Diversity-Multiplexing Tradeoff in Multiple-Relay Network–Part II: Multiple-Antenna Networks

Shahab Oveis Gharan, Alireza Bayesteh and Amir K. Khandani
Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, ON, Canada, N2L 3G1
email: {shahab,alireza,khandani}@cst.uwaterloo.ca

***Abstract*—This paper studies the setup of a multiple-relay network in which K half-duplex multiple-antenna relays assist in the transmission between a/several multiple-antenna source(s) and a multiple-antenna destination. Each two nodes are assumed to be either connected through a quasi-static Rayleigh fading channel, or disconnected. This paper is comprised of two parts. In this part of the paper, we study multiple-antenna multiple-relay network. We prove that the *Random Sequential* (RS) scheme proposed in [1], [2] achieves the maximum diversity gain in a general multiple-antenna multiple-relay network. Moreover, we show that utilizing independent random unitary matrix multiplication at the relay nodes enables the RS scheme to achieve better diversity-multiplexing tradeoff (DMT) results comparing with the traditional amplify-and-forward relaying. Indeed, using the RS scheme, we derive a new achievable DMT for the MIMO parallel relay network. Interestingly, it turns out that the DMT of the RS scheme is optimum for the MIMO half-duplex parallel 2-relay ($K = 2$) setup. Finally, we show that utilizing random unitary matrix multiplication also improves the DMT of the Non-Orthogonal amplify-and-forward relaying scheme of [3] in the MIMO single relay channel.**

I. Introduction

Recently, cooperative schemes and protocols have been proposed as candidates to exploit the spatial diversity offered by the relay networks (for example, see [3]–[7]). *Decode-and-Forward* (DF), *Amplify-and-Forward* (AF) and *Compress-and-Forward* (CF) relaying are the main relaying strategies utilized in the proposed relaying schemes. While DF and CF strategies are utilized in small-scale networks to obtain capacity results (for example, see [8]–[10]), the AF relaying turns out to be more suitable to exploit the cooperative diversity (for example, see [3], [4], [6], [11]).

In AF relaying, the relays are not supposed to decode the transmitted message. Hence, the relays consume less computing power and the end-to-end system expends much smaller amount of delay comparing with the other relaying strategies. Accordingly, AF relaying schemes are more suitable for the practical situations. Moreover, against the DF relaying, the AF relaying performance is not limited by the source-to-relay channel quality while, unlike the CF relaying, the parallel relays can yet exploit the power boosting advantage by coherently forwarding their received signals to the destination. Hence, they can asymptotically perform optimal in large scale parallel networks (see [7], [12], [13]).

While AF relaying is investigated well in the cooperative single-antenna networks, much is unknown about its potential for the multiple-antenna counterpart. Indeed, unlike the single-antenna scenario, in this case the AF multipliers are matrices rather than scalars. Hence, finding the optimum AF matrices becomes challenging.

A fundamental measure to evaluate the performance of the existing cooperative diversity schemes is the diversity-multiplexing tradeoff (DMT) which was first introduced by Zheng and Tse in the context of point-to-point MIMO fading channels [14]. Roughly speaking, the diversity-multiplexing tradeoff identifies the optimal compromise between the “transmission reliability” and the“data rate” in the high-SNR regime.

The non-orthogonal amplify-and-forward (NAF) scheme, first proposed by Nabar *et al.* in [15], has been further studied by Azarian *at al.* in [3] for the single-antenna multiple-relay setup. In addition to analyzing the DMT of the NAF scheme, reference [3] shows that NAF is the best in the class of AF strategies for single-antenna single-relay systems.

Recently, Yang and Belfiore in [11] study the DMT performance of the NAF scheme for the multiple-antenna parallel relay setup. Moreover, based on the non-vanishing determinant criterion, the authors construct a family of space-time code for the NAF scheme over MIMO setup. However, as shown in [11], the NAF scheme falls far from the DMT upper-bound in the multiple-antenna setup specially for small values of multiplexing gain. Indeed, even for the case of MIMO 2-hop single-relay setup, the NAF scheme is unable to achieve the maximum diversity gain of the system.

Yuksel *et al.* in [5] apply compress-and-forward (CF) strategy and show that CF achieves the DMT upper-bound for the multiple-antenna half-duplex single-relay system. However, in their proposed scheme, the relay node needs to know the CSI of all the channels in the network which may not be practical.

In this paper, we investigate the potential benefits of amplify-and-forward relaying in the multiple-antenna multiple-relay networks. For this purpose, we study the random sequential scheme proposed in [1], [2]. The key elements of the proposed scheme are: 1) signal transmission through sequential paths in the network, 2) path timing such that no non-causal interference is caused from the source of the future paths on the receiver of the current path, 3) multiplication by a random unitary matrix at each relay node, and 4) no signal boosting in amplify-and-forward relaying at the relay nodes, i.e. the received signal is amplified by a coefficient with the absolute value of at most 1. We prove that this scheme achieves the maximum diversity gain in a general multiple-antenna multiple-relay network. Furthermore, we derive the DMT of the RS scheme for multiple-antenna multiple-relay network. To accomplish this problem, we first study the full-duplex multiple-antenna 2-hop network with a single relay. We show that against the traditional AF relaying, the RS scheme can achieve the optimum DMT. Indeed, using the traditional AF relaying, there exists a chance that the eigenvectors corresponding to the large eignenvalues of the incoming channel matrix of the relay projects to the eigenvectors corresponding to the small eignenvalues of the relay's outgoing channel matrix. This event degrades the performance of traditional AF relaying in the MIMO setup. However, in the RS scheme, utilizing the random unitary matrix multiplication at the relay nodes for different time-slots, such an event is much more unlikely to happen. This fact will be elaborated more throughout the paper. Next, we study the MIMO parallel half-duplex relay network and by deriving the DMT, we show that the RS scheme improves the DMT of the traditional AF relaying schemes. Interestingly, it turns out that the DMT of the RS scheme is optimum for the MIMO half-duplex parallel 2-relay ($K = 2$) setup. Finally, we show that utilizing random unitary matrix multiplication also improves the DMT of the Non-Orthogonal amplify-and-forward relaying scheme of [3] in the MIMO single half-duplex relay channel.

For description of the system model and the RS scheme, the reader is recommended to see [2] or [1]. The rest of the paper is organized as follows. In section II, the proof of maximum diversity achievability of the RS scheme is explained and section III is dedicated to DMT analysis of the RS scheme in the multiple-antenna setup.

II. Maximum Diversity Achievability Proof in General Multi-Hop Multiple-Antenna Scenario

In this section, we consider our proposed RS scheme and prove that it achieves the maximum diversity gain between two end-points in a general multiple-antenna multi-hop network (no additional constraints imposed).

Theorem 1 *Consider a relay network with the connectivity graph $G = (V, E)$ and K relays, in which each two adjacent nodes are connected through a Rayleigh-fading channel. Assume that all the network nodes are equipped with multiple antenna. Then, by properly choosing the path sequence, the proposed RS scheme achieves the maximum diversity gain of the network which is equal to*

$$d_G = \min_{\mathcal{S}} w_G(\mathcal{S}), \tag{1}$$

where $\mathcal{S}$ is a cut-set on G[1].

Proof: First, we have to show that d_G is indeed an upper-bound on the diversity-gain of the network. This is shown in [2]. Now, we prove that this bound is indeed achievable by the RS scheme. First, we provide the path sequence needed to achieve the maximum diversity gain. Consider the graph $\hat{G} = (V, E, w)$ with the same set of vertices and edges as the graph G and the weight function w on the edges as $w_{\{a,b\}} = N_a N_b$. Consider the maximum-flow algorithm [16] on $\hat{G}$ from the source node 0 to the sink node $K + 1$. Since the weight function is integer over the edges, according to the Ford-Fulkerson Theorem [16], one can achieve the maximum flow which is equal to the minimum cut of $\hat{G}$ or d_G by the union of elements of a sequence $(\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_{d_G})$ of paths ($L = d_G$). We show that this family of paths are sufficient to achieve the maximum diversity.

