

Diversity-Multiplexing Gain Trade-off of a MIMO System with Relays

Chaitanya Rao and Babak Hassibi
 Department of Electrical Engineering
 California Institute of Technology
 Pasadena CA 91125 USA
 Email: {rao, hassibi}@systems.caltech.edu

Abstract—

We find the diversity-multiplexing gain trade-off of a multiple-antenna (MIMO) system with M transmit antennas, N receive antennas, R relay nodes, and with independent Rayleigh fading, in which the relays apply a distributed space-time code. In this two-stage scheme the trade-off is shown to coincide with that of a MIMO system with R transmit and $\min\{M, N\}$ receive antennas.

I. INTRODUCTION

There has been recent interest in the area of *cooperative diversity*, in which relay nodes provide diversity, setting up multiple independent paths between each transmitter and receiver [1], [2], [3]. In this way errors will only occur in the rare event that all of the links are in a deep fade.

In this paper we are interested in the high SNR behavior of a system with a single transmitter with M antennas, a single receiver with N antennas, and R single-antenna relay nodes to assist communication (see Fig. 1). We assume *half-duplex* communication, meaning that nodes cannot simultaneously transmit and receive information. The relays forward a unitary transform of the signal they receive over T channel uses. The diversity-multiplexing gain trade-off, first introduced by Zheng and Tse in [4] for MIMO systems, is seen as a way of comparing the rate versus reliability behavior in the high SNR regime. Since then it has been applied to systems with relays or more than one transmitter-receiver pair [5].

Following [4], we consider a family of codes of fixed block length and increasing signal-to-noise ratio (SNR) and say that the user supports a *multiplexing gain* of r if its data rate $\text{rate}(\text{SNR})$ satisfies $\lim_{\text{SNR} \rightarrow \infty} \frac{\text{rate}(\text{SNR})}{\log \text{SNR}} = r$. The family has a *diversity* d if the average probability of error P_e behaves as $\lim_{\text{SNR} \rightarrow \infty} \frac{\log P_e}{\log \text{SNR}} = -d$ (for which we shall write $P_e \doteq \text{SNR}^{-d}$).

The protocol for this system, whose trade-off behavior will be analyzed in this paper, was first described by Jing and Hassibi in [6]. This can be considered a generalization of [7] in which there were single-antenna nodes. In these works the pairwise error probability of the system was evaluated and conditions placed on the coding to achieve optimal performance. An explicit code construction is given in [8]. The single-antenna case has also been studied by Elia et al. [9] and Susinder Rajan and Sundar Rajan [10]. The coding used is known as *distributed space-time coding* as it is spread over the relays in the network.

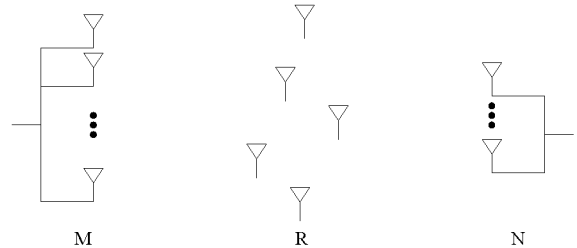


Fig. 1. Relay network with multiple antennas at transmitter and receiver.

Yang and Belfiore [11] consider the diversity-multiplexing trade-off of a MIMO system where the channel matrix H is the product of two independent complex Gaussian matrices. Such a model takes into account instances of lower rank channel matrices. The model we will work with reduces to the Rayleigh product channel in the special case that the relays do no processing of their received signal.

II. MODEL

Let f_{ij} be the fading coefficient from the i -th transmit antenna to the j -th relay, and g_{jk} be the fading coefficient from the j -th relay to the k -th transmit antenna. All fading coefficients are drawn from a $\mathcal{CN}(0, 1)$ distribution. Next define

$$f_i := \begin{bmatrix} f_{1i} \\ f_{2i} \\ \vdots \\ f_{Mi} \end{bmatrix}, \quad g_i := [g_{i1} \quad g_{i2} \quad \dots \quad g_{iN}] \quad (1)$$

as column and row vectors corresponding to connections to the i -th relay.

