Algebraic Approaches to Space-Time Code Construction for Multiple-Antenna Communication

U. Raviteja[1], I. Sharanappa[2], B. Vanamali[2] AND P. Vijay Kumar[2]

Abstract | A major challenge in wireless communications is overcoming the deleterious effects of fading, a phenomenon largely responsible for the seemingly inevitable dropped call. Multiple-antennas communication systems, commonly referred to as MIMO systems, employ multiple antennas at both transmitter and receiver, thereby creating a multitude of signalling pathways between transmitter and receiver. These multiple pathways give the signal a diversity advantage with which to combat fading. Apart from helping overcome the effects of fading, MIMO systems can also be shown to provide a manyfold increase in the amount of information that can be transmitted from transmitter to receiver. Not surprisingly, MIMO has played, and continues to play, a key role in the advancement of wireless communication.

Space-time codes are a reference to a signalling format in which information about the message is dispersed across both the spatial (or antenna) and time dimension. Algebraic techniques drawing from algebraic structures such as rings, fields and algebras, have been extensively employed in the construction of optimal space-time codes that enable the potential of MIMO communication to be realized, some of which have found their way into the IEEE wireless communication standards. In this tutorial article, reflecting the authors' interests in this area, we survey some of these techniques.

Introduction

While the use of multiple receive antennas goes back to the times of Marconi and Bose, the use of multiple transmit antennas is far more recent. Upon closer examination, it turns out that there is a tradeoff between the benefits of increase reliability and increased information rate offered by MIMO (short for multiple input, multiple output antenna) systems. We show in this article how space-time codes that optimally achieve this tradeoff can be designed. There are two general approaches to the design of efficient space-time codes and there is a different rate-reliability tradeoff to be achieved in the two cases.

Department of ECE, Indian Institute of Science, Bangalore - 560012, India

E-mail: [1]teju81@gmail.com, [2]{sharanappac, vanamali, vijay}@ece.iisc.ernet.in.

Under the first approach, space-time (ST) codes are designed to communicate a fixed amount of information across the channel, independent of the signal to noise ratio (SNR) of the channel. The tradeoff encountered here, is referred to as the rate-diversity tradeoff. Under the second approach, the rate of information transfer is indexed in a natural way to the operating SNR. Here again there is a tradeoff to be achieved, known as the diversity-multiplexing gain tradeoff. An overview of some of the algebraic techniques that have gone into the design of ST codes that are efficient when measured against the two tradeoffs is provided here.

In Section 2, we introduce the ST channel and discuss its fundamental limits. In Section 3, two approaches adopted by the ST research community for efficient code design are explained. In Sections 4 and 5, two types of constructions corresponding to the first approach are addressed. Space-time code design based on the second approach is explained in Section 6. We conclude the discussion in Section 7.

§ 1 Space-Time Channel

We will work throughout with the model of the ST channel presented below. Let n_t, n_r represent the number of transmit and receive antennas respectively. We assume that communication takes place in blocks, each block comprising a T-unit duration of time. Without loss of generality, the presentation will be in terms of a representation of the channel known as the equivalent, baseband, complex representation that abstracts away radio-frequency aspects of communication such as the carrier frequency and phase. For example, under the representation, the radio-frequency signal $A\cos(2\pi f_c t + \phi)$ of frequency f_c and phase shift ϕ, has the simpler representation $Ae^{\iota\phi}$, where $\iota = \sqrt{-1}$.

1.A Channel Model

We will assume that each block of T channel uses is processed independently. While this assumption does place a limit on performance, it has the important practical advantage of reducing the latency of communication. The transmitter addresses the channel by transmitting a T-length sequence of n_t-component vectors, known as a code matrix. Thus, each code matrix X is of size $(n_t \times T)$. The (i,j)th entry of a code matrix X represents the transmission by the ith transmit antenna during the jth channel use. The code matrix is drawn from a collection $\mathcal{X}$ of such matrices, known as a *ST code*. The word "space" as used here, is a reference to the spatial or antenna dimension. Similarly, let Y_{ij} represent the signal received by the ith receive antenna during the jth channel use. We will refer to the corresponding $(n_r \times T)$ matrix Y as the received matrix. Then the transmitted and received signals are related by the equation

$$Y = \theta HX + W \tag{1}$$

in which the $(n_r \times n_t)$ matrix H represents channel gains with H_{ij} denoting the channel gain along the path from jth transmit to ith receive antenna.

Channel fading is an inherent feature of most wireless communication and much of the activity in ST coding research is directed towards devising means of overcoming the deleterious

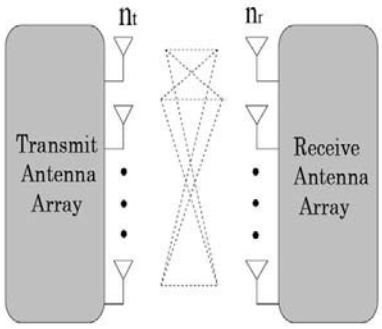

Figure 1: MIMO Channel Model

effects of fading. The effects of fading are modeled by allowing the channel gains H_{ij} to vary with time in random fashion. In our model, we make the block-fading (also known as quasi-static) assumption in which we assume that the channel matrix remains constant for the duration of a block, i.e., a duration of T channel uses. Across blocks, the channel matrices are assumed to be statistically independent.

The $(n_r \times T)$ matrix W represents the additive noise that is present in the channel. We assume, as is typical, that the components of W are independent and identically distributed (i.i.d.) and complex, circularly-symmetric Gaussian, i.e., that each component W_{ij} of the noise matrix W is distributed according to

$$\mathrm{P}_{W_{ij}}(z) = \frac{1}{\pi} e^{-|z|^2}, \quad \forall z \in \mathbb{C}. \tag{2}$$

The scalar $\theta = \sqrt{\frac{\rho}{n_t}}$ where the constant ρ represents the transmitted signal power and at the same time, SNR, since in our noise model, the variance of the complex noise variable is set equal to 1. The denominator n_t appearing in the expression for θ ensures that the transmitted power is the same regardless of the number of transmit antennas used. In referring to ρ as the SNR, we have also implicitly assumed that the variance $\mathbb{E}(|X_{ij}|^2)$ of the each code-matrix component is set equal to 1 in the code design.

1.B Decoder

In our channel model, the receiver is assumed to have full knowledge of the channel matrix H. We assume a decoder that, given a received matrix Y, selects the code matrix $\hat{X}$ that was most likely transmitted. Such a decoder is known in the literature as a maximum-likelihood (ML) decoder. It can be shown that the ML decoder minimizes the probability of decoding to an incorrect code matrix. This probability will be referred to as the *probability of codeword error* and denoted by P_e. Given the Gaussian nature of the noise, it can be shown that ML decisions result when, for a given received matrix Y, the decoder selects that code matrix $\hat{X}$ as having being transmitted, having the property that $H\hat{X}$ is closest in Euclidean distance to the received channel matrix Y, i.e., the decoded code matrix $\hat{X}$ minimizes the quantity

$$\|Y - \theta H\hat{X}\|_F^2, \tag{3}$$

called the ML metric, where $\|\mathrm{A}\|_F$ denotes the Frobenius norm of matrix A. Determining a closed-form expression for the probability of error of a ST communication system is typically difficult. We often settle for a simpler description that describes the behaviour of P_e for large values of SNR. Towards this end, we introduce the following convenient notation. We will write $f(\mathrm{SNR}) \doteq \mathrm{SNR}^{-b}$ if

$$b = -\lim_{\mathrm{SNR}\to\infty} \frac{\log(f(\mathrm{SNR}))}{\log(\mathrm{SNR})}. \tag{4}$$

Analogous definitions hold for the notation $\dot{\leq}$ and $\dot{\geq}$. If in a communication system where fading is encountered, the probability of error satisfies

$$P_e \doteq SNR^{-d} \tag{5}$$

we will say that the system has diversity d. As will be seen, the diversity d of a ST communication system can often be determined exactly.