Noting that the received signal at each node is multiplied by a random isotropically distributed unitary matrix, at the destination side we have

$$\mathbf{y}_{K+1,i} = \mathbf{H}_{K+1,\mathrm{p}_i(l_i-1)}\alpha_{i,l_i-1}\mathbf{U}_{i,l_i-1}\mathbf{H}_{\mathrm{p}_i(l_i-1),\mathrm{p}_i(l_i-2)}\times \alpha_{i,l_i-2}\mathbf{U}_{i,l_i-2}\cdots\alpha_{i,1}\mathbf{U}_{i,1}\mathbf{H}_{\mathrm{p}_i(1),0}\mathbf{x}_{0,i} + \sum_{j<i}\mathbf{X}_{i,j}\mathbf{x}_{0,j} + \sum_{j\leq i, m\leq l_j}\mathbf{Q}_{i,j,m}\mathbf{n}_{j,m}. \tag{2}$$

Here, $\mathbf{x}_{0,i}$ is the vector transmitted at the source side during the $s_{i,1}$'th slot as the input for the i'th path, $\mathbf{y}_{K+1,i}$ is the vector received at the destination side during the s_{i,l_i}'th slot as the output for i'th path, $\mathbf{U}_{i,j}$ denotes the multiplied unitary matrix at the $\mathrm{p}_i(j)$'th node of the ith path, $\mathbf{X}_{i,j}$ is the interference matrix which relates the input of the j'th path ($j < i$) to the output of the i'th path, $\mathbf{n}_{j,m}$ is the noise vector during the $s_{j,m}$'th slot at the $\mathrm{p}_j(m)$'th node of the network, and finally, $\mathbf{Q}_{i,k,m}$ is the matrix which relates $\mathbf{n}_{k,m}$ to $\mathbf{y}_{K+1,i}$. Notice that as the timing sequence satisfies the noncausal interference assumption, the summation terms in (2) do not exceed i. Defining $\mathbf{x}(s) = \left[\mathbf{x}_{0,1}^T(s)\,\mathbf{x}_{0,2}^T(s)\cdots\mathbf{x}_{0,L}^T(s)\right]^T$, $\mathbf{y}(s) = \left[\mathbf{y}_{K+1,1}^T(s)\,\mathbf{y}_{K+1,2}^T(s)\cdots\mathbf{y}_{K+1,L}^T(s)\right]^T$, and $\mathbf{n}(s) = \left[\mathbf{n}_{1,1}^T(s)\,\mathbf{n}_{1,2}^T(s)\cdots\mathbf{n}_{L,l_L}^T(s)\right]^T$, we have the following equivalent block lower-triangular matrix between the end nodes

$$\mathbf{y}(s) = \mathbf{H}_T\mathbf{x}(s) + \mathbf{Q}\mathbf{n}(s). \tag{3}$$

Here,

$$\mathbf{H}_T = \begin{pmatrix} \mathbf{X}_{1,1} & \mathbf{0} & \mathbf{0} & \ldots \\ \mathbf{X}_{2,1} & \mathbf{X}_{2,2} & \mathbf{0} & \ldots \\ \vdots & \vdots & \vdots & \ddots \\ \mathbf{X}_{L,1} & \mathbf{X}_{L,2} & \ldots & \mathbf{X}_{L,L} \end{pmatrix},$$

where

$$\mathbf{X}_{i,i} = \mathbf{H}_{K+1,\mathrm{p}_i(l_i-1)}\alpha_{i,l_i-1}\mathbf{U}_{i,l_i-1}\mathbf{H}_{\mathrm{p}_i(l_i-1),\mathrm{p}_i(l_i-2)}\times \alpha_{i,l_i-2}\mathbf{U}_{i,l_i-2}\cdots\alpha_{i,1}\mathbf{U}_{i,1}\mathbf{H}_{\mathrm{p}_i(1),0},$$

[1] cut-set and the weight of a cut-set ($w_G(\mathcal{S})$) are defined in [1], [2]

and

$$\mathbf{Q}=\begin{pmatrix} \mathbf{Q}_{1,1,1} & \dots & \mathbf{Q}_{1,1,l_1} & \mathbf{0} & \mathbf{0} & \dots \\ \mathbf{Q}_{2,1,1} & \dots & \mathbf{Q}_{2,1,l_1} & \dots & \mathbf{Q}_{2,2,l_2} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ \mathbf{Q}_{L,1,1} & \mathbf{Q}_{L,1,2} & \dots & \dots & \dots & \mathbf{Q}_{L,L,l_L} \end{pmatrix}.$$

Having (3), the outage probability can be written as

$$\mathbb{P}\{\mathcal{E}\} \quad = \quad \mathbb{P}\left\{\left|\mathbf{I}_L + P\mathbf{H}_T\mathbf{H}_T^H\mathbf{P}_n^{-1}\right| < 2^{SR}\right\}, \qquad (4)$$

where $\mathbf{P}_n = \mathbf{Q}\mathbf{Q}^H$. First, similar to the proof of theorem 3 in [2], we can easily show that $\alpha_{i,j} \doteq 1$ with probability 1[2], and also show that there exists a constant c which depends just on the topology of graph G and the path sequence such that $\mathbf{P}_n \preccurlyeq c\mathbf{I}_L$. Assume that for each $\{a,b\} \in E$, $\lambda_{\max}(\mathbf{H}_{a,b}) = P^{-\mu_{\{a,b\}}}$, where $\lambda_{\max}(\mathbf{A})$ denotes the greatest eigenvalue of $\mathbf{A}\mathbf{A}^H$. Also, assume that

$$\begin{aligned} \gamma_{i,j} \quad &\triangleq \quad \left|\mathbf{v}_{r,\max}^H\left(\mathbf{H}_{\{\mathrm{p}_i(j+1),\mathrm{p}_i(j)\}}\right)\mathbf{U}_{i,j}\times\right. \\ &\qquad \left.\mathbf{v}_{l,\max}\left(\mathbf{H}_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}}\mathbf{U}_{i,j-1}\dots\mathbf{H}_{\{\mathrm{p}_i(1),0\}}\right)\right|^2 \\ &= \quad P^{-\nu_{i,j}}, \end{aligned} \qquad (5)$$

where $\mathbf{v}_{l,\max}(\mathbf{A})$ and $\mathbf{v}_{r,\max}(\mathbf{A})$ denote the left and the right eigenvectors of $\mathbf{A}$ corresponding to $\lambda_{\max}(\mathbf{A})$, respectively. The outage probability can be upper-bounded as

$$\begin{aligned} \mathbb{P}\{\mathcal{E}\} &\overset{(a)}{\leq} \mathbb{P}\left\{\lambda_{\max}\left(\left(\mathbf{H}_T\mathbf{H}_T^H\mathbf{P}_n^{-1}\right)^{\frac{1}{2}}\right) \leq \left(2^{SR}-1\right)P^{-1}\right\} \\ &\overset{(b)}{\leq} \mathbb{P}\left\{\lambda_{\max}(\mathbf{H}_T) \leq c\left(2^{SR}-1\right)P^{-1}\right\} \\ &\overset{(c)}{\leq} \mathbb{P}\left\{\bigcap_{i=1}^{L}\left(\lambda_{\max}(\mathbf{X}_{i,i}) \leq c\left(2^{SR}-1\right)P^{-1}\right)\right\} \\ &\overset{(d)}{\leq} \mathbb{P}\left\{\bigcap_{i=1}^{L}\left(\sum_{j=1}^{l_i}\mu_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}} + \sum_{j=1}^{l_i-1}\nu_{i,j} \geq \right.\right. \\ &\qquad \left.\left. 1 - \log\frac{c\left(2^{SR}-1\right)}{P}\right)\right\} \\ &\overset{(e)}{\doteq} \mathbb{P}\left\{\bigcap_{i=1}^{L}\left(\sum_{j=1}^{l_i}\mu_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}} + \sum_{j=1}^{l_i-1}\nu_{i,j} \geq 1\right)\right\}. \end{aligned} \qquad (6)$$