Let s be the $T \times M$ matrix sent by the transmitter over T transmission steps and via its M antennas. That is, $s = [s_1 \quad s_2 \quad \dots \quad s_M]$, where s_i is the T -dimensional vector sent by the i -th transmit antenna.

In the first T time steps, the transmitter sends s to the relay nodes. The receiver is assumed to be inactive at this stage. The i -th relay node receives a faded version of s plus noise:

$$r_i = \sqrt{P} s f_i + v_i,$$

where v_i is a T -dimensional noise vector. Here we assume $\mathbf{E}\|v_i\|^2 = TM$, so P is a normalization constant proportional to the SNR at the receiver.

Next each relay left-multiplies its T -dimensional signal by a unitary matrix A_i (thus forming linear combinations of the rows, coding over time) and transmits the resulting vector to the receiver. The receiver obtains a $T \times N$ matrix: the sum of the R faded copies of s processed by the relay nodes that are received by N antennas over T further time steps:

$$\begin{aligned} X &= \sum_{i=1}^R A_i r_i g_i + w \\ &= \sqrt{P} \sum_{i=1}^R A_i s f_i g_i + \sum_{i=1}^R A_i v_i g_i + w \\ &= \sqrt{P} S H + W, \end{aligned} \quad (2)$$

where

$$S = [A_1 s \quad A_2 s \quad \dots \quad A_R s], \quad (3)$$

$$H = \begin{bmatrix} f_1 g_1 \\ f_2 g_2 \\ \vdots \\ f_R g_R \end{bmatrix}, \quad W = \sum_{i=1}^R A_i v_i g_i + w. \quad (4)$$

Here S is a $T \times MR$ transmission matrix, H is an $MR \times N$ matrix describing the channel connections as R outer products $f_i g_i$ stacked together, and W is a $T \times N$ noise matrix. The noise sources v_i from each relay and w at the receiver are also assumed to be independent complex Gaussian random variables with zero mean and unit variance for each time step. Equation (2) has the same form as a space-time code for a MIMO system. We assume only the receiver knows H , while the transmitter and relays do not have any channel knowledge. The difference between this system and a MIMO system is that there is no cooperation amongst the relays, and none of them have direct access to s . The code to describe s and the unitary matrices A_i form what is known as a *distributed space-time code* [6].

We will assume T is sufficiently large for codewords to be allowed to be spaced appropriately distant from each other. This will allow an error event to be outage dominated. The model differs from that of [7] in that we do not concern ourselves with power allocation at the relays. Any constants of proportionality become absorbed in the forthcoming high-SNR analysis.

III. MAIN RESULT AND PROOF

In this section we show the following trade-off behavior for the system just described.

Theorem 1: Consider a multiple-antenna system with M transmit antennas, N receive antennas and R relay nodes. The diversity-multiplexing gain trade-off of the two-stage scheme defined in the previous section (1)–(4) is given by the piecewise linear curve joining the points $(k, d(k))$, where $d(k) = (\min\{M, N\} - k)(R - k)$, $k = 1, \dots, \min\{M, N, R\}$.

That is, the trade-off curve coincides with that of a MIMO system with $\min\{M, N\}$ transmit and R receive antennas.

To carry out the trade-off analysis we proceed as follows.

- 1) Find the probability of outage taking into account that there is additional relay noise. We claim that at high SNR the outage behavior is the same as though there were no relay noise.
- 2) Apply the results of [11] to find the exponent of the probability density function of the joint eigenvalue distribution of the channel matrix.
- 3) Optimize this exponent subject to constraints that prevent exponential decay of the pdf — this gives the diversity by application of Laplace's method [4], that asymptotically an integral is close to its integrand having dominating exponent:

$$\int P^{\epsilon(x)} dx \doteq P^{\sup_x \epsilon}, \quad P \rightarrow \infty.$$

We assume $M \geq N$ initially.