1.C Fundamental Limits

Entropy and Capacity Consider an information source that outputs a sequence $\{X_i\}_{i=1}^{n}$ of i.i.d. discrete random variables, having probability density function $p_X(x)$, $x \in \mathcal{X}$, where $\mathcal{X}$ is the symbol alphabet of the source. The *entropy* of such a source is a measure of the uncertainty dispelled when a particular source symbol is received. When uncertainty is dispelled, we say that we have gained information. It can be shown via an axiomatic approach, that the entropy $H(X)$ of such a source is given by the expression

$$\begin{aligned} H(X) &= \int_{x\in\mathcal{X}} p_X(x)\log(\frac{1}{p_X(x)})dx \text{ bits} \quad \text{(continuous alphabet)}, \\ H(X) &= \sum_{x\in\mathcal{X}} p_X(x)\log(\frac{1}{p_X(x)}) \text{ bits} \quad \text{(discrete alphabet)}. \end{aligned}$$

In the expression, logs are taken to base 2. If, in the discrete-alphabet case, all outcomes are equally likely, i.e.,

$$p_X(x) = \frac{1}{|\mathcal{X}|}, \ \forall x \in \mathcal{X},$$

then

$$H(X) = \log(|\mathcal{X}|). \tag{6}$$

It follows from this, that $H(X) = 1$ if $|\mathcal{X}| = 2$ and both outcomes are equally likely and hence, a bit may be quantified as the amount of information conveyed or uncertainty dispelled upon being told of the outcome of the toss of a fair coin. The *capacity* C of a communication channel is a measure of the amount of information that can be reliably transmitted across the channel. By reliable communication, we mean communication in which the probability of decoding error can be made as small as one desires by suitably encoding data. A major result in information theory, due to Shannon [3] states that reliable communication is possible only as long as the entropy of the source is strictly less than the capacity of the communication channel, i.e., provided

$$H(X) < C.$$

Additive, White, gaussian Noise Channel A common example of a communication channel, the *additive, white, Gaussian noise channel* commonly abbreviated as the AWGN channel, is a channel in which the output Y of the channel, a scalar, is given by $Y = X + W$, where X is the input and W represents the additive Gaussian noise distributed according to

$$\mathrm{P}_W(z) = \frac{1}{\pi} e^{-|z|^2}, \quad \forall z \in \mathbb{C}.$$

It can be shown that for large values of the SNR, the capacity of the additive white Gaussian noise (AWGN) channel is given by

$$C_{\mathrm{AWGN}} \approx \log(\mathrm{SNR}). \tag{7}$$

Ergodic capacity of the Space-Time Channel The ergodic capacity C_{ergodic} of the ST channel modeled by (1) is the capacity of the ST channel for the case when it is permitted to code across blocks. By this we mean that information about a particular message is spread across potentially infinite number of communication blocks. It can be shown that at large SNR, the ergodic capacity takes on the form

$$C_{\mathrm{ergodic}} \approx \min\{n_t, n_r\} \log(\mathrm{SNR}). \tag{8}$$

Thus in a sense, the ST channel has the capability to carry an amount of information equal to $\min\{n_t, n_r\}$ times that which can be carried by the AWGN channel. This observation caused much excitement in communication circles when it was first announced since it meant that just by adding a number of transmit and receive antennas and incurring some additional communication complexity, the capacity of a communication channel could be increased manyfold. To a cellular operator, it meant that he could with two receive and two transmit antennas, serve twice as many customers as he could with just a single transmit and receive antenna.

Probability of Outage We now return to our communication setup wherein it is not permitted to code across blocks. Thus we limit our attention to a time span of T channel uses and accept that during this time interval, the matrix H appearing in (1) is a single realization of a set of $n_t n_r$ random variables. We will also assume that the receiver knows the channel matrix H, but that the transmitter does not, as the latter would require the presence of a feedback channel. Let R be the desired rate of communication in bits per channel use (bpcu). We would like to highlight two key differences between the above framework and the ergodic capacity framework. Shannon's channel capacity theorem, when applied to the ergodic ST channel states that as long as a code is attempting to transmit information at rate R, that is less than the ergodic channel capacity C, i.e., $R < C$, the probability of codeword error can be made to approach zero by encoding information across an ever increasing number of blocks, i.e., $P_e \to 0$, using appropriately designed codes with infinite codeword lengths [4]. For the block-fading model of the ST channel (1), the goal of driving $P_e(R) \to 0$ can never be achieved for any finite non-zero rate R that is chosen independent of the channel condition. This is due to the fact that there is always a certain probability that the channel realization H is unable to support the rate R of information transmission for any value of information rate R, however small. A channel that is unable to support the rate of attempted communication across it, is said to be "in outage" and the corresponding probability is denoted by $\mathrm{P}_{\mathrm{out}}(R)$. Thus, the capacity of the ST channel (1) in the strict Shannon sense [3] is zero. It can be shown via

an information-theoretic inequality known as Fano's inequality [4], that the probability of codeword error is lower bounded by the outage probability, i.e.,

$$P_e(R) \;\geq\; P_{out}(R). \tag{9}$$

Thus in contrast to the ergodic capacity scenario, where one can ensure that $P_e \to 0$, in the block-fading setup, the best performance one can hope to achieve with a code operating over the ST channel (1) is $P_e(R) \approx P_{out}(R)$. A second point of difference, which we will see later is that at high SNRs, explicit codes with short block lengths $T \geq n_t$ suffice to achieve the best performance. This is in contrast to Shannon's coding theorem [3] which require encoding via codes of infinite block length.

Not surprisingly, the central goal of ST communication is to come up with a coding scheme whose associated probability of codeword error $\mathrm{P_e}(R)$ is as close to the outage probability $\mathrm{P_{out}}(R)$ as possible.

§ 2 Two Approaches to Code Design

The ST community has in a sense, adopted two approaches to achieve this goal.

First Approach Under the first, it is assumed that the entries X_{ij} of each code matrix X are drawn from an alphabet $\mathcal{A}$ of size $q = |\,\mathcal{A}\,|$. To transmit at rate R bpcu, the code designer typically constructs a ST code $\mathcal{X}$ of size $|\,\mathcal{X}\,| = 2^{RT}$ and then selects each code matrix with equal likelihood in accordance with a given stream of input message symbols. It follows from (6), that the average amount of information transmitted per successful reception of the code matrix at the receiver end is equal to R bpcu. In this case, it can be shown that the probability of outage has a large SNR exponent equal to $(-1)n_t n_r$, i.e.,

$$P_{\mathrm{out}}(R) \;\doteq\; \mathrm{SNR}^{-n_t n_r}.$$

Second Approach Under the second approach, the rate of information transmitted is allowed to vary with the SNR and is given by an expression of the form

$$R \;=\; r\log(\mathrm{SNR}) \quad \text{bpcu}. \tag{10}$$

The parameter r in the expression, is called the multiplexing gain (MG). The expression for ergodic capacity given above in (8) suggests that the range of r for which the probability of error can be kept low is limited to the range $0 \leq r \leq \min\{n_t, n_r\}$ and this can be shown to be the case. Thus our interest is now in the probability of error performance of a ST code for MG in the range $0 \leq r \leq \min\{n_t, n_r\}$. Here again, the goal is to make the probability of error as close to the probability of outage as possible. For large values of SNR, let us define outage exponent $d_{\mathrm{out}}(r)$ by

$$P_{\mathrm{out}}(r) \;\doteq\; \mathrm{SNR}^{-d_{\mathrm{out}}(r)}. \tag{11}$$

We can similarly associate an error exponent $d_e(r)$ with the probability of error, i.e.,

$$P_e(r) \;\doteq\; \mathrm{SNR}^{-d_{\mathrm{e}}(r)}, \tag{12}$$

then the goal of ST code design under the second approach, is to design codes whose error exponent $d_{\mathrm{e}}(r)$ equals the outage exponent $d_{\mathrm{out}}(r)$.

§ 3 Fixed Rate Approach

In the fixed-rate setting, ST codes are designed to operate in the block-fading setting, at a fixed rate R bpcu. Since determining a closed-form expression for the probability P_e of codeword error is difficult, analysis typically focuses on a related measure that is more tractable, namely pairwise error probability.

3.A Pairwise Error Probability

The pairwise-error probability (PEP), $P_e(X_1 \to X_2)$ is defined as the probability that X_1 is transmitted, but that $\hat{X} = X_2$ is decoded as having been sent, under the assumption that the ST code is comprised of just the two code matrices X_1, X_2. Given the PEP for every pair (X_1, X_2), one can derive a simple upper bound on the codeword error probability P_e.

Given that we are employing an ML decoder, it is not surprising that the PEP is a function of the squared Euclidean distance d_E^2 between the received matrices HX_1, HX_2:

$$\begin{aligned} P_e(X_1 \to X_2|H) &= Q\left(\frac{d_E(H)}{\sqrt{2}}\right) \\ &\leq e^{-\frac{d_E^2(H)}{4}}, \end{aligned} \tag{13}$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{u^2}{2}} du. \tag{14}$$

This is the PEP for a given realization of the channel matrix H. By averaging over all possible realizations of H, one obtains $P_e(X_1 \to X_2)$. We now provide an alternate expression for d_E^2 which provides insight into the problem of code design for improved performance and which is easier to average over various channel realizations. Let $\triangle X := X_1 - X_2$. Then, since $\triangle X \triangle X^\dagger$ is a Hermitian matrix, we have the decomposition

$$\triangle X \triangle X^\dagger = V L V^\dagger,$$

where L is the diagonal matrix of eigenvalues of $\triangle X \triangle X^\dagger$ and V is a unitary matrix. We set $D = HV$. Then in terms of the components $\{l_j\}$ of L and d_{ij} of D, the squared Euclidean distance has the alternate expression:

$$\begin{aligned} d_E^2 &= \|\theta H X_1 - \theta H X_2\|_F^2 \\ &= \mathrm{Tr}(\theta^2 H \triangle X \triangle X^\dagger H^\dagger) \\ &= \theta^2 \sum_{j=1}^{n_t} l_j \sum_{i=1}^{n_r} \|d_{ij}\|^2 . \end{aligned} \tag{15}$$

Using (13) and (15), and averaging over all channel realizations H, one obtains the upper bound on PEP given below :

$$P_e(X_1 \to X_2) \leq \left[\frac{\mathrm{SNR}}{4n_t} l_{gm}\right]^{-\nu n_r}, \tag{16}$$

where $\nu = \text{rank}(\triangle X) = \text{rank}(\triangle X \triangle X^\dagger)$ and $l_{gm} = (\prod_{j=1}^{\nu} l_j)^{\frac{1}{\nu}}$ is the geometric mean of the non-zero eigenvalues of $\triangle X \triangle X^\dagger$. If PEP $\leq \frac{1}{(k\text{SNR})^d}$ then k is called the *coding gain* and d is called the *diversity gain* of the system. Thus from (16), we can see that the diversity gain is $n_r \nu$ and the coding gain is l_{gm}.

