}^{\epsilon}\right) = O\big(\exp(-\mu\,\mathsf{snr}^{\epsilon})\big)$ when $\mathsf{snr} \to \infty$. For instance, with i.i.d Rayleigh fading, it is readily shown that $\mu = \frac{1}{2}$. Now, let us define the event $\mathcal{B}_\epsilon \triangleq \{\|\boldsymbol{G}\|_{\text{F}}^2 > \mathsf{snr}^{\epsilon}\}$ and partition the outage event $\mathcal{O}$ as $\mathcal{O} = (\mathcal{O}\cap\mathcal{B}_\epsilon) \cup (\mathcal{O}\cap\bar{\mathcal{B}}_\epsilon)$. Then, the outage probability is written as

$$\begin{aligned}
&\mathbb{P}\left(\mathcal{O}\cap\mathcal{B}_\epsilon\right) + \mathbb{P}\left(\mathcal{O}\cap\bar{\mathcal{B}}_\epsilon\right) \\
&\quad\leq \mathbb{P}\left(\mathcal{B}_\epsilon\right) + \mathbb{P}\left(\mathcal{O}\cap\{(1+\mathsf{snr}^{\epsilon})\sigma^2\mathbf{I} \succeq \boldsymbol{\Sigma}_{\boldsymbol{z}} \succeq \sigma^2\mathbf{I}\}\right).
\end{aligned}$$

Since $\mathbb{P}\left(\mathcal{B}_\epsilon\right)$ decays exponentially with snr for any $\epsilon > 0$, i.e., $\mathbb{P}\left(\mathcal{B}_\epsilon\right) \doteq \mathsf{snr}^{-\infty}$, the outage probability is dominated by $\mathbb{P}\left(\mathcal{O}\cap\{(1+\mathsf{snr}^{\epsilon})\sigma^2\mathbf{I} \succeq \boldsymbol{\Sigma}_{\boldsymbol{z}} \succeq \sigma^2\mathbf{I}\}\right)$ in which we can make ϵ as close to 0 as possible. In other words, we can assume, without loss of generality, that $\boldsymbol{\Sigma}_{\boldsymbol{z}} = \sigma^2\mathbf{I}$, without causing any impact on the SNR exponent of the outage probability. Consequently, we can ignore the exact noise covariance, as far as the diversity-multiplexing tradeoff is concerned [3]. Here, the DMT of the proposed protocol depends uniquely on the time-varying equivalent channel matrix $\tilde{\boldsymbol{H}}[t] \triangleq \boldsymbol{G}\boldsymbol{\Delta}_t\boldsymbol{F}$ where the factor c is omitted for the same reason. We are now interested in the following average mutual information in bits per channel use

$$I_T(\mathsf{snr}) \triangleq \frac{1}{T}\sum_{t=1}^{T} \log\det\Big(\mathbf{I} + \mathsf{snr}\tilde{\boldsymbol{H}}[t]\tilde{\boldsymbol{H}}[t]^*\Big)$$

[1]With a slight abuse of terminology, the outage probability of a scheme means the outage probability of the equivalent channel created by the relaying scheme.

where we recall that $T = K^n$. Then, we can get the following chain of equalities

$$I_T(\mathsf{snr}) = \frac{1}{T}\sum_{t=1}^{T} \log\det\left(\mathbf{I} + \mathsf{snr}\boldsymbol{\Delta}_t \boldsymbol{F}\boldsymbol{F}^* \boldsymbol{\Delta}_t^* \boldsymbol{G}^*\boldsymbol{G}\right) \quad (2)$$

$$= \frac{1}{T}\sum_{t=1}^{T} \log\det\left(\mathbf{I} + \mathsf{snr}\boldsymbol{\Delta}_t \boldsymbol{P} \boldsymbol{\Delta}_t^* \boldsymbol{Q}\right)$$

$$= \frac{1}{K^n}\sum_{\theta_n\in\mathcal{A}_K}\cdots\sum_{\theta_1\in\mathcal{A}_K} \log\det\left(\mathbf{I} + \mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}} \boldsymbol{P} \boldsymbol{R}_{\boldsymbol{\theta}}^* \boldsymbol{Q}\right) \quad (3)$$

where (2) follows from Sylvester's identity $\log\det(\mathbf{I}+\boldsymbol{A}\boldsymbol{B}) = \log\det(\mathbf{I}+\boldsymbol{B}\boldsymbol{A})$, and we define $\boldsymbol{P} \triangleq \boldsymbol{F}\boldsymbol{F}^*$, $\boldsymbol{Q} \triangleq \boldsymbol{G}^*\boldsymbol{G}$, and $\boldsymbol{R}_{\boldsymbol{\theta}} \triangleq \mathrm{diag}(e^{j2\pi\theta_1},\ldots,e^{j2\pi\theta_n})$. Analyzing (3) directly being difficult in general, we try to derive insightful bounds instead. To that end, we need the following lemma.

Lemma 1: Some properties on $\det\left(\mathbf{I} + \mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}} \boldsymbol{P} \boldsymbol{R}_{\boldsymbol{\theta}}^* \boldsymbol{Q}\right)$ are listed as follows:

Property 1: The determinant $\det\left(\mathbf{I} + \mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}} \boldsymbol{P} \boldsymbol{R}_{\boldsymbol{\theta}}^* \boldsymbol{Q}\right)$ is a multilinear function of $e^{\pm j2\pi\theta_i}$, $i=1,\ldots,n$.[2]

Property 2: For any $\mathcal{S} \subseteq \{1,\ldots,n\}$,

$$\int_{\mathcal{U}^{|\mathcal{S}|}} \det\left(\mathbf{I} + \mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}} \boldsymbol{P} \boldsymbol{R}_{\boldsymbol{\theta}}^* \boldsymbol{Q}\right) \mathrm{d}\boldsymbol{\theta}_{\mathcal{S}}$$

is a multilinear function of $e^{\pm j2\pi\theta_i}$, $\forall\, i \in \bar{\mathcal{S}}$, $\bar{\mathcal{S}}$ being the complementary set of $\mathcal{S}$ in $\{1,\ldots,n\}$.

Property 3:

$$\int_{\mathcal{U}^n} \det\left(\mathbf{I} + \mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}} \boldsymbol{P} \boldsymbol{R}_{\boldsymbol{\theta}}^* \boldsymbol{Q}\right) \mathrm{d}\boldsymbol{\theta} = \sum_{\mathcal{I}\subseteq\{1,\ldots,n\}} \mathsf{snr}^{|\mathcal{I}|} \det\left(\boldsymbol{P}_{\mathcal{I}}\right)\det\left(\boldsymbol{Q}_{\mathcal{I}}\right) \quad (4)$$

where we recall that $\det(A_\emptyset) = 1$.

Proof: See Appendix A ∎

The following bounds on I_T are obtained. Since the proof is quite involved, it is deferred in Section IV.

Theorem 1: The mutual information I_T of the rotate-and-forward scheme with n relays is upper and lower bounded by

$$I^\star(\mathsf{snr}) + (n-1) \geq I_T(\mathsf{snr}) \geq \left(\frac{K-1}{K}\right)^{n-1} I^\star(\mathsf{snr}) - 2 \quad (5)$$

with

$$I^\star(\mathsf{snr}) \triangleq \log\left(\sum_{\mathcal{I}\subseteq\{1,\ldots,n\}} \mathsf{snr}^{|\mathcal{I}|} \det\left(\boldsymbol{P}_{\mathcal{I}}\right)\det\left(\boldsymbol{Q}_{\mathcal{I}}\right)\right) \quad (6)$$

where we recall that $T = K^n$.

Now, let us define the outage probabilities

$$\mathbb{P}_{\mathrm{out},T}(R) \triangleq \mathbb{P}\left(I_T(\mathsf{snr}) < R\right),$$
$$\mathbb{P}^\star_{\mathrm{out}}(R) \triangleq \mathbb{P}\left(I^\star(\mathsf{snr}) < R\right).$$

[2]Here, there is a slight abuse of terminology. Technically, we should say that $\det(\mathbf{I}+\mathrm{diag}\{\boldsymbol{v}_1\}\boldsymbol{P}\mathrm{diag}\{\boldsymbol{v}_2\}\boldsymbol{Q})$ is multilinear function of $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{C}^{n\times 1}$, and specify $\mathrm{diag}\{\boldsymbol{v}_1\} = \boldsymbol{R}_{\boldsymbol{\theta}}$ and $\mathrm{diag}\{\boldsymbol{v}_2\} = \boldsymbol{R}_{\boldsymbol{\theta}}^*$.

Using this and (5), we obtain

$$\mathbb{P}^\star_{\mathrm{out}}(r\log\mathsf{snr} - (n-1)) \leq \mathbb{P}_{\mathrm{out},T}(r\log\mathsf{snr}) \leq \mathbb{P}^\star_{\mathrm{out}}\left(\left(\frac{K}{K-1}\right)^{n-1}(r\log\mathsf{snr}+2)\right).$$

C. Diversity-Multiplexing Tradeoff

By neglecting the constant terms and applying (1), we obtain the following corollary.

Corollary 1: Let $d_T(r)$ and $d^\star(r)$ denote respectively the DMT of the RF scheme and the DMT corresponding to $I^\star$. Then, we have

$$d^\star(r) \geq d_T(r) \geq d^\star\left(\left(\frac{K}{K-1}\right)^{n-1} r\right) \quad (7)$$

whence

$$d_\infty(r) \triangleq \lim_{K\to\infty} d_T(r) = d^\star(r).$$

The corollary states that the DMT of the RF scheme is upper bounded by $d^\star(r)$. More importantly, it is shown that this upper bound can be approached with a large number of rotation angles. The main message from these results is that to characterize the DMT of the RF scheme, it is enough to consider $d^\star(r)$, that is, to analyze $I^\star$ defined in (6). As a matter of fact, $I^\star$ is much more tractable than I_T, as will be shown in the next section.

Here comes the main result of the paper. The proof will be provided separately in Section V.

Theorem 2: In a two-hop (m,n,p) relay channel with i.i.d. Rayleigh fading, we have

$$d^\star(r) = d_{\min\{m,p\},n}(r) \quad (8)$$

where $d_{m,n}(r)$ denotes the DMT of a classical $m\times n$ MIMO channel, i.e., a piece-wise linear function connecting $(k,(m-k)(n-k))$, $k=0,\ldots,\min\{m,n\}$. Furthermore, it can be achieved by the rotate-and-forward scheme when $K\to\infty$.

