Approximately-Universal Space-Time Codes for the Parallel, Multi-Block and Cooperative-Dynamic-Decode-and-Forward Channels

Petros Elia and P. Vijay Kumar *

Abstract

Explicit codes are constructed that achieve the diversity-multiplexing gain tradeoff of the cooperative-relay channel under the dynamic decode-and-forward protocol for any network size and for all numbers of transmit and receive antennas at the relays.

A particularly simple code construction that makes use of the Alamouti code as a basic building block is provided for the single relay case.

Along the way, we prove that space-time codes previously constructed in the literature for the block-fading and parallel channels are approximately universal, i.e., they achieve the DMT for any fading distribution. It is shown how approximate universality of these codes leads to the first DMT-optimum code construction for the general, MIMO-OFDM channel.

1 Introduction

Cooperative relay communication is a promising means of wireless communication in which cooperation is used to create a virtual transmit array between the source and the destination, thereby providing the much-needed diversity to combat the fading channel.

Consider a communication system in which there are a total of \( N + 1 \) nodes that cooperate in the communication between source node \( S \) and destination node

*Petros Elia and P. Vijay Kumar are with the Department of EE-Systems, University of Southern California, Los Angeles, CA 90089 (elia,vijayk)@usc.edu). This work was carried out while P. Vijay Kumar was on a leave of absence at the Indian Institute of Science Bangalore. This research is supported in part by NSF-ITR Grant CCR-0326628 and in part by the DRDO-IISc Program on Advanced Research in Mathematical Engineering.
D. The remaining \((N - 1)\) nodes thus act as relays. We follow the literature in making the assumptions listed below concerning the channel. Our description is in terms of the equivalent complex-baseband, discrete-time channel.

- All nodes have a single transmit and single receive antenna and are assumed to transmit synchronously.
- The number of channel uses \(T\) over which communication takes place is short enough to invoke the quasi-static assumption, i.e., the channel fading coefficients are fixed for the duration of the communication,
- We assume half-duplex operation at each node, i.e., at any given instant a node can either transmit or receive, but not do both.
- The noise vector at the receivers is assumed to be comprised of i.i.d., circularly symmetric complex gaussian \(\mathcal{CN}(0, \sigma^2)\) random variables.

2 The DDF Protocol

Under the DDF protocol, the source transmits for a total time duration of \(BT\) channel uses. This collection of \(BT\) channel uses is partitioned into \(B\) blocks with each block composed of \(T\) channel uses. Communication is slotted in the sense that each relay is constrained to commence transmission only at block boundaries. A relay will begin transmitting after listening for a time duration equal to \(b\) blocks only if the channel “seen” by the relay is good enough to enable it to decode the signal from the source with negligible error probability. We explain in more detail.

2.1 Notation and Expressions for the Received Signal

Initially, we will assume that each relay node has a single transmit antenna. The extension to arbitrary number of antennas is straightforward. We will similarly make the initial assumption that the destination node has a single receive antenna.

It will be convenient at times to regard the destination node as the \((N + 1)\)th relay, i.e., \(D \equiv R_{N+1}\) and the source as the first relay, i.e., \(S \equiv R_1\). The notation below is with respect to a fixed channel realization that lasts for the \(B\)-block duration.

Let \(x_b(n), 1 \leq b \leq B, n = 1, 2, \cdots, N\) denote the \(T\)-tuple transmitted by the \(n\)th node during the \(b\)th block. Since all nodes do not transmit in all blocks, we will make the assignment \(x_b(n) = \varnothing\), where we regard \(\varnothing\) as the “empty” vector to handle the case of no transmission. In particular, the vectors \(x_b(1), b = 1, 2, \cdots, B\) denote the \(B\) successive transmissions by the source.
Let us assume that up until the end of the $(b-1)$th block, we know which relays began transmitting and when. We will assume that once a relay has begun transmitting, it will keep on transmitting thereafter until the end of the $B$th block. Let $\mathcal{I}_k$ denote the set of indices of the relays that transmit during the $k$th block, $k = 1, 2, \cdots, B$. We will refer to $\mathcal{I}_k$ as the $k$th activation set. Clearly

$$\mathcal{I}_1 = \{1\}$$
$$\mathcal{I}_k \subseteq \mathcal{I}_{k+1}, \ 1 \leq k \leq (b-2).$$

We next proceed to determine for the relays not in $\mathcal{I}_{b-1}$, whether or not the time is right for them to begin transmission during the $b$th block. In other words, we will determine $\mathcal{I}_b$ given $\{\mathcal{I}_k\}_{k=1}^{b-1}$. Since $\mathcal{I}_1$ is known, this procedure will allow us to recursively determine the activation sets $\mathcal{I}_k$ for all $1 \leq k \leq B$.

We will begin by first identifying the signal received by such a relay during the $(b-1)$th block. Let $\zeta_b$, $1 \leq b \leq B$, denote the size of $\mathcal{I}_b$, i.e.,

$$|\mathcal{I}_b| = \zeta_b.$$  

Clearly,

$$1 = \zeta_1 \leq \zeta_2 \leq \cdots \leq \zeta_{b-1} \leq N.$$  

Let the elements of $\mathcal{I}_k$, $1 \leq k \leq (b-1)$, be given by

$$\mathcal{I}_k = \{1 = m_1, m_2, \cdots, m_{\zeta_k}\}.$$  

We use $h(m, n)$ to denote the fading coefficient between the $m$th and $n$th nodes. Let $n \notin \mathcal{I}_{b-1}$ and

$$h^k(n) = [h(m_1, n), h(m_2, n), \cdots, h(m_{\zeta_k}, n)]$$
$$X_k = \begin{bmatrix}
\mathcal{I}_k(m_1) \\
\mathcal{I}_k(m_2) \\
\vdots \\
\mathcal{I}_k(m_{\zeta_k})
\end{bmatrix}.$$  

Let

$$y^k(n) = [y_{(k, 1)}(n) \ y_{(k, 2)}(n) \ \cdots \ y_{(k, T)}(n)]$$
$$w^k(n) = [w_{(k, 1)}(n) \ w_{(k, 2)}(n) \ \cdots \ w_{(k, T)}(n)]$$

denote the received signal and noise vector at the $n$th node during the $k$th block. Then we have

$$y^k(n) = h^k(n)X_k + w^k(n).$$
Therefore the totality of the received signal at the $n$th node up until the end of the $(b - 1)$th block is given by

$$
\begin{align*}
\begin{bmatrix}
y_t^1(n) & \cdots & y_{b-1}^1(n)
\end{bmatrix} &= \begin{bmatrix}
h_t^1(n) & \cdots & h_{b-1}^1(n)
\end{bmatrix} \\
& \begin{bmatrix}
X_1 \\
\vdots \\
X_{b-1}
\end{bmatrix} \\
& + \begin{bmatrix}
w_t^1(n) & \cdots & w_{b-1}^1(n)
\end{bmatrix}.
\end{align*}
$$

(1)

Note that the vectors $h_l(n)$ as well as the matrices $X_l$, $1 \leq l \leq b - 1$ are in general, of different sizes.