- 4) We argue that the case $M < N$, the optimal trade-off can be obtained by disregarding $N - M$ receive antennas.

The maximum diversity of this scheme ($r = 0$) is already known [6] to be $R \min\{M, N\}$. The maximum multiplexing gain will be determined by the maximum dimension of the signal space that can be exploited: $\min\{M, N, R\}$. In the analysis we now will see what happens for intermediate values of r .

We remark that since this is a two stage protocol and that the direct link between transmitter and receiver is not used, all multiplexing gains obtained should strictly be divided by 2, since two channel uses are used to transmit information. Adopting the direct path for relay networks with slightly different models are considered in [5], [12] and [10] amongst others. Adopting lines similar to Yang and Belfiore [12], we conjecture the following, adopting a non-orthogonal amplify and forward protocol.

Conjecture 1: If $d(r)$ is the trade-off curve we obtain for our system in the following analysis, and $d_1(r)$ that of the direct path between transmitter and receiver (a MIMO link), then the overall diversity is lower-bounded by

$$d_{NAF}(r) \geq d_1(r) + d(2r).$$

The first term here corresponds to the source-destination link used all the time; the second to the source-relays-destination system used at half the rate.

A. Outage Calculation

Since the equivalent noise W in (4) consists of Gaussian noise amplified by g_i we need to verify that it does not adversely impact the diversity-multiplexing gain trade-off. First we determine the noise covariance. Considering the k -th row of W for $k = 1, \dots, T$ we have from (4) $W_k = \sum_{i=1}^R (A_i v_i)_k g_i + w_k$ where $(A_i v_i)_k$ and w_k denote the k -th row of $A_i v_i$ and w respectively. Then it can be shown that

$$\mathbf{E}_{v,w} W_k^* W_l = \delta_{kl} (I_N + G^* G) \quad \text{where } G = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_R \end{bmatrix}.$$

Note that this is the same for all time steps $k = 1, \dots, T$. Hence the instantaneous mutual information between transmitter and receiver is given by [13]

$$I(X; S) = \log \det(I_N + P(I_N + G^*G)^{-1}H^*R_S H) \quad (5)$$

where R_S is the $MR \times MR$ input covariance matrix of X for each time step.

We remark that S given by (3) has a particular structure, namely its columns span the M -dimensional space generated by the columns of s . It is known that with the channel unknown to the transmitter the mutual information (5) is maximized when $R_S = I_{MR}$. Hence we make that substitution here. Although S has rank M , R_S takes an average over all realizations of S and so is allowed to have full rank MR .

Since G is a Gaussian matrix, the probability that an eigenvalue of $I + G^*G$ is large is exponentially small. That is, $\lambda_{\max}(G^*G) \leq \text{tr}G^*G$ which is a χ^2 random variable with $2N^2$ degrees of freedom (equivalently a sum of N^2 exponential random variables). Hence it has an exponentially decaying tail and its probability of being large is exponentially small:

$$(1 + \lambda_{\max}(GG^*))^{-1} \doteq P^0.$$

We therefore have justified the following sequence of steps:

$$\begin{aligned} & P_{\text{outage}}(r \log P) \\ &= \min_{R_S: \text{tr}R_S \leq MR} \Pr(I(X; S) < r \log P) \\ &= \min_{R_S: \text{tr}R_S \leq MR} \Pr(\log \det(I_N + PR_W^{-1}H^*R_S H) < r \log P) \\ &\quad (\text{where } R_W = I + G^*G) \\ &\leq \Pr(\log \det(I_N + PR_W^{-1}H^*H) < r \log P) \\ &\doteq \Pr(\log \det(I_N + PH^*H) < r \log P) \end{aligned}$$

To proceed with the calculation of this probability we first observe that

$$\begin{aligned} H^*H &= [(f_1 g_1)^* \dots (f_R g_R)^*] \begin{bmatrix} f_1 g_1 \\ \vdots \\ f_R g_R \end{bmatrix} \\ &= \sum_{i=1}^R g_i^* f_i^* f_i g_i \\ &= \sum_{i=1}^R g_i^* \|f_i\|^2 g_i \\ &= G^* D G \quad \text{where } D = \text{diag}(\|f_1\|^2, \dots, \|f_R\|^2). \end{aligned}$$

This shows that the rank of H is the smaller of the ranks of G and D : $\min\{R, N\}$ which is independent of M . For the remainder of the analysis we will assume $M \geq N$. At the end of the section we will comment on the $M < N$ case.

