IST-4-027310 MEMBRANE

Deliverable D3.3

Diversity Multiplexing Tradeoff in Network Performance Terms

Contractual Date of Delivery to the CEC:	December 31st, 2007
Actual Date of Delivery to the CEC:	December 22nd, 2007
Author(s):	C. Akçaba[1], V. Morgenshtern[1], P. Kuppinger[1], H. Bölcskei[1], K. K. Leung[2], A. Alexiou[3]
Participant(s):	ETHZ[1], IMPERIAL[2], ALCATEL-LUCENT[3]
Workpackage:	WP3
Est. person months:	20
Security:	Public
Nature:	R
Version:	1.0
Total number of pages:	23

Abstract

We analyze fading relay networks, where a source-destination terminal pair communicates through a set of half-duplex single-antenna relays using four time-division-multiple-access (TDMA) based protocols with linear processing at the relay level. For each protocol, we derive the diversity-multiplexing (DM) tradeoff curve, and the sufficient conditions on the set of linear processing schemes and codebooks for achieving the DM-tradeoff curve. We show that these conditions are independent of the underlying fading distribution, and guarantee the achievability of the DM-tradeoff curve for any fading distribution. Our results show that the protocol with the highest degree of broadcasting and receive collision dominates the other protocols in terms of the DM-tradeoff performance. Further, we demonstrate that delay diversity and phase-rolling at the relay level are optimal with respect to the entire DM-tradeoff curve for each protocol, provided the family of codebooks, the delays and the modulation frequencies are chosen appropriately.

I. INTRODUCTION

This paper is concerned with improving link reliability in wireless networks. Specifically, we consider fading relay networks, where a source-destination terminal pair communicates through a set of K half-duplex single-antenna relays. We analyze a single-antenna source-destination pair communicating over two time slots using one of four time-division multiple access (TDMA)-based protocols. 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. We do not assume a particular kind of fading distribution for the channels in the network.

Previous work and contributions: Cooperative diversity was originally proposed to realize spatial diversity in a distributed fashion [1], [2], [3]. In [4], Laneman and Wornell propose a space-time coded cooperative diversity protocol which achieves full spatial diversity gain in the number of cooperating terminals. In [5] and [6], Nabar *et al.* analyze three TDMA-based protocols for single-relay fading channels and establish the superiority of the protocol that exhibits the highest level of receive collision and the highest level of degree of broadcasting (we shall call this protocol NBK-AF). In [7], Azarian *et al.* show that the NBK-AF protocol is diversity-multiplexing tradeoff (DMT) optimal for the single relay channel. However, the multi-relay extension of this protocol, the so-called *non-orthogonal amplify-and-forward* (NAF) protocol is optimal only for a certain class of time-division based strategies, i.e., only when all the relays are forced to receive during a listening phase, and transmit during a relaying phase where the durations of the listening and relaying phases are equal. We shall call such strategies *equal-duration-cooperation* (EDC). EDC-type strategies might be more suitable for implementation when compared to *variable-duration-cooperation* (VDC) strategies. Examples of VDC strategies are the slotted amplify-and-forward (SAF) protocol [8] and the dynamic decode-and-forward (DDF) protocol [7], which can both achieve a higher DMT curve than the EDC strategies can. In this work, we focus on EDC strategies where the relays are required to listen during the initial half of the source transmission, and do not consider VDC-type protocols with varying ratios of listening and relaying durations.

In this paper, we are interested in linear relay transmit diversity schemes that realize full distributed spatial diversity gain. Specific examples include phase-rolling [9] and delay diversity [10], [11] developed in the context of point-to-point multiple-antenna systems and adopted to relay networks in [12], [13], [14]. 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 [12], it is concluded, through simulation results that a K-relay cyclic delay diversity (CDD) system can achieve a diversity gain of K. In [13], it is demonstrated that phase-rolling at the relay level can achieve second-order diversity. In [15], it is shown that, for the three protocols analyzed in [6], any square complex orthogonal design achieves full diversity in the number of cooperating terminals. Recently, Rao and Hassibi derived the DMT curve for a K-relay Rayleigh-fading MIMO system; they considered multiantenna source and destination terminals, which communicate using a two-hop EDC-type protocol, and they assumed that there is no direct link present between the source and the destination and the relays use unitary transformations on their received signals prior to amplifying and forwarding [16].

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. Unlike prior work [15], [16], we do not restrict relays to use unitary transformations.
- While the (numerical) results in [12], [13] and the (analytical) results in [15] 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 [17]. DM-tradeoff curves for various protocols are derived using a novel technique that is used to calculate DM-tradeoffs in selective fading channels [18] and a set of methods that we describe in this manuscript. The novelty of the methods we use arise from their decreased dependence on the probability distribution of the underlying fading coefficients.

- The NAF protocol requires orthogonality between the relay transmissions [7] and the Rao-Hassibi protocol requires unitary transformations at relays [16], we show that neither orthogonality nor unitary relay transformations is required for achieving DM-tradeoff optimality.
- While the previous works in the literature are concerned with the special case of Rayleigh fading [7], [4], [15], [16], our results are valid for a general class of fading distributions. Finally, we establish the approximately universal code design criterion for the half-duplex relay channel for *any fading distribution*.

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 $\mathbf{Y}$. $\operatorname{rank} \mathbf{X}$ stands for the rank of $\mathbf{X}$. $\operatorname{Tr}(\mathbf{X})$, $\|\mathbf{X}\|_F$, 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. For $N \times N$ positive semi-definite matrices $\mathbf{A}, \mathbf{B}$, $\mathbf{A} \succ \mathbf{B}$ means $\lambda_i(\mathbf{A}) > \lambda_i(\mathbf{B})$ $(i = 0, 1, \ldots, N-1)$, $\mathbf{A} \succeq \mathbf{B}$ is similarly defined. $\mathbf{I}_N$ is the $N \times N$ identity matrix. $\mathbf{0}$ denotes the all zeros matrix of appropriate size. We say that the square matrices $\mathbf{X}$ and $\mathbf{Y}$ are orthogonal if $\langle \mathbf{X}, \mathbf{Y} \rangle = \operatorname{Tr}\big(\mathbf{X}\mathbf{Y}^H\big) = 0$. All logarithms are to the base 2 and $(a)^+ = \max(a, 0)$. $\operatorname{diag}(a_1, \ldots, a_N)$ denotes the $N \times N$ diagonal matrix with a_i on diagonal entry i. The $N \times N$ discrete Fourier transform (DFT) matrix, $\mathbf{F}$, has entries $F_{ln} = \frac{1}{\sqrt{N}} e^{-j\frac{2\pi}{N}(l-1)(n-1)}$. $X \sim \mathcal{CN}(0, \sigma^2)$ stands for a circularly symmetric complex Gaussian random variable (RV) with variance σ^2. $f(\rho) \doteq g(\rho)$ 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.

Preliminaries

The DM-tradeoff realized by a family (one at each SNR ρ) of codebooks $\mathcal{C}_r$ with rate $R = r \log \rho$, where $r \in [0, 1]$, is given by the function

$$d(r) = -\lim_{\rho \to \infty} \frac{\log P_e(\rho, r)}{\log \rho}$$

where $P_e(\rho, r)$ is the error probability obtained through maximum likelihood (ML) decoding. We say that $\mathcal{C}_r$ operates at multiplexing gain r. For a given SNR ρ, the codebook $\mathcal{C}_r(\rho) \in \mathcal{C}_r$ contains ρ^{2Nr} codewords $\mathbf{x}_i$. Following the framework in [17], we define the probability of outage at multiplexing gain r and SNR ρ as the probability of mutual information between the transmit signal $\mathbf{x}$ and the received signal $\mathbf{y}$ given the channel matrix falling below the target rate $r \log \rho$, i.e.

$$P_{\mathcal{O}}(\rho, r) = \mathbb{P}[I(\mathbf{y}; \mathbf{x} | \mathbf{M}) < r \log \rho] \,.$$

II. Protocol Descriptions and System Model

For the fading relay channel shown in Fig. 1, information is to be transmitted from the source terminal ($\mathcal{S}$) to the destination terminal ($\mathcal{D}$) with the assistance of K half-duplex relay terminals. We assume that all terminals are equipped with single antenna transmitters, and cannot receive and transmit simultaneously. The relay terminals perform linear operations on their received signal. In [5], [6] and [19], three different cooperative protocols are identified for the single-relay half-duplex fading channel, implementing varying degrees of receive and transmit collision. In what follows, we generalize these three protocols to the multi-relay case and analyze a fourth protocol, which might be more suited to certain scenarios. Throughout, we assume that communication takes place in time slots.

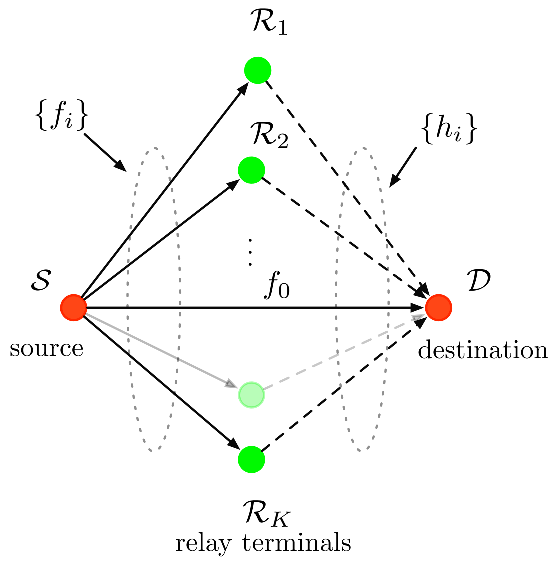

Fig. 1. Two-hop relay network with K single-antenna relay terminals.

Protocol 0 (P0): In this protocol, $\mathcal{D}$ is either unable to or chooses not to receive transmissions from $\mathcal{S}$. In practice, this assumption might be motivated by heavy shadowing of the $\mathcal{S} \to \mathcal{D}$[1] link. In the first time slot, $\mathcal{S}$ communicates with the relay terminals ($\{\mathcal{R}_k\}$). Each relay terminal $\mathcal{R}_k$ processes its received sequence using a linear transformation and transmits the result during the second time slot to $\mathcal{D}$, $\mathcal{S}$ remains silent during the second time slot.

Protocol 1 (P1): $\mathcal{S}$ transmits to both $\{\mathcal{R}_i\}$ and $\mathcal{D}$ during the first time slot. In the second time slot, each $\{\mathcal{R}_i\}$ processes its received sequence using a linear transformation and transmits the result simultaneously to $\mathcal{D}$. $\mathcal{S}$ also transmits new information to $\mathcal{D}$ during the second time slot.

Protocol 2 (P2): This protocol is identical to P1 except that $\mathcal{S}$ does not transmit during the second time slot.

Protocol 3 (P3): This protocol is identical to P1 except that $\mathcal{D}$ does not receive transmission from $\mathcal{S}$ during the first time slot.

The protocols are summarized in Table I. P1, P2 and P3 are the extended versions (to the multi-relay case) of protocols originally proposed for the (single) relay channel in [6], [5] and [19] respectively. The difference of P1 from the NAF protocol of [7] is that in the NAF protocol, the transmissions from different relays are orthogonal, i.e. no two relays transmit at the same time. All K relay terminals are allowed to transmit simultaneously in P1, requiring less overhead in terms of scheduling.