### 2.1.1 Signal at Destination

Since $D \equiv R_{N+1}$, by replacing $n$ by $(N + 1)$ and $b - 1$ by $B$ in equation (1) above, we recover the expression for the received signal at the destination during the $B$th block:

$$
\begin{align*}
\begin{bmatrix}
y_t^1(N + 1) & \cdots & y_{B}^1(N + 1)
\end{bmatrix} &= \begin{bmatrix}
h_t^1(N + 1) & \cdots & h_{B}^1(N + 1)
\end{bmatrix} \\
& \begin{bmatrix}
X_1 \\
\vdots \\
X_{B}
\end{bmatrix} \\
& + \begin{bmatrix}
w_t^1(N + 1) & \cdots & w_{B}^1(N + 1)
\end{bmatrix}.
\end{align*}
$$

(2)

### 2.1.2 Outage of Relay Node

From (3), we note that the channel “seen” by the $n$th relay node over the course of the first $b - 1$ blocks is the MISO (multiple-input single output) channel characterized by the matrix equation

$$
y = \begin{bmatrix}
h_t^1(n) & \cdots & h_{b-1}^1(n)
\end{bmatrix}x + w.
$$

(3)

The $n$th relay node can only hope to decode reliably at the end of the $(b - 1)$th block if at that point, it has sufficient mutual information to recover the transmitted signal whose information content equals $r B \log(\rho)$ bits. Here $r$ denotes the multiplexing gain, $\rho$ the signal to noise ratio, and $r \log(\rho)$ the rate of communication between source and destination [15]. If it does not have sufficient information, then we say that the relay is in outage. Thus the probability of outage $P_{\text{out}, n, b-1}(r)$ of
the $n$th relay node at the end of the $(b - 1)$th block is given by

$$P_{\text{out},n,b-1}(r) = \Pr \left( 1 + \rho \sum_{l=1}^{b-1} | h_l^T(n) |^2 < r \frac{BT}{(b-1)T} \log(\rho) \right),$$

Under the DDF protocol, the $n$th relay node at the end of block $b - 1$ uses this expression to decide whether or not it is ready to decode. If it is ready to decode, then it will proceed to do so and then begin transmitting from block $b$ onwards, i.e., $n \in \mathcal{I}_b$.

### 2.2 Performance under the DDF Protocol

In our analysis of the DDF protocol, we will make the assumption that if a relay does decode erroneously, then this error will propagate and cause the receiver to decode incorrectly as well. Thus the receiver at the destination will decode correctly if and only if in addition to the receiver at the decoder, the receivers at all intermediate nodes that have participated in relaying of the transmitted signal have also decoded correctly.

A lower bound on the probability of error of the DDF scheme can thus be derived by making the assumption that when the channel seen by a relay node is not in outage and the relay proceeds to decode the signal transmitted by the source, it will do so without error. Under this condition, the error probability of the DDF scheme, will be lower bounded by the probability of outage of the channel (2), seen by the destination. In Section 3.2, we will construct codes whose error performance at large SNR is equal to this lower bound, thereby establishing that this lower bound is indeed the error probability associated with the DMG tradeoff of the DDF protocol.

Let $\gamma$ denote the vector composed of the $\binom{N+1}{2}$ fading coefficients

$$\left\{ h(m,n) \mid n > m, \quad 1 \leq m \leq N, \quad 2 \leq n \leq (N + 1), \right\}$$

ordered lexicographically. We will use $\Gamma$ to denote the random vector of which $\gamma$ is a realization. The activations sets $\mathcal{I}_k$ are clearly a function of the channel realization $\gamma$. Writing $\mathcal{I}_k(\gamma)$ in place of $\mathcal{I}_k$ to emphasize this, let us define

$$\mathcal{I}(\gamma) = (\mathcal{I}_1(\gamma), \cdots, \mathcal{I}_B(\gamma)).$$
Let $A$ denote the collection of all possible activation sets. It follows that the error probability of the DDF scheme satisfies

$$P_e(r) \geq \sum_{\mathcal{I} \in A} \int_{\gamma \in \mathcal{R}(\mathcal{I})} p_r(\gamma) \, d\gamma$$

where

$$\mathcal{R}(\mathcal{I}) = \left\{ \gamma \mid \left(1 + \rho \sum_{b=1}^{B} |h^t_l(N + 1)|^2\right) < r \log(\rho) \right\}.$$ 

### 2.3 Notation to Aid in Code Analysis

Returning to the expression for the signal at the $n$th relay node up until the $(b-1)$th block in (1), we extend the vectors $h_k(n)$ and the matrices $X_k$ to be of equal size with a view towards the ST code construction to be presented in Section 3.2.

The vectors

$$\{h_k(n) \mid 1 \leq k \leq b-1, \ 1 \leq n \leq (N+1)\}$$

will be extended by zero padding, while the matrices $X_k, 1 \leq k \leq b-1$ will be padded with arbitrary row vectors. The extra row vectors can be chosen arbitrarily since the extended matrix $\hat{X}_k$ will be left multiplied by row vectors $\hat{h}_k(n)$ having zeros in the locations corresponding to the indices of the row vectors where padding of the matrix $X_k$ takes place.

We thus define, for $1 \leq k \leq b-1$,

$$\hat{h}_k(n) = [\hat{h}_k(1,n) \ \hat{h}_k(2,n) \ \cdots \ \hat{h}_k(N,n)]$$

where

$$\hat{h}_k(m,n) = \begin{cases} h(m,n) & m \in \mathcal{I}_k \\ 0 & \text{else.} \end{cases}$$

Also, let

$$\hat{X}_k = \begin{bmatrix} \hat{x}_k(1) \\ \hat{x}_k(2) \\ \vdots \\ \hat{x}_k(N) \end{bmatrix}$$

where

$$\hat{x}_k(m) = \begin{cases} x_k(m) & m \in \mathcal{I}_k \\ \text{arbitrary } n\text{-length vector} & \text{else.} \end{cases}$$
In terms of the extended vector and extended matrix notation, the received signal at the \( n \)th relay node, \( n \notin I_{b-1} \) and the destination can respectively be re-expressed in the form

\[
[y_{t1}^t(n) \cdots y_{tB-1}^t(n)] = [\hat{h}_{t1}^t(n) \cdots \hat{h}_{tB-1}^t(n)] \\
\begin{bmatrix}
\hat{X}_1 \\
\vdots \\
\hat{X}_{b-1}
\end{bmatrix} \\
+ [w_{t1}^t(n) \cdots w_{tB-1}^t(n)],
\]

(4)

\[
[y_{t1}^t(N+1) \cdots y_{tB}^t(N+1)] = [\hat{h}_{t1}^t(N+1) \cdots \hat{h}_{tB}^t(N+1)] \\
\begin{bmatrix}
\hat{X}_1 \\
\vdots \\
\hat{X}_B
\end{bmatrix} \\
+ [w_{t1}^t(N+1) \cdots w_{tB}^t(N+1)].
\]

(5)

In this representation, all vectors \( \hat{h}_l(n) \) are of the same size, \((1 \times T)\). The same comment also applies to the matrices \( \hat{X}_l, 1 \leq l \leq b-1 \), which are of size \((N \times T)\).

As will be shown in Section 4 below, ST codes that are approximately universal for an appropriate class of block-fading channels will be the building blocks of codes for the DDF protocol that attain the DMG performance of this channel. For this reason, a discussion on the block-fading channel is presented in the next two sections.