In the above equation, (a) follows from the fact that $1 + \lambda_{\max}(\mathbf{A}^{\frac{1}{2}}) \leq |\mathbf{I}+\mathbf{A}|$, for a positive semi-definite matrix $\mathbf{A}$. (b) results from $\mathbf{P}_n \preccurlyeq c\mathbf{I}_L$. (c) follows from the fact that $\lambda_{\max}(\mathbf{H}_T) \geq \max_i \lambda_{\max}(\mathbf{X}_{i,i})$. To obtain (d), we first notice that according to Lemma 2 which is described in the sequel, we have

$$\lambda_{\max}(\mathbf{AUB}) \geq \lambda_{\max}(\mathbf{A})\lambda_{\max}(\mathbf{A})\left|\mathbf{v}_{r,\max}^H(\mathbf{A})\mathbf{U}\mathbf{v}_{l,\max}(\mathbf{B})\right|^2$$

By recursively applying the above inequality, it follows that

$$\begin{aligned} \lambda_{\max}(\mathbf{X}_{i,i}) \quad &\geq \quad \lambda_{\max}\left(\mathbf{H}_{K+1,\mathrm{p}_i(l_i-1)}\right)\gamma_{i,l_i-1}\times \\ &\qquad \lambda_{\max}\left(\mathbf{H}_{\mathrm{p}_i(l_i-1),\mathrm{p}_i(l_i-2)}\right)\gamma_{i,l_i-2}\times \\ &\qquad \cdots\gamma_{i,1}\lambda_{\max}\left(\mathbf{H}_{\mathrm{p}_i(1),0}\right) \\ &= \quad \prod_{j=1}^{l_i}\lambda_{\max}\left(\mathbf{H}_{\mathrm{p}_i(j),\mathrm{p}_i(j-1)}\right)\prod_{j=1}^{l_i-1}\gamma_{i,j}. \end{aligned} \qquad (7)$$

Noting the definitions of $\mu_{\{i,j\}}$ and $\nu_{i,j}$, (d) easily follows. Finally, (e) results from the fact that as $P \to \infty$, the term $\log\frac{c\left(2^{SR}-1\right)}{P}$ can be ignored.

Since the left and the right unitary matrices resulting from the SVD of an i.i.d. complex Gaussian matrix are independent of its singular value matrix [17] and $\mathbf{U}_{i,j}$ is an independent isotropically distributed unitary matrix, we conclude that all the random variables in the set $\left\{\{\mu_e\}_{e\in E}, \{\nu_{i,j}\}_{1\leq i\leq L, 1\leq j<l_i}\right\}$ are mutually independent. From the probability distribution analysis of the singular values of circularly symmetric Gaussian matrices in [14], we can easily prove $\mathbb{P}\left\{\mu_e \geq \mu_e^0\right\} \doteq P^{-N_aN_b\mu_e^0} = P^{-w_e\mu_e^0}$. Similarly, as $\mathbf{U}_{i,j}$ is isotropically distributed, it can be shown that $\mathbb{P}\{\nu(i,j) \geq \nu_0(i,j)\} \doteq P^{-\nu_0(i,j)}$. Hence, defining $\boldsymbol{\mu} = [\mu_e]_{e\in E}^T$, $\boldsymbol{\nu} = [\nu_{i,j}]_{1\leq i\leq L,1\leq j<l_i}^T$, and $\mathbf{w} = [w_e]_{e\in E}$, we have

$$\mathbb{P}\{\boldsymbol{\mu} \geq \boldsymbol{\mu}_0, \boldsymbol{\nu} \geq \boldsymbol{\nu}_0\} \doteq P^{-(\mathbf{1}\cdot\boldsymbol{\nu}+\mathbf{w}\cdot\boldsymbol{\mu})}. \qquad (8)$$

Let us define $\mathcal{R}$ as the region in $\mathbb{R}^{|E|+\sum_{i=1}^{L}l_i-L}$ of the vectors $\left[\boldsymbol{\mu}^T\boldsymbol{\nu}^T\right]^T$ such that for all $1 \leq i \leq L$, we have $\sum_{j=1}^{l_i}\mu_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1)\}} + \sum_{j=1}^{l_i-1}\nu_{i,j} \geq 1$. Using the same argument as in the proof of Theorem 3 in [2], it can easily be shown that $\mathbb{P}\{\mathcal{R}\} = \mathbb{P}\left\{\mathcal{R}\bigcap\mathbb{R}_+^{|E|+\sum_{i=1}^{L}l_i-L}\right\}$. Hence, defining $\mathcal{R}_+ = \mathcal{R}\bigcap\mathbb{R}_+^{|E|+\sum_{i=1}^{L}l_i-L}$ and $d_0 = \min_{[\boldsymbol{\mu}^T\boldsymbol{\nu}^T]^T\in\mathcal{R}_+}\mathbf{w}\cdot\boldsymbol{\mu} + \mathbf{1}\cdot\boldsymbol{\nu}$, which can easily be verified to be bounded, and applying the argument of Lemma 4 which is explained in the sequel, we have

$$\mathbb{P}\{\mathcal{E}\} \dot{\leq} \mathbb{P}\{\mathcal{R}_+\} \doteq P^{-d_0}. \qquad (9)$$

To complete the proof, we have to show that $d_0 = d_G$, or equivalently, $d_0 = L$ (note that $L = d_G$). The value of d_0 is obtained from the following linear programming optimization problem

$$\begin{aligned} &\min \quad \mathbf{w}\cdot\boldsymbol{\mu} + \mathbf{1}\cdot\boldsymbol{\nu} \qquad (10) \\ &s.t. \quad \boldsymbol{\mu} \geq \mathbf{0}, \boldsymbol{\nu} \geq \mathbf{0}, \forall i \sum_{j=1}^{l_i}\mu_{\{\mathrm{p}_i(j),\mathrm{p}_i(j-1))\}} + \sum_{j=1}^{l_i-1}\nu_{i,j} \geq 1. \end{aligned}$$

According to the argument of linear programming [18], the solution of the above linear programming problem is equal to the solution of the dual problem which is

$$\begin{aligned} &\max \sum_{i=1}^{L} f_i \qquad (11) \\ &s.t. \quad \mathbf{0} \leq \mathbf{f} \leq \mathbf{1}, \forall e \in E, \sum_{e\in\mathrm{p}_i} f_i \leq w_e. \end{aligned}$$

[2] More precisely, with probability greater than $1 - P^{-\delta}$ for any $\delta > 0$.

Let us consider the solution $\mathbf{f}_0 = \mathbf{1}$ for (11). As the path sequence $(\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_L)$ consists of the paths that form the maximum flow in $\hat{G}$, we conclude that for every $e \in E$, we have $\sum_{e \in \mathrm{p}_i} 1 \leq w_e$. Hence, $\mathbf{f}_0$ is a feasible solution for (11). On the other hand, as for all feasible solutions $\mathbf{f}$ we have $\mathbf{f} \leq \mathbf{1}$, we conclude that $\mathbf{f}_0$ maximizes (11). Hence, we have

$$d_0 = \min \ \mathbf{w} \cdot \boldsymbol{\mu} + \mathbf{1} \cdot \boldsymbol{\nu} \overset{(a)}{=} \max \ \sum_{i=1}^{L} f_i = L = d_G. \quad (12)$$

Here, (a) results from duality of the primal and dual linear programming problems. This completes the proof. ■

Remark 1- It is worth noting that according to the proof of Theorem 1, any RS scheme achieves the maximum diversity of the wireless multiple-antenna multiple-relays network as long as its corresponding path sequence includes the paths $\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_{d_G}$ used in the proof of Theorem 1.

III. Diversity-Multiplexing Tradeoff

A. Two-Hop Single Relay Network

Here, we first consider a simple configuration of the two-hop relay network consisiting of a source, single full-duplex relay and a destination with no direct link between the source and the destination. The source, relay, and destination are supposed to be equipped with m, p, and n antenna, respectively. It can easily be shown that the decode-forward (DF) scheme achieves the optimum DMT for this configuration[3]. However, here, we show that the RS scheme which is based on AF relaying can achieve the optimum DMT as well.

The channel between the source and the relay is denoted by $\mathbf{H}$ and the channel between the relay and the destination is denoted by $\mathbf{G}$, which are assumed to be circularly symmetric zero-mean unit-variance Gaussian and remain fixed during the whole transmission period (quasi-static fading). It is assumed that the source and the relay have power constraints $\mathbb{E}\{\mathbf{x}_t^H \mathbf{x}_t\} \leq P$ and $\mathbb{E}\{\mathbf{x}_r^H \mathbf{x}_r\} \leq P$, respectively.