TABLE I
PROTOCOL DESCRIPTIONS

Protocol	1^{st} slot	2^{nd} slot
P0	$\mathcal{S}\to \{ \mathcal{R}_k \}$	$\{ \mathcal{R}_k \} \to \mathcal{D}$
P1	$\mathcal{S}\to \mathcal{D}, \{ \mathcal{R}_k \}$	$\mathcal{S}\to\mathcal{D}, \{ \mathcal{R}_k \} \to \mathcal{D}$
P2	$\mathcal{S}\to \mathcal{D}, \{ \mathcal{R}_k \}$	$\{ \mathcal{R}_k \} \to \mathcal{D}$
P3	$\mathcal{S}\to \{ \mathcal{R}_k \}$	$\mathcal{S}\to\mathcal{D}, \{ \mathcal{R}_k \} \to \mathcal{D}$

Assumptions and Signal Model: We consider a wireless network with $K+2$ single-antenna terminals, where a source terminal $\mathcal{S}$ communicates with a destination terminal $\mathcal{D}$. A set of K half-duplex relay terminals

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

$\mathcal{R}_k$ $(k=1,2,\ldots,K)$ assists $\mathcal{S}$ using one of the four TDMA based protocols. The channels $\mathcal{S}\rightarrow\mathcal{D}$, denoted as f_0, $\mathcal{S}\rightarrow\mathcal{R}_k$, denoted as f_k, and $\mathcal{R}_k\rightarrow\mathcal{D}$, denoted as h_k, $(k=1,2,\ldots,K)$, are i.i.d. with distribution obeying

$$\lim_{z\to 0}\frac{\log(\mathbb{P}[|h_k|^2\leq z])}{\log(z)}=a$$
$$\lim_{z\to\infty}\frac{\log(\mathbb{P}\left[|h_k|^2>z^{\delta}\right])}{\log(z)}=-\infty$$

for all $(k=1,2,\ldots,K)$ and

$$\lim_{z\to 0}\frac{\log(\mathbb{P}[|f_k|^2\leq z])}{\log(z)}=a$$
$$\lim_{z\to\infty}\frac{\log(\mathbb{P}\left[|f_k|^2>z^{\delta}\right])}{\log(z)}=-\infty$$

for all $(k=0,1,\ldots,K)$ and for every $\delta>0$. Many fading distributions, such as the Rayleigh distribution, fall within this class of distributions. 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 under one of the four protocols described earlier. We assume that $\mathcal{S}$ and the relay terminals do not have CSI, whereas $\mathcal{D}$ knows f_0 and f_k, h_k $(k=1,2,\ldots,K)$ perfectly. For simplicity, we further assume perfect synchronization and ignore the impact of pathloss.

Input-output relation for P1: In the following, we derive the signal model for P1. Signal models for other protocols are to follow from this derivation. The vectors $\mathbf{s}[t]$, $\mathbf{r}_k$, $\mathbf{d}[t]\in\mathbb{C}^N$ represent the N-dimensional transmitted signal in time slot t $(t=1,2)$, received signal at $\mathcal{R}_k$ in the first time slot[2], and received signal in time slot t at $\mathcal{D}$, respectively. The transmit signal $\mathbf{s}[t]$ obeys $\mathbb{E}\{(\mathbf{s}[t])^H\mathbf{s}[t]\}=N$ for $(t=1,2)$. Throughout the paper, we assume that $N\geq K+1$. The transmit power constraint at $\mathcal{S}$ is given by $N\rho$, and the power constraint at each relay is given by $N\rho/K$. The vectors $\mathbf{r}_k$ and $\mathbf{d}[1]$ are then given by

$$\mathbf{d}[1]=\sqrt{\rho}f_0\mathbf{s}[1]+\mathbf{z} \tag{1}$$
$$\mathbf{r}_k=\sqrt{\rho}f_k\mathbf{s}[1]+\mathbf{w}_k,\ \ k=1,2,\ldots,K \tag{2}$$

where ρ denotes the average signal-to-noise ratio (SNR) (for all links) and $\mathbf{w}_k$ is the N-dimensional noise vector at $\mathcal{R}_k$ for $(k=1,2,\ldots,K)$ and $\mathbf{z}$ is the N-dimensional noise vector at $\mathcal{D}$, with i.i.d. $\mathcal{CN}(0,1)$ entries. The $\mathbf{w}_k$'s are assumed to be independent across k as well.

Relay k ($\mathcal{R}_k$) applies the linear transformation $\mathbf{G}_k$ to $\mathbf{r}_k$, where the $N\times N$ matrix $\mathbf{G}_k$ obeys $\|\mathbf{G}_k\|_F^2=N$, $(k=1,2,\ldots,K)$. $\mathcal{R}_k$ scales the result to obey the per-relay power constraint $N\rho/K$ and retransmits $\sqrt{\frac{\rho}{NK(1+\mu\rho)}}\mathbf{G}_k\mathbf{r}_k$ where $\mu=\mathbb{E}\left\{|f_k|^2\right\}$. We justify this statement by the following steps

$$\begin{aligned}
&\mathbb{E}\left\{\left\|\sqrt{\frac{\rho}{NK(1+\mu\rho)}}\mathbf{G}_k\mathbf{r}_k\right\|_F^2\right\}\\
&=\frac{\rho}{NK(1+\mu\rho)}\mathbb{E}\left\{\|\mathbf{G}_k\mathbf{r}_k\|_F^2\right\}\\
&\leq\frac{\rho}{NK(1+\mu\rho)}\mathbb{E}\left\{\|\mathbf{G}_k\|_F^2\,\|\mathbf{r}_k\|_F^2\right\}\\
&=\frac{N\rho}{NK(1+\mu\rho)}\left(\mathbb{E}\{\rho|f_k|^2(\mathbf{s}[1])^H\mathbf{s}[1]\}+\mathbb{E}\{\mathbf{w}_k^H\mathbf{w}_k\}\right)\\
&=\frac{\rho}{K(1+\mu\rho)}\left(\rho\mu N+N\right)=\frac{N\rho}{K(1+\mu\rho)}\left(\rho\mu+1\right)=\frac{N\rho}{K}
\end{aligned} \tag{3}$$

[2]Since the relays only receive symbols in the first time slot, we drop the time index t and use $\mathbf{r}_k$ instead of $\mathbf{r}_k[1]$.

where (3) follows from the Cauchy-Schwarz inequality [20, Theorem 5.1.4]. This results in the overall input-output relation in the second time slot

$$\mathbf{d}[2] = \sqrt{\rho} f_0 \mathbf{s}[2] + \frac{\rho}{\sqrt{NK(1+\mu\rho)}} \sum_{k=1}^{K} h_k f_k \mathbf{G}_k \mathbf{s}[1] + \tilde{\mathbf{z}} \tag{4}$$

where the effective noise term $\tilde{\mathbf{z}}$ (when conditioned on $\mathbf{h}$) has mean $\mathbb{E}\{\tilde{\mathbf{z}}|\mathbf{h}\} = \mathbf{0}$ and covariance

$$\mathbf{R}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} = \mathbb{E}\{\tilde{\mathbf{z}}\tilde{\mathbf{z}}^H|\mathbf{h}\} = \mathbf{I}_N + \frac{\rho}{NK(1+\mu\rho)} \sum_{k=1}^{K} |h_k|^2 \mathbf{G}_k \mathbf{G}_k^H.$$

Since $\mathbf{R}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}}$ is Hermitian and positive definite, we can use the spectral theorem for Hermitian matrices to obtain $\mathbf{R}_{\tilde{\mathbf{z}}\tilde{\mathbf{z}}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$ where $\mathbf{U}$ is unitary and $\mathbf{\Lambda}$ is real and diagonal [20, Theorem 4.1.5]. We are interested in the mutual information under the assumption that $\mathcal{D}$ knows all the channels and the channel and noise statistics in the network perfectly, therefore $\mathcal{D}$ can whiten the noise by left-multiplying (4) with $\mathbf{V} = \mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{U}^H$ to obtain the effective input-output relation

$$\mathbf{d}[2] = \sqrt{\rho} f_0 \mathbf{V}\mathbf{s}[2] + \frac{\rho}{\sqrt{NK(1+\mu\rho)}} \sum_{k=1}^{K} h_k f_k \mathbf{V}\mathbf{G}_k \mathbf{s}[1] + \bar{\mathbf{z}}$$

where $\bar{\mathbf{z}}$ (when conditioned on $\mathbf{h}$ and given knowledge of $\mathbf{G}_k$ $k = (1, 2, \ldots, K)$ at the destination) is a circularly symmetric Gaussian noise vector with $\mathbb{E}\{\bar{\mathbf{z}}|\mathbf{h}\} = \mathbf{0}$ and $\mathbb{E}\{\bar{\mathbf{z}}\bar{\mathbf{z}}^H|\mathbf{h}\} = \mathbf{I}_N$. In the remainder of the paper, we shall be interested in the $\rho \to \infty$ case where $\frac{\rho}{\sqrt{NK(1+\mu\rho)}} \approx \sqrt{\frac{\rho}{\mu NK}}$. We can now rewrite the input-output relation for Protocol 1 as

$$\mathbf{y} = \sqrt{\rho}\mathbf{H}_1\mathbf{x} + \mathbf{n}. \tag{5}$$

where the vectors $\mathbf{x}$, $\mathbf{y}$, $\mathbf{n} \in \mathbb{C}^{2N}$ are defined as follows

$$\mathbf{y} = \begin{bmatrix} \mathbf{d}[1] \\ \mathbf{d}[2] \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} \mathbf{s}[1] \\ \mathbf{s}[2] \end{bmatrix} \quad \mathbf{n} = \begin{bmatrix} \tilde{\mathbf{z}} \\ \bar{\mathbf{z}} \end{bmatrix} \tag{6}$$

and the effective channel matrix $\mathbf{H}_1$ is given by

$$\mathbf{H}_1 = \begin{bmatrix} f_0\mathbf{I}_N & \mathbf{0}_N \\ \frac{1}{\sqrt{\mu NK}} \sum_{k=1}^{K} h_k f_k \mathbf{V}\mathbf{G}_k & f_0\mathbf{V} \end{bmatrix}.$$

Input-output relations for other protocols: The respective input output relations for P0, P2 and P3 can be derived from (5) and (6) . The $2N$-dimensional received signal vector $\mathbf{y}$ at destination under protocol p satisfies

$$\mathbf{y} = \sqrt{\rho}\mathbf{H}_p\mathbf{x} + \mathbf{n}. \tag{7}$$

where $\mathbf{x}$, $\mathbf{y}$, $\mathbf{n} \in \mathbb{C}^{2N}$ are as defined in (6) and the effective channel matrix for protocol p, $\mathbf{H}_p$, is

$$\mathbf{H}_0 = \begin{bmatrix} \mathbf{0}_N & \mathbf{0}_N \\ \frac{1}{\sqrt{\mu NK}} \sum_{k=1}^{K} h_k f_k \mathbf{V}\mathbf{G}_k & \mathbf{0}_N \end{bmatrix} \tag{8}$$

$$\mathbf{H}_2 = \begin{bmatrix} f_0\mathbf{I}_N & \mathbf{0}_N \\ \frac{1}{\sqrt{\mu NK}} \sum_{k=1}^{K} h_k f_k \mathbf{V}\mathbf{G}_k & \mathbf{0}_N \end{bmatrix} \tag{9}$$

$$\mathbf{H}_3 = \begin{bmatrix} \mathbf{0}_N & \mathbf{0}_N \\ \frac{1}{\sqrt{\mu NK}} \sum_{k=1}^{K} h_k f_k \mathbf{V}\mathbf{G}_k & f_0\mathbf{V} \end{bmatrix}.$$

As explained in [6], the four protocols investigated differ in the amount of receive collision and degree of broadcasting provided. It is therefore not surprising that the difference between any two protocols amounts to replacing entries of the effective block channel matrix. Since Protocol 1 encapsulates all other protocols, describing other protocols using the effective block channel matrix of Protocol 1 is effectively replacing the appropriate blocks with the all-zero matrix. Finally, we note that the lower triangular nature of the effective block channel matrices is a result of the half-duplex nature of the TDMA protocols considered.

III. Diversity Multiplexing Tradeoff

In this section, we derive the DM-tradeoff curves for the four protocols described under the assumptions stated in the previous section. The maximal mutual information for protocols $p = 0, 1, 2$ and 3 with effective block channel matrices $\mathbf{H}_p$ is achieved by i.i.d. Gaussian codebooks. Assuming an i.i.d. Gaussian codebook with covariance matrix $\mathbf{R}_{\mathbf{xx}} = \mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = \mathbf{I}_{2N}$, we have

$$I(\mathbf{y};\mathbf{x}|\mathbf{H}_p) = \frac{1}{2N}\log\det(\mathbf{I}_{2N} + \rho\mathbf{H}_p\mathbf{H}_p^H) \tag{10}$$

where the factor $1/2N$ is due to considering $2N$ source transmissions.