3 The Block-Fading Channel

3.1 Outage Probability

Consider the block-fading MIMO channel with \( n_t \) transmit and \( n_r \) receive antennas and \( B \) blocks, characterized by

\[
y_b = H_b x_b + w_b, \quad 1 \leq b \leq B.
\]

(6)
Thus each matrix $H_b$ is of size $(n_r \times n_t)$. The probability of outage of this channel is given by

$$P_{\text{out}}(r) = \Pr(\sum_{b=1}^{B} \log \det(I_{n_r} + \rho H_b H_b^\dagger) < r B \log(\rho))$$

$$= \Pr(\log \det(I_{B n_r} + \rho \Lambda_H A_H^\dagger) < r B \log(\rho))$$

$$= \Pr(\log \det(I_{B n_t} + \rho \Lambda_H A_H^\dagger) < r B \log(\rho))$$

where $\rho$ is the SNR and where $\Lambda_H$ is the $(B n_r \times B n_t)$ block diagonal matrix

$$\Lambda_H = \begin{bmatrix} H_1 & & \\ & H_2 & \\ & & \ddots \\ & & & H_B \end{bmatrix}.$$

In the above, $\hat{=} \leq \leq \hat{\geq}$ corresponds to exponential equality and inequality. For example, $y = \rho^x$ is used to indicate that $\lim_{\rho \to \infty} \frac{\log(y)}{\log(\rho)} = x$. Let $q = n_t B$ and let

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_q$$

be an ordering of the $q$ eigenvalues of $\Lambda_H^\dagger \Lambda_H$. Note that if $n_r < n_t$, then

$$\lambda_1 = \lambda_2 = \cdots = \lambda_{(n_t-n_r)B} = 0.$$ (8)

Let $\delta = ([n_t - n_r]B)^+$ where $(x)^+$ denotes $\max\{x, 0\}$, and let the $\alpha_i$ be defined by

$$\lambda_i = \rho^{-\alpha_i}, \quad \delta + 1 \leq i \leq q.$$ (7)

Then

$$P_{\text{out}}(r) = \Pr(\sum_{i=\delta+1}^{q} (1 - \alpha_i)^+ < r B).$$

We will now proceed to identify a ST code in the next section, Section 3.2, that is approximately universal for the class of block-fading channels, i.e., a code that achieves the D-MG tradeoff of the channel model in (6) for every statistical distribution of the fading coefficients $\{[H_b]_{i,j}\}$.

Similar construction of codes for such a setting have previously been identified in [20,26] and independently in [21]. We adopt the code-construction technique of these papers for the most part, although the construction presented here is slightly more general, for example, we permit the individual block codes to be rectangular and offer flexibility with respect to number of conjugate blocks employed. Most importantly though, our proof will establish the new result that these codes are approximately universal for the block-fading channel and parallel channels.
3.2 Approximately-Universal Codes for the Block-Fading Channel

3.2.1 Constructing the Appropriate Cyclic Division Algebra

Let $T$ be an integer satisfying $T \geq n_t$. Let $m \geq B$ be the smallest integer such that the gcd of $m, T$ equals 1, i.e., $(m, T) = 1$. Let $K, M$ be cyclic Galois extensions of $\mathbb{Q}(i)$ of degrees $m, T$ whose Galois groups are generated respectively by the automorphisms $\phi_1, \sigma_1$, i.e.,

$$\text{Gal}(K/\mathbb{Q}(i)) = < \phi_1 >$$
$$\text{Gal}(M/\mathbb{Q}(i)) = < \sigma_1 > .$$

Let $L$ be the composite of $K, M$, see Fig. 1. Then it is known that $L/\mathbb{Q}(i)$ is cyclic and that further,

$$\text{Gal}(L/\mathbb{Q}(i)) \cong \text{Gal}(K/\mathbb{Q}(i)) \times \text{Gal}(M/\mathbb{Q}(i)).$$

Thus every element of $\text{Gal}(L/\mathbb{Q}(i))$ can be associated with a pair $(\phi_1, \sigma_1)$ belonging to $\text{Gal}(K/\mathbb{Q}(i)) \times \text{Gal}(M/\mathbb{Q}(i))$. Let $\phi, \sigma$ be the automorphisms associated to the pairs $(\phi_1, \text{id}), (\text{id}, \sigma_1)$ respectively. Then $\phi, \sigma$ are the generators of the Galois groups $\text{Gal}(L/M)$, $\text{Gal}(L/K)$ respectively.

![Figure 1: Construction of the underlying cyclic-division algebra.](image)

Let $\gamma \in K$ be a non-norm element of the extension $L/K$, i.e., the smallest exponent $e$ for which $\gamma^e$ is the norm of an element of $L$ is $T$. Let $z$ be an indeterminate satisfying $z^T = \gamma$. Consider the $T$-dimensional vector space

$$D = \{ z^{T-1} \ell_{T-1} \oplus z^{T-2} \ell_{T-2} \oplus \cdots \ell_0 \mid \ell_i \in \mathbb{L} \} .$$

We define multiplication on $D$ by setting $\ell_i z = z\sigma(\ell_i)$ and extending in a natural fashion. This turns $D$ into a cyclic division algebra (CDA) whose center is $K$ and
having \( \mathbb{L} \) as a maximal subfield. See [13, 16] for an exposition of the relevant background on division algebras. Every element \( x = z^{T-1}\ell_{T-1} + z^{T-2}\ell_{T-2} + \cdots + \ell_0 \) in \( D \) has the regular representation

\[
X = \begin{bmatrix}
\ell_0 & \gamma\sigma(\ell_{T-1}) & \ldots & \gamma\sigma^{T-1}(\ell_1) \\
\ell_1 & \sigma(\ell_0) & \ldots & \gamma\sigma^{T-1}(\ell_2) \\
\vdots & \vdots & \ddots & \vdots \\
\ell_{T-1} & \sigma(\ell_{T-2}) & \ldots & \sigma^{T-1}(\ell_0)
\end{bmatrix}. \tag{9}
\]

The determinant of such a matrix is known to lie in \( \mathbb{K} \). Given a matrix \( X \) with components \( X_{i,j} \in \mathbb{L} \), we define \( \phi(X) \) to be the matrix over \( \mathbb{L} \) whose \( (i, j) \)th component is given by \( [\phi(X)]_{i,j} = \phi([X]_{i,j}) \). Note that in this case,

\[
\prod_{i=0}^{m-1} \det(\phi^i(X)) = \prod_{i=0}^{m-1} \phi^i(\det(X)) = \prod_{i=0}^{m-1} \phi^i_1(\det(X)) \in \mathbb{Q}(i).
\]

Hence if the elements \( \ell_i \) underlying the matrix \( X \) are in addition, restricted to lie in the ring \( \mathcal{O}_L \) of algebraic integers of \( \mathbb{L} \), then we have that

\[
\prod_{i=0}^{m-1} \det(\phi^i(X)) \in \mathbb{Z}(i)
\]

so that

\[
| \prod_{i=0}^{m-1} \det(\phi^i(X)) |^2 \geq 1. \tag{10}
\]

### 3.2.2 Space-time Code Construction on the CDA

Let \( \mathcal{X} \) be the rectangular \((n_t \times T)\) ST code comprised of the first \( n_t \) rows of the regular representations of the elements \( \sum_{i=0}^{T-1} z^i\ell_i \), where \( \ell_i \) are restricted to be of the form:

\[
\ell_i = \sum_{j=1}^{T} \ell_{i,j} \gamma_j, \quad \ell_{i,j} \in A_{QAM}
\]
where \( \{ \gamma_1, \cdots, \gamma_T \} \) are a basis for \( L/K \) and where

\[
\mathcal{A}_{\text{QAM}} = \{ a + ib \mid |a|, |b| \leq (M-1), \ a, b \ \text{odd} \} \subseteq \mathbb{Z}[i]
\]
denotes the QAM constellation of size \( M^2 \). Note that as a result, we have ensured that \( \ell_i \in \mathcal{O}_L \). Thus each code matrix in \( \mathcal{X} \) is of the row-deleted form

\[
X = \begin{bmatrix}
\ell_0 & \gamma \sigma(\ell_{T-1}) & \cdots & \gamma \sigma^{T-1}(\ell_1) \\
\ell_1 & \sigma(\ell_0) & \cdots & \gamma \sigma^{T-1}(\ell_2) \\
\vdots & \vdots & \ddots & \vdots \\
\ell_{n_t-1} & \gamma \sigma(\ell_{n_t-2}) & \cdots & \gamma \sigma^{T-1}(\ell_{n_t})
\end{bmatrix}, \quad (11)
\]

