Asymptotically Optimal Cooperative Wireless Networks with Reduced Signaling Complexity

Petros Elia, Frédérique Oggier, and P. Vijay Kumar

Abstract—This paper considers an orthogonal amplify-and-forward (OAF) protocol for cooperative relay communication over Rayleigh-fading channels in which the intermediate relays are permitted to linearly transform the received signal and where the source and relays transmit for equal time durations. The diversity-multiplexing gain (D-MG) tradeoff of the equivalent space-time channel associated to this protocol is determined and a cyclic-division-algebra-based D-MG optimal code constructed. The transmission or signaling alphabet of this code is the union of the QAM constellation and a rotated version of QAM. The size of this signaling alphabet is small in comparison with prior D-MG optimal constructions in the literature and is independent of the number of participating nodes in the network.

Index Terms—cooperative diversity, distributed space-time code, orthogonal amplify and forward, diversity-multiplexing gain tradeoff, space-time codes, cyclic division algebra codes.

I. INTRODUCTION

IN WIRELESS communications networks with fading, cooperative diversity protocols seek to provide MIMO-diversity benefits without requiring multiple transmit or receive antennas at any of the nodes in the network.

A. Existing Cooperative Diversity Schemes

Several cooperative diversity protocols were recently presented (for example in [4]-[9]). These can be separated into two main categories: the *amplify-and-forward* category, where the assisting nodes perform a linear operation on the signal vector they receive from the information source, before forwarding it, and the *decode-and-forward* category, where the assisting nodes try to decode the received signal, and eventually re-encode it before sending it again. While the error performance analysis in [7] focused on the diversity of the protocols, and the analysis in [9] focused on the capacity of the network, the works in [4], [5], [6] applied the diversity-multiplexing gain (D-MG) tradeoff as a means of evaluating the fundamental limitations in the error performance of the different cooperation protocols, in a similar manner that the same tradeoff was used by Zheng and Tse for the point to point MIMO case [13]. D-MG analysis also led to the introduction of the orthogonal *selection*-decode-and-forward

(O-SDAF, [5]) protocol, which asks the users to adapt to the channel's outage behavior. D-MG analysis also highlighted the benefits of the *non-orthogonal* amplify-and-forward (NAF) scheme proposed in [8] and generalized in [6], where the protocol's D-MG performance was established to be better than that of the O-SDAF and of the orthogonal amplify-and-forward (OAF) protocols, mainly due to the fact that it had the source transmitting continuously. As a consequence, the NAF protocol allowed for a non-zero diversity gain, even for the highest multiplexing gain regions.

The optimal D-MG performance of the protocols was established in the early papers through random coding arguments and the D-MG optimal implementation of the above O-SDAF and NAF protocols was thought to require infinite time duration and infinite decoding complexity.

Recently however, in ([10], [11]), D-MG optimal implementations of the OAF and NAF cooperation protocols were explicitly constructed to meet the corresponding protocol's high-SNR outage region, for any network size, in finite time duration with finite sphere-decoding complexity.

In the distributed space-time codes designed for cooperative relay communication, the signals transmitted by the source and intermediate relays are typically drawn from a common alphabet $\mathcal{S}$ apart from scale factors that adjust for transmission power. We will refer to $\mathcal{S}$ as the signaling alphabet of the distributed space-time code (DSTC).

Definition 1: A DSTC operating at R bits per channel use (bpcu) is said to have signaling set complexity C_s, if the cardinality of its signaling set $\mathcal{S}$ is:

$$|\mathcal{S}| = \kappa_s (2^R)^{C_s} \qquad (1)$$

where κ_s is a constant independent of the rate R.

In this work we present a reduced signaling complexity, D-MG optimal implementation of an OAF protocol (see [7], [10]) in which both source and relays transmit for equal durations of time. This implementation maintains a signaling complexity $C_s = 2$ that is independent of the number of users in the network, thereby avoiding the exponential increase in signaling alphabet incurred by other D-MG optimal constructions found in the literature, for example those found in [11] and [10]. This is discussed in greater detail in Section II-B.

Our code construction, like other prior constructions [11] and [10], is based cyclic division algebras (CDA). A key difference in our construction is that we make use of the matrix representation of a subset $\mathcal{F}$ of elements in the CDA which commute under multiplication. It turns out that the corresponding matrix representations can be simultaneously diagonalized.

Section II presents the network model we consider as well as the equivalent channel seen by the DSTC. An upper
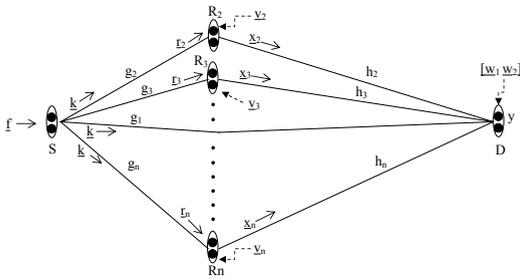
Fig. 1. Snapshot of a wireless network, where source terminal S utilizes its peers $(R_2, \cdots, R_n)$ for communicating with destination D.

bound on the D-MG tradeoff of the cooperative relay network operating under the OAF protocol described here is then presented. Section III shows that the upper bound on D-MG tradeoff derived in II-C is tight by presenting the construction of a CDA-based DSTC whose probability of error performance achieves the upper bound. Most proofs are found in the Appendix.

II. NETWORK AND CHANNEL MODELS

A. Network Model

As in [4], the network consists of a set

$$\mathcal{R} = \{R_1, R_2, \cdots, R_n, R_{n+1}\}$$

of $n+1$ different cooperating terminals/relays (see Figure 1), each with the ability to communicate over $n+1$ different orthogonal frequencies $\mathcal{F} = \{\nu_1, \nu_2, \cdots, \nu_n, \nu_{n+1}\}$. A certain relay R_i, wishing to communicate with relay $d(R_i)$, broadcasts its information over frequency ν_i. Depending on the availability of each intermediate relay, the set

$$D(R_i) \subset \{\mathcal{R} \setminus \{R_i \cup d(R_i)\}\} \qquad (2)$$

is then the set of all intermediate relays that cooperate with R_i. Consequently, each relay $R_j \in D(R_i)$ transmits a possibly modified version of the received signal over frequency ν_i. By the end of the transmission, $d(R_i)$ has received the information from R_i over frequency ν_i, essentially in a form of a superposition of faded versions of signals originating from R_i and from $D(R_i)$.

We consider the case where communication takes place in the presence of additive receiver noise, and in the presence of spatially independent quasi-static fading. Furthermore, we will assume complete knowledge of the fading channel at the receiver of the final destination, and no knowledge of the fading at the receivers of the assisting relays.

a) Assumptions: The overall rate of information transmission is R bpcu, which specifically means that for some time duration of T channel uses, from the beginning to the end of the communication between the source and the destination, the total number of information bits received by the destination (correctly or incorrectly), is RT.

It is assumed that each node has a single receive-transmit antenna operating under the half-duplex constraint, being able to either transmit or receive but not both simultaneously. This is due to practical considerations such as the large ratio between the transmission and reception powers at the relay antennas ([4], [5], [6]). Furthermore, no feedback is permitted to any of the transmitters.

All channels are assumed to be Rayleigh fading and all fade coefficients are assumed to be i.i.d., circularly-symmetric, complex Gaussian $\mathbb{CN}(0,1)$ random variables, i.e., they are i.i.d. with common density function

$$p(u) = \frac{1}{\pi} e^{-|u|^2}.$$

It is also assumed that all fading coefficients remain fixed for the duration of communication, i.e., that encoding over multiple channel realizations is not permitted.