Deriving the DM-tradeoff curve by evaluating the outage probability directly is difficult since the joint eigenvalue distribution of $\mathbf{H}_p\mathbf{H}_p^H$ does not seem to be available for most of fading distributions. We instead employ a technique developed in [18] for deriving DM-tradeoffs for selective fading channels. We start by rewriting $I(\mathbf{y};\mathbf{x}|\mathbf{H}_p)$

$$I(\mathbf{y};\mathbf{x}|\mathbf{H}_p) = \frac{1}{2N}\log\det\begin{bmatrix} \mathbf{A}_p & \mathbf{B}_p \\ \mathbf{B}_p^H & \mathbf{C}_p \end{bmatrix}. \tag{11}$$

where $\mathbf{A}_p$, $\mathbf{B}_p$ and $\mathbf{C}_p$ are $N \times N$ matrices for each protocol p. We note that the determinant of the above block matrix can be decomposed using the following equation [20] to be

$$\det\begin{bmatrix} \mathbf{A}_p & \mathbf{B}_p \\ \mathbf{B}_p^H & \mathbf{C}_p \end{bmatrix} = \det(\mathbf{A}_p)\det(\mathbf{C}_p - \mathbf{B}_p^H\mathbf{A}_p^{-1}\mathbf{B}_p). \tag{12}$$

We insert (12) into (10) to get

$$\begin{aligned} I(\mathbf{y};\mathbf{x}|\mathbf{H}_p) &= \frac{1}{2N}\log\left(\det(\mathbf{A}_p)\det\left(\mathbf{C}_p - \mathbf{B}_p^H\mathbf{A}_p^{-1}\mathbf{B}_p\right)\right) \\ &= \frac{1}{2N}\left(\sum_{n=0}^{N-1}\log\lambda_n(\mathbf{A}_p) + \sum_{n=0}^{N-1}\log\lambda_n\left(\mathbf{C}_p - \mathbf{B}_p^H\mathbf{A}_p^{-1}\mathbf{B}_p\right)\right). \end{aligned} \tag{13}$$

We note that $\mathbf{I}_{2N} + \mathbf{H}_p\mathbf{H}_p^H$ is a positive definite matrix, i.e. $\mathbf{I}_{2N} + \mathbf{H}_p\mathbf{H}_p^H \succ 0$. Hence, we have that $\mathbf{A}_p \succ 0$ and $\mathbf{C}_p - \mathbf{B}_p^H\mathbf{A}_p^{-1}\mathbf{B}_p \succ 0$ [20, Corollary 7.7.4] and therefore we can use Jensen's inequality to get an upper bound on $I(\mathbf{y};\mathbf{x}|\mathbf{H}_p)$ as follows

$$I(\mathbf{y};\mathbf{x}|\mathbf{H}_p) \le I_J(\mathbf{y};\mathbf{x}|\mathbf{H}_p) \tag{14}$$

where

$$\begin{aligned} &I_J(\mathbf{y};\mathbf{x}|\mathbf{H}_p) \\ &= \frac{1}{2}\log\frac{1}{N}\sum_{n=0}^{N-1}\lambda_n(\mathbf{A}_p) + \frac{1}{2}\log\frac{1}{N}\sum_{n=0}^{N-1}\lambda_n\left(\mathbf{C}_p - \mathbf{B}_p^H\mathbf{A}_p^{-1}\mathbf{B}_p\right) \\ &= \frac{1}{2}\log\frac{1}{N}\mathrm{Tr}(\mathbf{A}_p) + \frac{1}{2}\log\frac{1}{N}\mathrm{Tr}\left(\mathbf{C}_p - \mathbf{B}_p^H\mathbf{A}_p^{-1}\mathbf{B}_p\right) \\ &= \frac{1}{2}\log\frac{\mathrm{Tr}(\mathbf{A}_p)}{N} + \frac{1}{2}\log\left(\frac{\mathrm{Tr}(\mathbf{C}_p)}{N} - \frac{\mathrm{Tr}\left(\mathbf{B}_p^H\mathbf{A}_p^{-1}\mathbf{B}_p\right)}{N}\right). \end{aligned} \tag{15}$$

We will show that the DM-tradeoff of $I(\mathbf{y};\mathbf{x}|\mathbf{H}_p)$ equals that of $I_J(\mathbf{y};\mathbf{x}|\mathbf{H}_p)$. The importance of this bounding technique lies in the fact that the resulting expressions can be analyzed using simple probability theory. Next, we specify the protocol-dependent matrices $\mathbf{A}_p, \mathbf{B}_p$ and $\mathbf{C}_p$ for each protocol. These matrices do not describe any particular physical attribute of the channel by themselves, but their introduction simplifies the ensuing analysis. For Protocol 0, we have

$$\mathbf{A}_0 = \mathbf{I}_N \tag{16}$$

$$\mathbf{B}_0 = \mathbf{0}_N \tag{17}$$

$$\mathbf{C}_0 = \mathbf{I}_N + \frac{\rho}{\mu N K}\sum_{k=1}^{K}\sum_{j=1}^{K} h_k f_k h_j^* f_j^* \mathbf{V}\mathbf{G}_k\mathbf{G}_j^H\mathbf{V}^H. \tag{18}$$

Then for Protocol 1 we obtain

$$\mathbf{A}_1 = \left(1+\rho|f_0|^2\right)\mathbf{I}_N \tag{19}$$

$$\mathbf{B}_1 = \frac{\rho}{\sqrt{\mu N K}}\sum_{k=1}^{K} h_k^* f_k^* f_0 \mathbf{G}_k^H\mathbf{V}^H \tag{20}$$

$$\mathbf{C}_1 = \mathbf{C}_0 + \rho|f_0|^2\mathbf{\Lambda}^{-1}. \tag{21}$$

Since the P0 and P1 include all possible orderings of $\mathcal{S}$ and $\mathcal{R}$ transmissions, we can describe the matrices $\mathbf{A}_p, \mathbf{B}_p$ and $\mathbf{C}_p$ for P2 and P3 in terms of the matrices of P0 and P1. Therefore, the mutual information expression for P2 can be described using

$$\mathbf{A}_2 = \mathbf{A}_1 \tag{22}$$

$$\mathbf{B}_2 = \mathbf{B}_1 \tag{23}$$

$$\mathbf{C}_2 = \mathbf{C}_0. \tag{24}$$

Similarly, for P3, we have

$$\mathbf{A}_3 = \mathbf{A}_0 \tag{25}$$

$$\mathbf{B}_3 = \mathbf{B}_0 \tag{26}$$

$$\mathbf{C}_3 = \mathbf{C}_1. \tag{27}$$

In the following, we will need the normalized codeword difference matrix $\mathbf{\Phi}_p(\mathbf{\Delta x})$ for protocol p defined as

$$\mathbf{\Phi}_0(\mathbf{\Delta x}) = \begin{bmatrix} \mathbf{G}_1\mathbf{\Delta x}_1 & \mathbf{G}_2\mathbf{\Delta x}_1 & \cdots & \mathbf{G}_K\mathbf{\Delta x}_1 \end{bmatrix}$$

$$\mathbf{\Phi}_1(\mathbf{\Delta x}) = \begin{bmatrix} \mathbf{\Delta x}_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{G}_1\mathbf{\Delta x}_1 & \mathbf{G}_2\mathbf{\Delta x}_1 & \cdots & \mathbf{G}_K\mathbf{\Delta x}_1 & \mathbf{\Delta x}_2 \end{bmatrix}$$

$$\mathbf{\Phi}_2(\mathbf{\Delta x}) = \begin{bmatrix} \mathbf{\Delta x}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{G}_1\mathbf{\Delta x}_1 & \mathbf{G}_2\mathbf{\Delta x}_1 & \cdots & \mathbf{G}_K\mathbf{\Delta x}_1 \end{bmatrix}$$

$$\mathbf{\Phi}_3(\mathbf{\Delta x}) = \begin{bmatrix} \mathbf{G}_1\mathbf{\Delta x}_1 & \mathbf{G}_2\mathbf{\Delta x}_1 & \cdots & \mathbf{G}_K\mathbf{\Delta x}_1 & \mathbf{\Delta x}_2 \end{bmatrix}$$

where $\mathbf{\Delta x} = [\mathbf{\Delta x}_1\ \mathbf{\Delta x}_2]^T = \tilde{\mathbf{x}} - \hat{\mathbf{x}}$ for any two codewords $\tilde{\mathbf{x}}, \hat{\mathbf{x}}$ in codebook $\mathcal{C}_r(\rho)$. We will need the Gram matrix $\mathbf{K}(\mathbf{A})$ defined as

$$\mathbf{K}(\mathbf{A}) = \begin{bmatrix} \mathrm{Tr}\left(\mathbf{G}_1\mathbf{A}\mathbf{G}_1^H\right) & \cdots & \mathrm{Tr}\left(\mathbf{G}_K\mathbf{A}\mathbf{G}_1^H\right) \\ \vdots & \ddots & \vdots \\ \mathrm{Tr}\left(\mathbf{G}_1\mathbf{A}\mathbf{G}_K^H\right) & & \mathrm{Tr}\left(\mathbf{G}_K\mathbf{A}\mathbf{G}_K^H\right) \end{bmatrix}. \tag{28}$$

for a $N \times N$ positive definite matrix $\mathbf{A}$. Our main finding is given by the next theorem.

Theorem 1 (DM-Tradeoff for Protocols 0-3): For the half-duplex relay channel operating under protocol p and described by (7), the DM-tradeoff curve is given by

$$d_p(r) = \begin{cases} aK(1-2r)^+, & \text{if } p=0,\ r\in[0,1/2] \\ aK(1-2r)^+ + a(1-r)^+, & \text{if } p=1,\ r\in[0,1] \\ a(K+1)(1-2r)^+, & \text{if } p=2,\ r\in[0,1/2] \\ a(K+1)(1-2r)^+, & \text{if } p=3,\ r\in[0,1/2]. \end{cases}$$

Let $\{\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_K\}$ be a set of transformation matrices and $\mathcal{C}_r$ be a family of codebooks such that for any codebook $\mathcal{C}_r(\rho) \in \mathcal{C}_r$ and any two codewords $\tilde{\mathbf{x}}, \hat{\mathbf{x}} \in \mathcal{C}_r(\rho)$ the conditions

$$\operatorname{rank} \mathbf{\Phi}_p(\Delta\mathbf{x}) = \begin{cases} K & \text{if } p=0, \\ K+1 & \text{if } p=1, \\ K & \text{if } p=2, \\ K+1 & \text{if } p=3. \end{cases} \tag{29}$$

$$\operatorname{rank} \mathbf{K}(\mathbf{I}_N) = K \tag{30}$$

hold. Then the ML decoding error probability under protocol p satisfies

$$P_e(\rho, r) \doteq \rho^{-d_p(r)}. \tag{31}$$

Proof: See Appendix A ∎

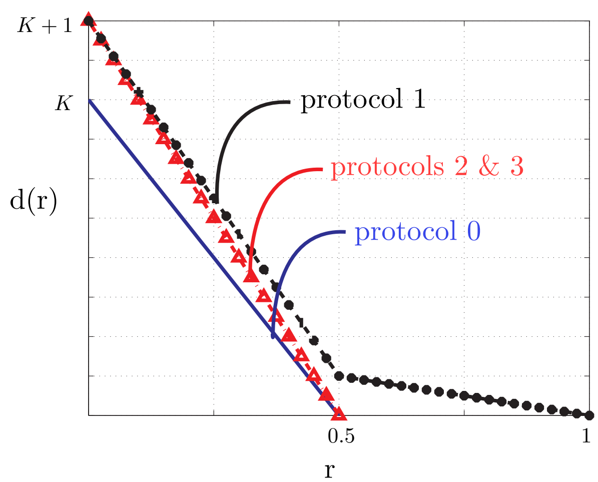

Fig. 2. Diversity-Multiplexing Tradeoff for protocols 0,1,2 and 3.

