# Distributed Space-Time Coding in Wireless Relay Networks

YINDI JING AND BABAK HASSIBI*

Department of Electrical Engineering

California Institute of Technology, Pasadena, CA 91125

April 27, 2006

**Abstract**

We apply the idea of space-time coding devised for multiple-antenna systems to the problem of communications over a wireless relay network with Rayleigh fading channels. We use a two-stage protocol, where in one stage the transmitter sends information and in the other, the relays encode their received signals into a "distributed" *linear dispersion (LD)* code, and then transmit the coded signals to the receive node. We show that for high SNR the *pairwise error probability (PEP)* behaves as $\left(\frac{\log P}{P}\right)^{\min\{T,R\}}$, with $T$ the coherence interval, that is, the number of symbol periods during which the channel keeps constant, $R$ the number of relay nodes, and $P$ the total transmit power. Thus, apart from the $\log P$ factor and assuming $T \geq R$, the system has the same diversity as a multiple-antenna system with $R$ transmit antennas, which is the same as assuming that the $R$ relays can fully cooperate and have full knowledge of the transmitted signal. We further show that for a network with a large number of relays and a fixed total transmit power across the entire network, the optimal power allocation is for the transmitter to expend half the power and for the relays to collectively expend the other half. We also show that at low and high SNR, the coding gain is the same as that of a multiple-antenna

*This work was supported in part by the National Science Foundation under grant no. CCR-0133818, by the office of Naval Research under grant no. N00014-02-1-0578, and by Caltech's Lee Center for Advanced Networking.

system with $R$ antennas. However, at intermediate SNR, it can be quite different, which has implications for the design of distributed space-time codes.

# 1 Introduction

It is known that multiple antennas can greatly increase the capacity and reliability of a wireless communication link in a fading environment using space-time codes [1, 2, 3, 4]. Recently, with the increasing interests in ad hoc networks, researchers have been looking for methods to exploit spatial diversity using antennas of different users in the network [5, 6, 7, 8, 9]. In [8], the authors exploit spatial diversity using the repetition and space-time algorithms. The mutual information and outage probability of the network are analyzed. However, in their model, the relays need to decode their received signals. In [9], a network with a single relay under different protocols is analyzed and second order spatial diversity is achieved. In [10], the authors use space-time codes based on the Hurwitz-Radon matrices and conjecture a diversity factor around $R/2$ from their simulations. Also, the simulations in [11] show that the use of Khatri-Rao codes lowers the average bit error rate. In this paper, we consider a relay network with fading and apply a LD space-time code [12] among the relays. The problem we are interested in is: "Can we increase the reliability of a wireless network by using space-time codes among the relays?"

More specifically, the focus of this paper is on the PEP analysis of wireless relay network. We investigate in the diversity gain and coding gain that can be achieved in a wireless relay network by having the relays cooperate distributively. Here, by diversity gain or diversity in brief, we mean the negative of the exponent of the SNR or transmit power in the PEP formula at high SNR regime. This definition is consistent with the diversity definition in multiple-antenna systems [4, 13]. It determines how fast the PEP decreases with the SNR or transmit power. The same as before, coding gain is the improvement in the PEP obtained by the code design.

The wireless relay network model we use is similar to those in [14, 15]. In [14], the authors show that the capacity of the wireless relay network with $R$ nodes behaves like $\log R$. In [15], a power efficiency that behaves like $\sqrt{R}$ is obtained. Both results are based on the assumption that every relay knows its local channels so that they can work coherently. Therefore, for the results of [14] and [15] to hold, the system should be synchronized at the carrier level. In this paper, we assume that the relays do not know the channel information. All we need is a much more reasonable assumption that the system is synchronized at the symbol level.

We use a two-step protocol for transmissions in the wireless relay network, where in the first step the transmitter sends information and in the other, the relays encode their received signals into a "distributed" LD code, and then transmit the coded signals to the receive node. A key feature of our work is that we do not require the relays to decode. Only simple signal processing is done at the relays. This has two main benefits: first, the operations at the relays are considerably simplified, and second, we can avoid imposing bottlenecks on the rate by requiring some relays to decode (See e.g., [16]).

Our work shows that in a wireless relay network with $R$ relays, coherence interval $T$, and a single transmit-and-receive pair, using LD space-time codes among the relays can achieve a diversity of $\min\{T,R\}\left(1-\frac{\log\log P}{\log P}\right)$, where $P$ is the total power used in the whole network. When $T \geq R$, the diversity gain is linear in the number of relays (size of the network) and is a function of the total transmit power. When $P$ is very large ($\log P \gg \log\log P$), the diversity is approximately $R$. The coding gain for very large $P$ is $\det{(S_k-S_l)^*(S_k-S_l)}$, where $S_k$ and $S_l$ are codewords in the distributed space-time code. Therefore, at very high SNR, the same diversity gain and coding gain are obtained as in the multiple-antenna case, which means that the system works as if the relays can fully cooperate and have full knowledge of the transmit signal. We then improve the diversity gain shown above and prove the optimality of the result. We also consider a more general type of LD codes which includes Alamouti's scheme as a special case. Although, the same diversity gains are achieved, the coding gain can be improved. Simulations are also provided, which verify our theoretical analysis.

The paper is organized as follows. In the following section, the network model and the two-step protocol are introduced. The distributed space-time code is explained in Section 3 and the PEP is calculated in Section 4. In Section 5, we derive the optimum power allocation based on the PEP. Sections 6 contains the main results of our work. The diversity gain and the coding gain are derived. To motivate our results, we first give a simple approximate derivation and then give the more involved rigorous derivation. In Section 7, we slightly improve the diversity gain obtained in Section 6 and prove the optimality of the new diversity result. A more general distributed LD space-time code is discussed in Section 8, and in Section 9 the diversity gain and coding gain for a special case are obtained, which coincide with those in Sections 6 and 7. We have simulated the performances of relay networks with random distributed LD space-time codes and have compared them with the performances of the same space-time codes used in multiple-antenna systems. The details of the simulations and the figures are given in Section 10. Section 11 provides the conclusion and future work. The proofs of the technical theorems and lemmas are given in the appendices.

# 2 System Model

We first introduce some notation used in the paper. For a complex matrix $A$, $\bar{A}, A^t$, and $A^*$ denote the conjugate, the transpose, and the conjugate transpose of $A$, respectively. $\det A$, $\operatorname{rank} A$, and $\operatorname{tr} A$ indicate the determinant, rank, and trace of $A$, respectively. $A_{Re}$ and $A_{Im}$ are the real and imaginary parts of $A$. $I_n$ denotes the $n \times n$ identity matrix and $0_{m,n}$ is the $m \times n$ matrix with all zero entries. We often omit the subscripts when there is no confusion. $\log$, $\log_2$, $\log_{10}$ indicate the natural logarithm, the base-2 logarithm, and the base-10 logarithm. $\|\cdot\|$ indicates the Frobenius norm. $g(x) = O(f(x))$ means that $\lim_{x\to\infty} \frac{g(x)}{f(x)}$ is a constant. $h(x) = o(f(x))$ means that $\lim_{x\to\infty} \frac{g(x)}{f(x)} = 0$.

Consider a wireless network with $R+2$ nodes which are placed randomly and independently according to some distribution. There is one transmit node and one receive node. All the other $R$ nodes work as relays. Every node has a single antenna, which can be used for both transmission and reception. Denote the channel from the transmitter to the $i$th relay as $f_i$, and the channel from the $i$th relay to the receiver as $g_i$. Assume that $f_i$ and $g_i$ are independent complex Gaussian random variables with zero-mean and unit-variance. If the fading coefficients $f_i$ and $g_i$ are known to relay $i$, it is proved in [14] and [15] that the capacity behaves like $\log R$ and a power efficiency that behaves like $\sqrt{R}$ can be obtained. However, these results rely on the assumption that the relays know their local connections, which requires the system to be synchronized at the carrier level. In this paper, we make the much more practical assumption that the relays are only coherent at the symbol level. We assume that the relays know only the statistical distribution of the channels. However, we make the assumption that the receiver knows all the fading coefficients $f_i$ and $g_i$. Its knowledge of the channels can be obtained by sending training signals from the relays and the transmitter. Many types of gains can be obtained from the network, for example, gains on the capacity and gains on the error rate. In this paper, we focus on the improvement in the error rate, more precisely, the pairwise error probability (PEP).

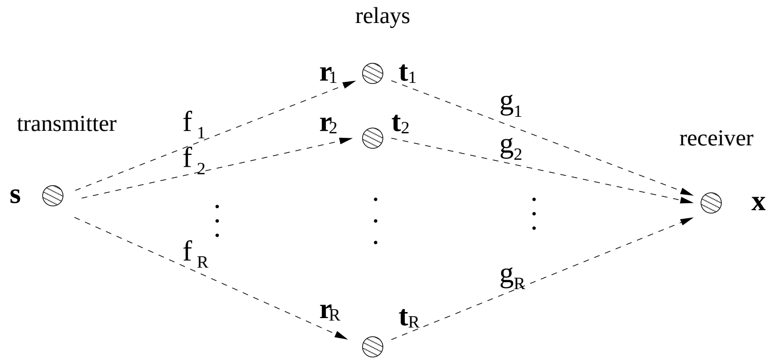


Figure 1: Wireless relay network

Assume that the transmitter wants to send the signal $\mathbf{s} = [s_1, \cdots, s_T]^t$ in the codebook $\{\mathbf{s}_1, \cdots, \mathbf{s}_L\}$ to the receiver, where $L$ is the cardinality of the codebook. $\mathbf{s}$ is normalized as

$$\mathrm{E}\,\mathbf{s}^*\mathbf{s} = 1. \tag{1}$$

The transmission is accomplished by the following two-step strategy, which is also shown in Fig 1. From time $1$ to $T$, the transmitter sends signals $\sqrt{P_1 T} s_1, \cdots, \sqrt{P_1 T} s_T$ to each relay. Based on the normalization of $\mathbf{s}$, the average power used at the transmitter for every transmission is $P_1$. The received signal at the $i$th relay at time $\tau$ is denoted as $r_{i,\tau}$, which is corrupted by both the fading $f_i$ and the noise $v_{i,\tau}$. From time $T+1$ to $2T$, the $i$th relay sends $t_{i,1}, \cdots, t_{i,T}$ to the receiver. We denote the received signal and noise at the receiver at time $\tau + T$ by $x_\tau$ and $w_\tau$ respectively. Assume that the noises are independent complex Gaussian random variables with zero-mean and unit-variance, that is, the distribution of $v_{i,\tau}, w_\tau$ are $\mathcal{CN}(0,1)$.