We proceed along lines similar to recent work by Yang and Belfiore [11] which was applied to the Rayleigh product channel. The non-zero eigenvalues of $H^*H = G^*DG$ are the same as those of QQ^* where $Q_{R \times N} = D^{1/2}G$. We know that given D , QQ^* has a *Wishart* distribution $QQ^* \sim \mathcal{W}_R(N, D)$. That is, the columns of Q are zero-mean independent complex Gaussian vectors having covariance matrix D .

The distribution of the ordered eigenvalues $\lambda_1 > \dots > \lambda_{\min\{N, R\}}$ of QQ^* given D is given by the following lemma, applying a result in [14].

Lemma 1: Let $\mu_1^2 > \dots > \mu_R^2$ be the ordered values of $\|f_1\|^2, \dots, \|f_R\|^2$, the diagonal elements of D , assumed to be distinct. The distribution of the ordered eigenvalues $\lambda_1 > \dots > \lambda_{\min\{N, R\}}$ of $QQ^* = D^{1/2}GG^*D^{1/2}$ for fixed D and for G a random $R \times N$ matrix with independent $\mathcal{CN}(0, 1)$ entries, is given by

$$p(\lambda|D) = \begin{cases} K_{R,N} \det[e^{-\lambda_j/\mu_i^2}] \prod_{i=1}^R \mu_i^{2(R-N-1)} \lambda_i^{N-R} \cdot \prod_{i<j}^R \frac{\lambda_i - \lambda_j}{\mu_i^2 - \mu_j^2}, & N \geq R \\ G_{R,N} \det[\Xi] \prod_{i<j}^R \frac{1}{\mu_i^2 - \mu_j^2} \prod_{i<j}^N (\lambda_i - \lambda_j), & N < R \end{cases} \quad (6)$$

where $K_{R,N}$ and $G_{R,N}$ are normalization constants, $[e^{-\lambda_j/\mu_i^2}]$ is a $(\min\{R, N\} \times \min\{R, N\})$ matrix with (i, j) entry $e^{-\lambda_j/\mu_i^2}$, and

$$\Xi = \begin{bmatrix} 1 & \mu_1^L & \dots & \mu_1^L & \mu_1^L e^{(-\lambda_1/\mu_1^2)} & \dots & \mu_1^L e^{(-\lambda_N/\mu_1^2)} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & \mu_R^L & \dots & \mu_R^L & \mu_R^L e^{(-\lambda_1/\mu_R^2)} & \dots & \mu_R^L e^{(-\lambda_N/\mu_R^2)} \end{bmatrix} \quad (7)$$

where $L = 2(R - N - 1)$.

With each entry being the sum of M independent exponential random variables, the joint distribution of the diagonal entries of D is

$$p(\mu^2) := p(\mu_1^2, \dots, \mu_R^2) = \prod_{i=1}^R \frac{e^{-\mu_i^2} \mu_i^{2(M-1)}}{(M-1)!}. \quad (8)$$

Define

$$\alpha_i = \frac{-\log \lambda_i}{\log P}, \quad \beta_i = \frac{-\log \mu_i^2}{\log P} \Rightarrow \lambda_i = P^{-\alpha_i}, \mu_i^2 = P^{-\beta_i}. \quad (9)$$

