

Cyclic Distributed Space-Time Codes for Wireless Relay Networks with no Channel Information

Frédérique Oggier and Babak Hassibi¹

Department of Electrical Engineering

California Institute of Technology

Pasadena, CA 91125

{frederique,hassibi}@systems.caltech.edu

Abstract

In this paper, we present a coding strategy for half duplex wireless relay networks, where we assume no channel knowledge at any of the transmitter, receiver or relays. The coding scheme uses distributed space-time coding, that is, the relay nodes cooperate to encode the transmitted signal so that the receiver senses a space-time codeword. It is inspired by noncoherent differential techniques. The proposed strategy is available for any number of relays nodes. It is analyzed, and shown to yield a diversity linear in the number of relays. We also study the resistance of the scheme to relay node failures, and show that a network with R relay nodes and d of them down behaves, as far as diversity is concerned, as a network with $R - d$ nodes. Finally, our construction can be easily generalized to the case where the transmitter and receiver nodes have several antennas.

I. PRELIMINARIES

Wireless relay networks have recently received a lot of attention. Coding strategies inspired by space-time coding techniques, where the transmit antennas cooperate to resist the fading, have been proposed, yielding *cooperative diversity* schemes, where the relay nodes form a virtual multiple antennas array to obtain the diversity advantage known to be achieved by MIMO systems [17], [1], [10], [2], [23], [5], [4]. These works have focused on different aspects of coding for wireless networks. In [2], the capacity of the network is computed, while a lot of work has been done on finding codes optimal with respect to the so-called diversity-multiplexing gain tradeoff [1], [23], [5].

¹This work was supported in part by NSF grant CCR-0133818, by Caltech's Lee Center for Advanced Networking and by a grant from the David and Lucille Packard Foundation. Some parts of this paper first appeared in Allerton 2006 and ICASSP07.

A. Previous work on distributed space-time coding

In [10], a strategy called *distributed space-time coding* has been presented. The idea is to have the relay nodes apply a simple operation on their received signal, in such a way that the signal at the receiver appears as a space-time code. This involves a two-step transmission, where the transmitter first broadcasts the information to the relays, and second, the relays forward the received signal after having performed a unitary matrix multiplication on the signal. The pairwise probability of error and the diversity gain of such coding strategy has been computed. Random distributed space-time codes are used, that is, the matrices used at the relays are generated randomly.

In [18], [14], [12], distributed space-time codes following the two-step protocol of [10] which reach the diversity with good coding gain have been proposed. In [11], it was argued that in order to increase the data rate of the network, one may use multiple antennas at both the transmitter and the receiver nodes. The two-step transmission of [10] has been generalized, and the diversity and coding gain of random distributed space-time codes in the multiple antenna nodes scenario has been computed. Algebraic distributed space-time codes for multiple antenna nodes reaching the diversity and having a better coding gain than random codes have been presented in [19]. Recently, the setting of [10] has been generalized in [16], where the constraint of unitary matrix multiplication at the relays has been relaxed.

In [4], a different two-step protocol to implement distributed space-time code has been presented. During the first phase, the source sends a signal and the relays decode their received signal if the channel was not in outage. During the second phase, the relays which decoded cooperate to encode a space-code. The receiver knows the received signals from both the phases, and decode accordingly. This gives a practical scheme, following the scheme of [17], where the code construction is based on information theoretic arguments.

All the above strategies assume a “coherent” channel in the sense that the receiver knows all the channel paths occurring during communication. Recently, different authors started to investigate coding for a “noncoherent” channel, assuming that none of the transmitter, relay nodes nor receiver, knows about the channel. A natural approach has been to recall what are the techniques used for the noncoherent MIMO channel.

B. Previous differential coding techniques

Unitary differential modulation is a technique to code for noncoherent MIMO channels [7], [9]. By asking the transmitter to send at each time t a codeword multiplied by what was sent at time $t - 1$,

differential modulation yields a decoding strategy that does not depend on the channel, and which is thus suitable for a noncoherent channel.

Using a similar approach for wireless networks has been investigated by several authors. In [24], differential modulation has been adapted for a decode-and-forward strategy. Differential distributed space-time coding when the nodes do not decode has been proposed independently in [15], [20], [13], in the single antenna case. The works in [15], [13] consider a joint design of the matrices at the relays with the transmitted signal, while the work in [20] suggests a construction where the matrices at the relays are fixed first, while the transmitted signal is optimized independently.

C. Contribution and organization of the paper

In this work, we are interested in designing a coding strategy for wireless networks where we assume no channel information, that is, the transmitter and the relays are assumed not to know the channel, and the receiver decodes with no knowledge of the different paths used during communication. We furthermore consider a network where nodes are small devices with few resources, so we do not assume that they are able to decode. Instead, they just do a simple operation on the received signal for which they do not need to know the fading. Our strategy is inspired by noncoherent MIMO *unitary differential modulation*. We will consider both the cases when the transmitter and receiver have one or several antennas. Our construction is valid for any number of relays, and any number of transmit/receive antennas.

We organize this paper as follows. In Section II, we start by recalling the wireless network we consider, and how distributed coding is performed. We then present a distributed coding strategy that emulates communication over a noncoherent MIMO channel. This allows us to define a differential coding strategy, described in Section III. For convenience, the idea behind differential modulation is recalled. In Section IV, we propose a mismatched decoder and show that this yields a diversity gain linear in the number of relays. We further analyze the behavior of the proposed coding strategy in case of node failures. In Section V, we discuss issues related to code constructions and we provide simulation results. Finally we show that our construction can be generalized to the case when both the transmitter and receiver nodes have several transmit antennas.

II. NONCOHERENT DISTRIBUTED SPACE-TIME CODING

A. Distributed space-time codes

Following the setting defined in [10], consider a wireless network with $R+2$ nodes which are randomly and independently distributed. Two nodes, a receiver and a transmitter, want to communicate, while the

R other nodes serve as relays. Every node is equipped with a single antenna. It can transmit and receive, but not simultaneously. The channels are denoted by f_i from the transmitter to the i th relay, and by g_i from the i th relay to the receiver (see Fig. 1). Both channels are assumed independent complex Gaussian $\mathcal{CN}(0, 1)$. We assume a coherence interval of length $T \geq R$ (there is no need to have more relays than coherence time, since it is shown in [10] that the diversity of the system depends on $\min\{T, R\}$). The total power of the system ρ is equally distributed between the transmitter and the relays, so that the transmitter has an energy of $P_1 = \rho/2$, while each relay has $P_2 = \rho/(2R)$.

The transmission is done in two steps:

- **Step 1: at the transmitter.** Let $\mathbf{s} = (s_1, \dots, s_T)^t$ be the signal to be sent, from the codebook $\{\mathbf{s}_1, \dots, \mathbf{s}_L\}$ of cardinality L . The vector $\mathbf{s}$ is normalized such that $E[\mathbf{s}^\dagger \mathbf{s}] = 1$. Let P_1 be the average power available for each transmission. From time 1 to T , the transmitter sends the signals $\sqrt{P_1 T} \mathbf{s}$ to each relay. The received signal at the i th relay is given by

$$\mathbf{r}_i = \sqrt{P_1 T} f_i \mathbf{s} + \mathbf{v}_i, \quad i = 1, \dots, R, \quad (1)$$

where $\mathbf{v}_i$ is the complex Gaussian noise $\mathcal{CN}(0, 1)$ at the i th relay.

- **Step 2: at the relays.** The i th relay multiplies its received signal by a unitary matrix A_i (see Remark 1), and sends from time $T + 1$ to $2T$ the signals $\mathbf{t}_i$ to the receiver, where

$$\mathbf{t}_i = \sqrt{\frac{P_2}{P_1 + 1}} A_i \mathbf{r}_i, \quad (2)$$

and the normalization factor is chosen so that $E[\mathbf{t}_i^\dagger \mathbf{t}_i] = P_2 T$. The signal at the receiver is given by:

$$\mathbf{y} = \sum_{i=1}^R g_i \mathbf{t}_i + \mathbf{w}, \quad (3)$$

where $\mathbf{w}$ is the complex Gaussian noise $\mathcal{CN}(0, 1)$ at the receiver.

Remark 1: In order to have an equitable protocol among different users and among different times instants, the matrices A_i are assumed unitary. This also guarantees that the noise $\mathbf{w}$ at the receiver remains temporally white. Note that it has been shown recently [16] that for that purpose, it is enough to require the $A_i \mathbf{v}_i$, $i = 1, \dots, R$ to be uncorrelated.

From (3), (2) and (1), the received signal is given by

$$\mathbf{y} = \sqrt{\frac{P_1 P_2 T}{P_1 + 1}} S H + W, \quad (4)$$

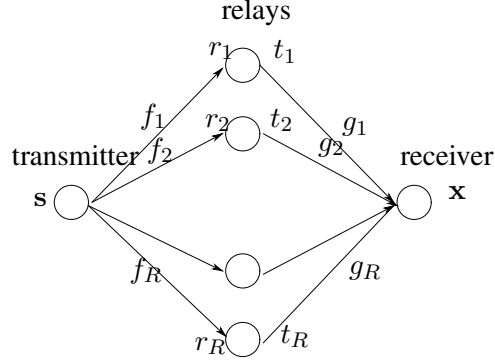


Fig. 1. The single antenna wireless relay network model

with

$$S = [A_1 \mathbf{s} \cdots A_R \mathbf{s}], \quad H = [f_1 g_1, \dots, f_R g_R]^t \quad (5)$$

and

$$W = \sqrt{\frac{P_2}{P_1 + 1}} \sum_{i=1}^R g_i A_i \mathbf{v}_i + \mathbf{w}.$$

The $T \times R$ matrix S works like a space-time code in a multiple-antenna system. It is called a *distributed space-time code* since it has been generated in a distributed way by the relay nodes.

B. A noncoherent channel

Let us now consider the equation

$$\mathbf{y} = \sqrt{\frac{P_1 P_2 T}{P_1 + 1}} S H + W,$$

derived above, but assume that none of the fadings are known. In a traditional noncoherent MIMO setting, it has been argued [6] that the transmitted codeword S has to be unitary.