Remark 3.1: The RF scheme achieves the *optimal* DMT in this setting. To see this, let us consider the following two *cuts* between the source and the destination: 1) source-relays cut is an $m\times n$ Rayleigh channel, and 2) relays-destination cut is an $n\times p$ Rayleigh channel. According to the information theoretic max-flow min-cut theorem [23], it is readily shown that the DMT of the end-to-end channel with *any* relaying strategy is dominated by the DMT of either cut, i.e., $\min\left\{d_{(m,n)}(r), d_{(n,p)}(r)\right\}$ that coincides with (8).

Remark 3.2: If the relays could cooperate perfectly (e.g., co-located relays), it would be possible to perform joint decoding and joint encoding to achieve the cut-set bound in a straightforward manner. However, Theorem 2 shows that even with linear distributed relaying, the cut-set bound can be achieved. To the authors' best knowledge, this is the first distributed linear scheme that achieves the optimal DMT in such a setting with any number of antennas.

Remark 3.3: While the RF scheme is designed here for the two-hop relay channel without direct source-destination link, it can also be applied to the non-orthogonal AF schemes [10], [24] to improve the performance. The key idea is to extend

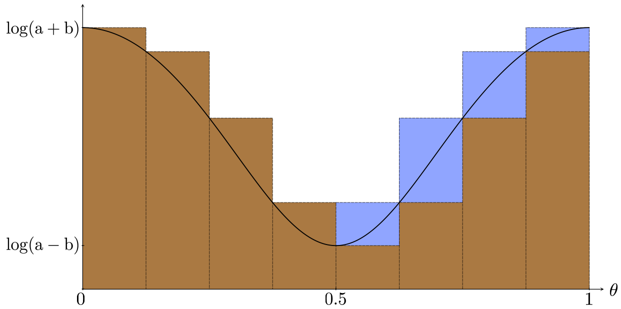

Fig. 2. An example of Riemann sum approximation of $\int_0^1 \log(a + b\cos(2\pi\theta))\,\mathrm{d}\theta$ with $K = 8$. The Riemann sum area is represented by the dark rectangles. The additional light rectangles are needed in order to cover the area of integration.

the AF relaying to RF relaying, that is, to introduce time-varying AF processing wherever multiple *distributed* relays are involved.

Remark 3.4: With the lower bound in (7), we obtain a sufficient condition of $K \to \infty$ for the DMT optimality of the RF scheme. But the converse is not true in general, i.e., K need not go to infinity for the RF scheme to achieve the optimal DMT. For example, it is shown in [25] that the flip-and-forward scheme, a particular case of rotate-and-forward with $K = 2$ is DMT optimal for any $(m, 2, p)$ channel with i.i.d. Rayleigh fading. Therefore, supported by this example, one may expect to have a better lower bound than (7). It is however still an open question.

Remark 3.5: The lower bound in (7) *is* tight for $r = 0$, for any $K \geq 2$. It means that the RF scheme achieves the maximum diversity gain whenever $K \geq 2$. On the other hand, since the rotations are independent of the relay-destination channel matrix G, the rank of the end-to-end channel matrix remains the same after the rotation with probability 1. Therefore, the maximum multiplexing gain is achieved for any K (e.g., $K = 1$ for the AF case).

IV. Proof of Theorem 1

Let us first prove the lower bound. For $K = 1$, the lower and upper bounds in (5) are trivial. Therefore, we assume $K \geq 2$ in the rest of the proof. To prove the proposition, we pick up (3) in which $\det(\mathbf{I} + \mathsf{snr}\boldsymbol{R_\theta P R_\theta^* Q})$ is multilinear function of $e^{\pm j2\pi\theta_i}$, $i = 1, \dots, n$. Furthermore, since $\boldsymbol{P}, \boldsymbol{Q} \succeq \mathbf{0}$, it is readily shown that $\det(\mathbf{I} + \mathsf{snr}\boldsymbol{R_\theta P R_\theta^* Q})$ is real and has the following form in terms of θ_i

$$\det(\mathbf{I} + \mathsf{snr}\boldsymbol{R_\theta P R_\theta^* Q}) = a_i + b_i\cos(2\pi\theta_i + \phi_i), \quad i = 1, \dots, n \tag{9}$$

where $a_i \geq 0$, $b_i \geq 0$, and $\phi_i \in [0, 2\pi)$ do not depend on θ_i. It is worth noting that

$$a_i = \int_0^1 \det(\mathbf{I} + \mathsf{snr}\boldsymbol{R_\theta P R_\theta^* Q})\,\mathrm{d}\theta_i \tag{10}$$

is also multilinear function of the rest of the θ_i. In addition, we can show that

$$a_i - b_i \geq 1, \quad \forall i. \tag{11}$$

Hence, we can take a closer look on the Riemann sum of θ_1 in (3)

$$\begin{aligned}
&\sum_{\theta_1 \in \mathcal{A}_K} \log\det(\mathbf{I} + \mathsf{snr}\boldsymbol{R_\theta P R_\theta^* Q}) \\
&\quad= \sum_{\theta_1 \in \mathcal{A}_K} \log(a_1 + b_1\cos(2\pi\theta_1 + \phi_1)) &\quad (12)\\
&\quad\geq K\int_0^1 \log(a_1 + b_1\cos(2\pi\theta_1 + \phi_1))\,\mathrm{d}\theta_1 \\
&\qquad - \big(\log(a_1 + b_1) - \log(a_1 - b_1)\big) &\quad (13)
\end{aligned}$$

where (12) is from (9); (13) is actually nothing but approximating the integral by a Riemann sum. To see this, let us consider an illustrative example in Fig. 2 where K is even and $\phi = 0$. The dark rectangles (Riemann sum (12)) together with the light rectangles (second term in (13)) are larger than the integrated area (first term in (13)). Furthermore, it can be verified that for general K and ϕ, the hatched area can only be smaller.[3] Therefore, (13) always holds.

Lemma 2: For any a, b with $a \geq b \geq 0$ and $a \neq 0$, we have

$$\log(a) - 1 \leq \int_0^1 \log(a + b\cos(2\pi\theta))\mathrm{d}\theta \leq \log(a). \tag{14}$$

Proof: For $b = 0$, (14) holds trivially. For $a \geq b > 0$, we use the following equality [26], [27],

$$\int_0^1 \log(a + b\cos(2\pi\theta))\mathrm{d}\theta = \log\left(\frac{a + \sqrt{a^2 - b^2}}{2}\right)$$

from which (14) is immediate. ■

By proceeding further from (13), we can get simpler lower bounds

$$\begin{aligned}
&\sum_{\theta_1 \in \mathcal{A}_K} \log\det(\mathbf{I} + \mathsf{snr}\boldsymbol{R_\theta P R_\theta^* Q}) \\
&\quad\geq K(\log(a_1) - 1) - \log(2a_1 - 1) &\quad (15)\\
&\quad\geq K(\log(a_1) - 1) - (\log(a_1) + 1) &\quad (16)\\
&\quad= (K-1)\log(a_1) - (K+1) &\quad (17)
\end{aligned}$$

where the first and second terms in (15) are from the lower bound in (14) and from (11), respectively; (16) is from $\log(2a_1 - 1) \leq \log(2a_1) = \log(a_1) + 1$. From (17) and (10), we get

$$\begin{aligned}
&\sum_{\theta_1 \in \mathcal{A}_K} \log\det(\mathbf{I} + \mathsf{snr}\boldsymbol{R_\theta P R_\theta^* Q}) \\
&\quad\geq (K-1)\log\left(\int_0^1 \det(\mathbf{I} + \mathsf{snr}\boldsymbol{R_\theta P R_\theta^* Q})\,\mathrm{d}\theta_1\right) \\
&\qquad - (K+1).
\end{aligned}$$

Since $\int_0^1 \det(\mathbf{I} + \mathsf{snr}\boldsymbol{R_\theta P R_\theta^* Q})\,\mathrm{d}\theta_1$ is real and multilinear on $e^{\pm j2\pi\theta_2}$ according to Property 2 from Lemma 1, it is also in the form $a_2' + b_2'\cos(2\pi\theta_2 + \phi_2')$ with

$$a_2' = \int_0^1\int_0^1 \det(\mathbf{I} + \mathsf{snr}\boldsymbol{R_\theta P R_\theta^* Q})\,\mathrm{d}\theta_1\mathrm{d}\theta_2. \tag{18}$$

[3] It can be proved by basic maths and is omitted for conciseness.

Therefore, with the same reasoning as above on θ_1, it is readily shown that

$$\sum_{\theta_1\in\mathcal{A}_K,}\sum_{\theta_2\in\mathcal{A}_K}\log\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{\theta}}^{*}\boldsymbol{Q}\right) \geq (K-1)^2\log\left(\int_0^1\int_0^1\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{\theta}}^{*}\boldsymbol{Q}\right)\mathrm{d}\theta_1\mathrm{d}\theta_2\right) - (K-1)(K+1)-(K+1)$$

and

$$\sum_{\theta_{n-1}\in\mathcal{A}_K}\cdots\sum_{\theta_1\in\mathcal{A}_K}\log\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{\theta}}^{*}\boldsymbol{Q}\right) \geq (K-1)^{n-1}\log\left(\int_{\mathcal{U}^{n-1}}\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{\theta}}^{*}\boldsymbol{Q}\right)\mathrm{d}\theta_1\cdots\mathrm{d}\theta_{n-1}\right) - (K+1)\sum_{k=0}^{n-2}(K-1)^k. \quad (19)$$

It is worth noting that the integral inside the logarithm in (19) does not depend on θ_n and can be simply replaced by $\int_{\mathcal{U}^n}\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{\theta}}^{*}\boldsymbol{Q}\right)\mathrm{d}\boldsymbol{\theta}$. Moreover, we have

$$\begin{aligned}(K+1)\sum_{k=0}^{n-2}(K-1)^k &= 1+K+(K+1)\sum_{k=1}^{n-2}(K-1)^k\\ &\leq 1+K+\sum_{k=1}^{n-2}(K^2-1)(K-1)^{k-1}\\ &\leq \sum_{k=0}^{n-1}K^k\\ &\leq \frac{K^n-1}{K-1}\\ &\leq 2K^{n-1}. \end{aligned}\quad (20)$$

Finally, combining (6), (3), (4), (19), and (20), we obtain the lower bound in (5).