In the traditional AF strategy, the received signal at the relay is multiplied by a constant α such that the power constraint at the relay is satisfied and transmitted to the destination. In this way, the received signal at the destination can be written as

$$\mathbf{y} = \alpha \mathbf{G}\mathbf{H}\mathbf{x}_t + \alpha \mathbf{G}\mathbf{n}_r + \mathbf{n}_d, \quad (13)$$

where $\mathbf{n}_r \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_p)$ and $\mathbf{n}_d \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_n)$ denote the noise vectors at the relay and the destination, respectively.

Lemma 1 *The DMT of the systems in (13) is equal to the DMT of the following system:*

$$\mathbf{y} = \alpha \mathbf{G}\mathbf{H}\mathbf{x}_t + \mathbf{n}_d, \quad (14)$$

Proof: See [19]. ■

[3]In fact, this configuration is a special case of the degraded relay channel studied by [9]. In [9], the authors show that the DF scheme achieves the capacity of the degraded relay channel.

A direct conclusion of Lemma 1 is that the DMT of the two-hop network can be expressed as the DMT of the product channel $\mathbf{GH}$. In other words, imposing the constraint $\alpha \leq 1$ does not change the DMT of the system. The DMT of the product channel is computed in [11]. Due to their result given in Proposition 1, the DMT of the product channel $\mathbf{A} = \mathbf{GH}$ is a piecewise-linear function connecting the points $(r, d_{\mathbf{A}}(r)), r = 0, \ldots, l$, where

$$d_{\mathbf{A}}(r) = (p-r)(q-r) - \frac{1}{2}\left\lfloor \frac{[(p-\Delta-r)^+]^2}{2} \right\rfloor, \quad (15)$$

$q = \min(m, n)$ and $\Delta = |m-n|$. On the other hand, the piecewise-linear function connecting the integer points $(k, (p-k)(q-k))$ can be easily derived as the upper-bound by considering each of the source-relay or the relay-destination cuts. Comparing (15) with the upper-bound, it follows that the traditional amplify-and-forward achieves the optimum DMT only when $r \geq p - \Delta$. This motivates us to use a variant of amplify-and-forward which achieves the optimum DMT in all cases. In fact, using the traditional AF, there are three sources of outage: (i) the outage in the source-relay link, (ii) the outage in the relay-destination link, and (iii) the projection of the eigenmodes of $\mathbf{H}$ over the eigenmodes of $\mathbf{G}$ is very small. More precisely, the matrix $\mathbf{V}^H(\mathbf{G})\mathbf{U}(\mathbf{H})$, in which $\mathbf{V}^H(\mathbf{G})$ denotes the right eigenvector matrix from the SVD of $\mathbf{G}$ and $\mathbf{U}(\mathbf{H})$ denotes the left eigenvector matrix from the SVD of $\mathbf{H}$, has very small eigenvalues. The extra term $\frac{1}{2}\left\lfloor \frac{[(p-\Delta-r)^+]^2}{2} \right\rfloor$ in (15) is due to the third source of outage. The first two outage events depend on the distribution of the eigenvalues of $\mathbf{H}$ and $\mathbf{G}$, while the third event depends solely on the *direction* of the eigenvectors of these two matrices. This suggests us that in order to eliminate the extra terms of $\frac{1}{2}\left\lfloor \frac{[(p-\Delta-r)^+]^2}{2} \right\rfloor$ one can multiply the received signal at the relay by $\alpha\boldsymbol{\Theta}$, for some $p \times p$ unitary matrix $\boldsymbol{\Theta}$ (for preserving the power constraint at the relay). However, it should be noted that in each transmission slot l, an independent random unitary matrix $\boldsymbol{\Theta}_l$ should be applied; otherwise, the performance of the systems does not change. It should be noted that the proposed RS scheme performs in this way. Indeed, as in this setup the source and the destination are connected only through one path, the RS scheme reduces to the following scheme: The source's message is sent through L slots by the same path; at the relay side, the received signal is multiplied by a random independent (through different slots) unitary matrix and following that, is multiplied by a scalar α such that $\alpha \leq 1$ and the power constant is satisfied and the result is sent in the next slot. At the destination, following receiving the signal of the slots $2, 3, \ldots, L+1$, the source's message is decoded. In the following theorem, we show that as long as L is above a certain threshold, the probability of the third outage event is negligible compared to the first two outage events and hence, the optimum DMT is achievable by the RS scheme.

Theorem 2 *Consider the two-hop network consisiting of a source with m antenna and a destination with n antenna which are connected through a full-duplex relay node with p antenna. Let us define $q = \min(m,n)$. Providing L is large enough such that $L \geq \min^2(p,q)\max(p,q)$, the RS scheme achieves the optimum DMT which is the piecewise-linear function connecting the points $(k,(p-k)(q-k)), k = 0,1,\dots,\min(p,q)$.*

Proof: Using Lemma 1, the DMT of the system using the proposed RS scheme is equal to the DMT of the following system:

$$\mathbf{Y} = \alpha \mathbf{\Omega} \mathbf{X}_t + \mathbf{N}_d, \tag{16}$$

where $\mathbf{X}_t \triangleq [\mathbf{x}_t(1), \cdots, \mathbf{x}_t(L)]^T$, $\mathbf{Y} = [\mathbf{y}(1), \cdots, \mathbf{y}(L)]^T$, and $\mathbf{N}_d \triangleq [\mathbf{n}_d(1), \cdots, \mathbf{n}_d(L)]^T$, in which $\mathbf{x}_t(l)$ denotes the transmitted signal vector in the lth slot, and $\mathbf{y}(l)$ and $\mathbf{n}_d(l)$ denote the received signal and the noise at the destination corresponding to the signal sent in this slot, and

$$\mathbf{\Omega} \triangleq \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_L \end{bmatrix}, \tag{17}$$

in which $\mathbf{A}_l \triangleq \mathbf{G}\mathbf{\Theta}_l\mathbf{H}$. Hence, the matrix of the end-to-end channel is a block diagonal matrix consisting of $\mathbf{A}_l$'s. Assuming that the transmitted signals in each block are independent of each other, the mutual information between the input and the output of (16) can be written as

$$\mathcal{I}(\mathbf{X}_t;\mathbf{Y}) = \sum_{l=1}^{L} \log \left| \mathbf{I} + \alpha^2 \frac{P}{M} \mathbf{A}_l \mathbf{A}_l^H \right|, \tag{18}$$

in which, it is assumed that $\mathbf{x}_t(l) \sim \mathcal{CN}(\mathbf{0}, \frac{P}{M}\mathbf{I}_m)$, $\forall l = 1, \cdots, k$. Using the above equation, the probability of outage can be written as

$$\begin{aligned} \mathbb{P}\{\mathcal{O}\} &= \mathbb{P}\left\{ \sum_{l=1}^{L} \log \left| \mathbf{I} + \alpha^2 \frac{P}{m} \mathbf{A}_l \mathbf{A}_l^H \right| < Lr \log(P) \right\} \\ &= \mathbb{P}\left\{ \sum_{l=1}^{L} \sum_{j=1}^{\min(p,q)} \log \left(1 + \alpha^2 \frac{P}{m} \lambda_j(\mathbf{A}_l) \right) < Lr\log(P) \right\}, \end{aligned}$$

where $\lambda_i(\mathbf{A})$ denotes the ith ordered eigenvalue of $\mathbf{A}^H\mathbf{A}$ ($\lambda_1 > \lambda_2 > \cdots > \lambda_{\min}$). Defining $\gamma_j(\mathbf{B}) \triangleq -\frac{\log(\lambda_j(\mathbf{B}))}{\log(P)}$ and $\delta \triangleq -\frac{\log(\alpha^2)}{\log(P)}$, we have

$$\mathbb{P}\{\mathcal{O}\} \doteq \mathbb{P}\left\{ \sum_{l=1}^{L} \sum_{j=1}^{\min(p,q)} (1 - \delta - \gamma_j(\mathbf{A}_l))^+ < Lr \right\}. \tag{19}$$

It is shown in [19] that $\alpha \doteq 1$, with probability one[4]. Accordingly, one can replace δ by zero in (19), which results in

$$\mathbb{P}\{\mathcal{O}\} \doteq \mathbb{P}\left\{ \sum_{l=1}^{L} \sum_{j=1}^{\min(p,q)} (1 - \gamma_j(\mathbf{A}_l))^+ < Lr \right\}. \tag{20}$$

Moreover, it is shown in [19] that we can assume $\gamma_j(\mathbf{A}_l) \geq 0$, $\forall j = 1, \cdots, \min(p,q)$, i.e. imposing the constraint $\gamma_j(\mathbf{A}_l) \geq 0$ does not change the DMT.