It follows from the discussion above, that from the point of view of minimizing the probability of pairwise error, the code designer should aim to design codes in such a way that the rank of the difference matrix $\triangle X = X_i - X_j$ be as large as possible and then, given the rank, making the product of the non-zero eigenvalues as large as possible. These design criteria are formalized below :

- *Rank Criterion* : Maximize

$$\min\{\text{Rank}(X_i - X_j) \mid X_i, X_j \in \mathcal{X},\ i \neq j\}.$$

  While it is clear that the maximum possible value of difference rank equals the number n_t of transmit antennas, as we shall see below, rate considerations may force us to work with a lesser value of minimum difference rank.

- *Determinant Criterion* : Maximize the minimum of the product of non-zero eigenvalues of the difference matrix $\triangle X = X_i - X_j$ as this maximizes the coding gain.

3.B Rate-Diversity Tradeoff

For a fixed size $|\mathcal{A}|$ of the code-symbol alphabet $\mathcal{A}$, the amount of diversity gain achievable is a function of the rate of information transmission. For the purposes of quantifying this tradeoff, it will be found more convenient to use a slightly different measure of information rate. Let $R_{\mathcal{A}}$ be the information measured in number of symbols from the alphabet $\mathcal{A}$ per channel use. This is equivalent to saying that the ST code is of size

$$|\mathcal{X}| = |\mathcal{A}|^{R_{\mathcal{A}} T}. \tag{17}$$

Since we also have

$$|\mathcal{X}| = 2^{RT},$$

we have the rate conversion equation,

$$R = R_{\mathcal{A}} \log(|\mathcal{A}|). \tag{18}$$

The rate-diversity tradeoff is a consequence of the Singleton bound below. This bound dictates the maximum possible rate $R_{\mathcal{A}}$ of a ST code $\mathcal{X}$ when the difference rank is constrained to be no smaller than ν, $1 \leq \nu \leq n_t$.

Theorem 1. *[5] Let $\mathcal{A}$ be a symbol alphabet of size $|\mathcal{A}| = q$ that is a subset $\mathcal{A} \subset \mathbb{F}$ of some field $\mathbb{F}$. Let $\mathcal{X}$ be a ST code over $\mathcal{A}$ such that the difference $X_1 - X_2$ for any pair of distinct matrices $X_1, X_2 \in \mathcal{X}$ has rank $\geq \nu$. Then the size of the ST code $\mathcal{X}$ is upper-bounded by*

$$|\mathcal{X}| \leq q^{T(n_t - \nu + 1)}.$$

Thus the code rate $R_{\mathcal{A}}$ must satisfy the upper bound

$$R_{\mathcal{A}} \leq (n_t - \nu + 1). \tag{19}$$

Proof: If the first $n_t - \nu + 1$ rows of the two distinct code matrices X_1, X_2 are identical, the difference matrix $\triangle X = X_1 - X_2$ will have rank $< \nu$. It follows that $\mathcal{X}$ cannot be of size larger than $|\mathcal{A}|^{T(n_t-\nu+1)}$. The result follows. □

Remark 1. (19) *is known as the* rate-diversity tradeoff [1]. *Note that the tradeoff is independent of the number of receive antennas n_r and for $n_t \leq T$, is independent of T as well. The maximum achievable diversity for a given rate $R_{\mathcal{A}}$ can be determined from the rate-diversity tradeoff as plotted in Fig. 2 for the case when $n_t = 4$.*

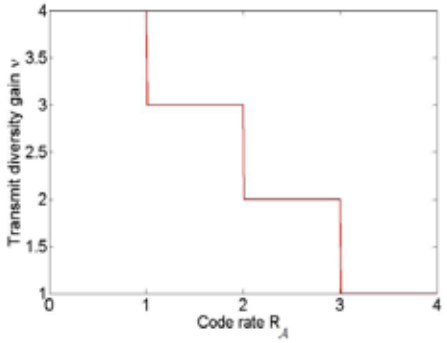

Figure 2: Rate Diversity Tradeoff for $n_t = 4$

3.B.1 Maximal-Rank Codes

Codes that achieve the rate-diversity tradeoff are known as maximal-rank codes. In this section we will show how maximal-rank codes can be constructed for the case when the code symbol alphabet is the popular 16-QAM constellation. The ST code matrices constructed in this section are square and hence correspond to the case when $n_t = T$. Starting with the construction of maximal-rank codes over the binary field, we construct maximal-rank codes over 16-QAM alphabet.

Additive Polynomials Polynomials of the form

$$L(x) \quad = \quad \sum_{i=0}^{R_{\mathcal{A}}-1} L_i x^{2^i}, \;\; L_i \in \mathbb{F}_{2^T}$$

are called additive polynomials because $L(x+y) = L(x) + L(y)$. It follows from this, that the collection of zeros (lying in $\mathbb{F}_{2^T}$) of an additive polynomial forms a vector space over the scalar binary field $\mathbb{F}_2$ of dimension not exceeding $R_{\mathcal{A}} - 1$. Such polynomials will be used in the constructions presented below of maximal-rank codes.

Maximal-Rank Binary Codes A maximal-rank-ν, $(n_t \times n_t)$ binary code $\mathcal{B}$, is a maximal collection of matrices over $\mathbb{F}_2$ such that the difference $X_1 - X_2$ of any two distinct matrices

[1] Not to be confused with the diversity-multiplexing gain tradeoff which also discusses the tradeoff between rate and reliability, but which applies only to the case when the operating rate is a function of SNR.

has rank $\geq \nu$. The size of the code must necessarily satisfy the Singleton bound and is hence given by:

$$|\mathcal{B}| \;=\; 2^{n_t-\nu+1}.$$

We show below, how additive polynomials can be used to construct maximal-rank codes. In turn, maximal-rank binary codes will enable the construction of ST codes achieving the rate-diversity tradeoff.

Theorem 2. *[5] For any $n_t \leq T < \infty$ and $1 \leq \nu \leq n_t$, let $R_{\mathcal{A}} = n_t - \nu + 1$ and define the set of code polynomials by*

$$\mathcal{L} = \{L(x) : L(x) = \sum_{j=0}^{R_{\mathcal{A}}-1} L_j x^{2^j}, L_j \in \mathbb{F}_{2^T}\} \tag{20}$$

Associate to every code polynomial L in $\mathcal{L}$, the vector

$$\underline{c}_L^T = [L(1), L(\alpha), ..., L(\alpha^{n_t-1})]$$

where α is a primitive element of $\mathbb{F}_{2^T}$. We associate with every code vector $\underline{c}_L$, the $(n_t \times T)$ code matrix

$$C_L = [\underline{L(1)}, \underline{L(\alpha)}, ..., \underline{L(\alpha^{n_t-1})}]^T$$

where, by $\underline{L(\alpha^j)}$ we mean, the representation of the element $L(\alpha^j)$ as a binary $(T \times 1)$ column vector. Then the collection

$$\mathcal{X} = \{C_L \mid L \in \mathcal{L}\}$$

is a binary, linear, maximal-rank-ν code.

Proof: We only need to establish that any difference matrix has rank $\geq \nu$ and for a linear code, it suffices to show that any code matrix C_L has rank $\geq \nu$. The theorem would then follow from the Singleton bound. Note that if $\underline{a}^T = [a_1, a_2, \cdots, a_{n_t}]$, then $\underline{a}$ belongs to the left null-space $\mathcal{N}(C_L)$ of C_L iff

$$\sum_{j=1}^{n_t} a_j L(\alpha^{j-1}) \;=\; L\left(\sum_{j=1}^{n_t} a_j \alpha^{j-1}\right) \;=\; 0.$$

Since, each such polynomial $L(x)$ has degree $\leq 2^{R_{\mathcal{A}}-1}$ it follows that $|\,\mathcal{N}(C_L)\,| = 2^{\dim(\mathcal{N}(C_L))} \leq 2^{R_{\mathcal{A}}-1}$. The theorem then follows, as the $\text{rank}(C_L) \;=\; n_t - \dim(\mathcal{N}(C_L)) \geq n_t - (R_{\mathcal{A}} - 1) \;=\; \nu$. □

The 16-QAM Alphabet The 16-QAM alphabet $\mathcal{A}_{\text{QAM}}$ is a popularly employed signal constellation in present-day communication systems and hence it is of interest to design ST codes over this alphabet. The constellation may be described as the collection of 16 points in the complex plane (see Fig. 3) given by

$$\mathcal{A}_{\text{QAM}} = \{a + \imath b \;\mid\; |a|, |b| \leq 3,\; a, b \text{ odd}\}, \tag{21}$$

where, M is even. The constellation has the alternate description

$$\mathcal{A}_{\text{QAM}} = \{(1+\imath)[\imath^a + 2\imath^b] \;\mid a, b \in \mathbb{Z}_4\},$$

which will prove useful in our construction of maximal-rank codes over this alphabet.