Let \( S \) be the \((Bn_t \times BT)\) ST code comprised of code matrices having the block diagonal form:

\[
S = \left\{ \theta \begin{bmatrix}
X \\
\phi(X) \\
\vdots \\
\phi^{B-1}(X)
\end{bmatrix}, \ X \in \mathcal{X} \right\}
\]

where \( \theta \) accounts for SNR normalization. When this code matrix is in use, the received signal over the block-fading channel is given by

\[
\begin{bmatrix}
Y_1 \\
Y_2 \\
\vdots \\
Y_B
\end{bmatrix} = \begin{bmatrix}
H_1 & H_2 & \cdots & H_B
\end{bmatrix} S + \begin{bmatrix}
W_1 \\
W_2 \\
\vdots \\
W_B
\end{bmatrix}, \quad (12)
\]

This can also be expressed in the form

\[
\begin{bmatrix}
Y_1 \\
Y_2 \\
\vdots \\
Y_B
\end{bmatrix} = \theta \begin{bmatrix}
H_1 & H_2 & \cdots & H_B
\end{bmatrix} \begin{bmatrix}
X \\
\phi(X) \\
\vdots \\
\phi^{B-1}(X)
\end{bmatrix} + \begin{bmatrix}
W_1 \\
W_2 \\
\vdots \\
W_B
\end{bmatrix}, \quad (14)
\]

in which the channel matrix is of block-diagonal form. This latter form is convenient when comparing the block-fading channel with the parallel channel.

### 3.2.3 Proof of Optimality

We will now show that the ST code \( S \) is approximately universal for the class of channel models described by (6). From information rate considerations we must have \((M^2)^{mT^2} = \rho^{rBT} \), i.e., \( M^2 = \rho^\frac{rBT}{mT} \). Also, since \( \theta^2 M^2 \cong \rho \), we have \( \theta^2 \cong \rho^{1 - \frac{rBT}{mT}} \).
We next examine the error performance of the code. Let

\[ H = [H_1 \ H_2 \ \cdots \ H_B] \]

and let \( \Delta S = S_1 - S_2 \) where \( S_1, S_2 \), are two distinct code matrices belonging to \( S \). We have the following expression for the squared Euclidean distance:

\[
d_E^2 = \theta^2 Tr(H \Delta S \Delta S^\dagger H^\dagger) \\
= \theta^2 Tr(\Lambda_H \Delta S \Delta S^\dagger \Lambda_H^\dagger) \\
\geq \theta^2 \sum_{i=1}^{n_t B} \lambda_i \mu_i
\]

by the mismatched eigenvalue bound (see \[13, 30\]), where \( \mu_1 \geq \mu_2 \geq \cdots \geq \mu_{n_t B} \) are the ordered eigenvalues of \( \Delta S \Delta S^\dagger \).

Let \( \Delta \hat{X} \) be the \( T \times T \) matrix that corresponds to \( \Delta X \) in the sense that it would have been the matrix obtained if in constructing the ST code \( X \), we had not deleted the bottom \( (T - n_t) \) rows from the regular representations of the elements \( \sum_{i=0}^{T-1} z^i \ell_i \). Correspondingly, let \( \Delta \hat{S} \) be the \( (BT \times BT) \) matrix

\[
\Delta \hat{S} = \begin{bmatrix}
\Delta \hat{X} \\
\phi(\Delta \hat{X}) \\
\ldots \\
\phi^{B-1}(\Delta \hat{X})
\end{bmatrix}.
\]

By the inclusion principle of Hermitian matrices, (see Theorem 4.3.15 of \[31\]) the smallest \( Bn_t \) eigenvalues of \( \Delta S \Delta S^\dagger \) are term-by-term larger than the corresponding smallest \( Bn_t \) eigenvalues of \( \Delta \hat{S} \Delta \hat{S}^\dagger \), i.e.,

\[
\mu_{Bn_t-k+1} \geq \nu_{BT-k+1}, \quad 1 \leq k \leq Bn_t,
\]

where \( \nu_1 \geq \nu_2 \geq \cdots \geq \nu_{BT} \) are the eigenvalues of \( \Delta \hat{S} \Delta \hat{S}^\dagger \).

Next, let \( \Delta \hat{\hat{S}} \) be the \( (mT \times mT) \) matrix obtained by further extending \( \Delta \hat{S} \) to include all \( m \) “conjugates” of \( \Delta \hat{X} \), i.e.,

\[
\Delta \hat{\hat{S}} = \begin{bmatrix}
\Delta \hat{X} \\
\phi(\Delta \hat{X}) \\
\ldots \\
\phi^{m-1}(\Delta \hat{X})
\end{bmatrix}.
\]

Let \( \hat{\nu}_i, 1 \leq \hat{\nu}_i \leq mT \), be the eigenvalues of \( \Delta \hat{\hat{S}}[\Delta \hat{\hat{S}}]^\dagger \) ordered such that

\[
\hat{\nu}_i = \nu_i, \quad 1 \leq i \leq BT.
\]
Then for every $1 \leq J \leq q = B n_t$, we have

\[
d_E^2 \geq \theta^2 \sum_{i=0}^{J} \lambda_{q-i} \mu_{q-i} \]
\[
\geq \theta^2 \sum_{i=0}^{J} \lambda_{q-i} \nu_{BT-i} \]
\[
= \theta^2 \sum_{i=0}^{J} \lambda_{q-i} \hat{\nu}_{BT-i} \]
\[
\geq \theta^2 \left( \prod_{i=0}^{J} \lambda_{q-i} \right) \frac{1}{\prod_{i=0}^{J} \hat{\nu}_{BT-i}^{\frac{1}{J+1}}} \]
\[
\geq \left( \frac{\prod_{i=0}^{mT} (\theta^2 \hat{\nu}_i)}{\prod_{i=1}^{BT-J+1} (\theta^2 \hat{\nu}_i) \prod_{i=BT+1}^{mT} (\theta^2 \hat{\nu}_i)} \right) \frac{1}{\prod_{i=0}^{J} \hat{\nu}_{BT-i}} \]
\[
= \left( \rho \sum_{i=0}^{J} \alpha_{q-i} (\theta^2)^{mT} \right)^{\frac{1}{J+1}} \geq \delta^{\frac{1}{J+1}} \]

where

\[
\delta_J = mT(1 - \frac{r B}{mT}) + J + 1 - mT - \sum_{i=0}^{J} \alpha_{q-i} \]
\[
= J + 1 - r B - \sum_{i=0}^{J} \alpha_{q-i} \]
\[
= \sum_{i=q-J}^{q} (1 - \alpha_i) - r B. \]

In this derivation we have made use of the fact that the product of the eigenvalues is equal to the determinant and of the non-vanishing determinant property enunciated in (10). We will now show that if the block-fading channel is not in outage, that for some $J$, $1 \leq J \leq q$, $\delta_J > 0$.

Suppose

\[
\sum_{i=1}^{q} (1 - \alpha_i) \geq (r + \epsilon) B. \quad (15)\]
Clearly for some \( i, \alpha_i < 1 \). Since \( \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_q \), assume \( \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_{q-J-1} \geq 1 \), and \( \alpha_i < 1, \ i \geq q - J \). Then (15) becomes

\[
\sum_{i=q-J}^{q} (1 - \alpha_i) \geq (r + \epsilon)B
\]

and this causes the corresponding \( \delta_J \geq \epsilon B \) hence leading to negligible error probability when not in outage. By taking the limit as \( \epsilon \to 0 \) we see that the probability of error is negligible in the no-outage region. This proves that the space-time code \( X \) achieves the DMG tradeoff regardless of the statistical distribution of the fading coefficients that comprise the matrices \( \{H_b\} \), i.e., proves approximate universality of the constructed ST code.

