Distributed Transmit Diversity in Relay Networks

Cemal Akçaba, Patrick Kuppinger and Helmut Bölcskei
Communication Technology Laboratory
ETH Zurich, Switzerland
Email: {cakcaba | patricku | boelcskei}@nari.ee.ethz.ch

***Abstract*— We analyze fading relay networks, where a single-antenna source-destination terminal pair communicates through a set of half-duplex single-antenna relays using a two-hop protocol with linear processing at the relay level. A family of relaying schemes is presented which achieves the optimal diversity-multiplexing (DM) tradeoff for all multiplexing gains. As a byproduct of our analysis, it follows that delay diversity and phase-rolling at the relay level are optimal with respect to the entire DM-tradeoff curve, provided the delays and the modulation frequencies are chosen appropriately.**

I. INTRODUCTION

Efficiently utilizing the available distributed spatial diversity in wireless networks is a challenging problem. In this paper, we consider fading relay networks, where a single-antenna source-destination terminal pair communicates through a set of K half-duplex single-antenna relays. We assume that there is no direct link between the source and the destination terminals and communication takes place using a two-hop protocol over two time slots. The source terminal and the relays do not have any channel state information (CSI), and the destination terminal knows all channels in the network perfectly.

Previous work: For setups similar to that described above, Laneman *et al.* [1] propose space-time coded cooperative diversity protocols achieving full spatial diversity gain (i.e., the diversity order equals the number of relay terminals). For the setup considered in this paper, Jing and Hassibi [2] analyze distributed linear dispersion space-time coding (STC) schemes and show that a diversity order equal to the number of relay terminals can be achieved. In [3], Azarian *et al*. derive the optimal diversity-multiplexing (DM) tradeoff curve for half-duplex relay networks and provide protocols achieving the entire tradeoff curve.

Contributions: In this paper, we are interested in simple linear relay transmit diversity schemes that realize full distributed spatial diversity gain. Specific examples include phase-rolling [4] and delay diversity [5], [6] developed in the context of point-to-point multiple-antenna systems and adopted to relay networks in [7], [8], [9]. Phase-rolling and delay diversity at the relay level are attractive from an implementation point-of-view as they convert distributed spatial diversity into time-diversity and frequency-diversity, respectively, which can be exploited using standard forward error correction over the resulting effective single-antenna point-to-point-channel. In [7], it is concluded, through simulation results that a K-relay delay diversity system can achieve a diversity gain of K. In [8], it is demonstrated that phase-rolling at the relay level can achieve second-order diversity. The contributions in this paper can be summarized as follows:

- We introduce a broad family of relay transmit diversity schemes encompassing delay diversity and phase-rolling as special cases.
- While the (numerical) results in [7], [8] are for the case of fixed rate (i.e., the rate does not scale with SNR), we provide a sufficient condition on the family of relay transmit diversity schemes introduced in this paper to achieve the entire DM-tradeoff curve as defined in [10]. The tools used to prove DM-tradeoff optimality are a method for computing the optimal DM-tradeoff curve in selective channels, introduced in [11], and a set of techniques described in [3].

Notation: The superscripts ^{T,H} and * stand for transpose, conjugate transpose, and conjugation, respectively. x_i represents the ith element of the column vector $\mathbf{x}$, and $X_{i,j}$ stands for the element in the ith row and jth column of the matrix $\mathbf{X}$. $\mathbf{X} \circ \mathbf{Y}$ denotes the Hadamard product of the matrices $\mathbf{X}$ and $\mathbf{Y}$. $\mathrm{rank}(\mathbf{X})$ stands for the rank of $\mathbf{X}$. $\mathrm{Tr}(\mathbf{X})$, $\|\mathbf{X}\|_F$, and $\lambda_i(\mathbf{X})$ $(i = 0, 1, \ldots, N-1)$ denote the trace, the Frobenius norm, and the ith eigenvalue (sorted in descending order) of $\mathbf{X}$, respectively. $\mathbf{I}_N$ is the $N \times N$ identity matrix. $\mathbf{0}$ denotes the all zeros vector of appropriate size. We say that the square matrices $\mathbf{X}$ and $\mathbf{Y}$ are orthogonal if $\langle \mathbf{X}, \mathbf{Y} \rangle = \mathrm{Tr}\big(\mathbf{X}\mathbf{Y}^H\big) = 0$. All logarithms are to the base 2 and $(a)^+ = \max(a, 0)$. $X \sim \mathcal{CN}(0, \sigma^2)$ stands for a circularly symmetric complex Gaussian random variable (RV) with variance σ^2. Let the RV X be parameterized by $\rho > 0$. The exponential order of X in ρ is defined as $v = -\frac{\log X}{\log \rho}$. $f(\rho) \doteq g(\rho)$ denotes exponential equality in ρ of the functions $f(\cdot)$ and $g(\cdot)$, i.e.,

$$\lim_{\rho\to\infty} \frac{\log f(\rho)}{\log \rho} = \lim_{\rho\to\infty} \frac{\log g(\rho)}{\log \rho}.$$

The symbols $\dot{\geq}$, $\dot{\leq}$, $\dot{>}$ and $\dot{<}$ are defined analogously. $\overset{d}{=}$ denotes equivalence in distribution.