Recall that here

$$S = [A_1 \mathbf{s}, \dots, A_R \mathbf{s}].$$

The aim is now to design the signal constellation $\{\mathbf{s}_1, \dots, \mathbf{s}_L\} \ni \mathbf{s}$ and the unitary matrices A_i , $i = 1, \dots, R$, so that the $T \times R$ matrix S is unitary, i.e., $S^\dagger S = \mathbf{I}_R$. Let

$$\mathbf{s}' = \frac{1}{\sqrt{T}}(1, \dots, 1)^t, \quad \mathbf{s}_i = U_i \mathbf{s}', \quad i = 1, \dots, L,$$

where U_i 's are $T \times T$ unitary matrices and $\mathbf{s}'$ is normalized so that $E[\mathbf{s}'^\dagger \mathbf{s}'] = 1$.

Assume now that there exists a $T \times T$ matrix M such that $MM^\dagger = T$. We can then choose the matrices A_i to be

$$A_i = \text{diag}(M_i), \quad i = 1, \dots, R,$$

where M_i denotes a column of M (recall that $T \geq R$). Let $U_j, j = 1, \dots, L$, be diagonal unitary matrices, thus commuting with all A_i . We have, when sending the codeword $\mathbf{s}_j$, that the corresponding distributed space-time codeword S_j is given by

$$S_j = [A_1 \mathbf{s}_j, \dots, A_R \mathbf{s}_j] = [U_j A_1 \mathbf{s}', \dots, U_j A_R \mathbf{s}'] = U_j M / \sqrt{T},$$

and $S_j^\dagger S_j = M^\dagger U_j^\dagger U_j M / T = \mathbf{I}_R$ for all transmitted signal $\mathbf{s}_j$. Let us keep in mind that the matrices A_i have to be unitary.

Such matrices M can be found in the class of Butson-Hadamard matrices (see for example [8]).

Definition 1: A *Generalized Butson-Hadamard* (GBH) matrix is a $T \times T$ matrix M with coefficients in a ring such that

$$MM^* = M^*M = T\mathbf{I}_T$$

where M^* is the transpose of the matrix of inverse elements of M : $m_{ij}^* = m_{ji}^{-1}$.

If the coefficients of M are chosen to be roots of unity, then $m_{ij}^{-1} = \overline{m_{ij}}$, i.e., the inverse is the conjugate, so that

$$MM^\dagger = M^\dagger M = T\mathbf{I}_T.$$

Furthermore, this implies that all matrices A_i are unitary, $i = 1, \dots, R$.

Example 1: Let $\zeta_3 = \exp(2i\pi/3)$ be a primitive 3rd root of unity. Then the matrix

$$M = \begin{pmatrix} 1 & 1 & 1 \\ 1 & \zeta_3 & \zeta_3^2 \\ 1 & \zeta_3^2 & \zeta_3 \end{pmatrix}$$

is a Butson-Hadamard matrix. It is easy to check that $MM^\dagger = 3\mathbf{I}_3$. Also

$$\text{diag}(M_1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{diag}(M_2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \zeta_3 & 0 \\ 0 & 0 & \zeta_3^2 \end{pmatrix}, \quad \text{diag}(M_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \zeta_3^2 & 0 \\ 0 & 0 & \zeta_3 \end{pmatrix}$$

are clearly unitary.

Other examples of such matrices can be found in [8]. Note that the tensor product of two GBH matrices

is again a GBH matrix. This is thus a convenient way of building GBH matrices for a any dimension T .

Example 2: The following tensor product

$$M \otimes M = \begin{pmatrix} 1 & 1 & 1 \\ 1 & \zeta_3 & \zeta_3^2 \\ 1 & \zeta_3^2 & \zeta_3 \end{pmatrix} \otimes \begin{pmatrix} 1 & 1 & 1 \\ 1 & \zeta_3 & \zeta_3^2 \\ 1 & \zeta_3^2 & \zeta_3 \end{pmatrix}$$

is a Butson-Hadamard matrix that can be used for a network with 9 relay nodes.

Note that our strategy has the advantage of having independent designs for the matrices A_i and U_j , unlike for example the scheme given in [15]. Thus, for a given number of relays R , the matrices are built once for all and given to the relays. The code design then simplifies to considering the matrices U_j . This question is addressed in Section V.

III. A DIFFERENTIAL CODING STRATEGY

In this section, we give a strategy to implement a differential distributed coding scheme (for the sake of completeness, the standard differential scheme for MIMO channel is recalled below). It is a priori not clear how to emulate differential coding in a distributed setting. Where should the differential encoding take place? One can imagine the relays cooperating to encode differentially, similarly to the coherent case where relays encode the space-time codes, as well as having the transmitter itself collaborating with the relays. However, the construction presented in the previous section clearly suggests the approach where the differential encoder is actually at the transmitter itself. The relays cooperate not to encode differentially, but to encode a unitary space-time code.

A. Differential MIMO coding

Consider a Rayleigh flat fading channel with M transmit antennas and N receive antennas, with *unknown channel information*. The channel is used in blocks of M channel uses, so that the transmitted signal can be represented as an $M \times M$ matrix $\mathbf{S}_t$, where $t = 0, 1 \dots$ represents the block channel use. If we assume that the channel is constant over M channel uses, we may write it as

$$\mathbf{Y}_t = \sqrt{\rho} \mathbf{S}_t \mathbf{H}_t + \mathbf{W}_t, \quad t = 0, 1, \dots \quad (6)$$

Here $\mathbf{H}_t$, the channel matrix, and $\mathbf{W}_t$, the noise matrix, are two $M \times N$ matrices with independent complex normal coefficients, and ρ is the expected SNR at each receiver antenna.

Differential unitary space-time modulation [7], [9] is implemented by sending at time t a new codeword multiplied by the signal transmitted at time $t - 1$. The transmitted signal $\mathbf{S}_t$ is thus (assuming $\mathbf{S}_0 = \mathbf{I}$)

$$\mathbf{S}_t = \mathbf{X}_{z_t} \mathbf{S}_{t-1}, \quad t = 1, 2, \dots, \quad (7)$$

where $z_t \in \{0, \dots, L - 1\}$ is the data to be transmitted, and $\mathcal{C} = \{\mathbf{X}_0, \dots, \mathbf{X}_{L-1}\}$ the constellation to be designed. It can be seen from the above equation that the codebook has to contain *unitary* matrices, to prevent $\mathbf{S}_t$ to tend either to zero or infinity.

If we further assume the channel constant for $2M$ consecutive uses, we get from (6) and (7) that

$$\begin{aligned} \mathbf{Y}_t &= \sqrt{\rho} \mathbf{X}_{z_t} \mathbf{S}_{t-1} \mathbf{H} + \mathbf{W}_t \\ &= \mathbf{X}_{z_t} (\mathbf{Y}_{t-1} - \mathbf{W}_{t-1}) + \mathbf{W}_t \\ &= \mathbf{X}_{z_t} \mathbf{Y}_{t-1} + \mathbf{W}'_t, \end{aligned}$$

where $\mathbf{W}'_t = \mathbf{W}_t - \mathbf{X}_{z_t} \mathbf{W}_{t-1}$ is statistically independent of $\mathbf{X}_{z_t}$, since $\mathbf{X}_{z_t}$ is unitary. Since the matrix $\mathbf{H}$ does not appear in the last equation, this means that differential modulation allows decoding without knowledge of the channel.

Remark 2: Note that in practice, the coherence interval is usually much larger than $2M$. In fact, what is often encountered is a continuously-fading channel. The assumption that the differential scheme exploits is that the channel is roughly constant over “any” $2M$ channel uses.

B. A differential encoder

It is straightforward to adapt the two-steps transmission described in Subsection II-A to allow differential encoding and decoding.

Assume that the transmitter wants to send at time $t + nT$ the data z_{t+nT} . It is encoded into a unitary matrix $U(z_{t+nT})$. We consider the following strategy:

- 1) Let $\mathbf{s}_t = U(z_t) \mathbf{s}'$ be the signal to be transmitted, where $\mathbf{s}' = (1, \dots, 1)^t / \sqrt{T}$ is normalized so that $E[\mathbf{s}_t^\dagger \mathbf{s}_t] = 1$. Let P_1 be the average energy available for each transmission. From time $t + 1$ to $t + n$, the transmitter sends the signal $\sqrt{P_1 T} \mathbf{s}_t$ to each relay. From time $t + T + 1$ to $t + 2T$, the signal to be transmitted is $\mathbf{s}_{t+T} = U(z_{t+T}) \mathbf{s}_t$.
- 2) At the i th relay, the received signals are (indexing the signals as a function of the time at which they have been sent)

$$\mathbf{r}_i(t) = \sqrt{P_1 T} f_i \mathbf{s}_t + \mathbf{v}_i(t), \quad i = 1, \dots, R,$$

and

$$\mathbf{r}_i(t+T) = \sqrt{P_1 T} f_i U(z_{t+T}) \mathbf{s}_t + \mathbf{v}_i(t+T), \quad i = 1, \dots, R.$$

- 3) The i th relay multiplies its received signal by a unitary matrix A_i , where A_i has been built using a Butson-Hadamard matrix as described in Subsection II-B. From time $t+T+1$ to $t+2T$, the transmitted signal is $\mathbf{t}_i(t) = \sqrt{\frac{P_2}{P_1+1}} A_i \mathbf{r}_i(t)$, and similarly from time $t+2T+1$ to $t+3T$: $\mathbf{t}_i(t+T) = \sqrt{\frac{P_2}{P_1+1}} A_i \mathbf{r}_i(t+T)$.
- 4) At time $t+2T$, resp. $t+3T$, the received signals are similar to (4)-(5):

$$\mathbf{y}(t) = \sqrt{c_\rho} \sum_{i=1}^R g_i f_i A_i \mathbf{s}_t + W(t) \quad (8)$$

$$\mathbf{y}(t+T) = \sqrt{c_\rho} \sum_{i=1}^R g_i f_i A_i U(z_{t+T}) \mathbf{s}_t + W(t+T), \quad (9)$$

where

$$W(t) = \sqrt{\frac{P_2}{P_1+1}} \sum_{i=1}^R g_i A_i \mathbf{v}_i(t) + \mathbf{w}(t) \text{ and } c_\rho = \frac{P_2 P_1 T}{P_1 + 1}.$$

Under the assumption that A_i and $U(z_{t+T})$ commute, for all i and for all possible choices of $U(z_{t+T})$, we can plug equation (8) into equation (9), which yields

$$\mathbf{y}(t+T) = U(z_{t+T}) \mathbf{y}(t) + [W(t+T) - U(z_{t+T}) W(t)]. \quad (10)$$

Note that the channel coefficients f_i and g_i do not appear in (10). Also, the assumption that A_i and $U(z_{t+T})$ commute is valid since both the unitary codewords U_j and the matrices A_i are chosen diagonal.