Then defining the vectors

$\alpha = (\alpha_1, \dots, \alpha_{\min\{N, R\}})$, $\beta = (\beta_1, \dots, \beta_R)$, we may write the joint eigenvalue distribution of α and β from (6)–(9) as:

$$p(\alpha, \beta) = \begin{cases} C_{R,M,N} (\log P)^{N+R} \cdot \prod_{i=1}^R P^{-(N-R+1)\alpha_i} P^{-(R+M-N-1)\beta_i} \cdot \prod_{i<j}^R \frac{P^{-\alpha_i} - P^{-\alpha_j}}{P^{-\beta_i} - P^{-\beta_j}} \cdot \exp\left(-\sum_{i=1}^R P^{-\beta_i}\right) \cdot \det[\exp(-P^{-(\alpha_j - \beta_i)})]_{i,j=1}^R, & N \geq R \\ D_{R,M,N} (\log P)^{N+R} \prod_{i=1}^N P^{-\alpha_i} \prod_{i=1}^R P^{-M\beta_i} \cdot \prod_{i<j}^R \frac{1}{P^{-\beta_i} - P^{-\beta_j}} \prod_{i<j}^N (P^{-\alpha_i} - P^{-\alpha_j}) \cdot \exp\left(-\sum_{i=1}^R P^{-\beta_i}\right) \cdot \det A & N < R \end{cases} \quad (10)$$

where A is the transformation of the matrix Ξ in (7) under this change of variables (9), while $C_{R,M,N}$ and $D_{R,M,N}$ are normalization constants.

Note that from the exponential factor, $p(\alpha, \beta)$ will decay exponentially unless $\beta_i \geq 0$ for $i = 1, \dots, R$.

The next task is to find the exponent of P in (10). To deal with the determinant factors we make use of the results proved by induction on the matrices' dimensions, in Lemmas 2 and 3 of [11]:

Lemma 2: We have the following expressions for the exponential orders of the determinants in (10).

$$\begin{aligned} & \det \left[\exp \left(-P^{-(\alpha_j - \beta_i)} \right) \right]_{i,j=1}^R \\ & \doteq \exp \left(- \sum_{i=1}^R P^{-(\alpha_i - \beta_i)} \right) P^{-\sum_{i < j} (\alpha_i - \beta_j)^+} \quad (11) \end{aligned}$$

This determinant decays exponentially in P unless $\alpha_i \geq \beta_i$.

$$\begin{aligned} \det A & \doteq \prod_{i=1}^{N+1} P^{-(R-N-1)\beta_i} \prod_{i=N+2}^R P^{-(R-i)\beta_i} \\ & \cdot \prod_{i=1}^N \prod_{j=N+1}^R P^{-(\alpha_i - \beta_j)^+} \\ & \cdot \prod_{i < j}^N P^{-(\alpha_i - \beta_j)^+} \exp \left(- \sum_{i=1}^N P^{-(\alpha_i - \beta_i)} \right) \quad (12) \end{aligned}$$

For $i = 1, \dots, N$ this decays exponentially in P unless $\alpha_i \geq \beta_i$ for $\beta_i \geq 0$.

Since $\lambda_i > \lambda_j$ and $\mu_i > \mu_j$ for $i < j$, we have $\alpha_i \leq \alpha_j$ and $\beta_i \leq \beta_j$ for $i < j$. Hence

$$\frac{P^{-\alpha_i} - P^{-\alpha_j}}{P^{-\beta_i} - P^{-\beta_j}} \doteq \frac{P^{-\alpha_i}}{P^{-\beta_i}} = P^{-(\alpha_i - \beta_i)}. \quad (13)$$