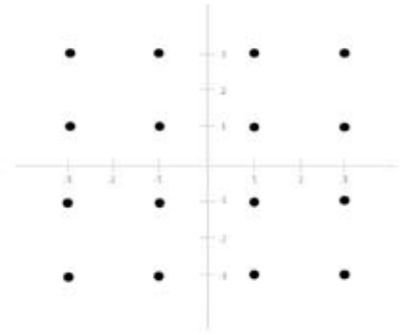

Figure 3: QAM Constellation

Maximal-Rank Codes over the QAM Alphabet The theorem below shows how a maximal-rank binary code can be used to put together a ST code $\mathcal{X}$ over the 16-QAM constellation $\mathcal{A}_{\text{QAM}}$ that achieves the Singleton bound.

Theorem 3. *[5] Let $\mathcal{C}$ be a maximal rank ν, $(n_t \times T)$ code over $\mathbb{F}_2$. Then, the ST code*

$$\mathcal{X} = \left\{ (1+\imath) \sum_{k=0}^{1} 2^k \, \imath^{C_{k,0}+2C_{k,1}} \right\}$$

where $C_{k,\ell} \in \mathcal{C}$, $k,\ell \in \{0,1\}$, also has the property that the difference rank between any two distinct code matrices is $\geq \nu$ and a code rate R that achieves the Singleton bound given by

$$R_{\mathcal{A}} \;=\; n_t - \nu + 1.$$

Proof: Let X, Y be two distinct code matrices given by

$$\begin{aligned} X &= (1+\imath)\left\{ \imath^{A_0+2A_1} + 2\imath^{B_0+2B_1} \right\}, \\ Y &= (1+\imath)\left\{ \imath^{C_0+2C_1} + 2\imath^{D_0+2D_1} \right\}. \end{aligned}$$

Our aim is to show if the matrices $\{A_i, B_i, C_i, D_i \mid i \in \{0,1\}\}$ are all drawn from a maximal-rank-ν code, then $X-Y$ also has rank $\geq \nu$. We will regard X,Y as matrices over the ring $Z[i]/8Z[i]$. This is a chain ring, i.e., a principal ideal ring possessing a unique maximal ideal $(1-\imath)$ by McDonald's theorem [6]. Also, any matrix $\imath^{P_0+2P_1}$ can be expanded in the form,

$$\begin{aligned} \imath^{P_0+2P_1} &= [1-(1-\imath)]^{P_0} \odot (-1)^{P_1} \\ &= [J - P_0(1-\imath)] \odot [J - 2P_1] \\ &= J - (1-\imath)P_0 - 2P_1 + 2(1-\imath)P_0 \odot P_1, \end{aligned}$$

where $\odot$ refers to Schur product between matrices. Thus, $\frac{\triangle}{1+\imath} = \frac{X-Y}{1+\imath}$ can be expanded as

$$\begin{aligned} &= \left[\imath^{A_0+2A_1} + 2\imath^{B_0+2B_1} \right] - \left[\imath^{C_0+2C_1} + 2\imath^{D_0+2D_1} \right] \\ &= (1-\imath)(C_0 - A_0) + 2(C_1 - A_1) + 2(1-\imath)\left[(A_0 \odot A_1 - C_0 \odot C_1) + (D_0 - B_0) \right] + 4(D_1 - B_1). \end{aligned}$$

Next, assume $C_0 \neq A_0$ and consider

$$\frac{\triangle}{(1+\iota)(1-\iota)} = (C_0 - A_0) + (1+\iota)(C_1 - A_1) + 2[(A_0 \odot A_1 - C_0 \odot C_1) + (D_0 - B_0)] + 2(1+\iota)(D_1 - B_1).$$

Since $(C_0 - A_0)$ is a matrix over $\mathbb{F}_2$ with rank $\geq \nu$, it follows that some $(\nu \times \nu)$ principal submatrix of $(C_0 - A_0)$ has non-zero determinant. For any $(m \times n)$ matrix G, we will use $G_{(S_1,S_2)}$ to denote the principal submatrix of G whose rows are indexed by elements in S_1 and columns are indexed by S_2. It follows that

$$\det\left[\frac{\triangle}{(1+\iota)(1-\iota)}\right]_{(S_1,S_2)} \neq 0 \quad \text{since,}$$
$$\det\left[C_0 - A_0\right]_{(S_1,S_2)} \neq 0 \ (\text{mod} 2),$$

where (S_1, S_2) identify the rows and columns of the principal submatrix having non-zero determinant. If $C_0 = A_0$, but $C_1 \neq A_1$, we consider instead $\frac{\triangle}{(1+\iota)^2(1-\iota)}$ and argue in a similar fashion. □

§ 4 Orthogonal Designs

A major and often overriding concern in communication systems is decoding complexity. A ST code maps a collection $\{u_i\}_{i=1}^{K}$ of message symbols onto a code matrix. In general, decoding requires us to identify the particular *set* of message symbols that yields the code matrix that is most likely to have been transmitted given the received signal matrix Y. Interestingly, at times, it is possible to construct ST codes in such a way that one can decode each message symbol u_i without regard to the values of the other symbols. Such ST codes are said to be *single-symbol decodable* and clearly incur significantly reduced decoding complexity.

Space-time codes derived from combinatorial objects known as *orthogonal designs* possess the property of single symbol decodability and the Alamouti code [7] is a prime example of such a code and possesses a particularly simple structure.

4.A Alamouti Code

The Alamouti code is a ST code that is designed for the case when there are $n_t = 2$ transmit antennas. Each code matrix in the Alamouti code has the form:

$$X = \begin{bmatrix} x_1 & -x_2^* \\ x_2 & x_1^* \end{bmatrix}, \tag{22}$$

where x_1, x_2 are drawn from an alphabet $\mathcal{A}$ that is a subset of the field $\mathbb{C}$ of complex numbers. The difference $X_1 - X_2$ between any two distinct matrices X_1, X_2 is of the form

$$\triangle X := \begin{bmatrix} \triangle x_1 & -\triangle x_2^* \\ \triangle x_2 & \triangle x_1^* \end{bmatrix}.$$

The difference matrix $\triangle X$ is of full rank since the columns of the matrix can be verified to be orthogonal and thus satisfies the rank criterion needed to obtain full diversity. From Theorem 1, we see that the Alamouti code has $n_t = 2$, $\nu = 2$, $R_{\mathcal{A}} = 1$ and is hence a $R_{\mathcal{A}} = 1$, full diversity code and thus achieves the rate-diversity tradeoff. The Alamouti code is now part of the WiMAX IEEE 802.16e wireless communication standard.

Single-Symbol Decodability We now go on to explain the reason why the Alamouti code is single-symbol decodable. We observe, first of all, that every code matrix X in the Alamouti code $\mathcal{X}$ has a representation as the linear combination, over the reals, of certain fixed matrices known as *dispersion matrices*. To see this, let us expand the symbols x_1, x_2 appearing in (22).

$$\begin{aligned} x_1 &= u_1 + iu_2, \\ x_2 &= u_3 + iu_4, \end{aligned}$$

where $u_i \in \mathbb{R}$, $1 \leq i \leq 4$. Then X can be written in the form

$$X = \sum_{i=1}^{4} u_i A_i, \tag{23}$$

where,

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad A_2 = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

The matrices A_i, $1 \leq i \leq 4$ are called dispersion matrices. We can easily check that these matrices satisfy

$$\begin{aligned} A_i A_i^\dagger &= I \quad \forall i = 1,2,3,4 \\ A_i A_j^\dagger + A_j A_i^\dagger &= [0] \quad j \neq i, \end{aligned} \tag{24}$$

and as we shall see, this is key to the single-symbol decodability of the code. The received signal at the receiver is given by

$$Y = HX + W,$$

where θ in (1) has been absorbed into X. From (24), it follows that the real inner product of any two matrices HA_i, HA_j $i \neq j$, $\langle HA_i, HA_j \rangle$ equals zero as shown below:

$$\begin{aligned} < HA_i, HA_j > &= \text{Re}\,[\,\text{Tr}(HA_i A_j^\dagger H^\dagger)] \\ &= \frac{1}{2}[\,\text{Tr}(HA_i A_j^\dagger H^\dagger) + (\,\text{Tr}(HA_i A_j^\dagger H^\dagger)\,)^*\,] \\ &= \frac{1}{2}[\,\text{Tr}(HA_i A_j^\dagger H^\dagger) + \text{Tr}(HA_j A_i^\dagger H^\dagger)\,] \\ &= \frac{1}{2}\text{Tr}(H(A_i A_j^\dagger + A_j A_i^\dagger)\,H^\dagger) \\ &= 0. \end{aligned}$$