IV. PAIRWISE PROBABILITY OF ERROR AND NODES FAILURES

A. A decoding rule

Emulating the point to point case, a natural candidate for the differential decoder is

$$\arg \min_{U_l, l=1, \dots, L} \|\mathbf{y}(t+T) - U_l \mathbf{y}(t)\|^2. \quad (11)$$

Let us restrict to the case where $T = R$. In order to analyze this strategy, we consider two instances of the noncoherent channel,

$$\mathbf{y}(t) = \sqrt{c_\rho} S(t) H + W(t)$$

where H is an $T \times T$ matrix unknown at both the transmitter and receiver, $S(t)$ is a $T \times T$ unitary matrix, and c_ρ is a constant which depends on the SNR ρ , that is,

$$\begin{pmatrix} \mathbf{y}(t) \\ \mathbf{y}(t+T) \end{pmatrix} = \sqrt{c_\rho} \begin{pmatrix} S(t) \\ S(t+T) \end{pmatrix} H + \begin{pmatrix} W(t) \\ W(t+T) \end{pmatrix}.$$

Since $S(t)$ and $S(t)\psi$ are indistinguishable for an arbitrary unitary $T \times T$ matrix ψ , we preprocess the signal so that

$$\begin{pmatrix} \mathbf{y}(t) \\ \mathbf{y}(t+T) \end{pmatrix} = \sqrt{c_\rho} \begin{pmatrix} \mathbf{I}_T \\ U_k \end{pmatrix} H + \begin{pmatrix} W(t) \\ W(t+T) \end{pmatrix},$$

for U_k a unitary matrix belonging to the codebook. To suit the network model, we have

$$W(t) = \sqrt{\frac{P_2}{P_1+1}} \sum_{i=1}^T A_i g_i \mathbf{v}_i(t) + \mathbf{w}(t)$$

and $H = D_g \mathbf{f}$, where $D_g = \text{diag}(g_1, \dots, g_T)$ and $\mathbf{f} = (f_1, \dots, f_T)^t$. Furthermore, we have $c_\rho = \frac{P_2 P_1 T}{P_1 + 1}$, and we denote $c'_\rho = \frac{P_2}{P_1 + 1}$. Recall that $P_1 = \rho/2$, $P_2 = \rho/(2R)$ and ρ is the total power of the system. Because of the two steps transmission, both the noise $W(t)$ and the channel matrix H contains products of Gaussian random variables, which makes a precise analysis difficult. In this work, we thus consider a mismatched decoder, and we will show that such a decoder already gives the diversity.

B. Pairwise probability of error

Let us now compute the pairwise error probability of decoding with a mismatched decoder. Let $\mathbf{I}_n$ denote the identity matrix of size n . Knowing $\mathbf{g} = (g_1, \dots, g_T)^t$, we have

$$\begin{aligned} E[W(t)W(t)^\dagger] &= E[W(t+T)W(t+T)^\dagger] \\ &= c'_\rho E\left[\sum_{i,j=1}^T A_i g_i \mathbf{v}_i(t) (A_j g_j \mathbf{v}_j(t))^\dagger\right] + E[W(t)W(t)^\dagger] \\ &= c'_\rho \sum_{i=1}^T A_i A_i^\dagger |g_i|^2 \mathbf{I}_T + \mathbf{I}_T \\ &= (c'_\rho \|\mathbf{g}\|^2 + 1) \mathbf{I}_T. \end{aligned}$$

Let $\mathbf{y} = [\mathbf{y}(t) \ \mathbf{y}(t+T)]^t$, we have

$$\Sigma := E[\mathbf{y}\mathbf{y}^\dagger] = \begin{pmatrix} c_\rho D_{|g|} + (c'_\rho \|\mathbf{g}\|^2 + 1) \mathbf{I}_T & c_\rho D_{|g|} U_k^\dagger \\ c_\rho D_{|g|} U_k & c_\rho D_{|g|} + (c'_\rho \|\mathbf{g}\|^2 + 1) \mathbf{I}_T \end{pmatrix},$$

where $D_{|g|} = \text{diag}(|g_1|^2, \dots, |g_T|^2)$.

The pairwise probability of error is given by

$$\begin{aligned} P(U_k \rightarrow U_l) &= P(\|\mathbf{y}(t+T) - U_k \mathbf{y}(t)\|^2 \geq \|\mathbf{y}(t+T) - U_l \mathbf{y}(t)\|^2 \mid U_k \text{ is sent}) \\ &= P(\|[-U_k \quad \mathbf{I}_T] \mathbf{y}\|^2 \geq \|[-U_l \quad \mathbf{I}_T] \mathbf{y}\|^2 \mid U_k \text{ is sent}) \\ &= P(\mathbf{y}^\dagger \mathbf{U} \mathbf{y} \geq 0 \mid U_k \text{ is sent}) \end{aligned}$$

with $k \neq l$ and

$$\mathbf{U} = \begin{pmatrix} 0 & U_l^\dagger - U_k^\dagger \\ U_l - U_k & 0 \end{pmatrix}. \quad (12)$$

We start by computing $P(\mathbf{y}^\dagger \mathbf{U} \mathbf{y} \geq 0)$ knowing $\mathbf{g}$.

Proposition 1: We have that, at high SNR,

$$P(\mathbf{y}^\dagger \mathbf{U} \mathbf{y} \geq 0 \mid \mathbf{g}) \leq \frac{1}{2} \det(\mathbf{I}_T + \frac{c_\rho}{8(\frac{1}{T}\|\mathbf{g}\|^2 + 1)} D_{|g|} (U_k - U_l)(U_k - U_l)^\dagger)^{-1}.$$

Proof: We have

$$\begin{aligned} P(\mathbf{y}^\dagger \mathbf{U} \mathbf{y} \geq 0 \mid \mathbf{g}) &= E[u(\mathbf{y}^\dagger \mathbf{U} \mathbf{y})] \\ &= E\left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega \mathbf{y}^\dagger \mathbf{U} \mathbf{y}}}{i\omega} d\omega\right] \end{aligned}$$

where u is the step function ($u(x) = 1$ if $x > 0$, $u(x) = 0$ else), and the second equality is the Fourier transform of u . Computing the expectation yields

$$\begin{aligned} P(\mathbf{y}^\dagger \mathbf{U} \mathbf{y} \geq 0 \mid \mathbf{g}) &= \frac{1}{2\pi} \int \int_{-\infty}^{\infty} \frac{e^{i\omega \mathbf{y}^\dagger \mathbf{U} \mathbf{y}}}{i\omega} \frac{e^{-\mathbf{y}^\dagger \Sigma^{-1} \mathbf{y}}}{(2\pi)^{2T} \det(\Sigma)} d\omega d\mathbf{y} \\ &= \frac{1}{(2\pi)^{2T+1}i} \int_{-\infty}^{\infty} \int \frac{e^{-\mathbf{y}^\dagger (-i\omega \mathbf{U} + \Sigma^{-1}) \mathbf{y}}}{\omega \det(\Sigma)} d\mathbf{y} d\omega. \end{aligned}$$

Since the exponent of the exponential is of the form $i\mathbf{y}^\dagger \omega \mathbf{U} \mathbf{y} - \mathbf{y}^\dagger \Sigma^{-1} \mathbf{y}$, with real part $-\mathbf{y}^\dagger \Sigma^{-1} \mathbf{y}$ which is negative, and imaginary part given by $i\mathbf{y}^\dagger \omega \mathbf{U} \mathbf{y}$ (recall that $\mathbf{U}$ is Hermitian), then this integral converges and we have

$$P(\mathbf{y}^\dagger \mathbf{U} \mathbf{y} \geq 0 \mid \mathbf{g}) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{1}{\omega \det(\mathbf{I} - i\omega \mathbf{U} \Sigma)} d\omega.$$

We have that the above determinant is given by

$$\prod_{k=1}^T (1 + i\omega c_\rho |g_k|^2 |u_{ik} - u_{jk}|^2 + \omega^2 |u_{ik} - u_{jk}|^2 (2ac_\rho |g_k|^2 + a^2)),$$

where $a = c'_\rho \|\mathbf{g}\|^2 + 1$. Since our goal is a diversity computation, we are interested in an very high SNR regime. Note that when ρ is big, $c'_\rho = \frac{P_2}{P_1+1} = \frac{\rho}{T(\rho+2)} \rightarrow 1/T$. Thus $a \rightarrow \|\mathbf{g}\|^2/T + 1$ and the term in a^2 does not depend on ρ , so that we have

$$\det(\mathbf{I} - i\omega \mathbf{U} \Sigma) \approx \prod_{k=1}^T c_\rho |g_k|^2 |u_{ik} - u_{jk}|^2 \left(\frac{1}{c_\rho |g_k|^2 |u_{ik} - u_{jk}|^2} + i\omega + \omega^2 2a \right).$$