The proof of the upper bound uses the same ideas as the lower bound. First, note that

$$\begin{aligned}I_T(\mathsf{snr}) &\leq \max_{\boldsymbol{\theta}\in\mathcal{U}^n}\log\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{\theta}}^{*}\boldsymbol{Q}\right)\\ &= \max_{\theta_n\in\mathcal{U}}\cdots\max_{\theta_1\in\mathcal{U}}\log\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{\theta}}^{*}\boldsymbol{Q}\right).\end{aligned}$$

From (9), we have

$$\begin{aligned}\max_{\theta_1\in\mathcal{U}}\log\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{\theta}}^{*}\boldsymbol{Q}\right) &= \log(a_1+b_1)\\ &\leq \log(a_1)+1\end{aligned}$$

where a_1 is defined in (10). As noted before, a_1 is also in the form $a_2'+b_2'\cos(2\pi\theta_2+\phi_2')$ with a_2' defined in (18). Therefore, we have

$$\max_{\theta_2\in\mathcal{U}}\max_{\theta_1\in\mathcal{U}}\log\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{\theta}}^{*}\boldsymbol{Q}\right) \leq \log\left(\int_0^1\int_0^1\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{\theta}}^{*}\boldsymbol{Q}\right)\mathrm{d}\theta_1\mathrm{d}\theta_2\right)+2$$

and then

$$\max_{\theta_{n-1}\in\mathcal{U}}\cdots\max_{\theta_1\in\mathcal{U}}\log\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{\theta}}^{*}\boldsymbol{Q}\right) \leq \log\left(\int_{\mathcal{U}^{n-1}}\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{\theta}}^{*}\boldsymbol{Q}\right)\mathrm{d}\theta_1\cdots\mathrm{d}\theta_{n-1}\right) + (n-1). \quad (21)$$

Again, note that the integral inside the logarithm in (21) does not depend on θ_n and can be simply replaced by $\int_{\mathcal{U}^n}\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{\theta}}^{*}\boldsymbol{Q}\right)\mathrm{d}\boldsymbol{\theta}$. Finally, the upper bound in (5) follows straightforwardly. This completes the proof of Theorem 1. ∎

V. Proof of Theorem 2

A. Preliminary: LQ decomposition of the GUE ensemble

In this section, we describe a new approach for analyzing the joint distribution of the entries of random matrices with complex Gaussian entries (GUE ensemble) performing the LQ decomposition of the matrix. This approach turns out to be the key for proving the DMT optimality of the rotate-and-forward scheme. Let $\boldsymbol{H}$ be the $n\times m$ channel matrix, with i.i.d. circularly symmetric complex Gaussian entries with unit variance. The idea is to permute the rows of matrix $\boldsymbol{H}$ such that the diagonal entries of lower triangular matrix L provide us the optimal DMT.

Lemma 3 ([28]): Suppose we have n i.i.d. samples $X_1, X_2, \ldots X_n$ from a continuous distribution with density $f_X(x)$. Then, the joint distribution of the maximum of the samples $X_{(1)}$ and the rest is

$$f_{X_{(1)},X_2,\ldots,X_n}(x_1,\ldots,x_n) = n\prod_{i=1}^{n}f_X(x_i), \quad x_1\geq x_i,\ \forall\, i\geq 2.$$

The LQ decomposition of matrix $\boldsymbol{H}$ can be done by the following procedure. Let $\boldsymbol{h}_i\in\mathbb{C}^{1\times m}$ be the i th row vector of matrix $\boldsymbol{H}$, with probability density function (pdf) $p_h(\boldsymbol{h}_i)$. The joint pdf of $\boldsymbol{H}$ is

$$p_H(\boldsymbol{H}) = \prod_{i=1}^{n}p_h(\boldsymbol{h}_i).$$

In first step, we permute the matrix such that the row with the largest norm would be the first row. By Lemma 3, putting the strongest vector in the first row will give us the following pdf

$$p(\tilde{\boldsymbol{h}}_1,\ldots,\tilde{\boldsymbol{h}}_n) = n\prod_{i=1}^{n}p_h(\tilde{\boldsymbol{h}}_i), \quad \|\tilde{\boldsymbol{h}}_1\|\geq\|\tilde{\boldsymbol{h}}_i\|,\ \forall\, i\geq 2,$$

and after marginalization

$$p(\|\tilde{\boldsymbol{h}}_1\|^2,\tilde{\boldsymbol{h}}_2,\ldots,\tilde{\boldsymbol{h}}_n) = n\,p_{\chi^2_{2n}}(\|\tilde{\boldsymbol{h}}_1\|^2)\prod_{i=2}^{n}p_h(\tilde{\boldsymbol{h}}_i), \quad \|\tilde{\boldsymbol{h}}_1\|\geq\|\tilde{\boldsymbol{h}}_i\|,\ \forall\, i\geq 2.$$

Performing the unitary transform $\boldsymbol{U}_1$ (the first row of $\boldsymbol{U}_1$ aligns with $\tilde{\boldsymbol{h}}_1$) on $\tilde{\boldsymbol{h}}_2, \ldots, \tilde{\boldsymbol{h}}_n$, the joint pdf is unchanged

$$p(\|\tilde{\boldsymbol{h}}_1\|^2, \tilde{\boldsymbol{h}}_2\boldsymbol{U}_1, \ldots, \tilde{\boldsymbol{h}}_n\boldsymbol{U}_1) = n\, p_{\chi^2_{2m}}(\|\tilde{\boldsymbol{h}}_1\|^2) \prod_{i=2}^{n} p_h(\tilde{\boldsymbol{h}}_i\boldsymbol{U}_1),$$
$$\|\tilde{\boldsymbol{h}}_1\| \geq \|\tilde{\boldsymbol{h}}_i\|, \ \forall\, i \geq 2.$$

Setting $l_{11} = \|\tilde{\boldsymbol{h}}_1\|$ and $\hat{\boldsymbol{h}}_i = \tilde{\boldsymbol{h}}_i\boldsymbol{U}_1$, $i \geq 2$, we have

$$p(l_{11}^2, \hat{\boldsymbol{h}}_2, \ldots, \hat{\boldsymbol{h}}_n) = n\, p_{\chi^2_{2m}}(l_{11}^2) \prod_{i=2}^{n} p_h(\hat{\boldsymbol{h}}_i),$$
$$l_{11} \geq \|\hat{\boldsymbol{h}}_i\|, \ \forall\, i \geq 2.$$

Now, let $\bar{\boldsymbol{h}}_2, \ldots, \bar{\boldsymbol{h}}_n$ be any ordered version of $\hat{\boldsymbol{h}}_2, \ldots, \hat{\boldsymbol{h}}_n$ such that $\bar{\boldsymbol{h}}_2$ has the largest norm of the $(2, \ldots, n)$ subvector, we have the pdf

$$p(l_{11}^2, \bar{\boldsymbol{h}}_2, \ldots, \bar{\boldsymbol{h}}_n) = n(n-1) p_{\chi^2_{2m}}(l_{11}^2) \prod_{i=2}^{n} p_h(\bar{\boldsymbol{h}}_i),$$
$$l_{11} \geq \|\bar{\boldsymbol{h}}_i\|, \ \forall\, i \geq 2,$$
$$\sum_{j=2}^{n} |\bar{h}_{2j}|^2 \geq \sum_{j=2}^{n} |\bar{h}_{ij}|^2, \ \forall\, i \geq 3.$$

We can now set $l_{21} = \bar{h}_{21}$, $l_{22}^2 = \sum_{j=2}^{n} |\bar{h}_{2j}|^2$, and perform the unitary transform $\boldsymbol{U}_2$ on the $(2, \ldots, n; 2, \ldots, n)$ submatrix and repeat the same procedure above

$$\begin{aligned} &p(l_{11}^2, l_{21}, l_{22}^2, \check{\boldsymbol{h}}_3, \ldots, \check{\boldsymbol{h}}_n) \\ &\quad = n(n-1) p_{\chi^2_{2n}}(l_{11}^2) p_{\mathcal{N}_\mathbb{C}}(l_{21}) p_{\chi^2_{2(m-1)}}(l_{22}^2) \prod_{i=3}^{n} p_h(\check{\boldsymbol{h}}_i), \end{aligned}$$
$$l_{11}^2 \geq |l_{21}|^2 + l_{22}^2,$$
$$l_{11}^2 \geq \|\check{\boldsymbol{h}}_i\|^2, \ \forall\, i \geq 3,$$
$$l_{22}^2 \geq \sum_{j=2}^{n} |\check{h}_{ij}|^2, \ \forall\, i \geq 3.$$

Now, if we continue, we will get a lower-triangular matrix $\boldsymbol{L}$ with the following distribution[4]

$$p_L(\boldsymbol{L}) = n! \prod_{i=1}^{n} p_{\chi^2_{2(m-i+1)}}(l_{ii}^2) \prod_{j<i} p_{\mathcal{N}_\mathbb{C}}(l_{ij}),$$
$$l_{ii}^2 \geq \sum_{j=i}^{k} |l_{kj}|^2, \ \forall\, k > i$$

where $p_{\mathcal{N}_\mathbb{C}}(\cdot)$ is the pdf of the standard circularly symmetric Gaussian distribution. Note that a similar expression (with the $n!$ factor) appears in the joint distribution of order statistics of n i.i.d. samples.

[4]With a slight abuse of notation, we use $p_L(\boldsymbol{L})$ to denote the joint pdf of $\{l_{ii}^2 : \ i = 1, \ldots, n\}$ and $\{l_{ij} : \ n \geq i > j \geq 1\}$. Similar notation with be applied to $p_R(\boldsymbol{R})$ later on.

B. Simultaneous LQ and QR decomposition

In this section, we perform a simultaneous LQ and QR decomposition on the two channel matrices $\boldsymbol{F} \in \mathbb{C}^{m\times n}$ and $\boldsymbol{G} \in \mathbb{C}^{n\times p}$ with i.i.d. $\mathcal{N}_\mathbb{C}(0,1)$ entries. Suppose by now that $\min\{m, p\} \geq n$. Let $\boldsymbol{f}_i$ and $\boldsymbol{g}_i$ be the i th row and column vectors of $\boldsymbol{F}$ and $\boldsymbol{G}$, respectively. The joint pdf of $\boldsymbol{F}$ and $\boldsymbol{G}$ is

$$\begin{aligned} p_F(\boldsymbol{F}) p_G(\boldsymbol{G}) &= \prod_{i=1}^{n} p_f(\boldsymbol{f}_i) \prod_{i=1}^{n} p_g(\boldsymbol{g}_i) \\ &= p\left(\boldsymbol{f}_1, \boldsymbol{g}_1, \ldots, \boldsymbol{f}_n, \boldsymbol{g}_n\right). \end{aligned}$$