Discussion

Theorem 1 shows that P1 dominates P0, P2 and P3 in terms of DM-tradeoff performance. Further, P1 achieves the entire optimal DM-tradeoff of the half-duplex relay channel for the $K = 1$ case [7]. Yang and Belfiore show that the DMT curve for amplify-and-forward protocols in which the relays transmit and receive in exactly half of the source transmission is upper bounded by

$$d(r) = K(1-2r)^+ + (1-r)^+ \; r \in [0,1] \tag{32}$$

for Rayleigh fading channels [8]. It follows that for Rayleigh fading channels, P1 combined with any family of codebooks satisfying the rank conditions in Theorem 1 achieves this upper bound. The DMT for the NAF protocol, proposed by Azarian *et al.*, also matches this upper bound in Rayleigh fading channels [7], however the NAF protocol attains orthogonality between the relay transmissions by allowing only one relay to transmit in any given time slot (concurrently with the source transmission). The NAF protocol is a special case of our formulation with $\mathbf{G}_k$'s being diagonal matrices with

$$[\mathbf{G}_k]_{ii} = \begin{cases} \sqrt{N}, \text{ if } \ i = k \\ 0, \text{ otherwise.} \end{cases} \tag{33}$$

It can be seen that for the NAF protocol, the full rank condition $\operatorname{rank} \mathbf{\Phi}_1(\mathbf{\Delta x}) = K + 1$ amounts to $[\mathbf{\Delta x}_1]_i \neq 0$ for all $(i = 1, 2, \ldots, N)$ and $[\mathbf{\Delta x}_2]_1 \neq 0$ for every codebook $\mathcal{C}_r(\rho) \in \mathcal{C}_r$ and any two codewords $\tilde{\mathbf{x}}, \hat{\mathbf{x}} \in \mathcal{C}_r(\rho)$. A relaying protocol is called an orthogonal scheme, if the set of transformation matrices $\mathcal{G} = \{\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_K\}$ is orthogonal in the inner product space defined by $\langle \mathbf{A}, \mathbf{B} \rangle = \operatorname{Tr}\big(\mathbf{A}\mathbf{B}^H\big)$, i.e.,

$$\langle \mathbf{G}_k, \mathbf{G}_j \rangle = \begin{cases} N, \text{ if } \ k = j \\ 0, \text{ otherwise} \end{cases}$$

for all $(j, k = 1, 2, \ldots, K)$. Therefore, the NAF protocol is an orthogonal scheme and it follows trivially that $\mathbf{K}(\mathbf{I}_N) = \mathbf{I}_K$. Our results show that there exists a whole family of linear processing schemes that achieve the DM-tradeoff curve of the NAF protocol given in Theorem 1. Furthermore, within the set of half duplex protocols, this is the best possible DMT.

P2 and P3 achieve the same tradeoff even though P2 maximizes the degree of broadcasting whereas P3 maximizes the degree of receive collision in the network. This suggests that (for this network) maximizing the degree of broadcasting and the degree of receive collision has the same DM-tradeoff performance. Further, the dominance of P1 shows that it is possible to maximize both the degree of broadcasting *a*nd the degree of receive collision in the network, and that this is the optimal mode of transmission from a DM-tradeoff point of view. P2 and P3 achieve a maximum multiplexing gain of $\frac{1}{2}$. This is due to the fact that $\mathcal{S}$ terminal does not transmit to $\mathcal{D}$ terminal in all available time slots. However, it turns out that for P2 and P3 the direct link does provide extra diversity gain. Due to half-duplex constraint and the lack of a direct link, P0 reaches a multiplexing gain of $\frac{1}{2}$.

It can be shown that the rank conditions on $\mathbf{\Phi}_0(\mathbf{\Delta x})$ and $\mathbf{\Phi}_3(\mathbf{\Delta x})$ imply $\operatorname{rank} \mathbf{K}(\mathbf{I}_N) = K$, however this is not the case for P1 and P3. As a simple example, consider a two relay scenario where relays are given the transformation matrices $\mathbf{G}_1 = \mathbf{G}_2 = \mathbf{I}_N$. Then $\mathbf{\Phi}_1(\mathbf{\Delta x})$ is given by

$$\mathbf{\Phi}_1(\mathbf{\Delta x}) = \left[\begin{array}{ccc} \mathbf{\Delta x}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{\Delta x}_1 & \mathbf{\Delta x}_1 & \mathbf{\Delta x}_2 \end{array} \right]$$

which can be made to have full rank by an appropriate choice of $\mathbf{\Delta x}_1$ and $\mathbf{\Delta x}_2$, but $\mathbf{K}(\mathbf{I})$ has rank 1 since

$$\mathbf{K}(\mathbf{I}_N) = N \left[\begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array} \right].$$

Unitary Designs

The matrices $\mathbf{G}_k$ can be chosen to be unitary for all $(k = 1, 2, \ldots, K)$. We give some specific examples of unitary designs that are also orthogonal in the next two subsections.

Cyclic Delay Diversity: We start by noting that the cyclic delay diversity scheme [12] can be cast into our framework by setting $\mathbf{G}_k = \mathbf{P}_k$ where $\mathbf{P}_k$ denotes the permutation matrix that, when applied to a vector $\mathbf{x}$, cyclically shifts the elements in $\mathbf{x}$ up by $k-1$ positions.With

$$\langle \mathbf{P}_k, \mathbf{P}_j \rangle = \begin{cases} N, & k = j \\ 0, & k \neq j \end{cases} \tag{34}$$

the condition $\operatorname{rank} \mathbf{\Phi}_0(\mathbf{\Delta x}) = K$ takes a particularly simple form, namely $(\mathbf{F\Delta x})_i \neq 0$ for all $i \in \{1, 2, \ldots, N\}$. To see this note that $\operatorname{rank} \mathbf{\Phi}_0(\mathbf{\Delta x}) = \operatorname{rank} \mathbf{F\Phi}_0(\mathbf{\Delta x})$ and $\mathbf{P}_k = \mathbf{F}^H \mathbf{\Lambda}_k \mathbf{F}$ where

$$\mathbf{\Lambda}_k = \operatorname{diag}\left(e^{j\theta_k[0]}, e^{j\theta_k[1]}, \cdots, e^{j\theta_k[N-1]}\right) \tag{35}$$

with $\theta_k[n] = \frac{2\pi n(k-1)}{N}$. Next, we have

$$\operatorname{rank} \mathbf{F\Phi}_0(\mathbf{\Delta x}) = \operatorname{rank} \mathbf{\Sigma}\, [\mathbf{l}_1\; \mathbf{l}_2\; \cdots\; \mathbf{l}_K] \tag{36}$$

where $\mathbf{\Sigma} = \operatorname{diag}\left((\mathbf{F}\Delta\mathbf{x})_1, (\mathbf{F}\Delta\mathbf{x})_2, \ldots, (\mathbf{F}\Delta\mathbf{x})_N\right)$ and $[\mathbf{l}_k]_i = e^{j\theta_k[i-1]}, i = 1, 2, \ldots, N, k = 1, 2, \ldots K$. As a consequence of (34), the columns of the matrix $[\mathbf{l}_1\; \mathbf{l}_2\; \cdots\; \mathbf{l}_K]$ are orthogonal and hence $\operatorname{rank} \mathbf{F\Phi}(\Delta\mathbf{x}) = K$ if $\mathbf{\Sigma}$ has full rank which is the case if $(\mathbf{F}\Delta\mathbf{x})_i \neq 0$ for all $i \in \{1, 2, \ldots, N\}$.

Phase-rolling: In the case of phase-rolling [13], [14], we have $\mathbf{G}_k = \mathbf{\Lambda}_i$. Again, the condition $\operatorname{rank} \mathbf{\Phi}_0(\mathbf{\Delta x}) = K$ takes a particularly simple form, namely $(\Delta\mathbf{x})_i \neq 0$ for all $i \in \{1, 2, \ldots, N\}$. The proof of this statement follows by considering $\mathbf{\Phi}_0(\mathbf{\Delta x})$ directly, putting $\operatorname{rank} \mathbf{\Phi}_0(\mathbf{\Delta x})$ into the form of the right-hand side of (36) and applying the remaining steps in the argument for the cyclic delay diversity case. While the (numerical) results in [12], [13] are for the $r = 0$ case, our analysis reveals optimality of cyclic delay diversity and phase-rolling for the entire DM-tradeoff curve, provided the codebooks satisfy the full-rank condition in Theorem 1. We finally note that cyclic delay diversity and phase-rolling are time-frequency duals of each other in the sense that the linear transformation matrices for the two schemes obey $\mathbf{G}_k = \mathbf{F}\mathbf{P}_k\mathbf{F}^H$.

Relation to approximately universal codes [21]: For the half-duplex relay channel investigated in this paper, a family of codes $\mathcal{C}_r$ achieves the DM-tradeoff curve of the underlying channel if

$$\eta_{\min}(\mathbf{\Phi}_0(\mathbf{\Delta x})^H \mathbf{\Phi}_0(\mathbf{\Delta x})) \dot{>} \rho^{-2r} \tag{37}$$

$$\eta_{\min}(\mathbf{\Phi}_1(\mathbf{\Delta x})^H \mathbf{\Phi}_1(\mathbf{\Delta x})) \dot{>} \rho^{-r} \tag{38}$$

$$\eta_{\min}(\mathbf{\Phi}_2(\mathbf{\Delta x})^H \mathbf{\Phi}_2(\mathbf{\Delta x})) \dot{>} \rho^{-r} \tag{39}$$

$$\eta_{\min}(\mathbf{\Phi}_3(\mathbf{\Delta x})^H \mathbf{\Phi}_3(\mathbf{\Delta x})) \dot{>} \rho^{-2r} \tag{40}$$

where $\eta_{\min}$ is the smallest eigenvalue of $(\mathbf{\Phi}_p(\Delta\mathbf{x}))^H \mathbf{\Phi}_p(\Delta\mathbf{x})$ over all $\Delta\mathbf{x} = \tilde{\mathbf{x}} - \hat{\mathbf{x}}$ with $\tilde{\mathbf{x}}, \hat{\mathbf{x}} \in \mathcal{C}_r(\rho)$. This result follows immediately from (93) in the proof of Theorem 1. However, the choice of the transformation matrices $\mathbf{G}_k$ effects the statistics of the underlying channel and therefore the DM-tradeoff curve. Hence, a further condition that guarantees the best possible DM-tradeoff curve for the underlying channel is required; namely we need

$$\operatorname{rank} \mathbf{K}(\mathbf{I}_N) = K. \tag{41}$$

Based on (37) and (41), we can conclude (using the same arguments as in Sec. IV. A in [18]) that any family of codes $\mathcal{C}_r$ satisfying (37) and (41) will also be approximately universal in the sense of [21, Th. 3.1]. We point out that the conditions $\eta_{\min}(\mathbf{\Phi}_0(\mathbf{\Delta x})^H \mathbf{\Phi}_0(\mathbf{\Delta x})) \dot{>} \rho^{-2r}$ and $\eta_{\min}(\mathbf{\Phi}_3(\mathbf{\Delta x})^H \mathbf{\Phi}_3(\mathbf{\Delta x})) \dot{>} \rho^{-2r}$ imply (41) for protocols 0 and 3, respectively. In [16], the authors consider unitary $\mathbf{G}_k$ and claim (without proof) that choosing "different" unitary matrices is sufficient for extracting diversity benefits. However, our analysis reveals that choosing different unitary matrices is not always enough (even when there is a rank condition on the normalized codeword difference); the unitary matrices should be chosen according to (41). For protocols 0 and 3, enforcing a rank condition on the normalized code difference matrix will guarantee (41), however this is not true for all possible protocols.

IV. CONCLUSIONS

We identify the conditions on linear processing at the relay level which guarantee DM-tradeoff achievability for various TDMA based protocols. Unlike the previous works in the literature, our results are valid for a large class of fading distributions, and therefore system architectures using these design criteria are robust against channel measurement errors. Relative to the previously proposed linear relaying architectures (for Rayleigh fading channels), the class of relaying schemes we introduce is less restrictive and encompasses a larger number of design choices.