II. SYSTEM MODEL

Preliminaries: We consider a wireless network with $K + 2$ single-antenna terminals, where a source terminal $\mathcal{S}$ communicates with a destination terminal $\mathcal{D}$ through a set of K half-duplex relay terminals $\mathcal{R}_i$ $(i = 1, 2, \ldots, K)$ (see Fig. 1). For the sake of simplicity, we assume that there is no direct link between $\mathcal{S}$ and

This research was supported by Nokia Research Center Helsinki, Finland and by the STREP project No.IST-027310 MEMBRANE within the Sixth Framework Programme of the European Commission.

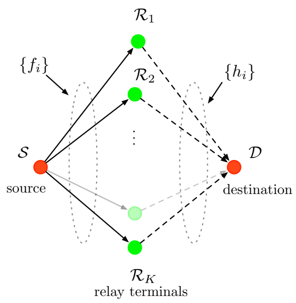

Fig. 1. Two-hop relay network with K single-antenna relay terminals.

$\mathcal{D}$. The channels[1] $\mathcal{S} \to \mathcal{R}_i$, denoted as f_i, and $\mathcal{R}_i \to \mathcal{D}$, denoted as h_i, $(i = 1, 2, \ldots, K)$, are i.i.d. $\mathcal{CN}(0, 1)$. We define the column vectors $\mathbf{f} = [f_1\ f_2\ \cdots\ f_K]^T$ and $\mathbf{h} = [h_1\ h_2\ \cdots\ h_K]^T$.

Communication takes place over two time slots. In the first time slot, $\mathcal{S}$ transmits N symbols consecutively. The relay terminals process the received length-N sequence using a linear transformation and transmit the result during the second time slot to $\mathcal{D}$, while $\mathcal{S}$ remains silent. We assume that $\mathcal{S}$ and the relay terminals do not have CSI, whereas $\mathcal{D}$ knows f_i, h_i $(i = 1, 2, \ldots, K)$ perfectly. For simplicity, we further assume perfect synchronization and ignore the impact of shadowing and pathloss.

Signal model: The vectors $\mathbf{x}$, $\mathbf{r}_i$, $\mathbf{y} \in \mathbb{C}^N$ represent the N-dimensional transmitted symbol sequence, received symbol sequence at $\mathcal{R}_i$, and received symbol sequence at $\mathcal{D}$, respectively. The vector $\mathbf{r}_i$ is then given by

$$\mathbf{r}_i = \sqrt{\rho} f_i \mathbf{x} + \mathbf{w}_i,\ \ i = 1, 2, \ldots, K \tag{1}$$

where ρ denotes the average signal-to-noise ratio (SNR) (for all links) and $\mathbf{w}_i$ is the N-dimensional noise vector at $\mathcal{R}_i$, with i.i.d. $\mathcal{CN}(0, 1)$ entries. The $\mathbf{w}_i$ are independent across i as well. We assume an i.i.d. Gaussian codebook with covariance matrix $\mathbf{R}_{\mathbf{xx}} = \mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = \mathbf{I}_N$.

$\mathcal{R}_i$ applies the unitary transformation $\mathbf{G}_i$ according to $\mathbf{q}_i = \mathbf{G}_i \mathbf{r}_i$ and scales the result to meet the power constraint $\mathbb{E}\{\|\mathbf{q}_i\|^2\} = 1$. This results in the overall input-output relation

$$\mathbf{y} = \sum_{i=1}^{K} \frac{\rho}{\sqrt{\rho + 1}} h_i f_i \mathbf{G}_i \mathbf{x} + \tilde{\mathbf{z}} \tag{2}$$

where the effective noise term $\tilde{\mathbf{z}}$ (when conditioned on $\mathbf{h}$) is circularly symmetric complex Gaussian with $\mathbb{E}\{\tilde{\mathbf{z}}|\mathbf{h}\} = \mathbf{0}$ and $\mathbb{E}\{\tilde{\mathbf{z}}\tilde{\mathbf{z}}^H|\mathbf{h}\} = N_o' \mathbf{I}_N$ where $N_o' = \left(1 + \frac{\rho}{\rho+1}\|\mathbf{h}\|^2\right)$.