Combining (10), (11), (12) and (13), and using the fact that constants and powers of $(\log P)$ do not contribute to the exponential order allows us to find the exponential order $\epsilon(\alpha, \beta)$ of $p(\alpha, \beta)$. We make use of the identity

$$\sum_{i < j}^R a_i := \sum_{j=1}^R \sum_{i < j} a_i = \sum_{i=1}^R (R-i)a_i.$$

For $N \geq R$,

$$\begin{aligned} \epsilon(\alpha, \beta) & = \sum_{i=1}^R (N+1-i)\alpha_i + \sum_{i=1}^R (M-N+i-1)\beta_i \\ & + \sum_{i < j} (\alpha_i - \beta_j)^+, \quad (14) \end{aligned}$$

while for $N < R$,

$$\begin{aligned} & \epsilon(\alpha, \beta) \\ & = \sum_{i=1}^N (N-i+1)\alpha_i + \sum_{i=1}^N (M-N+i-1)\beta_i \\ & + M \sum_{i=N+1}^R \beta_i + \sum_{i=1}^N \sum_{j=N+1}^R (\alpha_i - \beta_j)^+ + \sum_{i < j}^N (\alpha_i - \beta_j)^+. \quad (15) \end{aligned}$$

The diversity of our wireless scheme is then $\inf \epsilon(\alpha, \beta)$ where $\sum_{i=1}^N (1-\alpha_i)^+ < r$, $\alpha_1 \leq \alpha_2 \leq \dots \leq \alpha_{\min\{N,R\}}$, $\beta_1 \leq \beta_2 \leq \dots \leq \beta_R$ and $\alpha_i \geq \beta_i \geq 0$ for $i = 1, \dots, N$.

Next we find this infimum by optimizing (14) and (15) over α and β .

B. Optimization over α and β

To minimize (14) and (15) subject to the constraints just stated, we first fix $\alpha_1 \leq \dots \leq \alpha_{\min\{N,R\}}$ and find the optimal β . We will assume $M \geq N$.

1) *Case 1: $N \geq R$:* As done in [11] we begin with an initial configuration

$$0 \leq \beta_1 = \alpha_1 \leq \beta_2 = \alpha_2 \leq \dots \leq \beta_R = \alpha_R$$

so that $\sum_{i < j} (\alpha_i - \beta_j)^+ = 0$.

The sum (14) is of the form $\sum_{i=1}^R (a_i \alpha_i + b_i \beta_i)$. With the initial configuration just described we would have $a_i = N+1-i$ and $b_i = M-N+i-1 \geq 0$. For some fixed j , if β_j is decreased below α_i for some $i < j$, the term $(\alpha_i - \beta_j)^+$ is no longer zero. Here a_i increases by 1, b_j decreases by 1 and the overall sum (14) decreases provided b_j remains positive. How many α_i 's should β_j cross? Denote this number by c_j . Since b_j decreases by 1 each time from an initial value of $M-N+j-1$, we have $c_j = \min\{j-1, M-N+j-1\} = j-1$ since $M \geq N$. In other words, for $M \geq N$ it is optimal for β_j to be decreased all the way to zero. Substituting this into (14) leads to

$$\epsilon(\alpha, 0) = \sum_{i=1}^R (N+R+1-2i)\alpha_i \quad (16)$$

2) *Case 2: $N < R$:* Again we assume an initial configuration

$0 \leq \beta_1 = \alpha_1 \leq \beta_2 = \alpha_2 \leq \dots \leq \beta_N = \alpha_N \leq \beta_{N+1} \leq \dots \leq \beta_R$, so that the last two terms of (15) are 0.

This time (15) is of the form $\epsilon(\alpha, \beta) = \sum_{i=1}^N a_i \alpha_i + \sum_{j=1}^N b_j \beta_j + \sum_{N+1}^R b_j \beta_j$. With the initial configuration, $a_i = N-i+1$, $b_j = M-N+j-1$. Proceeding similarly to Case 1, but treating the cases $j \leq N$ and $j > N$ separately, we see that if β_j is decreased below α_i for some $i < j$, a_i increases by 1 while b_j decreases by 1. The overall sum (15) decreases provided b_j remains positive. If c_j is the number of α_i 's that b_j should cross, we have as before $c_j = \min\{j-1, M-N+j-1\} = j-1$. Hence β_j may cross all α_i 's all the way to zero to minimize the sum. Substituting this into (15) leads to