The noise vector at the receivers is assumed to be comprised of i.i.d., circularly-symmetric, complex Gaussian $\mathbb{CN}(0,1)$ random variables as well.

b) Notation: The notation $\doteq$ and $\dot{\leq}, \dot{\geq}$ corresponds to the exponential equality and inequalities describing the equivalence of $y \doteq \rho^x$ to $\lim_{\rho \to \infty} \frac{\log(y)}{\log(\rho)} = x$. Matrices, vectors, and scalars are respectively denoted by capital letters, underlined small letters, and small letters. x^* represents the complex conjugate of a scalar x, and $X^\dagger$ represents the conjugate transpose of some matrix X. $\|X\|_F^2$ represents the Frobenius norm of X, $|\underline{x}|^2$ the square of the magnitude of some vector $\underline{x}$, and $|x|^2$ denotes the square of the magnitude of some scalar x. For $\mathcal{Y}$ a set, $|\mathcal{Y}|$ is its cardinality. Furthermore, if $\mathcal{Y}$ is a set of scalars, vectors or matrices with entries from the complex numbers, $\Delta \mathcal{Y}$ denotes the set of all differences of such elements, where the difference is taken in a component-wise manner. The symbol $\mathbb{Z}$ represents the sets of integers, $\mathbb{Q}$ represents the rationals, and $\iota := \sqrt{-1}$.

c) Performance Measure: Performance of the DSTC constructed here will be given in the form of the diversity-multiplexing gain (D-MG) tradeoff. In a network where each user operates at rate R bpcu and at SNR ρ, performance is described in terms of the diversity gain

$$d(r) := - \lim_{\rho \to \infty} \log(\Pr(e)) / \log(\rho)$$

which is a function of the multiplexing gain

$$r := R / \log(\rho).$$

B. Channel Model Under the OAF protocol

Under this OAF protocol, communication takes place in two phases. In the broadcast phase, Phase I, lasting for n channel uses, the source broadcasts to the relays and destination. In the cooperation phase, Phase II, the relays broadcast to the destination, again for a duration of n channel uses. The source is assumed to remain silent in Phase II.

Let g_1 denote the fading coefficient from source S to destination D, $g_i, i = 2, 3, \ldots, n$ denote the fading coefficient between S and ith intermediate relay R_i and h_i the fading coefficient from R_i to D. As mentioned above, the h_i, g_i are assumed to be independent and have a $\mathbb{CN}(0,1)$ distribution. These fading coefficients remain constant throughout the transmission, and change in an i.i.d. manner for every new message vector.

The random vectors $\underline{w}_1$ and $\underline{v}_i$, $i = 2 \cdots, n$, represent the zero-mean, additive white Gaussian noise seen at the destination and the relays R_i in Phase I respectively. The noise vector at the destination in Phase II is denoted $\underline{w}_2$. The components w_{1j}, v_{ij}, w_{2j}, $j = 1, 2, \ldots, n$ are also assumed

to be i.i.d. $\mathbb{CN}(0,1)$. Under this OAF protocol, the source sequentially transmits during time $t = 1, 2, \cdots, n$, the n-length vector

$$\underline{s}^T = \theta \underline{f}^T = \theta [f_1 \ f_2 \ \cdots \ f_n] \tag{3}$$

which contains the information to be communicated to the destination. The f_i are drawn from some alphabet $\mathcal{A}$ and the scalar θ normalizes for energy.

The destination receives the n-length vector

$$\underline{r}_1^T = \theta g_1 \underline{f}^T + \underline{w}_1^T. \tag{4}$$

Each intermediate relay R_i, $i = 2, 3, \cdots, n$ receives the n-length vector

$$\underline{r}_i^T = \theta g_i \underline{f}^T + \underline{v}_i^T, \tag{5}$$

and then transmits

$$\underline{r}_i^T A_i, \tag{6}$$

where the A_i, $2 \leq i \leq n$ are chosen to be $n \times n$ matrices satisfying the Frobenius-norm constraint:

$$||A_j||_F^2 \ \leq \ \alpha^2 \ \dot{=} \ \rho^0, \tag{7}$$

for some $\alpha > 0$.

Consequently, the receiver up to time $t = 2n$ has received

$$[\underline{y}']^T \ = \ \theta \underline{z}^T X' + [\underline{w}']^T \tag{8}$$

where

$$\underline{z}^T \ = \ \begin{bmatrix} g_1 & g_2 h_2 & \cdots & g_n h_n \end{bmatrix}_{1 \times n} \tag{9}$$

$$[\underline{w}']^T \ = \ \begin{bmatrix} \underline{w}_1^T & \sum_{j=2}^n h_j \underline{v}_j^T A_j + \underline{w}_2^T \end{bmatrix}_{1 \times 2n}, \tag{10}$$

$$X' \ = \ \begin{bmatrix} \overbrace{\underline{f}^T}^{} & \overbrace{0 \ \cdots \ 0}^{n} \\ 0 \ \cdots \ 0 & \underline{f}^T A_2 \\ \vdots & \\ \underbrace{0 \ \cdots \ 0}_{n} & \underline{f}^T A_n \end{bmatrix}_{n \times 2n} . \tag{11}$$

Taking the transpose of both sides of equation (8) leads to

$$\underbrace{\begin{bmatrix} \underline{r}_1 \\ \underline{y}_1 \end{bmatrix}}_{\underline{y}'} = \underbrace{\begin{bmatrix} g_1 I_n \\ \sum_{j=2}^n g_j h_j A_j \end{bmatrix}}_{H} \underline{s} + \underbrace{\begin{bmatrix} \underline{w}_1 \\ \sum_{j=2}^n h_j A_j \underline{v}_j + \underline{w}_2 \end{bmatrix}}_{\underline{w}'} . \tag{12}$$

d) Signaling Complexity: One means of constructing a D-MG optimal DSTC in the setting of (12) is to employ a cyclic-division-algebra-based approximately-universal (see [14], [15]) space-time code $\mathcal{X}$ whose elements are $(n \times n)$ matrices X. If

$$X = \begin{bmatrix} \underline{x}_1 & \underline{x}_2 & \cdots & \underline{x}_n \end{bmatrix},$$

then one applies the OAF protocol to the column vectors $\underline{x}_i, i = 1, 2, \cdots, n$ in succession by setting $\underline{s} = \underline{x}_1$ followed by $\underline{s} = \underline{x}_2$ etc. Thus this communication would then take place over a total time duration of $n \times (2n) = 2n^2$ channel uses. It turns out the symbols of the matrix X would in such case, be drawn from an alphabet of size $(2^R)^{2n}$ where R is the desired rate of communication in bits per channel use. Thus the resultant signaling alphabet is of size exponential in the

number n of relays. From the definition of signaling complexity given in (1), it follows that the signaling complexity C_s for this coding scheme is equal to $2n$. The coding schemes presented in [10] and [11] (for the NAF protocol) similarly share signaling alphabets of size exponential in the number of relays.

In contrast, the coding scheme presented in the present paper has signaling complexity $C_s = 2$ independent of the number of cooperating relays, thus permitting the transmitted signal alphabet to be independent of the number of cooperating network nodes.

C. Upper Bound on the D-MG Tradeoff of the OAF Protocol

Lemma 2.1: The D-MG tradeoff of the above OAF protocol satisfies the upper bound:

$$d(r) \ \leq \ n(1 - 2r). \tag{13}$$

Proof: See Appendix I.   ∎

In the next section, we will identify a DSTC whose codeword error probability when employed over the OAF channel described by (8), satisfies

$$P_e(r) \ \dot{\leq} \ \rho^{-n(1-2r)}$$

thereby proving that the right hand side in equation (13) is indeed the D-MG tradeoff of the described OAF protocol.

III. A D-MG OPTIMAL CODE FOR THE OAF PROTOCOL

Our code construction is based on cyclic division algebras. Some background on these algebraic objects can be found in Appendix II and we will assume in the present section, that the reader is familiar with the terminology and notation introduced there.

A. Code Construction

For M even, let $\mathcal{A}_{\text{QAM}}$ denote the M^2-QAM constellation given by

$$\mathcal{A}_{\text{QAM}} = \{a + \imath b \ | \ |a|, |b| \leq M - 1, \ a, b \ \text{odd}\}. \tag{14}$$

Let $\mathcal{X}$ be the collection of matrices

$$\mathcal{X} \ = \ \{X(\underline{l}) \mid l_i \in \mathcal{A}_{\text{QAM}}\}, \tag{15}$$

where

$$X(\underline{l}) \ = \ \begin{bmatrix} l_0 & \gamma l_{n-1} & \cdots & \gamma l_1 \\ l_1 & l_0 & \cdots & \gamma l_2 \\ \vdots & l_1 & \cdots & \cdots \\ \vdots & \vdots & \vdots & \gamma l_{n-1} \\ l_{n-1} & l_{n-2} & \cdots & l_0 \end{bmatrix}$$

where $X(\underline{l})$ is the matrix representation of an element $l_0 + z l_1 + \cdots + z^{n-1} l_{n-1}$ belonging to a CDA and where γ is a non-norm element of unit magnitude $|\gamma| = 1$ belonging to $\mathbb{Q}(\imath)$, both as described in Appendix II. Often for the sake of convenience, we will abbreviate and write X in place of $X(\underline{l})$.