Since we will be interested in the mutual information under the assumption that $\mathcal{D}$ knows all the channels in the network perfectly, we can divide (2) by $\sqrt{N_o'}$ to obtain the effective input-output relation

$$\mathbf{y} = \frac{\rho}{\sqrt{1 + \rho(1 + \|\mathbf{h}\|^2)}} \sum_{i=1}^{K} h_i f_i \mathbf{G}_i \mathbf{x} + \mathbf{z} \tag{3}$$

[1] $\mathcal{A} \to \mathcal{B}$ denotes the link between terminals $\mathcal{A}$ and $\mathcal{B}$.

where $\mathbf{z}$ (when conditioned on $\mathbf{h}$) is a circularly symmetric Gaussian noise vector with $\mathbb{E}\{\mathbf{z}|\mathbf{h}\} = \mathbf{0}$ and $\mathbb{E}\{\mathbf{z}\mathbf{z}^H|\mathbf{h}\} = \mathbf{I}_N$. In the remainder of the paper, we shall be interested in the $\rho \to \infty$ case where $\frac{\rho}{\sqrt{1+\rho(1+\|\mathbf{h}\|^2)}} \approx \sqrt{\frac{\rho}{1+\|\mathbf{h}\|^2}}$. With $\mathbf{H}_{\text{eff}} = \sqrt{\frac{\rho}{1+\|\mathbf{h}\|^2}} \sum_{i=1}^{K} h_i f_i \mathbf{G}_i$, we can now rewrite the input-output relation (3) as

$$\mathbf{y} = \mathbf{H}_{\text{eff}} \mathbf{x} + \mathbf{z}. \tag{4}$$

III. Achieving The Optimal Diversity Multiplexing Tradeoff

Under the assumptions stated in the previous section, it follows that the mutual information of the effective channel in (4) is given by

$$I(\mathbf{y}; \mathbf{x}|\mathbf{H}_{\text{eff}}) = \frac{1}{2N} \sum_{n=0}^{N-1} \log(1 + \lambda_n(\mathbf{H}_{\text{eff}}\mathbf{H}_{\text{eff}}^H)) \tag{5}$$

where the factor $1/2$ is due to the half-duplex constraint. We shall next compute the optimal DM-tradeoff curve [10] for the effective channel $\mathbf{H}_{\text{eff}}$ and provide a sufficient condition on the unitary matrices $\mathbf{G}_i$ to achieve the entire tradeoff curve. Following the framework in [10], we define a channel outage event to occur if the mutual information does not support a target data rate of $R = r \log \rho$ (b/s/Hz). The probability of outage at multiplexing rate r and signal-to-noise ratio (SNR) ρ is then

$$P_{\mathcal{O}}(\rho, r) = \mathbb{P}[I(\mathbf{y}; \mathbf{x}|\mathbf{H}_{\text{eff}}) < r \log \rho]\,. \tag{6}$$

Directly analyzing (6) is challenging for the problem at hand as closed-form expressions for the eigenvalue distribution of $\mathbf{H}_{\text{eff}}$ do not seem to be available. However, noting that

$$I(\mathbf{y}; \mathbf{x}|\mathbf{H}_{\text{eff}}) \le I_J(\mathbf{y}; \mathbf{x}|\mathbf{H}_{\text{eff}}) \tag{7}$$

where

$$\begin{aligned} I_J(\mathbf{y}; \mathbf{x}|\mathbf{H}_{\text{eff}}) &= \frac{1}{2} \log\left(1 + \frac{1}{N} \sum_{n=0}^{N-1} \lambda_n(\mathbf{H}_{\text{eff}}\mathbf{H}_{\text{eff}}^H)\right) \\ &= \frac{1}{2} \log\left(1 + \frac{1}{N} \|\mathbf{H}_{\text{eff}}\|_F^2\right), \end{aligned} \tag{8}$$

we can resort to a technique developed in [11] to show that the DM-tradeoff corresponding to $I_J(\mathbf{y}; \mathbf{x}|\mathbf{H}_{\text{eff}})$ equals that corresponding to $I(\mathbf{y}; \mathbf{x}|\mathbf{H}_{\text{eff}})$. The significance of this result lies in the fact that the quantity $\|\mathbf{H}_{\text{eff}}\|_F^2$ lends itself nicely to analytical treatment.

In the following, we will need the Gramian matrix for the set $\mathcal{Q} = \{\mathbf{G}_1, \mathbf{G}_2 \ldots, \mathbf{G}_K\}$ defined as

$$\mathbf{K} = \frac{1}{N} \begin{bmatrix} \text{Tr}\left(\mathbf{G}_1\mathbf{G}_1^H\right) & \cdots & \text{Tr}\left(\mathbf{G}_K\mathbf{G}_1^H\right) \\ \text{Tr}\left(\mathbf{G}_1\mathbf{G}_2^H\right) & \cdots & \text{Tr}\left(\mathbf{G}_K\mathbf{G}_2^H\right) \\ \vdots & \vdots & \vdots \\ \text{Tr}\left(\mathbf{G}_1\mathbf{G}_K^H\right) & \cdots & \text{Tr}\left(\mathbf{G}_K\mathbf{G}_K^H\right) \end{bmatrix}. \tag{9}$$

Our main result can now be summarized as follows.

Theorem 1: For the half-duplex relay channel in (4), the optimal DM-tradeoff curve is given by

$$d(r) = K(1-2r)^+,\ r \in [0, \frac{1}{2}]. \tag{10}$$

Any linear relay processing scheme $\mathcal{Q} = \{\mathbf{G}_1, \mathbf{G}_2 \ldots, \mathbf{G}_K\}$ satisfying $\operatorname{rank}(\mathbf{K}) = K$ achieves the entire DM-tradeoff curve.