In order to compute the outage probability in (20), we need to find the statistical behavior of $\gamma_j(\mathbf{A}_l)$. Since we are interested in upper-bounding the outage probability, finding an upper-bound for $\gamma_j(\mathbf{A}_l)$, or equivalently, a lower-bound for $\lambda_j(\mathbf{A}_l)$ would be sufficient. This is performed in the following lemma:

Lemma 2 *Consider matrices $\mathbf{G}$ and $\mathbf{H}$ with the size of $m \times p$ and $p \times n$, respectively, and a $p \times p$ matrix $\mathbf{\Theta}$. Assume $\mathbf{G}$ and $\mathbf{H}$ are SVD decomposed as $\mathbf{G} = \mathbf{U}(\mathbf{G})\mathbf{\Lambda}^{\frac{1}{2}}(\mathbf{G})\mathbf{V}^H(\mathbf{G})$ and $\mathbf{H} = \mathbf{U}(\mathbf{H})\mathbf{\Lambda}^{\frac{1}{2}}(\mathbf{H})\mathbf{V}^H(\mathbf{H})$, respectively. We have*

$$\lambda_i(\mathbf{G}\mathbf{\Theta}\mathbf{H}) \geq \lambda_i(\mathbf{G})\lambda_i(\mathbf{H})\lambda_{\min}\left(\mathbf{V}^H_{(1,i)}(\mathbf{G})\mathbf{\Theta}\mathbf{U}_{(1,i)}(\mathbf{H})\right),$$

where $\lambda_i(\mathbf{A})$ and $\lambda_{\min}(\mathbf{A})$ denote the i'th largest eigenvalue and the minimum eigenvalue of $\mathbf{A}^H\mathbf{A}$, respectively, and $\mathbf{A}_{(a,b)}$ denotes the submatrix of $\mathbf{A}$ consisting of the $a, a+1, \dots, b$'th columns of $\mathbf{A}$.

Proof: See [19]. ■

The good thing about the above lemma is that $\lambda_i(\mathbf{A}_l)$ is related to $\lambda_i(\mathbf{G})$ and $\lambda_i(\mathbf{H})$, which fascilitates the subsequent derivations. A direct consequence of the above lemmas is that

$$\gamma_i(\mathbf{A}_l) \leq \gamma_i(\mathbf{G}) + \gamma_i(\mathbf{H}) + \gamma_{\min}(\mathbf{\Psi}_{i,l}), \tag{21}$$

where $\mathbf{\Psi}_{i,l} \triangleq \mathbf{V}^H_{(1,i)}(\mathbf{G})\mathbf{\Theta}_l\mathbf{U}_{(1,i)}(\mathbf{H})$. As the statistical behaviors of $\gamma_i(\mathbf{G})$ and $\gamma_i(\mathbf{H})$ are known from [14], it is sufficient to derive the asymptotic behavior of $\gamma_{\min}(\mathbf{\Psi}_{i,l})$ or equivalently, $\lambda_{\min}(\mathbf{\Psi}_{i,l})$, which is performed in the following lemma:

Lemma 3 *Assuming small enough ε, we have*

$$\mathbb{P}\left\{\lambda_{\min}(\mathbf{\Psi}_{i,l}) \leq \varepsilon\right\} \leq \eta \sqrt[i]{\varepsilon}, \tag{22}$$

for some constant η.

Proof: See [19]. ■

A direct consequence of the above lemma is that $\mathbb{P}\left\{\gamma_{\min}(\mathbf{\Psi}_{i,l}) > \theta\right\} \dot{\leq} P^{-\frac{\theta}{i}}$. Defining the $1 \times L$ vector $\boldsymbol{\psi}$ as $\psi(l) \triangleq \max_i \gamma_{\min}(\mathbf{\Psi}_{i,l})$, we have

$$\begin{aligned} \mathbb{P}\left\{\boldsymbol{\psi} \geq \boldsymbol{\psi}_0\right\} &\overset{(a)}{=} \prod_{l=1}^{L} \mathbb{P}\left\{\psi(l) \geq \psi_0(l)\right\} \\ &= \prod_{l=1}^{L} \mathbb{P}\left\{ \bigcup_{i=1}^{\min(p,q)} \left(\gamma_{\min}(\mathbf{\Psi}_{i,l}) \geq \psi_0(l)\right) \right\} \\ &\overset{(b)}{\dot{\leq}} P^{-\frac{\mathbf{1}\cdot\boldsymbol{\psi}_0}{\min(p,q)}} \end{aligned} \tag{23}$$

[4]More precisely, with probability greater than $1 - P^{-\delta}$ for any $\delta > 0$

As $\mathbf{\Theta}_l$'s are independent isotropic unitary matrices, their products with any possibly correlated set of unitary matrices constructs a set of independent isotropic unitary matrices [17]. Accordingly, $\mathbf{\Psi}_{i,l}$'s are independent for different values of l which results in (a). Also, (b) follows from Lemma 3 and the union bound inequality.

Let us define the $1 \times \min(p,q)$ vectors $\boldsymbol{\chi}(\mathbf{H}) \triangleq \left[\gamma_{\min(p,m)}(\mathbf{H}), \gamma_{\min(p,m)-1}(\mathbf{H}), \ldots, \gamma_{1+\min(p,m)-\min(p,q)}(\mathbf{H})\right]$ and $\boldsymbol{\chi}(\mathbf{G}) \triangleq \left[\gamma_{\min(p,n)}(\mathbf{G}), \ldots, \gamma_{1+\min(p,n)-\min(p,q)}(\mathbf{G})\right]$. Notice that these vectors include the log-values of the corresponding $\min(p,q)$ smallest eigenvalues of $\mathbf{H}\mathbf{H}^H$ and $\mathbf{G}\mathbf{G}^H$, respectively. Now, applying the result of Lemma 2 to (20), we can upper-bound the outage probability of the end-to-end channel as

$$
\begin{aligned}
&\mathbb{P}\{\mathcal{O}\}\dot{\leq} \\
&\mathbb{P}\left\{\sum_{l=1}^{L}\sum_{i=1}^{\min(p,q)}\left(1-\gamma_i(\mathbf{G})-\gamma_i(\mathbf{H})-\gamma_{\min}(\mathbf{\Psi}_{i,l})\right)^+ < Lr\right\} \\
&\dot{\leq}\mathbb{P}\left\{\sum_{l=1}^{L}\sum_{i=1}^{\min(p,q)}\left(1-\gamma_i(\mathbf{G})-\gamma_i(\mathbf{H})-\psi(l)\right)^+ < Lr\right\} \\
&\dot{\leq}\mathbb{P}\left\{\sum_{l=1}^{L}\sum_{i=1}^{\min(p,q)}\left(1-\chi_i(\mathbf{G})-\chi_i(\mathbf{H})-\psi(l)\right)^+ < Lr\right\} \quad (24)
\end{aligned}
$$

Here, the problem is that according to (23), we have an upper-bound for $\mathbb{P}\left\{\boldsymbol{\psi} \geq \boldsymbol{\psi}_0\right\}$ which is not necessarily sufficient to upper-bound the probability of the region of $(\boldsymbol{\psi}, \boldsymbol{\chi}(\mathbf{H}), \boldsymbol{\chi}(\mathbf{G}))$ that satisfies (24). Indeed, for this purpose, we need the following Lemma.