Projecting Y onto $\{HA_i \mid 1 \leq i \leq 4\}$ we get,

$$\begin{aligned} < Y, HA_i > &= \text{Re}\,[\,\text{Tr}(YA_i^\dagger H^\dagger)\,] \\ &= \text{Re}\,[\,\text{Tr}((H\sum_{j=1}^{4} u_j A_j + W)A_i^\dagger H^\dagger)\,] \\ &= u_i \text{Re}\,[\,\text{Tr}(HA_iA_i^\dagger H^\dagger)\,] + \text{Re}\,[\,\text{Tr}(WA_i^\dagger H^\dagger)\,] \\ &= u_i \|H\|_F^2 + \text{Re}\,[\,\text{Tr}(WA_i^\dagger H^\dagger)\,]. \end{aligned}$$

In this way, message symbol u_i has been isolated and can be decoded independently of the other message symbols. The code is single-symbol decodable since each of the other message symbols can be decoded in similar fashion. It is natural to wish to extend the Alamouti code to a larger number of antennas while maintaining rate $R_{\mathcal{A}} = 1$, i.e., we would hope to find in the case of a (3×3) ST code, a collection of 6 complex matrices $\{B_j \mid 1 \leq j \leq 6\}$ such that every code matrix can be expressed as a real linear combination of these 6 matrices $\{B_j\}$ and furthermore that the matrices satisfy the analogue given below, of condition (24) appearing above:

$$\begin{aligned} B_iB_i^\dagger &= I \quad \forall i = 1,2,3,...,6 \\ B_iB_j^\dagger + B_jB_i^\dagger &= [0] \quad j \neq i. \end{aligned} \tag{25}$$

But as shown below, these are precisely the defining relation of a complex orthogonal design.

4.B General Orthogonal Designs

Definition 1. *An $(n \times T)$ complex, linear-processing orthogonal design (C-LPOD), in the complex variables (x_i; $i = 1,2,...,K$) is an ($n \times T$) matrix P whose elements are linear combinations (typically, 0; $\pm x_i$; $\pm x_i^*$) of the variables x_i, x_i^*, that satisfies*

$$PP^\dagger = (\sum_{i=1}^{K} |x_i|^2) I. \tag{26}$$

We will say that the C-LPOD has rate = K/T.

By the linearity of P, we can express P in the form,

$$P = \sum_{i=1}^{2K} u_i A_i \tag{27}$$

for some complex $(n \times T)$ matrices A_i, where

$$\begin{aligned} u_i &= Re(x_i) = x_{i,I},\ 1 \leq i \leq K, \\ u_{k+i} &= Im(x_i) = x_{i,Q},\ 1 \leq i \leq K. \end{aligned}$$

Substituting (27) in (26) and equating variables we get,

$$\begin{aligned} A_iA_i^\dagger &= I \quad \forall i = 1,2,....2K \\ A_iA_j^\dagger + A_jA_i^\dagger &= [0] \quad j \neq i. \end{aligned} \tag{28}$$

The equations above imply in particular, that any non-trivial real linear combination $P = \sum_{i=1}^{2K} \lambda_i A_i$, is a scaled version of a unitary matrix as shown below :

$$\begin{aligned} PP^{\dagger} &= \left[\sum_{i=1}^{2K} \lambda_i A_i\right]\left[\sum_{j=1}^{2K} \lambda_j A_j\right]^{\dagger} \\ &= \sum_{i,j} \lambda_i \lambda_j A_i A_j^{\dagger} \\ &= \frac{1}{2}\sum_{i\neq j} \lambda_i \lambda_j (A_i A_j^{\dagger} + A_j A_i^{\dagger}) + \sum_i \lambda_i^2 A_i A_i^{\dagger} \\ &= \sum_{i}^{2K} \lambda_i^2 I. \end{aligned}$$

In particular, it follows that any real linear combination of the matrices $\{A_i\}$ is nonsingular and hence that the matrices $\{A_i\}$ are linearly independent. The theorem below by Adam *et. al.* places bounds on the number of such matrices $\{A_i\}$ that one can construct, of a given size.

Theorem 4 (Adams, Lax and Phillips)**.** *[8] Let $\mathbb{C}$ be complex field. Let $\{M_i \mid i = 1, 2, ..., K_n\}$ be $(n \times n)$ matrices over $\mathbb{C}$ having the property that every real linear combination*

$$\lambda_1 M_1 + \lambda_2 M_2 + ...\lambda_{K_n} M_{K_n}$$

of the matrices M_i is nonsingular. Let $\mathbb{C}(n)$ denote the largest integer K_n for which such a collection of matrices can be found. Then

$$\mathbb{C}(n) = \mathbb{C}(2^a(2b+1)) = 2a+2\,.$$

From Theorem 4, when $n = T = 2^a(2b+1)$ we have,

$$2K = K_n \leq \mathbb{C}(n) = (2a+2).$$

Thus we have that $K \leq a+1$. As a consequence, the complex symbol rate $R_{\mathcal{A}}$ of a $(n \times n)$ square orthogonal design has its maximum value limited by

$$\mathbb{C}(n)/(2n) = (a+1)/(2^a(2b+1)), \tag{29}$$

as shown in the table below:

n	$\mathbb{C}(n)$	MAX RATE
2	4	1
3	2	1/3
4	6	3/4
5	2	1/5
6	4	1/3
7	2	1/7
8	8	1/2

.

Note that it is not possible to construct a rate $R_{\mathcal{A}} = 1$ square orthogonal design for $n_t > 2$ transmit antennas.

4.C Orthogonal Designs from Clifford Algebras

Clifford algebras offer a means of constructing orthogonal designs as explained below.

Definition 2. *The Clifford Algebra Cliff_L is defined as the algebra over $\mathbb{R}$ generated by L objects γ_k, k=1,2,...,L, which are anti-commuting*

$$\gamma_k\gamma_j = -\gamma_j\gamma_k, \quad \forall k \neq j, \tag{30}$$

square roots of unity

$$\gamma_k^2 = -1, \quad \forall k = 1,2,...,L. \tag{31}$$

As a vector space over $\mathbb{R}$, Cliff_L has the basis

$$\mathcal{B}_L = \{\mathbb{1}_L\} \cup \{\gamma_k \mid k = 1,2,...,L\} \bigcup_{m=2}^{L} \{\prod_{i=1}^{m} \gamma_{k_i} \mid 1 \leq k_i \leq k_{i+1} \leq L\}. \tag{32}$$

Note that the product of more than L generators $\{\gamma_k\}$ can be reduced to a $\pm$ product involving each generator at most once by using the defining relations (30) and (31). To make the connection with ST codes, we turn to a complex representation of the elements of Cliff_L. Clearly, any such representation is completely specified by a representation of its basis elements which in turn are completely specified by a representation of the generators γ_ks. We are interested in particular, in unitary representations of γ_ks which when coupled with (31) will result in an skew-Hermitian representation of the γ_ks. Thus, in summary, we are looking for a representation of the L generators which will form a collection of L anti-commuting, skew-Hermitian, unitary matrices over the field $\mathbb{C}$ of complex numbers. The matrix representations G_i of the γ_i will then satisfy

$$\begin{aligned} G_i^2 &= -I \quad \text{and} \quad G_iG_j + G_jG_i = [0] \text{ or equivalently,} \\ G_i^\dagger &= -G_i \quad \text{and} \quad G_iG_j^\dagger + G_jG_i^\dagger = [0] \end{aligned}$$

and thus satisfy the conditions needed for constructing orthogonal designs (28). Thus orthogonal designs of different sizes can be constructed through the representation of Cliff_L for different L.

***Example* 1.** *A representation of anti-commuting, skew-Hermitian, unitary matrices over $\mathbb{C}$ for the case $L = 2$ is presented below:*

$$\mathcal{R}(\mathbb{1}_2) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} := \sigma_0 \quad \mathcal{R}(\gamma_1) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} := \sigma_1$$

$$\mathcal{R}(\gamma_2) = \begin{bmatrix} 0 & \imath \\ \imath & 0 \end{bmatrix} := \sigma_2 \quad \mathcal{R}(\gamma_1\gamma_2) = \begin{bmatrix} \imath & 0 \\ 0 & -\imath \end{bmatrix} := \sigma_3,$$

where $\mathcal{R}$ denotes the representation, σ_1, σ_2 and σ_3 are unitary skew-Hermitian and $\sigma_1\sigma_2 = \sigma_3$.

Starting with the unitary representation for Cliff_2 given in Example 1, the unitary representation of Cliff_L for any L can be derived. From Proposition $A.8$ of [9], Cliff_{2M-2} can be represented as a tensor product of $M-1$ copies of representations of Cliff_2. Thus the matrices which represent Cliff_{2M-2} are of size 2^{M-1}. We get an explicit unitary representation of Cliff_{2M-1} for all M as given by the Theorem $A.2$ of [9]. The unitary representation for Cliff_{2M-2} is also obtained from the representation matrices of Cliff_{2M-1} by Theorem $A.2$ of [9]. Thus, unitary representations of both Cliff_{2M-1} and Cliff_{2M-2} are of size 2^{M-1}.