Proof: See Appendix A. ■

Discussion: Theorem 1 shows that the DM-tradeoff properties of the half-duplex relay channel in (4) are equal to the "cooperative upper bound" (apart from the factor $1/2$ loss, which is due to the half-duplex constraint) corresponding to a system with one transmit and K cooperating receive antennas. The lack of cooperation and noise forwarding at the relay level, hence, do not impact the DM-tradeoff behavior, provided the matrices $\{\mathbf{G}_i\}_{i=1}^K$ are chosen such that $\operatorname{rank}(\mathbf{K}) = K$. Azarian *et al.* [3] propose a protocol that achieves the optimal DM-tradeoff of the amplify-and-forward half-duplex relay channel and attains orthogonality between the relay transmissions by allowing only one relay to transmit in any given time slot. Our results show that there exists a whole family of linear relay processing schemes that are DM-tradeoff optimal and that $\operatorname{rank}(\mathbf{K}) = K$ is sufficient to achieve tradeoff optimality. Another immediate conclusion that can be drawn from Theorem 1 is the DM-tradeoff optimality of cyclic delay diversity [7] and phase-rolling [8], [9] at the relay level, provided the delays and modulation frequencies are chosen appropriately. This can be seen by noting that the cyclic delay diversity scheme [7] can be cast into our framework by setting $\mathbf{G}_i = \mathbf{P}_i$ where $\mathbf{P}_i$ denotes the permutation matrix that when applied to a vector $\mathbf{x}$ cyclically shifts the elements in $\mathbf{x}$ up by i positions.With

$$\langle \mathbf{P}_i, \mathbf{P}_j \rangle = \begin{cases} N, & \text{for } i = j \\ 0, & \text{for } i \neq j \end{cases}$$

the Gramian of $\mathcal{Q}$ is given by $\mathbf{K} = \mathbf{I}_K$ which upon application of Theorem 1 concludes the argument. In the case of phase-rolling [8], [9], we have

$$\mathbf{G}_i = \begin{bmatrix} e^{j\theta_i[0]} & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & e^{j\theta_i[N-1]} \end{bmatrix}$$

which, choosing the modulation frequencies $\theta_i[n] = \frac{2\pi n(i-1)}{N}$ gives

$$\langle \mathbf{G}_i, \mathbf{G}_j \rangle = \begin{cases} N, & \text{for } i = j, \\ 0, & \text{for } i \neq j \end{cases}$$

and hence $\operatorname{rank}(\mathbf{K}) = K$, concludes the argument. While the (numerical) results in [7], [8] are for the $r = 0$ case, our analysis reveals optimality of cyclic delay diversity and phase-rolling for the entire tradeoff curve.

Numerical results: We illustrate our results numerically using a network consisting of $K = 4$ relays. Fig. 2 shows the outage probability (obtained through Monte Carlo simulation) as a function of SNR for the cyclic delay diversity scheme with $\mathbf{K} = \mathbf{I}_4$ and for four other linear relay processing schemes with $\operatorname{rank}(\mathbf{K}) = 1, 2, 3,$ and 4, respectively, and nonzero spread of the positive eigenvalues of $\mathbf{K}$. We observe that both the cyclic delay diversity scheme and the scheme with $\operatorname{rank}(\mathbf{K}) = 4$ achieve a diversity order of 4. However, the delay diversity scheme offers slightly better performance indicating that orthogonality between the $\mathbf{G}_i$ improves performance. Moreover, the numerical results suggest that the diversity order achieved by the other schemes is given by $\operatorname{rank}(\mathbf{K})$. However, at this point, we do not have a proof of the latter statement.

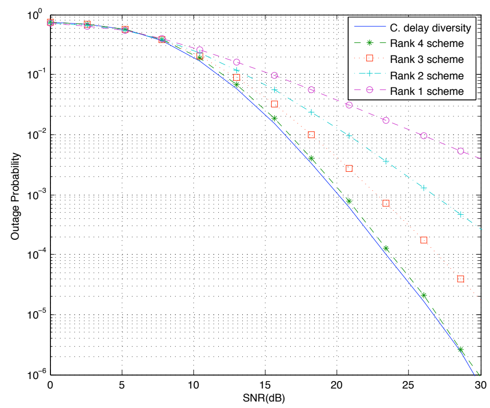

Fig. 2. Outage Probability vs. SNR for a delay diversity scheme and for four other schemes with different ranks.

IV. Conclusions

We introduced a family of linear relay processing schemes achieving the optimal DM-tradeoff of half-duplex relay channels. Cyclic delay diversity and phase-rolling were shown to be special (DM-tradeoff optimal) cases. Our results can be readily extended to account for the presence of a direct link between the source and the destination terminals. The DM-tradeoff framework seems to be too crude to distinguish between different full-rank schemes suggesting that a finer analysis is needed to quantify the impact of the eigenvalue spread of the Gramian matrix of the $\mathbf{G}_i$.