Similar to Section V-A, we put the vector pair $(\boldsymbol{f}_i, \boldsymbol{g}_i)$ in the first row and column of $\boldsymbol{F}$ and $\boldsymbol{G}$ if the product norm $\|\boldsymbol{f}_i\|\|\boldsymbol{g}_i\|$ is the largest. It will give us the following pdf

$$p(\tilde{\boldsymbol{f}}_1, \tilde{\boldsymbol{g}}_1, \ldots, \tilde{\boldsymbol{f}}_n, \tilde{\boldsymbol{g}}_n) = n \prod_{i=1}^{n} p_{\boldsymbol{f}}(\tilde{\boldsymbol{f}}_i) \prod_{i=1}^{n} p_g(\tilde{\boldsymbol{g}}_i),$$
$$\|\tilde{\boldsymbol{f}}_1\|\|\tilde{\boldsymbol{g}}_1\| \geq \|\tilde{\boldsymbol{f}}_i\|\|\tilde{\boldsymbol{g}}_i\|, \ \forall\, i \geq 2$$

and after marginalization

$$\begin{aligned} &p(\|\tilde{\boldsymbol{f}}_1\|^2, \|\tilde{\boldsymbol{g}}_1\|^2, \tilde{\boldsymbol{f}}_2, \tilde{\boldsymbol{g}}_2, \ldots, \tilde{\boldsymbol{f}}_n, \tilde{\boldsymbol{g}}_n) \\ &\quad = n\, p_{\chi^2_{2m}}(\|\tilde{\boldsymbol{f}}_1\|^2) p_{\chi^2_{2p}}(\|\tilde{\boldsymbol{g}}_1\|^2) \prod_{i=2}^{n} p_f(\tilde{\boldsymbol{f}}_i) \prod_{i=2}^{n} p_g(\tilde{\boldsymbol{g}}_i), \end{aligned}$$
$$\|\tilde{\boldsymbol{f}}_1\|\|\tilde{\boldsymbol{g}}_1\| \geq \|\tilde{\boldsymbol{f}}_i\|\|\tilde{\boldsymbol{g}}_i\|, \ \forall\, i \geq 2.$$

Performing the unitary transform $\boldsymbol{U}_1$ (the first column of $\boldsymbol{U}_1$ aligns with $\tilde{\boldsymbol{f}}_1$) on $\tilde{\boldsymbol{f}}_2, \ldots, \tilde{\boldsymbol{f}}_n$, and $\boldsymbol{V}_1$ (the first row of $\boldsymbol{V}_1$ aligns with $\tilde{\boldsymbol{g}}_1$), the joint pdf is unchanged

$$\begin{aligned} &p(\|\tilde{\boldsymbol{f}}_1\|^2, \|\tilde{\boldsymbol{g}}_1\|^2, \tilde{\boldsymbol{f}}_2\boldsymbol{U}_1, \boldsymbol{V}_1\tilde{\boldsymbol{g}}_2, \ldots, \tilde{\boldsymbol{f}}_n\boldsymbol{U}_1, \boldsymbol{V}_1\tilde{\boldsymbol{g}}_n) \\ &\quad = n\, p_{\chi^2_{2m}}(\|\tilde{\boldsymbol{f}}_1\|^2) p_{\chi^2_{2p}}(\|\tilde{\boldsymbol{g}}_1\|^2) \prod_{i=2}^{n} p_f(\tilde{\boldsymbol{f}}_i\boldsymbol{U}_1) \prod_{i=2}^{n} p_g(\boldsymbol{V}_1\tilde{\boldsymbol{g}}_i), \end{aligned}$$
$$\|\tilde{\boldsymbol{f}}_1\|\|\tilde{\boldsymbol{g}}_1\| \geq \|\tilde{\boldsymbol{f}}_i\|\|\tilde{\boldsymbol{g}}_i\|, \ \forall\, i \geq 2.$$

Setting $l_{11} = \|\tilde{\boldsymbol{f}}_1\|$ and $r_{11} = \|\tilde{\boldsymbol{g}}_1\|$ and $\hat{\boldsymbol{f}}_i = \tilde{\boldsymbol{f}}_i\boldsymbol{U}_1$, $\hat{\boldsymbol{g}}_i = V_1\tilde{\boldsymbol{g}}_i$, $i \geq 2$, we have

$$\begin{aligned} &p(l_{11}^2, r_{11}^2, \hat{\boldsymbol{f}}_2, \hat{\boldsymbol{g}}_2, \ldots, \hat{\boldsymbol{f}}_n, \hat{\boldsymbol{g}}_n) \\ &\quad = n\, p_{\chi^2_{2m}}(l_{11}^2) p_{\chi^2_{2p}}(r_{11}^2) \prod_{i=2}^{n} p_f(\hat{\boldsymbol{f}}_i) \prod_{i=2}^{n} p_g(\hat{\boldsymbol{g}}_i), \end{aligned}$$
$$l_{11} r_{11} \geq \|\hat{\boldsymbol{f}}_i\|\|\hat{\boldsymbol{g}}_i\|, \ \forall\, i \geq 2.$$

Now, as in Section V-A, we order the rest of the row-column vector pairs $(\hat{\boldsymbol{f}}_2, \hat{\boldsymbol{g}}_2), \ldots, (\hat{\boldsymbol{f}}_n, \hat{\boldsymbol{g}}_n)$ such that $(\bar{\boldsymbol{f}}_2, \bar{\boldsymbol{g}}_2)$ has the largest product norm of the $(2, \ldots, n)$ subvector. We obtain

$$\begin{aligned} &p(l_{11}^2, r_{11}^2, \bar{\boldsymbol{f}}_2, \bar{\boldsymbol{g}}_2, \ldots, \bar{\boldsymbol{f}}_n, \bar{\boldsymbol{g}}_n) \\ &\quad = n(n-1) p_{\chi^2_{2m}}(l_{11}^2) p_{\chi^2_{2p}}(r_{11}^2) \prod_{i=2}^{n} p_f(\bar{\boldsymbol{f}}_i) \prod_{i=2}^{n} p_g(\bar{\boldsymbol{g}}_i), \end{aligned}$$
$$l_{11} r_{11} \geq \|\bar{\boldsymbol{f}}_i\|\|\bar{\boldsymbol{g}}_i\|, \ \forall\, i \geq 2,$$
$$\sum_{j=2}^{n} |\bar{f}_{2j}|^2 \left(\sum_{j=2}^{n} |\bar{g}_{2j}|^2\right) \geq \sum_{j=2}^{n} |\bar{f}_{ij}|^2 \left(\sum_{j=2}^{n} |\bar{g}_{ij}|^2\right), \ \forall\, i \geq 3.$$

We can now set $l_{21} = \bar{f}_{21}$, $r_{12} = \bar{g}_{21}$, $l_{22}^2 = \sum_{j=2}^{n} |\bar{f}_{2j}|^2$, $r_{22}^2 = \sum_{j=2}^{n} |\bar{g}_{2j}|^2$, and perform the unitary transform $\boldsymbol{U}_2$ and $\boldsymbol{V}_2$ on the $(2,\ldots,n;2,\ldots,n)$ submatrices of $\boldsymbol{F}$ and $\boldsymbol{G}$, respectively. Now, if we continue, we will get a lower-triangular matrix $\boldsymbol{L}$ and a upper-triangular matrix $\boldsymbol{R}$ with the following joint distribution

$$\begin{aligned} & p_L(\boldsymbol{L})p_R(\boldsymbol{R}) \\ & = n! \prod_{i=1}^{n} p_{\chi^2_{2(m-i+1)}}(l_{ii}^2) p_{\chi^2_{2(p-i+1)}}(r_{ii}^2) \prod_{j<i} p_{\mathcal{N}_{\mathbb{C}}}(l_{ij}) p_{\mathcal{N}_{\mathbb{C}}}(r_{ji}), \\ & \qquad l_{ii}^2 r_{ii}^2 \geq \left(\sum_{j=i}^{k} |l_{kj}|^2\right)\left(\sum_{j=i}^{k} |r_{jk}|^2\right), \ \forall\, k > i. \end{aligned} \tag{22}$$

Keep in mind that the above constraint implies that $l_{ii}^2 r_{ii}^2 \geq |l_{kj}|^2 |r_{jk}|^2$, $\forall\, k \geq j \geq i$.

Observe that the permutation as well as the unitary transform on matrices $\boldsymbol{F}$ and $\boldsymbol{G}$ will not change the expression I_T in (3) for the mutual information. Let $\boldsymbol{\Pi}$ be the permutation matrix. Then, $\boldsymbol{F} = \boldsymbol{\Pi} \boldsymbol{L} \boldsymbol{Q}$ and $\boldsymbol{G} = \boldsymbol{Q} \boldsymbol{R} \boldsymbol{\Pi}^T$. Therefore,

$$\begin{aligned} & \det\left(\mathbf{I} + \mathsf{snr} \boldsymbol{R_\theta} \boldsymbol{F} \boldsymbol{F}^* \boldsymbol{R_\theta}^* \boldsymbol{G}^* \boldsymbol{G}\right) \\ & \qquad = \det\left(\mathbf{I} + \mathsf{snr} \boldsymbol{\Pi}^T \boldsymbol{R_\theta} \boldsymbol{\Pi} \boldsymbol{L} \boldsymbol{L}^* \boldsymbol{\Pi}^* \boldsymbol{R_\theta}^* \boldsymbol{\Pi} \boldsymbol{R}^* \boldsymbol{R}\right). \end{aligned}$$

Since $\boldsymbol{\Pi}$ is a permutation matrix, $\boldsymbol{\Pi}^T \boldsymbol{R_\theta} \boldsymbol{\Pi}$ has the same structure as $\boldsymbol{R_\theta}$ and (6) still holds.

C. DMT of the Rotate-and-Forward Scheme

Let us recall that we identify I_T with $I^\star$ in the high snr regime, since the latter is achievable when $K \to \infty$.

$$I^\star = \log\left(\sum_{\mathcal{I}} \mathsf{snr}^{|\mathcal{I}|} \det\left(\boldsymbol{F}\boldsymbol{F}^*\right)_{\mathcal{I}} \det\left(\boldsymbol{G}^*\boldsymbol{G}\right)_{\mathcal{I}}\right)$$

It is lower-bounded by

$$I^\star \geq \log\left(\sum_{k=0}^{n} \mathsf{snr}^{k} \det\left(\boldsymbol{H}\boldsymbol{H}^*\right)_{\mathcal{I}_k} \det\left(\boldsymbol{G}^*\boldsymbol{G}\right)_{\mathcal{I}_k}\right)$$

where $\mathcal{I}_k \triangleq \{1,\ldots,k\}$ and $\mathcal{I}_0 \triangleq \emptyset$. Noting that $\boldsymbol{F}\boldsymbol{F}^* = \boldsymbol{L}\boldsymbol{L}^*$ and $\boldsymbol{G}^*\boldsymbol{G} = \boldsymbol{R}^*\boldsymbol{R}$, we have

$$\begin{aligned} I^\star & \geq \log\left(\sum_{k=0}^{n} \mathsf{snr}^{k} \det\left(\boldsymbol{L}\boldsymbol{L}^*\right)_{\mathcal{I}_k} \det\left(\boldsymbol{R}^*\boldsymbol{R}\right)_{\mathcal{I}_k}\right) \\ & = \log\left(\sum_{k=0}^{n} \mathsf{snr}^{k} \prod_{i=1}^{k} l_{ii}^2 r_{ii}^2\right) \end{aligned}$$