Appendix A
Proof of Theorem

Proof: In the following, we will give a proof of the theorem for protocol 1. The proof of the theorem for other protocols can be derived from this proof. We need to evaluate (15) for P1

$$I_{\mathcal{J}}(\mathbf{y};\mathbf{x}|\mathbf{H}_1) = \frac{1}{2}\left(\log\frac{\mathrm{Tr}(\mathbf{A}_1)}{N} + \log\left(\frac{\mathrm{Tr}(\mathbf{C}_1)}{N} - \frac{\mathrm{Tr}\left(\mathbf{B}_1^H\mathbf{A}_1^{-1}\mathbf{B}_1\right)}{N}\right)\right). \tag{42}$$

First, $\mathrm{Tr}(\mathbf{A}_1)/N$ simply evaluates to

$$\frac{\mathrm{Tr}(\mathbf{A}_1)}{N} = 1+\rho|f_0|^2. \tag{43}$$

Next, for $\mathrm{Tr}(\mathbf{C}_1)/N$ we obtain

$$\frac{1}{N}\mathrm{Tr}(\mathbf{C}_1) = \frac{1}{N}\mathrm{Tr}(\mathbf{C}_0) + \frac{\rho|f_0|^2}{N}\mathrm{Tr}\left(\Lambda^{-1}\right) = 1+\frac{\rho}{\mu N^2 K}\left\|\mathbf{V}\left(\sum_{k=1}^{K} h_k f_k \mathbf{G}_k\right)\right\|_F^2 + \frac{\rho|f_0|^2}{N}\mathrm{Tr}\left(\mathbf{\Lambda}^{-1}\right) \tag{44}$$

since

$$\begin{aligned}\frac{1}{N}\mathrm{Tr}(\mathbf{C}_0) &= \frac{1}{N}\mathrm{Tr}(\mathbf{I}_N) + \mathrm{Tr}\left(\frac{\rho}{\mu N^2 K}\sum_{k=1}^{K}\sum_{j=1}^{K} h_k f_k h_j^* f_j^* \mathbf{V}\mathbf{G}_k\mathbf{G}_j^H\mathbf{V}^H\right)\\ &= 1+\frac{\rho}{\mu N^2 K}\left\|\mathbf{V}\left(\sum_{k=1}^{K} h_k f_k \mathbf{G}_k\right)\right\|_F^2.\end{aligned}$$

Finally, to evaluate $\mathrm{Tr}\left(\mathbf{B}^H\mathbf{A}^{-1}\mathbf{B}\right)$, we need to calculate the inverse of $\mathbf{A}_1$, which can be readily seen to be

$$\mathbf{A}_1^{-1} = \frac{1}{1+\rho|f_0|^2}\mathbf{I}_N. \tag{45}$$

Hence, $\mathrm{Tr}\left(\mathbf{B}^H\mathbf{A}^{-1}\mathbf{B}\right)/N$ is

$$\frac{1}{N}\mathrm{Tr}\left(\mathbf{B}^H\mathbf{A}^{-1}\mathbf{B}\right) = \frac{1}{N(1+\rho|f_0|^2)}\mathrm{Tr}\left(\mathbf{B}^H\mathbf{B}\right) \tag{46}$$

$$= \frac{\rho^2|f_0|^2}{\mu N^2 K(1+\rho|f_0|^2)}\mathrm{Tr}\left(\sum_{k=1}^{K}\sum_{j=1}^{K} h_k f_k h_j^* f_j^* \mathbf{V}\mathbf{G}_k\mathbf{G}_j^H\mathbf{V}^H\right)$$

$$= \frac{\rho^2|f_0|^2}{\mu N^2 K(1+\rho|f_0|^2)}\left\|\mathbf{V}\left(\sum_{k=1}^{K} h_k f_k \mathbf{G}_k\right)\right\|_F^2. \tag{47}$$

Then, we further obtain

$$\frac{1}{N}\mathrm{Tr}\left(\mathbf{C}_1 - \mathbf{B}_1^H\mathbf{A}_1^{-1}\mathbf{B}_1\right) = 1+\frac{\rho|f_0|^2}{N}\mathrm{Tr}\left(\mathbf{\Lambda}^{-1}\right) + \frac{\rho}{\mu N^2 K(1+\rho|f_0|^2)}\left\|\mathbf{V}\left(\sum_{k=1}^{K} h_k f_k \mathbf{G}_k\right)\right\|_F^2.$$

Denoting $\tilde{\mathbf{H}}_1 = \sum_{k=1}^{K} h_k f_k \mathbf{G}_k$ and combining the previous results yields

$$\begin{aligned} & I_{\mathcal{J}}(\mathbf{y};\mathbf{x}|\mathbf{H}_1) \\ & = \frac{1}{2}\log\left(1+\rho|f_0|^2\left(1+\frac{\mathrm{Tr}(\mathbf{\Lambda}^{-1})}{N}(1+\rho|f_0|^2)\right)+\frac{\rho}{\mu N^2 K}\left\|\mathbf{V}\tilde{\mathbf{H}}_1\right\|_F^2\right). \end{aligned}$$

Next, we compute a lower bound on $\mathbb{P}[\mathcal{J}]$ by upper-bounding $I_{\mathcal{J}}(\mathbf{y};\mathbf{x}|\mathbf{H}_1)$ as follows

$$I_{\mathcal{J}}(\mathbf{y};\mathbf{x}|\mathbf{H}_1) \leq I_{\mathcal{U}}$$

where

$$I_{\mathcal{U}} = \frac{1}{2}\log\left(\left(1+\rho|f_0|^2\right)^2 + \frac{\rho}{\mu}\tilde{\mathbf{h}}^H\tilde{\mathbf{h}}\right) \tag{48}$$

with $\tilde{\mathbf{h}} = \mathbf{h}\circ\mathbf{f}$. We have used the fact that $\mathrm{Tr}(\mathbf{\Lambda}^{-1}) \leq N$ in addition to the Cauchy-Schwarz inequality to arrive at the upper bound (48). To see that $\mathrm{Tr}(\mathbf{\Lambda}^{-1}) \leq N$, note that

$$\begin{aligned} \lambda_i(\mathbf{R}_{\tilde{z}\tilde{z}}) &= 1 + \frac{\rho}{NK(1+\mu\rho)}\lambda_i\left(\sum_{k=1}^{K}|h_k|^2\mathbf{G}_k\mathbf{G}_k^H\right) \\ &\geq 1. \end{aligned} \tag{49}$$

(49) follows from Weyl's theorem [20, 4.3.1] since $\lambda_i\left(\sum_{k=1}^{K}|h_k|^2\mathbf{G}_k\mathbf{G}_k^H\right) \geq 0$ for $(k=1,2,\ldots,K)$ and $(i = 1,2,\ldots,N)$. Using $\lambda_i(\mathbf{\Lambda}^{-1}) = \frac{1}{\lambda_{N-i}(\mathbf{R}_{\tilde{z}\tilde{z}})}$, we have $\lambda_i(\mathbf{\Lambda}^{-1}) \leq 1$, and therefore $\mathrm{Tr}(\mathbf{\Lambda}^{-1}) = \sum_{i=1}^{N}\lambda_i(\mathbf{\Lambda}^{-1}) \leq N$. To see that $\frac{\rho}{\mu N^2 K}\left\|\mathbf{V}\tilde{\mathbf{H}}_1\right\|_F^2 \leq \frac{\rho}{\mu}\tilde{\mathbf{h}}^H\tilde{\mathbf{h}}$, consider rewriting

$$\frac{\rho}{\mu N^2 K}\left\|\mathbf{V}\tilde{\mathbf{H}}_1\right\|_F^2 = \frac{\rho}{\mu N^2 K}\tilde{\mathbf{h}}^H\mathbf{K}(\mathbf{R}_{\tilde{z}\tilde{z}}^{-1})\tilde{\mathbf{h}} \tag{50}$$

$$\leq \frac{\rho}{\mu N^2 K}\lambda_{\max}(\mathbf{K}(\mathbf{R}_{\tilde{z}\tilde{z}}^{-1}))\tilde{\mathbf{h}}^H\tilde{\mathbf{h}} \tag{51}$$

$$\leq \frac{\rho}{\mu N^2 K}\mathrm{Tr}\big(\mathbf{K}(\mathbf{R}_{\tilde{z}\tilde{z}}^{-1})\big)\,\tilde{\mathbf{h}}^H\tilde{\mathbf{h}} \tag{52}$$

$$\leq \frac{\rho}{\mu N^2 K}N^2K\tilde{\mathbf{h}}^H\tilde{\mathbf{h}} = \frac{\rho}{\mu}\tilde{\mathbf{h}}^H\tilde{\mathbf{h}} \tag{53}$$

where $\mathbf{K}(\mathbf{A})$ is defined in (28). Now, (51) follows from the Rayleigh-Ritz inequality [20, Theorem 4.2.2], (52) follows since the Gram matrix $\mathbf{K}(\mathbf{R}_{\tilde{z}\tilde{z}}^{-1})$ has non-negative eigenvalues by [20, Theorem 7.2.10]. For (53), we have used $\mathrm{Tr}\big(\mathbf{G}_k\mathbf{R}_{\tilde{z}\tilde{z}}^{-1}\mathbf{G}_k^H\big) = \mathrm{Tr}\big(\mathbf{G}_k^H\mathbf{G}_k\mathbf{R}_{\tilde{z}\tilde{z}}^{-1}\big)$ and the positive semi-definiteness of matrices $\mathbf{G}_k^H\mathbf{G}_k$ $(k = 1,2,\ldots,K)$ and $\mathbf{R}_{\tilde{z}\tilde{z}}^{-1}$ to deduce $\mathrm{Tr}\big(\mathbf{G}_k^H\mathbf{G}_k\mathbf{R}_{\tilde{z}\tilde{z}}^{-1}\big) \leq \mathrm{Tr}\big(\mathbf{G}_k^H\mathbf{G}_k\big)\,\mathrm{Tr}\big(\mathbf{R}_{\tilde{z}\tilde{z}}^{-1}\big) \leq N^2$ and hence $\mathrm{Tr}\big(\mathbf{K}(\mathbf{R}_{\tilde{z}\tilde{z}}^{-1})\big) \leq N^2K$. Therefore

$$I_{\mathcal{J}}(\mathbf{y};\mathbf{x}|\mathbf{H}_1) \leq \frac{1}{2}\log\left(\left(1+\rho|f_0|^2\right)^2 + \frac{\rho}{\mu}\tilde{\mathbf{h}}^H\tilde{\mathbf{h}}\right) \tag{54}$$

Define the event

$$\mathcal{U} = \{f_0, \tilde{\mathbf{h}} : \left(1+\rho|f_0|^2\right)^2 + \frac{\rho}{\mu}\tilde{\mathbf{h}}^H\tilde{\mathbf{h}} < \rho^{2r}\}$$

Next, we calculate $\mathbb{P}[\mathcal{U}]$

$$\begin{aligned}
&\mathbb{P}[\mathcal{U}] \\
&\geq \mathbb{P}\left[\left(1+\rho|f_0|^2\right)^2 < \frac{\rho^{2r}}{2}\right] \mathbb{P}\left[\frac{\rho}{\mu}\tilde{\mathbf{h}}^H\tilde{\mathbf{h}} < \frac{\rho^{2r}}{2}\right] && (55)\\
&\doteq \mathbb{P}\left[\ 1+\rho|f_0|^2 < \frac{\rho^{r}}{2}\right] \mathbb{P}\left[\frac{\rho}{\mu}\tilde{\mathbf{h}}^H\tilde{\mathbf{h}} < \frac{\rho^{2r}}{2}\right] && (56)\\
&\doteq \mathbb{P}\left[\ |f_0|^2 < \rho^{r-1}\right] \mathbb{P}\left[\tilde{\mathbf{h}}^H\tilde{\mathbf{h}} < \rho^{2r-1}\right] && (57)\\
&\doteq \mathbb{P}\left[\ |f_0|^2 < \rho^{r-1}\right] \mathbb{P}\left[\sum_{k=1}^{K}|h_k|^2|f_k|^2 < \rho^{2r-1}\right] && (58)\\
&\doteq \mathbb{P}\left[\ |f_0|^2 < \rho^{r-1}\right] \left(\mathbb{P}\left[|h_1|^2|f_1|^2 < \rho^{2r-1}\right]\right)^K && (59)\\
&\doteq \rho^{-a(1-r)^+}\left(\rho^{-a(1-2r)^+}\right)^K && (60)\\
&= \rho^{-a(1-r)^+ - Ka(1-2r)^+} && (61)
\end{aligned}$$

where (55) follows from the independence of f_0 and $\tilde{\mathbf{h}}$, (57) is true since multiplicative and additive constants do not alter the limiting behavior in (57), (58) is by definition of $\tilde{\mathbf{h}}^H\tilde{\mathbf{h}}$. (59) follows due to independence of f_ih_i, and final step (60) is justified by Theorem 2 in Appendix B.