Application to Space-Time Code Construction We now consider the problem of code construction from orthogonal designs corresponding to unitary representations of Clifford algebras. Given a specific number n_t of antennas, we consider the representations for generators of Cliff_L where $L = 2\log_2 T + 1$ for $T = 2^{\lceil \log_2 n_t \rceil}$. The representations of the L generators of this Clifford algebra, together with the identity matrix, form a collection of $(L+1)$ matrices each of size $(T \times T)$. This collection of $(L+1)$ matrices can be converted into a set of $L+1$ unitary matrices $\{\beta_i \mid 1 \leq i \leq (L+1)\}$ over $\mathbb{C}$ satisfying the following conditions needed for complex orthogonal designs by multiplying them on the left by an arbitrary unitary $(T \times T)$ matrix β_1:

$$\begin{aligned} \beta_i \beta_i^\dagger &= I \quad \forall 1 \leq i \leq (L+1) \\ \beta_i \beta_j^\dagger + \beta_j \beta_i^\dagger &= [0] \quad \forall 1 \leq i \neq j \leq (L+1). \end{aligned} \tag{33}$$

Finally, the ST code for n_t antennas is obtained from these $(T \times T)$ matrices by deleting the same set of $(T - n_t)$ rows from each of these matrices. It can be easily shown that the submatrices thus obtained continue to satisfy the conditions (33). Thus, the complex symbol rate of a ST code for n_t transmit antennas, derived from unitary representation of clifford algebra is given by

$$\begin{aligned} R_{\mathcal{A}} &= \frac{\log_2 T + 1}{T}, \\ \text{i.e.,} \quad R_{\mathcal{A}} &= \frac{\lceil \log_2 n_t \rceil + 1}{2^{\lceil \log_2 n_t \rceil}}. \end{aligned} \tag{34}$$

From this, it follows that ST codes derived from Clifford algebras in this manner are optimal when n_t is a power of 2. For other values of n_t, rectangular codes obtained from Clifford algebras do not achieve the maximum rate achievable by a complex rectangular orthogonal design [10].

Remark 2. *Orthogonal designs permit maximal transmit diversity at the cost of rate. The rate can be recovered by relaxing the single symbol decodability condition to multi-symbol decodability. Clifford algebras can also be used for constructing multi-symbol decodable ST codes, see [22].*

§ 5 Fixed Multiplexing Gain Approach

We pursue in this section, a second approach to ST code design where the goal is to design a ST code that is designed to transmit at fixed MG r. This means that the rate in bpcu communicated by the ST code will vary with SNR as shown below:

$$R = r \log \text{SNR} \quad \text{bpcu}.$$

Since the rate is a function of the SNR, we speak of a *coding scheme*, by which is meant a family of codes indexed by SNR as opposed to dealing with a single code in the fixed-rate case.

5.A Diversity-Multiplexing Gain Tradeoff

It was shown in a landmark paper of Zheng and Tse [11] that there is a tradeoff between rate, represented by MG r, and reliability, measured by diversity gain $d(r)$. This tradeoff is termed the diversity-multiplexing gain tradeoff (DMT) and is characterized in the theorem below for the case of the Rayleigh-fading channel.

Theorem 5. *[11] Assume $T \geq n_t + n_r - 1$. The optimal (i.e., best possible) tradeoff curve $d^*(r)$ is given by the piecewise linear function connecting the points $(r, d^*(r))$, $r = 0, 1, \ldots, \min\{n_t, n_r\}$, where*

$$d^*(r) = (n_t - r)(n_r - r). \tag{35}$$

In particular,

$$\begin{aligned} \max_{d^*(r)>0} \{r\} &= \min\{n_t, n_r\}, \\ \max_{0 \leq r \leq \min\{n_t, n_r\}} d^*(r) &= d^*(0) = n_t n_r . \end{aligned}$$

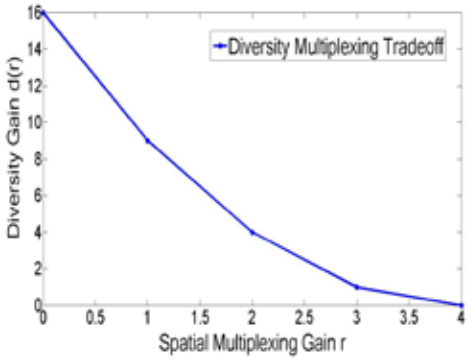

Figure 4: DMT curve for the case of $n_t = n_r = T = 4$.

Zheng and Tse [11] showed that random Gaussian codes achieve the DMT provided $T \geq n_t + n_r - 1$. The restriction $T \geq n_t + n_r - 1$ appearing in the theorem was relaxed to $T \geq n_t$ by Elia et. al. [12] who presented explicit ST code constructions with $T = n_t$, based on cyclic division algebras (CDAs) that achieve the DMT tradeoff. This construction is discussed in greater detail below.

Fig. 4 illustrates the DMT over the ST channel (1) for the case $n_t = n_r = T = 4$. An ST coding scheme that achieves the DMT is said to be DMT-optimal, or simply optimal. Zheng and Tse [11] also show that the DMT allows for the comparison of diverse ST coding schemes. For example, the tradeoff achieved by the Alamouti code is shown in Fig. 5 and it can be seen from the figure, that the Alamouti code does not in general, achieve the DMT [2].

[2]The Alamouti scheme is however, DMT-optimal for the case of two transmit antennas and a single receive antenna i.e $n_t = 2$, $n_r = 1$.

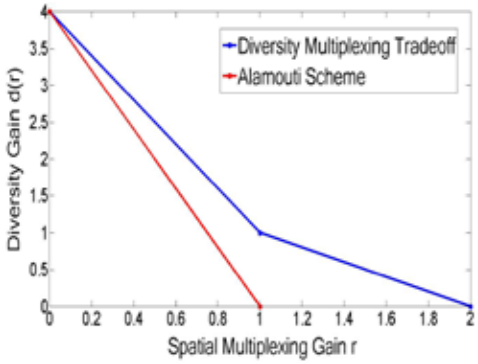

Figure 5: Comparing the optimal DMT curve for $n_t = n_r = T = 2$ with the DMT of the Alamouti code.

5.B Sufficient Criterion for DMT-Optimality

A sufficient condition for a ST code to be DMT-optimal over the Rayleigh-fading channel was established by Elia et. al. in [12]. The same condition is shown to be both necessary and sufficient for a code to be DMT-optimal over any statistical characterization of the fading channel in [13].

Theorem 6 (Sufficient Condition for DMT-Optimality)**.** *[12] An $(n_t \times T)$ ST code $\mathcal{X}$ operating at a MG r and satisfying*

1. *The energy constraint:*
$$\|\theta X\|_F^2 \leq T(SNR) \quad \forall X \in \mathcal{X}.$$
2. *The non-vanishing determinant (NVD) criterion:*
$$\min_{\substack{X_i,\, X_j \in \mathcal{X} \\ \Delta X = X_i - X_j \neq 0}} \det(\Delta X \Delta X^{\dagger}) \doteq SNR^0 \tag{36}$$

is DMT-optimal over the Rayleigh-fading channel.

This theorem can be proven by showing that the probability of error of a ST code satisfying the conditions laid out in the theorem is negligible for large SNR whenever the channel is not in outage for the rate of information being transmitted. As a result, the probability of codeword error of the ST code in the high SNR regime, is equal to the probability of outage, thereby causing the code to be DMT-optimal.

5.C DMT-Optimal Codes from CDAs

Space-time code construction from CDAs was first proposed by Sethuraman, Rajan and Shashidhar [14–17]. Motivated by considerations of coding gain for large values of information rate, Belfiore and Rekaya [18] introduced an approach for constructing CDA-based square $(n \times n)$ ST codes that satisfied the NVD property. Elia et. al. [12] then established the DMT-optimality of ST codes from CDA that possessed the NVD property.

5.D CDA Construction and Property

We will now show how one can go about constructing a CDA-based ST code that possesses the NVD property. We begin by examining the structure of CDAs and taking note of a key property. A convenient means of constructing a CDA is given below.

Proposition 1. *[14, 18, 21] Let $\mathbb{F}$ and $\mathbb{L}$ be fields such that $\mathbb{L}/\mathbb{F}$ is a Galois extension of degree n whose Galois group $G = Gal(\mathbb{L}/\mathbb{F})$ is cyclic, with generator σ. Let $\gamma \in \mathbb{F}^*$ be such that the smallest power t for which γ^t is the relative norm $N_{\mathbb{L}/\mathbb{F}}(u)$ of some element $u \in \mathbb{L}^*$ equals n. The element γ will be referred to as a non-norm element. Let z be an indeterminate such that $z^n = \gamma$ and impose the multiplication rule: $\ell z = z\sigma(\ell)$ for $\ell \in \mathbb{L}$. Then the direct sum $\mathbb{D} = (\mathbb{L}/\mathbb{F}, \sigma, \gamma)$ given by*

$$\mathbb{D} = \mathbb{L} \oplus z\mathbb{L} \oplus \cdots \oplus z^{n-1}\mathbb{L}$$

is a CDA having index n.

The CDA $\mathbb{D}$ is a vector space over its center $\mathbb{F}$ of dimension n^2. In addition, $\mathbb{D}$ is a right-vector space over the maximal subfield $\mathbb{L}$ of dimension n. Every element $d \in \mathbb{D}$ can be written in the form $d = \sum_{i=0}^{n-1} z^i \ell_i$ where $\ell_i \in \mathbb{L}$.