Appendix A
Proof of Theorem 1

We start by noting that an upper bound on the DM-tradeoff curve can be obtained by applying the broadcast cut-set bound [12] to the network in Fig. 1 and evaluating the corresponding DM-tradeoff. It is shown in [12], [13] that the broadcast cut amounts to a point-to-point link with a single transmit and K (cooperating) receive antennas. Taking into account the factor $1/2$ loss due to the half-duplex nature of the relay terminals, it follows from the results in [10] that the DM-tradeoff curve corresponding to the network analyzed in this paper is upper-bounded by

$$d(r) \leq K(1-2r)^+,\ r \in [0, \frac{1}{2}].$$

In the following, we shall show that this upper bound is achievable despite the lack of cooperation at the relay terminals provided the matrices $\mathbf{G}_i$ are chosen such that $\operatorname{rank}(\mathbf{K}) = K$. We start by noting that

$$P_{\mathcal{O}}(\rho, r) \geq \mathbb{P}[\mathcal{J}] = \mathbb{P}[I_J(\mathbf{y}; \mathbf{x}|\mathbf{H}_{\text{eff}}) < r \log \rho] \tag{11}$$

and defining the event

$$\mathcal{J} = \{\mathbf{H}_{\text{eff}} | I_J(\mathbf{y};\mathbf{x}|\mathbf{H}_{\text{eff}}) < r \log \rho\} \tag{12}$$

as "Jensen outage". Since

$$\begin{aligned}\frac{1}{N}\|\mathbf{H}_{\text{eff}}\|_F^2 &= \frac{\rho}{N(1+\|\mathbf{h}\|^2)}\text{Tr}\left(\sum_{i=1}^{K}\sum_{j=1}^{K}(h_i f_i)(h_j^* f_j^*)\mathbf{G}_i\mathbf{G}_j^H\right)\\ &= \frac{\rho}{1+\|\mathbf{h}\|^2}\tilde{\mathbf{h}}^H\mathbf{K}\tilde{\mathbf{h}},\end{aligned}$$

where $\tilde{\mathbf{h}} = \mathbf{h} \circ \mathbf{f}$, it follows that

$$\mathbb{P}[\mathcal{J}] = \mathbb{P}\left[\frac{1}{2}\log\left(1+\frac{\rho\tilde{\mathbf{h}}^H\mathbf{K}\tilde{\mathbf{h}}}{1+\|\mathbf{h}\|^2}\right) < r\log\rho\right]. \tag{13}$$

In what follows, we write $|h_i|^2 = \rho^{-u_i}$ and $|f_i|^2 = \rho^{-v_i}$ where u_i and v_i are RVs; the choice of this transformation will become clear later. Further, we define the events $\mathcal{A} = \{u_i, v_i | \ u_i > \delta, v_i > \delta, \ i \in \{1,2,\ldots,K\}\}$ and $\bar{\mathcal{A}} = \{u_i | \ u_i \le \delta, \ i \in \{1,2,\ldots,K\}\} \cup \{v_i | \ v_i \le \delta, \ i \in \{1,2,\ldots,K\}\}$ for a fixed $\delta > 0$. Using the law of total probability, we can write

$$\mathbb{P}[\mathcal{J}] = \mathbb{P}[\mathcal{A}]\,\mathbb{P}[\mathcal{J}|\mathcal{A}] + \mathbb{P}[\bar{\mathcal{A}}]\,\mathbb{P}[\mathcal{J}|\bar{\mathcal{A}}] \tag{14}$$

and bound $\mathbb{P}[\mathcal{J}]$ according to

$$\mathbb{P}[\mathcal{A}]\,\mathbb{P}[\mathcal{J}|\mathcal{A}] \le \mathbb{P}[\mathcal{J}] \le \mathbb{P}[\mathcal{A}]\,\mathbb{P}[\mathcal{J}|\mathcal{A}] + \mathbb{P}[\bar{\mathcal{A}}] \tag{15}$$

$$\mathbb{P}[\mathcal{A}]\,\mathbb{P}[\mathcal{J}|\mathcal{A}] \,\dot{\le}\, \mathbb{P}[\mathcal{J}] \,\dot{\le}\, \mathbb{P}[\mathcal{A}]\,\mathbb{P}[\mathcal{J}|\mathcal{A}] \tag{16}$$

$$\mathbb{P}[\mathcal{J}|\mathcal{A}] \,\dot{\le}\, \mathbb{P}[\mathcal{J}] \,\dot{\le}\, \mathbb{P}[\mathcal{J}|\mathcal{A}] \tag{17}$$

where (16) follows from the definition of the u_i and the v_i, their independence and by noting that $\lim_{\rho\to\infty}\frac{\log\mathbb{P}[\bar{\mathcal{A}}]}{\log\rho} = 0$. The double inequality (17) results from $\mathbb{P}[\mathcal{A}] = (1-\exp[-\rho^{-\delta}])^K$, $\lim_{\rho\to\infty}\frac{\log\mathbb{P}[\mathcal{A}]}{\log\rho} = -\delta K$, and the fact that δ can be made arbitrarily small. We have thus shown that $\mathbb{P}[\mathcal{J}] \doteq \mathbb{P}[\mathcal{J}|\mathcal{A}]$. Next, denoting the minimum and maximum eigenvalue of $\mathbf{K}$ as $\lambda_{\min}$ and $\lambda_{\max}$, respectively, we get the upper bound