By completing the squares, we get that $P(\mathbf{y}^\dagger \mathbf{U} \mathbf{y} \geq 0 | \mathbf{g})$ is given by

$$\frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{1}{\omega} \left(\prod_{k=1}^T c_\rho |g_k|^2 |u_{ik} - u_{jk}|^2 2a \left[\left(\omega + \frac{i}{4a} \right)^2 + c_k^2 \right] \right)^{-1} d\omega$$

where

$$c_k := \sqrt{\frac{1}{16a^2} + \frac{1}{c_\rho |g_k|^2 |u_{ik} - u_{jk}|^2}}.$$

Note that the above integral has poles in $\omega = -i(1/4a \pm c_k)$. Thus as long as $-i(1/4a - c_k) < \text{Im}(\omega) < -i(1/4a + c_k)$, the above integral is well-defined. We thus choose the following contour of integration, within the convergence region

$$\frac{1}{2\pi i} \int_{-\infty - \frac{i}{4a}}^{\infty - \frac{i}{4a}} \frac{1}{\omega} \left(\prod_{k=1}^T c_\rho |g_k|^2 |u_{ik} - u_{jk}|^2 2a \left[\left(\omega + \frac{i}{4a} \right)^2 + c_k^2 \right] \right)^{-1} d\omega$$

and with a change of variable, we get

$$\frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{1}{\omega - \frac{i}{4a}} \left(\prod_{k=1}^T c_\rho |g_k|^2 |u_{ik} - u_{jk}|^2 2a \left[\omega^2 + \frac{1}{16a^2} + \frac{1}{c_\rho |g_k|^2 |u_{ik} - u_{jk}|^2} \right] \right)^{-1} d\omega$$

Following [6], we obtain a bound on the probability of error that we know real by taking the real part of the above expression:

$$\begin{aligned} P(\mathbf{y}^\dagger \mathbf{U} \mathbf{y} \geq 0 | \mathbf{g}) &\lesssim \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{4a}{1 + 16a^2 \omega^2} \left(\prod_{k=1}^T c_\rho |g_k|^2 |u_{ik} - u_{jk}|^2 2a \left[\omega^2 + \frac{1}{16a^2} \right] + 1 \right)^{-1} d\omega \\ &\leq \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{4ad\omega}{1 + 16a^2 \omega^2} \left(\prod_{k=1}^T c_\rho |g_k|^2 |u_{ik} - u_{jk}|^2 \frac{1}{8a} + 1 \right)^{-1} \\ &= \frac{1}{2} \left(\prod_{k=1}^T c_\rho |g_k|^2 |u_{ik} - u_{jk}|^2 \frac{1}{8a} + 1 \right)^{-1} \\ &= \frac{1}{2} \det(\mathbf{I}_T + \frac{c_\rho}{8(\frac{1}{T}\|\mathbf{g}\|^2 + 1)} D_{|g|} (U_k - U_l)(U_k - U_l)^\dagger)^{-1}. \end{aligned}$$

■

Thus

$$P(U_k \rightarrow U_l) = E_{\mathbf{g}} P(\mathbf{y}^\dagger \mathbf{U} \mathbf{y} \geq 0 \mid \mathbf{g}) \leq E_{\mathbf{g}} \det(\mathbf{I}_T + \frac{c_p}{8(\frac{1}{T}\|\mathbf{g}\|^2 + 1)} D_{|\mathbf{g}|} (U_k - U_l)(U_k - U_l)^\dagger)^{-1}.$$

This bound on the pairwise probability of error is similar to the Chernoff bound obtained in [10, Theorem 1], where it has been proven, by computing the above expectation on $\mathbf{g} = (g_1, \dots, g_T)$, that the diversity gain is given by $\text{rank}((U_k - U_l)(U_k - U_l)^\dagger) \left(1 - \frac{\log \log P}{\log P}\right)$. Thus when $(U_k - U_l)(U_k - U_l)^\dagger$ is full rank (that is, the code is *fully diverse*), we get a diversity of

$$R \left(1 - \frac{\log \log P}{\log P}\right). \quad (13)$$

C. Analysis of node failures

In our network setting, relay nodes are small devices with few power. It is thus very likely that some of them may run out of battery, or may encounter a technical problem that will prevent them to communicate, at a given point of the transmission. In this part, we are interested in understanding how a coding strategy designed for R relay nodes will behave if d of them are down.

Let us consider one step of the differential scheme, thus involving two transmissions, say $\mathbf{y}(t)$ and $\mathbf{y}(t + T)$. Let us assume that $\mathbf{y}(t)$ is transmitted correctly, but one or several node failures happen while transmitting $\mathbf{y}(t + T)$. Denote by $\mathcal{D}$ the set of indices of relay nodes which are down during the transmission of $\mathbf{y}(t + T)$. In Equation (9) for $\mathbf{y}(t + T)$ we thus have

$$\sum_{i=1}^R g_i f_i A_i U(z_{t+T}) \mathbf{s}_t = \sum_{i \notin \mathcal{D}} g_i f_i A_i U(z_{t+T}) \mathbf{s}_t.$$

In Equation (8) for $\mathbf{y}(t)$, we can similarly write

$$\sum_{i=1}^R g_i f_i A_i \mathbf{s}_t = \sum_{i \notin \mathcal{D}} g_i f_i A_i \mathbf{s}_t + \sum_{i \in \mathcal{D}} g_i f_i A_i \mathbf{s}_t,$$

so that when we plug Equation (8) inside Equation (9), the term $\sum_{i \in \mathcal{D}} g_i f_i A_i \mathbf{s}_t$ adds to the noise, and the codeword $U_{z_{t+T}}$ will not be recovered.

However, during the next step, things are different. Consider now the step involving $\mathbf{y}(t + T)$ and $\mathbf{y}(t + 2T)$. Now, for both signals, the same set of relay nodes will be down. Thus communication will take place over a network where actually $R - |\mathcal{D}|$ are transmitting. The expression for $\mathbf{y}(t + T)$ and $\mathbf{y}(t + 2T)$ are similar to those obtained in Equations (8) and (9), except that the sums are over $i \in R \setminus \mathcal{D}$

R	L	rate	$\zeta(U)$	u
3	8	1	0.5134	$u = (1, 1, 3)$
3	63	1.99	0.3301	$u = (1, 17, 26)$
6	64	1	0.3792	$u = (1, 7, 15, 23, 25, 31)$
6	4096	2	0.1428	$u = (1, 599, 623, 1445, 1527, 1715)$
9	57	0.65	0.361	$u = (1, 4, 16, 7, 28, 55, 49, 25, 43)$

TABLE I
SOME CYCLIC CODES WITH GOOD DIVERSITY PRODUCT.

instead of $i = 1, \dots, R$. We thus expect the diversity of such systems to be

$$(R - d) \left(1 - \frac{\log \log P}{\log P} \right), \quad (14)$$

where $d = |\mathcal{D}|$. This is a positive result, since it says that the strategy is resistant to node failures.

V. CODE CONSTRUCTIONS AND SIMULATION RESULTS

In this section, we first consider code constructions, and show that though any random diagonal unitary matrix could be used to generate a codebook, codes designed to optimize the diversity product reach a much better coding gain. We then provide simulation results, also including the node failures scenario. We finally discuss the decoding issue.

A. Cyclic codes

Since the matrices A_i used at the relays are fixed for a given number of relays R , designing codes consists of constructing the diagonal unitary matrices $U(z)$ in which the data z to be sent will be encoded. Regarding the PEP computation (13), the matrices $U(z)$ have to satisfy the full diversity criterion:

$$\det(U(z_t) - U(z_{t'})) \neq 0, \quad t \neq t'.$$

A priori, any diagonal unitary matrix could be used to build the code, and we furthermore expect that such matrix generated randomly would yield full diversity. This makes our scheme attractive since it is thus available easily for any number of relay nodes R .

In order now to reach a good coding gain, one needs, similarly to the point to point case, to maximize the diversity product, given by

$$\zeta(U) = \min_{t \neq t'} \frac{1}{2} |\det(U(z_t) - U(z_{t'}))|^{1/R}.$$

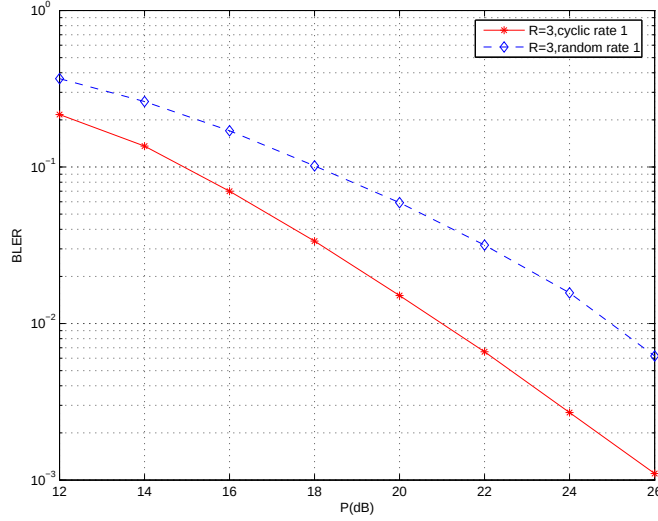


Fig. 2. A random rate 1 code versus a cyclic rate 1 code.

Diagonal unitary codes achieving good diversity product have already been studied in [7] for the MIMO case. They are called *cyclic codes*. Denote by $\zeta_L = \exp(2i\pi/L)$ a primitive L th root of unity. A cyclic code is parametrized as follows:

$$\begin{pmatrix} \zeta_L^{u_1 l} & & 0 \\ 0 & \ddots & 0 \\ 0 & & \zeta_L^{u_M l} \end{pmatrix}, \quad l = 0, \dots, L-1,$$

where L and $u = (u_1, \dots, u_M)$ have to be designed. Note that M denotes the number of antennas in the point to point case, or the number of relays R in the distributed case. Let us give an example with three relay nodes. Let $\zeta_{63} = \exp(2i\pi/63)$. The codebook is given by

$$\left\{ D^i, D = \begin{pmatrix} \zeta_{63} & 0 & 0 \\ 0 & \zeta_{63}^{17} & 0 \\ 0 & 0 & \zeta_{63}^{26} \end{pmatrix} \right\}$$

for $i = 1, \dots, 63$. Thus $L = 63$ and $u = (1, 17, 26)$.