Now, we find an upper bound on Jensen outage probability. We start by writing $|h_k|^2 = \rho^{-u_k}$ and $|f_k|^2 = \rho^{-v_k}$, where u_k and v_k are RVs; the choice of this transformation will become clear later. We define the event $\mathcal{A} = \{u_1, \ldots, u_K, v_0, v_1, \ldots, v_K | u_k \geq 0, v_0 \geq 0, v_k \geq 0\ \forall k \in \{1, 2, \ldots, K\}\}$ and its complementary event $\bar{\mathcal{A}}$ as the event where at least one u_k or v_k is strictly negative. Using the law of total probability, we can write

$$\mathbb{P}[\mathcal{J}] = \mathbb{P}[\mathcal{A}]\,\mathbb{P}[\mathcal{J}|\mathcal{A}] + \mathbb{P}\big[\bar{\mathcal{A}}\big]\,\mathbb{P}\big[\mathcal{J}|\bar{\mathcal{A}}\big] \tag{62}$$

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

$$\begin{aligned}
\mathbb{P}[\mathcal{J}] &\leq \mathbb{P}[\mathcal{A}]\,\mathbb{P}[\mathcal{J}|\mathcal{A}] + \mathbb{P}\big[\bar{\mathcal{A}}\big] && (63)\\
\mathbb{P}[\mathcal{J}] &\ \dot{\leq}\ \mathbb{P}[\mathcal{A}]\,\mathbb{P}[\mathcal{J}|\mathcal{A}] && (64)\\
\mathbb{P}[\mathcal{J}] &\ \dot{\leq}\ \mathbb{P}[\mathcal{J}|\mathcal{A}] && (65)
\end{aligned}$$

where (64) follows from the definition of the u_k and the v_k, their independence and by noting that $\mathbb{P}\big[\bar{\mathcal{A}}\big]$ decays exponentially fast in ρ and therefore can be ignored. The inequality (65) results from $\lim\limits_{\rho\to\infty} \frac{\log \mathbb{P}[\mathcal{A}]}{\log \rho} = 0$. Next, we have

$$\begin{aligned}
\lambda_i(\mathbf{R}_{\tilde{z}\tilde{z}}) \leq& 1 + \frac{\rho}{NK(1+\mu\rho)}\sum_{k=1}^{K}|h_k|^2 \mathrm{Tr}\big(\mathbf{G}_k\mathbf{G}_k^H\big) && (66)\\
=& 1 + \frac{\rho}{NK(1+\mu\rho)} N \sum_{k=1}^{K}|h_k|^2 && (67)\\
=& 1 + \frac{\rho}{K(1+\mu\rho)}\sum_{k=1}^{K}|h_k|^2 && (68)
\end{aligned}$$

where (66) follows from Weyl's Theorem [20, Theorem 4.3.2] and the fact that $\mathbf{G}_k\mathbf{G}_k^H$'s are positive semi-definite Hermitian matrices. The eigenvalues of $\mathbf{R}_{\tilde{z}\tilde{z}}^{-1}$ are therefore lower bounded by

$$\lambda_i(\mathbf{R}_{\tilde{z}\tilde{z}}^{-1}) \geq \frac{1}{1 + \frac{\rho}{K(1+\mu\rho)}\sum_{k=1}^{K}|h_k|^2}. \tag{69}$$

This leads to an upper bound on $\mathbb{P}[\mathcal{J}|\mathcal{A}]$

$$\mathbb{P}[\mathcal{J}|\mathcal{A}] \leq \mathbb{P}[\mathcal{L}] \tag{70}$$

where

$$\mathcal{L} = \{f_0, \tilde{\mathbf{h}} : \frac{\mu}{1+\mu}\left(1+\rho|f_0|^2\right)^2 + \frac{\lambda_{\min}\left(\mathbf{K}(\mathbf{I})\right)}{1+\mu}\rho\tilde{\mathbf{h}}^H\tilde{\mathbf{h}} < \rho^{2r}\}. \tag{71}$$

To see this, note that for $u_k \geq 0$, it follows that $0 \leq |h_k|^2 \leq 1$ since $|h_k|^2 = \rho^{-u_k}$ and, thus,

$$\lambda_i(\mathbf{R}_{\tilde{z}\tilde{z}}^{-1}) \geq \frac{1}{1+\frac{\rho}{K(1+\mu\rho)}\sum_{k=1}^{K}|h_k|^2} \tag{72}$$

$$\geq \frac{1}{1+\frac{\rho}{(1+\mu\rho)}} = \frac{1+\mu\rho}{1+(1+\mu)\rho}. \tag{73}$$

Therefore, the matrix $\mathbf{R}_{\tilde{z}\tilde{z}}^{-1}$ majorizes the matrix $\frac{1+\mu\rho}{1+(1+\mu)\rho}\mathbf{I}_N$ given $u_k \geq 0$ $(k = 1, 2, \ldots, K)$. It follows that

$$\mathrm{Tr}\big(\Lambda^{-1}\big) = \mathrm{Tr}\big(\mathbf{R}_{\tilde{z}\tilde{z}}^{-1}\big) \geq \frac{1+\mu\rho}{1+(1+\mu)\rho}N. \tag{74}$$

Thus, we have

$$1+\rho|f_0|^2\left(1+\frac{\mathrm{Tr}(\mathbf{\Lambda}^-1)}{N}(1+\rho|f_0|^2)\right) \tag{75}$$

$$\geq 1+\rho|f_0|^2\left(1+\frac{1+\mu\rho}{1+(1+\mu)\rho}(1+\rho|f_0|^2)\right) \tag{76}$$

$$\geq \frac{1+\mu\rho}{1+(1+\mu)\rho}\left(1+\rho|f_0|^2\right)^2 \tag{77}$$

$$= \frac{\mu}{1+\mu}\left(1+\rho|f_0|^2\right)^2 \tag{78}$$

where (76) follows from (74), (77) is due to $\frac{1+\mu\rho}{1+(1+\mu)\rho} \leq 1$ and (78) follows due to the high SNR assumption $\frac{1+\mu\rho}{1+(1+\mu)\rho} \approx \frac{\mu}{1+\mu}$. Furthermore, by [20, Observation 7.7.2], we have

$$\tilde{\mathbf{H}}_1^H\mathbf{R}_{\tilde{z}\tilde{z}}^{-1}\tilde{\mathbf{H}}_1 \succeq \frac{1+\mu\rho}{1+(1+\mu)\rho}\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_1 \tag{79}$$

which combined with [20, Corollary 7.7.4] gives

$$\mathrm{Tr}\Big(\tilde{\mathbf{H}}_1^H\mathbf{R}_{\tilde{z}\tilde{z}}^{-1}\tilde{\mathbf{H}}_1\Big) \geq \mathrm{Tr}\left(\frac{1+\mu\rho}{1+(1+\mu)\rho}\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_1\right). \tag{80}$$

Thus, the term $\left\|\mathbf{V}\tilde{\mathbf{H}}_1\right\|_F^2$ can be bounded as follows

$$\left\|\mathbf{V}\tilde{\mathbf{H}}_1\right\|_F^2 = \mathrm{Tr}\Big(\tilde{\mathbf{H}}_1^H\mathbf{V}^H\mathbf{V}\tilde{\mathbf{H}}_1\Big) = \mathrm{Tr}\Big(\tilde{\mathbf{H}}_1^H\mathbf{U}^H\mathbf{\Lambda^{-1}}\mathbf{U}\tilde{\mathbf{H}}_1\Big)$$

$$= \mathrm{Tr}\Big(\tilde{\mathbf{H}}_1^H\mathbf{R}_{\tilde{z}\tilde{z}}^{-1}\tilde{\mathbf{H}}_1\Big) \geq \frac{1+\mu\rho}{1+(1+\mu)\rho}\mathrm{Tr}\Big(\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_1\Big)$$

$$= \frac{\mu}{1+\mu}\mathrm{Tr}\Big(\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_1\Big) = \frac{\mu}{1+\mu}\tilde{\mathbf{h}}^H\mathbf{K}(\mathbf{I}_N)\tilde{\mathbf{h}} \tag{81}$$

$$\geq \frac{\mu\lambda_{\min}(\mathbf{K}(\mathbf{I}_N))}{1+\mu}\tilde{\mathbf{h}}^H\tilde{\mathbf{h}} \tag{82}$$

where (81) follows from the high SNR assumption and (82) follows from the Rayleigh-Ritz theorem [20, Theorem 4.2.2]. Since we assume $\operatorname{rank}\mathbf{K}(\mathbf{I}) = K$, it follows that $\lambda_{\min}(\mathbf{K}(\mathbf{I}_N)) > 0$. Now $\mathbb{P}[\mathcal{L}]$ can be calculated easily

$$\mathbb{P}[\mathcal{L}] = \mathbb{P}\left[\frac{\mu}{1+\mu}\left(1+\rho|f_0|^2\right)^2 + \frac{\lambda_{\min}\left(\mathbf{K}(\mathbf{I})\right)}{1+\mu}\rho\tilde{\mathbf{h}}^H\tilde{\mathbf{h}} < \rho^{2r}\right]$$

$$\doteq \mathbb{P}\left[\left(1+\rho|f_0|^2\right)^2 < \rho^{2r}\right]\mathbb{P}\left[\rho\tilde{\mathbf{h}}^H\tilde{\mathbf{h}} < \rho^{2r}\right] \tag{83}$$

$$\doteq \mathbb{P}\left[|f_0|^2 < \rho^{r-1}\right]\mathbb{P}\left[\tilde{\mathbf{h}}^H\tilde{\mathbf{h}} < \rho^{2r-1}\right] \tag{84}$$

$$\doteq \rho^{-a(1-r)^+}\left(\rho^{-a(1-2r)^+}\right)^K \tag{85}$$

$$= \rho^{-a(1-r)^+ - Ka(1-2r)^+}. \tag{86}$$

Therefore we have bounded $\mathbb{P}[\mathcal{J}]$ from above and below as follows

$$\mathbb{P}[\mathcal{U}] \leq \mathbb{P}[\mathcal{J}] \;\dot{\leq}\; \mathbb{P}[\mathcal{J}|\mathcal{A}] < \mathbb{P}[\mathcal{L}] \tag{87}$$

$$\rho^{-d(r)} \leq \mathbb{P}[\mathcal{J}] \;\dot{\leq}\; \rho^{-d(r)} \tag{88}$$

where $d(r) = a(1-r)^+ + Ka(1-2r)^+$ for $r \in [0,1]$ and hence

$$\mathbb{P}[\mathcal{J}] = P_{\mathcal{J}}(\rho, r) \doteq \rho^{-d(r)}. \tag{89}$$

Since $P_J(\rho, r) \leq P_{\mathcal{O}}(\rho, r)$ as a result of (14), and since the outage probability is a lower bound to the error probability achieved by any code [17], we have

$$P_J(\rho, r) \leq P_{\mathcal{O}}(\rho, r) \leq P_e(\rho, r). \tag{90}$$

Following the approach introduced in [18], we now complete the proof of the theorem by identifying a family of codes which has $P_e(\rho, r) \doteq P_J(\rho, r)$ and hence results in a DM-tradeoff curve which equals the "Jensen" DM-tradeoff curve derived above. We start by writing

$$P_e(\rho, r) = \mathbb{P}[\mathcal{L}]\,\mathbb{P}[\text{error}|\mathcal{L}] + \mathbb{P}\left[\text{error}, \bar{\mathcal{L}}\right]$$
$$\leq \mathbb{P}[\mathcal{L}] + \mathbb{P}\left[\text{error}, \bar{\mathcal{L}}\right].$$