Lemma 4 *Consider a fixed region $\mathcal{R} \subseteq [0,\infty)^n$. Assume that a uniformly continuous*[5] *non-negative function $f(\mathbf{x})$ ($f(\mathbf{x}) \geq 0$) is defined over $[0,\infty)^n$ such that for all $\mathbf{x} \in [0,\infty)^n$, we have $\mathbb{P}\left\{\mathbf{y} \geq \mathbf{x}\right\} \dot{\leq} P^{-f(\mathbf{x})}$ where $a(P) \dot{\leq} b(P)$ means $\lim_{P\to\infty}\frac{\log(a(P))}{\log(P)} \leq \lim_{P\to\infty}\frac{\log(b(P))}{\log(P)}$. Then, we have*

$$
\mathbb{P}\left\{\mathbf{x} \in \mathcal{R}\right\} \dot{\leq} P^{-\inf_{\mathbf{x}\in\mathcal{R}} f(\mathbf{x})}. \quad (25)
$$

Proof: See [19]. ∎

According to the upper-bound in (23) and the distribution of $\boldsymbol{\chi}(\mathbf{G}), \boldsymbol{\chi}(\mathbf{H})$ derived in [14], we have

$$
\begin{aligned}
&\mathbb{P}\left\{\boldsymbol{\psi} \geq \hat{\boldsymbol{\psi}}, \boldsymbol{\chi}(\mathbf{G}) \geq \hat{\boldsymbol{\chi}}(\mathbf{G}), \boldsymbol{\chi}(\mathbf{H}) \geq \hat{\boldsymbol{\chi}}(\mathbf{H})\right\} \dot{\leq} \\
&P^{-\frac{1}{\min(p,q)}\sum_{l=1}^{L}\hat{\psi}(l)-\sum_{i=1}^{\min(p,q)}(2i-1+|p-q|)(\hat{\chi}_i(\mathbf{G})+\hat{\chi}_i(\mathbf{H}))} \quad (26)
\end{aligned}
$$

Now, we can apply the result of Lemma 4 to the region defined in (24) and the upper-bound derived in (26). Accordingly, we have

$$
\mathbb{P}\left\{\mathcal{O}\right\} \dot{\leq} P^{-\min_{(\boldsymbol{\chi}(\mathbf{G}),\boldsymbol{\chi}(\mathbf{H}),\boldsymbol{\psi})\in\mathcal{R}} \frac{\sum_{l=1}^{L}\psi(l)}{\min(p,q)}+f(\boldsymbol{\chi}(\mathbf{G})+\boldsymbol{\chi}(\mathbf{H}))}, \quad (27)
$$

[5] A *uniformly continuous* function $f: \mathcal{M} \to \mathcal{N}$ where $\mathcal{M} \subseteq \mathbb{R}^m, \mathcal{N} \subseteq \mathbb{R}^n$ is a function that has the following property: for every ϵ, there exists a constant $g(\epsilon)$ such that for all $\mathbf{x}, \mathbf{y} \in \mathcal{M}, \|\mathbf{x}-\mathbf{y}\| \leq g(\epsilon)$, we have $\|f(\mathbf{x}) - f(\mathbf{y})\| \leq \epsilon$.

where $f : \mathbb{R}^{\min(p,q)} \to \mathbb{R}$ is defined as $f(\boldsymbol{\psi}) = \sum_{i=1}^{\min(p,q)}(2i-1+|p-q|)\psi_i$ and the region $\mathcal{R}$ is defined as

$$
\begin{aligned}
\mathcal{R} \quad \triangleq \quad & \left\{(\boldsymbol{\chi}(\mathbf{G}), \boldsymbol{\chi}(\mathbf{H}), \boldsymbol{\psi}) \middle| \boldsymbol{\psi} \geq \mathbf{0}, \chi_1(\mathbf{G}) \geq \cdots \geq \right. \\
& \chi_{\min(p,q)}(\mathbf{G}) \geq 0, \chi_1(\mathbf{H}) \geq \cdots \geq \chi_{\min(p,q)}(\mathbf{H}) \geq 0 \\
& \left. ,\sum_{l=1}^{L}\sum_{i=1}^{\min(p,q)}\left(1-\chi_i(\mathbf{G})-\chi_i(\mathbf{H})-\psi(l)\right)^+ \leq Lr\right\}
\end{aligned}
$$

Let us assume L is large enough such that $L \geq \min(p,q)\left(\sum_{i=1}^{\min(p,q)} 2i-1+|p-q|\right) = \min^2(p,q)\max(p,q)$. We define the $1 \times \min(p,q)$ vector $\boldsymbol{\varphi}$ as $\varphi_i \triangleq \chi_i(\mathbf{G}) + \chi_i(\mathbf{H}) + \frac{1}{L}\sum_{l=1}^{L}\psi(l)$. For each $(\boldsymbol{\chi}(\mathbf{G}), \boldsymbol{\chi}(\mathbf{H}), \boldsymbol{\psi}) \in \mathcal{R}$, we have

$$
\begin{aligned}
Lr \quad &\geq \quad \sum_{l=1}^{L}\sum_{i=1}^{\min(p,q)}\left(1-\chi_i(\mathbf{G})-\chi_i(\mathbf{H})-\psi(l)\right)^+ \\
&\geq \quad \sum_{i=1}^{\min(p,q)}\max\left\{0, \sum_{l=1}^{L}1-\chi_i(\mathbf{G})-\chi_i(\mathbf{H})-\psi(l)\right\} \\
&= \quad L\sum_{i=1}^{\min(p,q)}\left(1-\varphi_i\right)^+. \quad (28)
\end{aligned}
$$

On the other hand, according to (27) and the definition of $\boldsymbol{\varphi}$, we conclude that

$$
\mathbb{P}\{\mathcal{O}\} \dot{\leq} P^{-\min_{(\boldsymbol{\chi}(\mathbf{G}),\boldsymbol{\chi}(\mathbf{H}),\boldsymbol{\psi})\in\mathcal{R}}\sum_{i=1}^{\min(p,q)}(2i-1+|p-q|)\varphi_i}. \quad (29)
$$

Notice that according to the definition of $\boldsymbol{\varphi}$, we can easily conclude that $\varphi_1 \geq \cdots \geq \varphi_{\min(p,q)} \geq 0$. Hence, applying (28) and (29), we can upper-bound the outage probability as

$$
\mathbb{P}\{\mathcal{O}\} \dot{\leq} P^{-\min_{\boldsymbol{\varphi}\in\hat{\mathcal{R}}}\sum_{i=1}^{\min(p,q)}(2i-1+|p-q|)\varphi_i} \quad (30)
$$

where $\hat{\mathcal{R}}$ is defined as

$$
\hat{\mathcal{R}} \triangleq \left\{\boldsymbol{\varphi} \middle| \varphi_1 \geq \cdots \geq \varphi_{\min(p,q)} \geq 0, \sum_{i=1}^{\min(p,q)}\left(1-\varphi_i\right)^+ \leq r\right\} \quad (31)
$$

According to [14], (30) and (31) define the probability of outage from the rate $r\log(P)$ in an equivalent $p \times q$ MIMO point-to-point channel. Hence, we have $d_{RS}(r) \geq d_{p\times q}(r)$. On the other hand, due to the cut-set bound Theorem [20], we know that the DMT of the system is upper-bounded by the minimum of the DMT of the equivalent point-to-point $p \times m$ and $n \times p$ channels. Hence, $d_{RS}(r) \leq d_{opt}(r) = d_{p\times q}(r)$. Accordingly, we have $d_{RS}(r) = d_{opt}(r) = d_{p\times q}(r)$ which completes the proof. ∎

The statement of Theorem 2 can be generalized to multi-hop networks as follows.

Theorem 3 *Consider a multi-antenna multi-hop network consisiting of a single source and destination and full-duplex relays, with exactly one relay in each hop and assume all the nodes are equipped with N antennas. Providing L is* *large enough such that $L \geq N^3$, the RS scheme achieves the optimum DMT which is the piecewise-linear function connecting the points $(k,(N-k)^2), k=0,1,\dots,N$.*

Proof: See [19]. ■

B. Parallel Relay Network

Theorem 4 *Consider a multi-antenna parallel relay network consisiting of a source equipped with m antenna, a destination equipped with n antenna and K half-duplex relays each equipped with p antenna. Assume that there exists no direct link between the source and the destination. For any fixed $B \geq \min^2(p,q)\max(p,q)$, the RS scheme with $L=BK$, $S=BK+1$, the path sequence $\mathrm{Q} \equiv (\mathrm{q}_1,\dots,\mathrm{q}_K,\mathrm{q}_1,\dots,\mathrm{q}_K,\dots,\mathrm{q}_1,\dots,\mathrm{q}_K)$ where $\mathrm{q}_k \equiv (0,k,K+1)$ and the timing sequence $s_{i,j}=i+j-1$ achieves the diversity gain*

$$d_{RS}(r) \geq K d_{p\times q}\left(\left(1+\frac{1}{BK}\right)r\right), \tag{32}$$

where $q \triangleq \min(m,n)$ and $d_{p\times q}(r)$ denotes the diversity gain of the point-to-point MIMO $p\times q$ channel corresponding to the rate $r\log(P)$. Moreover, as $B\to\infty$, the RS scheme achieves the diversity gain $Kd_{p\times q}(r)$.