### 3.3 Analogous Results Hold for the Parallel Channel

By parallel channel we will mean the channel given by

\[
\begin{bmatrix}
Y_1 \\
Y_2 \\
\vdots \\
Y_B \\
\end{bmatrix} = \begin{bmatrix}
H_1 & & \\
& H_2 & \\
& & \ddots \\
& & & H_B \\
\end{bmatrix} \begin{bmatrix}
S \\
W_1 \\
W_2 \\
\vdots \\
W_B \\
\end{bmatrix},
\]

in which the channel matrix is of block-diagonal form. Consider the \((Bn_t \times T)\) space-time code \( S_{\text{par}} \) given by

\[
S_{\text{par}} = \left\{ \theta \begin{bmatrix}
X \\
\phi(X) \\
\vdots \\
\phi^{B-1}(X) \\
\end{bmatrix}, \ X \in \mathcal{X} \right\}
\]

which when used over the parallel channel leads to the equation below for the received signal at the receiver,

\[
\begin{bmatrix}
Y_1 \\
Y_2 \\
\vdots \\
Y_B \\
\end{bmatrix} = \theta \begin{bmatrix}
H_1 & & \\
& H_2 & \\
& & \ddots \\
& & & H_B \\
\end{bmatrix} \begin{bmatrix}
X \\
\phi(X) \\
\vdots \\
\phi^{B-1}(X) \\
\end{bmatrix} + \begin{bmatrix}
W_1 \\
W_2 \\
\vdots \\
W_B \\
\end{bmatrix}.
\]

Comparing this equation with the alternate expression for the block-fading channel given in (14) we see that the expressions are identical. There is one important
difference though. In the case of the block-fading channel, a rate requirement of $R$ bits per channel use translates into a space-time code $S$ of size $2^{RB_T} = \rho^{rB_T}$, whereas in the case of the parallel channel, the size of the corresponding ST code $S_{\text{par}}$ is required to be $2^{RT} = \rho^{rT}$.

It follows from this that by replacing $rB$ by $r$, one can similarly prove approximate universality of the code $S_{\text{par}}$ for the class of parallel channels. We omit the details.

### 3.4 DMT-optimal Codes for the General MIMO-OFDM Channel

The MIMO-OFDM channel can be regarded as a parallel channel in which each parallel block corresponds to a different subcarrier and can thus be represented in the form:

$$y_l = \theta H_l x_l + w_l, \quad 1 \leq l \leq Q,$$

where $Q$ is the number of OFDM tones or sub-carriers [25]. The matrices $H_l$ are correlated in general, with a correlation derived from the time-dispersion of the original ISI channel. Since the code $S_{\text{par}}$ is approximately universal, this means that the code $S_{\text{par}}$ is DMT optimal when used over the MIMO fading channel. When the code $S_{\text{par}}$ is used over the MIMO-OFDM channel, the received-signal equation will take on the form

$$
\begin{bmatrix}
Y_1 \\
Y_2 \\
\vdots \\
Y_Q
\end{bmatrix} = \theta
\begin{bmatrix}
H_1 & & \\
& H_2 & \\
& & \ddots \\
& & & H_Q
\end{bmatrix}
\begin{bmatrix}
X \\
\phi(X) \\
\vdots \\
\phi^{-1}(X)
\end{bmatrix}
+ 
\begin{bmatrix}
W_1 \\
W_2 \\
\vdots \\
W_Q
\end{bmatrix}.
$$

DMT-optimal codes for the OFDM channel have previously been constructed in [20] and [22]. In [20], the authors provide a proof only for the case when the matrices $H_n$ appearing along the diagonal are i.i.d. Rayleigh. The DMT-optimal construction in [22] is for the SIMO-OFDM case. We thus believe the results in this paper represent the first construction of DMT-optimal codes for the general OFDM-MIMO channel.

### 4 Codes Attaining the DMG of the DDF Protocol

We now show how ST codes constructed for the block-fading channel can be used to construct optimal codes under the DDF protocol.
We consider the DDF protocol as it applies to a communication system in which there are a total of $N + 1$ nodes that cooperate in the communication between source node $S$ and destination node $D$.

As in Sections 1, 2, under the DDF protocol, the source transmits for a total time duration of $BT$ channel uses. This collection of $BT$ channel uses is partitioned into $B$ blocks with each block composed of $T$ channel uses. Communication is slotted in the sense that each relay is constrained to commence transmission only at block boundaries. A relay will begin transmitting after listening for a time duration equal to $b$ blocks only if the channel “seen” by the relay is good enough to enable it to decode the signal from the source with negligible error probability.

Our coding strategy runs as follows. The role played by $n_t$ in the block-fading scenario is now played by the number $N$ which is the number of nodes in the network capable of transmitting to the destination. Let $X$ be the rectangular $(N \times T)$ ST code comprised of the first $N$ rows of the regular representations of the elements $\sum_{i=0}^{T-1} z^i \ell_i$, where $\ell_i$ are restricted to be of the form:

$$\ell_i = \sum_{j=1}^{T} \ell_{i,j} \gamma_j, \quad \ell_{i,j} \in \mathcal{A}_{QAM}.$$  

Let $D$ be the $(BN \times BT)$ ST code comprised of code matrices having the block diagonal form:

$$D = \left\{ \begin{bmatrix} X & \phi(X) & \cdots & \phi^{B-1}(X) \\ \vdots \end{bmatrix}, \quad X \in \mathcal{X} \right\}$$

where $\theta$ accounts for SNR normalization. The code to be used then has the following simple description. The source $S$ sends the first row of each of the matrices $X$, $\phi(X)$, $\cdots$, $\phi^{B-1}(X)$ in successive blocks. Let us assume that relay node $R_n$, $2 \leq n \leq N$, is not in outage for the first time at the conclusion of the $(b - 1)$th block. Then $R_n$ is ready to decode a the end of the $b - 1$th block. Thereafter, it proceeds to send in succession, the $n$th rows of the matrices $\phi^b(X)$, $\phi^{b+1}(X)$, $\cdots$, $\phi^{B-1}(X)$. Thus the appended matrices $\hat{X}_i$ appearing in (4), (5), correspond to the matrices $\phi^{i-1}(X)$ in (16).

It is easy to show using the results stated earlier relating to the block-fading channel that this coding strategy ensures that whenever a relay node decodes, it does so with negligible probability of error. The destination error probability is also similarly guaranteed to have error probability that is SNR-equivalent to the outage probability, thus proving DMG-optimality of the constructed ST code.
This follows since each relay node $R_n$, $n \notin \mathcal{I}_{b-1}$ “sees” a block-fading channel (see (4)) and the coding strategy we have adopted ensures that the code matrix carrying data from the nodes in $\mathcal{I}_{b-1}$ is DMT optimal for the corresponding block-fading channel. A similar statement is true for the relay $R_{N+1}$ that corresponds to the destination, since the corresponding channel equation is of the same block-fading form see (5).