Table I summarizes some good cyclic codes. The codes for 3 and 6 relays have been found in [7]. However, no cyclic code in bigger dimension is available, since there is no need for more than 6 antennas in a point to point setting. The code for 9 relays is the diagonal component of a fixed point free group based code [22].

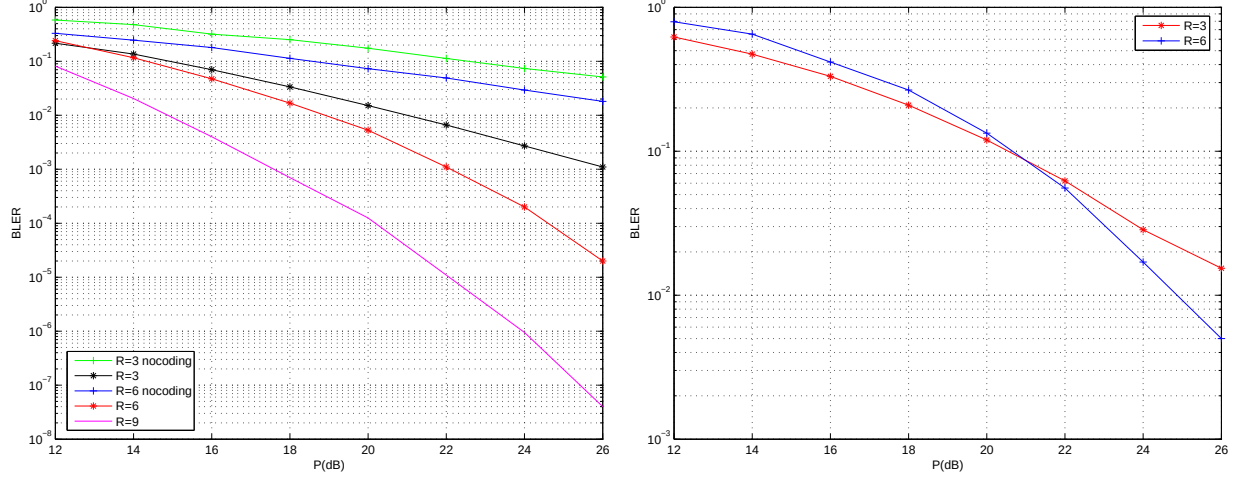


Fig. 3. On the left, rate 1 cyclic codes: with and without coding at the relays. On the right, rate 2 cyclic codes.

In Fig. 2, we compare a random code with a cyclic code. We plot the block error rate (BLER) as a function of the power of the system in dBs. We consider the cyclic code $u = (1, 1, 3)$ with rate 1, with a random code of rate 1. The random code consists of generating one random unitary diagonal matrix and take its powers from 1 to 8. The relays R_i , $i = 1, 2, 3$, use respectively the matrices

$$A_1 = \mathbf{I}_3, A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \zeta_3 & 0 \\ 0 & 0 & \zeta_3^2 \end{pmatrix}, A_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \zeta_3^2 & 0 \\ 0 & 0 & \zeta_3 \end{pmatrix}.$$

Both codes are fully diverse, but clearly, the cyclic code with high diversity product reaches a much better coding gain.

On the one hand, there is thus a real coding gain that could be achieved by choosing carefully the diagonal matrices. On the other hand, since random unitary diagonal matrices yield full diversity, the differential coding scheme we propose is available for any number of relay nodes.

B. Illustration of the diversity

Simulation results in Fig. 3 show the BLER as a function of P , the power of the system in dBs. To start with, Fig. 3 on the left first show what we already know: there is really a need for coding at the relays, without which we cannot get diversity. We see clearly that without coding at the relays, having 3 or 6 relays in the system does not change the curve and thus give no diversity.

Recall that the analysis of the PEP in (13) predicts a diversity which is linear in the number of relays.

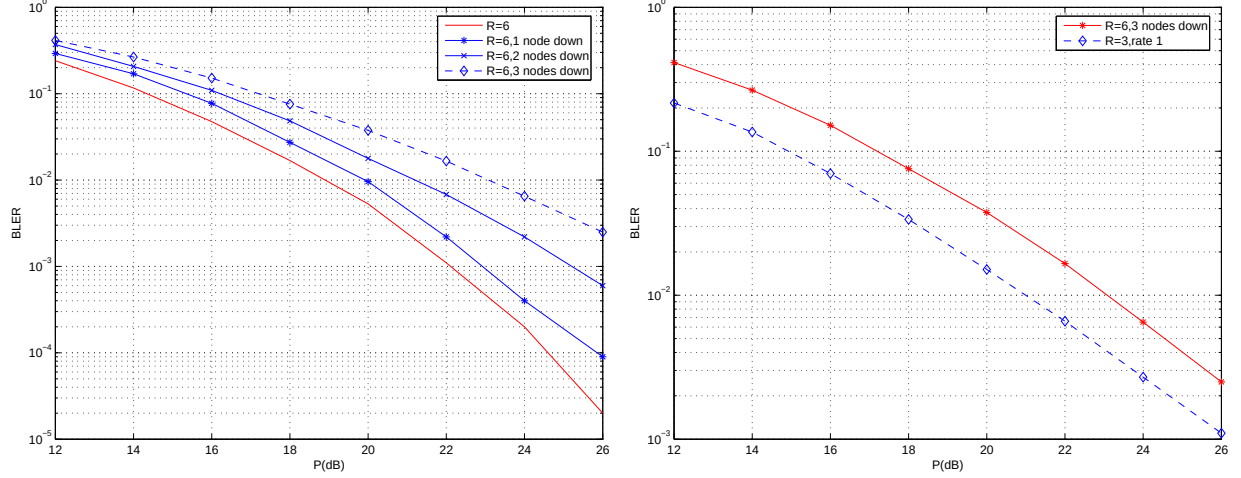


Fig. 4. On the left, a network of 6 relay nodes with up to 3 nodes down, on the right, a 6 nodes network with 3 nodes down versus a 3 nodes network.

Fig. 3 on the left further shows how the diversity is increasing, by going from 3 to 6, and then to 9 relays. The codes used are the two cyclic rate 1 codes given in Table I for resp. 3 and 6 relays, and the cyclic code for 9 relays with rate 0.65.

Fig. 3 on the right shows simulation results for rate 2 cyclic codes, with resp. 3 and 6 relay nodes. Clearly, increasing the rate decreases the performance, and in particular, the slope given by the diversity starts at high power. Note that this crossing of the two curves was already observed in the point to point case [7], but at a lower SNR.

C. Resistance to node failures

We finally give simulation results to illustrate the behavior in case of node failures. The analysis done in Section IV shows that as far as diversity is concerned, a network of R relay nodes having d relays not communicating behaves as a network with $R - d$ nodes. This behavior can be observed in Fig. 4. On the left, the performance of a network with 6 relay nodes is shown. Simulations are then done, assuming that 1, then 2, and finally 3 relay nodes are down. We see how the curves change linearly with the number of relay nodes.

Fig. 4 on the right compares a network with 6 relay nodes, but only 3 of them communicating, with a network having only 3 nodes. We observe the same diversity, which was predicted by the analysis in (14). However, there is a clear loss in coding gain, which is expected.

The simulation results of this section thus confirm a good behavior of the proposed strategy when

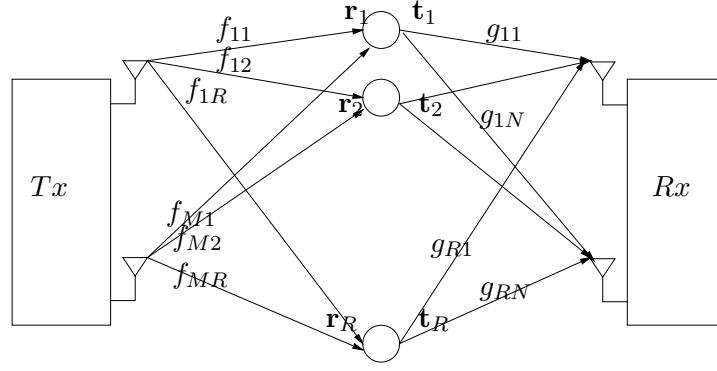


Fig. 5. A wireless network with multiple antenna nodes.

facing node failures.

D. Decoding issues

We finally briefly discuss decoding issues. Simulations shown in this paper have been done via exhaustive search. Indeed, for such rates, there is no real complexity problem in doing an exhaustive search. However, when increasing the number of relays or the rate, one may need to have a faster algorithm. Recall that the mismatched decoder (11) is given by

$$\arg \min_{U_l} \min_{l=1, \dots, L} \|\mathbf{y}(t+T) - U_l \mathbf{y}(t)\|^2 = \arg \min_{l=1, \dots, L} \sum_{k=1}^R |\mathbf{y}(t+T)_k - \zeta_L^{u_k l} \mathbf{y}(t)_k|^2.$$

This minimization problem has already been studied in [3], where the authors show how the problem can be expressed as a lattice reduction problem. They give an algorithm to solve it, which is faster than exhaustive search.

VI. GENERALIZATION TO MULTIPLE ANTENNAS CASE

In [11], a generalization of the network presented in Subsection II-A was given. In order to increase the data rate of the network, both the transmitter and receiver nodes are equipped with multiple antennas (see Fig. 5). Let M and N be the number of transmit resp. receive antennas.