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

$$\mathbb{P}\left[\text{error}, \bar{\mathcal{L}}\right] \leq \rho^{2Nr}\,\mathbb{P}\left[\hat{\mathbf{x}} \to \tilde{\mathbf{x}}, \bar{\mathcal{L}}\right] \tag{91}$$

where we used the fact that the codebook, $\mathcal{C}_r(\rho)$, contains ρ^{2Nr} codewords and $\mathbb{P}\left[\hat{\mathbf{x}} \to \tilde{\mathbf{x}}, \bar{\mathcal{L}}\right]$ denotes the maximum pairwise error probability (over all codeword pairs and all channels in $\bar{\mathcal{L}}$) for ML decoding. With $\Delta\mathbf{x} = \tilde{\mathbf{x}} - \hat{\mathbf{x}}$, we have

$$\mathbb{P}[\hat{\mathbf{x}} \to \tilde{\mathbf{x}}|\mathbf{H}_1] = Q\left(\sqrt{\frac{\rho}{2}}||\mathbf{H}_1\boldsymbol{\Delta}\mathbf{x}||_F\right)$$
$$\leq \exp\left[-\frac{\rho}{4}||\mathbf{H}_1\Delta\mathbf{x}||_F^2\right] \;\dot{\leq}\; \exp\left[-\frac{\rho\mu}{4(1+\mu)}\|\boldsymbol{\Phi}_1(\boldsymbol{\Delta}\mathbf{x})\tilde{\mathbf{f}}\|_F^2\right]$$
$$\leq \exp\left[-\frac{\rho\mu}{4(1+\mu)}\eta_{\min}(\rho)\|\tilde{\mathbf{f}}\|^2\right] \tag{92}$$

where $\eta_{\min}(\rho)$ is the minimum eigenvalue of $\mathbf{\Phi}_1(\mathbf{\Delta x})^H \mathbf{\Phi}_1(\mathbf{\Delta x})$, $\tilde{\mathbf{f}} = [f_0 \; h_1 f_1 \; h_2 f_2 \; \cdots \; h_K f_K]^T$ and (92) follows from applying the Rayleigh-Ritz theorem. Taking the expectation over the vector of all fading coefficients $\tilde{\mathbf{f}} \in \bar{\mathcal{L}}$, we have

$$\begin{aligned}\mathbb{P}\big[\hat{\mathbf{x}} \to \tilde{\mathbf{x}}, \bar{\mathcal{L}}\big] &\dot{\leq} \mathbb{E}_{\tilde{\mathbf{f}} \in \bar{\mathcal{L}}} \left\{ \exp\left[-\frac{\rho\mu}{4(1+\mu)} \eta_{\min}(\rho) \|\tilde{\mathbf{f}}\|^2 \right] \right\} \\ &\dot{\leq} \exp\left[-\frac{\mu \eta_{\min}(\rho)}{4\sqrt{\max(\mu, \lambda_{\min}(\mathbf{K}(\mathbf{I}_N)))}} \rho^r \right] \end{aligned} \tag{93}$$

where (93) follows since the event $\bar{\mathcal{L}}$ requires that $\rho \|\mathbf{f}\|^2 \geq \frac{1+\mu}{\sqrt{\max(\mu, \lambda_{\min}(\mathbf{K}(\mathbf{I}_N)))}} \rho^r$. Finally, inserting (93) into (91), we get

$$\mathbb{P}\big[\text{error}, \bar{\mathcal{L}}\big] \;\dot{\leq}\; \rho^{2Nr} \exp\left[-\frac{\mu \eta_{\min}(\rho)}{4\sqrt{\max(\mu, \lambda_{\min}(\mathbf{K}(\mathbf{I}_N)))}} \rho^r \right]. \tag{94}$$

The proof is complete since $\mathbf{\Phi}_1(\mathbf{\Delta x})$ has full rank for all $\Delta \mathbf{x}$ and for all codebooks in $\mathcal{C}_r$ and, hence, $\eta_{\min}(\rho) > 0$ which implies that (94) decays exponentially in ρ for all $r \in [0, 1]$. Summarizing our results, we obtain

$$\begin{aligned} P_e(\rho, r) &\leq \mathbb{P}[\mathcal{L}] + \mathbb{P}\big[\text{error}, \bar{\mathcal{L}}\big] \\ &\dot{\leq} \mathbb{P}[\mathcal{L}] + \rho^{2Nr} \exp\left[-\frac{\mu \eta_{\min}(\rho)}{4\sqrt{\max(\mu, \lambda_{\min}(\mathbf{K}(\mathbf{I}_N)))}} \rho^r \right] \\ &\dot{\leq} \mathbb{P}[\mathcal{L}] = P_J(\rho, r) \end{aligned}$$

which combined with (90) yields the desired result.

■

Appendix B
Some basic results on diversity

A. Series in Single Dimension

Lemma 1 (Taylor Series in Single Dimension): Assume function $f(x)$ obeys

$$f(x) = cx^\alpha + o(x^\alpha), x \to 0$$

where constant $c > 0$ and $\alpha > -1$. Then

$$\int_0^R f(x)dx = \frac{1}{1+\alpha} c R^{1+\alpha} + o(R^{1+\alpha}), R \to 0.$$

Proof:

$$\begin{aligned} \int_0^R f(x)dx &= \int_0^R cx^\alpha + o(x^\alpha)dx \\ &= \frac{1}{1+\alpha} c x^{1+\alpha}\Big|_0^R + \int_0^R o(x^\alpha)dx \\ &= \frac{1}{1+\alpha} c R^{1+\alpha} + \int_0^R o(x^\alpha)dx. \end{aligned} \tag{95}$$

It remains to prove that

$$\int_0^R o(x^\alpha) = o(R^{1+\alpha}). \tag{96}$$

In other words, we need the show that for any $\epsilon > 0$, there exists a R_0 such that (s.t.) for all $R < R_0$

$$\frac{\int_0^R o(x^\alpha)dx}{R^{1+\alpha}} < \epsilon. \tag{97}$$

By definition of $o(x^\alpha)$, for all $\epsilon > 0$ there exists an $x_0 \geq 0$ s.t. for all $x \leq x_0$, we have

$$\frac{o(x^\alpha)}{x^\alpha} < \epsilon. \tag{98}$$

Now, we fix $\epsilon > 0$. Let $R_0 = x_0$, then from (98) it follows that for all $R < R_0$

$$\frac{\int_0^R o(x^\alpha)dx}{R^{1+\alpha}} < \frac{\int_0^R \epsilon x^\alpha dx}{R^{1+\alpha}} < \epsilon \tag{99}$$

which concludes the argument. ■

Corollary 1: Assume that a RV X has pdf $f_X(x)$ s.t.

$$f_X(x) = cx^\alpha + o(x^\alpha), x \to 0$$

where $c > 0$ and $\alpha > -1$. Then,

$$\lim_{z \to 0} \frac{\log \mathbb{P}\big[X < z^\delta\big]}{\log z} = (1+\alpha)\delta \tag{100}$$

for all $\delta \in [0, 1]$.

Proof: By Lemma 1, with $R = z^\delta$. ■

B. The Diversity of Sum of Positive Random Variables

Lemma 2: Assume that independent RVs X and Y have pdfs $f_X(x)$ and $f_Y(y)$ s.t.

$$f_X(x) = c_x x^{\alpha_x} + o(x^{\alpha_x}), x \to 0$$
$$f_Y(y) = c_y y^{\alpha_y} + o(y^{\alpha_y}), y \to 0$$

where constants $c_x, c_y > 0$ and $\alpha_x, \alpha_y \geq 0$.

$$\lim_{z \to 0} \frac{\log[\mathbb{P}\big[X + Y < z^\delta\big]]}{\log z} = (2 + \alpha_x + \alpha_y)\delta \tag{101}$$

for all $\delta \in [0, 1]$.

Proof: We bound $\mathbb{P}\big[X + Y < z^\delta\big]$ from above as follows

$$\mathbb{P}\big[X + Y < z^\delta\big] \leq \mathbb{P}\big[X < z^\delta\big]\, \mathbb{P}\big[Y < z^\delta\big] \tag{102}$$

Then we have

$$\begin{aligned} \log[\mathbb{P}\big[X + Y < z^\delta\big]] &\leq \log[\mathbb{P}\big[X < z^\delta\big]\, \mathbb{P}\big[Y < z^\delta\big]] \\ &= \log[\mathbb{P}\big[X < z^\delta\big]] + \log[\mathbb{P}\big[Y < z^\delta\big]] \end{aligned}$$

Then diving by $\log z$, taking the limit on both sides and invoking Corollary 1, we get

$$\lim_{z \to 0} \frac{\log[\mathbb{P}\big[X + Y < z^\delta\big]]}{\log z} \geq (2 + \alpha_x + \alpha_y)\delta \tag{103}$$

Similar steps for the lower bound yield

$$\mathbb{P}\left[X < \frac{z^\delta}{2}\right] \mathbb{P}\left[Y < \frac{z^\delta}{2}\right] \leq \mathbb{P}\big[X + Y < z^\delta\big]$$

$$\log[\mathbb{P}\left[X < \frac{z^\delta}{2}\right] \mathbb{P}\left[Y < \frac{z^\delta}{2}\right]] \leq \log[\mathbb{P}\big[X + Y < z^\delta\big]].$$

Then diving by $\log z$, taking the limit on both sides and invoking Corollary 1, we get

$$(2+\alpha_x+\alpha_y)\delta \geq \lim_{z\to 0} \frac{\log[\mathbb{P}\left[X+Y<z^\delta\right]]}{\log z}. \tag{104}$$

Putting the upper (103) and lower bound (104) together completes the proof

$$\lim_{z\to 0} \frac{\log[\mathbb{P}\left[X+Y<z^\delta\right]]}{\log z} = (2+\alpha_x+\alpha_y)\delta$$

for all $\delta \in [0,1]$. ■

C. The Diversity of Product of Positive Random Variables

Theorem 2: Assume that independent RVs X and Y have pdfs $f_X(x)$ and $f_Y(y)$ s.t.

$$f_X(x) = c_x x^\alpha + o(x^\alpha), x \to 0$$
$$f_Y(y) = c_y y^\alpha + o(y^\alpha), y \to 0$$

where constants $c_x, c_y > 0$ and $\alpha \geq 0$. Then

$$\lim_{z\to 0} \frac{\log \mathbb{P}\left[XY \leq z^\delta\right]}{\log z} = (1+\alpha)\delta \tag{105}$$

for all $\delta \in [0,1]$.

Proof: Consider breaking $\mathbb{P}\left[XY \leq z^\delta\right]$ into five regions as shown in Fig. 3. Trivially,

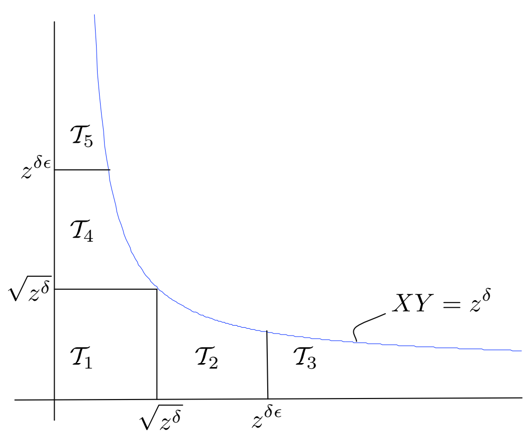

Fig. 3. Depiction of partitioning of $\mathbb{P}\left[XY \leq z^\delta\right]$.

$$\mathbb{P}\left[XY < z^\delta\right] = \sum_{i=1}^{5} \mathbb{P}[\mathcal{T}_i]$$

where events $\mathcal{T}_i$ $(i = 1, 2, \ldots, 5)$ are defined according to

$$\mathcal{T}_1 = \{x, y : X \leq \sqrt{z^\delta}, Y \leq \sqrt{z^\delta}\}$$
$$\mathcal{T}_2 = \{x, y : \sqrt{z^\delta} \leq X \leq z^{\delta\epsilon}, XY \leq z^\delta\}$$
$$\mathcal{T}_3 = \{x, y : z^{\delta\epsilon} \leq X, XY \leq z^\delta\}$$
$$\mathcal{T}_4 = \{x, y : \sqrt{z^\delta} \leq Y \leq z^{\delta\epsilon}, XY \leq z^\delta\}$$
$$\mathcal{T}_5 = \{x, y : z^{\delta\epsilon} \leq Y, XY \leq z^\delta\}.$$

Invoking lemma (1), we can directly arrive at

$$\mathbb{P}[\mathcal{T}_1] = \mathbb{P}\Big[X < \sqrt{z^\delta}\Big]\,\mathbb{P}\Big[Y < \sqrt{z^\delta}\Big] = c_1 z^{(1+\alpha)\delta} + o(z^{(1+\alpha)\delta})$$

where c_1 is a positive constant. To calculate $\mathbb{P}[\mathcal{T}_2]$ (and by symmetry $\mathbb{P}[\mathcal{T}_4]$), we substitute the pdfs for each variable with its series expansion by making use of the fact that for $\epsilon \in [0,1]$ both pdfs can be approximated with their series representation.

$$\begin{aligned}\mathbb{P}[\mathcal{T}_2] &= \int_{\sqrt{z^\delta}}^{z^{\delta\epsilon}} f_X(x) \int_0^{z^\delta/x} f_Y(y)dxdy \\ &= \int_{\sqrt{z^\delta}}^{z^{\delta\epsilon}} f_X(x)\left(\frac{c_y}{1+\alpha}\left(\frac{z^\delta}{x}\right)^{1+\alpha} + o\left(\left(\frac{z^\delta}{x}\right)^{1+\alpha}\right)\right)dx\end{aligned}$$