4.1 Example

We illustrate with an example. Consider a network in which there are a total of $N + 1 = 5$ nodes including source $S \equiv R_1$ and destination $D \equiv R_5$ nodes. Let the block length $T = 4$ and the number of blocks $B = 4$. Let us assume that at the end of the first block, relay $R_3$ is not in outage and therefore in a position to decode. Let us assume that relay $R_4$ is ready to decode at the end of the 3rd block and that relay $R_2$ is in outage throughout and thus does not participate in the communication. Thus

$$\mathcal{I}_1 = \{1\}, \quad \mathcal{I}_2 = \{1, 3\}, \quad \mathcal{I}_3 = \{1, 3\}, \quad \mathcal{I}_4 = \{1, 3, 4\}$$

here. In terms of the notation introduced in Section 2.1, the signal received by relays $R_2, R_3, R_4, R_5$ prior to decoding are given as follows. Decisions at each of the relays $R_3, R_4, R_5$ are based on the received signals

$$y_t^{x_1}(3) = h(1, 3)x_t^1(1) + w_t^1(3),$$

$$[y_t^2(4) y_t^3(4) y_t^4(4)] = [h(1, 4) h(1, 4) h(3, 4) h(3, 4) h(3, 4)]$$

$$\begin{bmatrix}
  x_t^1(1) \\
  x_t^2(1) \\
  x_t^2(3) \\
  x_t^3(1) \\
  x_t^3(3)
\end{bmatrix}$$

$$+ [w_t^2(4) w_t^3(4) w_t^4(4)]$$
\[
\begin{bmatrix}
[y^1_1(5) & y^1_2(5) & y^1_3(5) & y^1_4(5)] = \\
[h(1, 5) & h(1, 5) & h(3, 5) & h(1, 5) & h(3, 5) & h(1, 5) & h(3, 5) & h(4, 5)]
\end{bmatrix}
\]

\[
\begin{bmatrix}
+x^1_1(1) \\
x^1_2(1) \\
x^1_3(1) \\
x^1_4(1) \\
+x^2_1(3) \\
x^2_2(3) \\
x^2_3(3) \\
x^2_4(3) \\
+x^3_1(4) \\
x^3_2(4) \\
x^3_3(4) \\
x^3_4(4)
\end{bmatrix}
\]

respectively.

Set

\[
X(1) := \begin{bmatrix} x^1_1(1) \\ x^1_2(2) \\ x^1_3(3) \\ x^1_4(4) \end{bmatrix} = \begin{bmatrix} \ell_0 & \gamma\sigma(\ell_3) & \gamma\sigma^2(\ell_2) & \gamma\sigma^3(\ell_1) \\ \ell_1 & \ell_0 & \gamma\sigma^2(\ell_3) & \gamma\sigma^3(\ell_2) \\ \ell_2 & \gamma\sigma(\ell_1) & \ell_0 & \gamma\sigma^3(\ell_3) \\ \ell_3 & \gamma\sigma(\ell_1) & \gamma\sigma^2(\ell_2) & \ell_0 \end{bmatrix}
\]

\[
X(2) := \phi \{X(1)\}
\]
\[
X(3) := \phi^2 \{X(1)\}
\]
\[
X(4) := \phi^3 \{X(1)\}.
\]

The corresponding extended vectors are given by

\[
[\hat{h}^t_1(3)] = [h(1, 3) \ 0 \ 0 \ 0]
\]

\[
\hat{h}^t_1(4) = [h(1, 4) \ 0 \ 0 \ 0]
\]
\[
\hat{h}^t_2(4) = [h(1, 4) \ 0 \ h(3, 4) \ 0]
\]
\[
\hat{h}^t_3(4) = [h(1, 4) \ 0 \ h(3, 4) \ 0]
\]

\[
\hat{h}^t_1(5) = [h(1, 5) \ 0 \ 0 \ 0]
\]
\[
\hat{h}^t_2(5) = [h(1, 5) \ 0 \ h(3, 5) \ 0]
\]
\[
\hat{h}^t_3(5) = [h(1, 5) \ 0 \ h(3, 5) \ 0]
\]
\[
\hat{h}^t_4(5) = [h(1, 3) \ 0 \ h(3, 5) \ h(4, 5)]
\]
and we have

\[ [y_1^t(5) y_2^t(5) y_3^t(5)] = [\hat{h}_1^t(5) \hat{h}_2^t(5) \hat{h}_3^t(5) \hat{h}_4^t(5)] \begin{bmatrix} X(1) \\ X(2) \\ X(3) \\ X(4) \end{bmatrix}. \]

4.2 Extension to Multiple Antenna Case

The extension to the case of multiple antennas at each relay node and at the destination is straightforward.

5 An Alamouti-based Code for the case of a Single Relay

In this section, we provide a particularly simple code construction for the case when in addition to the source \( S \) and destination \( D \), there is a single relay antenna \( R_2 \). The basic building block for the distributed space-time code is an Alamouti code. A separate proof is given here as the proof given in previous sections does not apply here, primarily because the Alamouti code is not an approximately universal code.\(^1\)

5.1 Constructing the Appropriate Cyclic Division Algebra

The cyclic division algebra is constructed along the same lines as before, with some differences, for example \( \mathbb{Q} \) here plays the role of \( \mathbb{Q}(i) \) earlier.

Here the number of channel uses in each block equals 2, i.e., \( T = 2 \). Let \( m \geq B \) be the smallest integer such that \( (m, T) = 1 \), i.e., \( m \) is the smallest odd integer \( \geq B \). Set \( \mathbb{M} = \mathbb{Q}(i) \) and let \( \mathbb{K} \) be a cyclic Galois extension of \( \mathbb{Q} \) of degree \( m \). Let the Galois groups of \( \mathbb{K}/\mathbb{Q} \) and \( \mathbb{M}/\mathbb{Q} \) be generated respectively by the automorphisms \( \phi_1, \sigma_1 \), i.e.,

\[
\begin{align*}
\text{Gal}(\mathbb{K}/\mathbb{Q}) &= \langle \phi_1 \rangle \\
\text{Gal}(\mathbb{M}/\mathbb{Q}) &= \langle \sigma_1 \rangle.
\end{align*}
\]

Thus \( \sigma_1 \) corresponds to the complex conjugation operator. Let \( \mathbb{L} \) be the composite of \( \mathbb{K}, \mathbb{M} \), see Fig.\(^2\)

Then it is known that

\[
\text{Gal}(\mathbb{L}/\mathbb{Q}(i)) \cong \text{Gal}(\mathbb{K}/\mathbb{Q}(i)) \times \text{Gal}(\mathbb{M}/\mathbb{Q}(i)).
\]

\(^1\)This result was first presented at a poster session in Allerton 2006
Thus every element of \( \text{Gal}(L/Q(i)) \) can be associated with a pair \((\phi_1^i, \sigma_1^i)\) belonging to \( \text{Gal}(K/Q(i)) \times \text{Gal}(M/Q(i)) \). Let \( \phi, \sigma \) be the automorphisms associated to the pairs \((\phi_1, \text{id}), (\text{id}, \sigma_1)\) respectively. Then \( \phi, \sigma \) are the generators of the Galois groups \( \text{Gal}(L/M), \text{Gal}(L/K) \) respectively. Here again, \( \sigma \) is the complex conjugation operator operating on \( L \). We note that \( \sigma \) commutes with \( \phi \).

Let \( \gamma = -1 \). Then \( \gamma \) is a non-norm element of the extension \( L/K \), i.e., the smallest exponent \( e \) for which \( \gamma^e \) is the norm of an element of \( L \) is \( 2 \). This follows because the norm in \( L/K \) of any element in \( L \) is non-negative. Let \( z \) be an indeterminate satisfying \( z^2 = \gamma \). Consider the 2-dimensional vector space

\[ D = \{ z\ell_1 + \ell_0 \mid \ell_i \in L \}. \]

We define multiplication on \( D \) by setting \( \ell_iz = z\sigma(\ell_i) \) and extending in a natural fashion. This turns \( D \) into a CDA whose center is \( K \) and having \( L \) as a maximal subfield. Every element \( a = z\ell_1 + \ell_0 \) has the regular representation

\[ X = \begin{bmatrix} \ell_0 & -\ell_0^* \\ \ell_1^* & \ell_0^* \end{bmatrix} \quad (17) \]

which we recognize as the familiar Alamouti code matrix. The determinant of such a matrix is clearly real and thus lies in \( K \). Note also that the rows of every such matrix are orthogonal and hence the two eigenvalues of the matrix have the same magnitude.