Proof: See [19]. ■

In the following Theorem, we show that $Kd_{p\times q}(r)$ is the optimum DMT for the 2 relays half-duplex parallel relay network in which $m=n$.

Theorem 5 *Consider a multi-antenna parallel relay network consisiting of a source and a destination each equipped with m antennas, and $K=2$ half-duplex relays equipped with $n_k, k=1,2$ antennas. Assume that there exists no direct link between the source and the destination. Consider the RS scheme with $L=BK$, $S=BK+1$, and the path and timing sequences defined in Theorem 4. As $B\to\infty$, the RS scheme achieves the optimum DMT of the network.*

Proof: First, notice that according to the argument of Theorem 4, as $B\to\infty$, the RS scheme achieves the DMT $d_{RS,\infty}(r) \triangleq \min_{0\leq\nu\leq 2r} d_{m\times n_1}(\nu)+d_{m\times n_2}(2r-\nu)$. Now, to proof the Theorem, we just have to show that $d_{RS,\infty}(r)$ is indeed an upper-bound for the optimum DMT. According to the cut-set Theorem [20], we have an upper-bound for the capacity of the network for each channel realization. Hence, we can apply the cut-set Theorem to find an upper-bound for the optimum DMT. In general, for any general half-duplex relay network with K number of relays and any set $\{0\}\subseteq\mathcal{S}\subseteq\{0,1,\dots,K\}$, we say the network is in the *state* $\mathcal{S}$, if the network nodes in $\mathcal{S}$ are transmitting and the network nodes in $\mathcal{S}^c \triangleq \{0,1,\dots,K+1\}/\mathcal{S}$ are receiving. Notice that as the source is always transmitting and the destination is always receiving, we have $0\in\mathcal{S}, K+1\in\mathcal{S}^c$. Accordingly, we define a 1×2^K *state vector* $\boldsymbol{\rho}$ such that for any set $\mathcal{S}\subseteq\{1,2,\dots,K\}$, $\rho_{\mathcal{S}}$ shows the portion of time that the half-duplex relay network spends in the *state* $\mathcal{S}$ ($\sum_{\mathcal{S}\subseteq\{1,2,\dots,K\}}\rho_{\mathcal{S}}=1$). As the relay nodes and the source are assumed to have no channel state knowledge, we can assume that a fixed state vector $\boldsymbol{\rho}$ is associated with the strategy that achieves the optimum DMT. Denoting the outage event by $\mathcal{O}$, for any general half-duplex relay network consisiting of K relays, we have

$$\mathbb{P}\{\mathcal{O}\} \overset{(a)}{\geq} \min_{\boldsymbol{\rho}} \mathbb{P}\Bigg\{\bigcup_{\{0\}\subseteq\mathcal{T}\subseteq\{0,1,\dots,K\}}\Bigg(\sum_{\{0\}\subseteq\mathcal{S}\subseteq\{0,1,\dots,K\}} \rho_{\mathcal{S}} I\Big(X(\mathcal{S}\cap\mathcal{T});Y(\mathcal{S}^c\cap\mathcal{T}^c)|X(\mathcal{S}\cap\mathcal{T}^c)\Big) < r\log(P)\Bigg)\Bigg\}$$

$$\overset{(b)}{\doteq} \min_{\boldsymbol{\rho}} \max_{\{0\}\subseteq\mathcal{T}\subseteq\{0,1,\dots,K\}} \mathbb{P}\Bigg\{\sum_{\{0\}\subseteq\mathcal{S}\subseteq\{0,1,\dots,K\}} \rho_{\mathcal{S}}\times I\Big(X(\mathcal{S}\cap\mathcal{T});Y(\mathcal{S}^c\cap\mathcal{T}^c)|X(\mathcal{S}\cap\mathcal{T}^c)\Big) < r\log(P)\Bigg\}. \tag{33}$$

Here, (a) follows from the cut-set bound Theorem [20] and (b) follows from the union bound on the probability. Now, in our 2-relay parallel setup, let us define two sets $\mathcal{T}_1 \triangleq \{0,1\}$ and $\mathcal{T}_2 \triangleq \{0,2\}$ corresponding to two cut-sets. Moreover, let us define two events $\mathcal{O}_1$ and $\mathcal{O}_2$ as

$$\begin{aligned}\mathcal{O}_1 \triangleq\ & \left\{\left|\mathbf{I}_m + P\mathbf{G}_1\mathbf{G}_1^H\right| \leq \hat{\nu}\log(P),\right.\\ & \left.\left|\mathbf{I}_{n_2}+P\mathbf{H}_2\mathbf{H}_2^H\right| \leq (2r-\hat{\nu})\log(P)\right\},\end{aligned} \tag{34}$$

and

$$\begin{aligned}\mathcal{O}_2 \triangleq\ & \left\{\left|\mathbf{I}_{n_1} + P\mathbf{H}_1\mathbf{H}_1^H\right| \leq \hat{\nu}\log(P),\right.\\ & \left.\left|\mathbf{I}_{m}+P\mathbf{G}_2\mathbf{G}_2^H\right| \leq (2r-\hat{\nu})\log(P)\right\},\end{aligned} \tag{35}$$

where $\hat{\nu} \triangleq \underset{0\leq\nu\leq 2r}{\text{argmin}}\ d_{m\times n_1}(\nu)+d_{m\times n_2}(2r-\nu)$. Hence, in our setup,(33) can be simplified as

$$\begin{aligned}\mathbb{P}\{\mathcal{O}\} &\overset{(a)}{\geq} \min_{\boldsymbol{\rho}}\max\Bigg(\mathbb{P}\Bigg\{\sum_{\{0\}\subseteq\mathcal{S}\subseteq\{0,1,2\}}\rho_{\mathcal{S}}t_1(\mathcal{S}) \leq r\log(P)\Bigg\},\\ &\qquad \mathbb{P}\Bigg\{\sum_{\{0\}\subseteq\mathcal{S}\subseteq\{0,1,2\}}\rho_{\mathcal{S}}t_2(\mathcal{S}) \leq r\log(P)\Bigg\}\Bigg)\\ &\geq \min_{\boldsymbol{\rho}}\max\Bigg(\mathbb{P}\Bigg\{\Big(\rho_{\{0,1\}}+\rho_{\{0,1,2\}}\Big)\left|\mathbf{I}_m+P\mathbf{G}_1\mathbf{G}_1^H\right|\\ &\qquad +\Big(\rho_{\{0\}}+\rho_{\{0,1\}}\Big)\left|\mathbf{I}_{n_2}+P\mathbf{H}_2\mathbf{H}_2^H\right| \leq r\log(P)\Bigg\},\\ &\qquad \mathbb{P}\Bigg\{\Big(\rho_{\{0,2\}}+\rho_{\{0,1,2\}}\Big)\left|\mathbf{I}_m+P\mathbf{G}_2\mathbf{G}_2^H\right| +\\ &\qquad \Big(\rho_{\{0\}}+\rho_{\{0,2\}}\Big)\left|\mathbf{I}_{n_1}+P\mathbf{H}_1\mathbf{H}_1^H\right| \leq r\log(P)\Bigg\}\Bigg)\\ &\overset{(b)}{\geq} \min_{\boldsymbol{\rho}}\max\Big(\mathscr{Y}_1\mathbb{P}\{\mathcal{O}_1\},\mathscr{Y}_2 P\{\mathcal{O}_2\}\Big)\\ &\overset{(c)}{\dot{\geq}} P^{-d_{RS,\infty}(r)},\end{aligned} \tag{36}$$