We use the following notation:

$$\mathbf{v}_i = \begin{bmatrix} v_{i,1} \\ \vdots \\ v_{i,T} \end{bmatrix}, \quad \mathbf{r}_i = \begin{bmatrix} r_{i,1} \\ \vdots \\ r_{i,T} \end{bmatrix}, \quad \mathbf{t}_i = \begin{bmatrix} t_{i,1} \\ \vdots \\ t_{i,T} \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_T \end{bmatrix}, \text{ and } \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_T \end{bmatrix}.$$

If we assume a coherence interval of $T$, that is $f_i$ and $g_i$ keep constant for $T$ transmissions, clearly

$$\mathbf{r}_i = \sqrt{P_1 T} f_i \mathbf{s} + \mathbf{v}_i \quad \text{and} \quad \mathbf{x} = \sum_{i=1}^{R} g_i \mathbf{t}_i + \mathbf{w}. \tag{2}$$

There are two main differences between the wireless relay network given above and a multiple-antenna system with $R$ transmit antennas and one receive antenna [4, 13], although they both have $R$ independent transmission routes from the transmitter to the receiver. The first one is that in a multiple-antenna system, antennas of the transmitter can cooperate fully. In the considered network, they can cooperate only in a distributed fashion since the relays are different users. The other difference is that in the network, the relays observe only noisy versions of the transmit signal.

# 3 Distributed Space-Time Coding

From the above description, it is clear that if the transmission rate is sufficiently low, then all the relays can decode the transmit message. In this case, the relays can act as a multiple-antenna system with $R$ transmit antennas and therefore the communication from the relays to the receiver can achieve a diversity order of $R$. This approach, however, will require a substantial reduction of the rate and we will therefore not consider it. We will instead focus on the achievable diversity without requiring the relays to decode.[1]

In this paper, we use the idea of the LD space-time code [12] for multiple-antenna systems by designing the transmit signal at every relay as a linear function of its received signal:[2]

$$t_{i,\tau} = \sqrt{\frac{P_2}{P_1+1}} \sum_{t=1}^{T} a_{i,\tau t} r_{i,t} = \sqrt{\frac{P_2}{P_1+1}} [a_{i,\tau 1}, a_{i,\tau 2}, \cdots, a_{i,\tau T}] \mathbf{r}_i,$$

or in other words,

$$\mathbf{t}_i = \sqrt{\frac{P_2}{P_1+1}} A_i \mathbf{r}_i, \tag{3}$$

where

$$A_i = \begin{bmatrix} a_{i,11} & \cdots & a_{i,1T} \\ \vdots & \ddots & \vdots \\ a_{i,T1} & \cdots & a_{i,TT} \end{bmatrix}, \quad \text{for } i = 1, 2, \cdots, R.$$

While within the framework of LD codes, the $T \times T$ matrices $A_i$ can be quite arbitrary (apart from a Frobenius norm constraint), to have a protocol that is equitable among different users and

[1] A combination of requiring some relays to decode and others to not, may also be considered. However, in the interest of space, we shall not do so here.

[2] Note that the conjugate of $\mathbf{r}_i$ does not appear in (3). The case with $\overline{\mathbf{r}_i}$ is discussed in Section 8.

among different time instants, we shall henceforth assume that $A_i$ are unitary. As we shall presently see, this also simplifies the analysis considerably.

Now let's discuss the average transmit power at every relay. Because $\mathrm{E}\,\mathrm{tr}\,\mathbf{s}\mathbf{s}^* = 1$, $f_i, v_{i,j}$ are $\mathcal{CN}(0,1)$, and $f_i, s_i, v_{i,j}$ are independent, the average received power at relay $i$ is:

$$\mathrm{E}\,\mathbf{r}_i^*\mathbf{r}_i = \mathrm{E}\,\left(P_1T|f_i|^2\mathbf{s}^*\mathbf{s} + \mathbf{v}_i^*\mathbf{v}_i\right) = (P_1+1)T.$$

Therefore the average transmit power at relay $i$ is

$$\mathrm{E}\,\mathbf{t}_i^*\mathbf{t}_i = \frac{P_2}{P_1+1}\mathrm{E}\,(A_i\mathbf{r}_i)^*(A_i\mathbf{r}_i) = \frac{P_2}{P_1+1}\mathrm{E}\,\mathbf{r}_i^*\mathbf{r}_i = P_2T,$$

which explains our normalization in (3). $P_2$ is the average transmit power for one transmission at every relay.

Let us now focus on the received signal. Clearly from (2), the received signal can be calculated to be

$$\mathbf{x} = \sqrt{\frac{P_1P_2T}{P_1+1}}SH + W, \tag{4}$$

where we have defined

$$S = \begin{bmatrix} A_1\mathbf{s} & \cdots & A_R\mathbf{s} \end{bmatrix}, H = \begin{bmatrix} f_1g_1 \\ \vdots \\ f_Rg_R \end{bmatrix}, \text{ and } W = \sqrt{\frac{P_2}{P_1+1}}\sum_{i=1}^{R} g_iA_i\mathbf{v}_i + \mathbf{w}. \tag{5}$$

The $T \times R$ matrix $S$ in equation (4) works like the space-time code in a multiple-antenna system. We shall call it the *distributed space-time code* to emphasize that it has been generated in a distributed way by the relays, without having access to $\mathbf{s}$. $H$, which is $R \times 1$, is the equivalent channel matrix and $W$, which is $T \times 1$, is the equivalent noise. $W$ is clearly influenced by the choice of the space-time code. Using the unitarity of $A_i$, it is easy to get the normalization of $S$: $\mathrm{E}\,\mathrm{tr}\,S^*S = R$.

# 4 Pairwise Error Probability

Since $A_i$ are unitary and $w_j, v_{i,j}$ are independent Gaussian, $W$ is also Gaussian when $g_i$ are known. It is easy to see that $\mathrm{E}\,W = 0_{T,1}$ and $\mathrm{Var}\,W = \left(1 + \frac{P_2}{P_1+1}\sum_{i=1}^{R}|g_i|^2\right)I_T$. Thus, $W$ is both spatially and temporally white. Assume that $\mathrm{s}_k$ is transmitted. Define $S_k = [A_1\mathbf{s}_k, \cdots, A_R\mathbf{s}_k]$. Therefore, $S_k$ is an element in the distributed space-time code set. When both $f_i$ and $g_i$ are known, $\mathbf{x}|\mathbf{s}_k$ is also Gaussian with mean $\sqrt{\frac{P_1P_2T}{P_1+1}}S_kH$ and variance $\left(1+\frac{P_2}{P_1+1}\sum_{i=1}^R|g_i|^2\right)I_T$. Thus,

$$\mathrm{P}\left(\mathbf{x}|\mathbf{s}_k\right) = \frac{1}{\left[\pi\left(1+\frac{P_2}{P_1+1}\sum_{i=1}^R|g_i|^2\right)\right]^T} e^{-\frac{\left(\mathbf{x}-\sqrt{\frac{P_1P_2T}{P_1+1}}S_kH\right)^*\left(\mathbf{x}-\sqrt{\frac{P_1P_2T}{P_1+1}}S_kH\right)}{1+\frac{P_2}{P_1+1}\sum_{i=1}^R|g_i|^2}}.$$

The maximum-likelihood (ML) decoding of the system can be easily seen to be

$$\arg\max_{\mathbf{s}_k} \mathrm{P}\left(\mathbf{x}|\mathbf{s}_k\right) = \arg\min_{\mathbf{s}_k} \left\|\mathbf{x} - \sqrt{\frac{P_1P_2T}{P_1+1}}S_kH\right\|^2. \tag{6}$$

Since $S_k$ is linear in $\mathrm{s}_k$, by splitting the real and imaginary parts, the decoding is equivalent to the decoding of a real linear system. Therefore, sphere decoding can be used [17, 18].

**Theorem 1** (Chernoff bound on the PEP). *With the ML decoding in (6), the PEP, averaged over the channel coefficients, of mistaking* $\mathrm{s}_k$ *by* $\mathrm{s}_l$ *has the following Chernoff bound:*

$$\mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) \le \mathop{\mathrm{E}}_{f_i,g_i} e^{-\frac{P_1P_2T}{4\left(1+P_1+P_2\sum_{i=1}^R|g_i|^2\right)}H^*(S_k-S_l)^*(S_k-S_l)H}.$$

*By integrating over $f_i$, we can get the following inequality:*

$$\mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) \le \mathop{\mathrm{E}}_{g_i} \det{}^{-1}\left[I_R + \frac{P_1P_2T}{4\left(1+P_1+P_2\sum_{i=1}^R|g_i|^2\right)}(S_k-S_l)^*(S_k-S_l)\mathrm{diag}\left\{|g_1|^2,\cdots,|g_R|^2\right\}\right]. \tag{7}$$

**Proof:** See Appendix A. □

Let's compare (7) with the Chernoff bound on the PEP of a multiple-antenna system with $R$ transmit antennas and one receive antenna (the receiver knows the channel) [4, 13]:

$$\mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) \le \det{}^{-1}\left[I_R + \frac{PT}{4R}(S_k-S_l)^*(S_k-S_l)\right].$$

The difference is that now we need to do the expectations over $g_i$. Similar to the multiple-antenna case, the *full diversity* condition can be obtained from (7). It is easy to see that if $S_k - S_l$ drops rank, the upper bound in (7) increases. Therefore, the Chernoff bound is minimized when $S_k - S_l$ is full-rank, or equivalently, $\det(S_k-S_l)^*(S_k-S_l) \neq 0$ for any $1 \le k \neq l \le L$.

## 5 Optimum Power Allocation for Large $R$

In this section, we discuss the optimum power allocation between the transmitter and relays that minimizes the PEP. Because of the expectations over $g_i$, it is very difficult to obtain the exact solution. We shall therefore recourse to a heuristic argument. Note that $g = \sum_{i=1}^R |g_i|^2$ has the gamma distribution [19]:

$$p(g) = \frac{g^{R-1}e^{-g}}{(R-1)!},$$

whose mean and variance are both $R$. It is therefore reasonable to approximate $g$ by its mean, i.e., $\sum_{i=1}^R |g_i|^2 \approx R$, especially for large $R$. (By the law of large numbers, almost surely $g/R \to 1$ when $R \to \infty$.). Therefore, (7) becomes

$$\mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) \lesssim \mathop{\mathrm{E}}_{g_i} \det{}^{-1}\left[I_R + \frac{P_1P_2T}{4\left(1+P_1+P_2R\right)}(S_k - S_l)^*(S_k - S_l)\mathrm{diag}\left\{|g_1|^2, \cdots, |g_R|^2\right\}\right], \qquad (8)$$

which is minimized when $\frac{P_1P_2T}{4(1+P_1+P_2R)}$ is maximized.

Assume that the total power consumed in the whole network is $PT$ for transmissions of $T$ symbols. Since for every transmission, the power used at the transmitter and every relay are $P_1$ and $P_2$ respectively, we have $P = P_1 + RP_2$. Therefore,

$$\frac{P_1P_2T}{4\left(1+P_1+P_2R\right)} = \frac{P_1(P-P_1)T}{4R(1+P)} \le \frac{P^2T}{16R(1+P)},$$

with equality when

$$P_1 = \frac{P}{2} \quad \text{and} \quad P_2 = \frac{P}{2R}. \qquad (9)$$

Thus, the optimum power allocation is such that the transmitter uses half the total power and the relays share the other half. So, for large $R$, the relays spend only a very small amount of power to help the transmitter.