Given a matrix \( X \) with components \( X_{ij} \in L \), we define \( \phi(X) \) to be the matrix over \( L \) whose \((i, j)^{th}\) component is given by \( [\phi(X)]_{i,j} = \phi([X]_{i,j}) \). Note that in
this case,

\[
\prod_{i=0}^{m-1} \det(\phi^i(X)) = \prod_{i=0}^{m-1} \phi^i(\det(X)) \in \mathbb{Q}.
\]

Hence if the elements \( \ell_i \) underlying the matrix \( X \) are in addition, restricted to lie in the ring \( \mathcal{O}_L \) of algebraic integers of \( L \), then we have that

\[
\prod_{i=0}^{m-1} \det(\phi^i(X)) \in \mathbb{Z}
\]

so that

\[
|\prod_{i=0}^{m-1} \det(\phi^i(X))|^2 \geq 1. \tag{18}
\]

### 5.2 Channel Models and Outage

We assume as before that the total duration of communication is \( 2B \) channel uses, partitioned into \( B \) blocks of 2 channel uses each. Let us assume that at the end of the \((b-1)\)th block, the relay determines for the first time that it is not in outage and begins to transmit from block \( b \) onwards. The channel perceived by the relay antenna is given by

\[
y = h_{sr}x + w
\]

where \( w \) is the usual additive noise. The probability of outage of this channel is given by

\[
\Pr(\log(1 + \rho \ | h_{sr} \ |^2)) < \frac{2rB}{2(b-1)} \log(\rho)
\]

\[
= \Pr((1 - \beta_1)^+ < \frac{rB}{(b-1)})
\]

where \( \beta_1 \) is defined by

\[
| h_{sr} \ |^2 \overset{\triangle}{=} \rho^{-\beta_1}.
\]

The channel seen by the destination takes on the form

\[
y = \begin{cases} h_{sr} \cdots h_{sr} h_{rd} \cdots h_{sr} h_{rd} \hat{x} + w, & \text{if } \{(b-1) \text{ terms}\} \leq 2[B - (b-1)] \text{ terms} \\ h^t \hat{x} + w, & \text{otherwise} \end{cases}
\]
where we have defined
\[ h := [h_{sr} \cdots h_{sr} h_{sr} h_{rd} \cdots h_{sr} h_{rd}]. \]  
(19)

5.3 DMT-Optimal Code Construction

Optimal code construction proceeds as follows. The source transmits the \( B \) blocks
\[ \theta[A \phi(A) \cdots \phi^{B-1}(A)] \]
in succession, where
\[ A = [\ell_0 - \ell_1] \]
and where
\[ \ell_i = \sum_{j=1}^m \ell_{ij} \gamma_j \]
where \( \gamma_j \) is a basis for \( \mathbb{L}/\mathbb{Q}(i) \) and where \( \ell_{ij} \in \mathcal{A}_{QAM} \).

The relay transmits from block \( b \) onwards and its transmissions are of the form
\[ \theta[\phi^b(C) \phi^{b+1}(C) \cdots \phi^{B-1}(C)] \]
where
\[ C = [\ell_1 - \ell_0]. \]

The signal seen by the receiver at the destination is thus of the form
\[
y = \theta h\left[
\begin{array}{cccc}
A & \cdots & \phi^{b-1}(A) & \\
\cdots & \phi^b(A) & \phi^b(C) & \\
\phi^{b+1}(A) & \phi^{b+1}(C) & \cdots & \\
\phi^{B-1}(A) & \phi^{B-1}(C) & \cdots &
\end{array}
\right] + w,
\]
with \( h \) given in (19).
5.4 Proof of Optimality

We will show as in Section 4 that when either the relay or the destination is not in outage, the error probability incurred by this code is negligible i.e., of order \( \rho^{-\infty} \) thus proving DMT optimality of the code.

Note that from rate considerations, we must have

\[
(M^2)^{2m} = \rho^{2rB} \\
\therefore M^2 = \rho^{\frac{rB}{m}} \\
\theta^2 M^2 = \rho \\
\Rightarrow \theta^2 = \rho^{1 - \frac{rB}{m}}.
\]

5.4.1 Optimality in the Broadcast Phase

Let

\[
A_{\Delta QAM} = \{a + ib \mid a, b, \text{even}, \ 0 \leq |a|, |b| \leq 2(M - 1)\}.
\]

The Euclidean distance between the received matrices associated to code matrices \( X_1, X_2 \) is given by

\[
d_E^2(X_1, X_2) = |h_{sr}|^2 \theta^2 \sum_{i=0}^{b-1} \left[ |\phi^i(\ell_0)|^2 + |\phi^i(-\ell_1')|^2 \right], \ \ell_i \in A_{\Delta QAM}
\]

\[
= |h_{sr}|^2 \theta^2 \sum_{i=0}^{b-1} \phi^i(|\ell_0|^2 + |\ell_1|^2)
\]

\[
= |h_{sr}|^2 \theta^2 \sum_{i=0}^{b-1} \phi^i(\ell), \ \ell = |\ell_0|^2 + |\ell_1|^2
\]

\[
\geq |h_{sr}|^2 \theta^2 b \left[ \prod_{i=0}^{b-1} \phi^i(\ell) \right]^{\frac{1}{b+1}}
\]

\[
\leq |h_{sr}|^2 \left[ \prod_{i=0}^{m-1} \theta^2 \phi^i(\ell) \right]^{\frac{1}{b+1}}
\]

\[
= |h_{sr}|^2 \left( \frac{\rho^{m(1 - \frac{rB}{m})}}{\rho(m - b + 1)} \right)^{\frac{1}{b+1}}
\]

\[
= \rho^{1 - \beta_1 - \frac{rB}{b+1}}
\]
Consider the no-outage region associated to rate \( r + \epsilon \):

\[
(1 - \beta_1)^+ \geq (r + \epsilon) \frac{rB}{b-1}
\]

and it follows that the probability of error is negligible for all \( \epsilon > 0 \).

5.4.2 Code in the Cooperation Phase

Here, by making use of the orthogonality of the rows of the Alamouti code, we can bound the minimum Euclidean distance as follows:

\[
d_E^2(\Delta X) \geq \theta^2 |h_{sd}|^2 \sum_{i=0}^{b-1} \phi^i(|\ell_0|^2 + |\ell_1|^2) \\
+ \theta^2 (|h_{sd}|^2 + |h_{rd}|^2) \sum_{i=b}^{B-1} \phi^i(|\ell_0|^2 + |\ell_1|^2) \\
= \theta^2 |h_{sd}|^2 \sum_{i=0}^{B-1} \phi^i(\ell) + \theta^2 (|h_{rd}|^2) \sum_{i=b}^{B-1} \phi^i(\ell) \\
= \theta^2 |h_{sd}|^2 \sum_{i=0}^{b-1} \phi^i(\ell) + \theta^2 |h_{sd}|^2 \sum_{i=b}^{B-1} \phi^i(\ell).
\]

Let us abbreviate and write

\[
h_1 = h_{sr} \\
h_2 = \sqrt{|h_{sr}|^2 + |h_{rd}|^2} \\
\hat{h} = \begin{bmatrix} h_1 & \cdots & h_1 \\ \cdots & \cdots & \cdots \\ h_2 & \cdots & h_2 \\ \end{bmatrix}_{b-1\text{terms}} \begin{bmatrix} \ell \\ \phi(\ell) \\ \cdots \\ \phi^{B-1}(\ell) \\ \end{bmatrix}
\]

Then this squared Euclidean distance is also the squared Euclidean distance over the block-fading channel shown below:

\[
y = \theta \hat{h}^T \begin{bmatrix} \ell \\ \phi(\ell) \\ \cdots \\ \phi^{B-1}(\ell) \\ \end{bmatrix} + w.
\]

Note that since

\[
|\hat{h}|^2 = |\hat{h}'|^2
\]
both channels are in outage for precisely the same set of values of the fading coefficients $h_{sd}$, $h_{rd}$. On the other hand by our results in Section 3.2.3, the block diagonal code appearing in the equation above has negligible error probability when the channel is not in outage. This proves DMT optimality of the code in the cooperation phase as well.