$$\begin{aligned}\mathbb{P}[\mathcal{J}|\mathcal{A}] &\,\dot{\le}\, \mathbb{P}\left[\frac{1}{2}\log\left(1+\rho\frac{\lambda_{\min}}{1+K}\|\tilde{\mathbf{h}}\|^2\right) < r\log\rho\right] \quad (18)\\ &= \mathbb{P}\left[\frac{1}{2}\log\left(1+\rho\frac{\lambda_{\min}}{1+K}\sum_{i=1}^{K}|f_i|^2|h_i|^2\right) < r\log\rho\right] \quad (19)\\ &= \mathbb{P}\left[\frac{1}{2}\log\left(1+\frac{\lambda_{\min}}{1+K}\sum_{i=1}^{K}\rho^{1-v_i-u_i}\right) < r\log\rho\right], \quad (20)\end{aligned}$$

and the lower bound

$$\begin{aligned}\mathbb{P}[\mathcal{J}|\mathcal{A}] &\,\dot{\ge}\, \mathbb{P}\left[\frac{1}{2}\log\left(1+\rho\lambda_{\max}\|\tilde{\mathbf{h}}\|^2\right) < r\log\rho\right] \quad (21)\\ &= \mathbb{P}\left[\frac{1}{2}\log\left(1+\rho\lambda_{\max}\sum_{i=1}^{K}|f_i|^2|h_i|^2\right) < r\log\rho\right] \quad (22)\\ &= \mathbb{P}\left[\frac{1}{2}\log\left(1+\lambda_{\max}\sum_{i=1}^{K}\rho^{1-v_i-u_i}\right) < r\log\rho\right] \quad (23)\end{aligned}$$

where the key steps (18) and (21) follow from the Rayleigh-Ritz theorem [14] and the fact that $1 \le 1 + \sum_{i=1}^{K}\rho^{-u_i} \le 1 + K$ for $u_i > 0$ $(i = 1,2,\ldots,K)$ and $\rho > 1$. Note that due to the assumption $\text{rank}(\mathbf{K}) = K$, we have $\lambda_{\min} > 0$. We define the following events

$$\begin{aligned}\mathcal{B} &= \left\{u_i, v_i | \max_i(1-v_i-u_i) > 0\right\}\\ \mathcal{U} &= \left\{u_i, v_i \left| \frac{1}{2}\log\left(1+\frac{\lambda_{\min}\rho^{\max_i(1-v_i-u_i)}}{1+K}\right) < r\log\rho\right.\right\}\\ \mathcal{L} &= \left\{u_i, v_i \left| \frac{1}{2}\log\left(1+K\lambda_{\max}\rho^{\max_i(1-v_i-u_i)}\right) < r\log\rho\right.\right\}\end{aligned}$$

where the $\max$ is taken over $i = 1,2,\ldots,K$ in all three cases. With these definitions, we arrive at

$$\mathbb{P}[\mathcal{L}] \,\dot{\le}\, \mathbb{P}[\mathcal{J}|\mathcal{A}] \,\dot{\le}\, \mathbb{P}[\mathcal{U}] \tag{24}$$

$$\mathbb{P}[\mathcal{L}\cap\mathcal{B}] + \mathbb{P}[\mathcal{L}\cap\bar{\mathcal{B}}] \,\dot{\le}\, \mathbb{P}[\mathcal{J}|\mathcal{A}] \,\dot{\le}\, \mathbb{P}[\mathcal{U}\cap\mathcal{B}] + \mathbb{P}[\mathcal{U}\cap\bar{\mathcal{B}}]$$

$$\mathbb{P}[\mathcal{L}\cap\mathcal{B}] \,\dot{\le}\, \mathbb{P}[\mathcal{J}|\mathcal{A}] \,\dot{\le}\, \mathbb{P}[\mathcal{U}\cap\mathcal{B}] + \mathbb{P}[\bar{\mathcal{B}}] \tag{25}$$

where (24) follows from

$$\sum_{i=1}^{K}\rho^{1-v_i-u_i} \le K\rho^{\max_i(1-v_i-u_i)} \tag{26}$$

$$\sum_{i=1}^{K}\rho^{1-v_i-u_i} \ge \rho^{\max_i(1-v_i-u_i)} \tag{27}$$

and the RHS of (25) follows from conditioning with respect to event $\mathcal{B}$. Now we can expand $\mathbb{P}[\mathcal{U}\cap\mathcal{B}]$ as

$$\begin{aligned}\mathbb{P}[\mathcal{U}\cap\mathcal{B}] &= \mathbb{P}\left[0 < \max_i(1-v_i-u_i) < 2r+\epsilon_1\right] \quad (28)\\ &= \left(1-\frac{1}{\sqrt{\rho^{1-2r-\epsilon_1}}}\text{K}_1\left(\frac{1}{\sqrt{\rho^{1-2r-\epsilon_1}}}\right)\right)^K \quad (29)\\ &\doteq \rho^{-K(1-2r)^+} \quad (30)\end{aligned}$$