With this optimum power allocation, when $P \gg 1$,

$$\frac{P_1P_2T}{4\left(1+P_1+P_2\sum_{i=1}^R |g_i|^2\right)} \approx \frac{\frac{P}{2}\frac{P}{2R}T}{4\left(\frac{P}{2}+\frac{P}{2R}\sum_{i=1}^R |g_i|^2\right)} = \frac{PT}{8(R+\sum_{i=1}^R |g_i|^2)}.$$

(7) becomes

$$\mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) \lesssim \mathop{\mathrm{E}}_{g_i} \det{}^{-1}\left[I_R + \frac{PT}{8(R+\sum_{i=1}^R |g_i|^2)}(S_k - S_l)^*(S_k - S_l)\mathrm{diag}\left\{|g_1|^2, \cdots, |g_R|^2)\right\}\right]. (10)$$

It is easy to see that the expected receive SNR of the system is $\frac{P_1P_2T}{4\left(1+P_1+P_2\sum_{i=1}^R |g_i|^2\right)}$. Therefore, this optimal power allocation also maximizes the expected receive SNR for large $R$. We should emphasis that this power allocation only works for the wireless relay network described in Section 2, in which all channels are assumed to be i.i.d. Rayleigh and no path-loss is considered. It is obvious that it may not be optimal when the path-loss effect of the channels is considered.

# 6 Derivation of the Diversity

As mentioned earlier, to obtain the diversity we need to compute the expectations in (7). Since the calculation is detailed and gives little insight, to highlight the diversity result, we begin by giving a simple approximate derivation which leads to the same diversity result. As discussed in the previous section, when $R$ is large, $\sum_{i=1}^R |g_i|^2 \approx R$ with high probability. We use this approximation.

Define $M = (S_k - S_l)^*(S_k - S_l)$. We upper bound the PEP using the minimum nonzero singular value of $M$, which is denoted as $\sigma_{min}^2$. From (10),

$$\begin{aligned}
\mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) &\lesssim \mathop{\mathrm{E}}_{g_i} \det{}^{-1}\left(I_R + \frac{PT\sigma_{min}^2}{16R}\mathrm{diag}\left\{I_{\mathrm{rank}\,M}, 0\right\}\mathrm{diag}\left\{|g_1|^2, \cdots, |g_R|^2\right\}\right) \\
&= \mathop{\mathrm{E}}_{g_i} \prod_{i=1}^{\mathrm{rank}\,M}\left(1 + \frac{PT\sigma_{min}^2}{16R}|g_i|^2\right)^{-1} \\
&= \left[\int_0^\infty \left(1 + \frac{PT\sigma_{min}^2}{16R}x\right)^{-1} e^{-x}dx\right]^{\mathrm{rank}\,M} \\
&= \left(\frac{PT\sigma_{min}^2}{16R}\right)^{-\mathrm{rank}\,M}\left[-e^{\frac{16R}{PT\sigma_{min}^2}}\mathbf{Ei}\left(-\frac{16R}{PT\sigma_{min}^2}\right)\right]^{\mathrm{rank}\,M},
\end{aligned}$$

where

$$\mathbf{Ei}(\chi) = \int_{-\infty}^{\chi} \frac{e^t}{t}dt, \qquad \chi < 0$$

is the exponential integral function [20]. When $\chi < 0$,

$$\mathbf{Ei}(\chi) = c + \log(-\chi) + \sum_{k=1}^{\infty}\frac{(-1)^k\chi^k}{k \cdot k!}, \tag{11}$$

where $c$ is the Euler constant. Therefore, when $\log P \gg 1$, $e^{-\frac{16R}{PT\sigma_{min}^2}} = 1 + O\left(\frac{1}{P}\right) \approx 1$ and $-\mathbf{Ei}\left(-\frac{16R}{PT\sigma_{min}^2}\right) = \log P + O(1) \approx \log P$. Thus,

$$\mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) \lesssim \left(\frac{16R}{T\sigma_{min}^2}\right)^{\mathrm{rank}\,M}\left(\frac{\log P}{P}\right)^{\mathrm{rank}\,M} = \left(\frac{16R}{T\sigma_{min}^2}\right)^{\mathrm{rank}\,M} P^{-\mathrm{rank}\,M\left(1-\frac{\log\log P}{\log P}\right)}. \tag{12}$$

When $M$ is full rank, the diversity gain is $\min\{T,R\}\left(1-\frac{\log\log P}{\log P}\right)$. Therefore, similar to the multiple-antenna case, there is no point in having more relays than the coherence interval [4, 13]. Thus, we will henceforth always assume $T\geq R$. The diversity gain is therefore $R\left(1-\frac{\log\log P}{\log P}\right)$. (12) also shows that the PEP is smaller for bigger coherence interval $T$.

Now we give a rigorous derivation. Here is the main result.

**Theorem 2.** *Design the transmit signal at the $i$th relay as in (3), and use the power allocation in (9). Assume that $T\geq R$ and the distributed space-time code has full diversity. The PEP has the following Chernoff bound:*

$$\mathrm{P}\left(\mathbf{s}_k\to\mathbf{s}_l\right)\leq\sum_{r=0}^{R}\left(\frac{8}{PT}\right)^r M_r\left(1-e^{-x}\right)^{R-r}\sum_{j=0}^{r}B_{R+(R-k)x,x}(j,r)\left[-\mathbf{Ei}(-x)\right]^{r-j}.$$

*where*

$$M_r=\sum_{1\leq i_1<\cdots<i_r\leq R}\det{}^{-1}[M]_{i_1,\cdots,i_r}$$

*with $[M]_{i_1,\cdots,i_r}$ the $r\times r$ matrix composed of the $i_1,\cdots,i_r$ rows and columns of $M$ and*

$$\begin{aligned}B_{A,x}(j,r) &= \binom{r}{j}\sum_{i_1=1}^{r}\sum_{i_2=1}^{r-i_1}\cdots\sum_{i_j=1}^{r-i_1-\cdots-i_{j-1}}\binom{r}{i_1}\cdots\binom{r-i_1-\cdots-i_{j-1}}{i_j}\\ &\qquad\Gamma(i_1,x)\cdots\Gamma(i_j,x)A^{r-i_1-\cdots-i_j}.\end{aligned}\tag{13}$$

**Proof:** See Appendix B. □

**Corollary 1.** *If* $\log P\gg 1$,

$$\mathrm{P}\left(\mathbf{s}_k\to\mathbf{s}_l\right)\lesssim\frac{1}{P^R}\sum_{r=0}^{R}\left(\frac{8}{T}\right)^r M_r\sum_{j=0}^{r}B_{R,0}(r-j,r)\log^j P.\tag{14}$$

**Proof:** Set $x=\frac{1}{P}$[3]. When $\log P\gg 1$, since $\left[R+(R-k)\frac{1}{P}\right]^k=R^k+o(1)$, $-\mathbf{Ei}\left(-\frac{1}{P}\right)=\log P+o(1)$, $1-e^{-\frac{1}{P}}=\frac{1}{P}+o\left(\frac{1}{P}\right)$, and $\Gamma\left(i,\frac{1}{P}\right)=(i-1)!+o(1)$, (14) is obtained from (13) by omitting lower order terms of $P$. □

[3]Actually, this is not the optimal choice according to diversity gain. We can improve the diversity gain slightly by choosing an optimal x. However, the coding gain of that case is smaller than the coding gain in (14). The details will be discussed in Section 7.

**Corollary 2.** *If* $\log P \gg 1$ *and* $R \gg 1$,

$$\mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) \lesssim \frac{1}{P^R} \sum_{r=0}^{R} \left(\frac{8R}{T}\right)^r M_r \log^r P. \tag{15}$$

**Proof:** When $R \gg 1$, $B_{R,0}(0,r) >> B_{R,0}(l,r)$ for all $l > 0$ since $B_{R,0}(0,r) = R^r$ is the term with the highest order of $R$. Therefore, (15) is obtained from (14). $\square$

**Remarks:**

1. The highest order term of $P$ in (14) is the $r = j = R$ term:

$$\det{}^{-1} M \left(\frac{8R\log P}{TP}\right)^R = \det{}^{-1} M \left(\frac{8R}{T}\right)^R P^{-R\left(1-\frac{\log\log P}{\log P}\right)}, \tag{16}$$

Therefore, as in (12), distributed space-time coding achieves diversity gain $R\left(1 - \frac{\log\log P}{\log P}\right)$, which is linear in the number of relays. When $P$ is very large ($\log P \gg \log\log P$), $\frac{\log\log P}{\log P} \ll 1$, and a diversity gain about $R$ is obtained. This is the same as the diversity gain of a multiple-antenna system with $R$ transmit antennas and one receive antenna. Therefore, the relays work as if they fully cooperate and have full knowledge of the transmit signal. Generally, the diversity depends on the total transmit power $P$.

2. Note that in Theorem 2, we assume that $T \geq R$. For the general case, the rank of $M$ will be $\min\{T, R\}$ instead of $R$. By a similar argument, diversity $\min\{T, R\}\left(1 - \frac{\log\log P}{\log P}\right)$ will be obtained.

3. In a multiple-antenna system, if the transmit power (or SNR) is high, the PEP has the upper bound $\det{}^{-1} M \left(\frac{4R}{PT}\right)^R$. Comparing this with the highest order term given in (16), we can see the relay network has a performance that is

$$(3 + 10\log_{10}\log P)\, dB \tag{17}$$

worse. The 3dB difference is because in the network, each the transmitter and the relays use a half of the total power. It can be easily seen that if the total power used in the network is doubled, this 3dB difference will disappear. The second term, $10\log_{10}\log P$, is due to the diversity difference of the two cases. This analysis is verified by simulations in Section 10.

4. Corollary 2 also gives the coding gain for networks with large number of relays. When $\log P \gg 1$, the dominant term in (15) is (16). The coding gain is therefore $\det^{-1} M$, which is the same as that of the multiple-antenna case. When $P$ is not very large, the second term in (15), $\left(\frac{8R}{T}\right)^{R-1} \sum_{i=1}^{R} \det{}^{-1}[M]_{1,\cdots,i-1,i+1,\cdots,R} \frac{\log^{R-1} P}{P^R}$, cannot be ignored and even the $k = 3, 4, \cdots$ terms have non-neglectable contributions. Therefore, to have good performance, we want not only $\det M$ to be large but also $\det[M]_{i_1,\cdots,i_r}$ to be large for all $0 \leq r \leq R, 1 \leq i_1 < \cdots < i_r \leq R$. Note that

$$[M]_{i_1,\cdots,i_r} = ([S_i]_{i_1,\cdots,i_r} - [S_j]_{i_1,\cdots,i_r})^*([S_i]_{i_1,\cdots,i_r} - [S_j]_{i_1,\cdots,i_r}),$$

where $[S_i]_{i_1,\cdots,i_r} = (A_{i_1}\mathbf{s}_i, \cdots, A_{i_r}\mathbf{s}_i)$ is the space-time code when only the $i_1, \cdots, i_r$th relays are working. To have good performance when the transmit power is moderate, the distributed space-time code should be "scale-free" in the sense that it is still a good distributed space-time code when some of the relays are not working. In general, for networks with any $R$, the same conclusion can be obtained from (14).