The matrix $X(\underline{l})$ can be expressed in the form

$$X(\underline{l}) \;=\; \begin{bmatrix} \underline{l}^T B_1 \\ \underline{l}^T B_2 \\ \vdots \\ \underline{l}^T B_n \end{bmatrix} \tag{16}$$

where

$$\underline{l} \;=\; [l_0\; l_1\; \cdots\; l_{n-1}]^T,$$

and where each matrix B_i has just one nonzero entry , either 1 or γ, in each row and column. Since $|\gamma| = 1$, it follows that each matrix B_i is unitary. Next, set

$$\underline{f}^T \;=\; \underline{l}^T B_1 \tag{17}$$
$$A_j \;=\; B_1^\dagger B_j, \quad 2 \le j \le n. \tag{18}$$

Then each matrix A_i is also unitary and we can write

$$X(\underline{l}) \;=\; \begin{bmatrix} \underline{f}^T \\ \underline{f}^T A_2 \\ \vdots \\ \underline{f}^T A_n \end{bmatrix}. \tag{19}$$

Since the elements l_i belong to the subset $\mathcal{A}_{\text{QAM}}$ of $\mathbb{Z}[\imath]$, the matrices in $\mathcal{X}$ are the regular representations of elements in the *center* $\mathbb{F} = \mathbb{Q}(\imath)$ of the CDA and hence commute under multiplication, thus meeting a condition needed for simultaneous diagonalizability. As shown in Appendix IV, $\mathcal{X}$ is indeed a collection of normal matrices that can be simultaneously diagonalized by a unitary transformation S, i.e., every matrix $X \in \mathcal{X}$ can be expressed in the form

$$X \;=\; S^\dagger X_d S$$
$$\Leftrightarrow\; X_d \;=\; SXS^\dagger$$

for a suitable diagonal matrix X_d.

e) Usage of the code $\mathcal{X}$ in the OAF protocol: The DSTC to be employed in the OAF protocol is built upon the matrices in $\mathcal{X}$. Let $\underline{l}$, $\underline{l} \in \mathcal{A}_{\text{QAM}}$, represent the message vector. As we shall see below, in effect, the signals transmitted by the various relays correspond to different rows of the corresponding code matrix

$$X(\underline{l}) = \begin{bmatrix} l_0 & \gamma l_{n-1} & \dots & \gamma l_1 \\ l_1 & l_0 & \dots & \gamma l_2 \\ \vdots & l_1 & \dots & \dots \\ \vdots & \vdots & \vdots & \gamma l_{n-1} \\ l_{n-1} & l_{n-2} & \dots & l_0 \end{bmatrix}.$$

The source maps $\underline{l}$ onto $\underline{f}^T = \underline{l}^T B_1$ and transmits $\underline{f}^T$. The jth relay, $j = 2, 3, \ldots, n$ then applies the linear transformation represented by the unitary matrix A_j to the received vector $\underline{r}_j$.

The received signal at the destination is then given by:

$$\begin{bmatrix} \underline{r}_1^T & \underline{y}_1^T \end{bmatrix} \;=\; \theta \underline{z}^T \begin{bmatrix} \underline{f}^T & \underline{0}^T \\ \underline{0}^T & \underline{f}^T A_2 \\ \underline{0}^T & \underline{f}^T A_3 \\ \vdots & \vdots \\ \underline{0}^T & \underline{f}^T A_n \end{bmatrix}$$
$$+\; \begin{bmatrix} \underline{w}_1^T & \underline{w}_2^T + \sum_{j=2}^n h_j \underline{v}_j^T A_j \end{bmatrix} \tag{20}$$

where $\underline{z}^T$ is as defined in (9).

B. Proof of D-MG Optimality

Theorem 3.1 (Main Theorem): The D-MG tradeoff of the OAF protocol described here is given by

$$d(r) \;=\; n(1 - 2r) \tag{21}$$

and the CDA-based DSTC described in (15) and used as suggested by equations (19) and (20) is D-MG optimal with respect to this tradeoff.

Proof: Let $P_e(r)$ denote the probability of error of the CDA-based code when used over the OAF channel in accordance with (19) and (20). Our aim is to show that

$$P_e(r) \;\dot{\le}\; \rho^{-n(1-2r)}. \tag{22}$$

By Lemma 2.1, this will not only establish D-MG optimality of the code, it will also establish that the D-MG tradeoff of this OAF channel is precisely given by the right hand side of (13), thereby proving the theorem.

We begin by outlining a decoding procedure to be followed at the destination and showing that this decoding procedure has probability of codeword error upper bounded by (22).

Given received vector $\begin{bmatrix} \underline{r}_1^T & \underline{y}_1^T \end{bmatrix}$ as in (20), we instruct the receiver to add the two halves of the vector leading to

$$\begin{bmatrix} \underline{r}_1^T + \underline{y}_1^T \end{bmatrix} \;=\; \theta \underline{z}^T \begin{bmatrix} \underline{f}^T I_n \\ \underline{f}^T A_2 \\ \underline{f}^T A_3 \\ \vdots \\ \underline{f}^T A_n \end{bmatrix}$$
$$+\; \begin{bmatrix} \underline{w}_1^T + \underline{w}_2^T + \sum_{j=2}^n h_j \underline{v}_j^T A_j \end{bmatrix} \tag{23}$$

which with

$$X = \begin{bmatrix} \underline{f}^T \\ \underline{f}^T A_2 \\ \vdots \\ \underline{f}^T A_n \end{bmatrix} = \begin{bmatrix} \underline{l}^T B_1 \\ \underline{l}^T B_2 \\ \vdots \\ \underline{l}^T B_n \end{bmatrix} \tag{24}$$

can be rewritten in the form

$$\begin{bmatrix} \underline{r}_1^T + \underline{y}_1^T \end{bmatrix} \;=\; \theta \underline{z}^T X$$
$$+\underline{w}_1^T + \underline{w}_2^T + \sum_{j=2}^n h_j \underline{v}_j^T A_j . \tag{25}$$

Let us denote the noise vector appearing above by

$$\underline{n}' \;:=\; \underline{w}_1 + \underline{w}_2 + \sum_{j=2}^n h_j A_j^T \underline{v}_j.$$

The covariance of this noise vector is then given by

$$\mathbb{E}\left\{[\underline{n}'][\underline{n}']^\dagger\right\} = \left[2 + \sum_{j=2}^n |h_j|^2\right] I_n := \beta I_n \text{ (say)}.$$

Let $\underline{n}''$ denote the normalized noise vector

$$\underline{n}'' = \frac{1}{\sqrt{\beta}}\underline{n}'.$$

Then $\underline{n}''$ is white, distributed according to $\mathcal{CN}(0,1)$. Scaling both sides of (23) by $\sqrt{\beta}$ we obtain:

$$\frac{1}{\sqrt{\beta}}\left[\ \underline{r}_1^T + \underline{y}_1^T\ \right] = \frac{\theta}{\sqrt{\beta}}\underline{z}^T X + [\underline{n}'']^T. \qquad (26)$$

Making use of the fact that the collection of code matrices $\{X \in \mathcal{X}\}$ is a family of commuting normal matrices that can be simultaneously diagonalized as $X = S^\dagger X_d S$, we have:

$$\frac{1}{\sqrt{\beta}}\left[\ \underline{r}_1^T + \underline{y}_1^T\ \right] = \frac{\theta}{\sqrt{\beta}}\underline{z}^T S^\dagger X_d S + [\underline{n}'']^T.$$

Post-multiplying by $S^\dagger$ yields

$$\frac{1}{\sqrt{\beta}}\left[\ \underline{r}_1^T + \underline{y}_1^T\ \right] S^\dagger = \frac{\theta}{\sqrt{\beta}}\underline{z}^T S^\dagger X_d + [\underline{n}'']^T S^\dagger.$$

Since this is a unitary transformation, the transformed noise vector $[\underline{n}]^T = [\underline{n}'']^T S^\dagger$ remains white:

$$\underline{y}^T = \frac{\theta}{\sqrt{\beta}}\underline{q}^T X_d + \underline{n}^T \qquad (27)$$

where we have set

$$\underline{y}^T = \frac{1}{\sqrt{\beta}}\left[\ \underline{r}_1^T + \underline{y}_1^T\ \right] S^\dagger, \text{ and}$$

$$\underline{q}^T := [q_1, \cdots, q_n] = \underline{z}^T S^\dagger.$$

Equation (27) can now be expressed in the form of an equation for a parallel channel [14]:

$$\underline{y} = \frac{\theta}{\sqrt{\beta}}\begin{bmatrix} q_1 & & & \\ & q_2 & & \\ & & \ddots & \\ & & & q_n \end{bmatrix}\begin{bmatrix} X_{d,1} \\ X_{d,2} \\ \vdots \\ X_{d,n} \end{bmatrix} + \underline{n}, (28)$$

where $X_{d,i}$ are the diagonal elements of X_d.