The codeword $\mathbf{S}$ sent at the transmitter is normalized such that $E[\text{Tr}(\mathbf{S}\mathbf{S}^\dagger)] = M$. Channels are denoted by f_{mj} from the m th transmit antenna to the j th relay and g_{jn} from the j th relay to the n th receive antenna (see Fig. 5), which are assumed to be iid $\mathcal{CN}(0, 1)$. Similarly, let $\mathbf{v}_i$ and $\mathbf{W}_i$ be the noise vectors, with coefficients also iid $\mathcal{CN}(0, 1)$. Matrices at the relays are denoted by A_i , $i = 1, \dots, R$, and need to be unitary, as before. The power at the transmitter for each transmission is P_1 , thus P_1/M for each

antenna. Each relay has a power of P_2/R for each transmission. By defining $Y = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]$, $S = [A_1\mathbf{S} \ A_2\mathbf{S} \ \dots \ A_R\mathbf{S}]$, $\mathbf{g}_i = [g_{i1}, g_{i2}, \dots, g_{iN}]$, $H = [\mathbf{f}_1\mathbf{g}_1, \dots, \mathbf{f}_R\mathbf{g}_R]^T$, and

$$W = \left[\sqrt{\frac{P_2}{P_1+1}} \sum_{i=1}^R g_{i1} A_i \mathbf{v}_i + \mathbf{w}_1, \dots, \sqrt{\frac{P_2}{P_1+1}} \sum_{i=1}^R g_{iN} A_i \mathbf{v}_i + \mathbf{w}_N \right],$$

it was shown in [11] that the channel model can be written as

$$Y = \sqrt{\frac{P_1 P_2 T}{(P_1 + 1)M}} SH + W. \quad (15)$$

The received matrix Y is a $T \times N$ matrix, while S is a $T \times MR$ matrix, since the A_i , $i = 1, \dots, R$, are $T \times T$ matrices, and $\mathbf{S}$ is a $T \times M$ matrix.

We now show how the noncoherent scheme presented previously generalizes to the multiple antenna scenario.

A. The noncoherent channel

Similarly to what we have done in Subsection II-B, we now need to design the matrices A_i , $i = 1, \dots, R$ and the codebook such that $S = [A_1\mathbf{S}, \dots, A_R\mathbf{S}]$ is unitary, i.e., $S^\dagger S = \mathbf{I}_{MR}$.

First, we choose the codewords to be of the form

$$\mathbf{S}_j = U_j \mathbf{S}', \quad j = 1, \dots, L$$

where U_j is a unitary $T \times T$ matrix, and $\mathbf{S}'$ is an initial transmitted signal normalized to that $E[\text{Tr}(\mathbf{S}_j^\dagger \mathbf{S}_j)] = M$ for all $j = 1, \dots, L$. As for the one antenna case, we require that both the unitary matrices A_i , $i = 1, \dots, R$, and U_j , $j = 1, \dots, L$ be diagonal. Thus, denoting by $S_j = [A_1\mathbf{S}_j, \dots, A_R\mathbf{S}_j]$, we have

$$\begin{aligned} S_j^\dagger S_j &= [A_1 U_j \mathbf{S}', \dots, A_R U_j \mathbf{S}']^\dagger [A_1 U_j \mathbf{S}', \dots, A_R U_j \mathbf{S}'] \\ &= [A_1 \mathbf{S}', \dots, A_R \mathbf{S}']^\dagger U_j^\dagger U_j [A_1 \mathbf{S}', \dots, A_R \mathbf{S}'] \\ &= [A_1 \mathbf{S}', \dots, A_R \mathbf{S}']^\dagger [A_1 \mathbf{S}', \dots, A_R \mathbf{S}'] \end{aligned}$$

and the fact that S_j is unitary only depends on the choice of the matrices at the relays and the transmitted signal $\mathbf{S}'$. We thus design the matrices A_i and $\mathbf{S}'$ independently of the codebook of matrices U_j .

Suppose now that we have $T = MR$, with M the number of transmit antennas, R the number of relays, and T the coherence time. Choose a first generalized Butson-Hadamard (GBM) matrix A of size $R \times R$, and a second one B of size M (c.f. Definition 1), both with coefficients that are roots of unity.

Take

$$A_i = \text{diag}(\underbrace{a_{1i}, \dots, a_{1i}}_M, \dots, \underbrace{a_{Ri}, \dots, a_{Ri}}_M),$$

which is unitary for all i since all a_{ij} 's are roots of unity, and similarly

$$\mathbf{S}' = \frac{\sqrt{M}}{\|\mathbf{S}'\|_F} \begin{pmatrix} b_{1,1} & \dots & b_{1,M} \\ \vdots & & \vdots \\ b_{M,1} & & b_{M,M} \\ \vdots & & \vdots \\ b_{1,1} & \dots & b_{1,M} \\ \vdots & & \vdots \\ b_{M,1} & & b_{M,M} \end{pmatrix},$$

where $\frac{\sqrt{M}}{\|\mathbf{S}'\|_F}$ is the normalization factor such that $E[\text{Tr}(\mathbf{S}'^\dagger \mathbf{S}')] = M$. By definition of the Frobenius norm $\|\cdot\|_F$,

$$\|\mathbf{S}'\|_F = \sqrt{\sum_{i=1}^M \sum_{j=1}^T |s'_{i,j}|^2} = \sqrt{MT},$$

since all coefficients $s'_{i,j}$ are roots of unity. By construction

$$[A_1 \mathbf{S}', \dots, A_R \mathbf{S}'] = \frac{1}{\sqrt{T}} A \otimes B,$$

and

$$[A_1 \mathbf{S}', \dots, A_R \mathbf{S}']^\dagger [A_1 \mathbf{S}', \dots, A_R \mathbf{S}'] = \frac{RM}{T} \mathbf{I}_{RM} = \mathbf{I}_{RM},$$

since $A \otimes B$ is a GBH matrix of size RM .

Example 3: Let $M = 2$ and $R = 3$. Choose the following matrices:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & \zeta_3 & \zeta_3^2 \\ 1 & \zeta_3^2 & \zeta_3 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

The unitary matrices at the relays are given by

$$A_1 = \mathbf{I}_6, \quad A_2 = \text{diag}(1, 1, \zeta_3, \zeta_3, \zeta_3^2, \zeta_3^2), \quad A_3 = \text{diag}(1, 1, \zeta_3^2, \zeta_3^2, \zeta_3, \zeta_3),$$

where $\mathbf{I}_6$ denote the identity matrix of size 6, and the initial transmitted signal is

$$\mathbf{S}' = \sqrt{\frac{2}{12}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

B. Decoding and Diversity

A differential encoder is done similarly to the one antenna case, so that a natural candidate for the differential decoder is

$$\arg \min_{U_l, l=1, \dots, L} \|Y(t+T) - U_l Y(t)\|^2.$$

Let us restrict to the case where $T = MR$. Similarly to the single antenna case, in order to analyze this strategy, we consider two instances of the noncoherent channel,

$$Y(t) = \sqrt{c_\rho} S(t) H + W(t)$$

where H is an $T \times N$ matrix unknown at both the transmitter and receiver, $S(t)$ is a $T \times T$ unitary matrix, and c_ρ is a constant which depends on the SNR ρ , that is,

$$\begin{pmatrix} Y(t) \\ Y(t+T) \end{pmatrix} = \sqrt{c_\rho} \begin{pmatrix} S(t) \\ S(t+T) \end{pmatrix} H + \begin{pmatrix} W(t) \\ W(t+T) \end{pmatrix}.$$

Since $S(t)$ and $S(t)\psi$ are indistinguishable for an arbitrary unitary $T \times T$ matrix ψ , we preprocess the signal so that

$$\begin{pmatrix} Y(t) \\ Y(t+T) \end{pmatrix} = \sqrt{c_\rho} \begin{pmatrix} \mathbf{I}_T \\ U_k \end{pmatrix} H + \begin{pmatrix} W(t) \\ W(t+T) \end{pmatrix},$$

for U_k a unitary matrix belonging to the codebook. To suit the network model, we have

$$W(t) = \left[\sqrt{\frac{P_2}{P_1+1}} \sum_{i=1}^R g_{i1} A_i \mathbf{v}_i(t) + \mathbf{w}_1(t), \dots, \sqrt{\frac{P_2}{P_1+1}} \sum_{i=1}^R g_{iN} A_i \mathbf{v}_i(t) + \mathbf{w}_N(t) \right],$$

$H = [\mathbf{f}_1 \mathbf{g}_1, \dots, \mathbf{f}_R \mathbf{g}_R]^T$, with $\mathbf{f}_i = [f_{1i}, \dots, f_{Mi}]^T$ and $\mathbf{g}_i = [g_{i1}, g_{i2}, \dots, g_{iN}]$. Furthermore, we have $c_\rho = \frac{P_2 P_1 T}{(P_1+1)M}$, and we denote $c'_\rho = \frac{P_2}{P_1+1}$. Recall that $P_1 = \rho/2$, $P_2 = \rho/(2R)$ and ρ is the total power of the system. Since the columns of the matrix H are *dependent*, we cannot treat the channel equation in

matrix form and have to vectorize it to do the analysis. We denote by $\text{vec}(Y(t))$ the $TN \times 1$ vectorized version of $Y(t)$. Set $\mathbf{Y} = [\text{vec}(Y(t)), \text{vec}(Y(t+T))]^t$. The pairwise probability of error is thus given by

$$\begin{aligned}
& P(U_k \rightarrow U_l) \\
&= P(\|\text{vec}(Y(t+T)) - (U_k \otimes \mathbf{I}_N)\text{vec}(Y(t))\|^2 \geq \|\text{vec}(Y(t+1)) - (U_l \otimes \mathbf{I}_N)\text{vec}(Y(t))\|^2 \mid U_k \text{ sent}) \\
&= P(\|[-(U_k \otimes \mathbf{I}_N) \quad \mathbf{I}_{TN}] \mathbf{Y}\|^2 \geq \|[-(U_l \otimes \mathbf{I}_N) \quad \mathbf{I}_{TN}] \mathbf{Y}\|^2 \mid U_k \text{ is sent}) \\
&= P(\mathbf{Y}^\dagger \mathbf{\Delta} \mathbf{Y} \geq 0 \mid U_k \text{ is sent}) \\
&= E_{\mathbf{g}} P(\mathbf{Y}^\dagger \mathbf{\Delta} \mathbf{Y} \geq 0 \mid \mathbf{g})
\end{aligned}$$

with $k \neq l$, and

$$\mathbf{\Delta} = \begin{pmatrix} 0 & (U_l^\dagger - U_k^\dagger) \otimes \mathbf{I}_N \\ (U_l - U_k) \otimes \mathbf{I}_N & 0 \end{pmatrix}.$$

We now use the same approach as in the single antenna case and compute $P(\mathbf{Y}^\dagger \mathbf{\Delta} \mathbf{Y} \geq 0)$ knowing $\mathbf{g}$.