5. Now let's look at the low total transmit power case, that is, the $P \ll 1$ case. With the same approximation $\sum_{i=1}^{R} |g_i|^2 \approx R$, using the power allocation given in (9),

$$\frac{P_1 P_2 T}{4\left(1 + P_1 + P_2 \sum_{i=1}^{R} |g_i|^2\right)} \approx \frac{\frac{P}{2}\frac{P}{2R}T}{4\,(1+P)} = \frac{P^2 T}{16R}.$$

Therefore, (7) becomes

$$\begin{aligned} \mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) &\lesssim \operatorname*{E}_{g_i} \det{}^{-1}\left(I_R + \frac{P^2 T}{16R} M \mathrm{diag}\left\{|g_1|^2, \cdots, |g_R|^2\right\}\right) \\ &= \operatorname*{E}_{g_i}\left[1 + \frac{P^2 T}{16R}\mathrm{tr}\,\left(M\mathrm{diag}\left\{|g_1|^2, \cdots, |g_R|^2\right\}\right) + o(P^2)\right]^{-1} \\ &= \operatorname*{E}_{g_i}\left(1 - \frac{P^2 T}{16R}\sum_{i=1}^{R} m_{ii}|g_i|^2\right) + o(P^2) \\ &= \left(1 - \frac{P^2 T}{16R}\mathrm{tr}\, M\right) + o(P^2), \end{aligned}$$

where $m_{ii}$ is the $(i,i)$th entry of $M$. Therefore, as in the multiple-antenna case, the coding gain at low total transmit power is $\mathrm{tr}\, M$. The design criterion is to maximize $\mathrm{tr}\, M$.

6. In our model, there is no direct link between the transmitter and the receiver. Consider now that there is a direct fading channel between the transmitter and the receiver at step one.

It is easy to see that diversity $1+R\left(1-\frac{\log\log P}{\log P}\right)$ can be obtained if the new distributed space-time code $\left[\begin{array}{cccc} \mathbf{s} & A_1\mathbf{s} & \cdots & A_R\mathbf{s}\end{array}\right]$ is fully diverse. If this direct channel exists during the second step of transmission only, let the transmitter sends the same signal $\mathbf{s}$ at step two. Exactly the same distributed space-time code and thus the same diversity can be obtained. For the case that independent fading channels exist for both steps, we design the signal sent by the transmitter at step two as $A_{R+1}\mathbf{s}$ with $A_{R+1}$ a $T\times T$ unitary matrix. It is easy to prove that diversity $2+R\left(1-\frac{\log\log P}{\log P}\right)$ can be obtained if the distributed space-time code $\left[\begin{array}{cccc} \mathbf{s} & A_1\mathbf{s} & \cdots & A_{R+1}\mathbf{s}\end{array}\right]$ is fully diverse.

7. As mentioned in Section 2, the time slots for the two transmission steps of our protocol are equal. In general, we can use $T_1$ symbol periods for the first step and $T_2$ for the second. $A_i$ should therefore be $T_2\times T_1$ unitary matrices. When the distributed space-time code is fully diverse, we can prove that the achievable diversity is $\min\{T_2,R\}\left(1-\frac{\log\log P}{\log P}\right)$. For the case of $T_1<T_2$, although $T_1$ symbols are sent from the transmitter, at most $T_2$ of them can be independent for the distributed space-time code to be full diverse. Therefore, the is no benefit in having a longer time interval for the first step. On the other hand, if we prolong the second step and have $T_2>T_1$, the diversity can be improved when there are enough relays. However, the symbol rate of transmissions decreases. Therefore, having equal time slots for the two steps maximizes the symbol rate.

# 7 Improvement In Diversity Gain

In Theorem 2, we have chosen $x=1/P$. Although this choice gives an upper bound on the PEP, it is not the optimal choice in the sense that the diversity gain obtained from this upper bound is not maximized. We can improve the diversity slightly.

**Theorem 3.** *The best diversity gain that can be obtained using distributed space-time coding is $\alpha_0 R$, where $\alpha_0$ is the solution of*

$$\alpha+\frac{\log\alpha}{\log P}=1-\frac{\log\log P}{\log P}. \tag{18}$$

*For* $\log P \gg \log\log P$*, the PEP has the following upper bound*

$$\mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) \lesssim \sum_{r=0}^{R} \left(\frac{8}{T}\right)^r M_r \sum_{l=0}^{r} B_R(r-l,r) P^{-[\alpha_0 R + (1-\alpha_0)(r-l)]}. \tag{19}$$

*If* $R \gg 1$*,*

$$\mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) \lesssim \left[\sum_{r=0}^{R} \left(\frac{8R}{T}\right)^r M_r\right] P^{-\alpha_0 R}. \tag{20}$$

**Proof:** To save space, the proof of this theorem is omitted. For details, refer to [21]. $\square$

There is no closed form for the solution of equation (18). The following theorem gives a region of $\alpha_0$ and also gives some idea about how much improvement in diversity gain is obtained.

**Theorem 4.** *For* $P > e$*,*

$$1 - \frac{\log\log P}{\log P} < \alpha_0 < 1 - \frac{\log\log P}{\log P} + \frac{\log\log P}{\log P(\log P - \log\log P)}.$$

**Proof:** Refer to [21]. $\square$

Theorem 4 indicates that the PEP Chernoff bound of the distributed space-time codes decreases faster than $\sum_{r=0}^{R}\left(\frac{8R}{T}\right)^r M_r \left(\frac{\log P}{P}\right)^R$ and slower than $\sum_{r=0}^{R}\left(\frac{8R}{T}\right)^r M_r \left(\frac{(\log P)^{1-\frac{1}{\log P - \log\log P}}}{P}\right)^R$. When $\log P \gg \log\log P$, $1 - \frac{\log\log P}{\log P}$ is a very accurate approximation of $\alpha_0$. The improvement in the diversity is small.

Now let's compare the new upper bound in (20) with the one in (15). A slightly better diversity is obtained as discussed above. However, the coding gain in (20), which is $\left[\sum_{r=0}^{R}\left(\frac{8R}{T}\right)^r M_r\right]^{-1}$, is smaller than the coding gain of (15), which is $\det M$. To compare the two, we assume that $R = T$ and that the singular values of $M$ take their maximum value, $\sqrt{2}$. Therefore the coding gain of (20) is $\left[\sum_{k=0}^{R}\begin{pmatrix} R \\ k \end{pmatrix} 4^k\right]^{-1} = 5^{-R}$. The coding gain of (15) is $4^{-R}$. The upper bound in (15) is $0.97$dB better according to coding gain.

Therefore, when $P$ is large enough, the new upper bound is tighter than the previous one since it has a larger diversity. Otherwise, the previous bound is tighter since it has a larger coding gain.

# 8 The General Distributed Linear Dispersion Code

In this section, we work on a more general type of distributed linear dispersion space-time codes [12]. The transmit signal at the $i$th relay is designed as,

$$\mathbf{t}_i = \sqrt{\frac{P_2}{P_1+1}}(A_i\mathbf{r}_i + B_i\overline{\mathbf{r}_i}), \quad i = 1, 2, \cdots, R, \tag{21}$$

where $A_i, B_i$ are $T \times T$ complex matrices. By separating the real and imaginary parts, we can write (21) equivalently as

$$\begin{bmatrix} \mathbf{t}_{i,Re} \\ \mathbf{t}_{i,Im} \end{bmatrix} = \sqrt{\frac{P_2}{P_1+1}} \begin{bmatrix} A_{i,Re} + B_{i,Re} & -A_{i,Im} + B_{i,Im} \\ A_{i,Im} + B_{i,Im} & A_{i,Re} - B_{i,Re} \end{bmatrix} \begin{bmatrix} \mathbf{r}_{i,Re} \\ \mathbf{r}_{i,Im} \end{bmatrix}. \tag{22}$$

Similar as before, for fairness and simplicity, we assume that the $2T \times 2T$ matrix,

$$\begin{bmatrix} A_{i,Re} + B_{i,Re} & -A_{i,Im} + B_{i,Im} \\ A_{i,Im} + B_{i,Im} & A_{i,Re} - B_{i,Re} \end{bmatrix},$$

is orthogonal. Therefore, the expected transmit power per transmission at every relay is $P_2$.

After straightforward calculation, the following equivalent system equation can be obtained:

$$\hat{\mathbf{x}} = \sqrt{\frac{P_1P_2T}{P_1+1}}\mathcal{H}\hat{\mathbf{s}} + \mathcal{W},$$

where

$$\mathcal{H} = \sum_{i=1}^{R} \begin{bmatrix} g_{i,Re}I_T & -g_{i,Im}I_T \\ g_{i,Im}I_T & g_{i,Re}I_T \end{bmatrix} \begin{bmatrix} A_{i,Re} + B_{i,Re} & -A_{i,Im} + B_{i,Im} \\ A_{i,Im} + B_{i,Im} & A_{i,Re} - B_{i,Re} \end{bmatrix} \begin{bmatrix} f_{i,Re}I_T & -f_{i,Im}I_T \\ f_{i,Im}I_T & f_{i,Re}I_T \end{bmatrix},$$

is the equivalent channel matrix and

$$\mathcal{W} = \begin{bmatrix} \mathbf{w}_{Re} \\ \mathbf{w}_{Im} \end{bmatrix} + \sqrt{\frac{P_2}{P_1+1}} \sum_{i=1}^{R} \begin{bmatrix} g_{i,Re}I_T & -g_{i,Im}I_T \\ g_{i,Im}I_T & g_{i,Re}I_T \end{bmatrix} \begin{bmatrix} A_{i,Re} + B_{i,Re} & -A_{i,Im} + B_{i,Im} \\ A_{i,Im} + B_{i,Im} & A_{i,Re} - B_{i,Re} \end{bmatrix} \begin{bmatrix} \mathbf{v}_{i,Re} \\ \mathbf{v}_{i,Im} \end{bmatrix},$$

is the equivalent noise. For any $T \times 1$ complex vector $\mathbf{x}$, the $2T \times 1$ real vector $\hat{\mathbf{x}}$ is defined as $\begin{bmatrix} \mathbf{x}_{Re} \\ \mathbf{x}_{Im} \end{bmatrix}$.