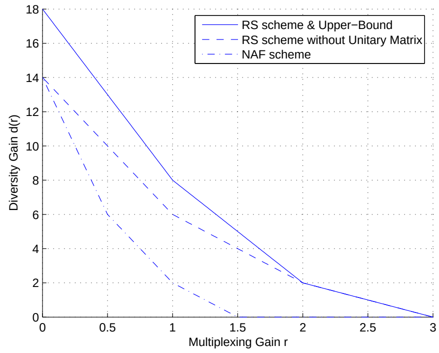

Fig. 1. Parallel relay network with $K = 2$ relays, each node with 3 antennas and no direct link between source and destination.

where $t_i(\mathcal{S}) \triangleq I\Big(X(\mathcal{S} \cap \mathcal{T}_i); Y(\mathcal{S}^c \cap \mathcal{T}_i^c) | X(\mathcal{S} \cap \mathcal{T}_i^c) \Big)$, $\mathscr{Y}_1 \triangleq \mathbf{1}\left[r - \left(\rho_{\{0,1,2\}} + \rho_{\{0,1\}}\right)\hat{\nu} - \left(\rho_{\{0,1\}} + \rho_{\{0\}}\right)(2r - \hat{\nu})\right]$, $\mathscr{Y}_2 \triangleq \mathbf{1}\left[r - \left(\rho_{\{0\}} + \rho_{\{0,2\}}\right)\hat{\nu} - \left(\rho_{\{0,2\}} + \rho_{\{0,1,2\}}\right)(2r - \hat{\nu})\right]$, and $\mathbf{1}[x] = 1$ for $x \geq 0$ and is 0 otherwise. Here, (a) results from taking the maximization of the right-hand side of (33) over $\mathcal{T}_1, \mathcal{T}_2$. (b) results from the facts that i) conditioned on $\mathcal{O}_1$ and assuming $r \geq \left(\rho_{\{0,1,2\}} + \rho_{\{0,1\}}\right)\hat{\nu} + \left(\rho_{\{0,1\}} + \rho_{\{0\}}\right)(2r - \hat{\nu})$, we have $\left(\rho_{\{0,1\}} + \rho_{\{0,1,2\}}\right)\left|\mathbf{I}_m + P\mathbf{G}_1\mathbf{G}_1^H\right| + \left(\rho_{\{0\}} + \rho_{\{0,1\}}\right)\left|\mathbf{I}_{n_2} + P\mathbf{H}_2\mathbf{H}_2^H\right| \leq r\log(P)$; and ii) conditioned on $\mathcal{O}_2$ and assuming $r \geq \left(\rho_{\{0\}} + \rho_{\{0,2\}}\right)\hat{\nu} + \left(\rho_{\{0,2\}} + \rho_{\{0,1,2\}}\right)(2r - \hat{\nu})$, we have $\left(\rho_{\{0,2\}} + \rho_{\{0,1,2\}}\right)\left|\mathbf{I}_m + P\mathbf{G}_2\mathbf{G}_2^H\right| + \left(\rho_{\{0\}} + \rho_{\{0,2\}}\right)\left|\mathbf{I}_{n_1} + P\mathbf{H}_1\mathbf{H}_1^H\right| \leq r\log(P)$. Knowing $\mathbb{P}\{\mathcal{O}_1\} = \mathbb{P}\{\mathcal{O}_2\} \doteq P^{-d_{m\times n_1}(\hat{\nu}) - d_{m\times n_2}(2r-\hat{\nu})} = P^{-d_{RS,\infty}(r)}$ and the fact that $\mathbf{1}[x] + \mathbf{1}[y] \geq \mathbf{1}[x+y]$ result in (c). (36) completes the proof of the Theorem. ■

However, if we do not apply random unitary matrix multiplication at the relay nodes, applying the proof-steps of Theorem 4, one can easily show that the RS scheme achieves the DMT $Kd_{\mathbf{GH}}(r)$ where $d_{\mathbf{GH}}(r)$ denotes the DMT of the product of the channel matrix $\mathbf{H}$ from the source to the relay and the channel matrix $\mathbf{G}$ from the relay to the destination (see (15)). Finally, applying the NAF scheme, one can easily show that the DMT $Kd_{\mathbf{GH}}(2r)$ is achievable.

C. Multiple-Antenna Single Relay Channel

[11] shows that in the multiple-antenna half-duplex single relay channel consisiting of the source, relay, and the destination equipped with m,p, and n antennas, respectively, the NAF protocol achieves the DMT $d_{NAF}(r) \geq d_{m\times n}(r) + d_{\mathbf{GH}}(2r)$. Here, we show that using random independent unitary marices also improves the DMT of the NAF scheme for the multiple-antenna single relay channel.

Theorem 6 *Consider the multiple-antenna half-duplex single relay channel consisiting of a source, a relay, and a destination equipped with m,p, and n antennas, respectively. Let us consider a modified NAF scheme in which the received signal is multiplied by a random unitary matrix at the relay node and finally, multiplied by a scalar $\alpha \leq 1$ such that the output power constraint is satisfied. Let us assume that the modified NAF scheme is applied for B consecutive slots. Assuming $B \geq \min^2(p,q)\max(p,q)$ where $q \triangleq \min(m,n)$, the modified NAF scheme achieves the DMT*

$$d_{MNAF}(r) \geq d_{m\times n}(r) + d_{p\times q}(2r). \tag{37}$$

Proof: See [19]. ■

REFERENCES

[1] Sh. Oveis Gharan, A. Bayesteh, and A. K. Khandani, "Diversity-multiplexing tradeoff in multiple-relay network-part I: proposed scheme and single-antenna networks," in *46th Allerton Conference on Communication, Control, and Computing*, 2008.

[2] ——, "Diversity-multiplexing tradeoff in multiple-relay network," *IEEE Trans. Inf. Theory*, april 2008, submitted for Publication, available online at http://arxiv.org/pdf/0803.3117.

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

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

[5] M. Yuksel and E. Erkip, "Cooperative wireless systems: a diversity-multiplexing tradeoff perspective," *IEEE Trans. Inform. Theory*, Aug. 2006, under Review.

[6] Sh. Yang and J.-C. Belfiore, "Towards the optimal amplify-and-forward cooperative diversity scheme," *IEEE Trans. Inform. Theory*, vol. 53, pp. 3114–3126, Sept. 2007.

[7] Sh. Oveis Gharan, A. Bayesteh, and A. K. Khandani, "Asymptotic analysis of amplify and forward relaying in a parallel MIMO relay network," *IEEE Trans. Inf. Theory*, 2006, revised, available online at http://cst.uwaterloo.ca/.

[8] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, pp. 3037–3063, Sept. 2005.

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

[10] S. N. D. A. S. Avestimehr and D. N. C. Tse, "Approximate capacity of Gaussian relay networks," in *IEEE Int. Symp. Information Theory*, 2008.

[11] Sh. Yang and J.-C. Belfiore, "Optimal space-time codes for the MIMO amplify-and-forward cooperative channel," *IEEE Trans. Inform. Theory*, vol. 53, no. 2, pp. 647–663, 2007.

[12] M. Gastpar and M. Vetterli, "On the capacity of large Gaussian relay networks," *IEEE Trans. Inf. Theory*, vol. 51, pp. 765–779, March 2005.

[13] H. Bolcskei, R. U. Nabar, O. Oyman, and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. on Wireless Communications*, vol. 5, no. 6, pp. 1433–1444, June 2006.

[14] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1073– 1096, May 2003.

[15] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler, "Fading relay channels: performance limits and space-time signal design," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.

[16] R. Diestel, *Graph Theory*. Springer, 2006.

[17] A. M. Tulino and S. Verdu, *Random Matrix Theory and Wireless Communications*. Now Publishers Inc, 2004.

[18] A. Schrijver, *Theory of Linear and Integer Programming*. Wiley, 1998.

[19] Sh. Oveis Gharan, A. Bayesteh, and A. K. Khandani, "On diversity-multiplexing tradeoff in multiple-antennas multiple-relay networks," University of Waterloo, Tech. Rep., July 2008, available online at http://cst.uwaterloo.ca/.

[20] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New york: Wiley, 1991.