The probability of outage of this parallel channel is given by

$$P_{\underline{q},\text{out}}(r) = \Pr\left\{\prod_{j=1}^n \frac{(1+\rho|q_j|^2)}{2+\sum_{i=2}^n |h_i|^2} < \rho^{2rn}\right\}$$

$$= \Pr\left\{\prod_{j=1}^n (1+\rho|q_j|^2) < \rho^{2rn}(2+\sum_{i=2}^n |h_i|^2)^n\right\}.$$

Since the probability that

$$|h_i|^2 \doteq \rho^\epsilon, \ \epsilon > 0$$

vanishes exponentially fast with ρ, it can be shown that

$$P_{\underline{q},\text{out}}(r) \doteq \Pr\left\{\prod_{j=1}^n (1+\rho|q_j|^2) < \rho^{2rn}\right\}.$$

In Appendix III it is shown that

$$P_{\underline{q},\text{out}}(r) \ \dot{\le}\ \rho^{-n(1-2r)}. \qquad (29)$$

A space-time code is said to be approximately universal [14] if it is D-MG optimal for every statistical characterization of the matrix of channel fading coefficients. We will make use of the following sufficient criterion for approximate universality over the parallel channel:

Theorem 3.2: [14] Consider a parallel channel given by

$$\underline{y} = \begin{bmatrix} q_1 & & & \\ & q_2 & & \\ & & \ddots & \\ & & & q_n \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \underline{n}, \qquad (30)$$

where q_i represent the fading coefficients and where the components of $\underline{n}$ are i.i.d. and $\mathcal{CN}(0,1)$ distributed. Then a space-time code $\mathcal{X}$ with code matrices X

$$X = \begin{bmatrix} \underline{x}_1^T \\ \underline{x}_2^T \\ \vdots \\ \underline{x}_n^T \end{bmatrix}$$

for this parallel channel is approximately universal iff the product of the row-norms of every code-difference matrix satisfies

$$\prod_{j=1}^n ||\underline{x}_j^T||^2 \ \dot{\ge}\ \rho^{n-r_p}$$

where ρ is the SNR and $r_p \log_2(\rho)$ is the rate of the code over the parallel channel in bpcu.

The columnar version of the diagonal code

$$\mathcal{X}_d = \left\{X_d = SXS^\dagger \mid X \in \mathcal{X}\right\}$$

turns out to satisfy this criterion of approximate universality. We regard the matrix

$$\frac{1}{\sqrt{\beta}}\begin{bmatrix} q_1 & & & \\ & q_2 & & \\ & & \ddots & \\ & & & q_n \end{bmatrix}$$

as representing the channel matrix of the parallel channel. Each code matrix consists of a single column $\theta\tilde{X}_d$ whose entries are the diagonal elements of the matrices θX_d. Let ΔX_d, $\Delta\tilde{X}_d$ denote the difference matrix in the diagonal and columnar versions respectively. Then the product of the squares of the row norms of $\theta\Delta\tilde{X}_d$ is given by

$$\prod_{i=1}^n |[\theta\Delta X_d]_{ii}|^2 = \theta^{2n}|\det\Delta X|^2. \qquad (31)$$

The code $\mathcal{X}$ must be of size ρ^{2rn} to deliver a rate of $r\log_2(\rho)$ bpcu over $2n$ uses of the MISO channel given by (20). Since

each code matrix represents the information content of n QAM symbols $(l_1, \cdots, l_n)$, it follows that the size M^2 of the QAM constellation must satisfy:

$$[M^2]^n \ \doteq \ \rho^{2rn}$$

i.e., $M^2 = \rho^{2r}$. Indeed this is the reason why the signaling complexity of the code constructed here equals 2. To ensure that the average energy per symbol equals ρ we must have

$$\theta^2 M^2 \ \doteq \ \rho \ \Rightarrow \ \theta^2 \ \doteq \ \rho^{1-2r}.$$

Since each code matrix X is drawn from a CDA-based code, its determinant is bounded below from (34) by

$$|\det(X)| \ = \ |\det(X_d)| \ = \ \prod_{i=1}^{n} |[X_d]_{ii}| \ \geq \ \frac{1}{|b|} \ \doteq \ \rho^0.$$

The same lower bound also applies to the magnitude of the determinant of the difference of two code matrices. It follows that we can lower bound the product of the row norms by

$$\prod_{i=1}^{n} |\,[\theta \Delta X_d]_{ii}\,|^2 \ = \ \theta^{2n}\,|\det \Delta X\,|^2$$
$$\dot{\geq} \ \rho^{n(1-2r)} \ = \ \rho^{n-2rn}.$$

From Theorem 3.2 with $r_p = 2rn$, it follows that the diagonal code when transmitting at rate $2rn$ is approximately universal for the parallel channel and therefore has probability of error equal to the probability of outage which from (29) is given by

$$P_e(r) \ \doteq \ P_{\text{out}}(r) \ \dot{\leq} \ \rho^{-n(1-2r)}.$$

However, the outage probability of the equivalent space-time channel formed by the cooperating relay network in conjunction with our DSTC is lower bounded from Lemma 2.1, by $\rho^{-n(1-2r)}$. This not only proves optimality of the DSTC constructed here, it also establishes that the outage probability bound derived in Lemma 2.1 is in fact the true value of outage probability. ∎

APPENDIX I
PROOF OF UPPER BOUND ON D-MG TRADEOFF
(LEMMA 2.1)

Our starting point is (12). We write the covariance matrices appearing in (12) as:

$$\Sigma_{\underline{w}'} \ = \ \mathbb{E}(\underline{w}'\underline{w}'^\dagger) = \begin{bmatrix} I_n & 0 \\ 0 & I_n + \sum_{j=2}^{n} |h_j|^2 A_j A_j^\dagger \end{bmatrix},$$
$$\Sigma_s \ = \ \mathbb{E}(\underline{s}\underline{s}^\dagger) \quad \text{with} \quad \text{Tr}(\Sigma_s) = n\rho$$

and we note that the mutual information

$$I(\underline{y}'; \underline{s}) = \log |I_{2n} + H\Sigma_s H^\dagger \Sigma_{\underline{w}'}^{-1}|$$

is bounded as [13]

$$\log |I_{2n} + \rho HH^\dagger \Sigma_{\underline{w}'}^{-1}| \leq \sup_{\text{Tr}(\Sigma_s) \leq n\rho} I(\underline{y}'; \underline{s})$$
$$\leq \log |I_{2n} + n\rho HH^\dagger \Sigma_{\underline{w}'}^{-1}|.$$