**Theorem 5** (ML decoder and PEP)**.** *Design the transmit signal at the $i$th relay as in (21). The ML decoding is*

$$\arg\max_{\mathbf{s}_i} \mathrm{P}\left(\mathbf{x}|\mathbf{s}_i\right) = \arg\min_{\mathbf{s}_i} \left\| \hat{\mathbf{x}} - \sqrt{\frac{P_1P_2T}{P_1+1}}\mathcal{H}\hat{\mathbf{s}}_i \right\|^2.$$

*Using the optimum power allocation given in (9), the PEP of mistaking* $\mathrm{s}_k$ *with* $\mathrm{s}_l$ *has the following Chernoff upper bound:*

$$\mathrm{P}\left(\mathbf{s}_k \rightarrow \mathbf{s}_l\right) \leq \underset{g_i}{\mathrm{E}} \det{}^{-1/2}\left(I_{2R} + \frac{PT}{8\left(R+\sum_{k=1}^{R}|g_k|^2\right)}\sum_{k=1}^{R}\mathcal{G}_k\mathcal{G}_k^t\right), \tag{23}$$

*where*

$$\mathcal{G}_k = \begin{bmatrix} g_{k,Re}I_T & -g_{k,Im}I_T \\ g_{k,Im}I_T & g_{k,Re}I_T \end{bmatrix} \begin{bmatrix} A_{i,Re}+B_{i,Re} & -A_{i,Im}+B_{i,Im} \\ A_{i,Im}+B_{i,Im} & A_{i,Re}-B_{i,Re} \end{bmatrix} \begin{bmatrix} (\mathbf{s}_k-\mathbf{s}_l)_{Re} & -(\mathbf{s}_k-\mathbf{s}_l)_{Im} \\ (\mathbf{s}_k-\mathbf{s}_l)_{Im} & (\mathbf{s}_k-\mathbf{s}_l)_{Re} \end{bmatrix}.$$

**Proof:** Refer to [21]. □

# 9 A Special Case

We have not yet been able to explicitly evaluate the expectation in (23). Our conjecture is that when $T \geq R$, the same diversity, $R\left(1-\frac{\log\log P}{\log P}\right)$, will be obtained. Here we give an analysis of a much simpler, but far from trivial, case: for any $i$, either $A_i = 0$ or $B_i = 0$. That is, each relay sends a signal that is linear in either its received signal or the conjugate of its received signal. It is clear to see that Alamouti's scheme is included in this case with $R = 2, A_1 = I_2, B_1 = 0, A_2 = 0$, and $B_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. The condition that $\begin{bmatrix} A_{i,Re}+B_{i,Re} & -A_{i,Im}+B_{i,Im} \\ A_{i,Im}+B_{i,Im} & A_{i,Re}-B_{i,Re} \end{bmatrix}$ is orthogonal becomes that $A_i$ is unitary if $B_i = 0$ and $B_i$ is unitary if $A_i = 0$.

**Theorem 6.** *Design the transmit signal at the* $i$*th relay as in (21). Use the optimum power allocation in (9). Further assume that for any* $i = 1, \cdots, R$, *either* $A_i = 0$ *or* $B_i = 0$. *The PEP of mistaking* $\mathrm{s}_i$ *with* $\mathrm{s}_j$ *has the following Chernoff upper bound:*

$$\mathrm{P}\left(\mathbf{s}_k \rightarrow \mathbf{s}_l\right) \lesssim \underset{g_i}{\mathrm{E}} \det{}^{-1}\left[I_R + \frac{PT}{8\left(R+\sum_{i=1}^{R}|g_i|^2\right)}(\hat{S}_k-\hat{S}_l)^*(\hat{S}_k-\hat{S}_l)\mathrm{diag}\left\{|g_1|^2,\cdots,|g_R|^2\right\}\right], \tag{24}$$

*where*

$$\hat{S}_k = \begin{bmatrix} A_1\mathbf{s}_k + B_1\overline{\mathbf{s}_k} & \cdots & A_R\mathbf{s}_k + B_R\overline{\mathbf{s}_k} \end{bmatrix} \tag{25}$$

*is a* $T \times R$ *matrix, which is the distributed space-time code.*

**Proof:** Refer to [21]. $\square$

(24) is exactly the same as (10) except that now the distributed space-time code is $\hat{S}$ instead of $S$. Therefore, by the same argument, the following theorem can be obtained.

**Theorem 7.** *Design the transmit signal at the $i$th relay as in (21). Use the optimum power allocation in (9). Assume $T \geq R$ and the distributed space-time code has full diversity. Define*

$$\hat{M} = (\hat{S}_k - \hat{S}_l)^*(\hat{S}_k - \hat{S}_l). \tag{26}$$

*If $\log P \gg 1$, the PEP has the following Chernoff bound:*

$$\mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) \lesssim \sum_{r=0}^{R} \left(\frac{8}{T}\right)^r \hat{M}_r \sum_{l=0}^{r} B_R(r-l,r) \frac{\log^l P}{P^R},$$

*where*

$$\hat{M}_r = \sum_{1 \leq i_1 < \cdots < i_r \leq R} \det{}^{-1} [\hat{M}]_{i_1,\cdots,i_r}.$$

*The best diversity gain that can be obtained is $\alpha_0 R$. When $\log P \gg \log\log P$,*

$$\mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) \lesssim \left[\sum_{r=0}^{R} \left(\frac{8}{T}\right)^r M_r \sum_{l=0}^{R} B_R(r-l,r)\right] P^{-[\alpha_0 R + (1-\alpha_0)(k-l)]}.$$

**Proof:** The same as the proofs of Theorems 2 and 3. $\square$

Therefore, the same diversity gain is obtained as in Section 6. The coding gain for $\log P \gg 1$ is $\det \hat{M}$. When $P$ is not very large, we want not only $\det \hat{M}$ to be large but also $\det[\hat{M}]_{i_1,\cdots,i_r}$ to be large for all $0 \leq r \leq R, 1 \leq i_1 < \cdots < i_r \leq R$. That is, to have good performance for a general transmit power, the distributed space-time code should have be "scale-free" in the sense that it is still a good code when some of the relays are not working. We can see from Theorem 7 that this general code does not improve the diversity gain of the system. However, from the definition of the new code in (25), it can improve the coding gain by code optimization.

# 10 Simulations

In this section, we give the simulated performances of the distributed space-time codes for different values of the coherence interval $T$, number of relays $R$, and total transmit power $P$. The fading coefficients between the transmitter and the relays, $f_i$, and between the receiver and the relays, $g_i$, are modeled as independent complex Gaussian random variables with zero-mean and unit-variance. The fading coefficients keep constant for $T$ channel uses. The noises at the relays and the receiver are also modeled as independent zero-mean unit-variance Gaussian additive noise. The block error rate (BLER), which corresponds to errors in decoding the vector of transmit signals $\mathbf{s}$, and the bit error rate (BER), which corresponds to errors in decoding $s_1, \cdots, s_T$, is demonstrated as the error events of interest. Note that a BLER may correspond to only a few bit errors.

The transmit signal at each relay is designed as in (3). We should remark that our goal here is to compare the performances of LD codes implemented distributively over wireless networks with the performances of the same codes in multiple-antenna systems. Therefore the actual design of the LD codes and their optimality is not an issue here: all that matters is that the codes should be the same.[4] Therefore, we generate $A_i$ randomly based on the isotropic distribution on the space of $T \times T$ unitary matrices. (It is certainly conceivable that the performance in the following figures can be improved by several dBs if $A_i$ are chosen optimally.)

The signals $s_1, \cdots, s_T$ are designed as independent $N^2$-QAM signals. Both the real and imaginary parts of $s_i$ are equal probably chosen from the $N$-PAM signal set:

$$\sqrt{\frac{6}{T(N^2-1)}}\{-(N-1)/2, \cdots, -1/2, 1/2, \cdots, (N-1)/2\},$$

where $N$ is a positive integer. The coefficient $\sqrt{\frac{6}{T(N^2-1)}}$ is used for the normalization of $\mathbf{s}$ given in formula (1). The number of possible transmit signals is $N^{2T}$. The rate of the code is, therefore,[5]

$$\frac{1}{2T} \log_2 N^{2T} = \log_2 N.$$

In the simulation of multiple-antenna systems, the number of transmit antennas is $R$ and the number of receive antennas is one. We also model the channels and noises as independent zero-mean unit-variance complex Gaussian random variables. As discussed before, the space-time code is the $T \times R$ matrix $S = \begin{bmatrix} A_1\mathbf{s} & \cdots & A_R\mathbf{s} \end{bmatrix}$. The rate of the space-time code is therefore $2\log_2 N$. In both systems, we use sphere decoding [17, 18] to obtain the ML results.

[4] The question of how to design optimal codes is an interesting one, but is beyond the scope of this paper.

[5] Due to the half-duplex protocol, $2T$ channel uses are needed for transmissions of $T$ symbols.

## 10.1 Performance of wireless networks with different $T$ and $R$

In Fig. 2, we compare the BER of relay networks for different coherence intervals $T$ and number of relays $R$. From the plot we can see that the bigger $R$, the faster the BER curve decreases, which verifies our analysis that the diversity is linear in $R$ when $T \geq R$. However, the slopes of the BER curves of networks with $T = R = 5$ and $T = 10, R = 5$ are the same when the transmit power is high. This verifies our result that the diversity only depends on $\min\{T, R\}$, which is always $R$ in our examples. Having a larger coherence interval but the same number of relays does not improve the diversity. According to the analysis in Sections 6, increasing $T$ can improve the coding gain. From the plot, we can see that the BER of the network with $T = 10, R = 5$ is about 1dB lower than that of the network with $T = R = 5$.

## 10.2 Performance comparisons of distributed space-time codes with space-time codes

In this subsection, we compare the performance of distributed space-time codes with those of space-time codes in two ways. In one, we assume that the average *total transmit power* for both systems is the same. (This is done since the noise and channel variances are everywhere normalized to unity.) In other words, the total transmit power in the network (summed over the transmitter and $R$ relays) is the same as the transmit power of the multiple-antenna system. In the other, we assume that the average *SNR at the receiver* is the same. Assuming that the total transmit power is $P$, in the distributed scheme, the average receive SNR can be calculated to be $\frac{P^2}{4(1+P)}$, and in the multiple-antenna setting it is $P$. Thus, we need roughly a $6$ dB increase in power to make the SNR of the relay network identical to that of the multiple-antenna system.

In the first example, $T = R = 5$ and $N = 2$. The BER and BLER curves are shown in Fig. 3 and 4. Fig. 3 shows the BER and BLER of the two systems with respect to the total transmit power. Fig. 4 shows the BER and BLER of the two systems with respect to the receive SNR. From the figures we can see that the performance of the multiple-antenna system is always better than that of the relay network at any power or SNR. This is what we expecte because in the multiple-antenna system, antennas of the transmitter can fully cooperate and have perfect information of the transmit signal. Also we can see from Fig. 3 that the BER and BLER curves of the multiple-antenna system decrease faster than those of the relay network. However, the differences of the slopes of the BER and BLER curves of the two systems are diminishing as the total transmit power goes higher. We can see this more clearly in Fig. 4. At low SNR regime, the BER and BLER curves of the multiple-antenna system decrease faster than those of the relay network. As SNR goes higher, the differences of the slopes of the BER and BLER curves vanishes, which indicates that the two systems have about the same diversity. This verifies our analysis of the diversity.