where $\epsilon_1 = \frac{\log(\frac{1+K}{\lambda_{\min}})}{\log\rho}$ and $\text{K}_1(\cdot)$ is the first-order modified Bessel function of the second kind. Eq. (29) follows from

$$\begin{aligned}&\mathbb{P}\left[0 < \max_i(1-v_i-u_i) < 2r+\epsilon_1\right]\\ &= \left(\mathbb{P}\left[|f_1|^2|h_1|^2 < \frac{1}{\rho^{1-2r-\epsilon_1}}\right]\right)^K - \left(\mathbb{P}\left[|f_1|^2|h_1|^2 < \frac{1}{\rho}\right]\right)^K\end{aligned}$$

and using the fact that the pdf of the square root of the product of two Rayleigh distributed RVs is given by $v(x) = x\text{K}_0(x)$ for $x > 0$, where $\text{K}_0(\cdot)$ is the zeroth-order modified Bessel function [15]. The exponential equality in (30) is proved using a Taylor series expansion of $f(x) = 1 - x\text{K}_1(x)$ around $x = 0$ and invoking asymptotic properties of the logarithmic function [16]. Similarly, one can show that

$$\mathbb{P}[\bar{\mathcal{B}}] = \left(1-\frac{1}{\sqrt{\rho}}\text{K}_1\left(\frac{1}{\sqrt{\rho}}\right)\right)^K \doteq 0. \tag{31}$$

To complete the proof, we need to establish that $\mathbb{P}[\mathcal{L} \cap \mathcal{B}]$ has the same exponential behavior in ρ as $\mathbb{P}[\mathcal{U} \cap \mathcal{B}]$. Using the same arguments as in (28)-(30), it follows readily that

$$\begin{aligned}\mathbb{P}[\mathcal{L} \cap \mathcal{B}] = &\ \mathbb{P}[0 < \max_i(1 - v_i - u_i) < 2r - \epsilon_2] &(32)\\ = &\left(1 - \frac{1}{\sqrt{\rho^{1-2r+\epsilon_2}}}\mathrm{K}_1\left(\frac{1}{\sqrt{\rho^{1-2r+\epsilon_2}}}\right)\right)^K &(33)\\ \doteq &\ \rho^{-K(1-2r)^+} &(34)\end{aligned}$$

where $\epsilon_2 = \frac{\log(K\lambda_{\max})}{\log \rho}$. We have thus shown that

$$\rho^{-K(1-2r)^+} \dot{\leq} \mathbb{P}[\mathcal{J}] \dot{\leq} \rho^{-K(1-2r)^+} \quad (35)$$

and hence

$$\mathbb{P}[\mathcal{J}] = P_J(\rho, r) \doteq \rho^{-K(1-2r)^+},\ r \in [0, \frac{1}{2}]. \quad (36)$$

Since $P_J(\rho, r) \leq P_{\mathcal{O}}(\rho, r)$ as a result of (7), and since the outage probability is a lower bound to the error probability achieved by any code [10], we have

$$P_J(\rho, r) \leq P_{\mathcal{O}}(\rho, r) \leq P_e(\rho, r). \quad (37)$$

Following the approach proposed in [11], we complete the proof of the theorem by identifying a code which has

$$P_e(\rho, r) \doteq P_J(\rho, r)$$

and hence a DM-tradeoff curve which equals the "Jensen" DM-tradeoff curve derived above. We start by writing

$$\begin{aligned}P_e(\rho, r) &= \mathbb{P}[\mathcal{J}]\,\mathbb{P}[\text{error}|\mathcal{J}] + \mathbb{P}\big[\text{error}, \bar{\mathcal{J}}\big] &(38)\\ &\leq \mathbb{P}[\mathcal{J}] + \mathbb{P}\big[\text{error}, \bar{\mathcal{J}}\big]. &(39)\end{aligned}$$

Next, we upper-bound $\mathbb{P}\big[\text{error}, \bar{\mathcal{J}}\big]$ through the union bound

$$\mathbb{P}\big[\text{error}, \bar{\mathcal{J}}\big] \leq \rho^{2Nr}\mathbb{P}\big[\mathbf{x}_0 \to \mathbf{x}_1, \bar{\mathcal{J}}\big] \quad (40)$$

where $\mathbb{P}\big[\mathbf{x}_0 \to \mathbf{x}_1, \bar{\mathcal{J}}\big]$ denotes the maximum pairwise error probability (over all codeword pairs) for ML decoding, and we used the fact that the codebook has ρ^{2Nr} codewords. With $\Delta\mathbf{x} = \mathbf{x}_1 - \mathbf{x}_0$, we have