In the high-SNR scale of interest, the two bounds are essentially the same and hence we assume below that

$$I(\underline{y}'; \underline{s}) = \log |I_{2n} + \rho HH^\dagger \Sigma_{\underline{w}'}^{-1}|$$

where

$$HH^\dagger \ = \ \begin{bmatrix} g_1 I_n \\ B \end{bmatrix} \begin{bmatrix} g_1^* I_n & B^\dagger \end{bmatrix} = \begin{bmatrix} |g_1|^2 I_n & g_1 B^\dagger \\ g_1^* B & BB^\dagger \end{bmatrix}$$

and where $B := \sum_{j=2}^{n} g_j h_j A_j$. We have

$$HH^\dagger \Sigma_{\underline{w}'}^{-1} = \begin{bmatrix} |g_1|^2 I_n & g_1 B^\dagger \\ g_1^* B & BB^\dagger \end{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & C^{-1} \end{bmatrix}$$

where $C = I_n + \sum_{j=2}^{n} |h_j|^2 A_j A_j^\dagger$. Furthermore

$$I_{2n} + \rho HH^\dagger \Sigma_{\underline{w}'}^{-1} \ = \ \begin{bmatrix} I_n(1+\rho\gamma_1) & \rho g_1 B^\dagger C^{-1} \\ \rho g_1^* B & I_n + \rho BB^\dagger C^{-1} \end{bmatrix}$$

$$\Rightarrow \ \begin{bmatrix} I_n(1+\rho\gamma_1) & \rho g_1 B^\dagger C^{-1} \\ 0 & I_n + \frac{\rho}{1+\rho\gamma_1} BB^\dagger C^{-1} \end{bmatrix}$$

upon row reduction (by multiplying the first row by $\frac{\rho g_1^*}{1+\rho\gamma_1} B$ and subtracting from the second row). Here we have set $|g_1|^2 = \gamma_1$ and $\gamma_j = |g_j h_j|^2$, $2 \leq j \leq n$.

Let

$$\mathcal{I} = |I_{2n} + \rho HH^\dagger \Sigma_{\underline{w}'}^{-1}|.$$

Then

$$\mathcal{I} = |I_n(1+\rho\gamma_1)| \cdot |I_n + \frac{\rho}{1+\rho\gamma_1} BB^\dagger C^{-1}|$$
$$= (1+\rho\gamma_1)^n |C^{-1}| \cdot |C + \frac{\rho}{1+\rho\gamma_1} BB^\dagger|$$
$$\leq (1+\rho\gamma_1)^n \cdot |I_n + \sum_{j=2}^{n} |h_j|^2 A_j A_j^\dagger|^{-1} \cdot$$
$$\cdot |I_n + \sum_{j=2}^{n} (|h_j|^2 + \frac{\rho}{1+\rho\gamma_1} |g_j h_j|^2) A_j A_j^\dagger|$$

where we have used the fact that by the Cauchy-Schwarz inequality,

$$\sum_{j,k} c_j c_k^* A_j A_k^\dagger \ \dot{\leq} \ \sum_{j} |c_j|^2 A_j A_j^\dagger$$

as non-negative definite (n.n.d.) matrices and

$$B \geq A \Rightarrow \log |I + B| \geq \log |I + A|$$

when A, B are n.n.d. matrices.

By the Frobenius-norm constraint, have

$$\sum_{j} |c_j|^2 A_j A_j^\dagger \leq (\sum_{j} |c_j|^2)\alpha^2 I_n.$$

Putting these inequalities together gives us

$$\mathcal{I} \ \dot{\leq} \ (1 + \rho\gamma_1 + \rho \sum_{j=2}^{n} \gamma_j)^n.$$

We here note that the outage region is defined by the set of all channel realizations under which

$$\log \mathcal{I} < 2rn \log(\rho).$$

We define $\{u_j, v_j\}$ by

$$\gamma_1 \ \doteq \ \rho^{-p_1}, \ \gamma_j \doteq \rho^{-(u_j + v_j)}, \ 2 \leq j \leq n, \ p_j = u_j + v_j.$$

We thus have the equivalent formulation of the outage region

$$\mathcal{O} = \Big\{ \{p_j\} \ : \ n\big\{ \max_{1 \leq j \leq n} (1 - p_j)^+ \big\} < 2rn \Big\}.$$

This means that for $d_{\text{out}}(r) := -\lim_{\rho \to \infty} \log(Pr(\log \mathcal{I}))/\log(\rho)$ then

$$d_{\text{out}}(r) = \inf \sum\nolimits_{j=1}^{n} p_j$$

where the infimum is taken over the outage region, i.e., in the region over which

$$n \min\{p_j\}_{j=1}^{n} > n - 2nr.$$

Consequently, $d_{\text{out}}(r) = n(1 - \frac{r2n}{n})$ and since we have an upper bound on the mutual information, it is the case that we have an upper bound on the optimal diversity gain. This bound is given by

$$d_{\text{out}}(r) \leq n(1 - 2r).$$

APPENDIX II
BACKGROUND ON CYCLIC DIVISION ALGEBRAS

A. The General CDA

Division algebras are rings with identity in which every nonzero element has a multiplicative inverse. The center $\mathbb{F}$ of any division algebra D, i.e., the subset comprising of all elements in D that commute with every element of D, is a field. The division algebra is a vector space over the center $\mathbb{F}$ of dimension n^2 for some integer n. A field $\mathbb{L}$ such that $\mathbb{F} \subset \mathbb{L} \subset D$ and such that no subfield of D contains $\mathbb{L}$ is called a *maximal subfield* of D. Every division algebra is also a vector space over a maximal subfield and the dimension of this vector space is the same for all maximal subfields and equal to n. This common dimension n is known as the *index* of the division algebra.

Division algebras in which the center $\mathbb{F}$ and a maximum subfield $\mathbb{L}$ are such that $\mathbb{L}/\mathbb{F}$ is a (finite) cyclic (Galois) extension are called Cyclic Division Algebras (CDA). CDAs have a simple characterization that aids in their construction, see [20], Proposition 11 of [21], or Theorem 1 of [19].

Let $\mathbb{F}$, $\mathbb{L}$ be number fields, with $\mathbb{L}$ a finite, cyclic Galois extension of $\mathbb{F}$ of degree n. Let σ denote the generator of the Galois group $\text{Gal}(\mathbb{L}/\mathbb{F})$. Let z be an indeterminate satisfying

$$\ell z \ = \ z \sigma(\ell) \quad \forall \ \ell \in \mathbb{L} \quad \text{and} \quad z^n = \gamma,$$

for some non-norm element $\gamma \in \mathbb{F}^*$, by which we mean some element γ having the property that the smallest positive integer t for which γ^t is the relative norm $N_{\mathbb{L}/\mathbb{F}}(u)$ of some element u in $\mathbb{L}^*$, is n. Then a CDA $D(\mathbb{L}/\mathbb{F}, \sigma, \gamma)$ with index n, center $\mathbb{F}$ and maximal subfield $\mathbb{L}$ is the set of all elements of the form

$$\sum_{i=0}^{n-1} z^i \ell_i, \quad \ell_i \in \mathbb{L}. \tag{32}$$

Moreover it is known that every CDA has this structure. It can be verified that D is a right vector space (i.e., scalars multiply vectors from the right) over the maximal subfield $\mathbb{L}$.

The matrix corresponding to an element $d \in D$ corresponds to the left multiplication by the element d in the division algebra. Let λ_d denote this operation, $\lambda_d : D \to D$, defined by

$$\lambda_d(e) = de, \ \forall \ e \in D.$$

It can be verified that λ_d is a $\mathbb{L}$-linear transformation of D. From (32), a natural choice of basis for the right-vector space D over $\mathbb{L}$ is $\{1, z, z^2, \ldots, z^{n-1}\}$. A typical element in the division algebra D is $d = \ell_0 + z\ell_1 + \cdots + z^{n-1}\ell_{n-1}$, where the $\ell_i \in \mathbb{L}$. By considering the effect of multiplying $d \times 1$, $d \times z$, $\ldots$, $d \times z^{n-1}$, one can show that the $\mathbb{L}$-linear transformation λ_d under this basis has the matrix representation

$$\begin{bmatrix} \ell_0 & \gamma\sigma(\ell_{n-1}) & \gamma\sigma^2(\ell_{n-2}) & \ldots & \gamma\sigma^{n-1}(\ell_1) \\ \ell_1 & \sigma(\ell_0) & \gamma\sigma^2(\ell_{n-1}) & \ldots & \gamma\sigma^{n-1}(\ell_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \ell_{n-1} & \sigma(\ell_{n-2}) & \sigma^2(\ell_{n-3}) & \ldots & \sigma^{n-1}(\ell_0) \end{bmatrix}, \tag{33}$$

known as the left regular representation of d. It is known that despite the fact that the entries of the left-regular representation belong to $\mathbb{L}$, the determinant of every such matrix lies in the subfield $\mathbb{F}$.