Now, we are ready to compute the exact DMT of the lower bound for the rotate-and-forward scheme. Gathering all pieces together, we obtain

$$\begin{aligned} & \mathbb{P}_{\text{out}}(r \log \mathsf{snr}) \\ & \qquad \doteq \mathbb{P}(I^* < r \log \mathsf{snr}) \\ & \qquad \dot{\leq} \mathbb{P}\left(\log\left(\sum_{k=0}^{n} \mathsf{snr}^{k} \prod_{i=1}^{k} l_{ii}^2 r_{ii}^2\right) < r \log \mathsf{snr}\right). \end{aligned} \tag{23}$$

1) Case $m = p > n$: Let us set $|l_{ij}|^2 = \mathsf{snr}^{-\alpha_{ij}}$ and $|r_{ij}|^2 = \mathsf{snr}^{-\beta_{ij}}$. The DMT of the above upper bound is the lower bound of $d^\star(r)$. It follows that

$$d_{\text{LB}}(r) = \min \sum_{i=1}^{n} (m+1-i)(\alpha_{ii} + \beta_{ii}) + \sum_{j<i} (\alpha_{ij} + \beta_{ji})$$

subject to

$$k - \sum_{i=1}^{k} (\alpha_{ii} + \beta_{ii}) \leq r,\ 1 \leq k \leq n \text{ (outage region)} \tag{24}$$

$$\begin{aligned} & \alpha_{ii} + \beta_{ii} \leq \alpha_{jk} + \beta_{kj},\ \forall\, j \geq k \geq i, \\ & \quad \alpha_{ij}, \beta_{ij} \geq 0,\ \forall\, i,j \text{ (pdf region)} \end{aligned} \tag{25}$$

where (24) is from the outage event in the upper bound (23); (25) is from the region in which (22) the pdf is defined.

Note that we can perform the variable changes $\eta_{ij} = \alpha_{ij} + \beta_{ji}$ and have

$$d_{\text{LB}}(r) = \min \sum_{j=1}^{n} (m+1-j)\eta_{jj} + \sum_{i>j} \eta_{ij} \tag{26}$$

subject to

$$\begin{aligned} & k - \sum_{i=1}^{k} \eta_{ii} \leq r, \quad k = 1,\ldots,n \\ & \eta_{ii} \leq \eta_{jk},\ \forall\, j \geq k \geq i, \quad \eta_{ij} \geq 0,\ \forall\, i,j \end{aligned} \tag{27}$$

which is exactly what we would get if we had only one matrix instead of two. In the following, we will find a lower bound on the (26) by

1) using the fact $\eta_{ij} \geq \eta_{jj}$ for $i < j$ from the pdf region (27), and
2) enlarging the pdf region by relaxing the constraints in (27) to $0 \leq \eta_{ii} \leq \eta_{jj}, \forall\, i \leq j$.

It is then readily shown that

$$d_{\text{LB}}(r) \geq \min \sum_{j=1}^{n} (2m+1-2j)\eta_{jj}$$

subject to

$$\begin{aligned} & k - \sum_{i=1}^{k} \eta_{ii} \leq r, \quad k = 1,\ldots,n \\ & 0 \leq \eta_{ii} \leq \eta_{jj},\ \forall\, j \geq i. \end{aligned}$$

And this optimization coincides perfectly with the one from the eigenvalue formulation for the classical MIMO channel in [3]. Therefore, the DMT optimality of the RF scheme is shown without directly solving this problem.

2) Case $m \neq p$: Let q denote the minimum of m and p. By the cutset bound, the diversity of the system is upper bounded by the diversity of each stage, i.e.,

$$d(r) \leq \min\{d_{m,n}(r), d_{n,p}(r)\} = d_{n,q}(r). \tag{28}$$

- If $m > p = q$, we can simply not send any signal in $m-p$ antennas at the source node. So we can apply the result giving us a lower bound on the DMT which matches the upper bound in (28). Therefore, $d(r) = d_{n,q}(r)$.

- If $q = m < p$, we can simply ignore the received signal in $p - m$ antennas at the destination node. Again, using the cutset bound, we deduce that $d(r) = d_{n,q}(r)$.

This completes the proof of Theorem 2. ■

VI. Towards Multiple Hops

A. Generalize the RF Scheme

Let us now consider an N-hop relay channel with $N-1$ layers of relays. As before, we assume m antennas at the source, p antennas at the destination, and $(n_1, n_2, \cdots, n_{N-1})$ antennas at the intermediate relaying layers, respectively. The channel matrices are denoted by $\boldsymbol{H}_1$, $\boldsymbol{H}_2$, …, $\boldsymbol{H}_N$ for the N hops respectively.

Then, it is straightforward to extend the two-hop RF scheme to the general N-hop case. Instead of using one DRS, each layer fixes a DRS $\{\boldsymbol{\Delta}_{i,t_i}\}_{t_i}$ with parameter K_i, $i = 1, \ldots, N-1$. Therefore, the RF scheme is a T-slot protocol with $T \triangleq \prod_{i=1}^{N-1} T_i$ and $T_i \triangleq K_i^{n_i}$. In each slot, the concatenated rotation sequence $(\boldsymbol{\Delta}_{1,t_1}, \ldots, \boldsymbol{\Delta}_{N-1,t_{N-1}})$ is different and all T different possibilities are run through in T time slots. It is readily shown that the average mutual information in bits per channel use can be obtained as a direct generalization of (2); the result is given in (29).

Now, we would like to bound $I_{T_1,\ldots,T_{N-1}}$ as we did in Theorem 1. To that end, we first consider the three-hop case.

1) The Three-Hop Case ($N = 3$): The expression (29) for the mutual information can be rewritten as (30), where we define $\boldsymbol{P}_1 \triangleq \boldsymbol{H}_1\boldsymbol{H}_1^*$ and $\boldsymbol{Q}_{(2)} \triangleq \boldsymbol{H}_2^*\boldsymbol{\Delta}_{2,t_2}^*\boldsymbol{H}_3^*\boldsymbol{H}_3\boldsymbol{\Delta}_{2,t_2}\boldsymbol{H}_2$. Let us denote the inner sum by $I_{\cdot,T_1}$ and we can bound it with the result from Theorem 1, i.e.,

$$I^\star_{(2)}(\mathsf{snr}) + (n_1-1) \geq I_{\cdot,T_1}(\mathsf{snr}) \geq \left(\frac{K_1-1}{K_1}\right)^{n_1-1} I^\star_{(2)}(\mathsf{snr}) - 2$$

with

$$I^\star_{(2)}(\mathsf{snr}) \triangleq \log\left(\sum_{\mathcal{I}\subseteq\{1,\ldots,n_1\}} \mathsf{snr}^{|\mathcal{I}|}\det(\boldsymbol{P}_1)_{\mathcal{I}}\det\left(\boldsymbol{Q}_{(2)}\right)_{\mathcal{I}}\right).$$

Now, it is sufficient to bound $\frac{1}{T_2}\sum_{t_2} I^\star_{(2)}(\mathsf{snr})$ in order to bound I_{T_2,T_1}. Note that only $\boldsymbol{Q}_{(2)}$ depends on $\boldsymbol{\Delta}_{2,t_2}$. Replacing $\boldsymbol{\Delta}_{2,t_2}$ with $\boldsymbol{R}_{\boldsymbol{\theta}}$, we notice that $\sum_{\mathcal{I}} \mathsf{snr}^{|\mathcal{I}|}\det(\boldsymbol{P}_1)_{\mathcal{I}}\det\left(\boldsymbol{Q}_{(2)}\right)_{\mathcal{I}}$ is real and multilinear function of $e^{\pm j2\pi\theta_i}$, $i = 1, \ldots, n_2$. Following exactly the same footsteps in Section IV, we obtain the following bounds by approximating the integral with Riemann sum n_2 times

$$\begin{aligned} I^\star_{(3)}(\mathsf{snr}) + (n_2-1) &\geq \frac{1}{T_2}\sum_{t_2=1}^{T_2} I^\star_{(2)}(\mathsf{snr}) \\ &\geq \left(\frac{K_2-1}{K_2}\right)^{n_2-1} I^\star_{(3)}(\mathsf{snr}) - 2 \end{aligned}$$

with

$$I^\star_{(3)}(\mathsf{snr}) \triangleq \log\int_{\mathcal{U}^{n_2}} \sum_{\mathcal{I}\subseteq\{1,\ldots,n_1\}} \mathsf{snr}^{|\mathcal{I}|}\det(\boldsymbol{P}_1)_{\mathcal{I}}\det\left(\boldsymbol{Q}_{(2)}\right)_{\mathcal{I}}\,\mathrm{d}\boldsymbol{\theta}.$$

Note that the above expression can be simplified since only $\det\left(\boldsymbol{Q}_{(2)}\right)_{\mathcal{I}}$ depends on $\boldsymbol{\theta}$. Thus, we have, for any $\epsilon > 0$,

$$\begin{aligned}
&\int_{\mathcal{U}^{n_2}} \det\left(\epsilon\mathbf{I} + \boldsymbol{Q}_{(2)}\right)_{\mathcal{I}}\mathrm{d}\boldsymbol{\theta} \\
&= \int_{\mathcal{U}^{n_2}} \det\left(\epsilon\mathbf{I} + (\boldsymbol{H}_2^*\boldsymbol{R}_{\boldsymbol{\theta}}^*\boldsymbol{H}_3^*\boldsymbol{H}_3\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{H}_2)_{\mathcal{I}}\right)\mathrm{d}\boldsymbol{\theta} \\
&= \int_{\mathcal{U}^{n_2}} \det\left(\epsilon\mathbf{I} + \boldsymbol{R}_{\boldsymbol{\theta}}^*\boldsymbol{H}_3^*\boldsymbol{H}_3\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{H}_{2,\cdot,\mathcal{I}}\boldsymbol{H}_{2,\cdot,\mathcal{I}}^*\right)\mathrm{d}\boldsymbol{\theta} \\
&= \int_{\mathcal{U}^{n_2}} \det\left(\epsilon\mathbf{I} + \boldsymbol{R}_{\boldsymbol{\theta}}^*\boldsymbol{Q}_3\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{S}_{2,\cdot,\mathcal{I}}\right)\mathrm{d}\boldsymbol{\theta} \\
&= \int_{\mathcal{U}^{n_2}} \epsilon^{|\mathcal{I}|}\det\left(\mathbf{I} + \epsilon^{-1}\boldsymbol{R}_{\boldsymbol{\theta}}^*\boldsymbol{Q}_3\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{S}_{2,\cdot,\mathcal{I}}\right)\mathrm{d}\boldsymbol{\theta} \\
&= \epsilon^{|\mathcal{I}|}\sum_{\mathcal{J}\subseteq\{1,\ldots,n_2\}} \epsilon^{-|\mathcal{J}|}\det(\boldsymbol{Q}_3)_{\mathcal{J}}\det(\boldsymbol{S}_{2,\mathcal{J},\mathcal{I}}) \qquad (31) \\
&= \sum_{\substack{\mathcal{J}\subseteq\{1,\ldots,n_2\} \\ |\mathcal{J}|\leq|\mathcal{I}|}} \epsilon^{|\mathcal{I}|-|\mathcal{J}|}\det(\boldsymbol{Q}_3)_{\mathcal{J}}\det(\boldsymbol{S}_{2,\mathcal{J},\mathcal{I}}) \qquad (32)
\end{aligned}$$