$$\epsilon(\alpha, 0) = \sum_{i=1}^N (N+R+1-2i)\alpha_i \quad (17)$$

Combining (16) and (17),

$$\epsilon(\alpha) = \sum_{i=1}^{\min\{N,R\}} (N+R+1-2i)\alpha_i. \quad (18)$$

Now we optimize this over α , so we assume $0 < \alpha_i < 1$ for $i = 1, \dots, \min\{N,R\}$. Initially suppose $r = 0$ which forces $\alpha_i = 1$ for all i . This corresponds to a maximum diversity of

$$\sum_{i=1}^{\min\{N,R\}} (N+R+1-2i) = NR.$$

As r increases, we are free to lower the values of α_i . Since the coefficients $(N + R - 1 - 2i)$ of α_i are positive and strictly decreasing in i , to minimize (18) it is optimal to push α_i to zero one at a time beginning with α_1 .

That is,

- For $r = 0$ set all $\alpha_i := 1$ achieving diversity NR .
- For $0 < r < 1$ push α_1 to zero, $\alpha_i = 1$ for $i > 1$. The diversity decreases by $N + R - 1$ to $NR - (N + R - 1) = (N - 1)(R - 1)$ for $r = 1$.
- For $1 < r < 2$ push α_2 to zero, while $\alpha_1 = 0$, $\alpha_i = 1$ for $i > 2$. The diversity decreases by $N + R - 3$ to $(N - 1)(R - 1) - (N + R - 3) = (N - 2)(R - 2)$ for $r = 2$.
- ...
- For $\min\{N, R\} - 1 < r < \min\{N, R\}$ push $\alpha_{\min\{N, R\}}$ to zero, while $\alpha_i = 0$ for $i < \min\{N, R\}$. The diversity decreases by $|N - R| + 1$ to zero for $r = \min\{N, R\}$.

We hence obtain a piecewise linear curve joining points $(k, (N - k)(R - k))$ for $k = 1, 2, \dots, \min\{N, R\}$. This is precisely the same as the optimal trade-off curve for a MIMO system with R transmit and N receive antennas. Next we will verify that this trade-off curve is achievable with a particular code construction.

C. Achieving the Trade-off

By adopting an analysis similar to that presented by Elia et al. in [9] one can construct an explicit coding scheme which satisfies a *non-vanishing determinant criterion*, and in doing so achieves the optimal trade-off. This property states that if S_1 and S_2 are two distinct codewords then

$$\det(S_1 - S_2)(S_1 - S_2)^* \doteq P^0. \quad (19)$$

That is, the determinant does not decay to zero as $P \rightarrow \infty$. Codes based on cyclic division algebras [8], [9], [10] can be made to have this property, where S_1 and S_2 each have the structure given by (3).

D. The Case $M < N$

The optimization of Section III-B would not have worked for the case $M < N$. In this instance we reason as follows.

Allowing for full cooperation amongst the relay nodes, the diversity of the system can be upper-bounded by the diversity obtained by the each of the first stage ($d_{\text{MIMO}(M,R)}(r)$) and second stage ($d_{\text{MIMO}(R,N)}(r)$) of transmission. That is,

$$\begin{aligned} d(r) &\leq \min\{d_{\text{MIMO}(M,R)}(r), d_{\text{MIMO}(R,N)}(r)\} \\ &= \min\{d_{\text{MIMO}(R,M)}(r), d_{\text{MIMO}(R,N)}(r)\} \\ &= d_{\text{MIMO}(R, \min\{M, N\})}(r). \end{aligned} \quad (20)$$

We have already seen this bound to be achievable in the $M \geq N$ case. When $M < N$, simply ignore the signal received at the final $N - M$ receive antennas. This allows us to apply the result we have proved and gives a lower bound on diversity:

$$\begin{aligned} d(r) &\geq d_{\text{MIMO}(M,R)}(r) \\ &= d_{\text{MIMO}(R, \min\{M, N\})}(r). \end{aligned} \quad (21)$$

Comparing (20) and (21), we conclude

$$d(r) = d_{\text{MIMO}(R, \min\{M, N\})}(r).$$

IV. DISCUSSION

We make the following summarizing remarks.