B. CDA Specific to the DSTC Construction

In the CDA relevant to the construction of the DSTC described in Section III, we choose

$$\mathbb{F} \ = \ \mathbb{Q}(\imath).$$

The field $\mathbb{L}$ can be taken to be any number field that is a cyclic Galois extension of $\mathbb{Q}(\imath)$ of degree n. A general construction, valid for any integer n, for such cyclic extensions can be found in [15]. The elements l_i are chosen to belong to $\mathcal{A}_{\text{QAM}} \subseteq \mathbb{Z}[\imath]$. A non-norm element $\gamma \in \mathbb{Q}(\imath)$ having unit magnitude, i.e., $|\gamma| = 1$, is chosen. Such a γ can always be found, see [17]. Under these conditions, it can be shown that the determinant $D(l(z))$ of the left-regular representation of an element

$$l(z) \ = \ \sum_{i=0}^{n-1} z^i l_i, \quad l_i \in \mathcal{A}_{\text{QAM}}$$

is of the form

$$D(\ l(z)\ ) \ = \ \frac{a(l(z))}{b}, \quad a(l(z)), b \ \in \mathbb{Z}[\imath]$$

where b is fixed and independent of the specific choice of elements $\{l_i\}$. This determinant is moreover, nonzero if any of the l_i is nonzero. As a result, we obtain the lower bound

$$|D(\ l(z)\ )| \ \geq \ \frac{1}{|b|}, \tag{34}$$

on the magnitude of the determinant, that applies to all $l(z))$ provided at least one $l_i \neq 0$.

We observe that the above remarks on the determinant also apply to the matrix that is the difference of the left-regular-representation of two distinct elements of the CDA, since the difference matrix is the left-regular representation of the difference element.

APPENDIX III
OUTAGE PROBABILITY OF THE PRODUCT-FADING,
PARALLEL CHANNEL

Our goal is to prove that the probability of outage of the parallel channel given by (28) satisfies

$$
\begin{aligned}
P_{\underline{q},\text{out}}(r) &= \Pr\left\{\prod_{j=1}^{n}\frac{(1+\rho|q_j|^2)}{2+\sum_{i=2}^{n}|h_i|^2} < \rho^{2rn}\right\} \\
&\dot{\leq} \rho^{-n(1-2r)}.
\end{aligned}
\tag{35}
$$

We begin with some preliminaries:

a) Density function of the product of two complex-Gaussian random variables: Let $z_1 = h_1$ and $z_j = g_j h_j$, $j = 2, \cdots, n$ where g_i, h_i are the fading coefficients introduced in Section II-B.

We will focus on the joint distribution of the real and imaginary parts of z_j, and for short, will write z, h, g in place of z_j, h_j, g_j. We have $h = h_R + \imath h_I, g = g_R + \imath g_I, z = z_R + \imath z_I$, with $z_R = h_R g_R - h_I g_I, z_I = h_R g_I + h_I g_R$.

The characteristic function $\phi_z(u)$ is given by:

$$
\begin{aligned}
\phi_{z_R,z_I}(u_R, u_I) &= \mathbb{E}_{g,h}\{e^{\imath[u_R z_R + u_I z_I]}\} \\
&= \mathbb{E}_{g,h}\{e^{\imath h_R[u_R g_R + u_I g_I] + \imath h_I[-u_R g_I + u_I g_R]}\} \\
&= \mathbb{E}_g\{e^{-\frac{1}{4}[(u_R g_R + u_I g_I)^2 + (-u_R g_I + u_I g_R)^2]}\} \\
&= \mathbb{E}_g\{e^{-\frac{1}{4}[|u|^2(g_R^2 + g_I^2)]}\} = (1 + |u|^2/4)^{-1}
\end{aligned}
$$

since

$$
\begin{aligned}
&\int e^{-\nu^2/(2\sigma^2)} e^{-\alpha\nu^2} (\sqrt{2\pi\sigma^2})^{-1} d\nu \\
&= \sqrt{2\pi/[2(\alpha + \frac{1}{2\sigma^2})]}/(\sqrt{2\pi\sigma^2}) = 1/\sqrt{1 + 2\alpha\sigma^2}.
\end{aligned}
$$

The characteristic function is directly related to the Fourier transform:

$$
\begin{aligned}
\phi_{z_R,z_I}(u_R, u_I) &= \mathbb{E}_z\{e^{\imath[u_R z_R + u_I z_I]}\} \\
&= \iint p_{z_R,z_I}(z_R, z_I) e^{\imath[u_R z_R + u_I z_I]} dz_R dz_I
\end{aligned}
$$

$$
\Rightarrow \phi_{z_R,z_I}(2\pi u_R, 2\pi u_I) = \mathcal{F}_{(u_R, u_I)}\{p_{z_R,z_I}(z_R, z_I)\}.
$$

The density function can thus be obtained by taking the Fourier inverse:

$$
\begin{aligned}
&p_{z_R,z_I}(z_R, z_I) \\
&= \iint \phi_{z_R,z_I}(2\pi u_R, 2\pi u_I) e^{-\imath[2\pi u_R z_R + 2\pi u_I z_I]} du_R du_I \\
&= \iint [1 + (2\pi)^2|u|^2/4]^{-1} e^{-\imath[2\pi u_R z_R + 2\pi u_I z_I]} du_R du_I \\
&= 2\pi \int \frac{1}{1 + (2\pi)^2|u|^2/4} J_0(2\pi|z||u|)|u|d|u| \\
&= 2\pi \int \frac{4}{4 + (2\pi)^2|u|^2} J_0(2\pi|z||u|)|u|d|u| \\
&= \frac{(4)}{(2\pi)^2} 2\pi \int \frac{1}{4 + s^2} J_0(2\pi\frac{|z|}{2\pi}s)s ds \\
&= \frac{4}{(2\pi)^2} 2\pi K_0(2\pi\frac{|z|}{2\pi}2) = \frac{4}{2\pi} K_0(2|z|)
\end{aligned}
$$

where $J_0(\cdot)$ is the Bessel function of the first kind and $K_0(\cdot)$ is the modified Bessel function of the second kind, both of order 0 [26]. We have used information about the Hankel transform taken from [24]. By independence of fading coefficients, we have that

$$
p_{Z_R,Z_I}(\underline{z}_R, \underline{z}_I) = \frac{1}{\pi} e^{-|z_1|^2} \prod_{j=2}^{n} \frac{4}{2\pi} K_0(2|z_j|).
$$

b) Change of Variables: Given a complex-valued vector $\underline{\mu}$, we define

$$
\hat{\underline{\mu}} = \begin{bmatrix} \underline{\mu}_R \\ \underline{\mu}_I \end{bmatrix}.
$$

Similarly given a complex-valued matrix S, we define

$$
\hat{S} = \begin{bmatrix} S_R & -S_I \\ S_I & S_R \end{bmatrix}.
$$

Since

$$
\underline{q}^T = \underline{z}^T S^\dagger \Rightarrow \underline{z} = S^T \underline{q}
$$

and S is unitary, it follows that

$$
\hat{\underline{z}} = \hat{S}^T \hat{\underline{q}}
$$

with $\hat{S}$ orthogonal and thus

$$
\begin{aligned}
p_Q(\underline{q}_R, \underline{q}_I) &= p_Z([S^T\underline{q}]_R, [S^T\underline{q}]_I) \\
&= \frac{1}{\pi} e^{-|\underline{s}_1^T\underline{q}|^2} \prod_{j=2}^{n} \frac{4}{2\pi} K_0(2|\underline{s}_j^T\underline{q}|),
\end{aligned}
$$

where $\underline{s}_i$ denotes the ith column vector of S.