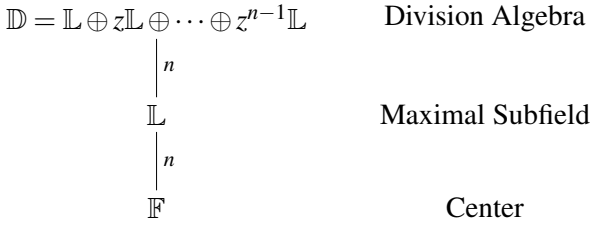

Figure 6: Structure of a CDA.

Left-Regular Representation A ST code $\mathcal{X}$ can be associated to $\mathbb{D} = (\mathbb{L}/\mathbb{F}, \sigma, \gamma)$ by selecting the set of matrices corresponding to the left-regular representation of elements of a finite subset of $\mathbb{D}$ [14]. The left-regular representation X of the element $x = \sum_{i=0}^{n-1} z^i \ell_i,\ \ell_i \in \mathbb{L}$ takes on the form

$$X = \begin{bmatrix} \ell_0 & \gamma\sigma(\ell_{n-1}) & \dots & \gamma\sigma^{n-1}(\ell_1) \\ \ell_1 & \sigma(\ell_0) & \dots & \gamma\sigma^{n-1}(\ell_2) \\ \vdots & \vdots & \vdots & \vdots \\ \ell_{n-1} & \sigma(\ell_{n-2}) & \dots & \sigma^{n-1}(\ell_0) \end{bmatrix}. \tag{37}$$

Note that the elements of X are drawn from the maximal subfield $\mathbb{L}$. The non-commutativity of the CDA endows these matrix representations with a very useful determinant property, namely that the determinant is always guaranteed to lie in the subfield $\mathbb{F}$ of $\mathbb{L}$ that is the center of the CDA.

Lemma 1. *Let X denote the $(n \times n)$ matrix that is the left-regular representation of the element*

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

Then $\det(X) \in \mathbb{F}$.

Proof. We have

$$\begin{aligned}\left(\sum_{i=0}^{n-1} z^i \ell_i\right)\cdot z &= \sum_{i=0}^{n-1} z^{i+1}\sigma(\ell_i)\,,\\ &= z\cdot\left(\sum_{i=0}^{n-1} z^i\sigma(\ell_i)\right)\,,\\ \therefore\ z^{-1}\cdot\left(\sum_{i=0}^{n-1} z^i \ell_i\right)\cdot z &= \sum_{i=0}^{n-1} z^i\sigma(\ell_i).\end{aligned}$$

It follows as a result, that the left-regular representations of $\sum_{i=0}^{n-1} z^i\ell_i$ and of $\sum_{i=0}^{n-1} z^i\sigma(\ell_i)$ are similar matrices and therefore have the same determinant. However, from inspection of (37), it follows that the left-regular representation of $\sum_{i=0}^{n-1} z^i\sigma(\ell_i)$ equals $\sigma(X)$ where X is the left regular representation of $\sum_{i=0}^{n-1} z^i\ell_i$. It follows that X and $\sigma(X)$ are similar and hence, have the same determinant, i.e.,

$$\begin{aligned}\det(\sigma(X)) &= \det(X)\\ \text{i.e.,}\ \ \sigma(\det(X)) &= \det(X),\end{aligned}$$

so that $\det(X)\in\mathbb{F}$. □

The construction begins by choosing $\mathbb{F}$ to be a number field. The particular choice for $\mathbb{F}$ depends on the choice of the underlying signal alphabet $\mathcal{A}$. For the rest of this paper, we will assume the M^2-QAM constellation given in (21). Since, $\mathcal{A}_{\text{QAM}}\subseteq\mathbb{Q}(\imath)$ it is natural to consider CDA with center $\mathbb{F}=\mathbb{Q}(\imath)$.

5.E CDA Possessing the NVD Criterion

Let $\mathcal{O}_{\mathbb{F}}=\mathbb{Z}[\imath]$ be the ring of integers in $\mathbb{F}$. Let $\mathcal{O}_{\mathbb{L}}$ be the integral closure of $\mathcal{O}_{\mathbb{F}}$ in $\mathbb{L}$. Let $\beta_i, i=1,\cdots,n$ be an integral basis for $\mathcal{O}_{\mathbb{L}}/\mathcal{O}_{\mathbb{F}}$. This means that the β_is are a basis for the vector space $\mathbb{L}/\mathbb{F}$ with the additional property that every element in $\mathcal{O}_{\mathbb{L}}$ can be expressed as a linear combination of the β_is with coefficients lying in $\mathcal{O}_{\mathbb{F}}$. We then choose a suitable non-norm element $\gamma\in\mathbb{F}^*$ and proceed to construct the CDA $\mathbb{D}(\mathbb{L}/\mathbb{F},\sigma,\gamma)$. The ST code can be constructed by only considering the left-regular representations of elements of the CDA of the form $\sum_{i=0}^{n-1} z^i\ell_i$ with $\ell_i\in\mathcal{O}_{\mathbb{L}}$ (See Fig. 7). The determinant of the difference of any two code matrices belongs to $\mathbb{F}\cap\mathcal{O}_{\mathbb{L}}=\mathcal{O}_{\mathbb{F}}$ by virtue of Lemma 1. Since, $\mathcal{O}_{\mathbb{F}}=\mathbb{Z}[\imath]$ it follows that the squared magnitude of any of non-zero element of $\mathcal{O}_{\mathbb{F}}$ is an integer ≥ 1. As a result, we have that

$$\min_{\substack{\Delta X=X_i-X_j\neq 0\\ X_i,X_j\in\mathcal{X}}}\det(\Delta X\Delta X^\dagger)\ \dot{\geq}\ \text{SNR}^0, \tag{38}$$

i.e., the ST code so constructed, possesses the NVD property. From Theorem 6, it follows that CDA-based square $(n\times n)$ ST codes with the NVD property achieve the DMT.

***Example* 2.** *The Golden code [23] is an example for a CDA-based (2×2) ST code that is DMT-optimal. A variation of this code is now part of the WiMAX IEEE 802.16e wireless communication standard. The CDA associated with the Golden code is of the form $\mathbb{D}=(\mathbb{L}=$*

$\mathbb{Q}(\iota,\theta)/\mathbb{Q}(\iota),\sigma,\iota)$ *where* $\theta=\frac{1+\sqrt{5}}{2}$ *and* $\sigma:\theta\to\bar{\theta}=\frac{1-\sqrt{5}}{2}$. *Each codeword matrix* $X\in\mathcal{X}$ *of the Golden code is of the form*

$$X=\frac{1}{\sqrt{5}}\begin{bmatrix} l_0 & l_1 \\ \iota\cdot\sigma(l_1) & \sigma(l_0)\end{bmatrix}, \tag{39}$$

obtained by restricting each ℓ_i *to the form* $\ell_i=(1+\iota-\iota\theta)(a+b\theta)$, *where* $a,b\in\mathbb{Z}[\iota]$ *and* $i=0,1$. *The non-norm element in this CDA is given by* $\gamma=\sqrt{\iota}$.

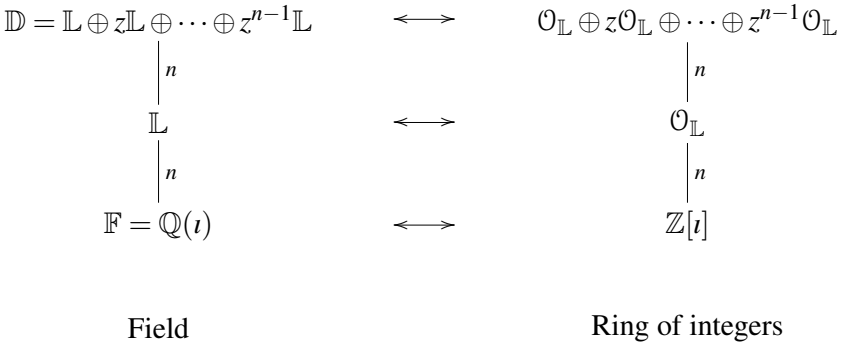

Figure 7: Construction of a CDA that yields ST codes over the QAM constellation possessing the NVD property.

From the discussion in Section 5.D, it follows that the problem of constructing DMT-optimal ST codes from CDAs reduces to one of identifying cyclic Galois extensions $\mathbb{L}$ of arbitrary degree n over $\mathbb{F}=\mathbb{Q}(\iota)$ containing a suitable non-norm element γ. The lemma by Kiran et. al. [19] given below, simplifies the task of identifying a non-norm element γ.

Lemma 2. *[19] Let* $\mathbb{L}$ *be a cyclic extension of a number field* $\mathbb{F}$. *Let* $\mathcal{O}_{\mathbb{F}}$ *denote the ring of integers of* $\mathbb{F}$. *Let* $\mathfrak{p}$ *be a prime ideal of* $\mathcal{O}_{\mathbb{F}}$ *that remains inert in the extension* $\mathbb{L}/\mathbb{F}$ *and let* $\gamma\in\mathfrak{p}\setminus\mathfrak{p}^2$. *Then* γ *is a non-norm element.*

Thus, for constructing DMT-optimal square $(n\times n)$ ST codes from CDA, it is sufficient to construct cyclic extensions of $\mathbb{Q}(\iota)$ of degree n such that $\mathcal{O}_{\mathbb{F}}$ contains a prime ideal $\mathfrak{p}$ that remains inert in the extension. Elia et. al. [12] provided two explicit constructions for CDA based DMT-optimal square $(n\times n)$ ST codes valid for any $n=T=n_t$. One of these constructions is presented in the theorem below.