Fig. 5 and Fig. 6 show the performances of the two systems with $T = R = 10$ and $N = 2$. Fig. 5 shows the BER and BLER of the two systems with respect to the total transmit power. Fig. 6 shows the BER and BLER of the two systems with respect to the receive SNR. We can see from the figures that the slopes of the BER and BLER curves for the wireless relay network approach the slopes of the BER and BLER curves of the multiple-antenna systems when the transmit power increases.

In Fig. 5, at the BER of $10^{-5}$, the transmit power used in the network is about of 24.5 dB. Our analysis of (17) indicates that the performance of the relay network should be $10.5$dB worse. Reading from the plot, we get a $8$dB difference. This verifies the correctness and tightness of our upper bound.

Finally, we give an example with $T \neq R$. In this example, $T = 10$, $R = 5$ and $N = 2$. Performances of both the relay network and the multiple-antenna system with respect to the total transmit power are shown in Fig. 7. The same phenomenon can be observed.

# 11 Conclusion and Future Work

In this paper, we propose the use of LD space-time codes in a wireless relay network. We assume that the transmitter and relays do not know the channel realizations but only their statistical distribution. The ML decoding and PEP at the receiver are analyzed. The main result is that diversity $\min\{T,R\}\left(1-\frac{\log\log P}{\log P}\right)$ can be achieved, which shows that when $T \geq R$ and the average total transmit power is very high ($\log P \gg \log\log P$), the relay network has almost the same diversity as a multiple-antenna system with $R$ transmit antennas and one receive antenna. We further show that the leading order term in the PEP behaves as $\left(\frac{8R\log P}{PT}\right)^R \det{}^{-1}(S_k-S_l)^*(S_k-S_l)$, which compared to $\left(\frac{4R}{PT}\right)^R \det{}^{-1}(S_k-S_l)^*(S_k-S_l)$, the PEP of a space-time code, shows the loss of performance due to the fact that the code is implemented distributively and the relays have no knowledge of the transmit symbols. We also observe that the high SNR coding gain, $\det(S_k - S_l)^*(S_k - S_l)$, is the same as what arises in space-time coding. The same is true at low SNR where $\operatorname{tr}(S_k - S_l)^*(S_k - S_l)$ should be maximized.

We then continue investigating the diversity gain of distributed space-time coding. At high total transmit power, we improve the diversity gain achieved in Section 6 slightly (by an order no larger than $O\left(\frac{\log\log P}{\log^2 P}\right)$). Furthermore, we discuss a more general type of distributed space-time linear dispersion codes: the transmit signal from each relay is a linear combination of both its received signal and the conjugate of its received signal. For a special case, which includes the Alamouti's scheme, the same diversity gains can be obtained. Simulation results on randomly generated distributed space-time codes are demonstrated, which verifies our results.

There are several directions for future work that can be envisioned. One is to study the outage capacity of our scheme. Another is to determine whether the diversity, $\min\{T,R\}\left(1-\frac{\log\log P}{\log P}\right)$, can be improved by other coding methods. We conjecture that it cannot. Another interesting question is to study the design of distributed space-time codes. For this the PEP expression (15) in Corollary 2 should be useful. In fact, relay networks provide an opportunity for the design of space-time codes with a large number of transmit antennas, since $R$ can be quite large. Finally, it should be interesting to see whether differential space-time coding techniques can be generalized to the distributed setting. We believe that Cayley codes [22] are a suitable candidate for this.

# A Proof of Theorem 1

**Proof:** The PEP of mistaking $\mathbf{s}_k$ by $\mathbf{s}_l$ has the following Chernoff upper bound [23, 13]:

$$\mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) \le \mathrm{E}\, e^{\lambda(\ln \mathrm{P}(\mathbf{x}|\mathbf{s}_l) - \ln \mathrm{P}(\mathbf{x}|\mathbf{s}_k))}.$$

Since $\mathbf{s}_k$ is transmitted, $\mathbf{x} = \sqrt{\frac{P_1P_2T}{P_1+1}}S_kH + W$. From (6),

$$\begin{aligned}
&\ln \mathrm{P}\left(\mathbf{x}|\mathbf{s}_l\right) - \ln \mathrm{P}\left(\mathbf{x}|\mathbf{s}_k\right)\\
= &-\frac{\left[\frac{P_1P_2T}{P_1+1}H^*(S_k-S_l)^*(S_k-S_l)H + \sqrt{\frac{P_1P_2T}{P_1+1}}H^*(S_k-S_l)^*W + \sqrt{\frac{P_1P_2T}{P_1+1}}W^*(S_k-S_l)H\right]}{1+\frac{P_2}{P_1+1}\sum_{i=1}^R |g_i|^2}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) &\leq \underset{f_i,g_i,W}{\mathrm{E}}\, e^{-\frac{\lambda}{1+\frac{P_2}{P_1+1}\sum_{i=1}^R |g_i|^2}\left[\frac{P_1P_2T}{P_1+1}H^*(S_k-S_l)^*(S_k-S_l)H+\sqrt{\frac{P_1P_2T}{P_1+1}}H^*(S_k-S_l)^*W+\sqrt{\frac{P_1P_2T}{P_1+1}}W^*(S_k-S_l)H\right]} \\
&= \underset{f_i,g_i}{\mathrm{E}}\, e^{-\frac{\lambda(1-\lambda)\frac{P_1P_2T}{1+P_1}}{1+\frac{P_2}{1+P_1}\sum_{i=1}^R |g_i|^2}H^*(S_k-S_l)^*(S_k-S_l)H} \int \frac{e^{-\frac{\left(\lambda\sqrt{\frac{P_1P_2T}{P_1+1}}(S_k-S_l)H+W\right)^*\left(\lambda\sqrt{\frac{P_1P_2T}{P_1+1}}(S_k-S_l)H+W\right)}{1+\frac{P_2}{P_1+1}\sum_{i=1}^R |g_i|^2}}}{\left[\pi\left(1+\frac{P_2}{P_1+1}\sum_{i=1}^R |g_i|^2\right)\right]^T} dW \\
&= \underset{f_i,g_i}{\mathrm{E}}\, e^{-\frac{\lambda(1-\lambda)P_1P_2T}{1+P_1+P_2\sum_{i=1}^R |g_i|^2}H^*(S_k-S_l)^*(S_k-S_l)H}.
\end{aligned}$$

Choose $\lambda = \frac{1}{2}$ which maximizes $\lambda(1-\lambda) = \frac{1}{4}$ and therefore minimizes the right-hand side of the above formula. We have

$$\mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) \leq \underset{f_i,g_i}{\mathrm{E}}\, e^{-\frac{P_1P_2T}{4(1+P_1+P_2\sum_{i=1}^R |g_i|^2)}H^*(S_k-S_l)^*(S_k-S_l)H}. \tag{A.1}$$

This is the first upper bound in Theorem 1. To obtain the second upper bound we need to calculate the expectation over $f_i$. Notice that $H = G\mathbf{f}$, where $G = \mathrm{diag}\{g_1, \cdots, g_R\}$ and $\mathbf{f} = [f_1, \cdots, f_R]^t$. (A.1) becomes

$$\begin{aligned}
\mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) &\leq \underset{f_i,g_i}{\mathrm{E}}\, e^{-\frac{P_1P_2T}{4(1+P_1+P_2\sum_{i=1}^R |g_i|^2)}\mathbf{f}^*G^*(S_k-S_l)^*(S_k-S_l)G\mathbf{f}} \\
&= \underset{g_i}{\mathrm{E}} \int \frac{1}{\pi^R} e^{-\frac{P_1P_2T}{4(1+P_1+P_2\sum_{i=1}^R |g_i|^2)}\mathbf{f}^*G^*(S_k-S_l)^*(S_k-S_l)G\mathbf{f}} e^{-\mathbf{f}^*\mathbf{f}} d\mathbf{f} \\
&= \underset{g_i}{\mathrm{E}} \det\left[I_R + \frac{P_1P_2T}{4\left(1+P_1+P_2\sum_{i=1}^R |g_i|^2\right)} G^*(S_k-S_l)^*(S_k-S_l)G\right]^{-1} \\
&= \underset{g_i}{\mathrm{E}} \det\left[I_R + \frac{P_1P_2T}{4\left(1+P_1+P_2\sum_{i=1}^R |g_i|^2\right)} (S_k-S_l)^*(S_k-S_l)\mathrm{diag}\left\{|g_1|^2, \cdots, |g_R|^2\right\}\right]^{-1}
\end{aligned}$$

as desired. □

# B Proof of Theorem 2

**Proof:** Before proving the theorem, we first give a lemma that is needed.

**Lemma 1.** *If $A$ is a constant,*

$$\int_x^\infty \cdots \int_x^\infty \left(A + \sum_{i=1}^k \lambda_i\right)^k \frac{e^{-\lambda_1} \cdots e^{-\lambda_k}}{\lambda_1 \cdots \lambda_k} d\lambda_1 \cdots d\lambda_k = \sum_{j=0}^k B_{A,x}(j,k)\left[-\mathbf{Ei}(-x)\right]^{k-j}, \tag{B.1}$$

*where $\Gamma(\alpha, \chi) = \int_\chi^\infty e^{-t}t^{\alpha-1}dt$ is the incomplete gamma function [20].*

**Proof:** See Appendix C. $\square$

From (10), we need to upper bound

$$\int_0^\infty \cdots \int_0^\infty \det\left[I_R + \frac{PT}{8\left(R+\sum_{i=1}^R \lambda_i\right)} M\mathrm{diag}\left\{\lambda_1, \cdots, \lambda_R\right\}\right]^{-1} e^{-\lambda_1} \cdots e^{-\lambda_R} d\lambda_1 \cdots d\lambda_R,$$

where we have defined $\lambda_i = |g_i|^2$. Therefore, $\lambda_i$ is a random variable with exponential distribution $p_{\lambda_i}(x) = e^{-x}$. We upper bound this by breaking every integral into two parts: the integration from $0$ to an arbitrary positive number $x$ and the integration from $x$ to $\infty$, and then upper bound every one of the resulting $2^R$ terms. That is,

$$\begin{aligned}
& \mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) \\
\lesssim\ & \left(\int_0^x + \int_x^\infty\right) \cdots \left(\int_0^x + \int_x^\infty\right) \det\left[I_R + \frac{PT}{8\left(R+\sum_{i=1}^R \lambda_i\right)} M\mathrm{diag}\left\{\lambda_1, \cdots, \lambda_R\right\}\right]^{-1} e^{-\lambda_1} \cdots e^{-\lambda_R} d\lambda_1 \cdots d\lambda_R \\
=\ & \sum_{r=0}^R \sum_{1\le i_1<\cdots<i_r\le R} T_{i_1,\cdots,i_r},
\end{aligned}$$

where

$$T_{i_1,\cdots,i_r} = \underbrace{\int \cdots \int}_{\substack{\text{the } i_1,\cdots i_r\text{th integrals} \\ \text{are from } x \text{ to } \infty, \\ \text{all others are from } 0 \text{ to } x}} \det\left[I_R + \frac{PT}{8\left(R+\sum_{i=1}^R \lambda_i\right)} M\mathrm{diag}\left\{\lambda_1, \cdots, \lambda_R\right\}\right]^{-1} e^{-\lambda_1} \cdots e^{-\lambda_R} d\lambda_1 \cdots d\lambda_R.$$