Next, switching to polar coordinates, we get:

$$
\begin{aligned}
(q_{jR}, q_{jI}) &= (r_j \cos(\theta_j), r_j \sin(\theta_j)), \ 1 \leq j \leq n \\
q_j &:= r_j e^{\imath\theta_j}, \\
p_{R,\Theta}(\underline{r}, \underline{\theta}) &= r_1 \frac{1}{\pi} e^{-|\underline{s}_1^T\underline{q}|^2} \prod_{j=2}^{n} \frac{4}{2\pi} r_j K_0(2|\underline{s}_j^T\underline{q}|).
\end{aligned}
$$

A final change of variables $(r_j, \theta_j) \rightarrow (\rho^{-\alpha_j/2}, \theta_j)$ (i.e., $|q_j|^2 = \rho^{-\alpha_j}$), gives us

$$
p_{\underline{\alpha},\underline{\Theta}}(\underline{\alpha}, \underline{\theta}) = (\log\rho)^n \rho^{-\alpha_1} \frac{1}{2\pi} e^{-|\underline{s}_1^T\underline{q}|^2} \prod_{j=2}^{n} \frac{2}{2\pi} \rho^{-\alpha_j} K_0(2|\underline{s}_j^T\underline{q}|).
\tag{36}
$$

c) Outage probability: The parallel channel characterized by

$$
\underline{y} = \begin{bmatrix} q_1 & & & \\ & q_2 & & \\ & & \ddots & \\ & & & q_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \underline{w}
$$

has outage region given by

$$
\mathcal{O} = \{(\underline{\alpha}, \underline{\theta}) : \prod_{j=1}^{n}(1 + \rho^{1-\alpha_j}) < \rho^{r_p}\},
\tag{37}
$$

where $r_p \log(\rho)$ is the desired rate in bits per channel use over the parallel channel and a corresponding probability of outage

given by

$$P_{\text{out},\underline{q}}(r_p) = \int_{\mathcal{O}} p_{\underline{\alpha},\underline{\Theta}}(\underline{\alpha},\underline{\theta}) \, d\underline{\alpha} \, d\underline{\theta} : \doteq \rho^{-d_{\text{out},\underline{q}}(r_p)} \qquad (38)$$

where $p_{\underline{\alpha},\underline{\Theta}}(\underline{\alpha},\underline{\theta})$ is as in (36).

We first show that in the above integral, it is sufficient to consider the region where $\alpha_j \geq 0, \forall j$. Recall that $|z_i|^2 = \rho^{-p_i}$. Let $\min\{p_i\} = p_0$, $\min\{\alpha_i\} = \alpha_0$ and assume $\alpha_0 < 0$. Since

$$\sum_{i=1}^{n} |q_i|^2 = \sum_{i=1}^{n} |z_i|^2$$

it follows that for large ρ, $p_0 = \alpha_0$. Consider the term $K_0(2|\underline{s}_{i_0}^T \underline{q}|) = K_0(2|z_{i_0}|)$ where $z_{i_0} = \rho^{-p_0}$, $p_0 > 0$. For large values of the argument $|z_{i_0}| := x$, we have the bound $|\sqrt{x}e^x K_0(x) - P(1/x)| < \epsilon$ where ϵ is a constant much smaller than 1, and $P()$ is a polynomial of degree greater than 2 (see [25, Section 6.6]). This implies that $|K_0(x) - x^{-1/2}e^{-x}P(1/x)| < \epsilon x^{-1/2}e^{-x}$. It follows from the presence of the exponential term e^{-x} and an application of Varadhan's Lemma, that for the purposes of determining $d_{\text{out},\underline{q}}(r)$ it suffices to restrict attention to the region $\alpha_i \geq 0$.

Accordingly, we define the effective outage region

$$\mathcal{O}' = \mathcal{O} \bigcap \{(\underline{\alpha},\underline{\theta}) \ : \ \alpha_i \geq 0, \ \forall i\}. \qquad (39)$$

For any $(\underline{\alpha},\underline{\theta})$, we have $e^{-|\underline{s}_1^T \underline{q}|^2} \leq 1 \doteq \rho^0$. We will now proceed to establish that

$$K_0(2|\underline{s}_j^T \underline{q}|) \overset{.}{\leq} \rho^0 \quad \text{all} \quad (\underline{\alpha},\underline{\theta}) \in \mathcal{O}'. \qquad (40)$$

We list below two properties of $K_0(\cdot)$ that will prove useful:

1)
$$\lim_{x \to 0} \frac{K_0(x)}{\ln(2/x)} = 1.$$

2) $K_0(x)$ is monotonically decreasing with increasing x, $x \geq 0$.

Let $\epsilon > 0$ and define

$$\mathcal{O}'' = \mathcal{O}' \bigcap \{(\underline{\alpha},\underline{\theta}) \ : \ 2|\underline{s}_j^T \underline{q}| \leq S_{\max}\rho^{-\epsilon}\}$$

where $S_{\max} = \max\{|S_{ij}| \mid 1 \leq i,j \leq n\}$. Then for $(\underline{\alpha},\underline{\theta}) \in \mathcal{O}''$, we can make the approximation

$$K_0(2|\underline{s}_j^T \underline{q}|) \doteq \ln\left(\frac{2}{2|\underline{s}_j^T \underline{q}|}\right) \doteq \rho^0.$$

For $(\underline{\alpha},\underline{\theta}) \in \mathcal{O}' \setminus \mathcal{O}''$, we have that

$$2|\underline{s}_j^T \underline{q}| > S_{\max}\rho^{-\epsilon}$$

and it follows from the monotonically decreasing property of $K_0(x)$ that $K_0(2|\underline{s}_j^T \underline{q}|) \overset{.}{\leq} \ln(\frac{2}{S_{\max}\rho^{-\epsilon}}) \doteq \rho^0$. Substituting (36), (37) and (40) into (38), we conclude that for the parallel channel defined by diagonal matrix $\text{diag}(q_1,q_2,\ldots,q_n)$,

$$P_{\text{out},\underline{q}}(r_p) \doteq \rho^{-d_{out}(r_p)}$$

where

$$d_{out}(r_p) \geq \inf_{\sum_{i=1}^{n}(1-\alpha_i)^+ < r_p} \sum_{i=1}^{n} \alpha_i$$
$$= (n - r_p)$$

so that

$$P_{\text{out},\underline{q}}(r_p) \overset{.}{\leq} \rho^{-(n-r_p)}.$$

Equation (35) now follows by inspecting (35) and (37) and consequently setting $r_p = 2rn$.

APPENDIX IV
PROOF OF SIMULTANEOUS DIAGONALIZABILITY

Our goal here is to show that the matrices X belonging to the CDA-based code $\mathcal{X}$ (see (15)) are simultaneously diagonalizable, i.e., can be diagonalized by a single unitary matrix S. We will do so by explicitly constructing this matrix.

Let ζ_n be the complex, primitive nth root of unity given by $\zeta_n = e^{2i\pi/n}$. Let the complex number ϕ be given by $\phi = \zeta_n\gamma^{1/n}$. Then ϕ is a nth root of γ, that is $\phi^n = \gamma$. In the field of complex numbers, the equation $x^n = \gamma$ has the n solutions $\zeta_n^i\gamma^{1/n}$, $i = 1,\ldots,n$. For $i = 1,2,\ldots,n$, let σ_i denote the mapping given by:

$$\sigma_i(\phi) = \zeta_n^i\gamma^{1/n},$$
$$\sigma_i(\phi^k) = \zeta_n^{ki}\gamma^{k/n}, \ k = 2,3,\ldots,n.$$

Note that $\sigma_i(\phi^k) = [\sigma_i(\phi)]^k$ and recall that a matrix X in the code $\mathcal{X}$ is of the form

$$X = \begin{bmatrix} l_0 & \gamma l_{n-1} & \cdots & \gamma l_1 \\ l_1 & l_0 & & \gamma l_2 \\ \vdots & \vdots & & \vdots \\ l_{n-1} & l_{n-2} & & l_0 \end{bmatrix}, \quad l_i \in \mathcal{A}_{\text{QAM}}.$$

Proposition 1: The n vectors

$$[1, \sigma_i(\phi), \sigma_i(\phi^2),\ldots,\sigma_i(\phi^{n-1})], \quad i = 1,\ldots,n$$

are eigenvectors of each matrix $X \in \mathcal{X}$. .