$$\begin{aligned}\mathbb{P}[\mathbf{x}_0 \to \mathbf{x}_1 | \mathbf{H}_{\text{eff}} = \mathbf{H}] &= Q\left(\frac{||\mathbf{H}\Delta\mathbf{x}||_F}{\sqrt{2}}\right) &(41)\\ &\leq \exp\left[-\frac{||\mathbf{H}\Delta\mathbf{x}||_F^2}{4}\right] &(42)\\ &\dot{\leq} \exp\left[-\frac{N\Delta x_{\min}^2 \lambda_{\min} \sum_{i=1}^{K} \rho^{1-v_i-u_i}}{4(1+K)}\right] &(43)\end{aligned}$$

where $\Delta x_{\min}^2 = \min_i ||\Delta x_i||^2$, and (43) follows from noting that $\Delta\mathbf{x}$ is Gaussian and hence $\mathbf{U}\Delta\mathbf{x} \overset{d}{=} \Delta\mathbf{x}$ for unitary $\mathbf{U}$, from applying the Rayleigh-Ritz theorem, and from inserting $|h_i|^2 = \rho^{-u_i}$ and $|f_i|^2 = \rho^{-v_i}$ into (42). Next, we have

$$\begin{aligned}&\mathbb{P}\big[\mathbf{x}_0 \to \mathbf{x}_1, \bar{\mathcal{J}}\big] &(44)\\ &\leq \mathbb{E}\left\{\exp\left[-\frac{N\Delta x_{\min}^2 \lambda_{\min} \sum_{i=1}^{K} \rho^{1-v_i-u_i}}{4(1+K)}\right] \Big| \bar{\mathcal{J}}\right\} \mathbb{P}\big[\bar{\mathcal{J}}\big]\\ &\leq \exp\left[-\frac{N\Delta x_{\min}^2 \lambda_{\min} \rho^{2r}}{4(1+K)}\right] \mathbb{P}\big[\bar{\mathcal{J}}\big] &(45)\end{aligned}$$

where (45) follows since the event $\bar{\mathcal{J}}$ requires that $\sum_{i=1}^{K} \rho^{1-v_i-u_i} \geq \rho^{2r}$. Finally, inserting (45) into (40), and using $\mathbb{P}\big[\bar{\mathcal{J}}\big] \leq 1$, we get

$$\mathbb{P}\big[\text{error}, \bar{\mathcal{J}}\big] \dot{\leq} \rho^{2Nr} \exp\left[-\frac{N\Delta x_{\min}^2 \lambda_{\min} \rho^{2r}}{4(1+K)}\right]. \quad (46)$$

Choosing the codebook such that $\Delta x_{\min} > 0$ the proof is complete since (46) decays exponentially fast in ρ for any $r > 0$ and hence

$$\begin{aligned}P_e(\rho, r) &\leq \mathbb{P}[\mathcal{J}] + \mathbb{P}\big[\text{error}, \bar{\mathcal{J}}\big]\\ &\dot{\leq} \mathbb{P}[\mathcal{J}] + \rho^{2Nr} \exp\left[-\frac{N}{4(1+K)} \Delta x_{\min}^2 \lambda_{\min} \rho^{2r}\right]\\ &\dot{\leq} \mathbb{P}[\mathcal{J}] = P_J(\rho, r).\end{aligned}$$

which combined with (37) yields the desired result.

References

[1] 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.

[2] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. Wireless Comm.*, vol. 5, no. 12, Dec. 2006.

[3] 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.

[4] J. Guey, M. P. Fitz, M. R. Bell, and W. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," in *Proc. IEEE VTC*, vol. 1, Atlanta, GA, Apr./May 1996, pp. 136–140.

[5] A. Wittneben, "Basestation modulation diversity for digital SIMULCAST," in *Proc. IEEE VTC*, St. Louis, MO, May 1991, pp. 848–853.

[6] N. Seshadri and J. H. Winters, "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity," *Int. J. Wireless Information Networks*, vol. 1, pp. 49–60, Jan. 1994.

[7] S. B. Slimane and A. Osseiran, "Relay communication with delay diversity for future communication systems," in *Proc. IEEE Vehicular Technology Conference*, Sept. 2006.

[8] I. Hammerström, M. Kuhn, and A. Wittneben, "Cooperative diversity by relay phase rotations in block fading environments," in *Proc. Fifth IEEE Workshop on Signal Processing Advances in Wireless Communications, SPAWC 2004*, July 2004, pp. 293–297.

[9] P. Kuppinger, "Transformation of distributed spatial into temporal diversity by relay phase rotations," M.Sc. Thesis, Imperial College London, Sept. 2006.

[10] 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.

[11] P. Coronel and H. Bölcskei, "Diversity multiplexing tradeoff in selective MIMO fading channels," in *Proc. IEEE ISIT, submitted*, 2007.

[12] M. Gastpar and M. Vetterli, "On the capacity of large Gaussian relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 765–779, March 2005.

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

[14] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York, NY: Cambridge Press, 1985.

[15] J. Salo, H. El-Sallabi, and P. Vainikainen, "The distribution of the product of independent Rayleigh random variables," *IEEE Trans. Antennas and Propagation*, vol. 54, no. 2, pp. 639–643, Feb. 2006.

[16] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1965.