Without loss of generality, we calculate $T_{1,\cdots,r}$, which is

$$\underbrace{\int_x^\infty \cdots \int_x^\infty}_{r} \underbrace{\int_0^x \cdots \int_0^x}_{R-r} \det\left[I_R + \frac{PT}{8\left(R+\sum_{i=1}^R \lambda_i\right)} M\mathrm{diag}\left\{\lambda_1, \cdots, \lambda_R\right\}\right]^{-1} e^{-\lambda_1} \cdots e^{-\lambda_R} d\lambda_1 \cdots d\lambda_R.$$

Since $M > 0$, for any $0 < \lambda_{r+1}, \cdots, \lambda_R < x$,

$$\begin{aligned}
& \det\left[I_R + \frac{PT}{8\left(R+\sum_{i=1}^R \lambda_i\right)} M\mathrm{diag}\left\{\lambda_1, \cdots, \lambda_R\right\}\right] \\
>\ & \det\left[I_R + \frac{PT}{8\left(R+(R-r)x+\sum_{i=1}^r \lambda_i\right)} M\mathrm{diag}\left\{\lambda_1, \cdots, \lambda_r, 0, \cdots, 0\right\}\right] \\
>\ & \det\left\{\frac{PT}{8\left[R+(R-r)x+\sum_{i=1}^r \lambda_i\right]} [M]_{1,\cdots,r} \mathrm{diag}\left\{\lambda_1, \cdots, \lambda_r\right\}\right\} \\
=\ & \left\{\frac{PT}{8\left[R+(R-r)x+\sum_{i=1}^r \lambda_i\right]}\right\}^r \det[M]_{1,\cdots,r} \lambda_1 \cdots \lambda_r.
\end{aligned}$$

Therefore,

$$T_{1,\cdots,r} \;<\; \left(\frac{8}{PT}\right)^r \det{}^{-1}[M]_{1,\cdots,r} \int_0^x \cdots \int_0^x e^{-\lambda_{r+1}} \cdots e^{-\lambda_R} d\lambda_{r+1} \cdots d\lambda_R$$
$$\int_x^\infty \cdots \int_x^\infty \left[R + (R-k)x + \sum_{i=1}^r \lambda_i\right]^r \frac{e^{-\lambda_1} \cdots e^{-\lambda_r}}{\lambda_1 \cdots \lambda_r} d\lambda_1 \cdots d\lambda_r.$$

Using Lemma 1,

$$T_{1,\cdots,r} < \left(\frac{8}{PT}\right)^r \det{}^{-1}[M]_{1,\cdots,r} \left(1 - e^{-x}\right)^{R-r} \sum_{j=0}^k B_{R+(R-k)x,x}(j,r) \left[-\mathbf{Ei}(-x)\right]^{r-j}.$$

In general,

$$T_{i_1,\cdots,i_r} < \left(\frac{8}{PT}\right)^r \det{}^{-1}[M]_{i_1,\cdots,i_r} \left(1 - e^{-x}\right)^{R-r} \sum_{j=0}^k B_{R+(R-r)x,x}(j,r) \left[-\mathbf{Ei}(-x)\right]^{r-j}.$$

Therefore,

$$\mathrm{P}\left(\mathbf{s}_k \to \mathbf{s}_l\right) \le \sum_{r=0}^R \left(\frac{8}{PT}\right)^r \left(\sum_{1 \le i_1 < \cdots < i_r \le R} \det{}^{-1}[M]_{i_1,\cdots,i_r}\right) \left(1 - e^{-x}\right)^{R-r} \sum_{j=0}^r B_{R+(R-k)x,x}(j,r) \left[-\mathbf{Ei}(-x)\right]^{r-j}.$$

$\square$

# C Proof of Lemma 1

**Proof:** We want to explicitly evaluate

$$I \equiv \int_x^\infty \cdots \int_x^\infty \left(A + \sum_{i=1}^k \lambda_i\right)^k \frac{e^{-\lambda_1} e^{-\lambda_2} \cdots e^{-\lambda_k}}{\lambda_1 \cdots \lambda_k} \, d\lambda_1 \cdots d\lambda_k.$$

Consider the expansion of $\left(A + \sum_{i=1}^k \lambda_i\right)^k$ into monomial terms. We have

$$\left(A + \sum_{i=1}^k \lambda_i\right)^k = \sum_{j=0}^k \left(\sum_{1 \le l_1 < \cdots < l_j \le k} \sum_{i_1=1}^k \sum_{i_2=1}^{k-i_1} \cdots \sum_{i_j=1}^{k-i_1-\cdots-i_{j-1}} C(i_1,\ldots,i_j) \lambda_{l_1}^{i_1} \lambda_{l_2}^{i_2} \cdots \lambda_{l_j}^{i_j} A^{k-i_1-\cdots-i_j}\right),$$

where $j$ denotes how many $\lambda$'s are present, $l_1,\ldots,l_j$ are the subscripts of the $j$ $\lambda$'s that appears, $i_m \ge 1$ indicates that $\lambda_{l_m}$ is taken to the $i_m$th power ( the summation should be $\sum\limits_{\substack{i_1,\ldots,i_j \ge 1 \\ \sum i_m \le k}}$, which is equivalent to $\sum_{i_1=1}^k \sum_{i_2=1}^{k-i_1} \cdots \sum_{i_j=1}^{k-i_1-\cdots-i_{j-1}}$. if we sum $i_1$ first, then $i_2$, etc. ), and finally

$$C(i_1,\ldots,i_j) = \binom{k}{i_1}\binom{k-i_1}{i_2} \cdots \binom{k-i_1-\cdots-i_{j-1}}{i_j}$$

counts how many times the term $\lambda_{l_1}^{i_1}\lambda_{l_2}^{i_2}\cdots\lambda_{l_j}^{i_j}A^{k-i_1-\cdots-i_j}$ appears in the expansion.

Thus we have

$$I = \sum_{j=0}^{k} \sum_{1\le l_1<\cdots<l_j\le k} \sum_{i_1=1}^{k} \cdots \sum_{i_j=1}^{k-i_1-\cdots-i_{j-1}} C(i_1,\ldots,i_j) I(j;l_1,\ldots,l_j;i_1,\ldots,i_j)$$

where

$$I(j;l_1,\ldots,l_j;i_1,\ldots,i_j) \equiv \int_x^\infty \cdots \int_x^\infty \lambda_{l_1}^{i_1}\lambda_{l_2}^{i_2}\cdots\lambda_{l_j}^{i_j}A^{k-i_1-\cdots-i_j}\frac{e^{-\lambda_1}\cdots e^{-\lambda_k}}{\lambda_1\cdots\lambda_k}\,d\lambda_1\cdots d\lambda_k.$$

We compute

$$\begin{aligned} I(j;l_1,\ldots,l_j;i_1,\ldots,i_j) &= A^{k-i_1-\cdots-i_j}\left(\prod_{m=1}^{j}\int_x^\infty \lambda_{l_m}^{i_m-1}e^{-\lambda_{l_m}}\,d\lambda_{l_m}\right)\prod_{i\neq i_1,\ldots i_j}\int_x^\infty \frac{e^{-\lambda_i}}{\lambda_i}\,d\lambda_i \\ &= A^{k-i_1-\cdots-i_j}\left(\prod_{m=1}^{j}\Gamma(i_m,x)\right)\left[-\mathbf{Ei}(-x)\right]^{k-j}. \end{aligned}$$

Note that the result is independent of $l_1,\ldots,l_j$. Finally adding the terms up, we have

$$\begin{aligned} I &= \sum_{j=0}^{k}\sum_{1\le l_1<\cdots<l_j\le k}\sum_{i_1=1}^{k}\cdots\sum_{i_j=1}^{k-i_1-\cdots-i_{j-1}} C(i_1,\ldots,i_j)A^{k-i_1-\cdots-i_j}\left[-\mathbf{Ei}(-x)\right]^{k-j}\prod_{m=1}^{j}\Gamma(i_m,x) \\ &= \sum_{j=0}^{k}\left[\left(\sum_{1\le l_1<\cdots<l_j\le k} 1\right)\left(\sum_{i_1=1}^{k}\cdots\sum_{i_j=1}^{k-i_1-\cdots-i_{j-1}} C(i_1,\ldots,i_j)A^{k-i_1-\cdots-i_j}\Gamma(i_1,x)\cdots\Gamma(i_j,x)\right)\right] \\ &\qquad \left[-\mathbf{Ei}(-x)\right]^{k-j} \\ &= \sum_{j=0}^{k}\left[\binom{k}{j}\sum_{i_1=1}^{k}\cdots\sum_{i_j=1}^{k-i_1-\cdots-i_{j-1}}\binom{k}{i_1}\cdots\binom{k-i_1-\cdots-i_{j-1}}{i_j}\Gamma(i_1,x)\cdots\Gamma(i_j,x)A^{k-i_1-\cdots-i_j}\right] \\ &\qquad \left[-\mathbf{Ei}(-x)\right]^{k-j} \\ &\equiv \sum_{j=0}^{k} B_{A,x}(j,k)\left[-\mathbf{Ei}(-x)\right]^{k-j}. \end{aligned}$$

Thus ends the proof. $\square$

# References


[1] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *Eur. Trans. Telecom.*, vol. 10, pp. 585–595, Nov. 1999.

[2] T. L. Marzetta and B. M. Hochwald, “Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading,” *IEEE Trans. Info. Theory*, vol. 45, pp. 139–157, Jan. 1999.

[3] G. J. Foschini, “Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas,” *Bell Labs. Tech. Jour.*, vol. 1, no. 2, pp. 41–59, 1996.

[4] V. Tarokh, N. Seshadri, and A. R. Calderbank, “Space-time codes for high data rate wireless communication: Performance criterion and code construction,” *IEEE Trans. Info. Theory*, vol. 44, pp. 744–765, 1998.

[5] A. Sendonaris, E. Erkip, and B. Aazhang, “User cooperation diversity-part I: System description,” *IEEE Trans. Comm.*, vol. 51, pp. 1927–1938, Nov. 2003.