Let the integer n be factored as follows:

$$n \;=\; 2^{e_0}\prod_{i=1}^{r}p_i^{e_i} \;=\; 2^{e_0}n_1 \tag{40}$$

where the $\{p_i\}$ are distinct odd primes. Given an integer $m\geq 3$, let $\omega_m \;=\; \exp(\iota\frac{2\pi}{m})$.

Theorem 7. *[12] Let* n *be given as in* (40). *Let* p^e *be the smallest prime power such that* $n_1|\phi(p^e)$. *Let* G *be the Galois group of* $\mathbb{Q}(\omega_{p^e})/\mathbb{Q}$. *Let* H *be a subgroup of* G *of size* $\phi(p^e)/n_1$. *Let* $\mathbb{K}$ *be the fixed field of* H. *Let* $\mathbb{M}$ *be the compositum of* $\mathbb{K}$ *and* $\mathbb{Q}(\iota)$ *and* $\mathbb{L}$ *be the compositum of* $\mathbb{M}$ *and* $\mathbb{Q}(\omega_{2^{e_0+2}})$. *Then* $\mathbb{L}$ *is the desired cyclic extension of* $\mathbb{Q}(\iota)$ *of degree* n *(Fig. 8).*

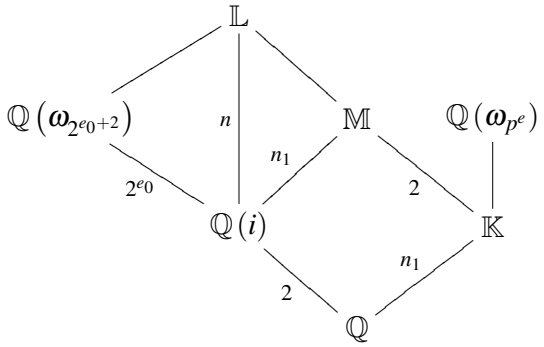

Figure 8: Illustrating the CDA construction in Theorem 7.

Let $\rho \in \mathbb{Z}_{p^e}^*$ be a generator of the cyclic group $\mathbb{Z}_{p^e}^*$. Let q be a rational prime such that

$$q = \begin{cases} \rho & (\text{mod } p^e) \\ 5 & (\text{mod } 2^{e_0+2}) \end{cases}. \tag{41}$$

Let β be a prime ideal of $\mathbb{Z}[\imath]$ lying above $q\mathbb{Z}$ in $\mathbb{Q}(\imath)/\mathbb{Q}$. Then β is the desired prime that remains inert in the extension $\mathbb{L}/\mathbb{Q}(\imath)$.

§ 6 Conclusion

In this paper, we considered the problem of code construction for block-fading space-time channel. Two approaches were considered corresponding respectively, to communicating at fixed rate and communicating at a rate indexed by the SNR as determined by a fixed value of the multiplexing gain. In each case, a tradeoff between rate and reliability was seen to hold. The challenge in code design is consequently one of designing efficient codes that optimally achieve this tradeoff. Algebraic approaches that made use of principal ideal rings, Clifford algebras and cyclic division algebras to come up with efficient code designs have been remarkably successful in meeting this challenge and some of these approaches have been discussed here.

References

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

[2] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Y. Kuo, "Signal design for transmitter diversity wireless communication systems over rayleigh fading channels," *Proc. IEEE VTC 96,* 2000, pp. 136-140.

[3] C. E. Shannon, "A mathematical theory of communication," *Bell Syst.Tech. J.,* vol. 27, pp. 379-423, 623-656, July, October, 1948.

[4] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: John Wiley and Sons, 1991.

[5] H.-F. Lu and P. V. Kumar, "Rate-diversity trade-off of space-time codes with fixed alphabet and optimal constructions for PSK modulation," *IEEE Trans. Info. Theory,* vol. 49, no. 10, pp. 2747-2751, Oct. 2003.

[6] G. Ganske and B. R. McDonald, "Finite Local Rings " in *Rocky Mountain Journal of Mathematics* Volume 3, Number 4, Fall 1973.

[7] S. M. Alamouti, "A simple transmit diversity scheme for wireless communications," *IEEE Journal on Selected Areas in Communication,* pp. 1451-1458, Oct. 1998.

[8] J. F. Adams, P. D. Lax, and R. S. Phillips, "On matrices whose real linear combinations are nonsingular," in *Proc. Amer. Math. Soc.,* vol. 16, pp. 318-322, 1965.

[9] Tirkkonen, O., Hottinen, A., "Square-matrix embeddable space-time block codes for complex signal constellations," *IEEE Trans. Info. Theory,* vol. 48, no. 2, pp. 384-395, Feb. 2002.

[10] X.-B. Liang, "Orthogonal designs with maximal rates," *IEEE Trans. Info. Theory,* vol.49, pp. 2468-2503, Oct. 2003.

[11] L. Zheng and D. Tse, "Diversity and Multiplexing: A Fundamental Tradeoff in Multiple-Antenna Channels," *IEEE Trans. Info. Theory*, vol. 49, no. 5, pp. 1073-1096, May 2003.

[12] Petros 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. Info. Theory,* vol. 52, No. 9, pp. 3689-3884, Sep. 2006.

[13] S. Tavildar and P. Viswanath, "Approximately universal codes over slow fading channels,"*IEEE Trans. Info. Theory*, vol. 52, pp. 3233-3258, Jul. 2006.

[14] 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.

[15] B. A. Sethuraman and B.Sundar Rajan, "Full-rank, full-rate STBCs from division algebras," *Proc. IEEE Info. Theory Workshop*, pp. 69-72, Oct. 20 - 25, 2002, Bangalore, India.

[16] B. A. Sethuraman and B.Sundar Rajan, "An algebraic description of orthogonal designs and the uniqueness of the Alamouti code," *Proc. IEEE Global Telecom. Conf.*, Taipei, pp. 1088-1092, 2002.

[17] V. Shashidhar, B.Sundar Rajan and B.A.Sethuraman, "STBCs using capacity achieving designs from cyclic division algebras," *Proc. IEEE Global Telecomm. Conf.*, 1-5 Dec., San Francisco, 2003, Vol. 4, pp. 1957-1962.

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

[19] Kiran.T. and B.Sundar Rajan, "STBC-schemes with non-vanishing determinant for certain number of transmit antennas," *IEEE Trans. Info. Theory*, vol. 51, pp. 2984 - 2992, August 2005.

[20] Paulo Ribenboim, *Classical theory of Algebraic Numbers*, New York: Springer-Verlag: Universitext, 2001.

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

[22] S. Karmakar and B. S. Rajan, “Multi-group decodable STBCs from Clifford algebras,” in *IEEE Trans. Info. Theory,* vol. 55, no. 1, pp. 223-231, Jan. 2009.

[23] J.-C. Belfiore, G. Rekaya, E. Viterbo: "The Golden Code: A 2×2 Full-Rate Space-Time Code with Non-Vanishing Determinants," *IEEE Trans. on Inform. Theory*, vol. 51, no. 4, pp. 1432-1436, Apr. 2005.

Received 14 February 2011; Revised 28 February 2011

Sharanappa C Ijeri completed his B.Tech degree from National Institute of Technology, Surathkal in 2008. He worked for Analog Devices for a year as design engineer. He is currently pursuing his M.E. in telecommunications at the Indian Institute of Science, Bangalore. His research interests include wireless communication, wireless sensor networks, system biology.

Vanamali Bhat completed his B.Tech degree from National Institute of Technology, Surathkal in 2009. He is currently pursuing his M.E. in telecommunications at the Indian Institute of Science, Bangalore. His research interests include signal processing, wireless communication and wireless sensor networks.

U. Raviteja completed his Masters (MS by research) from Indian Institute of Technology, Madras in 2008. His masters thesis discusses the design and construction of a system of nested LDPC codes for key reconciliation. He is currently pursuing his Ph.D. at the Indian Institute of Science, Bangalore. His research interests include signal processing, machine learning, and wireless sensor networks.

P. Vijay Kumar received the B.Tech. and M.Tech. degrees from the Indian Institutes of Technology (Kharagpur and Kanpur),and the Ph.D. Degree from the University of Southern California (USC) in 1983, all in Electrical Engineering. From 1983-2003 he was on the faculty of the EE-Systems Department at USC. Since 2003 he has been on the faculty of the Indian Institute of Science, Bangalore and also holds the position of adjunct research professor at USC. His current research interests include codes for distributed storage, distributed function computation, sensor networks and space-time codes for MIMO and cooperative communication networks. He is a fellow of the IEEE and an ISI highly-cited author. He is co-recipient of the 1995 IEEE Information Theory Society prize paper award as well as of a best paper award at the DCOSS 2008 conference on sensor networks.