where we define $\boldsymbol{Q}_3 \triangleq \boldsymbol{H}_3^*\boldsymbol{H}_3$; $\boldsymbol{H}_{2,\mathcal{J},\mathcal{I}}$ is the submatrix of $\boldsymbol{H}_2$ formed by the rows and columns with indices in $\mathcal{J}$ and $\mathcal{I}$, respectively; in particular, we obtain $\boldsymbol{H}_{2,\cdot,\mathcal{I}}$ if we take all the rows in $\boldsymbol{H}_2$; $\boldsymbol{S}_{2,\mathcal{J},\mathcal{I}} \triangleq \boldsymbol{H}_{2,\mathcal{J},\mathcal{I}}\boldsymbol{H}_{2,\mathcal{J},\mathcal{I}}^*$; (31) is from (4); we impose $|\mathcal{J}| \leq |\mathcal{I}|$ in (32) since $\det(\boldsymbol{S}_{2,\mathcal{J},\mathcal{I}}) = 0$ otherwise.

By letting $\epsilon \to 0$, we obtain

$$\int_{\mathcal{U}^{n_2}} \det\left(\boldsymbol{Q}_{(2)}\right)_{\mathcal{I}}\mathrm{d}\boldsymbol{\theta} = \sum_{\substack{\mathcal{J}\subseteq\{1,\ldots,n_2\} \\ |\mathcal{J}|=|\mathcal{I}|}} \det(\boldsymbol{Q}_3)_{\mathcal{J}}\det(\boldsymbol{S}_{2,\mathcal{J},\mathcal{I}}).$$

Finally, we have the following upper and lower bounds on I_{T_2,T_1}

$$\begin{aligned} I^\star_{(3)}(\mathsf{snr}) + (n_1-1) + (n_2-1) &\geq I_{T_2,T_1} \\ &\geq \left(\frac{K_1-1}{K_1}\right)^{n_1-1}\left(\frac{K_2-1}{K_2}\right)^{n_2-1} I^\star_{(3)}(\mathsf{snr}) - 4 \qquad (33) \end{aligned}$$

with

$$I^\star_{(3)}(\mathsf{snr}) \triangleq \log \sum_{\substack{\mathcal{I}\subseteq\{1,\ldots,n_1\} \\ \mathcal{J}\subseteq\{1,\ldots,n_2\} \\ |\mathcal{I}|=|\mathcal{J}|}} \mathsf{snr}^{|\mathcal{I}|}\det(\boldsymbol{Q}_3)_{\mathcal{J}}\det(\boldsymbol{S}_{2,\mathcal{J},\mathcal{I}})\det(\boldsymbol{P}_1)_{\mathcal{I}}.$$

Note that to obtain the lower bound in (33), we used the fact that $\frac{K_1-1}{K_1} \leq 1$.

2) The General Case: In the general case, lower and upper bounds on $I_{T_{N-1},\ldots,T_1}$ are summarized by the following theorem.

Theorem 3: In an N-hop layered relay channel, the average mutual information of the proposed rotate-and-forward scheme is lower and upper bounded as

$$\begin{aligned} I^\star_{(N)}(\mathsf{snr}) + \sum_{i=1}^{N-1}(n_i-1) &\geq I_{T_{N-1},\ldots,T_1}(\mathsf{snr}) \\ &\geq \prod_{i=1}^{N-1}\left(\frac{K_i-1}{K_i}\right)^{n_i-1} I^\star_{(N)}(\mathsf{snr}) - 2(N-1) \end{aligned}$$

$$I_{T_{N-1},\dots,T_1} = \frac{1}{T}\sum_{t_{N-1}=1}^{T_{N-1}}\cdots\sum_{t_1=1}^{T_1}\log\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{H}_N\boldsymbol{\Delta}_{N-1,t_{N-1}}\boldsymbol{H}_{N-1}\cdots\boldsymbol{\Delta}_{1,t_1}\boldsymbol{H}_1\boldsymbol{H}_1^*\cdots\boldsymbol{\Delta}_{N-1,t_{N-1}}^*\boldsymbol{H}_N^*\right) \tag{29}$$

$$\begin{aligned} I_{T_2,T_1} &= \frac{1}{T}\sum_{t_2=1}^{T_2}\sum_{t_1=1}^{T_1}\log\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{H}_3\boldsymbol{\Delta}_{2,t_2}\boldsymbol{H}_2\boldsymbol{\Delta}_{1,t_1}\boldsymbol{H}_1\boldsymbol{H}_1^*\boldsymbol{\Delta}_{1,t_1}^*\boldsymbol{H}_2^*\boldsymbol{\Delta}_{2,t_2}^*\boldsymbol{H}_3^*\right) \\ &= \frac{1}{T_2}\sum_{t_2=1}^{T_2}\left(\frac{1}{T_1}\sum_{t_1=1}^{T_1}\log\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{\Delta}_{1,t_1}\boldsymbol{P}_1\boldsymbol{\Delta}_{1,t_1}^*\boldsymbol{Q}_{(2)}\right)\right) \end{aligned} \tag{30}$$

---

with

$$I^\star_{(N)}(\mathsf{snr}) \triangleq \log\left(\sum_{\substack{\mathcal{I}_i\subseteq\{1,\dots,n_i\},\ \forall i\\ |\mathcal{I}_i|=|\mathcal{I}_j|,\ \forall i\neq j}}\mathsf{snr}^{|\mathcal{I}_1|}\det(\boldsymbol{P}_1)_{\mathcal{I}_1} \times\det(\boldsymbol{Q}_N)_{\mathcal{I}_{N-1}}\prod_{i=2}^{N-1}\det\left(\boldsymbol{S}_{i,\mathcal{I}_i,\mathcal{I}_{i-1}}\right)\right). \tag{34}$$

Proof: We prove the theorem by induction in a straightforward manner. For $N=2$ and $N=3$, (34) holds. Now, let us suppose that (34) holds for N, we can rewrite

$$I_{T_N,\dots,T_1} = \frac{1}{T_N}\sum_{t_N=1}^{T_N} I_{T_{N-1},\dots,T_1,\cdot}$$

where $I_{T_{N-1},\dots,T_1,\cdot}$ is defined similarly as $I_{T_1,\cdot}$ in the previous case ($N=3$). Since (34) holds for N, it can be applied to bound $I_{T_{N-1},\dots,T_1}$. By repeating exactly the same steps in the case $N=3$, we can prove that (34) also holds for $N+1$. ■

B. DMT Analysis

Now, we prove that, in the multi-hop case with i.i.d. circularly symmetric complex Gaussian distributed channel coefficients, the rotate-and-forward scheme is DMT optimal when the number of antennas in each relay layer is equal to 2. To this end, we lower bound the expression (34) by only considering equal subsets as stated below.

$$I^\star_{(N)}(\mathsf{snr}) = \log\left(\sum_{\mathcal{I}\subseteq\{1,2\}}\mathsf{snr}^{|\mathcal{I}|}\det(\boldsymbol{P}_1)_{\mathcal{I}}\det(\boldsymbol{Q}_N)_{\mathcal{I}} \times\prod_{i=2}^{N-1}\det(\boldsymbol{S}_{i,\mathcal{I},\mathcal{I}})\right).$$

We can rewrite the terms in the $\log(\cdot)$ function explicity as

$$1+\mathsf{snr}\,\|\boldsymbol{h}_{1,1}\|^2\left(\prod_{i=2}^{N-1}|h_{i,11}|^2\right)\|\boldsymbol{h}_{N,1}\|^2 +\mathsf{snr}\|\boldsymbol{h}_{1,2}\|^2\left(\prod_{i=2}^{N-1}|h_{i,22}|^2\right)\|\boldsymbol{h}_{N,2}\|^2 +\mathsf{snr}^2\prod_{i=1}^{N}\|\boldsymbol{h}_{i,1}\|^2\,\|\boldsymbol{h}_{i,2}\|^2u_i$$

where $\boldsymbol{h}_{i,j}$ denotes the vector of the j th row of the channel matrix $\boldsymbol{H}_i$ for $1\le i\le N-1$, and $h_{i,jk}$ is the entry corresponding to the j th row and k th column of matrix $\boldsymbol{H}_i$. Note that $\sum_k|h_{i,jk}|^2=\|\boldsymbol{h}_{i,j}\|^2$; $\boldsymbol{h}_{N,j}$ denotes the j th column of matrix $\boldsymbol{H}_N$. Furthermore, the u_i are independent random variables uniformly distributed on the interval $[0,1]$.

Now, using the method of the previous sections, we can compute the DMT of this lower bound, which turns out to be optimal. Let us operate the following change of random variables:

$$\begin{aligned} u_i &= \mathsf{snr}^{-\gamma_i}, \\ \|\boldsymbol{h}_{i,j}\|^2 &= \mathsf{snr}^{-\alpha_{i,j}}, \\ |h_{i,jk}|^2 &= \mathsf{snr}^{-\beta_{i,jk}}. \end{aligned}$$

Using again the Laplace integration method, we obtain that the DMT of the lower bound is the solution of the following optimization problem:

$$d(r)=\min\left\{2(\alpha_{1,1}+\alpha_{1,2}+\alpha_{N,1}+\alpha_{N,2}) +\sum_{i=2}^{N-1}(\beta_{i,11}+\beta_{i,12}+\beta_{i,21}+\beta_{i,22})+\sum_{i=1}^{N}\gamma_i\right\}$$

subject to

$$\max\left\{0,\ 1-\alpha_{1,1}-\sum_{i=2}^{N-1}\beta_{i,11}-\alpha_{N,1},\ 1-\alpha_{1,2}-\sum_{i=2}^{N-1}\beta_{i,22}-\alpha_{N,2},\ 2-\alpha_{1,1}-\alpha_{1,2}-\alpha_{N,1}-\alpha_{N,2} -\sum_{i=2}^{N-1}\left(\min\{\beta_{i,11},\beta_{i,12}\}+\min\{\beta_{i,21},\beta_{i,22}\}\right)-\sum_{i=1}^{N}\gamma_i\right\}<r.$$

It is easy to check that the solution of this optimization problem leads to the optimal DMT $d_{2,2}(r)$. As an illustration, the dominating outage event for $r=0$ occurs, e.g., when $\alpha_{1,1}=\alpha_{1,2}=1$, which corresponds to the situation where all the entries of $\boldsymbol{H}_1$ are small. For $r=1$, outage occurs, e.g., when $\gamma_1=1$, which corresponds to the situation where $\boldsymbol{H}_1$ is essentially rank one.