- Large diversity benefits can be reaped by increasing the number of relay antennas, provided they maintain independently fading channels. Implementing cheap relay nodes may be easier to do than adding antennas at the transmitter or receiver.
- The trade-off behavior is a function of $\min\{M, N\}$ so there is no point in having more transmit than receive antennas or vice versa.
- Compared with the results of Yang and Belfiore [11] for the product Rayleigh channel, there is a significant benefit of having different unitary matrices at the relays. The Rayleigh product channel's result applies to our model in the case A_i all equal to the identity matrix, and in that case, the diversity is at most MN if $R > \max\{M, N\}$ (i.e. no growth in R).

ACKNOWLEDGMENT

This work is supported in part by the National Science Foundation under grant nos. CCR-0133818 and CCR-0326554, by the David Lucille Packard Foundation, and by Caltech's Lee Center for Advanced Networking.

REFERENCES

- [1] A. Sendonaris, E. Erkip, and B. Aazhang. User cooperation diversity, Part I: System description. *IEEE Trans. Commun.*, 51:1927–1938, Nov. 2003.
- [2] N. Laneman, D. Tse, and G. Wornell. Cooperative diversity in wireless networks: Efficient protocols and outage behavior. *IEEE Trans. Inform. Theory*, 50:3062–3080, Dec. 2004.
- [3] R. Nabar, H. Bolcskei, and F. Kneubuhler. Fading relay channels: performance limits and space-time signal design. *IEEE Journ. Select. Areas Commun.*, 22:1099–1109, Aug. 2004.
- [4] L. Zheng and D. Tse. Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels. *IEEE Trans. Inform. Theory*, 49:1073–1096, May 2003.
- [5] K. Azarian, H. El Gamal, and P. Schniter. On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels. *IEEE Trans. Inform. Theory*, 51:4152–4172, Dec. 2005.
- [6] Y. Jing and B. Hassibi. Cooperative diversity in wireless relay networks with multiple-antenna nodes. *IEEE International Symposium on Information Theory: ISIT'05*, pages 815–819, Sep. 2005.
- [7] Y. Jing and B. Hassibi. Distributed space-time coding in wireless relay networks. *IEEE Trans. Wireless. Comm.*, 5:3524–3536, Dec. 2006.
- [8] F. Oggier and B. Hassibi. An algebraic coding scheme for wireless relay networks with multiple-antenna nodes. *Submitted to IEEE Trans. Signal. Proc.*, Mar. 2006.
- [9] P. Elia, P. Raj Kumar, S. Pawar, P. Vijay Kumar, and H.-F. Lu. Explicit space-time codes achieving the diversity-multiplexing gain tradeoff. *IEEE Trans. Inform. Theory*, 52:3869–3884, Sep. 2006.
- [10] G. Susinder Rajan and M. Sundar Rajan. A non-orthogonal distributed space-time coded protocol, part i: Signal model and design criteria. *Proceedings of 2006 IEEE Information Theory Workshop (ITW'06)*, Oct. 2006.
- [11] S. Yang and J.-C. Belfiore. Diversity-multiplexing tradeoff of double scattering MIMO channels. *Submitted to: IEEE Trans. Inform. Theory*, Mar. 2006.
- [12] S. Yang and J.-C. Belfiore. Optimal space-time codes for the MIMO amplify-and-forward cooperative channel. *IEEE Trans. Inform. Theory*, 53:647–663, Feb. 2007.
- [13] E. Telatar. Capacity of multi-antenna Gaussian channels. *Europ. Trans. on Telecomm.*, 10:585–595, Nov/Dec. 1999.
- [14] A. Tulino and S. Verdú. *Random Matrix Theory and Wireless Communications*. Now Publishers, 2004.