[6] A. Sendonaris, E. Erkip, and B. Aazhang, “User cooperation diversity-part II: Implementation aspects and performance analysis,” *IEEE Trans. Comm.*, vol. 51, pp. 1939–1948, Nov. 2003.

[7] Y. Tang and M. C. Valenti, “Coded transmit macrodiversity: block space-time codes over distributed antennas,” in *Vehicular Tech. Conf., 2001. VTC 2001 Spring. IEEE VTS 53rd*, vol. 2, pp. 1435 – 1438, May 2001.

[8] J. N. Laneman and G. W. Wornell, “Distributed space-time-coded protocols for exploiting cooperative diversity in wireless betwork,” *IEEE. Trans. Info. Theory*, vol. 49, pp. 2415–2425, Oct. 2003.

[9] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler, “Fading relay channels: Performance limits and space-time signal design,” *IEEE Jour. on Selected Areas in Comm.*, pp. 1099 – 1109, Aug. 2004.

[10] Y. Hua, Y. Mei, and Y. Chang, “Wireless antennas-making wireless communications perform like wireline communications,” in *IEEE AP-S Topical Conf. on Wireless Comm. Tech.*, Oct. 2003.

[11] Y. Chang and Y. Hua, “Application of space-time linear block codes to parallel wireless relays in mobile ad hoc networks,” in *the 36th Asilomar Conf. on Signals, Systems and Computers*, Nov. 2003.

[12] B. Hassibi and B. Hochwald, “High-rate codes that are linear in space and time,” *IEEE Trans. Info. Theory*, vol. 48, pp. 1804–1824, July 2002.

[13] B. M. Hochwald and T. L. Marzetta, “Unitary space-time modulation for multiple-antenna communication in Rayleigh flat-fading,” *IEEE Trans. Info. Theory*, vol. 46, pp. 543–564, Mar. 2000.

[14] M. Gasptar and M. Vetterli, "On the capacity of wireless networks: the relay case," in *IEEE Infocom*, June 2002.

[15] A. F. Dana and B. Hassibi, "On the power-efficiency of sensory and ad-hoc wireless networks," *Submitted to IEEE Trans. Info. Theory*, 2003.

[16] A. F. Dana, M. Sharif, R. Gowaikar, B. Hassibi, and M. Effros, "Is broadcase plus multi-access optimal for gaussian wireless network?," in *the 36th Asilomar Conf. on Signals, Systems and Computers*, Nov. 2003.

[17] M. O. Damen, K. Abed-Meraim, and M. S. Lemdani, "Further results on the sphere decoder algorithm," *Submitted to IEEE Trans. Info. Theory*, 2000.

[18] B. Hassibi and H. Vikalo, "On the expected complexity of integer least-squares problems," in *the 2002 IEEE International Conf. on Acoustics, Speech, and Signal Processing*, April 2002.

[19] M. Evans, N. Hastings, and B. Peacock, *Statistical Distributions*. Wiley, 2nd ed., 1993.

[20] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*. Academic Press, 6nd ed., 2000.

[21] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *Technical report*, 2004.

[22] B. Hassibi and B. Hochwald, "Cayley differential unitary space-time codes," *IEEE Trans. Info. Theory*, pp. 1458–1503, June 2002.

[23] H. L. V. Trees, *Detection, Estimation, and Modulation Theory–Part I*. New York: Wiley, 1968.

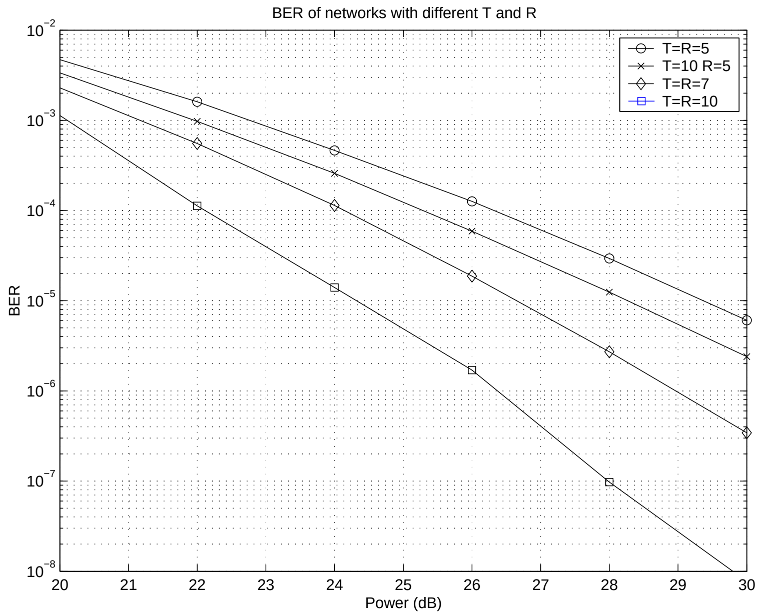


Figure 2: The BER comparison wireless networks with different $T$ and $R$.

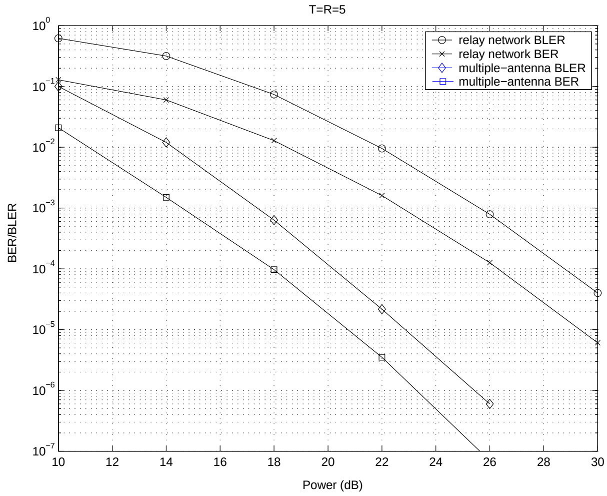


Figure 3: The comparison of the relay network with the multiple-antenna system with $T = R = 5$ and the same total transmit power.

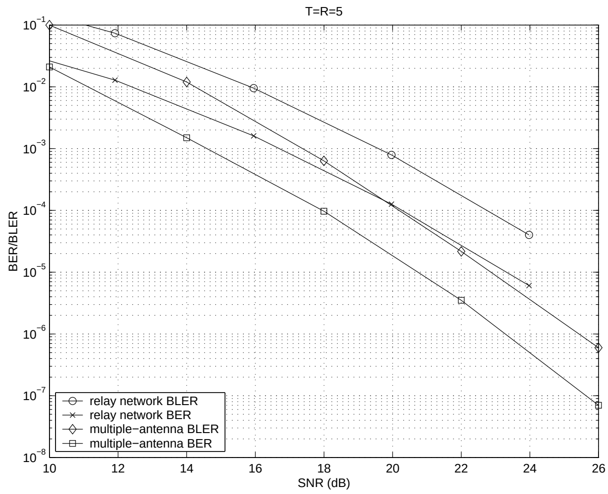


Figure 4: The comparison of the relay network with the multiple-antenna system with $T = R = 5$ and the same receive SNR.

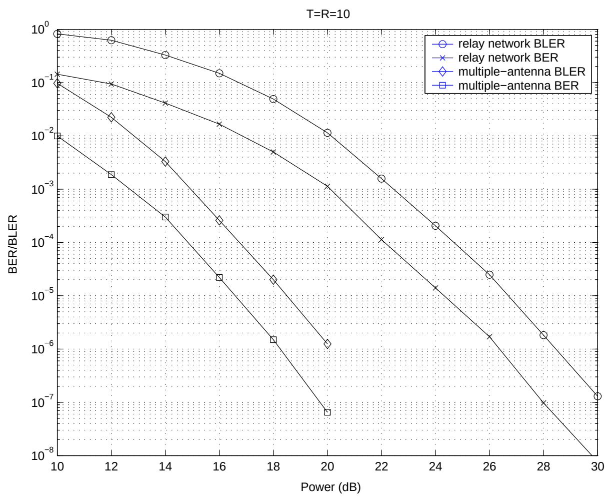


Figure 5: The comparison of the relay network with the multiple-antenna system with $T = R = 10$ and the same total transmit power.

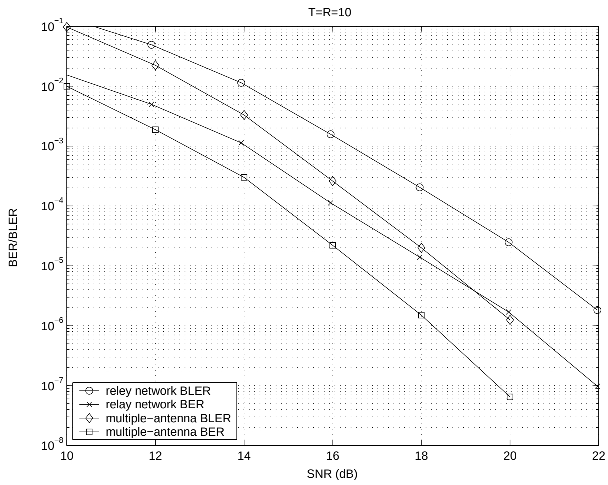


Figure 6: The comparison of the relay network with the multiple-antenna system with $T = R = 10$ and the same receive SNR.

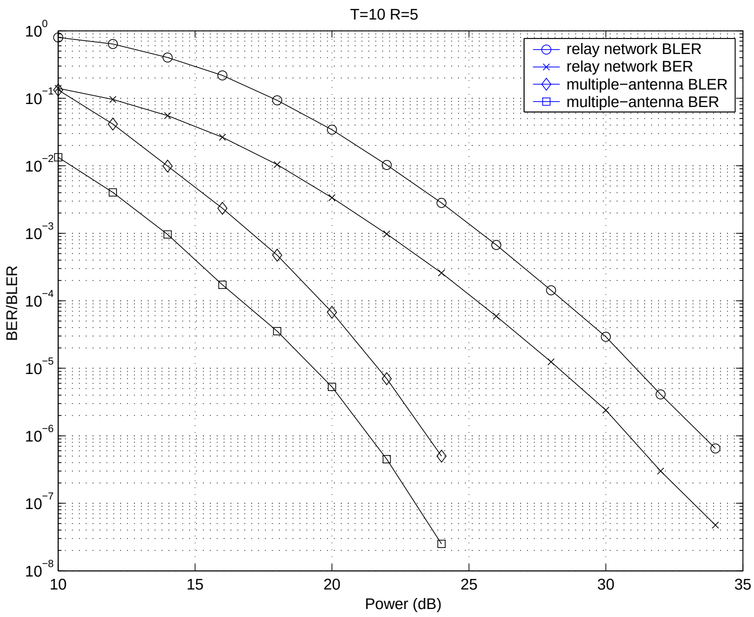


Figure 7: The comparison of the relay network with the multiple-antenna system with $T = 10, R = 5$ and the same total transmit power.