Acknowledgement

The authors would like to thank Hsiao-feng (Francis) Lu and Jean-Claude Belfiore for useful discussions.

References

[1] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, “Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior,” IEEE Trans. Info. Theory, vol. 50, no. 12, pp. 3062-3080, Dec. 2004.

[2] K. Azarian, H. El Gamal, and P. Schniter, “On the Achievable Diversity-Multiplexing Tradeoff in Half-Duplex Cooperative Channels,” IEEE Trans. Info. Theory, vol. 51, no. 12, pp. 4152–4172, Dec. 2005.

[3] J. N. Laneman and G. W. Wornell, “Distributed Space–Time-Coded Protocols for Exploiting Cooperative Diversity in Wireless Networks,” IEEE Trans. Info. Theory, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[4] Y. Jing and B. Hassibi, “Distributed space-time coding in wireless relay networks - Part I: basic diversity results,” submitted to IEEE Trans. Wireless Comm., 2004.

[5] S. Yang and J.-C. Belfiore, “On Slotted Amplify-and-Forward Cooperative Diversity Schemes,” in Proc. IEEE Int. Symp. Info. Theory (ISIT 2006), pp. 2446, July 9–14, Seattle 2006.

[6] S. Yang and J.-C. Belfiore, “Optimal Space-Time Codes for the MIMO Amplify-and-Forward Cooperative Channel,” submitted to IEEE Trans. Wireless Comm., Sept. 2005. Available: http://arxiv.org/pdf/cs.IT/0509006

[7] S. Yang and J.-C. Belfiore, “Towards the Optimal Amplify-and-Forward Cooperative Diversity Scheme,” submitted to IEEE Trans. Wireless Comm., Mar. 2006. Available: http://arxiv.org/pdf/cs.IT/0603123
[8] R. U. Nabar, H. Bölcskei, and F. W. Kneubühler, “Fading Relay Channels: Performance Limits and Space-Time Signal Design,” *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.

[9] A. Sendonaris, E. Erkip, and B. Aazhang, “User Cooperation Diversity–Part I: System Description,” *IEEE Trans. Commun.*, vol. 51, no.11, pp. 1927–1938, Nov. 2003.

[10] A. Sendonaris, E. Erkip, and B. Aazhang, “User Cooperation Diversity–Part II: Implementation Aspects and Performance Analysis,” *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

[11] M. Yuksel and E. Erkip, “Cooperative Wireless Systems: A Diversity-Multiplexing Tradeoff Perspective,” submitted to *IEEE Trans. Info. Theory*, Sept. 2006. Available: http://arxiv.org/pdf/cs.IT/0609122

[12] F. Oggier, G. Rekaya, J. C. Belfiore, E. Viterbo, “Perfect space-time block codes,” *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 3885–3902, Sep. 2006.

[13] P. Elia, K. Raj Kumar, S. A. Pawar, P. Vijay Kumar, and H-F. Lu, “Explicit, minimum-delay space-time codes achieving the diversity-multiplexing gain tradeoff,” *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 3869–3884, Sep. 2006.

[14] P. Elia and P. Vijay Kumar, “Explicit unified D-MG optimal construction for the dynamic decode-and-forward cooperative wireless network,” Proceedings. 43nd Allerton conf. on Comm., Control & Computing, Sep. 2006.

[15] L. Zheng and D. Tse, “Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels,” *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[16] B. A. Sethuraman and B. Sundar Rajan and V. Shashidhar, “Full-diversity, high-rate, space-time block codes from division algebras,” *IEEE Trans. Inform. Theory*, vol. 49, pp. 2596–2616, Oct. 2003.

[17] Kiran.T. and B.Sundar Rajan, “STBC-schemes with non-vanishing determinant for certain number of transmit antennas,” *IEEE Trans. Inform. Theory*, vol. 51, no. 8, pp. 2984–2992, Aug. 2005.

[18] J.-C. Belfiore and G. Rekaya, “Quaternionic lattices for space-time coding,” *Proc. IEEE Information Theory Workshop (ITW 2003)*, Paris, Apr. 2003.
[19] J.-C. Belfiore, G. Rekaya and E. Viterbo, “The Golden code: a $2 \times 2$ full-rate space-time code with non-vanishing determinants,” Proc. IEEE Int. Symp. Inform. Th (ISIT 2004).

[20] S. Yang, J.-C. Belfiore and G. Rekaya-Ben Othman, “Perfect space-time block codes for parallel MIMO channels,” Proc. IEEE Int. Symp. Inform. Th (ISIT 2006).

[21] Hsiao-feng (Francis) Lu, “Explicit Construction of Multi-Block Space-Time Codes that Achieve the Diversity-Multiplexing Gain Tradeoff,” Proc. IEEE Int. Symp. Inform. Th (ISIT 2006).

[22] Hsiao-feng (Francis) Lu, “Optimal Code Constructions for SIMO-OFDM Frequency Selective Fading Channels, Proc. IEEE Information Theory Workshop, July 1-6, Bergen, Norway, 2007.

[23] S. A. Pawar, K. Raj Kumar, P. Elia, B. A. Sethuraman and P. Vijay Kumar, “Minimum-delay space-time codes achieving the DMD tradeoff of the MIMO-ARQ channel,” Submitted to IEEE Trans. Inform. Theory, Oct. 2005.

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

[25] P. Coronel and H. Bölcskei, “Diversity-Multiplexing Tradeoff in Selective-Fading MIMO Channels,” Proc. International Symp. Inform. Th., Nice, France, June 24-29, 2007.

[26] Sheng Yang and Jean-Claude Belfiore, “Optimal Space-Time Codes for the MIMO Amplify-and-Forward Cooperative Channel,” IEEE Trans. Inform. Theory, vol. 53, pp.647-663, Feb. 2007.

[27] H. El Gamal, G. Caire and M.O. Damen, “Lattice coding and decoding achieve the optimal diversity-multiplexing tradeoff of MIMO channels,” IEEE Trans. Inform. Theory, vol. 50, no. 6, pp. 968–985, Jun. 2004.

[28] H. Yao, G.W. Wornell, “Structured space-time block codes with optimal diversity-multiplexing tradeoff and minimum delay,” GLOBECOM’03.IEEE, vol. 4, pp. 1941–1945, Dec. 2003.

[29] Genyuan Wang and Xiang-Gen Xia, “On optimal multi-layer cyclotomic space-time code designs,” IEEE Trans. Inform. Theory, vol. 51, no. 3, pp. 1102–1135, Mar. 2005.
[30] C. Köse and R. D. Wesel, “Universal space-time trellis codes,” *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2717–2727, Oct. 2003.

[31] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, 1999.