Remark 6.1: The analysis performed here only applies to the case where each relay has 2 antennas. We believe that the result can be extended to the general case with arbitrary number of antennas at the relays, with the same conclusion. However, in order to establish this result, the knowledge of the joint distribution of the subdeterminants

$$\det(\boldsymbol{S}_{i,\mathcal{I},\mathcal{I}}) = \det(\boldsymbol{H}_{i,\mathcal{I},\mathcal{I}}\boldsymbol{H}^*_{i,\mathcal{I},\mathcal{I}})$$

for all $\mathcal{I} = \{1,\dots,k\}$, $k = 1,\dots,n_i$ would be required, which remains unfortunately out of reach beyond the case $n_i = 2$.

VII. Decode-and-Forward with Distributed Rotation

In the following, we show another application of the distributed rotation in the context of decode-and-forward relaying. As will be pointed out later on, the relaying complexity, especially the signaling, is reduced thanks to the distributed rotation.

A. Protocol Description

As in [6], [13], we consider a wireless channel with a single-antenna source-destination pair and multiple half-duplex relays

$$y_{\mathrm{R},i}[t] = g_i x[t] + z_{\mathrm{R},i}[t],$$
$$y_{\mathrm{D}}[t] = \sum_{i=1}^{n} h_i x_i[t] + z_{\mathrm{D}}[t]$$

where x, x_i, $y_{\mathrm{R},i}$, and y_{D} denote the transmitted signal from the source, transmitted signal from the i th relay, received signal at the i th relay, and received signal at the destination, respectively; g_i and h_i are the channel gains between the source and the i th relay and between the i th relay and the destination, respectively; $z_{\mathrm{R},i}$ and z_{D} are independent AWGN.

We propose a decode-and-forward scheme based on a fixed DRS $\{\boldsymbol{\Delta}_t\}$, $t = 1,\dots,K^n$. The two-slot protocol works as follows. The length of each slot is $T = K^n$ symbols time. During the first slot, the source broadcasts a codeword $\boldsymbol{x} \in \mathbb{C}^{T\times 1}$ that belongs to a code $\mathcal{C}_T$ with rate R bits per channel use (BPCU), i.e., $|\mathcal{C}_T| = 2^{TR}$. At the end of the first slot, each relay tries to decode the message. Let $\mathcal{D}$ denote the set of indices of succeeding relays and $\bar{\mathcal{D}}$ the failing ones. During the second slot, the failing relays remain silent. For each succeeding relay $i \in \mathcal{D}$, the transmitted signal is

$$x_i[t] = e^{j2\pi\varphi_{i,t}}\, x[t], \quad t = 1,\dots,T$$

where $\varphi_{i,t}$ defined as in Definition 1. The received signal at the destination is

$$\begin{aligned} y[t] &= \sum_{i\in\mathcal{D}} h_i x_i[t] + z_{\mathrm{D}}[t] \\ &= \tilde{h}_{\mathcal{D}}[t] x[t] + z_{\mathrm{D}}[t] \end{aligned}$$

with the equivalent fast fading channel gain

$$\tilde{h}_{\mathcal{D}}[t] \triangleq \boldsymbol{h}_{\mathcal{D}}^T\, \boldsymbol{\Delta}_{t,\mathcal{D}}\, \mathbf{1}_{|\mathcal{D}|}$$

where $\boldsymbol{h}_{\mathcal{D}} \in \mathbb{C}^{|\mathcal{D}|\times 1}$ is a vector of $\{h_i\}_{i\in\mathcal{D}}$; $\boldsymbol{\Delta}_{t,\mathcal{D}} \triangleq \operatorname{diag}\left\{e^{j2\pi\varphi_{i,t}},\ i\in\mathcal{D}\right\}$.

B. Outage Analysis

First, the end-to-end outage probability of the equivalent channel is

$$\begin{aligned} &\mathbb{P}\left(\frac{1}{2T}\sum_{t=1}^{T}\log\left(1+\mathsf{snr}\left|\tilde{h}_{\mathcal{D}}[t]\right|^2\right) < R\right) \\ &= \mathbb{P}\Bigg(\underbrace{\frac{1}{T}\sum_{t=1}^{T}\log\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{\Delta}_{t,\mathcal{D}}\mathbf{1}\cdot\mathbf{1}^*\boldsymbol{\Delta}^*_{t,\mathcal{D}}\boldsymbol{h}^*_{\mathcal{D}}\boldsymbol{h}^T_{\mathcal{D}}\right)}_{I_T(\mathsf{snr})} < 2R\Bigg) \\ &= \mathbb{P}\left(\mathcal{O}_{\mathcal{D}}(R)\right) \end{aligned}$$

where I_T has exactly the same form as defined in (3) with different channel matrices. We can thus reuse the results obtained before and get the following theorem.

Theorem 4: The diversity-multiplexing tradeoff of the proposed decode-and-forward scheme with distributed rotation is

$$d(r) = n(1-2r)^+,$$

for i.i.d. Rayleigh fading channel when $K\to\infty$.

Proof: The end-to-end outage probability can be developped as

$$\mathbb{P}\left(\mathcal{O}_{\mathcal{D}}(R)\right) = \sum_{\mathsf{D}\subseteq\{1,\dots,n\}} \mathbb{P}\left(\mathcal{O}_{\mathcal{D}}(R)\ \middle|\ \mathcal{D}=\mathsf{D}\right)\mathbb{P}\left(\mathcal{D}=\mathsf{D}\right).$$

As shown in [6], $\mathcal{D}=\mathsf{D}$ means that all $n-|\mathsf{D}|$ source-relay channels are in outage. Thus, we have

$$\mathbb{P}\left(\mathcal{D}=\mathsf{D}\right) \doteq \mathsf{snr}^{-(n-|\mathsf{D}|)(1-2r)^+},\ \forall\,\mathsf{D}. \tag{35}$$

Now, we would like to show

$$\mathbb{P}\left(\mathcal{O}_{\mathcal{D}}(R)\ \middle|\ \mathcal{D}=\mathsf{D}\right) \doteq \mathsf{snr}^{-|\mathsf{D}|(1-2r)^+}. \tag{36}$$

To this end, we apply Corollary 1, (6), and (7) for $K\to\infty$, and we have

$$\mathbb{P}\left(\mathcal{O}_{\mathcal{D}}(R)\ \middle|\ \mathcal{D}=\mathsf{D}\right) \doteq \mathbb{P}\left(I^\star(\mathsf{snr}) < 2R\right) \tag{37}$$

where $I^\star$ is defined in (6) with $\boldsymbol{P} = \mathsf{snr}\,\mathbf{1}_{\mathcal{D}}\mathbf{1}^*_{\mathcal{D}}$ and $\boldsymbol{Q} = \boldsymbol{h}^*_{\mathcal{D}}\boldsymbol{h}^T_{\mathcal{D}}$. Since $\boldsymbol{P}$ is of rank one, it follows that

$$I^\star(\mathsf{snr}) = \log\left(1+\mathsf{snr}\|\boldsymbol{h}_{\mathcal{D}}\|^2\right). \tag{38}$$

Plugging (38) into (37), (36) is straightforward. Finally, from (35) and (36), the theorem is proved. ■

The proposed scheme is a fixed relaying scheme in that relaying functions are decided before any communication and do not depend on the channel condition. Hence, it is a flexible cooperation scheme. Theorem 4 says that same DMT performance as the distributed space-time coding proposed in [6] is achieved. As opposed to conventional distributed space-time coding scheme, no signaling on the decoding status of each relay is needed. As a matter of fact, the destination only need to know the equivalent scalar channel gain $\tilde{h}_{\mathcal{D}}[t]$. The latter can be estimated as in any fast fading channel.

Remark 7.1: Although we only consider the orthogonal DF scheme in this paper, the DRS can be applied to the non-orthogonal DF scheme with little modification. With two-slot non-orthogonal DF scheme, the source transmits a codeword of length $2T$ composed of two parts $\begin{bmatrix}\boldsymbol{x}_1 \ \boldsymbol{x}_2\end{bmatrix}$. After receiving $\boldsymbol{x}_1$, the relays try to decode the message. For the succeeding relays, $\boldsymbol{x}_2$ can be anticipated and the DRS is applied to $\boldsymbol{x}_2$. In this way, $\boldsymbol{x}_2$ goes through an artificial fast fading channel as in the orthogonal case. Also note that exactly the same idea can be applied to the dynamic decode-and-forward scheme [10] with multiple relays.

Remark 7.2: Since the destination and the source do not need to know the existence of the relays and that the performance can only be improved with the presence of the relays, the proposed scheme in the non-orthogonal case is an *oblivious* cooperative scheme [29].

VIII. Conclusions

We have proposed a framework of distributed rotation for cooperative relaying and shown that even simple time-varying linear processing can recover spatial diversity. The framework has been applied to both linear and nonlinear relaying schemes. Thanks to the tractability of the proposed schemes, we have proved that the optimal diversity-multiplexing tradeoff of some non trivial channel setting can be achieved with linear relaying. Furthermore, an oblivious decode-and-forward scheme based on distributed rotation has been proposed.

Acknowledgement

The authors would like to thank Emre Telatar for inspiring comments at the early stage of this work.