Then we plug in $f_X(x) = c_x x^\alpha + o(x^\alpha)$ to get

$$\mathbb{P}[\mathcal{T}_2] \leq z^{(1+\alpha)\delta}\int_{\sqrt{z^\delta}}^{z^{\delta\epsilon}} \frac{c_3}{x}dx + \int_{\sqrt{z^\delta}}^{z^{\delta\epsilon}} o\left(\frac{z^{(1+\alpha)\delta}}{x}\right)dx \tag{106}$$

By definition of $o\Big(\frac{z^{(1+\alpha)\delta}}{x}\Big)$, for all $\gamma > 0$ there exists an $\frac{z^{(1+\alpha)\delta}}{x} \geq 0$ s.t. for all $\frac{z^{(1+\alpha)\delta}}{x} \leq z_0$, we have

$$\frac{o\Big(\frac{z^{(1+\alpha)\delta}}{x}\Big)}{\frac{z^{(1+\alpha)\delta}}{x}} < \gamma. \tag{107}$$

Now, we fix $\gamma > 0$. Let $v = z_0$, then from (107) it follows that for all $z^\delta < v^{\frac{1}{0.5+\alpha}}$

$$\frac{\int_{\sqrt{z^\delta}}^{z^{\delta\epsilon}} o\Big(\frac{z^{(1+\alpha)\delta}}{x}\Big)dx}{z^{(1+\alpha)\delta}} < \frac{\int_{\sqrt{z^\delta}}^{z^{\delta\epsilon}} \gamma\frac{z^{(1+\alpha)\delta}}{x}dx}{z^{(1+\alpha)\delta}} < \gamma(\epsilon - .5)(\ln z^\delta). \tag{108}$$

Combining the above, we get

$$\begin{aligned}\mathbb{P}[\mathcal{T}_2] &\leq z^{(1+\alpha)\delta}\int_{\sqrt{z^\delta}}^{z^{\delta\epsilon}} \frac{c_3}{x}dx + o\Big(z^{(1+\alpha)\delta}(\epsilon - .5)(\ln z^\delta)\Big) \\ &= z^{(1+\alpha)\delta}\left(\ln z^{\delta\epsilon} - \ln\sqrt{z^\delta}\right) + o\Big(z^{(1+\alpha)\delta}(\epsilon - .5)(\ln z^\delta)\Big) \\ &= z^{(1+\alpha)\delta}(\epsilon - .5)(\ln z^\delta) + o\Big(z^{(1+\alpha)\delta}(\epsilon - .5)(\ln z^\delta)\Big).\end{aligned}$$

There is ζ_1, ζ_2 s.t. for all $\zeta_1, \zeta_2 > 0$, we have

$$z^{(1+\alpha)\delta+\zeta_1} < z^{(1+\alpha)\delta}(\epsilon - .5)(\ln z^\delta) < z^{(1+\alpha)\delta-\zeta_2}. \tag{109}$$

Therefore we conclude that

$$\mathbb{P}[\mathcal{T}_2] \leq c_4 z^{(1+\alpha)\delta} + o(z^{(1+\alpha)\delta}) \tag{110}$$

where c_4 is a positive constant.

Now we only need to upper bound $\mathbb{P}[\mathcal{T}_3]$

$$\begin{aligned}\mathbb{P}[\mathcal{T}_3] &= \int_{z^{\delta\epsilon}}^{\infty} f_X(x)\int_0^{z^\delta/x} f_Y(y)dydx \qquad (111)\\ &\leq c_5 z^{(1+\alpha)\delta(1-\epsilon)} + o(z^{(1+\alpha)\delta(1-\epsilon)}). \qquad (112)\end{aligned}$$

Since (111) holds for every positive ϵ, which can be made arbitrarily small, we conclude that

$$\mathbb{P}[\mathcal{T}_3] \leq c_4 z^{(1+\alpha)\delta} + o(z^{(1+\alpha)\delta}). \tag{113}$$

Hence we have

$$c_6 z^{(1+\alpha)\delta} + o(z^{(1+\alpha)\delta}) \leq \mathbb{P}\big[XY < z^{\delta}\big] \leq c_5 z^{(1+\alpha)\delta} + o(z^{(1+\alpha)\delta})$$

for some positive constants c_6 and c_5, we conclude that

$$\lim_{z\to 0} \frac{\log \mathbb{P}\big[XY \leq z^{\delta}\big]}{\log z} = (1+\alpha)\delta. \tag{114}$$

■

APPENDIX C
DIVERSITY ORDER OF INCOHERENT SUMS

As a simple case of rank deficient $\mathbf{K}$ consider the scenario when B relays use identical transformation matrices. If one writes the modified signal model for this scenario, one would get

$$\mathbf{y} = \frac{\rho}{\sqrt{1+\rho(1+\|\mathbf{h}\|^2)}} \left(\sum_{i=1}^{K-B} h_i f_i \mathbf{G}_i + \left(\sum_{j=K-B+1}^{K} h_j f_j \right) \mathbf{G}_K \right) \mathbf{x} + \mathbf{z} \tag{115}$$

The analysis for the DM-tradeoff of this scheme is very similar to that of Theorem 1, except that we have to calculate the diversity order of the incoherent sum $\left(\sum_{j=K-B+1}^{K} h_j f_j\right)$ for the reduced-rank problem. In what follows, we show that for a certain sub-class of random variables the incoherent sum $\left(\sum_{j=K-B+1}^{K} h_j f_j\right)$ has the same diversity order as each one of the terms included in the sum.

Definition: A real random variable X is said to be subgaussian if

$$\mathbb{P}[|X| > t] \leq e^{-c_1 t^2} \tag{116}$$

for some positive constant $c_1 > 0$ and all $t \geq 0$. In the following, we need two properties of subgaussian RVs.

Lemma 3 (Vershynin - Moment Generating Function): Let X be a zero-mean RV. Then, the following are equivalent;

1) X is subgaussian.
2) $\mathbb{E}[e^{tX}] \leq e^{ct^2}$ $\forall t \geq 0$.

Proof: For a proof, see [22]. ■

Theorem 3 (Vershynin- Sum of Subgaussian RVs): Let $X_1, X_2, \ldots X_n$ be independent random variables. Also let $a_1, a_2, \ldots a_n \in \mathbb{R}$ be such that $\sum_k a_k^2 = 1$. Then $\sum_k a_k X_k$ is a subgaussian RV.

Proof: Let $Y = \sum_k a_k X_k$.

$$\mathbb{E}\, e^{tY} = \mathbb{E}\, e^{t\sum_k a_k X_k} = \mathbb{E} \prod_k e^{t a_k X_k} = \prod_k \mathbb{E}\, e^{t a_k X_k} \tag{117}$$

By the subgaussianity of X_k, $\mathbb{E}\, e^{t a_k X_k} \leq e^{ct^2}$, for all $t \geq 0$ and for all k. Therefore,

$$\mathbb{E}\, e^{tY} = \mathbb{E}\, e^{t\sum_k a_k X_k} \leq \prod_k e^{ct^2 a_k^2} = e^{ct^2} \tag{118}$$

$$\mathbb{E}\, e^{tY} \leq e^{ct^2}. \tag{119}$$

Hence by Lemma 3, we conclude the argument. ■

Theorem 4 (Diversity order of magnitude of subgaussian sums): Assume that independent zero-mean subgaussian RVs X_1 and X_2 are such that

$$\mathbb{P}[|a_1X_1 + U| < z_1] \leq \mathbb{P}[|a_1X_1| < z_1] \tag{120}$$
$$\mathbb{P}[|a_2X_2 + V| < z_2] \leq \mathbb{P}[|a_2X_2| < z_2] \tag{121}$$

for all U and V and all $z_1, z_2 \geq 0$. Further $|X_1|$ and $|X_2|$ have pdfs $f_{|X_1|}(x)$ and $f_{|X_2|}(x)$ s.t.

$$f_{|X_k|}(x) = c_x x + o(x^2), x \to 0$$

where constants $c_x > 0$ for all k. Then

$$\lim_{z\to 0} \frac{\log \mathbb{P}\left[|a_1X_1 + a_2X_2| < z^\delta\right]}{\log z} = 2\delta. \tag{122}$$

where $a_1^2 + a_2^2 = 1$.

Proof:

$$\mathbb{P}\left[|a_1X_1 + a_2X_2| < z^\delta\right] = \int\int_{|a_1x_1+a_2x_2|<z^\delta} f_{X_1}(x_1)f_{X_2}(x_2)dx_1dx_2 \tag{123}$$

We know from the assumption in the theorem that for any fixed x_2 we have

$$\int_{|a_1x_1+a_2x_2|<z^\delta} f_{X_1}(x_1)dx_1 \leq \int_{|a_1x_1|<z^\delta} f_{X_1}(x_1)dx_1. \tag{124}$$

Hence we have,

$$\mathbb{P}\left[|a_1X_1 + a_2X_2| < z^\delta\right] \leq \int\int_{|a_1x_1|<z^\delta} f_{X_1}(x_1)dx_1f_{X_2}(x_2)dx_2 \tag{125}$$
$$= \mathbb{P}\left[|a_1X_1| < z^\delta\right]. \tag{126}$$

The lower bound follows from subgaussianity of $a_1X_1 + a_2X_2$. We have that

$$\mathbb{P}\left[|a_1X_1 + a_2X_2| > z^\delta\right] \leq e^{-cz^{2\delta}} \tag{127}$$
$$\mathbb{P}\left[|a_1X + a_2X_2| \leq z^\delta\right] \geq 1 - e^{-cz^{2\delta}} \tag{128}$$

Putting the upper and lower bounds together we have

$$1 - e^{-cz^{2\delta}} \leq \mathbb{P}\left[|a_1X_1 + a_2X_2| < z^\delta\right] \leq \mathbb{P}\left[|a_1X_1| < z^\delta\right] \tag{129}$$
$$\log(1 - e^{-cz^{2\delta}}) \leq \log \mathbb{P}\left[|a_1X_1 + a_2X_2| < z^\delta\right] \leq \log \mathbb{P}\left[|a_1X_1| < z^\delta\right] \tag{130}$$

For $z < 1$, we further have

$$\frac{\log(1 - e^{-cz^{2\delta}})}{\log z} \geq \frac{\log \mathbb{P}\left[|a_1X_1 + a_2X_2| < z^\delta\right]}{\log z} \geq \frac{\log \mathbb{P}\left[|a_1X_1| < z^\delta\right]}{\log z} \tag{131}$$

Finally, taking the limit and invoking Corollary 1, we have

$$\lim_{z\to 0} \frac{\log \mathbb{P}\left[|a_1X_1 + a_2X_2| < z^\delta\right]}{\log z} = 2\delta. \tag{132}$$

■

REFERENCES

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

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

[3] J. N. Laneman, G. W. Wornell, and D. N. C. Tse, “An efficient protocol for realizing cooperative diversity in wireless networks,” in *Proc. IEEE ISIT*, Washington, DC, June 2001, p. 294.

[4] 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.

[5] R. U. Nabar and H. Bölcskei, “Space-time signal design for fading relay channels,” in *Proc. IEEE GLOBECOM*, vol. 4, San Francisco, CA, Dec. 2003, pp. 1952–1956.

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

[7] 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.

[8] S. Yang and J.-C. Belfiore, “Towards the optimal amplify-and-forward cooperative diversity scheme,” *IEEE Trans. Inf. Theory*, 2007, to appear.

[9] 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.

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

[11] 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.

[12] S. B. Slimane and A. Osseiran, “Relay communication with delay diversity for future communication systems,” in *Proc. IEEE VTC (Fall)*, Sept. 2006, pp. 1–5.

[13] 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)*, July 2004, pp. 293–297.

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

[15] G. S. Rajan and B. S. Rajan, “A non-orthogonal distributed space-time coded protocol part 1: Signal model and design criteria,” in *Proc. of the 2007 IEEE Information Theory Workshop on Information Theory for Wireless Networks*, 2006, pp. 385–389.

[16] C. Rao and B. Hassibi, “Diversity-multiplexing gain trade-off of a MIMO system with relays,” in *Proc. of the 2007 IEEE Information Theory Workshop on Information Theory for Wireless Networks*, 2007, pp. 47–52.

[17] 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.

[18] P. Coronel and H. Bölcskei, “Diversity-multiplexing tradeoff in selective-fading MIMO channels,” in *Proc. IEEE ISIT*, Nice, France, June 2007.

[19] 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.

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

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

[22] R. Vershynin, “Non-asymptotic theory of random matrices,” *Lecture Notes*, no. 5, January 2007.