Proposition 2: We have that

$$P(\mathbf{Y}^\dagger \mathbf{\Delta} \mathbf{Y} \geq 0 \mid \mathbf{g}) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{1}{\omega \det(\mathbf{I}_{2TN} - i\omega \mathbf{\Delta} \Sigma)} d\omega,$$

where Σ is defined by (16).

Proof: We have

$$\begin{aligned}
P(\mathbf{Y}^\dagger \mathbf{\Delta} \mathbf{Y} \geq 0) &= E[u(\mathbf{Y}^\dagger \mathbf{\Delta} \mathbf{Y})] \\
&= E \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\mathbf{Y}^\dagger \mathbf{\Delta} \mathbf{Y}}}{i\omega} d\omega \right]
\end{aligned}$$

where u is the step function ($u(x) = 1$ if $x > 0$, $u(x) = 0$ else), and the second equality is the Fourier transform of u . Computing the expectation yields

$$\begin{aligned}
P(\mathbf{Y}^\dagger \mathbf{\Delta} \mathbf{Y} \geq 0 \mid \mathbf{g}) &= \frac{1}{2\pi} \int \int_{-\infty}^{\infty} \frac{e^{i\mathbf{Y}^\dagger \mathbf{\Delta} \mathbf{Y}}}{i\omega} \frac{e^{-\mathbf{Y}^\dagger \Sigma^{-1} \mathbf{Y}}}{(2\pi)^{2T} \det(\Sigma)} d\omega d\mathbf{Y} \\
&= \frac{1}{(2\pi)^{2TN+1} i} \int_{-\infty}^{\infty} \int \frac{e^{-\mathbf{Y}^{vec\dagger} (-i\omega \mathbf{\Delta} + \Sigma^{-1}) \mathbf{Y}^{vec}}}{\omega \det(\Sigma)} d\mathbf{Y} d\omega,
\end{aligned}$$

where $\Sigma = E[\mathbf{Y} \mathbf{Y}^\dagger]$ is the covariance matrix of $\mathbf{Y}$. Since the matrix H can be written

$$H = [\mathcal{G}_1 \mathbf{f}, \dots, \mathcal{G}_N \mathbf{f}]$$

with $\mathbf{f} = [\mathbf{f}_1, \dots, \mathbf{f}_R]^t$ and $\mathcal{G}_n = \text{diag}(g_{1n}\mathbf{I}_M, \dots, g_{Rn}\mathbf{I}_M)$, we have that

$$\Sigma = \begin{pmatrix} (c'_\rho \sum_{i=1}^R |g_{i1}|^2 + 1)\mathbf{I}_T & & \\ & \ddots & \\ & & (c'_\rho \sum_{i=1}^R |g_{iN}|^2 + 1)\mathbf{I}_T \end{pmatrix} \otimes \mathbf{I}_2 + c_\rho \begin{pmatrix} \mathcal{G} \\ \mathcal{G}U_k \end{pmatrix} \begin{pmatrix} \mathcal{G}^\dagger & \mathcal{G}^\dagger U_k^\dagger \end{pmatrix} \quad (16)$$

with

$$\mathcal{G} = \begin{pmatrix} \mathcal{G}_1 \\ \vdots \\ \mathcal{G}_N \end{pmatrix}.$$

The imaginary part of $-\mathbf{Y}^\dagger(-i\omega\mathbf{\Delta} + \Sigma^{-1})\mathbf{Y}$ is $i\mathbf{Y}^\dagger\omega\mathbf{\Delta}\mathbf{Y}$ since $\omega\mathbf{\Delta}$ is Hermitian, while Σ^{-1} is positive definite. Thus the above integral converges, and we have the desired result. $\blacksquare$

We now derive the diversity in the asymptotic regime where we assume that both the power of the system ρ and the number of relays R are big. Recall that

$$\begin{aligned} P(U_k \rightarrow U_l) &= E_{\mathbf{g}} P(\mathbf{Y}^\dagger \mathbf{\Delta} \mathbf{Y} \geq 0 \mid \mathbf{g}) \\ &= E_{\mathbf{g}|\mathbf{g} \in \mathcal{A}} P(\mathbf{Y}^\dagger \mathbf{\Delta} \mathbf{Y} \geq 0) + E_{\mathbf{g}|\mathbf{g} \notin \mathcal{A}} P(\mathbf{Y}^\dagger \mathbf{\Delta} \mathbf{Y} \geq 0), \end{aligned}$$

where

$$\mathcal{A} = \{\mathbf{g} \mid |(1/R) \sum_{i=1}^R |g_{ij}|^2 - 1| < \epsilon \text{ for all } j\}.$$

When $R \rightarrow \infty$, we now show that $P(\mathbf{g} \in \mathcal{A}) \rightarrow 1$ for all $\epsilon > 0$, which will thus imply that

$$E_{\mathbf{g}|\mathbf{g} \notin \mathcal{A}} P(\mathbf{Y}^\dagger \mathbf{\Delta} \mathbf{Y} \geq 0) \rightarrow 0.$$

For any $j = 1, \dots, N$, we have

$$P(|(1/R) \sum_{i=1}^R |g_{ij}|^2 - 1| < \epsilon) = 1 - P(|(1/R) \sum_{i=1}^R |g_{ij}|^2 - 1| \geq \epsilon).$$

Since $\gamma := \sum_{i=1}^R |g_{ij}|^2$ has a gamma distribution, we have that

$$P(\gamma \geq R + \epsilon R) = \int_{R(1+\epsilon)}^{\infty} \frac{\gamma^{R-1} e^{-\gamma}}{(R-1)!} d\gamma,$$

which can be bounded, using a Chernoff bound, by

$$P(\gamma \geq R + \epsilon R) \leq E[e^{\alpha(\gamma - R(1+\epsilon))}], \quad \alpha > 0$$

$$\begin{aligned}
&= e^{-\alpha R(1+\epsilon)} \int_0^\infty \frac{\gamma^{R-1} e^{-\gamma(1-\alpha)}}{(R-1)!} d\gamma, \quad 0 < \alpha < 1 \\
&= e^{-\alpha R(1+\epsilon)} \int_0^\infty \left(\frac{u}{1-\alpha} \right)^{R-1} \frac{e^{-u}}{(R-1)!} \frac{du}{1-\alpha}, \quad u = (1-\alpha)\gamma \\
&= \left(\frac{e^{-\alpha(1+\epsilon)}}{1-\alpha} \right)^R.
\end{aligned}$$

By computing the derivative of the above expression, we find that the optimal α is given by $1 - \alpha = 1/(1 + \epsilon)$, that is, $\alpha = \epsilon/(1 + \epsilon)$, so that

$$1 - P(\gamma \geq R + \epsilon R) \geq 1 - ((1 + \epsilon)e^{-\epsilon})^R.$$

Since $(1 + \epsilon)e^{-\epsilon} < 1$ for $\epsilon > 0$, we have that $(1 + \epsilon)e^{-\epsilon} \rightarrow 0$ with an exponential decay.

Proposition 3: We have that

$$P(\mathbf{Y}^\dagger \mathbf{\Delta} \mathbf{Y} \geq 0 \mid \mathbf{g} \in \mathcal{A}) \leq \frac{1}{2} \det \left(\mathbf{I}_T + \frac{c_\rho}{16} \sum_{j=1}^N \mathcal{G}_j \mathcal{G}_j^\dagger (U_k - U_l)(U_k - U_l)^\dagger \right)^{-1}$$

Proof: We have that $c'_\rho \sum_{i=1}^R |g_{ij}|^2 + 1 \approx 2$ since $c'_\rho = \rho/(R(\rho + 2))$ and we are integrating over $\mathcal{A}$, which yields

$$\Sigma \approx 2\mathbf{I}_{2TN} + c_\rho \begin{pmatrix} \mathcal{G} \\ \mathcal{G}U_k \end{pmatrix} \begin{pmatrix} \mathcal{G}^\dagger & \mathcal{G}^\dagger U_k^\dagger \end{pmatrix},$$

and what we need to compute is

$$\det \begin{pmatrix} \mathbf{I}_{TN} - i\omega c_\rho \mathcal{G} \Delta^\dagger U_k \mathcal{G}^\dagger & -i\omega [2\mathbf{\Delta}^\dagger + c_\rho \mathcal{G} \Delta^\dagger \mathcal{G}^\dagger] \\ -i\omega [2\mathbf{\Delta} + c_\rho \mathcal{G} \Delta \mathcal{G}^\dagger] & \mathbf{I}_{TN} - i\omega c_\rho \mathcal{G} \Delta U_k^\dagger \mathcal{G}^\dagger \end{pmatrix}$$

where $\Delta = U^l - U^k$. Since $\mathbf{I}_{TN} - i\omega c_\rho \mathcal{G} \Delta U_k^\dagger \mathcal{G}^\dagger$ and $-i\omega [2\mathbf{\Delta} + c_\rho \mathcal{G} \Delta \mathcal{G}^\dagger]$ commute, we have that this determinant is given by

$$\det \left(\mathbf{I}_{TN} + c_\rho \mathcal{G} [i\omega \Delta \Delta^\dagger + \frac{4\omega^2 \mathbf{\Delta}^\dagger \mathbf{\Delta}}{c_\rho} + 2\omega^2 2\Delta \Delta^\dagger] \mathcal{G}^\dagger \right).$$