Proof: Let $x = l_0 + l_1\phi + l_2\phi^2 + \ldots + l_{n-1}\phi^{n-1}$. First notice that since $\phi^n = \gamma$,

$$[1, \phi, \phi^2,\ldots,\phi^{n-1}]X = x[1, \phi, \phi^2,\ldots,\phi^{n-1}],$$

so that $[1, \phi, \phi^2,\ldots,\phi^{n-1}]$ is an eigenvector of F associated to the eigenvalue x. Similarly, since $[\sigma_i(\phi)]^n = \gamma$, we have

$$[1, \sigma_i(\phi), \sigma_i(\phi^2),\ldots,\sigma_i(\phi^{n-1})]X$$
$$= \sigma_i(x)[1, \sigma_i(\phi), \sigma_i(\phi^2),\ldots,\sigma_i(\phi^{n-1})],$$

where $\sigma_i(x) = l_0 + l_1\sigma_i(\phi) + l_2\sigma_i(\phi^2) + \ldots + l_{n-1}\sigma_i(\phi^{n-1})$, $i = 1,\ldots,n$. ∎

Let S be the matrix whose rows are scaled versions of these n eigenvectors as shown below:

$$S = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & \sigma_1(\phi) & \sigma_1(\phi^2) & \cdots & \sigma_1(\phi^{n-1}) \\ 1 & \sigma_2(\phi) & \sigma_2(\phi^2) & \cdots & \sigma_2(\phi^{n-1}) \\ \vdots & & & & \vdots \\ 1 & \sigma_n(\phi) & \sigma_n(\phi^2) & \cdots & \sigma_n(\phi^{n-1}) \end{bmatrix}.$$

It follows that $SX = X_d S$ where X_d is the diagonal matrix $X_d = \text{diag}(x, \sigma_1(x),\ldots,\sigma_{n-1}(x))$. We now show that S is a unitary matrix, from which it follows that we can write

$$X = S^\dagger X_d S \qquad (41)$$

and S is then the common unitary diagonalizing matrix for the matrices in $\mathcal{X}$ we are looking for.

Lemma 4.1: The matrix S is unitary, that is $S^\dagger S = S S^\dagger = I_n$.

Proof: The element in the ith row and j column of S, $1 \le i, j \le n$, is given by

$$\sigma_i(\phi^{j-1}) = \sigma_i(\zeta_n^{j-1} \gamma^{(j-1)/n}) = \zeta_n^{i(j-1)} \gamma^{(j-1)/n}.$$

As a result it follows that we can decompose S according to

$$S = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & \zeta_n & \cdots & \zeta_n^{n-1} \\ 1 & \zeta_n^2 & \cdots & (\zeta_n^2)^{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \zeta_n^{n-1} & \cdots & (\zeta_n^{(n-1)})^{n-1} \\ 1 & 1 & \cdots & 1 \end{bmatrix} \quad (42)$$

$$\times \begin{bmatrix} 1 & & \cdots & \\ & \gamma^{1/n} & & \\ \vdots & \vdots & \vdots & \vdots \\ & & & \gamma^{(n-1)/n} \end{bmatrix}. \quad (43)$$

The first matrix on the right is a row-permuted version of the familiar unitary Fourier-transform matrix and the second is also unitary since $|\gamma| = 1$. It follows that their product S is unitary as well. ∎

REFERENCES

[1] 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.

[2] 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.

[3] A. Sendonaris, E. Erkip, and B. Aazhang, "Increasing uplink capacity via user cooperation diversity," in *Proc. IEEE Int. Symp. Information Theory*, Cambridge, MA, Aug. 1998.

[4] 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.

[5] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062-3080, Dec. 2004.

[6] 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.

[7] Y. Jing and B. Hassibi, "Distributed space-time coding in wireless relay networks," to appear in *IEEE Trans. on Wireless Communications*, 2006.

[8] R. U. Nabar, H. Bölcskei, F. W. Kneubühler, "Fading relay channels: performance limits and space-time design", *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, Aug. 2004.

[9] 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.

[10] 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.

[11] S. Yang and J.-C. Belfiore, "Optimal space-time codes for the amplify-and-forward cooperative channel, " to appear in *IEEE Trans. Inform. Theory*.

[12] P. Elia *Asymptotic Universal Optimality in Wireless Multi-Antenna Communications and Wireless Networks*, Ph.D. thesis, USC, 2006.

[13] 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.

[14] S. Tavildar and P. Viswanath, "Approximately universal codes over slow fading channels," *IEEE Trans. on Inform. Theory*, vol. 52, no. 7, pp. 3233-3258, July 2006.

[15] P. Elia, K. Raj Kumar, Sameer A. Pawar, P. Vijay Kumar and Hsiao-feng Lu, "Explicit, minimum-delay space-time codes achieving the diversity-multiplexing gain tradeoff," *IEEE Trans. Inform. Theory*, vol. 52, no. 9, September 2006.

[16] F. Oggier, G. Rekaya, J.-C. Belfiore, E. Viterbo, "Perfect space-time block codes," *IEEE Trans. Inform. Theory*, vol. 52, no. 9, September 2006.

[17] P. Elia, B. Sethuraman, and P. Vijay Kumar, "Perfect space-time codes with minimum and non-minimum delay for any number of nntenas," submitted to *IEEE Trans. Info. Theory*, Dec. 2005. Available: http://arxiv.org/pdf/cs.IT/0512023.

[18] M. Marcus, and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Dover Publications, New-York, 1992.

[19] J.-C. Belfiore and G. Rekaya, "Quaternionic lattices for space-time coding," in *Proc. IEEE Inform. Theory Workshop.*, Paris, 31 March - 4 April 2003.

[20] A. A. Albert, *Structure of Algebras*, Coll. Publ., Vol. 24, Amer. Math. Soc., Providence, R. I., 1961.

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

[22] P. Camion, *Codes and Association Schemes*, Chapter 18 of Handbook of Coding Theory, North Holland, pp. 1441-1556, 1998.

[23] A. Dembo, O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd edition, Springer-Verlag, New York, 1998.

[24] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series and Products* Fifth edition, Academic Press, 1994.

[25] William H. Press, Saul A. Teukolsky, William T. Vetterling and Brian P. Flannery, *Numerical recipes in C. The Art of Scientific Computing* Second edition, Cambridge University Press, 1988-1992.

[26] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, Dover Publications, New York, 1965.

Petros Elia received the B.Sc. degree in electrical engineering from the Illinois Institute of Technology, Chicago, in 1997. During 1997-99 he worked as a consultant for MIOD Technologies, Chicago. He received the M.Sc. and Ph.D. degrees in electrical engineering from the University of Southern California, Los Angeles, in 2001 and 2006, respectively. His current research interests include communication and cooperation in wireless networks, MIMO systems and information theory.

P. Vijay Kumar (S'80-M'82-SM'01-F'02) received the B.Tech. and M.Tech. degrees from the Indian Institutes of Technology (Charagpur and Kanpur, India) and the Ph.D. degree from the University of Southern California (USC), Los Angeles, in 1983, all in electrical engineering. Since 1983, he has been on the faculty of the Electrical Engineering-Systems Department of USC, He is presently on leave of absence from USC at the Indian Institute of Science, Bangalore. His research interests include error-correcting codes and signal design for wireless andoptical communication. Prof. Kumar was an Associate Editor for Coding Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY from 1993 to 1996. In 1994, he received the USC School-of-Engineering Senior Research Award for contributions to coding theory. He is a corecipient of the IEEE Information Theory Society 1995 Prize Paper Award.

Frederique Oggier was born in Switzerland in 1977. She received her degree (Diplome) in Mathematics and Computer Science in 2000 from the University of Geneva. She then joined the Swiss Federal Institute of Technology, Lausanne (EPFL), where she graduated from the Doctoral School in Communication Systems (2001), and completed her Ph.D. thesis in Mathematics (2005). She is now a postdoctoral visitor at the California Institute of Technology (CalTech). She has been visiting Cornell University, Ithaca, NY, and AT&T Shannon Labs, Florham Park, NJ. Her current research interests are in applied algebra (in particular lattice theory, algebraic number theory and noncommutative algebras) to coding problems appearing in wireless communications, such as Space-Time Coding (both coherent and non-coherent) and coding for wireless networks.