Appendix

A. Proofs of Lemma 1

To prove the lemma, we first note that

$$\begin{aligned}
&\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{\theta}}^*\boldsymbol{Q}\right)\\
&\quad=\sum_{\mathcal{I}\subseteq\{1,\ldots,n\}}\mathsf{snr}^{|\mathcal{I}|}\det\left(\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{\theta}}^*\boldsymbol{Q}\right)_{\mathcal{I}}\\
&\quad=\sum_{\mathcal{I}\subseteq\{1,\ldots,n\}}\mathsf{snr}^{|\mathcal{I}|}\det\left(\boldsymbol{R}_{\boldsymbol{\theta}_{\mathcal{I}}}\boldsymbol{P}_{\mathcal{I},\cdot}\boldsymbol{R}_{\boldsymbol{\theta}}^*\boldsymbol{Q}_{\cdot,\mathcal{I}}\right) &(39)\\
&\quad=\sum_{\mathcal{I}\subseteq\{1,\ldots,n\}}\mathsf{snr}^{|\mathcal{I}|}\prod_{i\in\mathcal{I}}e^{j2\pi\theta_i}\det\left(\boldsymbol{P}_{\mathcal{I},\cdot}\boldsymbol{R}_{\boldsymbol{\theta}}^*\boldsymbol{Q}_{\cdot,\mathcal{I}}\right) &(40)
\end{aligned}$$

where $\boldsymbol{R}_{\boldsymbol{\theta}_{\mathcal{I}}}\triangleq\operatorname{diag}\left\{e^{j2\pi\theta_i}, i\in\mathcal{I}\right\}$; $\boldsymbol{P}_{\mathcal{I},\cdot}$ denotes the submatrix of $\boldsymbol{P}$ formed by the rows indexed by $\mathcal{I}$; $\boldsymbol{Q}_{\cdot,\mathcal{I}}$ is similarly defined with the columns. In fact, one can show that, for any $m\le n$, $\det\left(\boldsymbol{A}_{m\times n}\boldsymbol{R}_{\boldsymbol{\theta}}^*\boldsymbol{B}_{n\times m}\right)$ is a multilinear function of $e^{-j2\pi\theta_i}$, $i=1,\ldots,n$. This can be proved by induction on n, which we do not detail here. As such, we can deduce from (40) that $\det\left(\mathbf{I}+\mathsf{snr}\boldsymbol{R}_{\boldsymbol{\theta}}\boldsymbol{P}\boldsymbol{R}_{\boldsymbol{\theta}}^*\boldsymbol{Q}\right)$ is a multilinear function of $e^{\pm j2\pi\theta_i}$, $i=1,\ldots,n$, which completes the proof for Property 1. Property 2 is a direct consequence of Property 1. By integrating θ_i over $\mathcal{U}=[0,1)$, terms containing θ_i in the original multilinear function disappear. Hence, a new multilinear function of the rest of the θ's is obtained.

To prove Property 3, it is enough to show that

$$\int_{\mathcal{U}^n}\det\left(\boldsymbol{R}_{\boldsymbol{\theta}_{\mathcal{I}}}\boldsymbol{P}_{\mathcal{I},\cdot}\boldsymbol{R}_{\boldsymbol{\theta}}^*\boldsymbol{Q}_{\cdot,\mathcal{I}}\right)\mathrm{d}\boldsymbol{\theta}=\det\left(\boldsymbol{P}_{\mathcal{I}}\right)\det\left(\boldsymbol{Q}_{\mathcal{I}}\right) \quad (41)$$

from which and (39) we obtain (4). In order to prove (41), we rewrite

$$\boldsymbol{R}_{\boldsymbol{\theta}_{\mathcal{I}}}\boldsymbol{P}_{\mathcal{I},\cdot}\boldsymbol{R}_{\boldsymbol{\theta}}^*\boldsymbol{Q}_{\cdot,\mathcal{I}}=\boldsymbol{R}_{\boldsymbol{\theta}_{\mathcal{I}}}\left(\boldsymbol{P}_{\mathcal{I}}\boldsymbol{R}_{\boldsymbol{\theta}_{\mathcal{I}}}^*\boldsymbol{Q}_{\mathcal{I}}+\boldsymbol{P}_{\mathcal{I},\bar{\mathcal{I}}}\boldsymbol{R}_{\boldsymbol{\theta}_{\bar{\mathcal{I}}}}^*\boldsymbol{Q}_{\bar{\mathcal{I}},\mathcal{I}}\right). \quad (42)$$

Then, we use the following equalities

$$\det\left(\boldsymbol{A}+e^{j2\pi\theta_i}\boldsymbol{u}\,\boldsymbol{v}^*\right)=\det\left(\boldsymbol{A}\right)+e^{j2\pi\theta_i}\boldsymbol{u}\left(\operatorname{adj}\left(\boldsymbol{A}\right)\right)\boldsymbol{v}^*,$$

$\forall\,\boldsymbol{A},\boldsymbol{u},\boldsymbol{v}$, where $\operatorname{adj}\left(\boldsymbol{A}\right)$ is the adjugate matrix of $\boldsymbol{A}$, to show that

$$\int_0^1\det\left(\boldsymbol{A}+e^{j2\pi\theta_i}\boldsymbol{u}\,\boldsymbol{v}^*\right)\mathrm{d}\theta_i=\det\left(\boldsymbol{A}\right). \quad (43)$$

Finally, it follows that

$$\begin{aligned}
&\int_{\mathcal{U}^n}\det\left(\boldsymbol{R}_{\boldsymbol{\theta}_{\mathcal{I}}}\boldsymbol{P}_{\mathcal{I},\cdot}\boldsymbol{R}_{\boldsymbol{\theta}}^*\boldsymbol{Q}_{\cdot,\mathcal{I}}\right)\mathrm{d}\boldsymbol{\theta}\\
&=\int_{\mathcal{U}^{|\mathcal{I}|}}\det\left(\boldsymbol{R}_{\boldsymbol{\theta}_{\mathcal{I}}}\right)\\
&\quad\times\left(\int_{\mathcal{U}^{|\bar{\mathcal{I}}|}}\det\left(\boldsymbol{P}_{\mathcal{I}}\boldsymbol{R}_{\boldsymbol{\theta}_{\mathcal{I}}}^*\boldsymbol{Q}_{\mathcal{I}}+\boldsymbol{P}_{\mathcal{I},\bar{\mathcal{I}}}\boldsymbol{R}_{\boldsymbol{\theta}_{\bar{\mathcal{I}}}}^*\boldsymbol{Q}_{\bar{\mathcal{I}},\mathcal{I}}\right)\mathrm{d}\boldsymbol{\theta}_{\bar{\mathcal{I}}}\right)\mathrm{d}\boldsymbol{\theta}_{\mathcal{I}}\\
&=\int_{\mathcal{U}^{|\mathcal{I}|}}\det\left(\boldsymbol{R}_{\boldsymbol{\theta}_{\mathcal{I}}}\right)\det\left(\boldsymbol{P}_{\mathcal{I}}\boldsymbol{R}_{\boldsymbol{\theta}_{\mathcal{I}}}^*\boldsymbol{Q}_{\mathcal{I}}\right)\mathrm{d}\boldsymbol{\theta}_{\mathcal{I}} &(44)\\
&=\det\left(\boldsymbol{P}_{\mathcal{I}}\right)\det\left(\boldsymbol{Q}_{\mathcal{I}}\right)
\end{aligned}$$

where (44) is obtained by applying (42) and (43). More precisely, we rewrite

$$\boldsymbol{P}_{\mathcal{I},\bar{\mathcal{I}}}\boldsymbol{R}_{\boldsymbol{\theta}_{\bar{\mathcal{I}}}}^*\boldsymbol{Q}_{\bar{\mathcal{I}},\mathcal{I}}=\sum_{i\in\bar{\mathcal{I}}}e^{-j2\pi\theta_i}\boldsymbol{u}_i\boldsymbol{v}_i^*$$

with $\boldsymbol{u}_i$ and $\boldsymbol{v}_i^*$ the ith column of $\boldsymbol{P}_{\mathcal{I},\bar{\mathcal{I}}}$ and the ith row of $\boldsymbol{Q}_{\bar{\mathcal{I}},\mathcal{I}}$, respectively, and then apply (43) successively with all θ_i in $\boldsymbol{\theta}_{\bar{\mathcal{I}}}$.

References

[1] J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[2] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun., ETT*, vol. 10, no. 6, pp. 585–596, Nov. 1999.

[3] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[4] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—Part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[5] ——, "User cooperation diversity—Part II: Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

[6] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[7] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[8] T. Hunter, S. Sanayei, and A. Nosratinia, "Outage analysis of coded cooperation," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 375–391, Feb. 2006.

[9] R. U. Nabar, H. Bölcskei, and F. W. Kneubühler, "Fading relay channels: Performance limits and space-time signal design," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.

[10] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4152–4172, Dec. 2005.

[11] S. Yang and J.-C. Belfiore, "Towards the optimal amplify-and-forward cooperative diversity scheme," *IEEE Trans. Inf. Theory*, vol. 53, no. 9, pp. 3114–3126, Sep. 2007.

[12] M. Yuksel and E. Erkip, “Multi-antenna cooperative wireless systems: A diversity multiplexing tradeoff perspective,” *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3371–3393, Oct. 2007.

[13] Y. Jing and B. Hassibi, “Distributed space-time coding in wireless relay networks,” *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3524–3536, Dec. 2006.

[14] S. Yang and J.-C. Belfiore, “Diversity of MIMO multi-hop relay channels,” Aug. 2007. Internal Report. [Online]. Available: http://arxiv.org/abs/0708.0386

[15] P. Elia, K. Vinodh, M. Anand, and P. V. Kumar, “D-MG tradeoff and optimal codes for a class of AF and DF cooperative communication protocols,” *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3161–3185, Jul. 2009.

[16] T. M. Cover and A. El Gamal, “Capacity theorems for the relay channel,” *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[17] G. Kramer, M. Gastpar, and P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[18] S. Borade, L. Zheng, and R. Gallager, “Amplify and forward in wireless relay networks: Rate, diversity and network size,” *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3302–3318, Oct. 2007.

[19] S. Avestimehr, S. Diggavi, and D. N. C. Tse, “Wireless network information flow: A deterministic approach,” *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1872–1905, Jun. 2011.

[20] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, “Noisy network coding,” *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3132–3152, May 2011.

[21] R. Kolte and A. Özgür, “Generalized diversity-multiplexing tradeoff of half-duplex relay networks,” in *IEEE Int. Symp. Inf. Theory*, Jul. 2013, pp. 2755–2759.

[22] S. Tavildar and P. Viswanath, “Approximately universal codes over slow fading channels,” *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3233–3258, Jul. 2006.

[23] T. M. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[24] S. Yang and J.-C. Belfiore, “Optimal space-time codes for the MIMO amplify-and-forward cooperative channel,” *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 647–663, Feb. 2007.

[25] R. Pedarsani, O. Lévêque, and S. Yang, “Flip-and-forward achieves the optimal diversity-multiplexing tradeoff for the two-hop MIMO relay channel, with two relay antennas,” in *Proc. CROWNCOM*, Cannes, France, 2010.

[26] W. Gröbner and N. Hofreiter, *Integraltafel, Teil II, Bestimmte Integrale*. Wien and Innsbruck: Springer-Verlag, 1958.

[27] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products (Seventh Edition)*. Elsevier Academic Press, 2007.

[28] R. V. Hogg and A. T. Craig, *Introduction to mathematical statistics*. Macmilan, 1987.

[29] M. Katz and S. Shamai (Shitz), “Cooperative schemes for a source and an occasional nearby relay in wireless networks,” *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5138–5160, Nov. 2009.