Since $\rho \rightarrow \infty$, we have that $\frac{4\omega^2 \mathbf{\Delta}^\dagger \mathbf{\Delta}}{c_\rho} \rightarrow 0$, and the determinant simplifies to

$$\begin{aligned}
\det \left(\mathbf{I}_{TN} + c_\rho \mathcal{G} \Delta \Delta^\dagger [i\omega + 4\omega^2] \mathcal{G}^\dagger \right) &= \det \left(\mathbf{I}_T + c_\rho \mathcal{G} \Delta \Delta^\dagger [i\omega + 4\omega^2] \sum_{j=1}^N \mathcal{G}_j^\dagger \mathcal{G}_j \right) \\
&= \prod_{k=1}^T \left(1 + c_\rho \Delta_k \Delta_k^\dagger [i\omega + 4\omega^2] \sum_{j=1}^N \mathcal{G}_{jk}^\dagger \mathcal{G}_{jk} \right)
\end{aligned}$$

$$\begin{aligned}
&= \prod_{k=1}^T 4c_\rho \Delta_k \Delta_k^\dagger \sum_{j=1}^N \mathcal{G}_{jk}^\dagger \mathcal{G}_{jk} \left(\frac{1}{4c_\rho \Delta_k \Delta_k^\dagger \sum_{j=1}^N \mathcal{G}_{jk}^\dagger \mathcal{G}_{jk}} + \frac{i\omega}{4} + \omega^2 \right) \\
&= \prod_{k=1}^T 4c_\rho \Delta_k \Delta_k^\dagger \sum_{j=1}^N \mathcal{G}_{jk}^\dagger \mathcal{G}_{jk} \left(\left(\omega + \frac{i}{8} \right)^2 + c_k^2 \right)
\end{aligned}$$

where

$$c_k = \sqrt{\frac{1}{64} + \frac{1}{4c_\rho \Delta_k \Delta_k^\dagger \sum_{j=1}^N \mathcal{G}_{jk}^\dagger \mathcal{G}_{jk}}}.$$

From now on, we can follow step by step the computation done for the single antenna case, which finally yields that

$$P(\mathbf{Y}^\dagger \mathbf{\Delta} \mathbf{Y} \geq 0 | \mathbf{g}) \leq \frac{1}{2} \left(\prod_{k=1}^T c_\rho \Delta_k \Delta_k^\dagger \sum_{j=1}^N \mathcal{G}_{jk}^\dagger \mathcal{G}_{jk} \frac{1}{16} + 1 \right)^{-1},$$

which concludes the proof. ■

Thus

$$P(U_k \rightarrow U_l) = P(\mathbf{Y}^\dagger \mathbf{\Delta} \mathbf{Y} \geq 0) \leq E_{\mathbf{g} | \mathbf{g} \in \mathcal{A}} \det(\mathbf{I}_T + \frac{\rho}{32} \sum_{j=1}^N \mathcal{G}_j \mathcal{G}_j^\dagger (U_k - U_l)(U_k - U_l)^\dagger)^{-1}.$$

We now use the computation of the above expectation done in [11] to conclude that the diversity gain is given by

$$d = \begin{cases} \min\{M, N\}R & M \neq N \\ MR \left(1 - \frac{1}{M} \frac{\log \log P}{\log P} \right) & M = N, \end{cases}$$

assuming again that $(U_k - U_l)(U_k - U_l)^\dagger$ is full rank (that is, the code is *fully diverse*).

Fig. 6 shows simulation results comparing the performance of a 6 nodes network with one antenna at transmitter and receiver to a 3 nodes network with 2 antennas at both transmitter and receiver nodes. The x axis shows the power of the system in dBs, and the y axis the block error rate (BLER). We are interested in the diversity of both systems. Note that both networks use the same codebook, meaning that the 2 antennas network is operating at half the rate, explaining its better behaviour. The diversity of the 2 antennas network is however slightly better, which is expected looking at the expression of the diversity for both cases, since we have $R \left(1 - \frac{\log \log P}{\log P} \right)$ versus $2R \left(1 - \frac{1}{2} \frac{\log \log P}{\log P} \right)$.

VII. CONCLUSION

We considered the problem of coding over a wireless relay network. While existing schemes heavily rely on the knowledge of the channel, either at both relays and receiver, or at least at the receiver, we presented a scheme that requires no channel knowledge. Furthermore, it is available for any number of relay nodes.

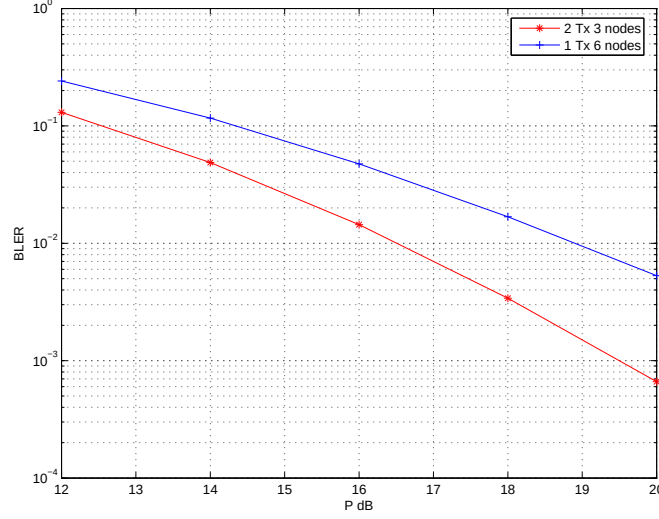


Fig. 6. A 6 nodes network with 1 antenna versus a 3 nodes network with 2 antennas.

This scheme is based on two ideas: distributed space-time coding, where relay nodes cooperate to encode the data, and differential MIMO coding, a popular technique to code over a noncoherent MIMO channel. We analyzed this strategy and showed that the diversity of the system depends on the number of relay nodes. Actually, when confronted to d node failures, a R nodes network behaves, in terms of diversity, as a network with $R - d$. Finally, we extended our construction to the case where both the transmitter and receiver nodes have several antennas.

ACKNOWLEDGMENTS

The authors would like to thank M. Varanasi and Y. Jing for sending their preprints.

REFERENCES

- [1] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Transactions on Information Theory*, vol. 51, no. 12, Dec. 2005.
- [2] H. Bölcskei, R. U. Nabar, Ö. Oyman, and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks", *IEEE Trans. Wireless Communications*, vol. 5, no. 6, June 2006.
- [3] K. L. Clarkson, W. Sweldens and A. Zheng, "Fast Multiple Antenna Differential Decoding", *IEEE Trans. Commun.*, vol. 49, no 2, 2001.
- [4] P. Dayal and M. K. Varanasi, "Distributed QAM-Based Space-Time Block Codes for Efficient Cooperative Multiple-Access Communication", submitted to *IEEE Trans. on Information Theory*, March 2006.
- [5] P. Elia, K. Vinodh, M. Anand and P. V. Kumar, "D-MG Tradeoff and optimal codes for a class of AF and DF cooperative communication protocols," submitted to the *IEEE Trans. Inform. Theory*, available at: <http://arxiv.org/pdf/cs.IT/0611156>.

- [6] B. M. Hochwald, T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading", *IEEE Transactions on Information Theory*, March 2000.
- [7] B. Hochwald and W. Sweldens, "Differential unitary space time modulation", *IEEE Trans. Commun.*, vol. 48, Dec. 2000.
- [8] K. J. Horadam, "A generalised Hadamard transform", in the proceedings of *ISIT 05*, Adelaide.
- [9] B. Hughes, "Differential space-time modulation," *IEEE Trans. Inform. Theory*, vol 46, Nov. 2000.
- [10] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," *IEEE Trans. on Wireless Communications*, vol. 5, no 12, December 2006.
- [11] Y. Jing and B. Hassibi, "Cooperative Diversity in Wireless Relay Networks with Multiple-Antenna Nodes", submitted to *IEEE Transactions on Signal Processing*.
- [12] Y. Jing and H. Jafarkhani, "Using Orthogonal and Quasi-Orthogonal Designs in Wireless Relay Networks", in the proceedings of *Globecom 2006*.
- [13] Y. Jing and H. Jafarkhani, "Distributed Differential Space-Time Coding for Wireless Relay Networks", submitted to *IEEE Trans. on Communications*.
- [14] T. Kiran and B. Sundar Rajan, "Distributed space-time codes with reduced decoding complexity," in the proceedings of *ISIT 06*, Seattle.
- [15] T. Kiran and B. Sundar Rajan, "Partially-coherent distributed space-time codes with differential encoder and decoder," in the proceedings of *ISIT 06*, Seattle.
- [16] G. Susinder Rajan, B. Sundar Rajan, "A Non-Orthogonal Distributed Space-Time Coded Protocol Part I: Signal Model and Design Criteria", proceedings of *ITW06*, Chengdu.
- [17] J. N. Laneman and G. W. Wornell, "Distributed Space-Time Coded Protocols for Exploiting Cooperative Diversity in Wireless Networks," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415-2525, Oct. 2003.
- [18] F. Oggier and B. Hassibi, "An Algebraic Family of Distributed Space-Time Codes for Wireless Relay Networks", in the proceedings of *ISIT 2006*.
- [19] F. Oggier and B. Hassibi, "An Algebraic Coding Scheme for Wireless Relay Networks with Multiple-Antenna Nodes", submitted to *IEEE Trans. Signal Processing*, March 2006.
- [20] F. Oggier and B. Hassibi, "A Coding Strategy for Wireless Networks with no Channel Information", in the proceedings of *Allerton conference 2006*.
- [21] F. Oggier and B. Hassibi, "A Coding Scheme for Wireless Networks with Multiple Antenna Nodes and no Channel Information", to appear in *ICASSP 2007*.
- [22] A. Shokrollahi, B. Hassibi, B. M. Hochwald, W. Sweldens, "Representation Theory for High-Rate Multiple-Antenna Code Design", *IEEE Trans. Information Theory*, vol. 47, no 6, Sept. 2001.
- [23] S. Yang and J.-C. Belfiore, "Optimal Space-Time Codes for the Amplify-and-Forward Cooperative Channel," *IEEE Trans. Inform. Theory*, vol. 53, no 2, February 2007.
- [24] S. Yiu, R. Schober, L. Lampe, "Differential Distributed Space-Time Block Coding", *IEEE PacRim conference*, 2005.